

République Algérienne Démocratique et Populaire
Ministère de l'enseignement supérieur et de la recherche scientifique
Université Mouloud Mammeri de Tizi-Ouzou
Faculté de Génie Electrique et Informatique
Département Informatique



THESE DE DOCTORAT LMD

DISCIPLINE : Informatique

Option : Ingénieries des Connaissances Appliquées à la Recherche d'information
et à la E-éducation

Présentée par

Samia LAZIB Ep. CHOUAKI

Sujet

Un système pour la E-Orientation Scolaire Intégrant l'Analyse d'Opinion

Jury d'examen composé de :

- | | | | |
|------------------------------------|-------------|--------------|--------------|
| • Mr SI MOHAMMED Malik | Professeur | UMMTO | Président |
| • Mme BOUARAB Née DAHMANI Farida | Professeure | UMMTO | Rapporteur |
| • Mme DRIAS Habiba | Professeure | USTHB, Alger | Examinatrice |
| • Mme AIT EL HADJ Née SOUAM Fatiha | MCA | UMMTO | Examinatrice |
| • Mr AMROUN Kamal | MCA | UAM, Bejaïa | Examineur |

Soutenue le 12 Mars 2020

Remerciements

En premier lieu, je remercie Dieu le tout puissant pour m'avoir donné la volonté et la force pour accomplir ce travail.

Je remercie aussi le Comité de la Formation Doctorale ICARE pour l'opportunité qu'il nous a offert afin d'accéder à ce doctorat.

Je tiens à remercier chaleureusement, ma directrice de recherche, Madame **Farida Bourarb-Dahmani**, Professeur à l'université de Tizi-Ouzou (UMMTO), pour ses conseils toujours pertinents, pour sa manière à la fois rigoureuse et agréable de mener les travaux du projet, pour son appui et ses encouragements dans les moments d'impasse de la recherche. Son aide précieuse, son soutien et sa disponibilité m'ont permis à faire les premiers pas dans cette recherche et à l'amener à maturité. Qu'elle soit rassurée de ma profonde gratitude et ma parfaite considération.

J'adresse mes cordiaux remerciements aux membres du jury :

- Monsieur **Malik Si Mohammed**, Professeur à l'UMMTO de l'honneur qu'il me fait en présidant le jury d'examen de cette thèse ;
- Madame **Habiba Drias**, Professeur à l'USTHB pour m'avoir fait l'honneur de bien vouloir porter intérêt à ce travail ;
- Madame **Fatiha Ait-El-hadj Née Souam**, Maître de Conférences à l'UMMTO pour son aimable participation au jury d'évaluation de ce travail de recherche ;
- Mr **Kamal Amroun**, Maître de Conférences à l'UAMB pour avoir accepté aimablement de juger cette thèse.

Ma reconnaissance s'adresse aussi à tous les membres de mon équipe de recherche, Mme Tassadit Berkane, Maître Assistante à l'UMMTO, d'avoir relu mon manuscrit avec beaucoup de précision et de dévouement, Dr Lynda Haddadi pour son soutien et ses encouragements, Dr Nacera Hamid pour ses relectures de mes travaux en anglais,

Mes remerciements s'adressent aussi à tous lycéens, leurs enseignants et leurs parents ayant prêté leur concours lors de la réalisation des enquêtes menées.

Je n'oublierais pas de remercier Mme Sadia Belkhir, Maître de Conférences à l'UMMTO pour ses relectures de mes travaux en anglais, Mme Louiza Marouf Khelefane, Professeur à l'UMMTO pour ses orientations, Mme Fatiha Tali, Maître de Conférences et chercheuse au sein du labo UMR-EFTS de Toulouse Jean Jaurès en France, pour ses relectures et conseils.

Je remercie affectueusement et intensément tous les membres de ma famille pour m'avoir soutenu et assisté tout au long de l'accomplissement de ce travail.

A tous ceux qui m'ont aidée et soutenue durant cette thèse de près ou de loin et que je n'ai pas cités, qu'ils trouvent ici l'expression de ma profonde reconnaissance.

*Je dédie cette thèse à
ma famille et ma belle famille*

Résumé

L'analyse d'opinion est l'une des grandes réalisations de la dernière décennie dans le domaine des technologies langagières. La prise en compte de l'opinion dans les processus de prise de décision est devenue un élément crucial de réussite. Cette technologie se voit alors introduite de manière très répandue dans de nombreuses disciplines, notamment dans celles de l'éducation.

L'orientation scolaire vise à améliorer la qualité des systèmes éducatifs et celle des rapports entre éducation et vie professionnelle. Cependant, dans la plus part des cas, l'orientation s'appuie sur les moyennes des notes. Les notes seules, ne peuvent pas satisfaire les attentes de tous les intervenants. Aussi, le système d'orientation est conçu de manière à conférer un monopole à l'administration sans associer tous les concernés : enfants, familles et enseignants. Dès lors, la mise en place de mécanismes renforçant les critères de l'orientation scolaire est plus qu'une nécessité.

Dans cette thèse, nous avons proposé un système de E-orientation scolaire, OPinOR (Opinion pour Orientation), où le processus d'orientation s'articule sur trois dimensions : la dimension opinion comprenant les avis et les idées de l'élève balisés par le taux de certitude qui vient renforcer la crédibilité de cette dimension, la dimension personnelle incluant tout ce qui est natif chez l'élève comme les traits de personnalité, les capacités, les intelligences, etc. La troisième est la dimension capitalisation basée sur les acquisitions de l'élève par ses propres efforts. Le système proposé permet la génération de carte d'aide à l'orientation utilisant un processus de cinq principaux modules : Le module d'initialisation de paramètres, le module de construction de la ressource lexicale DICO, le module de traitement de l'opinion, le module de mesure de certitude et enfin le module de synthèse d'orientation. Afin de valider nos propositions, nous avons développé deux prototypes : le premier pour la validation des propositions sans l'introduction de mesure de certitude et le second a permis l'évaluation de l'amélioration de l'orientation après introduction de la prise en compte de la certitude dans la synthèse de l'opinion. Les résultats de cette validation sont très intéressants.

Mots clés: Orientation Scolaire, E-Orientation Scolaire, Analyse D'opinion, Opinion Formatée, Opinion Libre, Analyse D'opinion Pour L'éducation, Synthèse D'opinion, Opinion Des Elèves, Mesure De Certitude, Opinion Fiable, Modélisation De L'opinion, Stabilité De L'opinion.

Abstract

Opinion analysis is one of the great achievements of the last decade in the field of language technologies. The price of attention in the decision-making process has become a crucial element of success. This technology is then widely introduced in many disciplines, particularly in education

Educational guidance aims to improve the quality of education systems and the relationship between education and working life. However, in most cases, the orientation is based on the averages of marks. Whereas, marks, alone cannot meet the expectations of all stakeholders. In addition, the guidance system is designed to confer a monopoly to the administration without involving all stakeholders: children, families and teachers Therefore, the establishment of mechanisms that strengthen the criteria for educational guidance is more than necessary.

In this thesis, we have proposed an educational guidance aided system, OPinOR (OPinion for ORientation). The orientation process is based on three dimensions: the opinion dimension including the opinions and ideas of the student marked up by a certainty rate that reinforces the credibility of this dimension, the personal dimension including everything that is native to the student, such as personality traits, abilities, intelligences, etc. The third one is the capitalization dimension based on the student's acquisitions through his or her own efforts. The proposed system allows the generation of orientation card using five main modules: Parameter initialization module, the building module of the DICO lexical resource, the opinion-processing module, the certainty measurement module and finally the orientation synthesis module. In order to validate our proposals, we have developed two prototypes: the first one, for the validation of the proposals without introduction of the certainty measure and the second one for the validation of the orientation improvement after introduction of certainty in the synthesis of opinion. The results of this validation are very interesting.

Key words: Educational Guidance, E-Educational Guidance, Opinion Mining, Formatted opinion, Free Opinion, Opinion Mining For Education, Opinion Synthesis, Pupil's Opinions, Certainty Measure, Reliable Opinion, Opinion Modeling, Opinion Stability.

Sommaire

REMERCIEMENTS	I
RESUME	III
ABSTRACT	IV
LISTE DES FIGURES	XI
LISTE DES TABLEAUX	XIII
INTRODUCTION GENERALE	1
1. CONTEXTE GENERAL	2
2. PROBLEMATIQUE	3
3. OBJECTIFS	3
4. NOTRE CONTRIBUTION	4
5. PLAN DE LA THESE	4
PARTIE A : ETUDE BIBLIOGRAPHIQUE	6
CHAPITRE I : ANALYSE D'OPINION	7
INTRODUCTION	7
I.2 L'ANALYSE D'OPINION- DEFINITIONS	7
I.2.1 FOUILLE DE TEXTE & FOUILLE DE DONNEES	7
I.2.2 LA FOUILLE D'OPINION (OPINION MINING)	7
I.2.3 DEFINITION	8
I.3 LES TACHES DE L'ANALYSE D'OPINION	9
I.3.1-CLASSIFICATION DE SUBJECTIVITE	9
I.3.2 CLASSIFICATION D'OPINION	10
<i>I.3.2.1 Détermination de la polarité</i>	10
<i>I.3.2.2 Classification d'opinion pour texte vague</i>	10
<i>I.3.2.3 Texte multilingue et analyse d'opinion</i>	10
<i>I.3.2.4 Classification d'opinion inter-domaines</i>	11
I.3.3 MESURE DE L'INTERET DES CRITIQUES	11
I.3.4 DETECTION DE SPAM D'OPINION	11
I.3.5 CREATION DE LEXIQUE D'OPINION	12
<i>I.3.5.1 Approches par induction lexicale</i>	12

<i>1.3.5.2 Approches fondées sur les ontologies</i>	14
1.3.6 EXTRACTION D'ASPECTS	14
I.4 LES NIVEAUX D'ANALYSE D'OPINION	15
I.5 LES APPROCHES DE CLASSIFICATION DANS L'ANALYSE D'OPINION	15
1.5.1 LES APPROCHES SYMBOLIQUES	15
1.5.2 LES APPROCHES STATISTIQUES	16
1.5.3 LES APPROCHES HYBRIDES	17
1.5.4 COMPARAISON ENTRE LES DIFFERENTES APPROCHES D'ANALYSE D'OPINION	18
1.5.6 MESURES DE QUALITE D'UNE CLASSIFICATION DE TEXTES	20
I.6 WORDNET, WOLF ET SENTIWORDNET	22
1.6.1. WORDNET	22
<i>1.6.1.1 Définition</i>	22
<i>1.6.1.2 Notion de synset</i>	22
<i>1.6.1.3. Fréquence des lemmes</i>	24
<i>1.6.1.4. Mesures de similarité</i>	25
<i>1.6.1.5. Mappage entre différentes versions</i>	25
<i>1.6.1.6- WordNets pour d'autres langues que l'anglais</i>	25
1.6.2 WOLF	26
1.6.3 SENTIWORDNET	27
I.7 MODELISATION DE L'OPINION	28
1.7.1 MODELE DE LIU	29
1.7.2 MODELE DE L'APPRaisal	29
I.8 APPLICATION DE L'ANALYSE D'OPINION	30
1.8.1 L'ANALYSE D'OPINION DANS LE DOMAINE COMMERCIAL	31
<i>1.8.1.1 Le point de vue du client individuel</i>	31
<i>1.8.1.2 Le point de vue des entreprises</i>	31
<i>1.8.1.3 Le point de vue des agences de publicité</i>	32
1.8.2 L'ANALYSE D'OPINION DANS LE DOMAINE POLITIQUE	32
1.8.3 L'ANALYSE D'OPINION POUR LE MARCHE BOURSIER	33
1.8.4 L'ANALYSE D'OPINION POUR L'EDUCATION	33
<i>1.8.4.1 Analyse d'opinion pour améliorer la qualité de l'enseignement</i>	33
<i>1.8.4.2 Analyse d'opinion pour la recherche scientifique et communautaire</i>	35
<i>1.8.4.3 Analyse d'opinion pour recueillir l'opinion publique sur l'enseignement</i>	35
CONCLUSION	36

CHAPITRE II : L'ORIENTATION ET L'E-ORIENTATION SCOLAIRE.....	37
INTRODUCTION	37
II.2 DEFINITIONS DE QUELQUES NOTIONS DE BASE	37
II.2.1 L'EDUCATION	37
II.2.2 LA PEDAGOGIE	38
II.2.3 LA DIDACTIQUE.....	39
II.3 L'ORIENTATION SCOLAIRE	40
II.3.1 DEFINITION	40
II.3.2 LES OBJECTIFS ET ROLE DE L'ORIENTATION SCOLAIRE	41
II.3.3 LES ENJEUX DE L'ORIENTATION SCOLAIRE	41
<i>II.3.3.1 Les enjeux collectifs</i>	<i>42</i>
<i>II.3.3.2 Les enjeux individuels</i>	<i>42</i>
II.3.4 MOYENS DE L'ORIENTATION SCOLAIRE	42
II.3.4.1 L'observation	42
II.3.4.2 La rencontre – Les entretiens individuels.....	42
II.3.4.3 Les séminaires	43
II.3.4.4 L'immersion et les stages en milieu professionnel.....	43
II.3.4.5 L'éducation au choix - L'éducation à l'orientation.....	43
II.3.4.6 Les tests psychologiques – Styles d'apprentissage	44
II.4 LES TIC ET E-ORIENTATION SCOLAIRE	46
II.4.1 L'E-ORIENTATION.....	46
II.4.2 APPORTS DES TIC POUR L'ORIENTATION SCOLAIRE	46
II.4.3 LES LOGICIELS D'AIDE A L'ORIENTATION.....	47
II.5 L'ORIENTATION SCOLAIRE A TRAVERS LE MONDE	49
II.5.1 L'ORIENTATION SCOLAIRE EN ALGERIE	49
<i>II.5.1.1 Le système éducatif Algérien</i>	<i>49</i>
<i>II.5.1.2 L'orientation scolaire en Algérie</i>	<i>54</i>
II.5.2 L'ORIENTATION SCOLAIRE AUX USA	57
II.5.3 L'ORIENTATION SCOLAIRE AU JAPON	58
II.5.4 L'ORIENTATION SCOLAIRE EN FRANCE	60
II.6 CONTRAINTES DE L'ORIENTATION SCOLAIRE	61
CONCLUSION	63
PARTIE B : PROPOSITIONS ET LEURS VALIDATIONS.....	64

CHAPITRE III : PROPOSITION D'UN SYSTEME DE E-ORIENTATION SCOLAIRE BASE SUR L'ANALYSE D'OPINION	65
INTRODUCTION	65
III.2 CONTEXTE GENERAL DE NOS PROPOSITIONS.....	65
III.3 ARCHITECTURE DU SYSTEME OPINOR.....	67
III.4 MODULE DE PARAMETRAGE.....	68
III.5 MODULE DE CONSTRUCTION DE LA RESSOURCE LEXICALE DICO	71
III.6 MODULE DE TRAITEMENT DE L'OPINION	73
III.6.1 MODULE DE RECUEIL D'OPINIONS	74
III.6.2 MODULE D'EXTRACTION D'OPINIONS	75
<i>III.6.2.1 Extraction de l'opinion formatée.....</i>	<i>75</i>
<i>III.6.2.2 Extraction de l'opinion libre.....</i>	<i>75</i>
<i>III.6.3 Module d'Amplification de l'opinion par le taux de certitude</i>	<i>80</i>
<i>III.6.4 Classification de l'opinion</i>	<i>81</i>
III.7 MODULE DE MESURE DE LA CERTITUDE DE L'EMETTEUR D'OPINION	81
III.7.1 MODELISATION DE LA CERTITUDE DE L'EMETTEUR D'OPINION.....	81
<i>III.7.1.1 Modélisation de l'opinion</i>	<i>82</i>
<i>III.7.1.2 Les patrons lexico-syntaxiques</i>	<i>82</i>
III.7.2 PREPARATION DES DONNEES.....	83
<i>III.7.2.1 classification des verbes d'opinions dans DICO</i>	<i>83</i>
<i>III.7.2.2 Patrons lexico syntaxique utilisés</i>	<i>84</i>
III.7.3 MESURE DE LA CERTITUDE DE L'EMETTEUR D'OPINION	85
III.8 MODULE DE SYNTHESE DE L'ORIENTATION	87
III.8.1 MODULE DE SYNTHESE D'OPINION	88
<i>III.8.1.1 Moyenne périodique des opinions.....</i>	<i>88</i>
<i>III.8.1.2 Opinion par domaine :.....</i>	<i>88</i>
<i>III.8.1.3 Opinion générale :</i>	<i>89</i>
III.8.2 MODULE DE CALCUL DU SCORE D'ORIENTATION	90
III.9 STRUCTURES DE L'ALGORITHME DU PROCESSUS PROPOSE	92
CONCLUSION	92
CHAPITRE IV : EXPERIMENTATION ET VALIDATION DES PROPOSITIONS	93
INTRODUCTION	93

IV.2 EVALUATION DU MODULE D'EXTRACTION DE L'OPINION LIBRE	93
IV.3 EVALUATION DU SYSTEME OPINOR SANS PRISE EN COMPTE DE LA CERTITUDE	94
IV.3.1 PROTOCOLE D'EVALUATION.....	94
IV.3.2 ECHANTILLON	95
IV.3.3 RESULTATS.....	98
<i>IV.3.3.1 Cas de la moyenne des notes seules.....</i>	<i>100</i>
<i>IV.3.3.2 Cas de la moyenne des opinions seules.....</i>	<i>101</i>
<i>IV.3.3.3 Cas des aptitudes seules.....</i>	<i>101</i>
<i>IV.3.3.4 Cas des trois paramètres combinés.....</i>	<i>102</i>
IV.3.4 INTERPRETATION.....	103
IV.4 EVALUATION DE LA MESURE DE CERTITUDE ET SON IMPACT SUR L'ORIENTATION SCOLAIRE	103
IV.4.1 PROTOCOLE D'EVALUATION.....	103
IV.4.2 EVALUATION	105
CONCLUSION	109
CHAPITRE V : SYNTHÈSE ET DISCUSSION DE NOS PROPOSITIONS	111
INTRODUCTION	111
V.2. SYNTHÈSE ET DISCUSSION	111
V.3. POINTS FORTS ET LIMITES DE NOS PROPOSITIONS.....	112
V.3.1 POINTS FORTS DE NOTRE PROPOSITION	113
V.3.2 POINTS FAIBLES DE NOS PROPOSITIONS	114
V.4 IMPACTS DE NOTRE PROPOSITION	115
V.4.1 IMPACT SUR L'APPRENTISSAGE ET L'ENSEIGNEMENT	116
V.4.2 IMPACT LA CITOYENNETE	116
V.4.3 IMPACT SUR LA RECHERCHE EN E-EDUCATION VERSUS E-ORIENTATION	116
CONCLUSION	117
CONCLUSION GENERALE ET PERSPECTIVES DE RECHERCHE	118
REFERENCES BIBLIOGRAPHIQUES ET WEBOGRAPHIES	122
REFERENCES BIBLIOGRAPHIQUES	123
REFERENCES WEBOGRAPHIES.....	139
ANNEXES	142
ANNEXES 1 – QUESTIONNAIRE POUR RECUEILLIR LES APTITUDES DE L'ELEVE.	143

ANNEXES 1-BIS LES INTELLIGENCES MULTIPLES ET LES PROFESSIONS CORRESPONDANTES.....	147
ANNEXE 2 : LISTE DES COMMUNICATIONS ET PUBLICATIONS PERSONNELLES.....	148
ANNEXE 3 : CERTIFICATIONS DE PARTICIPATION AUX CONFERENCES.....	149

Liste des figures

Figure 1 Axes de l'Analyse d'opinion (Ravi & Ravi, 2015).....	9
Figure 2 Techniques d'analyse d'opinion (Hemmatian & Sohrabi, 2017).....	18
Figure 3 Aperçu sur la structuration de WordNet.....	23
Figure 4 Exemple illustratif du synset (ID : "02086723 – n ") dans PWN.....	24
Figure 5 Aperçu sur la structure de SentiWordNet.....	28
Figure 6 Exemple de scores de polarité dans SentiWordNet (Chaumartin, 2007).	28
Figure 7 Vue globale du système de l'Appraisal (Zhang & Ferrari, 2014).	30
Figure 8 Domaines d'applications de l'analyse d'opinion.	31
Figure 9 Interaction entre l'orientation scolaire, la pédagogie et la didactique.	39
Figure 10 Rôles de l'orientation scolaire.	41
Figure 11 Structuration du système éducatif algérien (MEN, 2016).....	50
Figure 12 Structures de bases du système universitaire Algérien (MESRS, 2018).....	53
Figure 13 Inscription aux cycles d'études supérieures en Algérie (MESRS, 2018).....	55
Figure 14 Critères de l'orientation scolaire en Algérie (MESRS, 2018).	56
Figure 15 Dimensions de OPinOR.	66
Figure 16 Architecture d'OPinOR.	68
Figure 17 Liaisons pondérées des matières aux domaines d'études.	69
Figure 18 Liaisons pondérées des aptitudes des élèves aux domaines d'études.....	70
Figure 19 Construction de la ressource lexicale DICO.	72
Figure 20 Structure de la ressource lexicale DICO.	73
Figure 21 Architecture du module de traitement de l'opinion.....	73
Figure 22 Partie du questionnaire utilisé pour le recueil d'opinions.	74
Figure 23 Digramme d'extraction de l'opinion libre.	75
Figure 24 Algorithme de la procédure de prétraitement du texte brut.....	76
Figure 25 Algorithme de la procédure de desembugisation.	77
Figure 26 Algorithme d'extraction de l'opinion libre.....	79
Figure 27 Modélisation de l'opinion.	82
Figure 28 La ressource lexicale avec la classification des verbes et adverbes.	84
Figure 29 Diagramme de fonctionnement du module de mesure de certitude.	85
Figure 30 Architecture du module de synthèse d'orientation.	87
Figure 31 Niveaux de synthèse d'opinion.	90

Figure 32 Carte d'aide à l'orientation détaillée.....	91
Figure 33 Organigramme principal d'OPinOR.	92
Figure 34 Domaines d'études universitaires considérés.....	96
Figure 35 Liste des matières enseignées considérées dans cette étude.....	96
Figure 36 Pondération des notes de matières (α_i) par domaines d'études.....	97
Figure 37 Codification et Pondération des Aptitudes de l'élève.....	98
Figure 38 Carte d'aide à l'orientation synthétisée basée sur les notes du BAC.....	99
Figure 39 Carte d'aide à l'orientation synthétisée avec les trois critères.....	100
Figure 40 Stabilité et certitude de l'opinion.....	107

Liste des tableaux

Tableau 1 Comparaison des approches de l'analyse d'opinion.....	19
Tableau 2 Matrice de confusion pour la classification binaire.....	20
Tableau 3 Critère de performances pour la multi classification.....	21
Tableau 4 Inteligences multiples de Gardner.....	45
Tableau 5 Valeurs des polarités et classification de l'opinion.....	70
Tableau 6 Exemples illustratifs de calcul de l'opinion libre (VOL).....	78
Tableau 7 Liste des Patrons lexico-syntaxiques implémentés.....	84
Tableau 8 Taux de certitude des PLS selon CVA.....	85
Tableau 9 Exemple d'amplification de l'opinion par la certitude.....	86
Tableau 10 Matrice de confusion Obtenue.....	93
Tableau 11 Interprétation des valeurs du coefficient Kappa.....	95
Tableau 12 Matrice de similitudes - Cas des moyennes des notes.....	100
Tableau 13. Matrice de similitudes - Cas des opinions.....	101
Tableau 14 Matrice de similitudes - Cas des aptitudes.....	102
Tableau 15 Matrice de similitudes - Cas des trois paramètres.....	102
Tableau 16 Synthèse des valeurs du coefficient Kappa obtenues.....	103
Tableau 17 Les résultats obtenus pour un élève.....	106
Tableau 18 Valeur de la stabilité de l'opinion de l'élève.....	106
Tableau 19 Stabilité et moyenne de certitude de l'opinion des élèves.....	107
Tableau 20 Couverture en certitude.....	108
Tableau 21 Matrices de similitude avec et sans amplification par la certitude.....	109
Tableau 22 Interprétation des valeurs du coefficient Kappa obtenues.....	109

Introduction générale

1. Contexte général

L'orientation scolaire est la détermination de la meilleure voie scolaire ou professionnelle pour l'élève, en particulier en fin de cycle, après inventaire de ses aptitudes et de ses motivations (Berou, 2010). Elle vise, à la fois, l'amélioration de la qualité des systèmes éducatifs et celle des rapports entre éducation et vie professionnelle (Meunier, 2008). Elle détermine, en effet, les chances d'une insertion professionnelle réussie.

Cependant, dans la plupart des cas, l'orientation consiste à trier et classer les élèves en fonction de leurs seuls résultats scolaires dans les savoirs abstraits. Elle s'appuie sur leurs moyennes des notes et non sur leurs potentiels, ni leurs envies, et encore moins sur la réalité du marché du travail. Cette méthode présente des insuffisances qui ont été démontrées depuis longtemps par diverses études, car les notes seules ne peuvent satisfaire sans confusion les fonctions d'orientation, de régulation ou de certification, endossées par l'évaluation aux différents moments d'un apprentissage (Maccario, 1987). Il y a lieu de signaler aussi que le système d'orientation est conçu pour conférer un monopole à l'administration qui n'en assure pas toujours un service de qualité, au lieu d'être porté par tous les concernés : élèves, parents, enseignants. D'autant plus que l'orientation scolaire est une thématique importante pour les parents dans laquelle ils se disent très impliqués (Masson, 1997).

Dès lors, la mise en place de mécanismes renforçant les critères de l'orientation scolaire est plus qu'une nécessité. Ainsi, le concept d'une éducation au choix est introduit afin de donner aux élèves méthodes et connaissances pour les aider à devenir acteurs de leur orientation (Pugin, 2008).

Aussi, le sentiment d'efficacité personnelle et les attentes de résultats dans le comportement humain jouent un rôle très important, notamment au sujet des activités que les personnes choisissent de réaliser ou d'éviter (Bandura, 1986). Les fondements de cette théorie ont été testés et approuvés à travers la TSCOSP¹ et ont conclu que les sentiments d'efficacité personnelle rendent les personnes capables de donner une direction à leur orientation scolaire et professionnelle (W Lent, 2008). Le sentiment d'efficacité personnelle se traduit aussi par une opinion positive. En effet, la théorie de l'opinion peut donner la possibilité de comprendre l'importance de la prise en compte des opinions des élèves dans le processus de l'orientation pour l'acceptabilité des choix qui en découlent (Tarde, 1989). De ce fait, les opinions des élèves vis-à-vis des matières étudiées et de leurs apports dans le choix du domaine d'études, peuvent

¹ TSCOSP: Théorie Sociale Cognitive de l'Orientation Scolaire et Professionnelle.

de ce point de vue déterminer la portée future de ce choix dans la réussite scolaire et professionnelle, autrement dit son adoption et son acceptabilité par les élèves. Dans nos travaux, nous nous intéressons à l'amélioration de l'orientation scolaire notamment par l'intégration de l'opinion comme critère de prise de décision.

2. Problématique

Les recherches menées sur l'orientation scolaire ont montré que le fait de se baser uniquement sur les notes dans ce processus est insuffisant pour répondre aux attentes des élèves (Maccario, 1987). Ce processus n'associe pas toujours toutes les parties concernées, les parents d'élèves se retrouvent totalement à la marge de ce processus. L'administration s'accapare la totalité de ce processus sans offrir pour autant un service de qualité (Masson, 1997). Les élèves aussi se retrouvent presque à l'écart de ce processus, car ils ne peuvent y contribuer que par leurs évaluations scolaires. Ni leurs avis, encore moins leurs aptitudes, ne sont pris en compte dans ce processus qui est un pivot dans leur vie. Le processus d'orientation n'assure pas toujours l'équité attendue par ses acteurs. Aussi, l'opinion des divers intervenants n'est pas prise en compte dans ce processus, alors qu'elle représente un paramètre efficace dans les systèmes de prise de décisions. Cependant, l'opinion est un élément versatile, parfois il devient difficile de s'appuyer sur cet élément à son état brut. Il nécessite un raffinement pour être plus crédible.

3. Objectifs

Afin de répondre à la problématique posée ci haut, notre thèse se fixe les objectifs suivants :

- Enrichir les critères mis en œuvre dans les processus de l'orientation scolaire afin de permettre aux élèves de suivre un parcours scolaire ou professionnelle répondant aux mieux à leurs aspirations.
- Introduire l'opinion des élèves en tant que critère explicite dans le processus d'orientation.
- Introduire le style d'apprentissage des élèves, que nous désignons par aptitude, afin de mieux cerner les profils de compétences des métiers les plus propices.
- Introduire les opinions d'acteurs majeurs dans ce processus, tels que les parents et les enseignants.
- Proposer un moyen de mettre en évidence la crédibilité de l'opinion.

- Offrir aux acteurs du processus de l'orientation scolaire une plate-forme pour la génération automatisée de cartes d'orientation.

4. Notre contribution

Afin d'atteindre les objectifs fixés et remédier ainsi à la problématique soulevée, nous proposons un système de e-orientation basé sur l'analyse d'opinion. Son objectif principal est d'améliorer la qualité de l'orientation à la fin de cycles scolaires à l'instar du choix des domaines d'études universitaires. Il se base à la fois sur les opinions des différents intervenants à savoir : les élèves, leurs enseignants et leurs parents, ainsi que les aptitudes et les évaluations scolaires des élèves. Ces opinions sont recueillies périodiquement, grâce à des enquêtes menées auprès des différents intervenants. Ces enquêtes ont pour but de capter l'opinion formatée, à travers des questions fermées à échelle d'évaluation, et l'opinion libre, à travers des questions ouvertes. Les opinions ainsi recueillies feront l'objet d'analyse grâce au calcul statistique et technique d'analyse de sentiments pour extraire la polarité de l'opinion attisée à l'égard de chaque matière suivie par les élèves. Ces opinions seront ensuite amplifiées par des coefficients de certitude afin de renforcer leur crédibilité. Elles seront synthétisées afin d'aboutir à une valeur d'opinion par domaine. Ces valeurs d'opinion, conjuguées aux évaluations scolaires ainsi qu'aux aptitudes, représentées par les intelligences multiples de Gardner, seront exploitées dans le calcul du score d'orientation comme base d'élaboration des cartes d'aide à l'orientation. Ces cartes serviront à guider les élèves vers les domaines d'études qui leurs sont le plus favorables.

5. Plan de la thèse

Cette thèse est composée de deux parties : la première concerne l'étude bibliographique et la seconde porte sur nos propositions et leurs validations.

- **Partie A : Etude Bibliographique**

Cette partie fournit un état de l'art des différents concepts sur lesquels reposent nos propositions. Il est composé de deux chapitres.

- **Chapitre I. L'analyse d'opinion** : Ce premier chapitre est consacré à la présentation de généralités sur l'analyse d'opinion. Nous introduisons particulièrement les approches et les modèles utilisés dans ce domaine.

- **Chapitre II. L'orientation scolaire :** Dans ce second chapitre, nous mettons l'accent sur l'orientation scolaire en général et l'E-orientation en particulier. Nous évoquons les contraintes de cette discipline.

- **Partie B : Propositions et leurs validations**

Dans cette partie nous présentons nos propositions et leurs validations. Elle est composée de trois chapitres.

- **Chapitre III. Propositions d'un système de l'E-orientation basé sur l'analyse d'opinion :** Nous présentons dans ce chapitre nos propositions relatives à l'intégration de l'analyse d'opinion dans le processus de l'orientation scolaire. Nous introduisons les différents modules constituant le système proposé dénommé OPinOR (OPinion pour l'ORientation).
- **Chapitre IV. Expérimentation et validation de nos propositions :** Dans ce chapitre nous validons nos propositions via deux prototypes. Nous décrivons les expérimentations conduites.
- **Chapitre V. Discussion et impact de nos propositions :** Ce dernier chapitre, est dédié à la synthèse et à la discussion de nos propositions. Aussi il met en exergue les points forts et les faiblesses de nos propositions. Il présente enfin les différents impacts de nos propositions.

Partie A : Etude Bibliographique

Chapitre I : Analyse d'opinion

Introduction

L'analyse d'opinion est l'une des grandes réalisations de la dernière décennie dans le domaine des technologies langagières. La prise en compte de l'opinion dans les processus de prise de décision est devenue un élément crucial de réussite. Cette technologie se voit alors introduite de manière très répandue dans de nombreux domaines. Dans ce premier chapitre, nous présentons l'analyse d'opinion, un élément pivot de notre travail.

I.2 L'analyse d'opinion- Définitions

I.2.1 Fouille de texte & fouille de données

La fouille de textes, ou text mining en anglais(TM), est l'extraction de connaissances à partir de textes en langage naturel. La fouille de données, Data mining en Anglais (DM) est la base de la fouille de textes au sens où cette dernière est l'extension du même but et du même processus vers des données textuelles. En général le DM travaille sur des données structurées et stockées dans des bases de données. En revanche, le TM travaille sur des données textuelles non structurées. Donc l'objectif de la fouille de textes est le traitement de grandes quantités d'information qui sont disponibles sous une forme textuelle non structurée (Polanco, 2001). Notons, que la fouille d'opinion est un sous domaine de la fouille de texte.

I.2.2 La fouille d'opinion (Opinion mining)

La fouille d'opinions est un problème de traitement automatique de textes. Lorsqu'on parle de fouille d'opinion, on s'intéresse au traitement automatique des opinions, des sentiments et de la subjectivité dans les textes. La fouille d'opinions relativement à la fouille de textes ou de données est un problème plus difficile. En effet, il est plus facile de dire de manière automatique « Ce document parle de ... » (extraction de faits) que « Ce document dit que ... » (extraction d'opinions). La fouille d'opinions est un domaine de recherche devenu populaire depuis une dizaine d'années. Une explication à ce succès est l'exploitation de plus en plus poussée de l'apprentissage artificiel dans le traitement de l'information, notamment la catégorisation de textes. Une autre explication est l'avènement de l'internet social et les intérêts socio-économiques qu'il soulève (Pang & Lee, 2008; Sebastiani, 2002).

I.2.3 Définition

L'analyse d'opinion, appelée en anglais «Sentiment Analysis ou Opinion Mining ou encore subjectivity analysis », est le domaine d'étude qui analyse les opinions, les sentiments, les évaluations, les attitudes et les émotions des gens exprimés dans un texte envers des entités telles que les produits, les services, les organisations, les individus, les questions, les événements, les sujets et leurs attributs. Il représente un grand espace de problèmes. Son objectif est de détecter des informations subjectives contenues dans diverses sources et de déterminer l'état d'esprit d'une personne à l'égard d'un problème, d'une question...etc (Montoyo, et al., 2012; Medhat, et al., 2014). Notons que dans la réalité, les mots opinion, sentiment, point de vue, et croyance sont utilisés de façon interchangeable, alors qu'il y a des différences subtiles entre eux comme nous le voyons à travers les définitions suivantes :

- **L'opinion** représente la manière de penser sur un sujet en donnant un jugement personnel mais qui n'implique pas que ce jugement soit obligatoirement juste. On peut dire qu'une opinion est une pensée encore ouverte à la contestation. Exemple : « je pense que les mathématiques sont difficiles».
- **Le sentiment** indique l'expression d'une émotion ressentie, c'est une expression qui ne peut être que personnelle.
- **Le point de vue** est une opinion subjective très affirmée.
- **La croyance** dénote l'acceptation délibérée et l'assentiment intellectuel. Exemple : Une croyance ferme, qui élimine tout espoir de changement dans la plate-forme de son parti.

Pour simplifier la présentation, tout au long de ce document, nous utilisons le terme opinion pour désigner les termes ci-dessus

La source de données pour l'analyse d'opinion peut provenir de deux canaux :

- Canal online comme les critiques, les blogs, les commentaires.
- Canal offline comme les articles de presse, sondage d'opinions. etc.

L'analyse d'opinion peut être étudiée selon trois axes principaux à savoir : les tâches de l'analyse d'opinion, les approches qu'elle met en œuvre et ses applications comme illustré dans la figure 1. Ces trois axes sont décrits dans les sections ci-après.

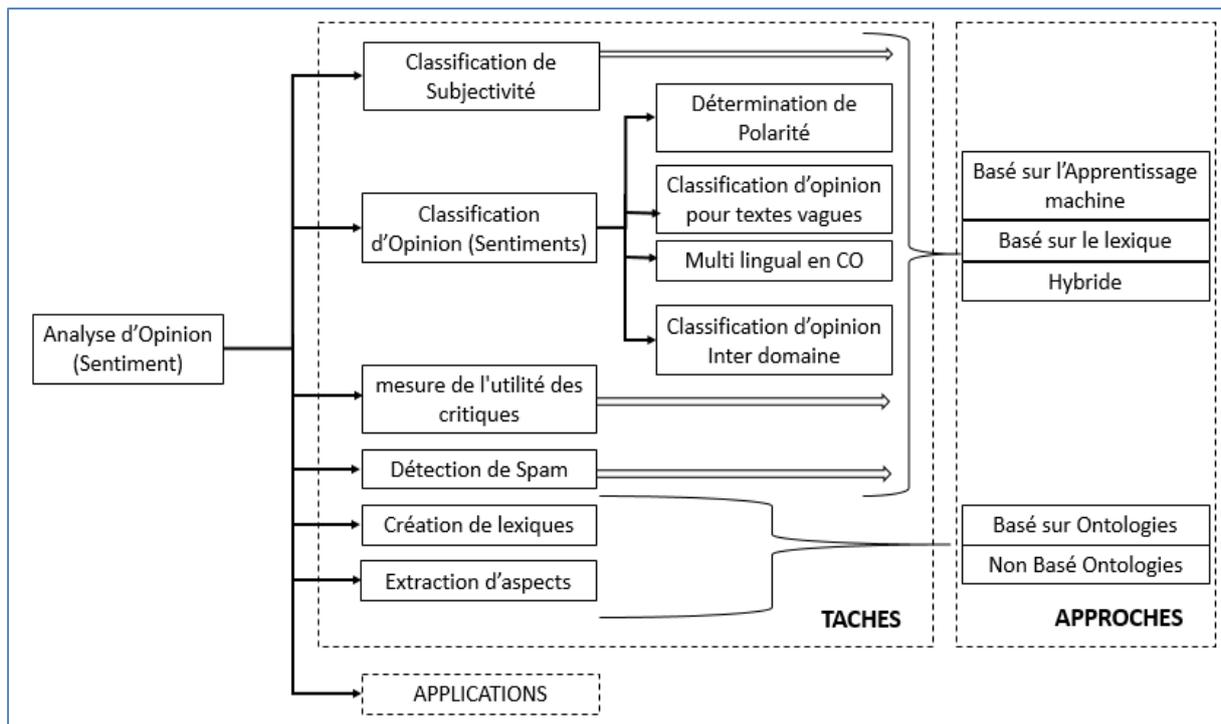


Figure 1 Axes de l'Analyse d'opinion (Ravi & Ravi, 2015).

I.3 Les Tâches de l'analyse d'opinion

L'analyse d'opinion se compose de plusieurs tâches, qu'il est utile ou non de mettre en œuvre selon les applications visées (Marchand, 2015). Les tâches de l'analyse d'opinion peuvent se définir par la classification de la subjectivité (Riloff & Wiebe, 2003), La classification du sentiment (Pang, et al., 2002) et des tâches complémentaires telles que l'extraction du porteur d'opinion ainsi que l'extraction d'entités et d'aspects de l'opinion (Yi, et al., 2003). Nous détaillons dans ce qui suit les tâches principales de l'analyse d'opinion comme illustrées dans la figure 1.

I.3.1-Classification de subjectivité

La classification de la subjectivité concerne la détection des "États privés" - un terme qui englobe les sentiments, les opinions, les émotions, les évaluations, les croyances et les spéculations (Montoyo, et al., 2012). L'analyse de subjectivité est définie comme la reconnaissance d'un langage orienté sur l'opinion afin de le distinguer d'un langage objectif. Le problème lié à la distinction entre instances subjectives et instances objectives s'est souvent révélé plus difficile que la classification ultérieure de la polarité. Par conséquent, les améliorations de la classification de subjectivité devraient avoir un impact positif sur la classification des sentiments. La classification de subjectivité est réalisée à l'aide d'approches

basées sur l'apprentissage automatique (Wang, et al., 2011; Pang & Lee, 2004) et d'approches lexicales ou encore des approches multilingues (Banea, et al., 2014; Molina-González, et al., 2013; Bravo-Marquez, et al., 2014; Xuan, et al., 2012).

I.3.2 Classification d'opinion

La classification d'opinion est la détermination de l'orientation des opinions d'un texte donné dans deux classes ou plus. Elle s'effectue dans différentes classes : binaire, ternaire, n-aire sous la forme d'étoiles (Zhang, 2008) et «pouce levé» ou «pouce baissé» (Turney, 2002; Pang, et al., 2002). Nous reviendrons plus en détail, sur les approches mises en œuvre pour cette tâche dans la section I.5.

I.3.2.1 Détermination de la polarité

La classification des opinions concerne la détermination de la polarité d'une phrase, qu'elle exprime un sentiment positif, négatif ou neutre à l'égard d'un sujet. Par conséquent, la classification de sentiments est également appelée détermination de la polarité. La détermination de la polarité a été effectuée pour les critiques de produits, les forums, les blogs, les articles de presse et les micro-blogs (Ravi & Ravi, 2015).

I.3.2.2 Classification d'opinion pour texte vague

L'ambiguïté et l'imprécision sont considérées comme des problèmes majeurs dans le domaine de l'analyse d'opinion. En effet, les critiques et les commentaires des internautes sont souvent rédigés dans un style flou par rapport aux textes standards. Ils expriment souvent un sarcasme (moquerie ou ironie), une rhétorique ou une métaphore. La détection d'une telle expression recourt généralement à des techniques reposant sur l'approche floue (Ravi & Ravi, 2015).

I.3.2.3 Texte multilingue et analyse d'opinion

Différentes langues à travers le monde ont des degrés différents de pouvoir expressif concernant les opinions. L'analyse d'opinion multilingue peut être réalisée à l'aide de deux approches différentes (Mihalcea & Banea, 2007) :

- Approche basée sur le lexique, dans laquelle un classifieur de subjectivité pour le langage cible est généré en traduisant le lexique existant dans une autre langue.
- Approche basée sur un corpus, où un corpus annoté de subjectivité pour la langue cible est construit par projection, formant un classifieur statistique sur le corpus résultant.

I.3.2.4 Classification d'opinion inter-domaines

L'analyse d'opinion inter-domaines est devenue un problème de recherche intéressant. Cette tâche est assez rude en raison de la grande fluctuation de la subjectivité selon les domaines. La classification de l'opinion inter-domaines nécessite au moins deux domaines : le domaine source sur lequel un classifieur doit être entraîné et le domaine cible sur lequel il doit être testé. Les études menées dans ce sens ont été réalisées à l'aide d'approches basées sur le lexique (Cho, et al., 2014; Weichselbraun, et al., 2013), sur l'apprentissage automatique (Tan, et al., 2009) ou encore sur les approches hybrides (Bollegala, et al., 2013).

I.3.3 Mesure de l'intérêt des critiques

À l'ère de l'information contemporaine, les responsables du marché des affaires s'intéressent de plus en plus à la promotion de leurs produits et services. Pour ce faire, certains d'entre eux engagent de faux utilisateurs pour rédiger de fausses critiques (Tsytsarau & Palpanas, 2012; Liu, 2012). Bien que ces fausses critiques n'aient pas d'effet à long terme sur la vente du produit, elles peuvent, tout de même, amplifier la vente initiale. Ainsi, la mesure de l'utilité des critiques ont suscité tant d'attention de la part de la communauté des chercheurs (Liu, 2012). Le traitement de la mesure de l'utilité des critiques se base sur les approches de l'apprentissage automatique (Huang & Yen , 2013; Ngo-Ye & Sinha, 2014; Krishnamoorthy, 2015) et des approches hybrides (Min & Park, 2012; Purnawirawan, et al., 2012).

I.3.4 Détection de spam d'opinion

La popularité croissante du commerce électronique et des critiques en ligne a généré, comme indiqué précédemment, le phénomène de la rédaction de fausses critiques par des professionnels dits spammeur (Heydari, et al., 2015). Ce phénomène peut être entretenu par un personnel de marketing, un fabricant, un fournisseur de services, un leader ou un producteur de films, dans le but de promouvoir un produit de mauvaise qualité ou pour discréditer un produit de bonne qualité. La majeure partie des techniques de détection de spam dans l'opinion dépend de trois fonctionnalités liées à la fausse critique à savoir (Liu, 2012) :

- Le contenu de la critique donné généralement par le texte de cette dernière ;
- Les métadonnées de la critique représentées par l'identifiant utilisateur, l'adresse IP, la géolocalisation, la date et l'heure d'écriture, le nombre d'étoiles, ...etc. ;
- La connaissance réelle du produit critiqué par exemple une critique qui rabaisse une marque connue pour montrer une marque à peine connue, peut être considérée comme un spam.

I.3.5 Création de Lexique d'opinion

Les mots de sentiments, ou lexiques d'opinion, sont utilisés par les individus, pour exprimer des opinions positives ou négatives. Par exemple, des mots tels que «bon», «beau» et «étonnant» induisent un sentiment positif chez l'humain et des mots comme «mauvais», «laid» et «effrayant» sont les mots à polarité négative (Liu, 2015; Liu, 2012). La polarité de mot désigne le sentiment et l'évaluation que le mot induit dans l'esprit. Il convient de noter que la majorité des mots de sentiments sont principalement des adjectifs et des adverbes. Bien que des noms tels que «corbeille» et «ordures» et des verbes comme «détester» et «aimer» ont aussi un sentiment (Liu, 2012). Dans le lexique de l'analyse d'opinion, chaque mot de sentiment est accompagné de la polarité de ce mot. Ce lexique peut être pondéré ou non. Généralement, ce poids est une probabilité pour que ce mot indique le niveau de positivité ou de négativité. Par conséquent, une autre tâche de l'analyse d'opinion consiste en la création de lexique d'opinion. Un lexique, dans ce cas, est un vocabulaire de mots d'opinion avec des valeurs de polarités respectives.

La création d'un lexique d'opinion peut se faire manuellement ou automatiquement. En effet, ce lexique peut être construit manuellement, comme par exemple pour General Inquirer.3. Cependant, la construction manuelle d'un lexique d'opinion prend beaucoup de temps et rend difficile l'obtention d'une couverture élevée pour la langue considérée. Par exemple, le lexique de General Inquirer ne définit qu'environ 3 600 entrées. Cette approche souffre également de biais d'annotateur : le même annotateur peut attribuer au même mot différentes polarités d'opinion à des moments différents, ou encore, des annotateurs différents peuvent attribuer au même mot différentes polarités d'opinion (Hung & Lin, 2013).

Un lexique d'opinion peut, également, être construit de manière automatique soit par des approches par induction lexicale ou bien par des approches basées sur des ontologies.

I.3.5.1 Approches par induction lexicale

Dans ces approches, une liste initiale de mots, également appelée mots d'origine (seed en anglais) est construite manuellement, associés à des polarités d'opinion non ambiguës. Cette liste est étendue à l'aide des synonymes et des antonymes des mots d'origine issus de dictionnaires en ligne tel que WordNet (Miller, 1995) ou de corpus. Ces approches restent tributaires de la taille de la liste initiale qui délimite significativement la couverture du lexique obtenu. Ces approches sont décrites ci-dessous :

- **Approche basée sur un dictionnaire**

L'utilisation de la méthode basée sur un dictionnaire pour compiler les opinions est

courante. Puisque les dictionnaires contiennent essentiellement des listes de synonymes. La procédure de cette méthode est comme décrit précédemment :

- Premièrement, afin de créer un petit ensemble, quelques mots d'opinion seront identifiés manuellement et auront une orientation sémantique (polarité) positive ou négative.
- Ensuite, l'algorithme aide à étendre cette collection en effectuant une recherche dans WordNet et d'autres dictionnaires en ligne pour trouver des synonymes et des antonymes (Liu, 2012).
- Ce processus se poursuit jusqu'à ce qu'aucun nouveau mot ne puisse être détecté.

Dans (Qiu, et al., 2011; Xia, et al., 2015), on peut voir quelques exemples d'utilisation de cette méthode. Dans (Pandarachalil, et al., 2015), elle a été utilisée pour analyser les sentiments de Twitter.

- Approche basée sur le corpus

Dans le domaine de l'analyse d'opinion, l'orientation sémantique d'un mot, dite aussi polarité, dépend du domaine. La création d'un ensemble de mots d'opinion propre à un domaine suppose que les polarités de ces mots soient appropriées au domaine. L'approche la plus indiquée pour ce faire est celle connue sous le nom d'approche basée sur un corpus (Molina-González, et al., 2015). Cette méthode repose sur des règles syntaxiques. L'idée principale de cette méthode a été proposée par (Hatzivassiloglou & McKeown, 1997), dans laquelle une liste d'adjectifs contenant les polarités est établie, puis les nouveaux adjectifs et leurs polarités sont déterminés à l'aide d'autres règles syntaxiques. Quelques règles sont également appliquées sur les termes de connectivité, telles que : 'et', 'mais', etc. appelées cohérence des opinions. L'une de ces règles concerne le mot de conjonction «et», de sorte que la jonction des adjectifs ne change pas dans la polarité initiale.

Exemple : Considérons la phrase : “*ce film est bon et attrayant*” ; Le mot “et” a fait que deux adjectifs “bon” et “attrayant” se rejoignent, ce qui selon cette règle, lorsque le mot “bon” est connu comme positif, «Attrayant» est également considéré comme positif (Liu, 2012).

L'absence de sources linguistiques pour d'autres langues autres que l'anglais, est l'une des problématiques les plus importantes dans le domaine de l'analyse d'opinion. Le

développement de lexiques est essentiel pour la classification d'opinion dans d'autres langues. Les auteurs dans (Molina-González, et al., 2015) ont abordé l'élaboration d'un lexique d'opinion et incorporé des informations de domaine pour les systèmes de classification en langue espagnole. Leurs résultats ont indiqué l'intégrité du lexique construit. Les auteurs dans (Liao, et al., 2016) ont proposé une nouvelle approche qui intègre un lexique de domaine avec des groupes de caractéristiques utilisant la syntaxe et la sémantique. Les résultats de leurs expériences sur des données réelles ont montré que l'approche proposée surpassait d'autres bases d'extraction de cibles d'opinion. L'approche basée sur le corpus aide à résoudre le problème de la recherche de mots d'opinion avec un contexte d'orientation spécifique, mais ses performances varient dans différents domaines, car un mot peut être positif dans un domaine et négatif dans un autre.

Il a été prouvé que la méthode basée sur les dictionnaires est plus utile car il est difficile d'assurer une grande couverture de la langue anglaise bien que celle-ci reste la mieux étudiée de par le monde (Liu & Zhang , 2012).

I.3.5.2 Approches fondées sur les ontologies

Les ontologies sont des spécifications explicites de conceptualisations, fournissant une représentation formelle de la connaissance, permettant, ainsi, de raisonner (Gruber, 1995). Les ontologies sont considérées meilleures que les taxonomies ou les systèmes de gestion de bases de données relationnelles, car elles capturent également l'association sémantique entre les concepts et les relations. Par conséquent, l'analyse d'opinion s'oriente rapidement vers des approches ontologiques pour représenter la base de connaissances du sens commun des mots du langage (Balahur, et al., 2012; Cambria, et al., 2010; Trainor, et al., 2014).

I.3.6 Extraction d'aspects

Une des tâches principales de l'analyse d'opinion est l'extraction de ses aspects (Rana & Cheah, 2016). La classification par niveau d'aspect offre de meilleures performances mais une condition préalable pour l'utiliser, c'est obtenir d'abord, ces aspects. La plupart des recherches dans le domaine de l'extraction d'aspects ont été concentré sur les revues en ligne (Lv, et al., 2017). Les techniques d'extraction d'aspects de l'opinion peuvent être classées en quatre catégories (Hemmatian & Sohrabi, 2017) :

- Extraction basée sur la fréquence des phrases nominales et des noms ;
- Extraction basée sur l'exploitation des relations entre les mots d'opinion et les aspects ;

- Extraction basée sur les techniques d'apprentissage supervisé ;
- Extraction basée sur la modélisation du thème de l'opinion.

I.4 Les niveaux d'analyse d'opinion

Toutes les tâches de l'analyse d'opinion peuvent se pratiquer à différents niveaux selon les applications envisagées :

- **Le niveau d'analyse de document** qui consiste à déterminer la polarité de l'opinion sur un document donné dans sa totalité (Pang, et al., 2002).
- **Le niveau d'analyse de phrase** qui détecte des phrases subjectives dans un document à partir d'un mélange de phrases objectives et subjectives. Puis, déterminer la polarité de chaque phrase subjective. (Ravi & Ravi, 2015).
- **Le niveau d'analyse d'entités et d'aspects** qui effectue une analyse plus fine. Au lieu de se contenter à examiner les éléments de langage tels que documents, paragraphes, phrases, le niveau d'aspect regarde directement l'opinion elle-même. Il est basé sur l'idée que l'opinion se compose d'un sentiment positif ou négatif, et d'une cible d'opinion (Hemmatian & Sohrabi, 2017).

Il y a aussi des niveaux plus élémentaires comme le niveau de mot, niveau de concept, basé sur les liens, niveau de clause et niveau de sens (Ravi & Ravi, 2015).

I.5 Les approches de classification dans l'analyse d'opinion

Il existe trois approches pour l'analyse d'opinions qui donnent lieu à des techniques variées de classification et de traitement de l'opinion. La figure 2 montre une vue d'ensemble sur ces différentes techniques venant compléter les approches de l'analyse d'opinion.

L'étape la plus importante et la plus critique de l'analyse d'opinion consiste à choisir une approche appropriée afin de classer les sentiments. Les méthodes de classification proposées dans la littérature distinguent trois approches pour l'analyse d'opinions (Cf. fig. 1), les approches symboliques, statistiques et hybrides.

I.5.1 Les approches symboliques

Appelées aussi approches basées sur le lexique. Elles sont utilisées pour calculer l'orientation des documents en fonction de la polarité des mots et des phrases dans les

documents. Ces méthodes se basent sur des lexiques d'opinion ou ressources lexicales. Elles utilisent des dictionnaires de mots subjectifs pour lesquels la tonalité (polarité) peut être pré-codée. Ces dictionnaires peuvent être généraux comme par exemple : General Inquirer, SentiWordNet, Opinion Finder, NTU, etc. ou construits manuellement, ou encore, générés automatiquement à partir de corpus (Cf. I.3.5).

Dans ces approches, un lexique d'opinion peut être recherché grâce à l'analyse de texte. De plus, dans certains cas, le traitement du langage naturel est utilisé pour trouver la structure syntaxique et aider à trouver les relations sémantiques (Moreo, et al., 2012; Di Caro & Grella, 2013). Une polarité est associée a priori à chacun de ces mots. Il est donné ensuite au document un score d'opinion calculé en fonction de la présence de mots issus de ces dictionnaires dans le texte (Boullier & Lohard, 2012; Martineau, et al., 2011; Farek, 2009). Les méthodes symboliques reposent généralement sur trois étapes :

- Extraction, à partir du document, des phrases contenant des mots d'opinion ;
- Estimation ou calcul de la polarité de chaque phrase selon une formule ou un modèle arrêté ;
- Classement du document en fonction de la polarité moyenne des phrases.

I.5.2 Les approches statistiques

Les approches statistiques sont des techniques d'apprentissage automatique, en anglais machine learning (ML), qui se base sur des approches mathématiques et statistiques pour donner aux ordinateurs la capacité d' « apprendre » à partir de données. L'apprentissage automatique comporte généralement deux phases. La première dite « d'apprentissage » ou « d'entraînement » intervient lors de la phase de conception du système. Cette phase consiste à concevoir un modèle à partir de données disponibles et en nombre fini. La seconde phase correspond à la mise en production : le modèle étant déterminé, de nouvelles données peuvent alors être soumises afin d'obtenir le résultat correspondant à la tâche souhaitée.

Ces approches regroupent les documents (ou les mots) dans deux axes de classification, soit dans l'opposition (subjectif-objectif), soit dans la distinction des opinions subjectives dans l'opposition (positif-négatif) (Maurel, et al., 2008).

Trois techniques d'apprentissage automatique sont utilisées pour classer les opinions (Hemmatian & Sohrabi, 2017) :

- **Méthodes d'apprentissage supervisées** : Dans l'apprentissage supervisé, le processus d'apprentissage est effectué à l'aide des données d'une formation dans lequel la valeur de sortie est spécifiée pour toute entrée et le système essaie d'apprendre une fonction, en

mappant l'entrée sur la sortie, Avec cette méthode, les classes sont initialement spécifiées et les données d'apprentissage sont attribuées à une classe spécifique. Les méthodes d'apprentissage supervisé reposent sur deux types d'approches de classification (Boullier & Lohard, 2012; Martineau, et al., 2011) :

- **Probabiliste** : Ces méthodes sont dérivées de modèles probabilistes qui offrent un moyen systématique de classification statistique dans les domaines tels que le langage naturel. Par conséquent, il a une application réelle dans l'analyse d'opinion. Naïve Bayes, Réseaux Bayésiens et Entropie Maximale sont les méthodes d'analyse d'opinion les plus connues dans cette catégorie.
 - **Non probabiliste** : Il existe aussi des méthodes d'apprentissage automatique supervisé non probabilistes. Les classifieurs de cette catégorie, les plus connus et les plus utilisés en analyse d'opinion sont : les réseaux de neurone, machine à vecteurs de support (SVM), le plus proche voisin, arbre de décision et les méthodes basées sur des règles.
- **Méthodes semi-supervisées** : Ces approches reposent sur l'utilisation de grande quantité de données non étiquetées, ainsi qu'une petite quantité de données étiquetées, pour créer des classifieurs qui offrent généralement de meilleures performances.
 - **Méthodes non supervisées (clustering)** : Ces méthodes visent à trouver une structure dans les collections de données qui n'ont pas encore été classées et pour lesquelles aucune classe prédéfinie n'existe déjà. En d'autres termes, le regroupement (clustering) consiste à placer des données dans différents groupes dans lesquels les membres de chaque groupe sont similaires entre eux d'un point de vue particulier (Rajput & Solanki, 2016). L'application du regroupement dans l'analyse d'opinion a été introduite dans (Li & Liu, 2012). On note que cette méthode met en œuvre deux types de regroupement (clustering) : clustering de partitions et clustering hiérarchique.

I.5.3 Les approches hybrides

Les approches hybrides combinent les points forts des deux approches précédentes. Elles prennent en compte tout le traitement linguistique des approches symboliques avant de lancer le processus d'apprentissage comme dans les approches statistiques (Hermann , 2010). La combinaison des approches symboliques et statistiques a donné des résultats plus précis que chacune des approches employées séparément (Maurel, et al., 2007).

La figure 2 résume les techniques de classification pour l'analyse d'opinion.

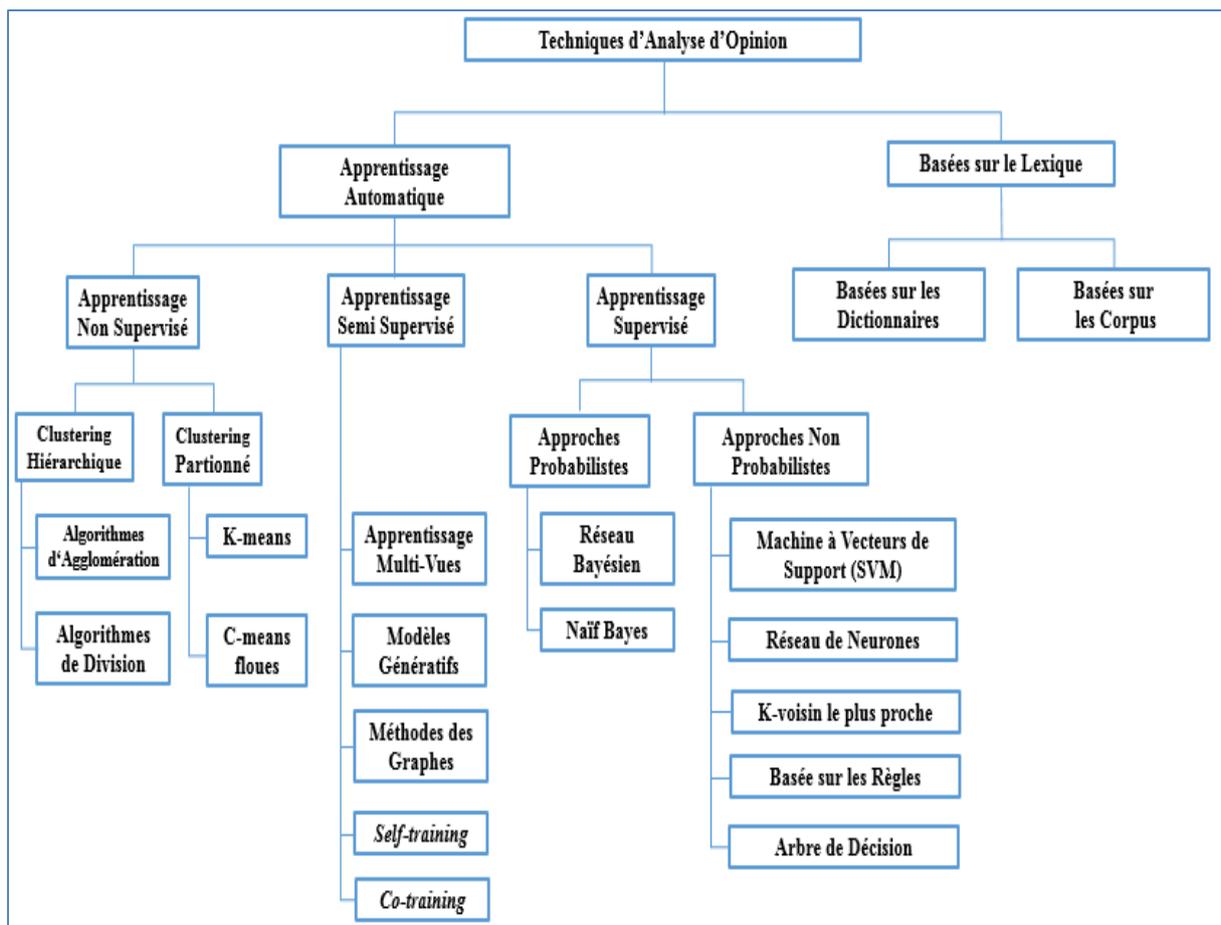


Figure 2 Techniques d'analyse d'opinion (Hemmatian & Sohrabi, 2017).

1.5.4 Comparaison entre les différentes approches d'analyse d'opinion

Il est difficile de comparer entre les différentes approches de classification d'opinion. L'approche basée uniquement sur le lexique a souvent pour le client un coût trop élevé. Les méthodes statistiques sont très utiles, mais non suffisantes. La combinaison des deux approches lexicales et statistiques semble donner des résultats plus précis que chacune des approches utilisées séparément (Maurel, et al., 2007). Nous présentons dans le tableau 1 la comparaison élaborée par (Hemmatian & Sohrabi, 2017) des différentes approches mises en œuvre dans l'analyse d'opinion.

Tableau 1 Comparaison des approches de l'analyse d'opinion².

Approches d'analyse d'opinion		Efficacité	Justesse (Accuracy)	Forces	Faiblesses
<i>Apprentissage Automatique</i>	Apprentissage Supervisé	Lent	Très élevée	Possède la capacité d'analyser de nombreuses catégories.	Dépendance à la formation de documents étiquetés.
				Efficacité dans la classification de la subjectivité.	Nécessite un effort humain considérable et des experts linguistiques. Coût élevé
				Résistant au bruit	Temps d'entraînement élevé et données dimensionnelles élevées.
	Apprentissage Semi Supervisé	Moyenne	Elevée	Bonne performance notamment lors d'ambiguïtés.	Difficulté de classification en présence de bruit.
				Aide à atteindre de plus haute précision avec seulement un petit effort humain d'annotation.	
	Apprentissage Non Supervisé (Clustering)	Rapide	Moyenne	Ne nécessite pas une grande participation humaine.	Le nombre de clusters dans la plupart des cas sont inconnus.
Efficace et largement applicable.				La justesse peut parfois être relativement faible.	
				Manque de résistance au bruit.	
<i>Basée sur le lexique</i>	Basée sur les dictionnaires	Très rapide	Relativement faible	Ne nécessite pas de formation d'échantillons annotés.	Difficulté de trouver des mots d'opinion avec des orientations spécifique à des domaines autre que ceux du dictionnaire.
				Accès facile au lexique d'opinion et leur orientation.	Incompatibilité dans les textes avec une certaine dépendance sémantique.
				Fournit de meilleurs résultats sur des domaines moins segmentés.	Moins précise lors d'analyse de différents domaines.
	Basée sur les corpus	Très rapide	Relativement faible	La capacité de trouver les mots d'opinion avec l'orientation spécifique du contexte.	Performances variables en raison de l'étendue du domaine du lexique.
				Donne de bons résultats lorsque l'analyse porte sur des domaines différents.	Difficulté de fournir des textes volumineux avec la capacité de couvrir tous les mots du texte.
					Ne peut pas être utilisée seule.

² Traduite de la source (Hemmatian & Sohrabi, 2017).

I.5.6 Mesures de qualité d'une classification de textes

L'évaluation de la performance des méthodes de classification de texte et d'analyse d'opinion emprunte des critères de la recherche d'information (Baeza-Yates & Ribeiro-Neto , 1999; Olson & Delen , 2008). Ces critères comprennent selon (Alfaro , et al., 2016), les mesures : de la justesse (accuracy), de la précision, du rappel, du taux d'erreur et du F-score, ayant toujours leurs valeurs comprises entre 0 et 1 et la meilleure performance est celle qui se rapproche de 1.

Pour les problèmes de classification binaire, la mesure de performance recourt à l'utilisation de la matrice de confusion comme illustré dans le tableau 2. Les lignes représentent les valeurs réelles des classes pendant que les colonnes indiquent les valeurs prédites de chaque classe. Les valeurs de Tp et Tn dénotent les nombres des vrais positifs et négatifs correctement classés, pendant que Fp et Fn définissent, respectivement, les vrais négatifs et positifs mal classés (Hossin & Sulaiman, 2015).

Tableau 2 Matrice de confusion pour la classification binaire.

		Classes Prédites	
		Valeurs positives	Valeurs Négatives
Classes Réelle	Valeurs positives	Tp	Fn
	Valeurs Négatives	Fp	Tn

- **Mesure d'exactitude (Accuracy) :**

L'exactitude représente la proportion des résultats bien classés sur l'ensemble des résultats. Elle se calcule grâce à la formule (I.1).

$$E = \frac{Tp + Tn}{Tp + Tn + Fp + Fn} \quad (I.1)$$

- **Mesure de précision :** La précision est une mesure de pertinence de la classification. La précision est le nombre de réponses correctes pour une classe, sur le nombre de réponses fournies pour cette même classe, comme indiqué par la formule (I.2).

$$P = \frac{Tp}{Tp + Fp} \quad (I.2)$$

- **Mesure de rappel** : Le rappel est une mesure de couverture de la classification. Le rappel est le nombre de réponses correctes, pour une classe, sur le nombre de réponses attendues pour cette même classe, comme illustré par la formule (I.3).

$$R = \frac{Tp}{Tp + Fn} \quad (I.3)$$

- **Le taux d'erreur** : Il représente le pourcentage des fausses prédictions en rapport au nombre total de prédictions. La formule (I.4) correspond au calcul de ce taux.

$$TE = \frac{Fp + Fn}{Tp + Tn + Fp + Fn} \quad (I.4)$$

- **F-score** : F-score est la mesure la plus utilisée en classification d'opinions. Son évaluation permet de tenir compte à la fois de la précision ainsi que du rappel. Il se mesure à l'aide de la formule (I.5).

$$F1 = \frac{2 * (P * R)}{(P + R)} \quad (I.5)$$

Le tableau 3 illustre ces critères étendus aux problématiques multi classes.

Tableau 3 Critère de performances pour la multi classification.

N° Formule	Critère	Formule	Description
I.6	Exactitude Moyenne	$E = \frac{\sum_{i=1}^N \frac{Tp_i + Tn_i}{Tp_i + Tn_i + Fn_i + Fp_i}}{N}$	L'exactitude moyenne de toutes les classes.
I.7	Taux d'erreur Moyen	$TE = \frac{\sum_{i=1}^N \frac{Fp_i + Fn_i}{Tp_i + Tn_i + Fn_i + Fp_i}}{N}$	Le taux d'erreur moyen de toutes les classes.
I.8	Précision Moyenne	$P = \frac{\sum_{i=1}^N \frac{Tp_i}{Tp_i + Fp_i}}{N}$	La précision moyenne de toutes les classes.
I.9	Rappel Moyen	$R = \frac{\sum_{i=1}^N \frac{Tp_i}{Tp_i + Tn_i}}{N}$	Le rappel moyen de toutes les classes.
I.10	F1-Score moyen	$F1 = 2 * \frac{(P_N * R_N)}{(P_N + R_N)}$	Le F1-Score moyen de toutes les classes.
N représente le nombre de classes (C ₁ ,... , C _N) ; Tp _i représente les vrais positifs pour la classe C _i ; Tn _i : représente les vrais négatifs pour la classe ; C _i ; Fn _i représente les faux négatifs pour la classe C _i ; Fp _i représente les faux positifs pour la classe C _i .			

I.6 WordNet, WOLF et SentiWordNet

I.6.1. WORDNET

I.6.1.1 Définition

WordNet est une base de données lexicale électronique développée depuis 1985 par des linguistes du laboratoire des sciences cognitives de l'université de Princeton. C'est un réseau sémantique de la langue anglaise, qui contient des noms, des verbes, des adjectifs, et des adverbes (Miller, et al., 1990; Miller & Fellbaum, 1991; Fellbaum, 1998), dont le sens des mots est représenté d'une façon différente de celle des dictionnaires traditionnels. La conception de WordNet est basée sur des théories de la représentation des connaissances mentales qui montrent que les humains enregistrent, dans leur mémoire, les mots et leurs concepts associés d'une façon hiérarchique, en utilisant la relation d'inclusion (Collins & Quillian, 1969). Son but est de répertorier, classier et mettre en relation de diverses manières le contenu sémantique et lexical de la langue anglaise (Chaumartin, 2007; Fellbaum, 1999). Outre la création de WordNet, les chercheurs ont eu également pour objectif de soutenir et de promouvoir l'analyse automatique de textes pour des applications dans le domaine de l'intelligence artificielle (Wilson, et al., 2005).

I.6.1.2 Notion de synset

Le synset (*synonym set*) est la composante atomique sur laquelle repose WordNet. Il est composé d'un ensemble de synonymes, groupes de mots ou de phrases qui expriment le même concept, dénotant, ainsi, un sens ou un usage particulier. Un synset a un identifiant unique et contient donc un certain nombre de littéraux, qui sont approximativement des lemmes (simples ou composés), des termes voire des collocations, qui tous peuvent exprimer le sens représenté par le synset. Les synsets sont reliés entre eux par des relations lexicales et/ou sémantiques (Sagot, 2017). Les noms et verbes sont organisés en hiérarchies. Des relations d'hyponymie (« est-un ») et d'hyponymie relient les « ancêtres » des noms et des verbes avec leurs « spécialisations ». Au niveau racine, ces hiérarchies sont organisées en types de base. L'organisation des adjectifs est différente. Un sens « tête » joue un rôle d'attracteur ; des adjectifs « satellites » lui sont reliés par des relations de synonymie. On a donc une partition de l'ensemble des adjectifs en petits groupes. Les adverbes sont le plus souvent définis par les adjectifs dont ils dérivent. Ils héritent donc de la structure des adjectifs (Chaumartin, 2007; Fellbaum, 1999). La figure 3 présente un aperçu sur la structuration de Princeton WordNet (PWN).

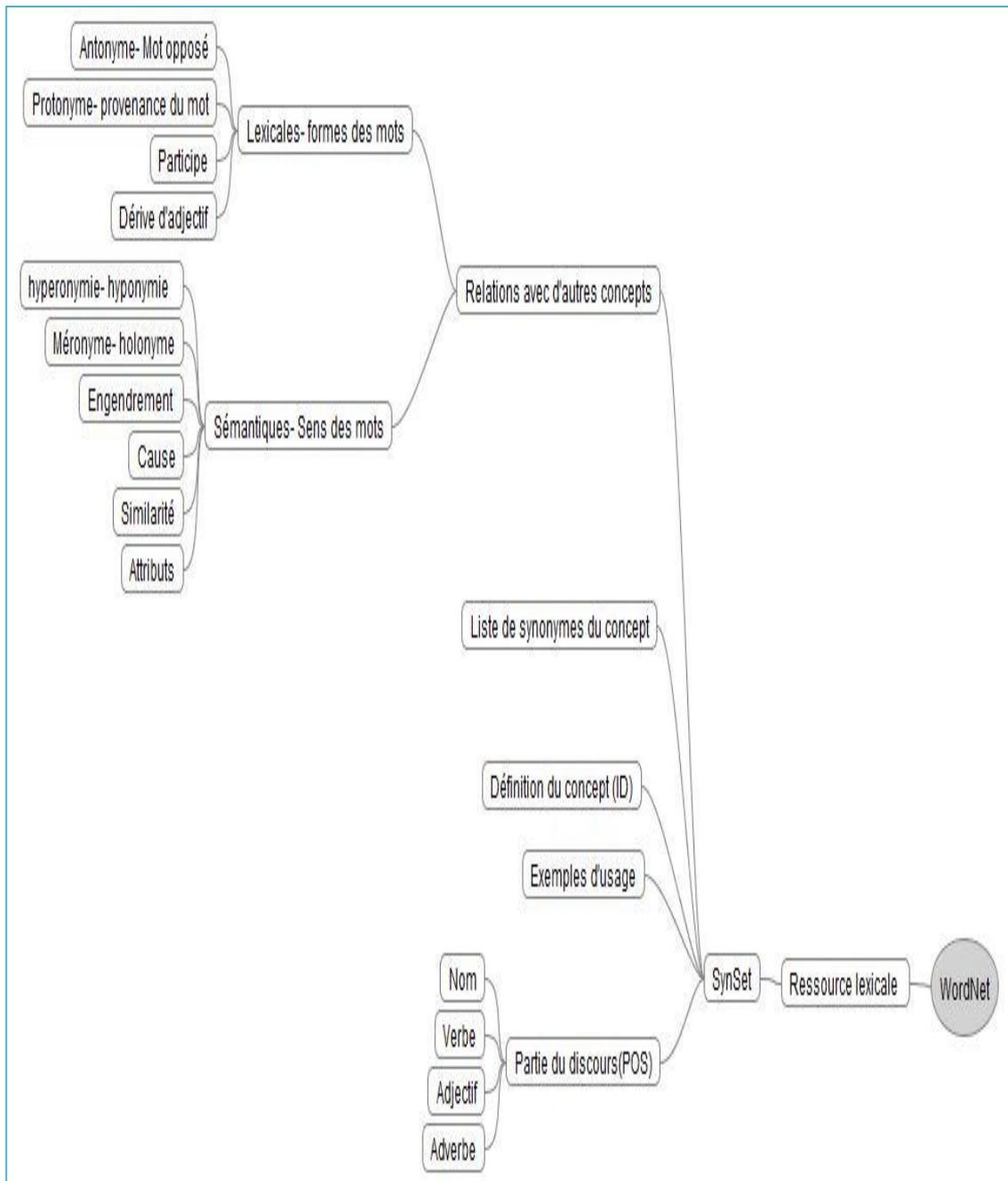


Figure 3 Aperçu sur la structuration de WordNet.

Nous citons comme exemple, dans la version 3.1 du Princeton WordNet (PWN), le synset nominal d'identifiant $ID = "02086723 - n"$ qui contient :

- Les synonymes : *{dog, domestic dog, Canis familiaris}*.
- La définition : *(a member of the genus Canis [...] that has been domesticated by man since prehistoric times; occurs in many breeds);*
- Un exemple d'emploi : *(the dog barked all night).*

Ce synset a deux hypéronymes :

- Le synset : (*ID= "2085998 -n", SYN= {canine, canid } ;*
- Le synset : (*ID="01320032 -n", SYN= {domestic animal, domesticated animal}*).

Il a un certain nombre d'hyponymes, dont par exemple :

- Le synset : (*ID="02089774 -n", SYN= {hunting dog} ;*
- Le synset : (*ID= "02113929 -n" SYN= {Newfoundland, Newfoundland dog}*).

La figure 4 illustre cet exemple.

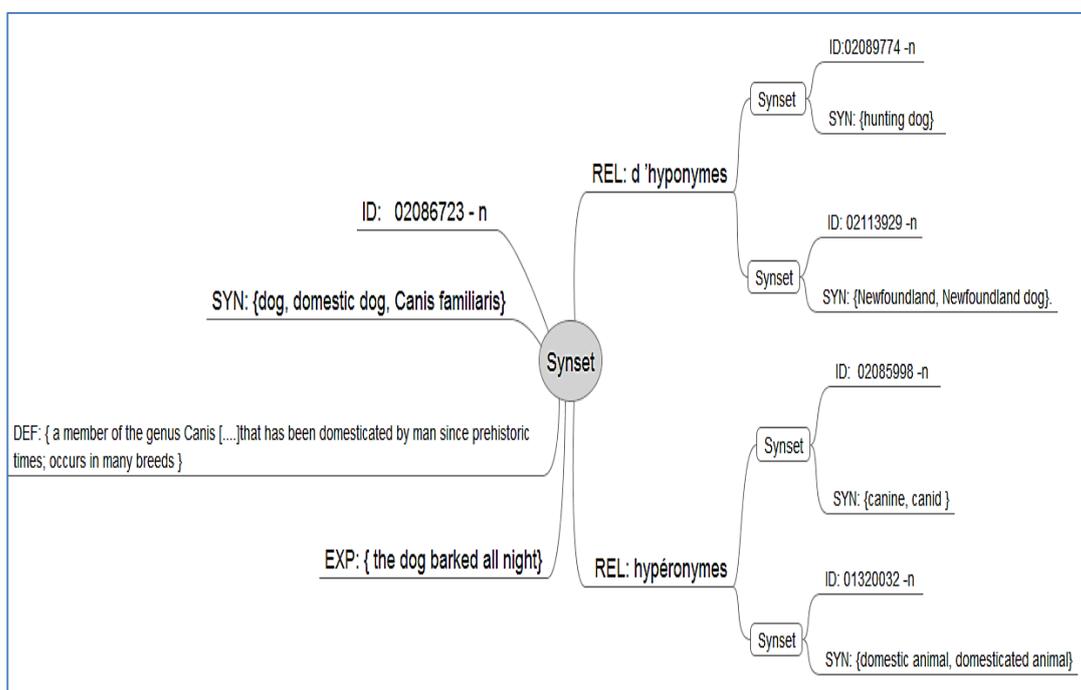


Figure 4 Exemple illustratif du synset (ID : "02086723 - n ") dans PWN.

I.6.1.3. Fréquence des lemmes

WordNet permet d'avoir la fréquence d'apparition pour chaque lemme définissant un synset. Il indique le nombre de fois qu'un mot apparaît dans un sens spécifique (synset). Il est alors possible de calculer pour un nom ou un verbe le contenu informationnel qui est la somme cumulée des fréquences d'un synset et de ses hyponymes au sein d'un sous arbre de la hiérarchie.

I.6.1.4. Mesures de similarité

WordNet permet, à travers l'ontologie qu'il fournit, la définition de métriques heuristiques de « distance sémantique » entre les synsets. Cette métrique quantifie la similarité entre deux concepts (Synset) en mesurant la distance à parcourir dans le graphe. Elle peut être combinée ou non avec le Contenu Informationnel et servir notamment à la désambiguïsation lexicale. (Pedersen, et al., 2004) y présentent une implémentation basée sur WordNet appelée WordNet:Similarity.

I.6.1.5. Mappage entre différentes versions

Il existe une correspondance des identifiants de synsets entre les différentes versions de WordNet. Ce mappage est indispensable pour assurer une traçabilité avec la version la plus récente. Le site Web de Princeton n'offre de mappage « officiel » que pour les noms et les verbes. D'autres sites proposent également des correspondances (construites automatiquement) pour les adjectifs et adverbes.

I.6.1.6- WordNets pour d'autres langues que l'anglais

La construction automatique de WordNet pour d'autres langues peut se faire selon deux approches à savoir :

- **L'approche par fusion** : Dans ce cas, un WordNet pour une langue donnée est créé indépendamment des autres WordNets existants, en exploitant au mieux des ressources monolingues disponibles ; dans un second temps, le WordNet ainsi créé peut être aligné avec les WordNets disponibles pour d'autres langues (Rudnicka, et al., 2012).
- **L'approche par extension** : Pour ce cas, l'inventaire de sens du PWN est conservé (mêmes identifiants de synsets, mêmes relations entre synsets). Cette approche recherche, alors, à peupler les synsets avec des littéraux de la langue cible, par exemple, par désambiguïsation et traduction des littéraux anglais présents dans le PWN (Vossen, 1998; Sagot, 2017).

Nous citons ci-dessous quelques exemples de WordNet construits pour d'autres langues que l'anglais. Nous avons choisi les plus illustres d'entre eux tout en mettant l'accent sur le WOLF (**W**OrdNet **L**ibre **F**rançais).

- **EuroWordNet** : Est une base de données pour plusieurs langues européennes. La phase initiale du projet s'est achevée en 1999, avec la conception de sa base de données, ainsi que la définition des types de relations, d'un haut d'ontologie et d'un Index-Inter-Langues (basé sur la version 1.5 du WordNet de Princeton). EuroWordNet a produit des

WordNets pour plusieurs langues (le néerlandais, l'italien, l'espagnol, l'allemand, le français, le tchèque et l'estonien). Les langues sont reliées ensemble par l'intermédiaire de l'Index-Inter-Langues. Il est ainsi possible de passer des mots dans une langue aux mêmes mots dans n'importe quelle autre langue. EuroWordNet permet donc une recherche d'information monolingue ou multilingue. On peut regretter que, contrairement à la version de Princeton, EuroWordNet ne soit pas distribué librement. Cela explique certainement sa diffusion beaucoup moins importante.

- **BalkaNet** : BalkaNet prolonge la base de données d'EuroWordNet avec d'autres langues européennes : tchèque, roumain, grec, turc, bulgare, et serbe.
- **WordNet Arabe** : L'arabe, langue sémitique parlée par environ 300 millions de personnes dans le monde, est elle aussi concerné par les avancées des WordNets. Les auteurs dans (ElKateb, et al., 2006), ont initié un projet visant à créer une ressource lexicale pour l'arabe moderne standard basé sur le PWN (Fellbaum, 1998). Leur objectif est de développer une ressource linguistique avec une base sémantique formelle profonde en utilisant les méthodes développées pour EuroWordNet (Vossen, 1998; ElKateb, et al., 2006). L'Arabic WordNet est utilisé par de récentes applications de PNL telles que la recherche d'informations, le E-learning en arabe, les applications basées sur la sémantique et la recherche conceptuelle (Regragui, et al., 2016).
- **Open Multi lingual WordNet** (Bond & Paik, 2012) : Il normalise et fusionne tous les WordNets dont la redistribution est autorisée par leurs auteurs respectifs, et inclut à ce jour des WordNets pour 27 langues.

I.6.2 WOLF

WOLF Est une ressource linguistique, développée depuis 2008 pour le français par l'INRIA³. Elle est en licence libre, d'où sa dénomination WOLF ((WOrdnet Libre Français). Elle est construite à partir du Princeton WordNet et de diverses ressources lexicales multilingues. L'approche retenue par (Sagot, 2017) pour la construction du WOLF est l'approche par extension basée sur :

- La construction d'un lexique bilingue ou multilingue à partir de ressources structurées (dictionnaires et lexiques généraux ou spécialisés), de ressources semi -structurées (articles Wikipédia) et de ressources non structurées (corpus alignés).

³ INRIA : [Institut national de recherche en informatique et en automatique](http://www.inria.fr).

- Le rapprochement des lexiques bilingues ou multilingues ainsi construit en procédant à :
 - La désambiguïsation des entrées multilingues issues de corpus alignés ;
 - La désambiguïsation des entrées issues de dictionnaires traitant le cas des entrées monosémiques. Cette étape a donné lieu à la construction de la première version du WOLF ;
 - La désambiguïsation des entrées issues de dictionnaires traitant le cas des entrées polysémiques, produisant l'extension initiale du WOLF ;
 - L'identification automatique et suppression manuelle des entrées erronées.
 - Mappage avec WordNet pour assurer l'alignement des synsets de WOLF à ceux de PWN.

WOLF 0.6 est intégré à la plate-forme Open Word Multilingue ouverte, il constitue une ressource de WordNet multilingue Babel Net (Medagoda, et al., 2015; Sagot, 2017).

I.6.3 SentiWordNet

SentiWordNet (Esuli & Sebastiani, 2006) est une ressource lexicale de la langue anglaise dédiée à l'analyse d'opinion. Il est considéré comme l'un des outils les plus prometteurs pour la détermination du score de sentiment (Baccianella, et al., 2010). SentiWordNet assigne à chaque synset de WordNet 2.0 deux valeurs : Positivité (PosScore), Négativité (NegScore), contenues dans l'intervalle [0,1]. La valeur d'objectivité (ObjScore) est calculée selon la formule : $ObjScore = 1 - (PosScore + NegScore)$. Les polarités respectives de chaque synset respectent l'égalité : $PosScore + NegScore + ObjScore = 1$.

Les synsets de SentiWordNet sont mappés sur ceux de WordNet, en d'autres termes, chaque synset possède le même identifiant dans les deux ressources (Baccianella, et al., 2010).

L'hypothèse qui permet le passage des termes en synsets est que différents sens du même terme peuvent avoir des scores (polarités) différents liées à l'opinion (Miller, 1995; Khan, et al., 2016; Hemmatian & Sohrabi, 2017). Il y a environ 117 569 synsets disponibles qui sont tous uni-grammes. La figure 5 donne un aperçu sur la structure de SentiWordNet.

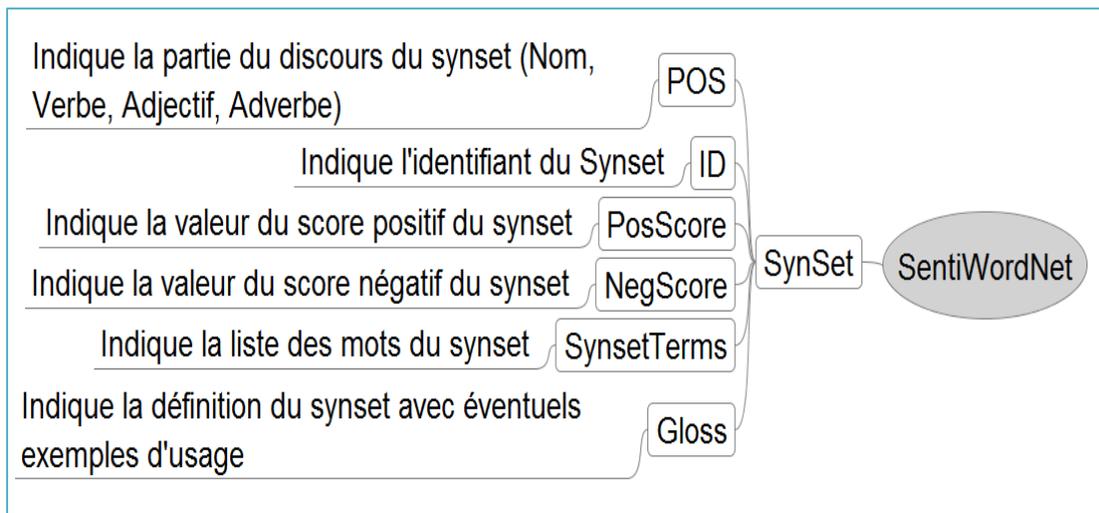


Figure 5 Aperçu sur la structure de SentiWordNet⁴.

Cette ressource a été créée d'une façon semi-automatisée, en mixant des techniques linguistiques et statistiques (utilisation de classifieurs). Avec cette classification, des scores de polarité ont été attribués à chaque synset. Par conséquent, un même mot peut avoir plusieurs scores relatifs à chacun de ses sens. La figure 6 illustre le cas de l'adjectif « estimable ».

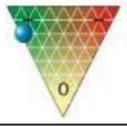
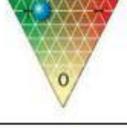
	P = 0 N = 0 O = 1	COMPUTABLE#1 ESTIMABLE#3 <i>may be computed or estimated; "a calculable risk"; "computable odds"; "estimable assets"</i>
	P = 0,75 N = 0 O = 0,25	ESTIMABLE#1 <i>deserving of respect or high regard</i>
	P = 0,625 N = 0,25 O = 0.125	HONORABLE#5 GOOD#4 RESPECTABLE#2 ESTIMABLE#2 <i>deserving of esteem and respect; "all respectable companies give guarantees"; "ruined the family's good name"</i>

Figure 6 Exemple de scores de polarité dans SentiWordNet (Chaumartin, 2007).

I.7 Modélisation de l'opinion

L'analyse d'opinion utilise souvent des procédures de classification de textes ou de phrases dans subjectif / objectif, puis attribue une polarité positive, négative ou nulle aux entités jugées subjectives. Cependant, il est nécessaire de préciser ce qui doit être pris en compte dans l'analyse, en indiquant la composition sémantique de l'opinion. C'est pourquoi il est nécessaire

⁴ https://github.com/aesuli/SentiWordNet/blob/master/data/SentiWordNet_3.0.0.txt

de structurer et de modéliser ce concept afin de faciliter son traitement. Nous présentons dans cette section deux modèles d'opinion qui restent les références dans ce domaine.

I.7.1 Modèle de Liu

Bing Liu (Liu, 2010) de l'Université de l'Illinois à Chicago, a modélisé l'opinion à l'aide d'un vocabulaire décrivant ses attributs, avec la proposition d'un quintuple qui est une référence dans le domaine de l'analyse d'opinion. Ce quintuple (e, a, s, h, t) représente respectivement :

- La cible de l'opinion (entité e) ;
- L'aspect recherché (aspect a) ;
- Le marqueur de l'opinion (Polarité du sentiment s) ;
- L'émetteur de l'opinion (Holder h) ;
- Le moment de l'émission de cette opinion (Temps t).

De ce fait, la tâche de l'analyse d'opinion est décomposée en sous-tâches pour détecter chaque élément de ce quintuple (Lark, et al., 2015).

I.7.2 Modèle de l'Appraisal

La théorie de l'Appraisal a également contribué à la modélisation de l'opinion. Cette théorie est, selon (Gardin, 2009; Martin & White, 2003), la manière dont les évaluations où les sentiments sont exprimées (opinions et réactions émotionnelles). L'Appraisal est représenté sous la forme d'un graphe appelé système comme illustré par la figure 7, qui modélise les différents choix du locuteur dans les fonctions qu'il peut exprimer (significations) et dans la façon dont elles sont exprimées (formulations). Les trois systèmes d'évaluation sont :

- **Attitude** : L'attitude est le système qui décrit les réactions émotionnelles (Affect) et les opinions divergentes selon qu'elles soient liées à des comportements (Jugement) ou à d'autres entités (Appréciation).
- **Engagement** : L'engagement est le système qui concerne le degré d'acceptation par le locuteur de sa proposition, ce qui inclut le discours rapporté.
- **Gradation** : La gradation est le système qui rend compte de l'intensité des évaluations exprimées, ainsi que du degré de modulation des propositions par l'engagement. Ces deux sous-systèmes sont :
 - La Force, qui correspond à une gradation en fonction de l'intensité ou de la quantité ;
 - Le Focus, qui modélise le degré de prototypage.

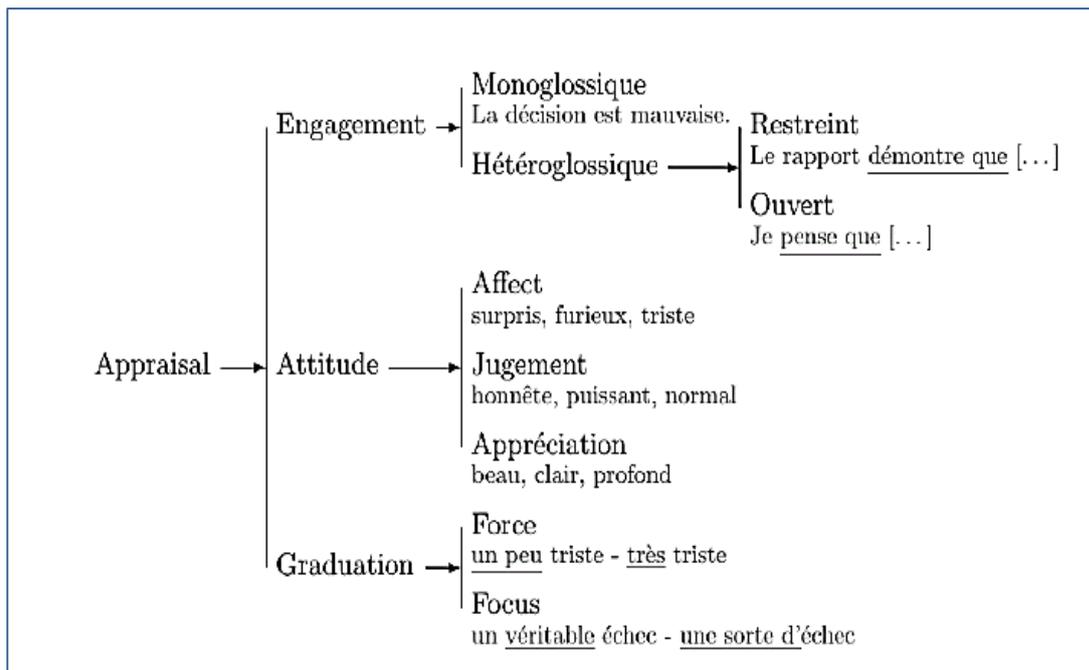


Figure 7 Vue globale du système de l'Appraisal (Zhang & Ferrari, 2014).

I.8 Application de l'analyse d'opinion

L'analyse d'opinion peut trouver de nombreuses applications dans le domaine de la prédiction et de la supervision. Elle est développée pour répondre principalement à la question : "Qu'est-ce que les autres pensent ?". La réponse à cette question est un élément d'information crucial dans les processus de prise de décision que ce soit dans le monde politique, économique, culturel ou social. Nous présentons dans cette section les applications majeures de l'analyse d'opinion comme illustré dans la figure 8.

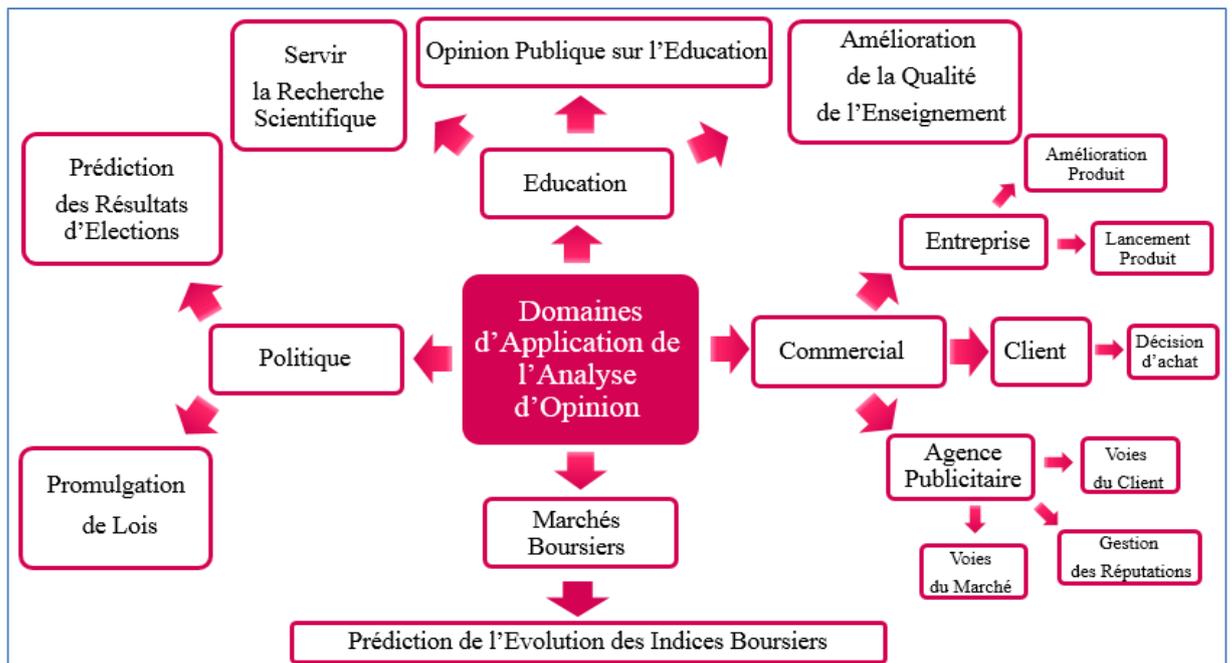


Figure 8 Domaines d'applications de l'analyse d'opinion.

I.8.1 L'analyse d'opinion dans le domaine commercial

L'utilisation de l'analyse d'opinion dans le domaine commercial (Chen, et al., 2014; Marrese-Taylor, et al., 2014; Jeyapriya & Selvi, 2015; Li, et al., 2012; Luo, et al., 2015) est importante à trois points de vue:

I.8.1.1 Le point de vue du client individuel

Lorsqu'un client veut acquérir un produit, il recherche les avis d'autres clients ayant déjà acheté celui-ci. Il devient alors utile de résumer les opinions de ces derniers et de les fournir au client afin de l'aider à prendre sa décision. Le client pourra ainsi, comparer facilement les produits ayant un résumé d'opinions.

I.8.1.2 Le point de vue des entreprises

L'analyse d'opinion permet aussi aux entreprises d'améliorer leurs produits. Les informations qui en découlent, sont utilisées aussi bien pour le marketing et l'évaluation du produit, que pour la conception et le développement du produit. Les entreprises de fabrication peuvent même augmenter, diminuer ou modifier leur production, en fonction, des opinions de leurs clients (Hemmatian & Sohrabi, 2017). L'utilisation la plus courante de l'analyse d'opinion se trouve, généralement, dans les entreprises du commerce électronique. En effet, ces dernières, à travers leurs sites web, invitent leurs utilisateurs à faire part de leurs expériences en matière d'achat de leurs produits ainsi que ses qualités. Par exemple, <http://www.flipkart.com> est un site web de vente en ligne où les utilisateurs évaluent de manière critique les produits qu'ils ont

achetés (Rajput & Solanki, 2016).

I.8.1.3 Le point de vue des agences de publicité

Les opinions sont importantes pour les agences de publicité, notamment pour situer la tendance du marché. L'exploration de la raison de l'opinion est extrêmement importante afin de déterminer la raison pour laquelle les consommateurs aiment ou n'aiment pas le produit (Tang, et al., 2009). Les applications de l'analyse d'opinion de ce point de vue se matérialisent selon (Rajput & Solanki, 2016) dans :

- **Les voix des clients (VOC - Voice of Customers)** qui sont une technique d'étude du marché permettant de décrire le processus approfondi de saisie des intentions, des désirs, des antipathies et des attentes des clients.
- **Les voix du marché (VOM-Voice of Market (VOM))** qui sont des techniques permettant aux entreprises, d'enquêter, non seulement sur leurs propres clients, mais aussi sur ceux de leurs principaux concurrents. La détection de ces informations le plus tôt possible aide à mener des campagnes marketing directes et ciblées.
- **La gestion de la réputation de marque (BRM- Brand Reputation Management)** aide à comprendre comment la perception du public d'une marque donnée change de manière positive ou négative. La variation après un événement peut être étudiée à l'aide de l'analyse d'opinion.

I.8.2 L'analyse d'opinion dans le domaine politique

Outre les commentaires sur la vente et l'achat de biens, sur l'emploi généralisé des services internet, les utilisateurs peuvent également commenter diverses questions politiques, sociales, religieuses et culturelles. La collecte et l'analyse de ces commentaires aident énormément les politiciens, les responsables de questions sociales ou les activistes religieux et culturels à prendre les décisions appropriées pour améliorer la vie sociale de la communauté. L'une des applications les plus importantes dans ces domaines est lors des élections politiques. A cette occasion, les individus du corps électoral peuvent recourir à l'analyse d'opinion afin de connaître les sentiments des uns et des autres, élément crucial pour prendre une décision lors de leur vote. Aussi, l'analyse d'opinion peut être utilisée pour détecter l'opinion publique à l'égard d'une politique. Elle permet ainsi d'élaborer des lois en adéquation avec les attentes et tendance du grand public (Haseena , 2014; Rajput & Solanki, 2016).

I.8.3 L'analyse d'opinion pour le marché boursier

Réaliser une croissance économique durable et à long terme nécessite une allocation optimale des ressources au niveau de l'économie nationale. Ceci est impossible sans l'aide d'informations et de connaissances appropriées. L'investissement dans les actions fournies en bourse est l'une des options rentables sur le marché des capitaux. Selon (Bollen, et al., 2011; Nofer & Hinz, 2015; Bing, et al., 2014), l'analyse d'opinion est utilisée pour prédire le marché boursier. Par exemple, l'analyse des commentaires quotidiens de Twitter, à l'aide d'*Opinion Finder* et de *GPOMS*, deux outils importants de suivi de l'activité boursière, appartenant à (Bollen, et al., 2011), a montré la corrélation avec les changements quotidiens de la moyenne des valeurs de clôture relatives à la Dow Jones industrie (Hemmatian & Sohrabi, 2017).

I.8.4 L'analyse d'opinion pour l'éducation

Dans cette thèse, nous nous intéressons à l'analyse d'opinion dans le domaine de l'éducation. En effet plusieurs travaux ont été menés dans ce domaine. Lors de notre étude bibliographique nous avons remarqué que les travaux menés dans cet axe de recherche peuvent être classés en trois catégories à savoir :

I.8.4.1 Analyse d'opinion pour améliorer la qualité de l'enseignement

Dans (Colace, et al., 2014), les auteurs ont mené une étude d'analyse des sentiments où ils ont mesuré l'humeur d'une salle de classe pour plusieurs cours. Cela a permis à des enseignants de corriger leur façon de présenter le cours de telle sorte à garder une bonne humeur au sein de leur classe. Afin de démontrer l'utilité de l'analyse des sentiments dans le domaine de l'éducation, les auteurs dans (Koehler , et al., 2015) ont étudié les tweets relatifs à deux conférences en l'occurrence SITE et EDUCAUSE. Pour comparer l'humeur qui règne dans chacune d'elles, ils ont mesuré le degré de satisfaction des participants. Bien que les résultats obtenus ne mettent pas en évidence une différence apparente entre les deux conférences, ils citent toutefois un certain nombre d'applications possibles pour cette technique émergente. Ils préconisent l'utilisation de ce procédé dans l'enseignement en ligne afin de surveiller l'humeur des apprenants à travers l'analyse des messages échangés par ces derniers lors de forums ou dans les réseaux sociaux. Une alerte est ainsi envoyée à l'instructeur à chaque changement brusque des sentiments pour que ce dernier puisse intervenir à temps pour garder les apprenants dans le cours.

Les auteurs dans (Ortigosa, et al., 2014) ont eu une autre idée de l'application de l'analyse des sentiments des élèves afin d'adapter les évaluations à leurs humeurs en vue d'optimiser les résultats des apprentissages en ligne.

Les auteurs dans (Altrabsheh, et al., 2013) ont proposé un système d'analyse de sentiment qui recueille les tweets des étudiants durant les cours à des laps de temps réguliers. Ces tweets sont analysés grâce aux méthodes Naïve Bayes et Support Vector Machine individuellement ou combinées. Les sentiments ambiants sont renvoyés à l'enseignant sous forme d'émoticônes. Cela va permettre à l'enseignant de mieux expliquer son cours.

Dans (Wen, et al., 2014), les auteurs ont utilisé l'analyse de sentiment dans des MOOC. En effet, ils se sont intéressés à étudier la relation entre les sentiments exprimés par les apprenants envers les cours qu'ils suivent et le taux d'abandon de ces derniers. Ils n'ont pas constaté une influence apparente entre les sentiments et l'abandon, toutefois, ils ont noté une cohérence entre le nombre d'abandons de cours particuliers et les sentiments exprimés dans le contexte de ce type de cours. D'ailleurs, ils préconisent la prudence lors du recours à l'analyse de sentiments notamment lorsque le texte à analyser est bruité.

Dans (Balahadia, et al., 2016), les auteurs ont développé un outil d'évaluation de la performance d'un enseignant à l'aide de l'analyse d'opinion. Leur étude peut aider à déterminer l'efficacité des membres du corps professoral sur la base des réactions positives et négatives des étudiants.

Dans (Oramas Bustillos, et al., 2019), les chercheurs décrivent un cadre général permettant d'intégrer l'analyse d'opinion dans un environnement d'apprentissage intelligent (ILE). Les résultats de leurs évaluations ont contribué à améliorer les supports d'apprentissage et les exercices de programmation dans le ILE (Java Sensei Java Sensei).

Dans (Nitin, et al., 2015), les auteurs ont mis au point un système d'exploration de retour (feedback) d'étudiants (SFMS) basé sur l'analyse de texte et une approche d'analyse d'opinion afin de fournir aux enseignants une analyse quantifiée et exhaustive de ces retours (feedback) et de leur octroyer des informations sur leurs pratiques d'enseignement. Cette étude vient en aide aux parties prenantes dans les cycles d'amélioration des cours.

Dans (Kechaou, et al., 2011), les auteurs ont appliqué une méthode d'analyse d'opinion dans le but d'aider les développeurs web à améliorer et à promouvoir la qualité des services pertinents. Leurs résultats expérimentaux ont montré que l'analyse d'opinion devient plus stimulante lorsqu'elle est réalisée pour des blogs de cyber apprentissage.

I.8.4.2 Analyse d'opinion pour la recherche scientifique et communautaire

L'analyse d'opinion est une opportunité prouvée dans la recherche scientifique. En effet l'étude menée par (Baucom, et al., 2013) a montré qu'il est possible de définir la géolocalisation des internautes à travers les messages qu'ils échangent sur le net. Les auteurs dans (O'Connor, et al., 2010) affirme que les opinions recueillies par des sondages et celles exprimées sur les réseaux sociaux sont fortement corrélées. La combinaison de ces deux résultats permet à la communauté scientifique de mener des études statistiques à des coûts très raisonnables grâce à l'analyse d'opinion sur les réseaux sociaux.

L'analyse de sentiments offre un outil potentiel dans l'étude des pratiques des communautés éducatives, y compris les communautés de perfectionnement professionnel, les communautés d'enseignants informels et les communautés universitaires. Ces analyses pourraient éclairer sur la façon dont les communautés changent et évoluent au fil du temps, la façon dont elles diffèrent d'un site à un autre. Elles peuvent renseigner sur la manière dont les communautés répondent à des événements particuliers comme par exemple, l'introduction de nouvelles normes (Eckert, 2006).

Dans (Chvanova, et al., 2016), les auteurs ont utilisé les techniques d'analyse d'opinion afin d'examiner les rôles des réseaux sociaux dans l'éducation et l'image de l'enseignant dans un réseau social. Les résultats de leur étude ont montré que les utilisateurs (étudiants et enseignants) sont habitués à percevoir les réseaux sociaux comme une opportunité pour une liberté d'expression illimitée.

I.8.4.3 Analyse d'opinion pour recueillir l'opinion publique sur l'enseignement

La recherche menée dans (Invento, et al., 2017), utilise l'approche Lexicon et Naïve Bayes avec Code Switching pour analyser des textes anglais-Cebuano trouvés sur des sites sociaux et des articles de presse en ligne concernant l'impact de la maternelle à la 12e année aux Philippines. Les résultats ont montré que le sentiment des Philippins aux yeux de ce programme est généralement neutre. Cependant, le sentiment sur la préparation du pays concernant le programme est généralement positif.

Dans (Hamzah & Widyastuti, 2016), les auteurs ont utilisé HMM-POS Tagger pour recueillir les commentaires et les suggestions permettant d'extraire l'opinion de la population universitaire. Cette étude a également montré la capacité de détecter une opinion à partir d'un texte de commentaires et de détecter la cible de l'opinion. Dans (Khomonenko, et al., 2017), les auteurs ont proposé une approche permettant d'identifier l'opinion publique sur les bases de données volumineuses extraites de réseaux sociaux afin de détecter l'opinion publique sur la

qualité des services éducatifs. Les résultats de leur étude ont mis en évidence un écart important entre les indicateurs attendus et ceux obtenus. Ceci a permis à l'administration d'envisager des actions correctives en vue d'améliorer l'opinion publique à son égard.

L'étude réalisée dans (Omar, et al., 2017), vise à utiliser Twitter pour analyser les sentiments. Les auteurs diffusent automatiquement des données en temps réel à l'aide de l'API Twitter afin de recueillir les sentiments du public concernant l'éducation. Ils utilisent, également, une enquête pour capter les opinions du public. Leurs résultats aident à surmonter les frustrations liées à la mise en œuvre des politiques et des réformes de l'éducation en tenant compte des points de vue et des opinions du public.

Conclusion

Dans ce premier chapitre nous introduisons l'analyse d'opinion comme branche du text-mining. Après une définition succincte, nous introduisons les attributs et les tâches de l'analyse d'opinion. Nous enchaînons avec la présentation des différentes approches utilisées dans ce domaine. Nous présentons quelques ressources lexicales de référence en mettant l'accent sur celle destinée à l'analyse d'opinion, SentiWordNet. Nous introduisons également, la modélisation de l'opinion avant de finir avec les applications de l'analyse d'opinion.

Dans le chapitre suivant, nous nous intéressons à l'orientation scolaire, ses contraintes et les facteurs pouvant la booster.

Chapitre II : L'orientation et l'E-orientation scolaire

Introduction

L'orientation scolaire et professionnelle est une question fondamentale dans le secteur de l'éducation. Elle détermine fortement les chances d'une insertion professionnelle réussie sur le marché du travail devenue de plus en plus difficile de nos jours. Elle est, toutefois, entravée par des défaillances liées principalement au manque de précision des techniques qu'elle met en œuvre dans son fonctionnement. Ces techniques se basent essentiellement sur des évaluations de connaissances purement objectives, alors que les notes ne peuvent satisfaire sans confusion, les fonctions de l'orientation, de régulation ou de certification, endossées par l'évaluation aux différents moments d'un apprentissage (Maccario, 1987). Dès lors, la mise en place de mécanismes renforçant les évaluations scolaires s'avère plus qu'une nécessité. L'orientation scolaire, de nos jours, ne peut être isolée de l'évolution que connaissent les sciences de l'éducation. En effet, l'essor de l'e-éducation impacte toutes les branches de l'éducation.

Notre étude s'insère dans la proposition d'un système d'e-orientation, destiné dans un premier temps à être mis en œuvre dans les établissements scolaires mais qui s'adaptera aux besoins des systèmes de l'e-éducation.

II.2 Définitions de quelques notions de base

Avant d'introduire la notion d'orientation scolaire, nous présentons ci-dessus quelques notions de base relatives au domaine de l'éducation.

II.2.1 L'éducation

Le terme éducation, malgré sa proximité évidente de la vie contemporaine, prête à plusieurs significations, en fonction des contextes et des centres d'intérêt. Il désigne, de manière générale, une action que l'adulte exerce sur l'enfant, en vue d'orienter ou d'infléchir son comportement. Il peut s'appliquer à une action formelle et académique, ou à une action informelle individuelle ou sociale. Très souvent, on parle d'éducation familiale, d'éducation sociale, d'éducation académique, etc., pour signifier que l'éducation intègre différents cadres, dimensions et acteurs (Sikali, 2009).

L'éducation apparaît nécessairement comme une action exercée par différents acteurs, visant la formation totale de l'homme par le développement de ses potentialités et leurs valorisations par le groupe social auquel il participe (Pieron & Salinas, 1969) (Bloch , et al., 1997). L'éducation se définit, alors, comme une action formative qui prend racine dans la société et le corps social et se développe dans les systèmes formalisés/formels. Dans ce sens, l'éducation constitue un processus où l'action de plusieurs acteurs convergent au travers (Littré, 1984) de :

- L'action d'élever un enfant, un jeune homme ;
- L'acquisition d'un ensemble d'habitudes intellectuelles ou manuelles ;
- Le développement d'un ensemble de qualités morales.

A partir de ces définitions, il peut être retenu que l'éducation constitue une action de transmission de l'héritage social et intellectuel à autrui, qui a également en filigrane un idéal moral. C'est cette acception de l'éducation qui place l'orientation scolaire au centre des préoccupations éducatives (Sikali, 2009).

II.2.2 La pédagogie

La pédagogie porte sur l'éducation ou l'action éducative. A première vue, cette notion semble familière. Pourtant, sa définition ne coule pas de source. Certains auteurs l'ont tantôt abordé comme l'art d'enseigner, comme technique de transmission de la connaissance, comme science de l'enseignement ou encore comme philosophie de l'éducation. Bien souvent elle est comprise comme l'un ou plusieurs de ces éléments (Sikali, 2009). Selon (Hubert, 1970) : « la pédagogie a pour objet d'élaborer une doctrine de l'éducation, à la fois théorique et pratique comme doctrine de la moralité dont elle est le prolongement et qui n'est exclusivement ni une science, ni une technique, ni une philosophie, ni un art, mais tout cela ensemble et ordonné selon des articulations logiques ». Elle se définit aussi comme étant une science de l'éducation qui met en évidence les méthodologies des pratiques d'éducation qui intègrent des dimensions affectives et des dimensions inconscientes. (Bloch , et al., 1997). Elle se manifeste notamment dans la relation entre l'enseignant et l'élève lors de l'action de formation.

Avec l'évolution des sciences et la complexification des apprentissages, plusieurs modèles et procédés pédagogiques sont nés, avec des niveaux variables de pertinence et d'efficacité. L'étude de ces nouveaux procédés et leur applicabilité a conduit à l'émergence de la didactique.

II.2.3 La didactique

Est une science qui a pour objet l'étude des méthodes et des théories de l'enseignement. Elle porte sur les méthodes ou les pratiques d'enseignement d'un savoir donné. Elle est mise en évidence dans la relation qui lie l'enseignant aux savoirs lors de l'action de l'enseignement. Notons que les notions « didactique » et « pédagogie » ne s'opposent pas mais sont complémentaires d'une même réalité : l'éducation. Elles s'intéressent toutes les deux : au savoir comme objet d'étude et au couple enseignant et élèves comme acteurs.

L'orientation scolaire vient couronner les scénarios mis en œuvre par la didactique et la pédagogie afin de permettre à l'élève de passer d'un palier à un autre dans le cursus éducatif. Elle s'appuie sur l'évaluation de la bonne acquisition des savoirs. De ce fait, elles se trouvent en interaction continue, comme illustré dans la figure 9, de sorte que :

- La pédagogie met en œuvre les conditions nécessaires à l'orientation scolaire, en garantissant notamment des formations de qualité aux élèves.
- La didactique contribue à la réussite de l'orientation scolaire par l'octroi d'une bonne qualité de l'enseignement.
- l'orientation scolaire peut agir sur la didactique et la pédagogie en amont pour actionner des remédiations et en aval en vue d'améliorer le rendement.

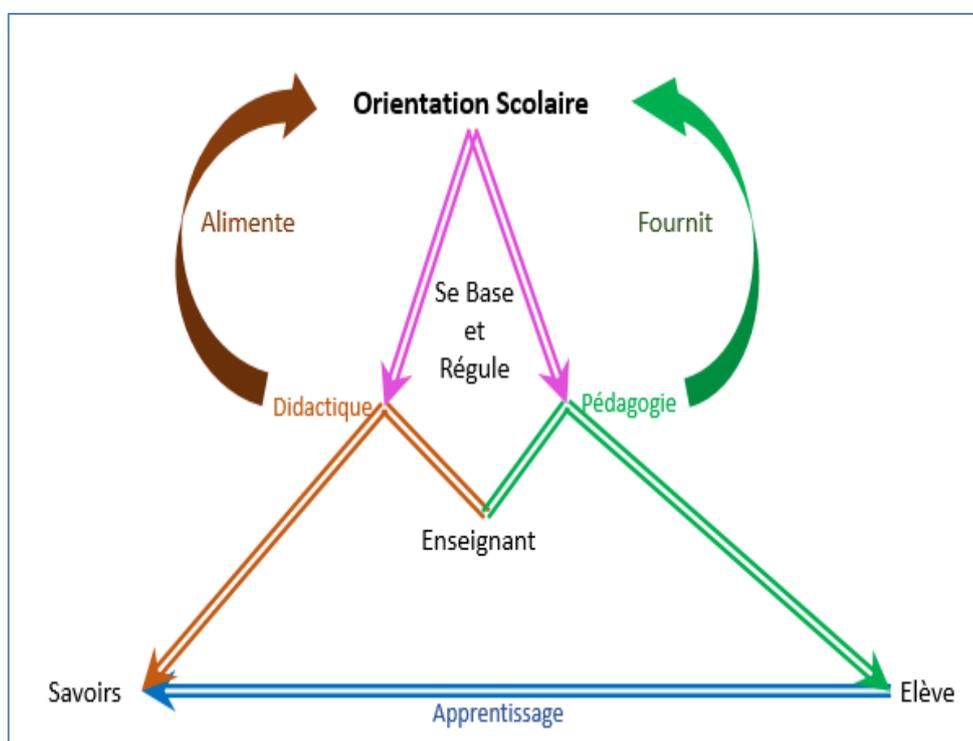


Figure 9 Interaction entre l'orientation scolaire, la pédagogie et la didactique.

Les acteurs impliqués dans l'acte d'apprentissage et d'enseignement ont chacun son opinion. L'enseignant en tant que didacticien et pédagogue a une idée précise de ses élèves. Il porte son opinion lors des passages d'une classe à une autre après évaluation des connaissances acquises par les élèves. De son côté, l'élève peut avoir une appréciation des matières enseignées, des domaines d'études ou encore de ses enseignants. Les différentes opinions des uns et des autres peuvent influencer l'orientation scolaire.

II.3 L'orientation scolaire

II.3.1 Définition

L'orientation désigne toute forme d'aide apportée à un individu pour lui permettre de découvrir ses capacités, ses possibilités, ses prédispositions et ses propres tendances et à assimiler et comprendre le milieu et l'environnement dans lequel il vit. L'orientation permet à l'individu d'être en mesure de résoudre les contraintes qu'il risque de rencontrer dans sa vie scolaire, professionnelle ou sociale (Ben Flisse, 2014).

Elle est apparue dans les années 1960, elle « permet à l'individu de se mettre en capacité de prendre conscience de ses caractéristiques personnelles et de les développer en vue du choix de ses études, de ses formations et de ses activités professionnelles, dans toutes les conjonctures de son existence, avec le souci conjoint du devenir collectif solidaire et de l'épanouissement de sa personnalité et de sa responsabilité » (UNESCO, 2001).

L'orientation scolaire est un processus qui s'occupe des préoccupations de l'individu et de ses styles comportementaux en vue de lui garantir un confort psychologique, une harmonie et une adaptation à son domaine scolaire (Demeuse & Lafontaine, 2005).

L'orientation scolaire peut se définir comme étant l'aide et conseil de l'élève afin de le guider vers le domaine d'étude qui lui correspond le mieux. Ceci à travers le décryptage de sa personnalité et la compréhension de son comportement en vue de l'aider à faire le bon choix du domaine d'étude qui lui garantira la poursuite avec succès de ses études d'une part, et la réussite dans sa vie active et sociale d'autre part.

Le besoin d'assurer un service d'orientation se fait sentir dès l'étape de l'enseignement moyen pour assurer un placement correct précocement aux élèves lors de leur passage au niveau secondaire.

II.3.2 Les objectifs et rôle de l'orientation scolaire

Les objectifs majeurs de l'orientation scolaire selon (Demeuse & Lafontaine, 2005) peuvent se résumer comme suit :

- Promouvoir la confiance en soi et le développement de la personnalité de chacun des élèves ;
- Amener tous les élèves à s'approprier des savoirs et à acquérir des compétences qui les rendent aptes à apprendre toute leur vie et à prendre une place active dans la vie économique, sociale et culturelle ;
- Préparer tous les élèves à être des citoyens responsables, capables de contribuer au développement d'une société démocratique, solidaire, pluraliste et ouverte aux autres cultures ;
- Assurer à tous les élèves des chances égales d'émancipation sociale.

Le rôle de l'orientation scolaire dans la réalisation de ses objectifs est d'offrir de l'aide aux intervenants de ce processus comme illustré dans la figure 10. En effet, elle doit aider les élèves à mieux comprendre l'environnement qui les entoure. Elle doit aussi, les assister dans la compréhension de soi afin qu'ils puissent bien réussir et mieux s'adapter à leur environnement. L'orientation scolaire doit offrir également un accompagnement pour les élèves lors de la prise de décisions et de résolution de contraintes.

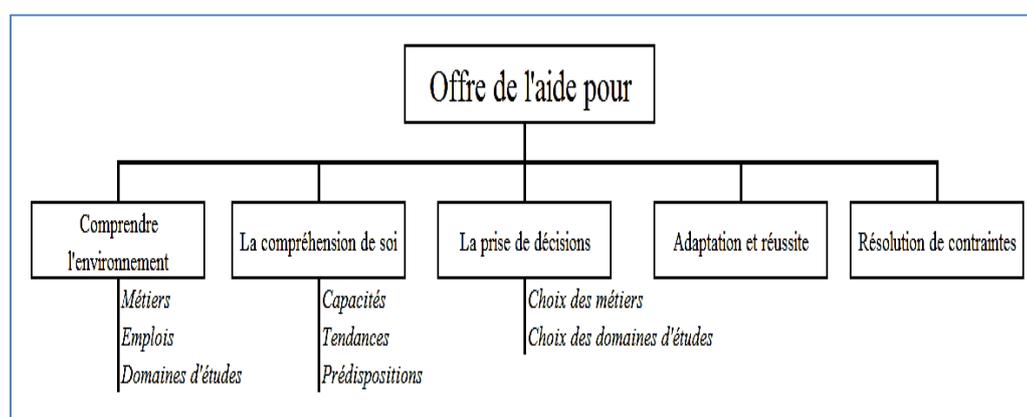


Figure 10 Rôles de l'orientation scolaire.

II.3.3 Les enjeux de l'orientation scolaire

Les enjeux de l'orientation sont doubles, enjeux collectifs et enjeux individuels, et parfois antagonistes (Richit, 2014).

II.3.3.1 Les enjeux collectifs

Ils se traduisent par des pratiques de sélection et de répartition, il s'agit de gérer au mieux les flux pour adapter les élèves au nombre de places dans les différents établissements de formation et aux débouchés prévus par les milieux économiques.

II.3.3.2 Les enjeux individuels

Ils s'illustrent par l'aspiration de chacun à réussir sa vie, et pour cela à développer ses compétences ; il s'agit pour chaque élève de trouver des études et un métier où il pourra s'épanouir, et pour ce faire il devra pouvoir bénéficier d'informations et de conseils.

II.3.4 Moyens de l'orientation scolaire

Pour pouvoir assurer une orientation adéquate, il est nécessaire d'évaluer les capacités des élèves, leurs tendances et leurs prédispositions personnelles. Il est, aussi, important d'observer leurs conduites, leurs passés et leurs présents pour cerner tous les facteurs influents sur leurs personnalités. Pour ce faire, les chargés de l'orientation recourent à un ensemble de moyens décrits ci-dessous.

II.3.4.1 L'observation

C'est le plus ancien outil de recueil d'informations sur les individus dans le domaine de l'orientation scolaire. Elle est utilisée pour observer :

- Le comportement de l'élève dans son environnement ;
- Sa capacité d'adaptation avec le travail scolaire (studieux, concentration) ;
- Sa capitalisation scolaire en détectant les côtés faibles ;
- Ses capacités d'apprentissage, de compréhension, de réflexion ;
- Son état physique et son état de santé ;
- Ses intérêts et loisirs ainsi que ses tendances.

II.3.4.2 La rencontre – Les entretiens individuels

Elle consiste à rencontrer l'élève et à discuter des difficultés qu'il rencontre, en vue, de l'aider à en trouver des solutions. Pour l'orientation scolaire, l'objectif est d'aider l'élève dans son choix lorsque cette décision lui est difficile à prendre (Guichard & Huteau, 2001). Dans ce cas, deux approches peuvent être adoptées : une approche existentielle et une approche vocationnelle.

- **L'approche existentielle** est centrée sur le « soi » du consultant. Les questions fondamentales sont les suivantes : qui suis-je vraiment et qui voudrais-je être ?

Comment être celui que je voudrais être ? Le soi du consultant est le point central de cette approche : c'est pourquoi elle est qualifiée d' « existentielle ».

- **L'approche vocationnelle** est centrée sur les « choix » d'études et de carrière. La question fondamentale est la suivante : quelles sont les voies qui me conviennent le mieux ? Autrement dit : quelles sont les voies qui me permettraient de devenir celui ou celle que je voudrais devenir ? La question de l'orientation est au centre de cette approche : elle est, alors, qualifiée de « vocationnelle ».

Dans nos sociétés où l'on considère que l'exercice d'une profession est un moyen de construire celui que l'on veut être, les deux approches ci-dessus évoquées sont intrinsèquement liées.

II.3.4.3 Les séminaires

Les séminaires centrés sur la problématique de la construction d'un projet personnel et professionnel constituent un bon moyen pour toucher rapidement un groupe de personnes confronté aux mêmes questionnements.

II.3.4.4 L'immersion et les stages en milieu professionnel

L'immersion et les stages en milieu professionnel sont un excellent moyen de confronter son projet personnel et professionnel à la réalité de la profession sur le terrain. Ils sont en quelque sorte un moyen de validation de la démarche de construction du projet professionnel.

II.3.4.5 L'éducation au choix - L'éducation à l'orientation

L'Éducation à l'Orientation (EAO) consiste en une éducation au choix : elle vise à donner aux élèves des méthodes et des connaissances pour les aider à devenir acteurs de leur orientation. Elle contribue au développement de la personnalité et de l'autonomie des élèves. Elle est fondée sur trois axes à savoir, la connaissance de soi, la connaissance de l'environnement socio-économique et la connaissance des formations (Leblanc, 2017).

L'éducation au choix insiste sur la responsabilité des individus dans leur orientation tout au long de leur vie. L'école est concernée par l'orientation des jeunes. Pour favoriser la prise de décision par l'élève, l'orientation scolaire est passée d'une approche prescriptive à une approche éducative. Dans l'approche prescriptive, l'adulte est responsable, c'est lui qui oriente, et l'élève fait ce qu'on lui dit de faire. Dans l'approche éducative, l'élève est responsable, c'est

lui qui s'oriente, et le rôle de l'adulte dans l'école consiste à aider l'élève à développer des compétences lui permettant de faire lui-même ses choix (Richit, 2014).

II.3.4.6 Les tests psychologiques – Styles d'apprentissage

C'est un moyen objectif qui aide les chargés de l'orientation à découvrir des côtés obscurs de l'élève en vue de solutionner des situations complexes telles que le choix du domaine d'étude universitaire. Ces tests peuvent être menés pour détecter les capacités de l'élève, ses niveaux d'intelligence ainsi que ses tendances et la nature de sa personnalité. Il convient de noter que les tests les plus fréquents ou les plus indiqués dans le cadre de l'orientation sont ceux destinés à définir les styles d'apprentissage ou les intelligences multiples de l'élève.

Le style d'apprentissage d'un individu, c'est son mode personnel de saisie et de traitement de l'information. En pratique, et en d'autres termes, le style d'apprentissage c'est la manière préférentielle d'aborder et de résoudre un problème (Hawk & Amit, , 2007). Des recherches dans le domaine de l'éducation ont montré que les élèves ont tendance à apprendre en se basant sur leur propre style d'apprentissage. Ce concept a, en effet, plusieurs applications dans le domaine de l'éducation notamment dans l'évaluation. Une recherche récente l'a intégré dans les évaluations dans les MOOC⁵ (Haddadi, et al., 2018 -c). Plusieurs auteurs, notamment Howard Gardner, David Kolb, Thomas Armstrong, etc. ont contribué à l'étude des styles d'apprentissages. Howard Gardner propose plusieurs types d'intelligence qu'il nomme intelligences multiples définissant des styles d'apprentissage (Gardner, 2011; Gardner, 2006). Ces intelligences de Gardner ont été reprises par Garceau dans (Garceau, 2010), où elle suggère des domaines d'études favorables pour chaque style d'apprentissage à travers les résultats d'un questionnaire (Cf. annexe 1). A titre d'exemple, un individu qui a une intelligence interpersonnelle très développée peut être un excellent politicien, avocat, directeur, enseignant ou encore commerçant. Pour prétendre à être architecte, l'individu doit disposer, à des degrés divers, de compétences d'ordre spatial, logico-mathématique, kinesthésique et interpersonnel. Le tableau 4 illustre les intelligences multiples de Gardner avec leurs descriptions.

⁵ MOOC : Massive Open Online Courses.

Tableau 4 Inteligences multiples de Gardner.

Intelligence	Description
Verbale/linguistique	Capacité d'utiliser les mots efficacement, oralement ou par écrit. Aimer écrire, lire, parler, expliquer, comprendre les consignes, convaincre ou encore raconter des histoires.
Logique/mathématique	Capacité de déduction, d'utiliser les nombres efficacement, de bien raisonner et de penser de façon abstraite et logique. Aimer classifier, catégoriser, observer, résoudre des problèmes, calculer, quantifier.
Visuelle/spatiale	Capacité de bien percevoir le monde spatial et visuel, de se faire une image mentale juste des choses, de penser en trois dimensions. Aimer travailler avec des objets, construire, assembler, fabriquer, imaginer, concevoir des plans, inventer des choses.
Interpersonnelle	Capacité de comprendre les motivations, les sentiments des autres et d'entrer en relation avec eux, empathie, sensibilité aux expressions verbales et non verbales d'autrui. Aimer aider, enseigner, influencer.
Intra-personnelle	Capacité de se comprendre et d'utiliser cette compréhension pour mieux vivre. Aimer résoudre ses problèmes personnels, aider à résoudre ceux de l'humanité.
Corporelle/kinesthésique	Capacité d'apprendre par la pratique, de bien manipuler les objets, utiliser son corps d'une manière fine et élaborée, à s'exprimer à travers le mouvement, d'être habile avec les objets, disposer d'une excellente coordination corporelle. (Visio-motrice, dextérité manuelle, motricité fine, etc.).
Musicale	Capacité de «saisir» le rythme et les sons de la musique, c'est-à-dire une aptitude à produire et à apprécier un rythme, une tonalité et un timbre, appréciation des formes d'expression musicale. Aimer chanter ou jouer d'un instrument.
Naturaliste	Capacité de comprendre, de classer et d'expliquer la nature (plantes, animaux, phénomènes naturels), être capable de reconnaître, classifier les individus, les espèces, intérêt pour l'écologie et l'environnement.

II.4 Les TIC et E-Orientation Scolaire

Le terme E-éducation est utilisé pour désigner des scénarios pédagogiques s'appuyant sur les TIC⁶. Le 'e' renvoie à 'électronique' ou 'en ligne' pour permettre d'inclure divers formules d'introduction de TIC ou TICE⁷ dans les processus de formation / enseignement / apprentissage pour, à priori améliorer l'accessibilité et/ou la qualité de l'apprentissage. De manière générale, un système de la E-éducation est caractérisé par des supports de cours (en ligne ou hors ligne), un apprentissage individuel ou collectif, la présence ou non d'un tuteur et des possibilités de guidage (aides). Dans ces systèmes, la différenciation et l'adaptation des parcours notamment le 'quand' et 'où' apprendre est devenu possible, ce qui renforce l'autonomie des apprenants. L'évolution des systèmes d'E-éducation donne lieu à trois grandes catégories : les systèmes destinés à l'auto-apprentissage, le E-learning et les systèmes basés sur l'approche connectiviste. Ces catégories sont identifiées en se basant sur deux principaux critères : le degré d'interaction et le degré de liberté des apprenants. L'évolution des systèmes de l'E-éducation à travers le temps est accompagnée d'une augmentation du degré d'interaction entre les apprenants d'une part et la souplesse du guidage des apprenants (Haddadi, 2018 -b).

II.4.1 L'E-orientation

Compte tenu de ce qui précède, nous définissons l'E-orientation comme la branche de l'E-éducation qui met en jeu les technologies émergentes au service de l'amélioration de la qualité de l'orientation scolaire. De ce fait, la E-orientation ne va pas se limiter à automatiser le processus de l'orientation en tant que système d'information mais elle intègre des procédés d'intelligence artificielle notamment l'extraction et la synthèse de connaissances.

II.4.2 Apports des TIC pour l'orientation scolaire

Le développement des TIC a permis de concevoir des outils pour aider à l'orientation scolaire et à l'insertion professionnelle. En effet, on voit émerger des outils se basant principalement sur soit les possibilités de l'informatique traditionnelle, soit les possibilités du multimédia et de l'Internet, soit encore sur les ressources développées autour de l'intelligence artificielle et les systèmes experts (Devauchelle , 2007). Les TIC qui ont progressivement envahi le champ de l'orientation scolaire, sont aujourd'hui incontournables. L'utilisation des TIC dans l'orientation est intéressante à différents points de vue. En effet, ils offrent divers avantages généraux parmi lesquels nous pouvons citer (CSE, 2000; Matmati, 2001) :

⁶ Technologies de l'information et de la communication.

⁷ Technologies de l'information et de la communication en éducation.

- une interactivité importante ;
- des sources de documentation actualisées et de taille conséquente disponibles en un seul clic.
- l'analyse des flux et le suivi individualisé des élèves, la documentation et la mise à disposition de l'information ;
- l'aide personnalisée à la connaissance de soi, aux choix et à la décision, utilisant des logiciels à caractère interactif et/ou psychologique.

Les TIC ont un impact considérable sur la fonction des conseillers d'orientation. Elles leur permettent notamment :

- De gérer de façon automatique des fiches d'orientation qui leur offre un gain de temps et une facilité de centralisation des données.
- D'avoir des informations mises à jour sur les filières de formation qu'offrent les universités et les grandes écoles toutes connectées de nos jours à internet ;
- D'utiliser les nouvelles méthodes d'orientation pratiquées ailleurs, méthodes disponibles gratuitement sur internet ;
- D'échanger des expériences avec les collègues d'ici et d'ailleurs à travers des courriers électroniques, des forums, des plates-formes... ;
- D'avoir une aisance dans le travail ;
- De pallier l'insuffisance de la documentation dans les établissements scolaires à travers des sites web, des brochures, des publications et de journaux diffusés sur internet et dont l'abonnement est gratuit ;
- D'obtenir les programmes de formation, des informations sur les concours (date, épreuves et corrections...).

II.4.3 Les logiciels d'aide à l'orientation

Les outils développés pour automatiser l'orientation scolaire sont dits logiciel d'aide à l'orientation (LAO). Ils peuvent être classés en fonction de leurs objectifs opérationnels qu'on leur assigne (Leblanc, 2017):

- **Outils centrés sur un apport de connaissances extérieures** (sur les formations et emplois) : Dans cette catégorie, l'outil a pour rôle de fournir de l'information à partir d'une base de données ou de travailler sur des stéréotypes de représentations professionnelles. Leur utilisation, individuelle ou collective, peut être partagée, dans

un cadre d'intervention concertée, entre les conseillers d'orientation-psychologues et d'autres membres de l'équipe éducative.

- **Outils centrés sur l'évaluation de la personne et l'image de soi** : Dans cette catégorie, l'outil est généralement producteur d'un profil à caractère psychologique. Son usage relève de la seule compétence d'un conseiller d'orientation-psychologue et s'inscrit dans le cadre d'une démarche globale qui comprend : une demande (à analyser), une investigation, la restitution de résultats et une synthèse. Ces logiciels peuvent avoir deux finalités :
 - révéler un potentiel : le jeune est maintenu dans une conduite passive ;
 - provoquer des changements, en faisant apparaître des éléments de confrontation dont le jeune pourra se saisir pour développer sa réflexion.

- **Outils composites** : Dans cette catégorie, l'outil juxtapose des modules d'information sur la connaissance des études et des métiers, à des modules de connaissance de soi et de ses intérêts. L'usage de ce genre d'outils nécessite également l'accompagnement des conseillers d'orientation psychologues. Dans cette catégorie, l'outil s'inscrit dans une démarche psychopédagogique qui articule la connaissance de soi, des études et des métiers, dans une perspective d'éducation à l'orientation.

Dans (Feral, 2013), l'auteur cite quelques exemples de LAO :

- **Itinéraires** : Logiciel d'évaluation des motivations professionnelles et des connaissances de base en mathématiques et en français. Système spécialement conçu pour répondre aux problèmes de l'orientation des jeunes vers une formation professionnelle.
- **Aideo - Aide à la décision d'orientation** : L'objectif de ce logiciel est de proposer l'offre de service la plus adaptée à la situation de l'utilisateur à partir des éléments de diagnostic et du paramétrage réalisé par la structure d'aide et d'accompagnement. Outil d'aide à la décision pour les professionnels de l'insertion et de l'orientation,
- **Valeurs - Logiciel d'aide au choix professionnel** : Le logiciel Valeurs offre à l'utilisateur la possibilité d'établir son profil en fonction de douze valeurs répertoriées en deux grandes dimensions : changement/stabilité - affirmation de soi/dépassement de soi. C'est un outil d'aide pour toutes les démarches

d'accompagnement, de bilan, d'aide à la réflexion professionnelle et à la décision. Valeurs est basé sur la Théorie universelle des valeurs de Schwartz. Quarante-six questions permettent à l'utilisateur d'établir son profil. Le logiciel lui offre en second lieu une analyse et un répertoire des métiers susceptibles de lui convenir en fonction des valeurs qui lui correspondent.

II.5 L'orientation scolaire à travers le monde

La question de l'orientation scolaire est d'une grande importance. Il est important de décrire la réalité de l'orientation scolaire et éducative à travers le monde. Nous citons le cas de l'Algérie pour l'Afrique, des USA pour l'Amérique, le Japon pour l'Asie et la France pour l'Europe (Berou, 2010; Ben Flisse, 2014).

II.5.1 L'orientation scolaire en Algérie

Avant d'exposer l'orientation scolaire en Algérie, il est plus judicieux de commencer par présenter son système éducatif.

II.5.1.1 Le système éducatif Algérien

Le système éducatif algérien est articulé autour de plusieurs niveaux : préparatoire, fondamental (primaire, et moyen), secondaire (général et technologique), professionnel et l'enseignement supérieur. Aussi, il y a la formation continue, un palier non négligeable, qui est assurée par l'université de la formation continue (JO 04, 2008).

L'organisation et la gestion des niveaux primaire, moyen et secondaire sont assurées par le ministère de l'éducation nationale. Le niveau professionnel est confié au ministère de la formation et de l'enseignement professionnels. L'enseignement supérieur est à la charge du ministère de l'enseignement supérieur et de la recherche scientifique. La figure 11 illustre les différents paliers du système éducatifs algérien.

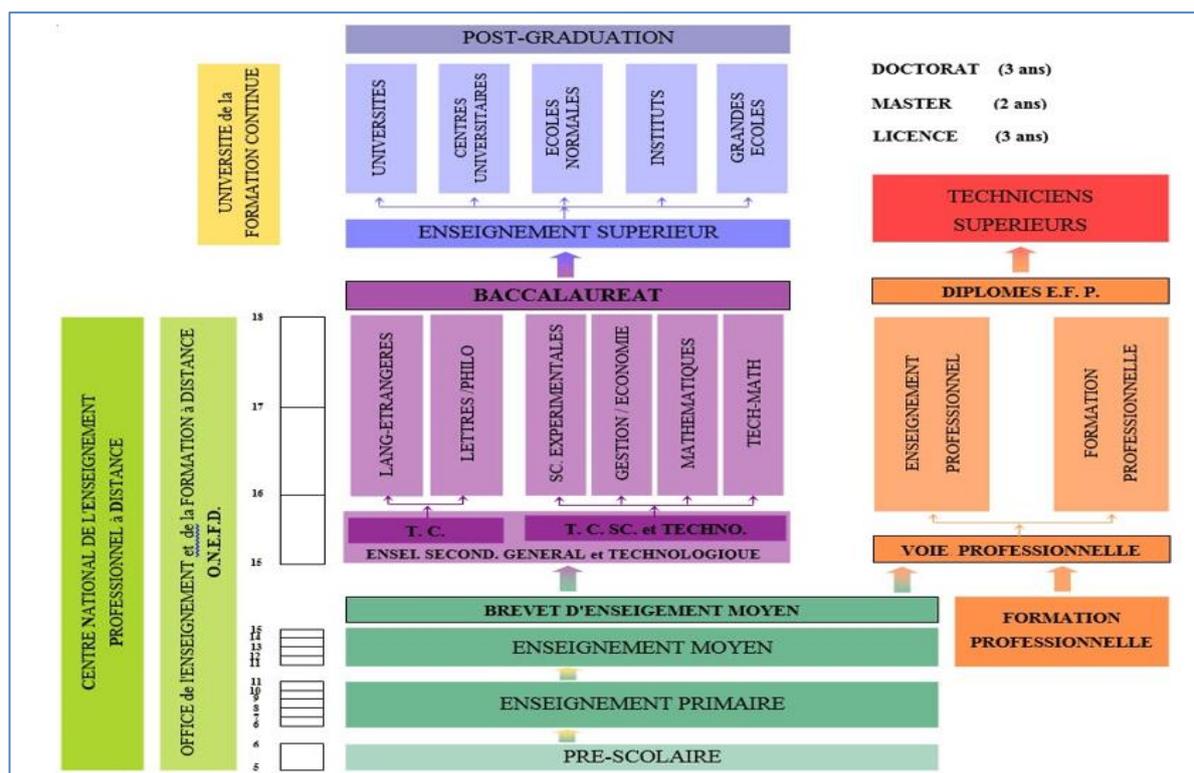


Figure 11 Structuration du système éducatif algérien (MEN, 2016).

L'éducation préparatoire a pour but de préparer les enfants âgés de trois (03) à six (06) ans, à l'accès à l'enseignement primaire. Elle regroupe les différents stades de prise en charge socio-éducative de ces enfants. Elle est dispensée dans des écoles préparatoires, des jardins d'enfants et des classes enfantines ouvertes au sein d'écoles primaires (JO 04, 2008). Son objectif principal est de les aider à développer leur pratique du langage à travers des situations de communication induites par les activités proposées et le jeu. Elle a aussi pour but de les initier aux premiers éléments de lecture, d'écriture et de calcul à travers des activités attrayantes et des jeux appropriés. Mais vu le caractère non obligatoire de l'éducation préscolaire, l'état a pris l'engagement de veiller au développement de cette éducation et poursuivre la généralisation avec le concours des institutions, administrations et établissements publics, des associations ainsi que du secteur privé (JO 04, 2008).

L'enseignement Fondamental (primaire et moyen) est l'étape de scolarité obligatoire dont la durée est de neuf (09) ans. Avant la réforme de 2008, il était organisé en trois cycles de trois ans : le cycle de base, le cycle d'éveil et le cycle d'orientation. Depuis la rentrée scolaire 2003/2004 il comprend (JO 04, 2008):

- **L'enseignement primaire**, d'une durée de cinq (05) ans. L'âge d'admission à l'école primaire est fixé à six (06) ans révolus sauf si une dérogation d'âge est

accordée selon des conditions fixées par le ministère de l'éducation nationale. La fin de scolarité dans l'enseignement primaire est sanctionnée par un examen final ouvrant droit à la délivrance d'une attestation de succès.

- **L'enseignement moyen** d'une durée de quatre (04) ans. À la fin de la scolarité dans l'enseignement moyen et après un examen final ouvrant droit à l'obtention d'un diplôme appelé « brevet d'enseignement moyen », l'élève est admis automatiquement en 1^{re} année secondaire général et technologique ou vers l'enseignement professionnel. Les élèves non admis ont la possibilité de rejoindre soit la formation professionnelle, soit la vie active, s'ils ont atteint l'âge de seize (16) ans révolus.

L'objectif de l'enseignement fondamental est de doter les élèves d'outils d'apprentissage essentiels que sont : la lecture, l'écriture et le calcul et de leur permettre d'acquérir des compétences qui les rendent aptes à apprendre tout au long de leur vie, ainsi que renforcer leur identité en harmonie avec les valeurs et traditions sociales, spirituelles et éthiques issues de l'héritage culturel commun.

L'enseignement fondamental aspire à offrir aux élèves des possibilités de s'imprégner des valeurs de la citoyenneté et des exigences de la vie en société et d'apprendre à observer, analyser, raisonner et résoudre des problèmes.

Il a aussi pour but de permettre aux élèves de comprendre le monde vivant et inerte, ainsi que les processus technologiques de fabrication et de production, et de développer leur sensibilité, leur curiosité, leur imagination, leur créativité et d'aiguiser leur sens esthétique leur esprit critique et de s'initier aux nouvelles technologies de l'information et de la communication et à leurs applications élémentaires.

L'enseignement fondamentale a pour cible de favoriser chez les élèves l'épanouissement harmonieux de leurs corps et de développer leurs capacités physiques et manuelles, d'encourager l'esprit d'initiative, le goût de l'effort, la persévérance et l'endurance. Aussi, de leurs permettre d'avoir une ouverture sur les civilisations et les cultures étrangères et d'accepter les différences et de coexister pacifiquement avec les autres peuples et de poursuivre des études ou des formations ultérieures (JO 04, 2008).

L'enseignement secondaire est d'une durée de trois (03) ans, il est dispensé dans des lycées. Il comprend l'enseignement secondaire général et l'enseignement secondaire technologique. Il est organisé en tronc commun en première année, et en filières à compter de la deuxième année. La fin de la scolarité dans ce cycle, est sanctionnée par le baccalauréat de l'enseignement secondaire (JO 04, 2008).

L'enseignement secondaire a pour missions, outre la poursuite des objectifs généraux de l'enseignement fondamental (JO 04, 2008) de :

- Consolider et d'approfondir les connaissances acquises dans les différents champs disciplinaires ;
- Développer les méthodes et les capacités de travail personnel et de travail en équipe et de cultiver les facultés d'analyse, de synthèse, de raisonnement, de jugement, de communication et de prise de responsabilités ;
- Offrir des parcours diversifiés permettant la spécialisation progressive dans les différentes filières en rapport avec les choix et les aptitudes des élèves ;
- Préparer les élèves à la poursuite d'études ou de formations supérieures.

La structuration de l'enseignement secondaire est basée sur deux grands tronc commun d'une (01) année chacun :

- **Un tronc commun lettres** avec deux (02) filières en deuxième et troisième année secondaire :
 - lettres/philosophie ;
 - langues étrangères.
- **Un tronc commun sciences et technologie** avec quatre (04) filières en deuxième et troisième année secondaire :
 - Mathématiques ;
 - Sciences expérimentales ;
 - Gestion-économie ;
 - Technique mathématiques avec quatre (04) options :
 - Génie électrique ;
 - Génie civil ;
 - Génie mécanique ;
 - Génie des procédés.

L'enseignement supérieur (universitaire) en Algérie peut se faire dans trois types d'établissements :

- **Université** : Etablissement avec de grands effectifs d'étudiants, répartis dans des facultés, des instituts et des départements. Plusieurs domaines et filières de formation sont proposés, ainsi que des formations en sciences médicales et vétérinaires. On peut y dispenser des filières à recrutement national, des masters à cursus intégré de licence (avec des moyennes d'accès élevées) ou des licences professionnalisantes.
- **Centre Universitaire** : Etablissement dispensant diverses formations avec un effectif d'étudiants relativement important, répartis dans des instituts et des départements.
- **Ecole Supérieure** : La formation dans les écoles supérieures se décline, par domaine, en deux étapes. La première se caractérise par deux (02) années de formation en classe préparatoire. La deuxième étape constitue le deuxième cycle des écoles supérieures qui se prépare en trois (03) années. L'accès au second cycle des écoles supérieures est conditionné par la réussite au concours national.

Les structures de bases du système universitaire algérien sont resumées dans la figure 12.

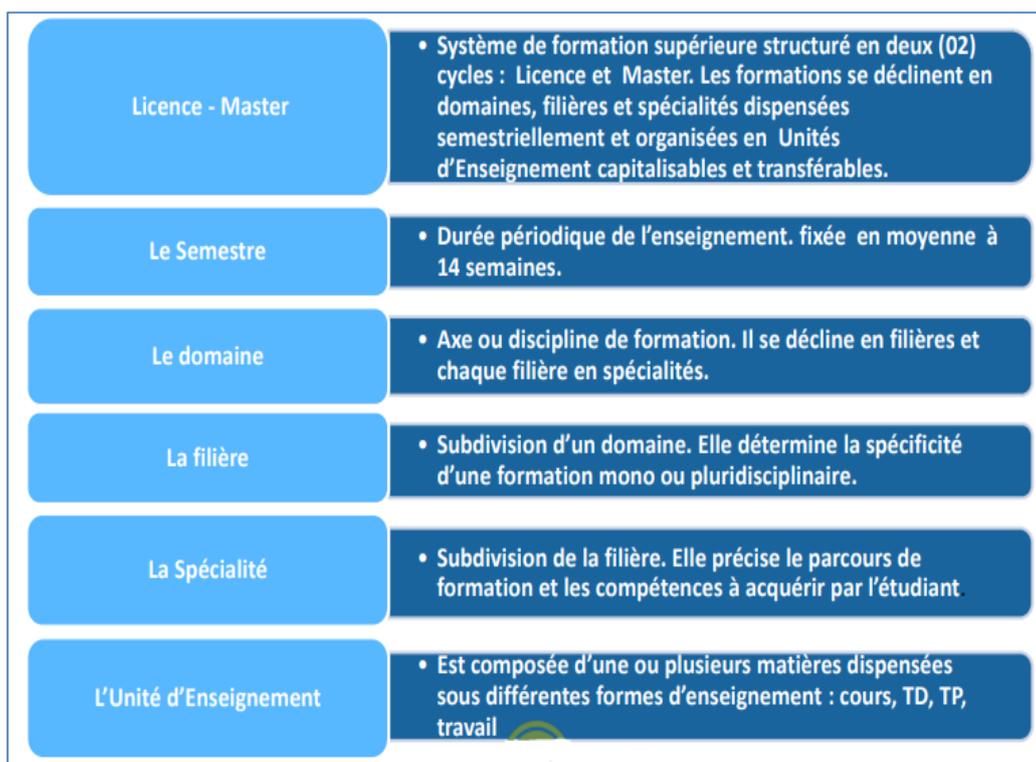


Figure 12 Structures de bases du système universitaire Algérien (MESRS, 2018).

II.5.1.2 L'orientation scolaire en Algérie

L'orientation scolaire, en Algérie, est portée par des équipes pédagogiques constituées d'enseignants, d'administratifs, d'inspecteurs et de conseillers d'orientation et de guidance scolaire. Ces équipes investissent des efforts et des comportements très actifs afin d'adopter des normes pédagogiques dans le processus d'orientation, d'une part et d'asseoir des pratiques d'évaluation et de remédiation auxquelles seront adossées les questions d'orientation et de guidance, d'autre part.

Les équipes d'orientation visent à renforcer le dispositif de l'orientation scolaire dans le système éducatif, en vue d'une prise en charge précoce des dimensions socio-psychologiques des élèves et celles relatives à leurs vocations. Elles ambitionnent, également, à développer plus spécifiquement ce dispositif à partir de la première année moyenne, pour accompagner et aider l'élève dans la construction progressive de son projet personnel et l'impliquer dans ses choix scolaires et professionnels. Elles œuvrent aussi à la consolidation de la méthodologie de l'orientation progressive vers les deux tronc communs de la première année secondaire et les différentes filières de la deuxième année secondaire. Elles veillent à sensibiliser sur la question de l'orientation dès le début de l'année scolaire, en organisant les opérations d'information nécessaires. Aussi, elles incitent les conseillers d'orientation et de guidance scolaire à intensifier les entretiens avec les élèves, afin de procéder à la remédiation des cas socio-psychologiques.

L'orientation scolaire en Algérie intervient dans trois phases principales du cursus éducatif à savoir :

- **L'orientation des élèves de quatrième année moyenne vers la première année de tronc commun** secondaire qui se fait sur la base des notes obtenues durant les deux dernières années ainsi que les souhaits des élèves à condition de réussir à l'examen du Brevet de l'Enseignement Moyen (BEM).
- **L'orientation des élèves des tronc communs** de la première année secondaire vers les spécialités techniques ou générales de l'enseignement secondaire se fait en fin d'année, selon leurs souhaits et leurs résultats.
- **L'orientation des élèves vers le cycle universitaire** après l'obtention de l'examen du baccalauréat, les nouveaux bacheliers sont invités à faire leurs préinscriptions en ligne sur une plateforme dédiée (PROGRES). La figure 13 illustre les étapes de préinscription et celles de confirmation par inscription.

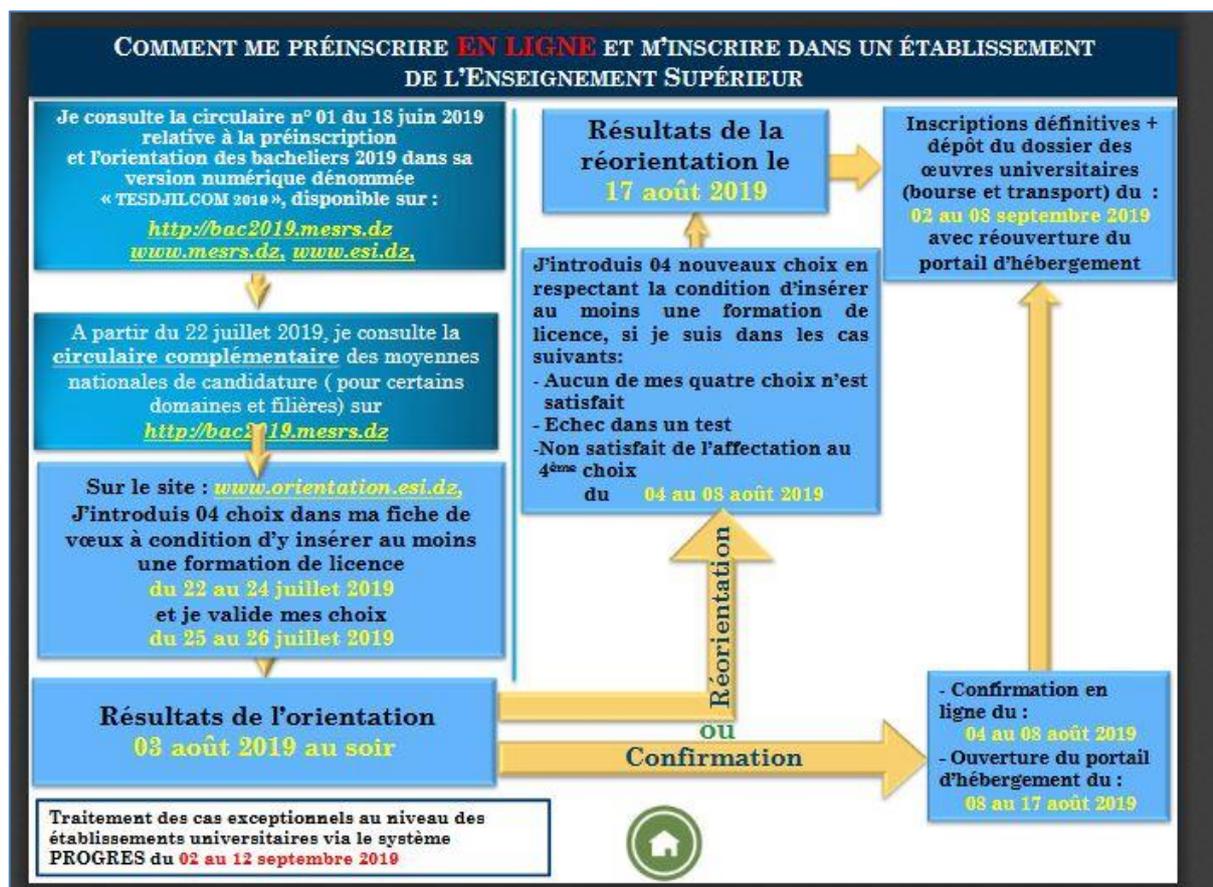


Figure 13 Inscription aux cycles d'études supérieures en Algérie (MESRS, 2018).

Dans cette plateforme, le nouveau bachelier est invité à indiquer ses choix d'étude dans une liste prédéfinie de domaines d'études universitaire, correspondant à la moyenne obtenue dans l'examen du baccalauréat ainsi qu'à sa filière d'étude secondaire. L'élève se verra orienté selon sa moyenne et la disponibilité des places pédagogiques relatives aux domaines d'études choisis. On dénote que les principaux critères d'orientation sur cette plateforme sont la moyenne générale obtenue à l'examen du baccalauréat et la disponibilité des places pédagogiques comme illustré dans la figure 14.

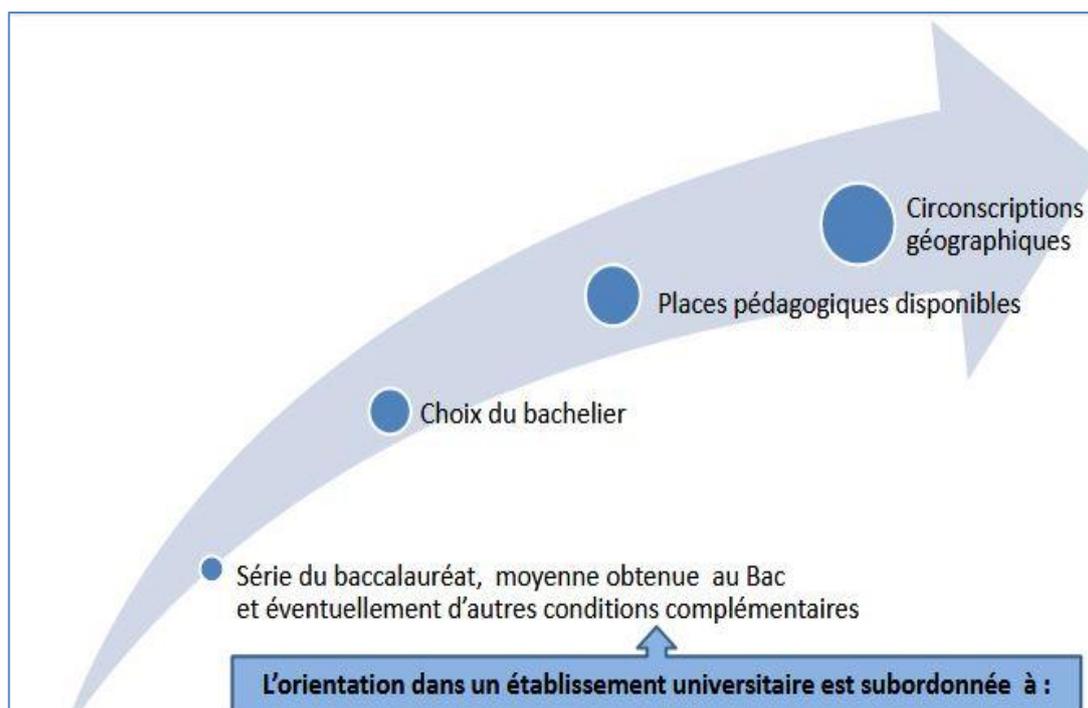


Figure 14 Critères de l'orientation scolaire en Algérie (MESRS, 2018).

En Algérie, la question de l'orientation scolaire est d'une importance capitale. Le ministère de l'éducation nationale conscient de l'inadéquation de l'orientation imposée, a mis l'accent sur la nécessité de changer les anciennes pratiques, en adoptant une orientation acceptée voire souhaitée par l'élève. Il préconise que l'orientation des élèves doit se faire de façon à leur permettre d'identifier leurs choix par rapport à leur avenir qui ne dépend pas uniquement des études supérieures, mais qu'il y a d'autres chances de réussite, notamment à travers la formation professionnelle. Il rappelle toutefois que : « *l'opération de l'orientation doit prendre en considération les compétences personnelles des élèves ainsi que les potentialités offertes par le marché du travail* ». Aussi, « *l'établissement scolaire doit devenir un espace qui offre aux élèves toutes les données sur le cursus scolaire, afin de lui permettre de choisir entre l'enseignement secondaire et l'enseignement professionnel* ».

Dans cette optique, un groupe de travail, regroupant des spécialistes des secteurs de l'éducation, de la formation professionnelle et de l'enseignement supérieur, a été installé. Son rôle est de mettre en place une feuille de route pour une orientation efficace des élèves. Aussi, le ministère de l'éducation est engagé dans la voie de la modernisation et, à ce titre, l'observatoire national de l'éducation et de la formation est mis sur rails. Cet observatoire a pour mission :

- D'observer le fonctionnement du système national d'enseignement dans toutes ses composantes ;

- D'analyser les facteurs déterminant des situations d'enseignement et d'apprentissage, l'évaluation et la qualité des prestations pédagogiques, ainsi que les performances des enseignants et des apprenants ;
- D'émettre des propositions et des mesures correctives ou d'améliorations inscrites dans le cadre de la loi d'orientation scolaire. Ces propositions doivent être en adéquation avec l'école du troisième millénaire, laquelle est ouverte sur le monde et reflète les grandes valeurs humanistes ».

La communauté éducative doit s'engager dans cette voie afin de permettre l'édification d'une société du savoir (Ettouahria, 2015).

II.5.2 L'orientation scolaire aux USA

L'orientation scolaire a débuté aux USA quand Truman L. Kelly a présenté sa thèse de doctorat intitulée « Educational Guidance » en 1914 à la Colomby University et comme la société américaine est hétérogène et les systèmes scolaires étaient diversifiés et propres à chaque localité, il était nécessaire que le plan d'orientation et de conseil adopté par les autorités éducatives fédérales soit souple et adapté à chaque localité. La condition est qu'il doit être construit sur des bases solides et claires afin qu'il puisse réaliser son objectif. Il doit aussi se doter d'une administration intelligente, audacieuse, avec un large esprit, polie et ayant une profonde volonté de maîtriser les outils de mesure et d'évaluation basés sur les technologies de l'information. Cela dans le but d'aider le conseil de l'orientation qui se trouve au niveau de chaque région. Ce conseil a pour mission d'accompagner chaque établissement scolaire dans les tâches de l'orientation. Sachant que l'équipe d'orientation à ce niveau est constituée du directeur de l'établissement, des enseignants, le conseiller d'orientation, le psychologue, le sociologue et le médecin scolaire.

Il est important de signaler ici que « Educational Guidance » ou « Scholar Counseling » porte un intérêt à la croissance globale de chaque apprenant que ce soit ses prédispositions intellectuelles, ses états d'esprit (émotionnels), ses tendances ou encore ses compétences personnelles. Elle lui apporte de l'aide pour atteindre les objectifs suivants :

- Préservation de la vie et sa protection des dangers ;
- Renforcer la confiance en soi et l'équilibre psychologique et la stabilité et la maturité émotionnelle ;
- Renforcer son réalisme et l'écarter des fausses idées et imaginations ;
- Organiser les plannings d'études et de récréation pour le faire profiter de ses temps libre ;

- Fixer ses objectifs et trajectoires pour le post lycée ;
- Respecter la valeur du travail, quel que soit sa nature ou son revenu ;
- Développer ses capacités à acquérir une polyvalence en suivant les stages (formation, entraînement...) nécessaires pour occuper des postes disponibles ;
- Développer son autonomie pour être en mesure de prendre les décisions adéquates, notamment celles relatives au choix de son métier ou études qui s'adaptent à ses capacités et aspirations ;
- Assurer le juste accord et la bonne adaptation entre les études actuelles et le métier d'avenir et de garantir leur alignement.

On comprend que l' « Educational Guidance » aux USA se concentre sur l'orientation scolaire et professionnelle des élèves selon leurs capacités et tendances pour leur assurer une vie active autonome et productive.

L'orientation scolaire dans ce pays respecte l'adage qui dit que « les compétences et les expériences de l'élève doivent renforcer chez lui la compréhension de lui-même et celle de son environnement pour lui permettre une parfaite adaptation ».

Aux USA, tous les élèves suivent leurs études jusqu'au lycée où ils doivent se spécialiser dans les filières académiques (générales), professionnelles, commerciales, universito-preparatoire et cela en fonction de leur capacités et de leurs intérêts. Dans l'opération de l'orientation scolaire, ils recourent à des tests psychologiques pour détecter les prédispositions, les capacités et les tendances de chaque élève en plus des évaluations de ses connaissances. Ces tests permettent aux élèves de prendre conscience que leurs chances de réussite sont plus importantes dans un domaine plutôt qu'un autre.

Il est important de noter aussi l'existence de lycées spécialisés aux USA qui exigent de leurs élèves la présentation de certificats d'aptitudes et de compétences.

Il est important de noter que tous les services d'orientation aux USA sont continuels au niveau de toutes les écoles que ce soit celles à quatre (04) ans après l'école primaire ou celles à huit (08) ans ou encore celles à trois (03) ans. Par conséquent, l'orientation accompagne l'élève durant tout son cursus scolaire et se poursuit même dans sa vie active.

II.5.3 L'orientation scolaire au Japon

Avant de parler de l'orientation scolaire au Japon, il est important de présenter son système éducatif bâti après la deuxième guerre mondiale qui est composé de trois cycles : primaire, secondaire et universitaire.

- **Le cycle primaire** : Il dure six (06) ans. Il est précédé par la période des jardins d'enfants ;

- **Le cycle secondaire** (Lycée) : Il a une durée de six (06) ans. Il est composé de deux phases :
 - **La phase secondaire inférieure** qui dure trois (03) ans. Tous les élèves du primaire sont admis à cette phase. La politique d'orientation en vigueur relative à la fin de la phase secondaire inférieure se base sur deux facteurs : l'admission à l'examen des capacités scolaires et le registre de l'élève des trois dernières années. Les élèves ayant obtenu les meilleures notes à l'examen et dont le registre confirme leur autonomie, choisissent leur orientation de plein gré, les autres seront orientés selon leurs capacités ;
 - **La phase secondaire supérieure** où sont admis tous les élèves de la phase précédente reçus à l'examen de capacités scolaires. Sa durée est de trois (03) ans. Cette phase est subdivisée en deux :
 - **Enseignement secondaire général** : le pourcentage d'admission est de 74% ;
 - **Enseignement secondaire (technique)** avec un pourcentage d'admission de 26%. Il offre trois spécialités : agriculture, industrie, commerce. A noter que 80% de ces spécialités sont embauchées à la fin de leur cycle, les 20% restant poursuivent leur études dans les facultés techniques.

L'orientation scolaire au Japon est continue dans le cycle secondaire et elle se fait par des équipes constituées des directeurs, enseignants principaux dans l'optique d'assurer la liberté du choix aux élèves.

- **Le cycle universitaire** : Il englobe les petites facultés, les facultés techniques et les universités. La durée minimale est de quatre (04) années.

Le Japon a réussi à offrir une scolarité à tous ses enfants en garantissant la continuité des études au maximum selon les capacités, les ambitions et les prédispositions des élèves. Ce pays a réussi le challenge de garantir l'évolution technologique. L'intérêt porté à l'enseignement technique a assuré une main d'œuvre qualifiée prête à l'exploitation. Sachant qu'il existe deux formules pour cet enseignement : jour et nuit.

Dans le sillage des efforts consentis dans l'amélioration du système éducatif japonais et son alignement aux besoins du monde du travail, un intérêt particulier a été accordé à l'orientation scolaire et professionnelle communément. Ce qui a poussé le haut conseil de l'éducation créé en 1984 à insister sur l'importance de cette orientation dans les établissements scolaires et secondaires en mettant l'accent sur l'amélioration de sa qualité. C'est pourquoi un panel pluridisciplinaire est offert aux élèves leur garantissant ainsi des branches d'études répondant à leurs ambitions, capacités et leurs besoins. Les spécialistes de l'orientation ont

toujours œuvré à satisfaire les attentes des élèves, l'opinion publique et l'élite. Ajouté à cela, ils se sont intéressés de très près à la qualité de l'enseignement notamment le volet de l'orientation en introduisant des innovations garantissant la liberté et la concurrence. Les développements réalisés dans le système éducatif permettent à ce pays de garder son haut niveau économique et social.

II.5.4 L'orientation scolaire en France

Les évolutions économiques et politiques qu'a connu l'Europe dans l'histoire ancienne ont influencé les systèmes éducatifs français à travers notamment les réformes qui ont toujours aspirées à être à la hauteur des évolutions industrielles et technologiques qui exigent des individus capables d'affronter toutes sortes de situations et d'y réagir positivement et de vaincre les effets de la mondialisation.

L'orientation scolaire en France se base sur les principes de l'éducation nationale présentés comme suit (Berou, 2010) :

- L'enseignement démocratique qui garantit l'égalité des chances pour tous les enfants du peuple, riche ou pauvre, dans les villes ou en dehors, les minorités et émigrants, fille ou garçon sans tenir compte des religions ;
- Largesses des offres d'enseignements : littérature, économie, technologie, sciences industrielle, naturelle sociale professionnelle etc. ;
- Liberté des apprenants dans le choix de leur spécialité convenant à leurs prédispositions et leurs capacités ;
- Relever le défi de la révolution technologique et informatique pour assurer le leadership mondial ;
- Intégration de la vision internationale dans les diverses étapes de l'enseignement pour renforcer la coopération, l'accord et la paix à travers le monde ;
- Développement du préscolaire ;
- Offre de chances et possibilités scolaires qui correspondent aux capacités des élèves et leurs tendances.

L'orientation scolaire dans le système éducatif français débute avec l'observation des élèves dès les premières années de l'enseignement moyen. Le conseil de classe chargé de l'orientation est constitué de l'enseignant principal et d'autres enseignants, le médecin scolaire, l'assistant social et le conseiller psychologique. Ce conseil décide de l'orientation à réserver à chaque élève.

A la fin du cycle moyen, le conseil de niveau constitué des enseignants principaux du même niveau et des conseillers psychologiques, oriente les élèves à des lycées académiques ou techniques ou des écoles techniques ou encore vers la vie active. En classe de terminal, ce conseil va continuer à ajuster l'orientation faite en début de cycle secondaire. Les chargés de l'orientation communiquent sur les différents établissements d'études supérieurs ainsi que les concours et examens d'accès.

La France accorde une importance particulière à l'orientation scolaire et professionnelle dans le sens où elle en fait un processus continu pour la construction d'un projet individuel (personnel) de formation et d'intégration social et professionnelle. Ce projet est adopté par les élèves à partir du cycle moyen et tout au long du cycle secondaire. Il se base sur les vœux, les tendances et les capacités de l'élève associant les parents, les enseignants et tous les intervenants dans l'enseignement et sa qualité.

II.6 Contraintes de l'orientation scolaire

Maillon faible du système éducatif, les conseillers à l'orientation sont souvent frustrés de ne pouvoir jouer pleinement leur rôle. Ils n'ont pas toujours, la possibilité d'accompagner les élèves, notamment ceux qui sont en fin de cycle. C'est, en effet, à ce moment que beaucoup d'élèves se retrouvent désarmés, ne sachant vers quelle filière se tourner, ni vers quelles études universitaires se pencher, lorsqu'il s'agit de nouveaux bacheliers. Pourtant, c'est le cœur même du métier de conseiller à l'orientation. Au fil des années, ces derniers se sont retrouvés à accomplir plus de tâches administratives que de conseil auprès des élèves (Imès, 2018).

Dans la plupart des établissements, le travail d'aide à l'orientation scolaire se résume très souvent à la recherche de documentation sur les métiers à l'aide des outils papier et informatique disponibles et éventuellement à la visite du conseiller (e) d'orientation sur recommandation des enseignants ou systématiquement. Il est noté que le souci de l'orientation est réel, mais que les communautés éducatives sont extrêmement démunies sur plusieurs plans (Devauchelle, 2007) :

- La connaissance de la vie professionnelle et des parcours professionnels ;
- La capacité à l'aide individualisée et à l'accompagnement du projet personnel et professionnel de l'élève ;
- La difficulté rencontrée par les établissements au sujet de la possibilité de placer les élèves en stage dans les entreprises ;
- Parfois, la diversité des dispositifs en place montre bien la difficulté à globaliser la question et le renvoi vers chaque établissement de la réflexion sur cette question.

A la fin du cycle obligatoire, à savoir le moyen, beaucoup de collégiens se retrouvent à la croisée des chemins. Certains font face au dilemme du choix de la filière au niveau du secondaire tandis que, pour d'autres, c'est le saut vers l'inconnu puisqu'ils n'ont pas pu obtenir la moyenne requise pour accéder au lycée. C'est à ce niveau-là que l'orientation prend toute son importance. Il s'agit de guider les futurs lycéens vers la filière dans laquelle ils ont le plus de chance de s'épanouir mais également d'éviter de perdre dans la nature ceux qui n'ont pas les capacités de poursuivre leurs études. Les conseillers à l'orientation peuvent à ce moment proposer les alternatives que sont les centres de formation professionnels.

Au niveau du lycée, il s'agit d'anticiper et de guider au mieux les lycéens dans le choix de leurs futures études universitaires. Très souvent, les nouveaux bacheliers n'ont pas suffisamment de connaissance sur les possibilités qui peuvent s'offrir à eux. Ils ne s'intéressent à la question qu'à l'approche de la période d'orientation. Parfois, le laps de temps qui leur est laissé pour s'inscrire ne leur donne pas la possibilité de prendre connaissance de l'ensemble de l'offre de formation universitaire. Le conseiller à l'orientation a pour rôle de, justement, les aider à construire un projet et à se donner les moyens de le réussir en faisant les bons choix au bon moment. Très souvent, un mauvais choix est à l'origine de l'abandon des études. Ceci peut être évité à condition que le projet fasse l'objet d'un bon accompagnement. (Imès, 2018).

Les difficultés rencontrées par les élèves dans leur orientation constituent des sources de stress. En effet, selon (Lacoste, et al., 2005), la nécessité pour les élèves du collège et du lycée de faire des choix d'orientation entre les différentes filières ou les différentes professions qui leurs sont proposées par l'institution selon leurs capacités, leurs résultats scolaires ou selon des contraintes socio-économiques est une source de difficultés pour 85,6 % des adolescents scolarisés entre 12 et 20 ans. Ces difficultés d'origine externe ou interne, favorisent un état de stress chez les élèves. Les résultats montrent que ce sont en particulier les difficultés internes comme le doute de soi qui expliquent le niveau de stress perçu. Avec l'avancée en âge, les élèves tendent à être de plus en plus stressés et se défendent en externalisant leurs difficultés d'orientation.

La notion d'égalité et, par extension celle d'inégalité dans les trajectoires scolaires et professionnelles, constituent un idéal des sociétés modernes démocratiques, qui refusent l'héritage de statuts sociaux au profit de la valorisation des qualités personnelles permettant de garantir l'efficacité et l'équité d'une société (Vignoli, 2012).

On note que l'orientation scolaire est un processus complexe et difficilement factorisable puisqu'il dépend des individus. Il est ainsi, difficile de formaliser et d'automatiser tous les

aspects liés à l'orientation scolaire. L'intelligence artificielle (IA) notamment les heuristiques d'extraction de connaissances peuvent intervenir pour aider à relever certaines contraintes de l'orientation scolaire.

Conclusion

Dans ce chapitre nous avons présenté l'orientation scolaire qui est une branche cruciale dans le domaine de l'éducation. Nous avons présenté la notion de l'E-orientation qui s'accommode à celle de l'e-éducation. Nous avons donné quelques définitions de l'orientation scolaire, ses objectifs, les moyens qu'elle met en œuvre. Nous avons, ensuite introduit les logiciels d'aide à l'orientation LAO. Nous avons présenté l'orientation scolaire à travers le monde, un pays pour chaque continent en insistant sur le cas Algérien. Nous avons aussi introduit les contraintes liés à l'orientation scolaire.

A travers notre étude bibliographique, nous avons constaté que l'orientation scolaire est une question capitale dans le domaine de l'éducation. Pour répondre aux besoins actuels, elle met en œuvre les outils offerts par le développement technologique notamment les TIC. Les logiciels d'aide à l'orientation n'ont pas encore exploité de manière significative les atouts de l'intelligence artificielle notamment l'analyse d'opinion en tant que telle, bien que celle-ci soit un facteur important dans les systèmes d'aide à la décision. C'est dans ce cadre que s'insère notre étude, qui consiste à proposer un système d'E-orientation scolaire intégrant l'analyse d'opinion. Dans le chapitre suivant, nous détaillons cette étude.

Partie B :

Propositions et leurs validations

Chapitre III : Proposition D'un Système De E-Orientation Scolaire Basé Sur l'Analyse d'Opinion

Introduction

L'orientation scolaire est un processus de prise de décision. Sa réussite est étroitement liée aux critères qu'elle met en œuvre afin de guider les élèves tout au long de leur cursus scolaire. Pour l'améliorer, nous proposons un système de E-orientation scolaire basé sur l'analyse d'opinion. Ce paramètre est largement introduit dans divers domaines et a prouvé son efficacité notamment dans les systèmes de prise de décision. Nous proposons d'intégrer l'opinion des intervenants dans le processus de l'orientation scolaire à savoir : les élèves, leurs parents ainsi que leurs enseignants. Ce système intègre également un deuxième critère qui est le style d'apprentissage de l'élève. Ceci en vue de donner toutes les chances de réussite à l'élève aussi bien lors de son parcours scolaire et étudiantin que dans sa vie professionnelle à travers un système d'aide à l'orientation des élèves en fin de cycle.

Dans ce système, l'opinion est modélisée, recueillie, analysée, amplifiée par un taux de certitude et synthétisée en vue d'aboutir à un score d'orientation optimale qui englobe à la fois, les moyennes des notes, les aptitudes représentant les styles d'apprentissage des élèves et l'opinion des différents intervenants (élèves, parents et enseignants). La carte d'aide à l'orientation de l'élève présente les domaines d'études triés selon les scores calculés. L'élève peut ainsi être orienté vers le domaine le mieux classé.

III.2 Contexte général de nos propositions

Nous proposons un système d'aide à l'orientation scolaire OPinOR (OPinion pour l'ORIENTATION), où le processus d'orientation se base sur trois dimensions (Cf. figure 15) : la dimension opinion, la dimension personnelle et la dimension capitalisation.

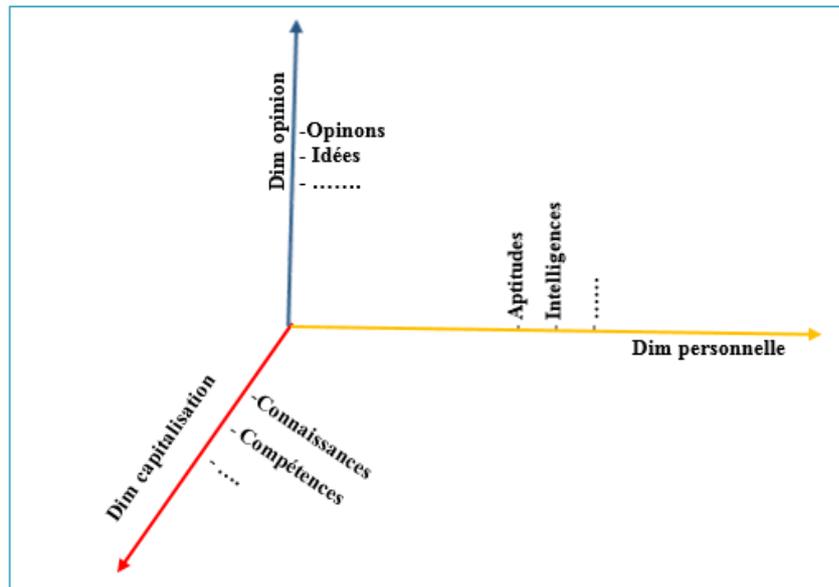


Figure 15 Dimensions de OPinOR.

La dimension opinion : est une dimension explicite, qui inclut les aspects exprimés par l'élève comme son opinion, ses idées, ses attentes et ses motivations. Dans notre étude, on s'intéresse à l'extraction de connaissances à partir de l'opinion des intervenants dans le processus de l'orientation. Cette connaissance explicite est utilisée pour recommander les domaines d'études les plus propices à chaque élève.

La dimension personnelle : est une dimension implicite qui concerne tout ce qui est natif chez l'élève, comme les traits de personnalité, les aptitudes, les intelligences, etc. Dans notre cas, on s'intéresse aux styles d'apprentissage de l'élève définis par les intelligences multiples de Gardner.

La dimension capitalisation : est une dimension basée sur les acquisitions de l'élève par ses propres efforts. On s'intéresse particulièrement à ses performances académiques représentées par les notes obtenues lors des évaluations notamment en fins de cycle comme le baccalauréat.

Le processus d'aide à l'orientation scolaire proposé est basé sur l'analyse d'opinion. Son objectif principal est d'améliorer la qualité de l'orientation à la fin de chaque cycle scolaire à l'instar de l'orientation universitaire ou professionnelle à la fin du cycle secondaire. Il se base à la fois sur les évaluations scolaires des élèves, leurs aptitudes, leurs opinions ainsi que celles de leurs enseignants et de leurs parents dans le choix des domaines d'études. Ce processus est conçu de sorte à recueillir périodiquement d'une part les opinions des élèves et celles leurs parents au sujet des matières dispensées et d'autre part l'opinion des enseignants de ces mêmes matières au sujet de chaque élève.

Pour ce faire, des enquêtes sont organisées pour recueillir une «opinion formatée» à travers des questions à échelle d'évaluation et, une «opinion libre » à travers des questions ouvertes. L'opinion formatée fera l'objet d'une analyse grâce au calcul statistique comme la moyenne des notes. L'opinion libre, quant à elle, fera appel aux techniques d'analyse d'opinion pour extraire la polarité de l'opinion émise à l'égard de chaque matière. Les résultats obtenus sont synthétisés pour aboutir à une valeur d'opinion par domaine. Ces valeurs d'opinion, conjuguées aux aptitudes et évaluations scolaires sont exploitées pour orienter les élèves vers les filières qui leur sont le plus favorables (Lazib, et al., 2016 -a; Lazib, et al., 2016 -b; Lazib, et al., 2019).

Bien que l'opinion soit un paramètre qui a démontré son efficacité dans les systèmes de prise de décision et de recommandation, elle reste, toutefois, changeante et incertaine. C'est pourquoi, dans cette étude, nous nous sommes penchés, aussi, sur la crédibilité de cette opinion. Nous avons alors étudié dans (Lazib & Bouarab-Dahmani, 2017) l'engagement de l'élève lors de l'expression de son opinion. Les résultats obtenus améliorent de manière assez sensible les valeurs de l'opinion et permettent de distinguer l'opinion triviale de l'opinion fondée à travers le calcul d'un taux de certitude de l'opinion. Nous présentons dans ce qui suit les détails de nos propositions.

III.3 Architecture du système OPinOR

Le système d'aide à l'orientation scolaire proposé est composé de cinq principaux modules : Le module de paramétrage, le module de construction de la ressource lexicale DICO, le module de traitement de l'opinion, le module de mesure de certitude et enfin le module de synthèse d'orientation (Cf. Figure 16). Les entrées de ce processus sont : les notes et les aptitudes des élèves ainsi que les opinions des différents intervenants. Il retourne en résultat la carte d'orientation pour chaque élève lui préconisant les domaines d'études propices. Nous détaillons dans ce qui suit ces différents modules.

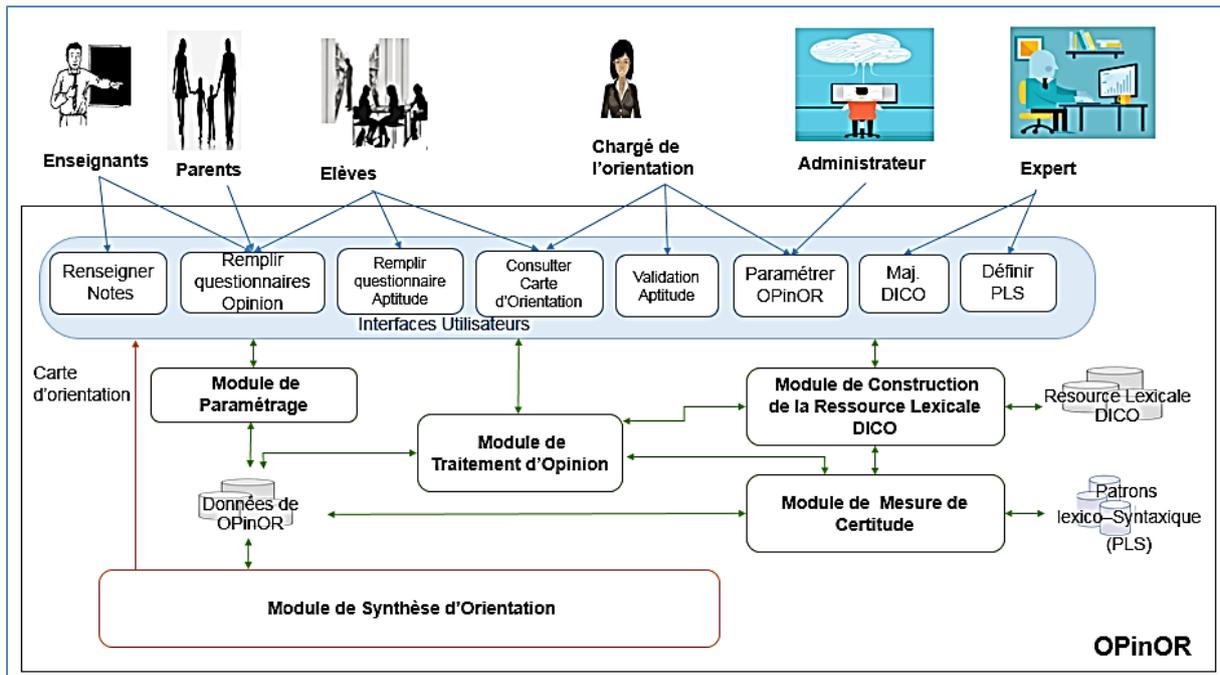


Figure 16 Architecture d'OPinOR.

III.4 Module de paramétrage

Ce module permet d'initialiser et de tenir à jour tous les paramètres nécessaires au fonctionnement du système proposé. Il permet notamment :

- De définir la liste des intervenants (élèves, parents et enseignants), la liste des matières dispensées et celle des domaines d'études universitaires ou professionnelles, bien qu'on se limite dans cette thèse à l'orientation universitaire.
- De connecter les matières dispensées aux domaines d'études universitaires. Nous proposons à ce niveau de pondérer par des coefficients notés (α_i), les notes des matières pré-requises de chaque domaine d'étude. Ces coefficients définissent l'importance des notes obtenues dans la matière (m_i) pour le domaine (D_j) ciblé. Nous proposons de définir selon le même principe, une pondération des matières dispensées relativement à l'opinion de chaque intervenant. En effet, pour pouvoir étudier un domaine universitaire particulier, il est important que l'opinion de l'élève envers les matières pré-requises à ce domaine, soient positives. A partir de là, nous recommandons de pondérer l'opinion des intervenants dans le processus de l'orientation scolaire avec des fonctions affines de α_i notées $a\alpha + b$. Nous assignons les fonctions f , g , h respectivement pour les élèves, leur parents et leurs enseignants. Nous proposons de définir les valeurs des paramètres a et b de ces fonctions selon l'influence de l'intervenant dans l'orientation. On suppose que la plus grande valeur doit être attribuée à

l'opinion de l'élève, viennent ensuite celle de ses enseignants et celle de ses parents. Toutefois, une étude en collaboration avec les sciences de l'éducation est souhaitable à ce niveau afin de couvrir tous les aspects relatifs aux choix de ces paramètres. La figure 17 illustre la connexion des matières aux domaines d'études avec les coefficients α_i , et les fonctions f , g , h .

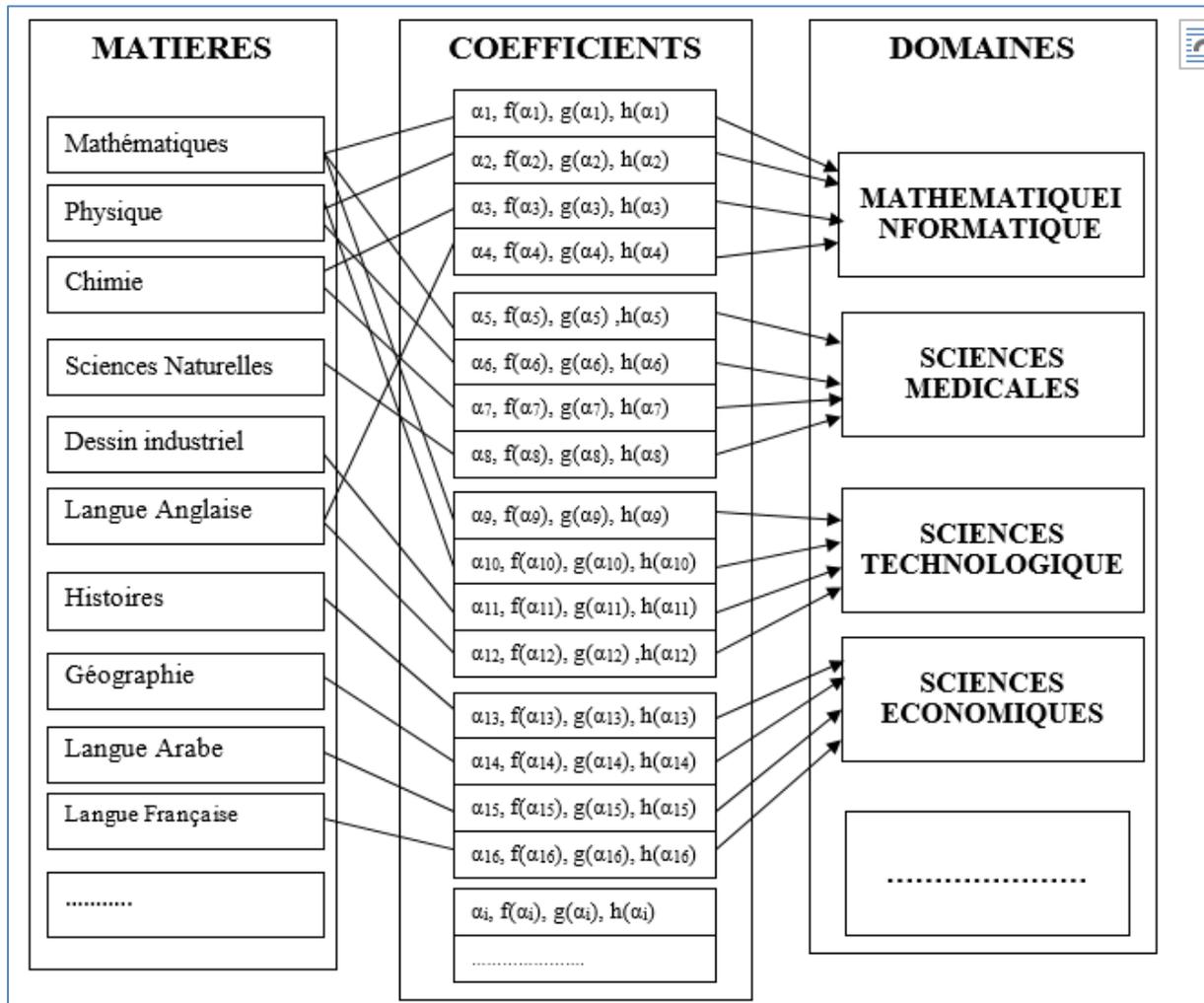


Figure 17 Liaisons pondérées des matières aux domaines d'études.

- D'introduire la liste des aptitudes représentées par les intelligences multiples de Gardner (Gardner, 2011).
- De rattacher ces aptitudes aux domaines d'études selon l'étude menée par (Garceau, 2010), en pondérant les aptitudes avec des coefficients (γ_i) (Cf. fig.18).
- De recueillir les aptitudes implicites de chaque élève grâce au questionnaire proposé par Garceau (Garceau, 2010) (Cf. annexe 1). Ensuite, de faire valider la liste des aptitudes recueillies de chaque élève par les conseillers d'orientation.

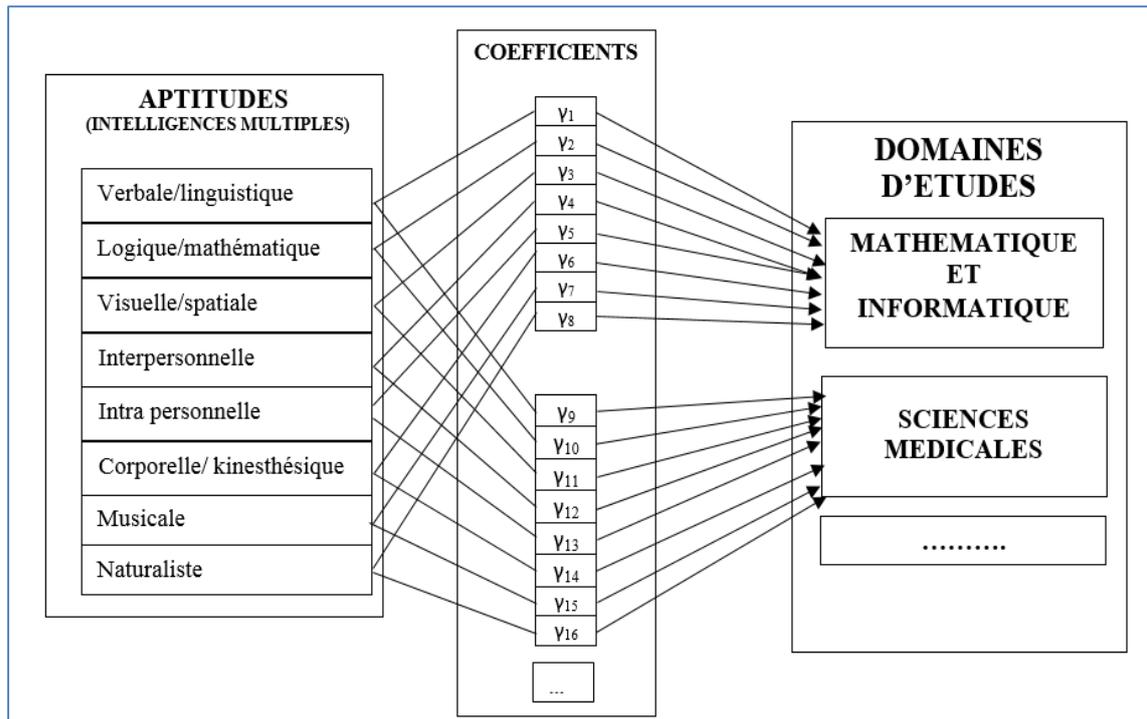


Figure 18 Liaisons pondérées des aptitudes des élèves aux domaines d'études.

- D'indiquer les valeurs des polarités d'opinion (1 : très défavorable, 2 : défavorable, 3 : neutre, 4 : favorable, 5 : très favorable) selon l'échelle de Linkert (Boone & Boone, 2012). Le tableau 5 illustre la classification de l'opinion selon les valeurs de ses polarités ;

Tableau 5 Valeurs des polarités et classification de l'opinion.

	Classification de l'opinion				
	Polarité Négative		Polarité Neutre	Polarité Positive	
	Très Défavorable	Défavorable	Neutre (objective)	Favorable	Très Favorable
Valeur Opinion Formatée	1	2	3	4	5
Valeur Opinion Libre	1		3	5	
Valeurs d'Opinion Amplifiée	~0	2	3	4	~6

- De fixer les types de questions à fournir aux intervenants. Notons que le questionnaire soumis aux intervenants lors des enquêtes comporte des questions fermées et des questions ouvertes. Les questions fermées sont présentées de façon à avoir des réponses préétablies sous formes de cases à cocher correspondant aux valeurs de l'opinion (Cf. Tab. 5). Les

questions ouvertes auront des réponses sous forme de texte libre. Le traitement de ces réponses permet d'extraire respectivement l'opinion formatée et l'opinion libre ;

- De fixer l'échelle de notation des évaluations scolaires selon les systèmes numériques ou alphabétiques ;
- De définir aussi les coefficients de pondération de l'opinion libre (φ) et de l'opinion formatée (θ), lors du calcul de l'opinion globale de chaque intervenant ;
- D'indiquer les coefficients de pondération de l'opinion par domaine, lors de la synthèse de l'opinion générale : (α) pour l'opinion des élèves, (β) pour l'opinion des enseignants et (λ) pour l'opinion des parents. La figure 31 illustre les niveaux de synthèse de l'opinion ;
- De fixer les coefficients du score d'orientation : (ω) pour l'opinion générale, (σ) pour les moyennes des notes et (γ) pour les aptitudes des élèves.
- En fin de fixer également le seuil de certitude admis (SC).

Nous reviendrons avec plus de détails sur l'utilité de ces paramètres dans les modules suivants.

III.5 Module de construction de la ressource lexicale DICO

L'analyse d'opinion recourt, dans son traitement, à différentes approches comme nous l'avons indiqué dans le chapitre I. Il existe trois approches principales : les approches symboliques dites aussi lexicales, elles reposent sur l'utilisation de ressources lexicales et de dictionnaires. Les approches statistiques qui recourent à l'apprentissage automatique. Les approches hybrides, elles mixent entre les deux approches sus citées. Dans notre cas, nous avons choisi d'utiliser une approche symbolique basée sur une ressource lexicale. Pour mettre en œuvre cette méthode, il a fallu construire la ressource que nous appellerons dans la suite de ce travail DICO.

DICO constitue notre ressource lexicale de la langue française dédiée à l'analyse d'opinion pour l'éducation. La construction de cette ressource s'est basée sur SentiWordNet et WOLF (Khan , et al., 2016; Hemmatian & Sohrabi, 2017). DICO est composée, en partie, des synsets de WOLF 0.6 avec les polarités correspondantes données par SentiWordNet 3.0. Cette opération repose sur les hypothèses suivantes (Medagoda, et al., 2015) :

- Le sens des mots dans les langues (Français et Anglais) sont identiques ;
- Les scores de polarité calculés pour être utilisés dans l'analyse d'opinion exprimée en langue anglaise sont considérés être identiques lorsqu'il s'agit des mêmes mots en français ;

- Finalement, les POS (Part Of Speech) partie du discours en français, dans les deux langues sont considérées être équivalentes.

La Figure 19 illustre une partie de cette opération.

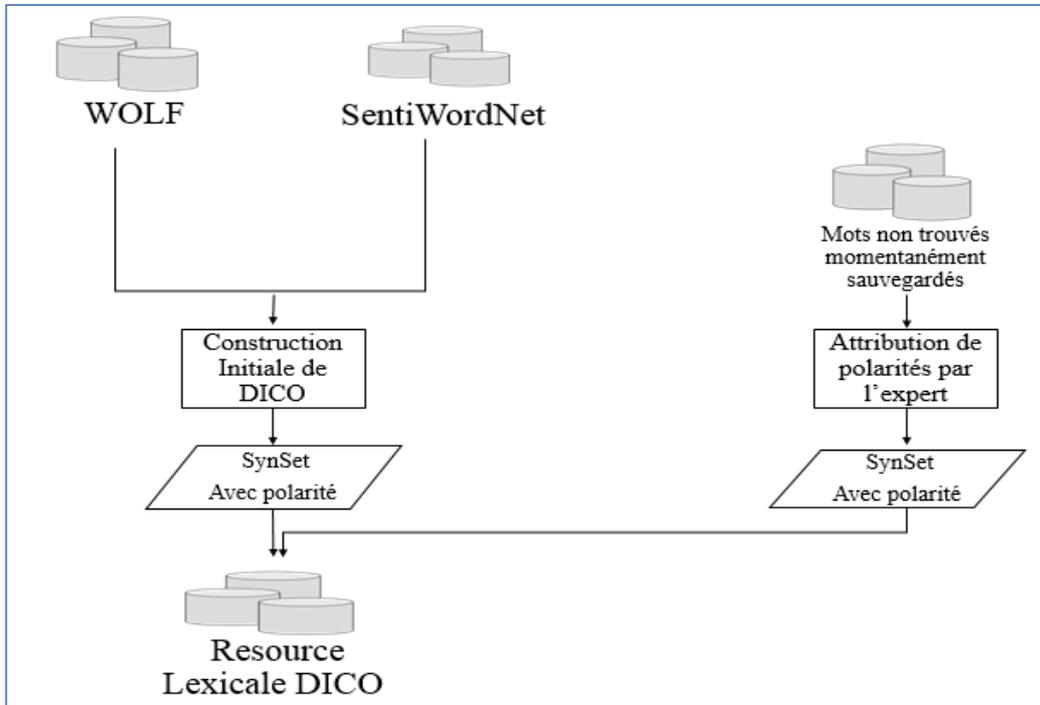


Figure 19 Construction de la ressource lexicale DICO.

Ce module intervient dans la construction initiale de DICO ainsi que sa mise à jour. En effet, lors du traitement de l'opinion libre, si des mots ne se trouvent pas dans DICO, ce module permet de les ajouter après configuration de leurs paramètres (polarités...) par un expert (Cf. figure 23).

Le système que nous avons proposé ne se limite pas uniquement à définir les polarités du texte mais on s'est proposé également de mesurer le taux de certitude contenu dans ces textes. Pour cela il a fallu ajouter une classification des verbes en verbes d'attitude ou d'opinion et autres. Les verbes d'opinion sont eux aussi classés en trois catégories qui indiquent l'ampleur de la certitude (forte, moyenne et faible). Il a fallu alors mettre à jour la ressource lexicale afin de pouvoir intégrer cette notion qui sera détaillée dans la section III.7. La figure 20 illustre la structure de la ressource lexicale DICO.

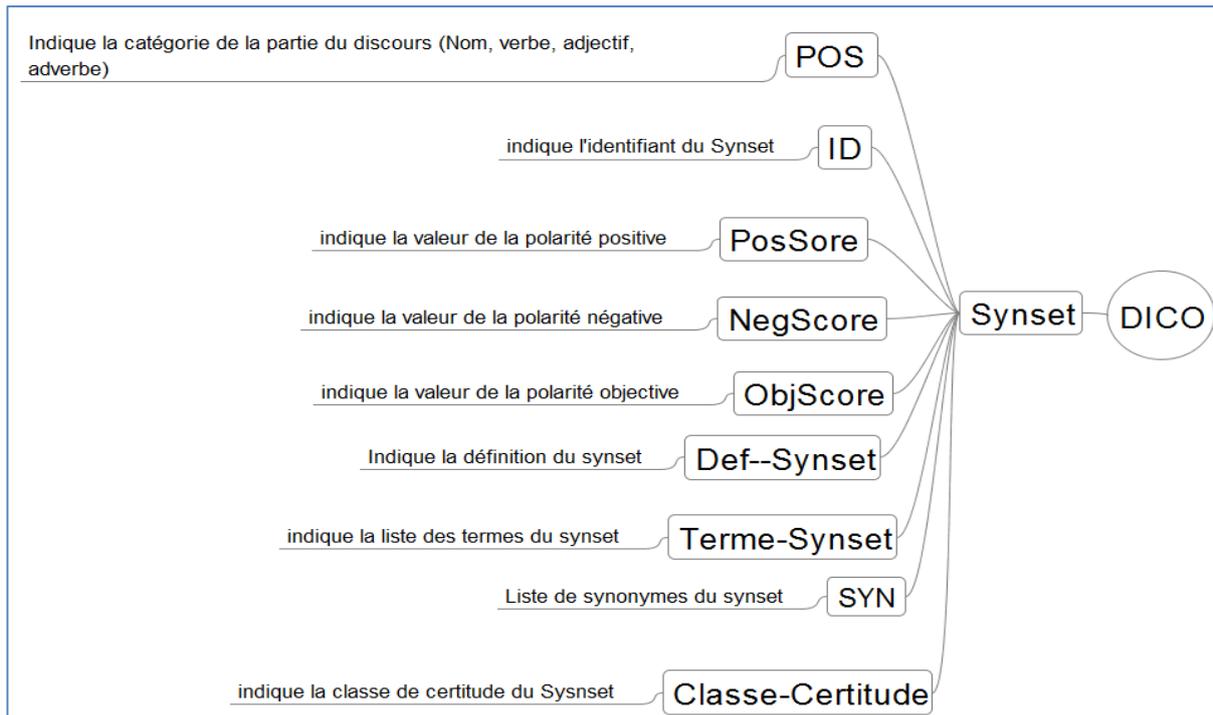


Figure 20 Structure de la ressource lexicale DICO.

III.6 Module de Traitement de l'opinion

Ce module constitue le noyau de notre proposition. En effet c'est dans cette partie du système que les opinions des différents intervenants sont recueillies, analysées, amplifiées par des coefficients de certitude et sauvegardées pour leur utilisation au moment de l'orientation.

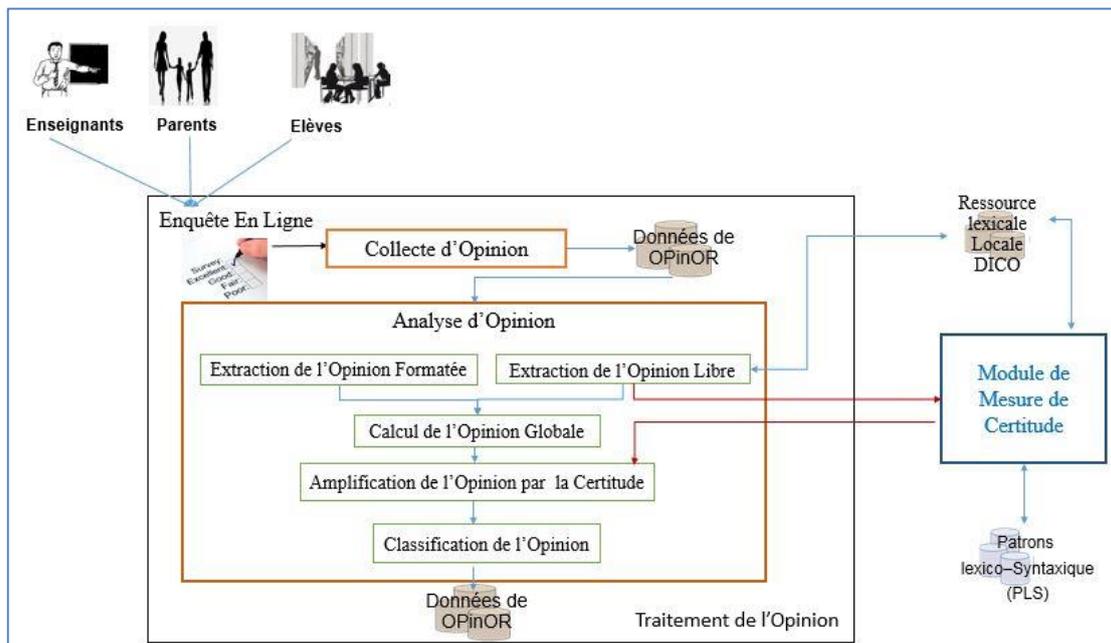


Figure 21 Architecture du module de traitement de l'opinion.

La figure 21 illustre le fonctionnement de ce module. Nous présentons dans ce qui suit les sous modules de ce dernier.

III.6.1 Module de recueil d'opinions

Dans ce module, les questionnaires destinés à la collecte d'opinions chaque fin de trimestre sont créés. Les questionnaires destinés aux élèves et à leurs parents sont constituées de questions fermées et de questions ouvertes sur chaque matière enseignée. Ceux destinés aux enseignants comportent les mêmes types de questions sauf qu'elles portent sur les élèves. La figure 22 illustre un écran de l'espace élève dédié à renseigner les questionnaires d'opinion sur les matières étudiées.

Figure 22 Partie du questionnaire utilisé pour le recueil d'opinions.

Ce module offre aussi la possibilité aux élèves de s'exprimer dans des forums restreints relatifs à chaque matière enseignée. Il permet aussi à chaque élève d'exprimer à la fin du cycle secondaire, son choix explicite des domaines d'études universitaires et ou professionnelles qu'il voudra suivre. Ce choix se fait dans une liste de domaines avec une argumentation de chaque choix en texte libre.

III.6.2 Module d'extraction d'opinions

Ce module reçoit les réponses aux questionnaires collectées précédemment afin de procéder à l'extraction de l'opinion formatée et de l'opinion libre. Les sous modules chargés de ces deux opérations sont décrits ci-après.

III.6.2.1 Extraction de l'opinion formatée

L'opinion formatée (OF_T) est extraite à partir des réponses aux questions fermées ($reqf_T$) relatives à la période (T) selon la formule (III.1) :

$$OF_T = \begin{cases} 1 & \text{si } reqf_T = \text{Très défavorable} \\ 2 & \text{si } reqf_T = \text{Défavorable} \\ 3 & \text{si } reqf_T = \text{Neutre} \\ 4 & \text{si } reqf_T = \text{Favorable} \\ 5 & \text{si } reqf_T = \text{Très favorable} \end{cases} \quad (III. 1)$$

Nous calculons ainsi durant la période "T" :

- OF_{EMT} : l'opinion formatée de l'élève "E" au sujet de la matière "M" ;
- OF_{SEMT} : l'opinion formatée de l'enseignant "S" de la matière "M" au sujet de l'élève "E" ;
- OF_{PEMT} : l'opinion formatée du parent "P" de l'élève "E" au sujet de la matière "M".

III.6.2.2 Extraction de l'opinion libre

L'extraction de l'opinion textuelle libre, que nous appelons dans ce rapport opinion libre, est basée sur la ressource lexicale DICO. Ce module calcule le score de polarité de l'opinion de chaque réponse libre. La figure 23 illustre le fonctionnement de ce module.

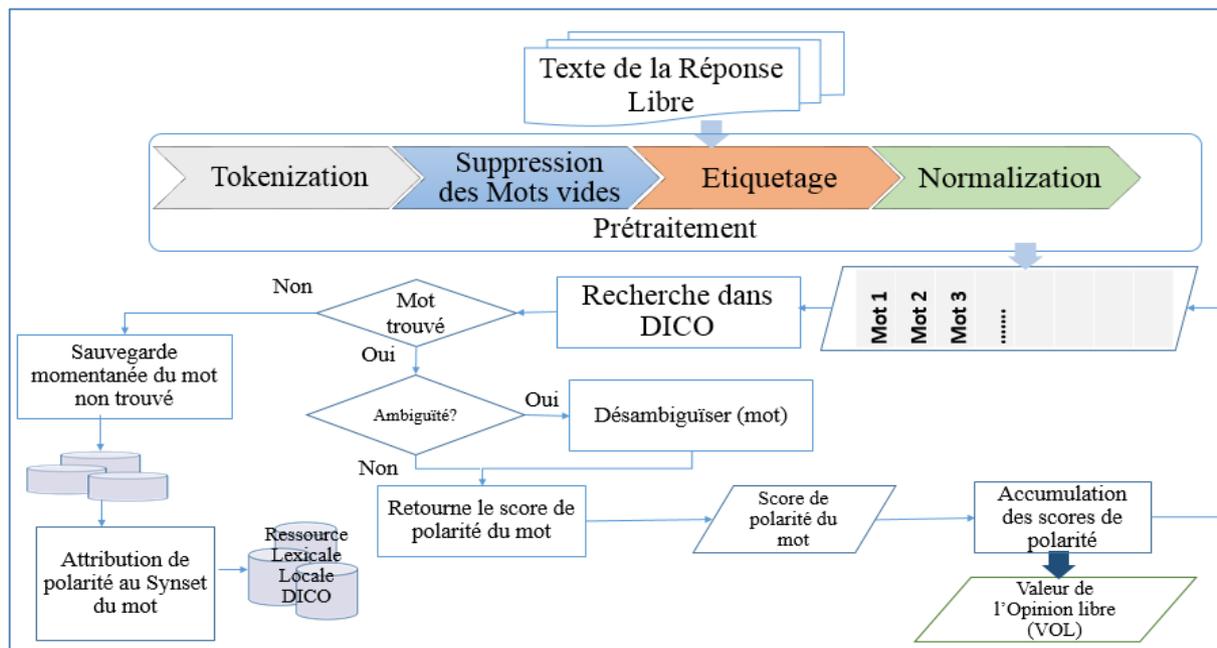


Figure 23 Digramme d'extraction de l'opinion libre.

Le texte des réponses libres, dans sa forme brute, nécessite un prétraitement avant de pouvoir être analysé. Il subit, alors un prétraitement de langage naturel qui consiste en la tokenisation, l'élimination des mots vides, étiquetages (partie du discours), et la normalisation. Ces opérations sont décrites dans ce qui suit (Kreutzer & Witte, 2013):

- La tokenisation est utilisée pour découper la phrase en mots en supprimant généralement les ponctuations.
- Les mots vides sont éliminés des mots qui n'ont pas une incidence sur le résultat de l'analyse, ils sont prédéfinis et ils sont éliminés tout au long de l'analyse.
- L'étiquetage (POS⁸ tagging) est l'opération qui permet d'identifier les différentes parties du discours contenues dans le texte. C'est une étape essentielle dans le traitement du langage naturel.
- La normalisation, est un ensemble d'opérations de stemming et/ou de lemmatisation.
 - Le stemming ou la racinisation en français, est l'opération qui ramène chaque mot à sa forme racine en ignorant toutes ses autres formes POS.
 - La lemmatisation est l'opération qui consiste à remplacer un mot par sa forme canonique. Pour un verbe, le lemme est la forme à l'infinitif. Pour un nom ou un adjectif, le lemme est la forme au masculin singulier.

La figure 24 illustre l'algorithme de la procédure de prétraitement du texte. A noter que ce texte une fois prétraité, il est sauvegardé pour son utilisation ultérieure à des fins d'extraction de certitude.

```

1: Procedure Pretreatment (FA: text, FA_Tab: List of Word with POS);
2: Begin
3:   Tokenization (FA,FA_Tab);
4:   Stop_Word_Removal (FA_Tab);
5:   Tagging (FA_Tab);
6:   Normalization (FA_TAB);
7: End /* Fin de la procédure de prétraitements*/
    
```

Figure 24 Algorithme de la procédure de prétraitement du texte brut.

Après le prétraitement, le texte est découpé en mots étiquetés. Chaque mot est recherché dans DICO. S'il est trouvé, son Synset est retourné en renvoyant ses scores de polarité (positive, négative, neutre). Sinon, la valeur zéro est renvoyée, dans ce cas le mot sera sauvegardé pour

⁸ POS : Part of speech.

une attribution de polarité par un expert et son insertion ultérieure dans DICO tel qu'illustré dans les figures 19 et 23.

Dans le cas où plusieurs Synset correspondent au même mot, le processus propose de recourir à un algorithme de désambiguïsation basé sur l'algorithme de LESK qui possède des applications développées pour WordNet (Lesk, 1986). Notons que cet algorithme est développé par M. Lesk en 1986. Il consiste en une méthode de désambiguïsation qui compte le nombre de mots communs entre les définitions d'un mot (généralement trouvées dans un dictionnaire électronique) et les définitions des mots de son contexte. Le sens retenu correspond à la définition pour laquelle on compte le plus de mots communs avec le contexte (Vasilescu, 2003).

L'algorithme que nous proposons dans la figure 25 repose sur la version simplifiée de l'algorithme de Lesk (Vasilescu, 2003).

```

1: Procedure Desambiguisation (W: Word, Contexte : List_of_tagued_words,
L_Sens: liste_de_synset, SW: ID_Synset);
2: /* W : mot ambigu, Contexte : liste des mots du texte comportant W, L_sens : liste
des synsets de DICO comportant le mot W dont il faut choisir le meilleurs, SW :
identifiant du synset retenu)
3: VAR Best_score, sup :Integer;
4:     S: Synset;
5:     LS: text; LSP: List_of_tagued_words;
6: Begin
7:   Best_score:=0;
8:   SW:=L_Sence.ID;
9:   Sup:=0;
10:  Foreach S in L_sens Do
11:    Begin
12:      LS := Extraire_word (S) ; /* extraction des mots de la définition et des
Synonymes du synset S*/
13:      LSP := Pretreatment (LS, LSP) ;
13:      Sup := Superposition (Contexte, LSP) ;
14:      /*Calcul du nombre de mots semblables entre le contexte du mot à
désambiguïser avec les mots du synset S */
15:      If Best_Score < Sup Then
16:        Begin
17:          Best_Score:=Sup;
18:          SW:=S.ID;
19:        End
20:    End
21:  End /*Fin de la procédure de Désambiguïsation*/

```

Figure 25 Algorithme de la procedure de desambiguisation.

Les scores de polarité correspondant à chacun des mots du texte ainsi retournés sont cumulés pour obtenir les scores de polarité de l'ensemble du texte de la réponse libre. On note Pos (S) le score positif du texte, Neg (S) le score négatif du texte et Obj (S) le score neutre (objectif) du texte.

La classification intervient à la fin du calcul de ces scores selon la formule III.2 :

$$VOL = \begin{cases} -1 & \text{Si } Neg(S) \geq Obj(S) \text{ et } Neg(S) > Pos(S). \\ 0 & \text{Si } (Obj(S) > Neg(S) \text{ et } Obj(S) > Pos(S)) \\ & \text{ou } (Neg(S) = Pos(S)). \\ +1 & \text{Si } Pos(S) \geq Obj(S) \text{ et } Pos(S) > Neg(S). \end{cases} \quad (III.2)$$

Pour mieux illustrer le fonctionnement de ce module, le tableau 6 présente des exemples de calcul de la valeur de l'opinion libre (VOL) pour chacun des cas : positif, négatif, neutre et un cas de négation. A noter que le seul cas de négation traité dans le cadre de cette étude est le *ne-pas*.

Tableau 6 Exemples illustratifs de calcul de l'opinion libre (VOL).

Texte en Français	Après Prétraitement	Score de Polarité			VOL	Classification
		POS-Score	NEG-Score	OBJ-Score		
Les mathématiques c'est difficile	Mathématiques	0,500	0,125	0,375	-1	Négatif
	Etre	0,500	0,125	0,375		
	Difficile	0,000	1,000	0,000		
	Score de polarité du texte	1,000	1,250	0,750		
cette matière est excitante	Matière	0,500	0,125	0,375	1	Positif
	Etre	0,500	0,013	0,488		
	Excitante	0,750	0,000	0,250		
	Score de polarité du texte	1,750	0,138	1,113		
Les mathématiques sont une science exacte	Mathématiques	0,500	0,125	0,375	0	Neutre
	Etre	0,500	0,125	0,375		
	science	0,000	0,000	1,000		
	exacte	0,625	0,000	0,375		
	Score de polarité du texte	1,625	0,250	2,125		
La physique n'est pas ma favorite	Physique	0,625	0	0,375	-1	Négatif
	Ne	0	1	0		
	Etre	0,5	0,125	0,375		
	Pas	0	1	0		
	favorite	0,875	0	0,125		
	Score de polarité du texte	2	2,125	0,875		

La figure 26 illustre l'algorithme du module d'extraction de l'opinion libre.

```

Algorithm 2 Extract Free Opinion
1: Input FA: Text; /*Free answer */
2: Output VOL: Integer; FA_Tab: List of Word with POS;
3: VAR FA_Tab: List of tagged Words;
4:   S_Positif, S_Negatif, S_Neutral :integer;
5:   IDS: ID_Synset;
6:   L_S : liste_de_synset ;
7: Begin
8:   VOL :=0 ; S_Positif :=0 ; S_Negatif :=0 ; S_Neutral :=0 ;
9:   Pretreatment (FA, FA_Tab);
10: Foreach Word in FA_Tab Do
11:   Begin
12:   If Found(Word in DICO) Then
13:     Begin
14:     If Unique (Word in DICO) Then
15:       Begin
16:       S_Positif:= S_Positif + synset(Word, PosScore);
17:       S_Negatif:= S_Negatif+ synset(Word, NegScore);
18:       S_Neutral:= S_Neutral+ synset(Word, ObjScore);
19:     End
20:     Else
21:       Begin
22:       L_S := Liste_synset (Word, Dico) ; /*liste des synset correspondants au
23:                                     mot dans DICO*/
24:       Desambiguisation (Word, FA_Tab, L_S, IDS);
25:       S_Positif:= S_Positif + synset(ID_S, PosScore);
26:       S_Negatif:= S_Negatif+ synset(ID_S, NegScore);
27:       S_Neutral:= S_Neutral+ synset(ID_S, ObjScore);
28:     End
29:     Else
30:       Begin
31:       Store(Word, Temporary);
32:       Attribute_Polarity(SynSet (Word),Polarity(Pos, Neg, Neut));
33:       Store(Synset(Word, Polarity), DICO);
34:     End
35:     End
36:     If (S_Positif>S_Negatif) and (S_Positif>=S_Neutral) then VOL=1
37:     Else If (S_Negatif>S_Positif) and (S_Negatif>=S_Neutral) then VOL=-1
38:     Else If (S_Neutral >S_Positif) and (S_Neutral >S_Negatif) then VOL= 0
39:     Else If (S_Positif=S_Negatif) then VOL= 0
40:   End

```

Figure 26 Algorithme d'extraction de l'opinion libre.

Après extraction de la polarité du texte d'une réponse à une question ouverte, celle-ci est extrapolée dans la même échelle que celle de l'opinion formatée selon la formule III.3.

$$OL_T = \begin{cases} 5 & \text{si } VOL > 0 \\ 3 & \text{si } VOL = 0 \\ 1 & \text{si } VOL < 0 \end{cases} \quad (\text{III. 3})$$

Nous calculons ainsi durant la période "T" :

- OL_{EMT}: L'opinion libre de l'élève "E" au sujet de la matière "M" ;

- OL_{SEMT} : L'opinion libre de l'enseignant "S" de la matière "M" au sujet de l'élève "E" ;
- OL_{PEMT} : L'opinion libre du parent "P" de l'élève "E" au sujet de la matière "M".

Après l'extraction des valeurs de l'opinion libre et de l'opinion formatée de chaque intervenant, son opinion globale est calculée selon la formule III.4 :

$$\begin{cases} OG_T = \theta * OF_T + \varphi * OL_T \\ \theta: \text{coefficient de l'opinion formatée} \\ \varphi: \text{coefficients de l'opinion libre} \end{cases} \quad (III.4)$$

Nous calculons ainsi durant la période "T" :

- OG_{EMT} : L'opinion globale de l'élève "E" au sujet de la matière "M" ;
- OG_{SEMT} : L'opinion globale de l'enseignant "S" de la matière "M" au sujet de l'élève "E" ;
- OG_{PEMT} : L'opinion globale du parent "P" de l'élève "E" au sujet de la matière "M".

III.6.3 Module d'Amplification de l'opinion par le taux de certitude

Les valeurs d'opinions obtenues peuvent subir une amplification par un taux de certitude de leur émetteur. Ce taux de certitude représente, en fait, le taux d'engagement de l'émetteur d'opinion étudié dans (Lazib & Bouarab-Dahmani, 2017). Le module de calcul de certitude reçoit en entrée le texte prétraité des réponses libres et renvoie un coefficient d'amplification (P) basé sur le taux de certitude (CR) exprimé dans chaque texte comme cela est détaillé dans la section IV.7. Afin de prendre en compte ce taux d'amplification dans le processus d'orientation nous proposons d'amplifier les opinions globales calculées ci haut. Le but étant de pouvoir distinguer les opinions exprimées avec certitude de celles qui ne le sont pas. On ajoute à l'opinion globale le taux d'amplification lorsque celle-ci est positive. On le soustrait lorsqu'elle est négative. Le cas de l'opinion neutre n'est pas concerné par l'amplification. L'amplification de l'opinion est donnée par la formule III.5.

$$OGA_T = \begin{cases} OG_T + P & \text{Si } 3.5 \leq OG_T < 5 \\ OG_T - P & \text{Si } 1 \leq OG_T \leq 2.5 \\ OG_T & \text{Si } 2.5 < OG_T < 3.5 \end{cases} \quad (III.5)$$

Où :

- P représente le coefficient d'amplification par la certitude.

Les données calculées à ce stade seront sauvegardées en vue de leur exploitation au moment de l'orientation.

III.6.4 Classification de l'opinion

La dernière étape du traitement de l'opinion est consacrée à la classification de cette dernière. Comme l'opinion globale a subi une amplification par le coefficient (P) et comme ce dernier est compris dans l'intervalle] 0,1[, l'échelle de classification de l'opinion sera étendue à l'intervalle] 0 , 6[, (Cf. tab. 5). La classification de l'opinion se fait selon la formule III.6 :

$$Polarity = \begin{cases} \text{Positive} & \text{Si } 3.5 \leq OGA_T < 6 \\ \text{Neutral} & \text{Si } 2.5 < OGA_T < 3.5 \\ \text{Négatif} & \text{Si } 0 < OGA_T \leq 2.5 \end{cases} \quad (\text{III. 6})$$

Cette classification va servir pour définir les matières que l'élève préfère ainsi que celles qu'il réfute totalement. Ceci, dans l'optique de favoriser ou éliminer certains domaines d'études pour lesquels ces matières sont importantes.

III.7 Module de Mesure de la Certitude de l'Émetteur d'Opinion

Comme nous l'avons introduit dans la section III.2, le traitement de l'opinion, consiste à déterminer sa polarité d'une part, et à examiner sa crédibilité d'autre part. Nous avons supposé qu'une opinion est plus crédible lorsque son émetteur l'exprime avec engagement de sa part, la théorie de l'Appraisal a mis en évidence cet aspect. Dans notre cas on s'est intéressé à cet aspect à travers la mesure de la certitude exprimée dans les textes d'opinion. Cette mesure revient à calculer un taux de certitude qui servira à fixer le coefficient d'amplification de l'opinion. Ainsi, les opinions exprimées avec certitude se verront amplifiées lors du calcul du score d'orientation.

Pour ce faire, il a fallu modéliser l'opinion afin d'intégrer la notion d'engagement de son émetteur, préparer les données en introduisant le paramètre de certitude et mettre en œuvre le processus de mesure du taux d'engagement de l'émetteur d'opinion qui se traduit par le taux de certitude contenu dans les textes des réponses libres. On présente ci-après ces différentes étapes.

III.7.1 Modélisation de la certitude de l'émetteur d'opinion

Pour mettre en œuvre le module de mesure de certitude nous avons élargi la modélisation de l'opinion en adoptant un nouvel attribut. Nous avons introduit, dans la ressource lexicale

DICO, la notion de classification des verbes d'opinion. Aussi, nous avons intégré un ensemble de patrons lexico-syntaxiques pour extraire la valeur de la certitude correspondant à chaque texte.

III.7.1.1 Modélisation de l'opinion

Pour intégrer la notion d'engagement de l'émetteur dans le traitement de l'opinion, que nous mesurons à travers le taux de certitude, nous avons proposé une modélisation de l'opinion inspirée à la fois du modèle de Liu et de la théorie de l'Appraisal introduits dans la section I.7. Nous avons étendu le quintuple de Liu (Liu, 2010) auquel nous avons ajouté l'engagement du locuteur inspiré de la théorie de l'Appraisal (Gardin, 2009; Martin & White, 2003) . Le modèle obtenu est le sextuple (O, fj, Soj, h, t, C) correspondant à :

- L'objet cible de l'opinion (O) ;
- Caractéristiques ou composants de l'objet (fj) ;
- La polarité de l'opinion (Soj) ;
- L'émetteur d'opinion (h) ;
- Le temps où l'opinion est exprimée (t) ;
- L'engagement de l'émetteur d'opinion (C).

La figure 27 illustre cette modélisation avec les instanciations correspondantes.

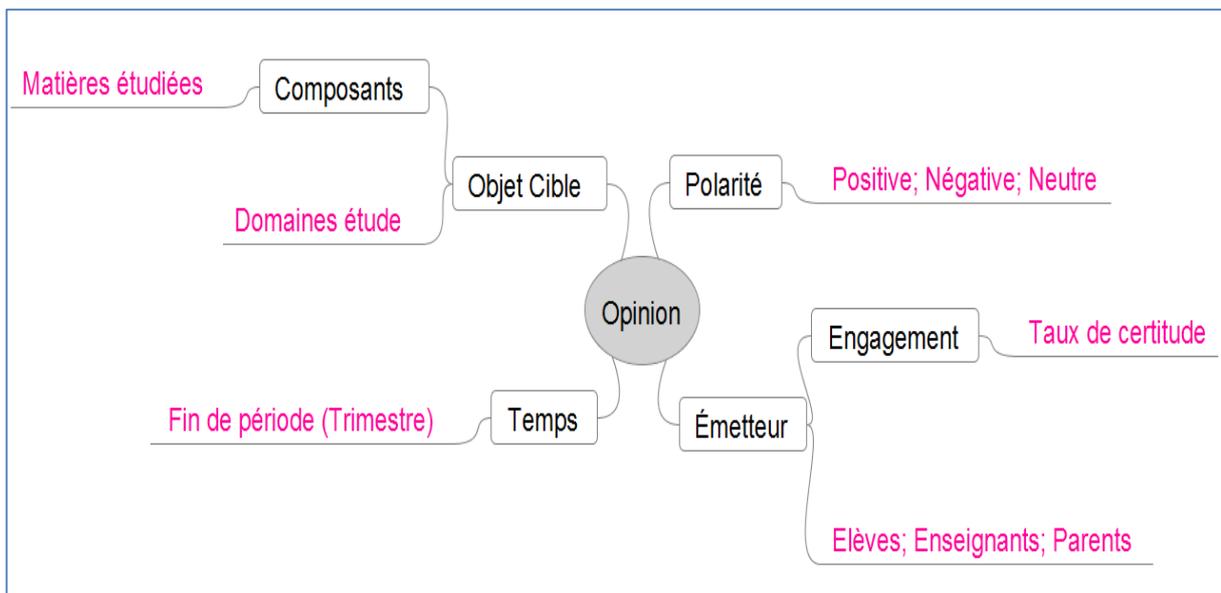


Figure 27 Modélisation de l'opinion.

III.7.1.2 Les patrons lexico-syntaxiques

Les patrons lexico-syntaxiques sont des structures représentant des schémas récurrents du langage. Ils sont utilisés dans le domaine du traitement automatique de la langue pour repérer

des schémas langagiers dont ils sont l'abstraction (Dragos & Jaulent, 2010). Il existe trois modèles pour élaborer les patrons lexico-syntaxiques :

- **Les structures prédicat-argument** : Ce modèle se présente sous forme de triplet comportant le verbe et ses arguments. On le retrouve généralement dans les approches d'extraction d'information basée sur les verbes. Sa limite réside dans le fait qu'il ne permet pas de traiter des formes langagières non basées sur le verbe (Wattarujekrit, et al., 2004).
- **Le modèle de sous arbres** : Ce modèle vient enrichir le précédent avec une représentation sous forme d'arbre dont la racine est le verbe. Ses nœuds représentent le reste des mots, les branches représentent les liens syntaxiques existants. Malgré cela, ce modèle présente les mêmes limites que le précédent (Greenwood, et al., 2005).
- **Les chemins de dépendance** : Ce modèle repose sur les chemins de dépendances qui sont construits sur la base des résultats fournis par un parseur. Un parseur construit un arbre de dépendance en identifiant les liens syntaxiques existants entre les différents mots d'une phrase. Un chemin de dépendance représente la succession des arcs reliant deux mots dans un arbre. A l'inverse de ses prédécesseurs, ce modèle n'impose pas de contraintes de modélisation. C'est pourquoi, il est possible d'élucider différents types de patrons véhiculant l'information recherchée, indépendamment des catégories constituantes (Dragos & Jaulent, 2010).

III.7.2 Préparation des données

III.7.2.1 classification des verbes d'opinions dans DICO

La ressource lexicale DICO, comme présenté dans la section III.5, a été améliorée en introduisant la notion de classification des verbes et adverbes en deux catégories : verbes d'attitude et adverbes modaux. Ces catégories marquent l'engagement des locuteurs lors de l'expression de ses opinions. Une liste de verbes d'attitude qui dénote de la certitude est pondérée selon la méthode décrite dans (Guerry, et al., 1993). Elle attribue pour chaque élément un taux de certitude selon l'échelle de valeurs suivante :

- Possible : taux de certitude inférieur à 50%,
- Probable : taux de certitude contenu entre 50% et 90%,
- Sure : taux de certitude contenu entre 90% et 100%.

Dans le cadre de cette étude, on s'est limité aux verbes d'attitude, les adverbes ne sont pas étudiés. La figure 28 illustre la classification des verbes et adverbes opérés dans la ressource lexicale.

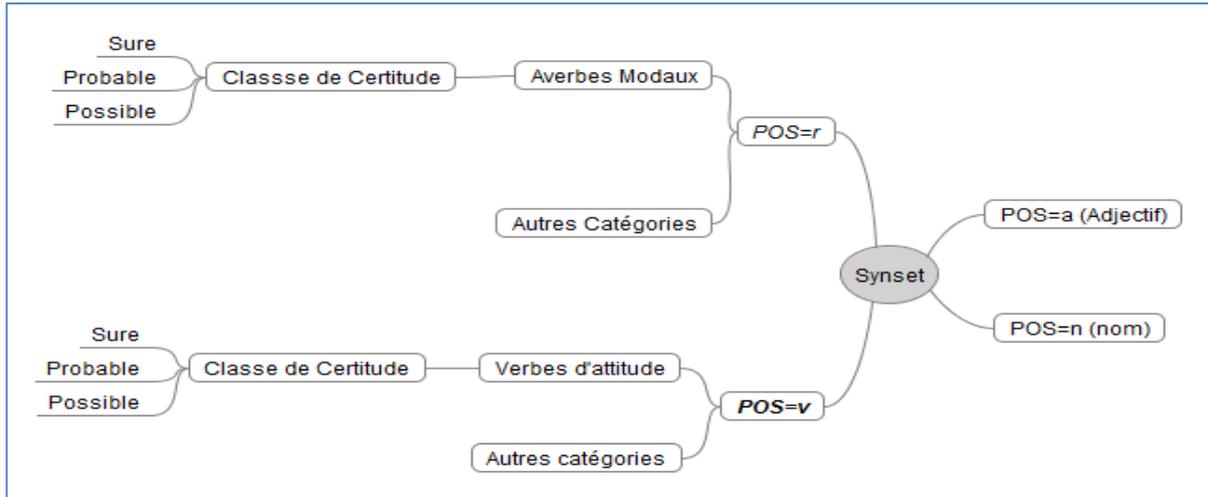


Figure 28 La ressource lexicale avec la classification des verbes et adverbes.

III.7.2.2 Patrons lexico syntaxique utilisés

Pour pouvoir attribuer à une phrase un taux de certitude, nous utilisons les patrons lexico syntaxiques. Le modèle retenu est le chemin de dépendance car, comme déjà indiqué précédemment, il n'impose pas de contraintes particulières relatives aux catégories syntaxiques. Bien qu'actuellement, on s'intéresse à la recherche et l'extraction de certitude de texte en se basant sur des verbes mais cela n'écarte pas d'autres pistes basées sur les adjectifs ou les adverbes. Nous recherchons d'abord la présence de verbes d'attitude dans la phrase puis le patron lexico-syntaxique qui lui correspond. Notons que les patrons implémentés sont illustrés dans le tableau 7.

Tableau 7 Liste des Patrons lexico-syntaxiques implémentés.

N° Patron	Patron	Verbe d'attitude (complément)	La certitude émise (sens)
1	[je]	Savoir [complément]	Une forte certitude
2	[Nous]	Savoir [complément]	Une forte certitude partagée (relative)
3	[Ils/Il/Elles/Elle]	Savoir [complément]	Une faible certitude basée sur une forte certitude des autres
1	[je]	Vouloir [complément]	Une certitude probable
2	[Nous]	Vouloir [complément]	Une certitude probable partagée (relative)
3	[Ils/Il/Elles/Elle]	Vouloir [complément]	Une faible certitude basée sur une certitude probable des autres.
1	[je]	Croire [complément]	Une certitude possible non prouvée.
2	[Nous]	Croire [complément]	une certitude possible partagée (relative).
3	[Ils/Il/Elles/Elle]	Croire [complément].	Une faible certitude basée sur une certitude possible des autres

Le taux de certitude attribué à chaque patron (modèle) représente le taux d'engagement de l'émetteur d'opinion exprimée dans la phrase. Chaque patron décrit précédemment aura un taux de certitude pour chaque classe de verbe d'attitude. Le tableau 8 illustre l'attribution de taux de certitude des patrons lexico-syntaxiques (PLS) selon les classes de verbes d'attitude (CVA).

Tableau 8 Taux de certitude des PLS selon CVA.

Classe CVA	Patron 1 [Je]	Patron 2 [Nous]	Patron 3 [Ils/ II/Elles/ Elle]
Sure	0,98	0,96	0,92
Probable	0,9	0,8	0,75
Possible	0,36	0,42	0,66

III.7.3 Mesure de la certitude de l'émetteur d'opinion

Le module de mesure de certitude interagit avec le module de traitement de l'opinion (Cf. fig. 16 & 21). Ce dernier lui soumet le texte étiqueté des réponses libres recueillies lors des enquêtes d'opinion et il reçoit un coefficient d'amplification d'opinion (P). La figure 29 illustre le diagramme de fonctionnement de ce module.

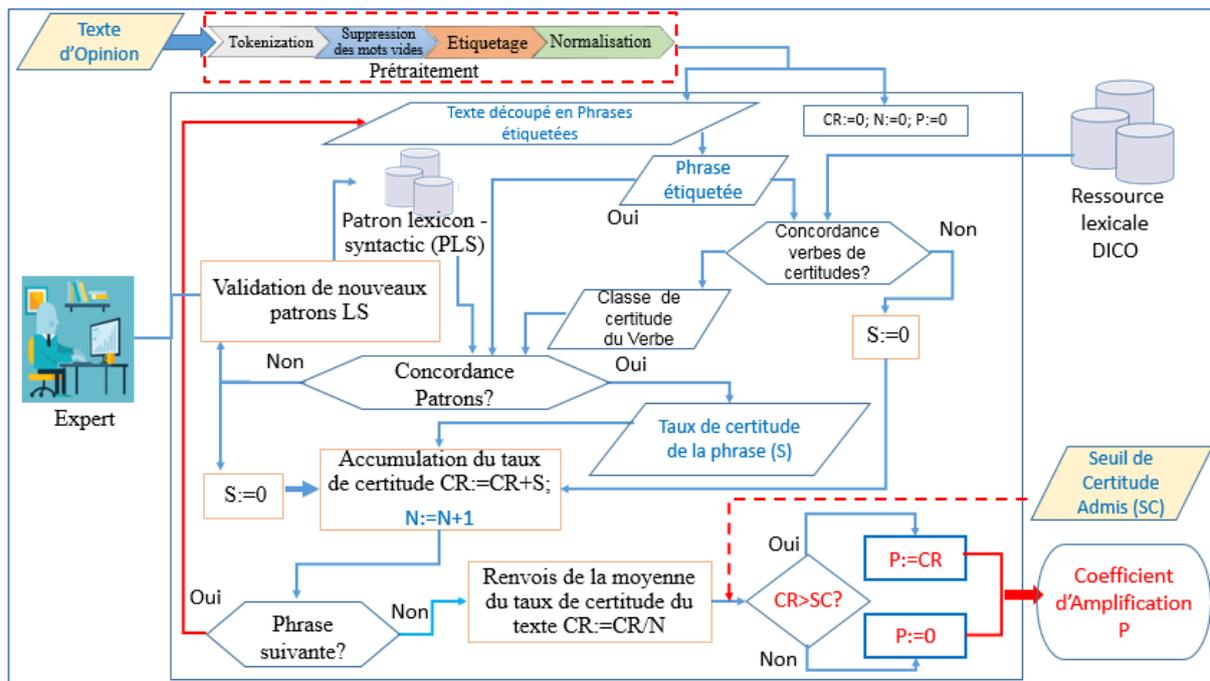


Figure 29 Diagramme de fonctionnement du module de mesure de certitude.

Rappelons que le prétraitement a consisté en : découpage en phrases, tokenisation, suppression de mots vides, étiquetage et normalisation. Le résultat fourni est un étiquetage de

chaque phrase du texte (nom, verbe, adverbe, adjectif). Le taux de certitude (CR) et le nombre de phrases du texte (N) sont initialisés à zéro. Pour chaque phrase, le processus recherche la présence d'un verbe d'attitude correspondant à ceux décrits dans la ressource lexicale DICO (Cf. fig. 29), s'il en trouve, il retourne sa classe de certitude (sûre, probable, possible). Le traitement se poursuit par la recherche d'une concordance avec un des patrons lexico syntaxiques implémentés. Si une concordance avec un patron est trouvée, la valeur du taux de certitude est retournée (Cf. tab. 8), sinon la valeur zéro est retournée. Dans ce dernier cas, la phrase sera sauvegardée pour une éventuelle validation d'un nouveau patron lexico syntaxique. Le processus continue ses itérations jusqu'à épuisement de toutes les phrases du texte. La moyenne du taux de certitude (CR) est ensuite calculée. Cette moyenne est comparée à un seuil de certitude admis (SC) défini préalablement qui permet de calculer le coefficient de l'amplification de l'opinion (P) (Cf. figure 29).

A titre d'exemple, le tableau 9 présente les résultats du traitement de réponses d'élèves obtenues durant les enquêtes d'opinion où on met l'accent sur l'amplification de l'opinion calculée avec le taux de certitude. A noter que le seuil de certitude admis (SC) est fixé à 0.5.

Tableau 9 Exemple d'amplification de l'opinion par la certitude.

Phrases		<i>Je veux améliorer mes connaissances en mathématique</i>	<i>Nous savons que la physique est intéressante.</i>	<i>Mes parents pensent que je suis bon en physique</i>
Paramètres de Certitude	Verbe	Vouloir	Savoir	Penser
	Classe	Probable	Sure	Possible
	Patron	[Je]	[Nous]	[Ils/Elle/Elles]
Taux de Certitude (CR)		0,9	0,96	0,66
Coefficient d'amplification (P)		0,9	0,96	0,66
Valeur d'Opinion sans Certitude		4,2	4.2	3,4
Opinion avec Certitude		5,1	5.16	3.4

On note à ce niveau, une différence remarquable entre l'opinion ordinaire et celle amplifiée par le coefficient de certitude. Cela octroie un moyen de distinguer les opinions selon le niveau de certitude qui peut renseigner sur la crédibilité de celle-ci. La prise en compte de cette amplification lors de la synthèse d'opinion par domaine d'étude va améliorer le score d'orientation, car cette amplification de l'opinion va permettre d'augmenter les opinions positives et de décroître les opinions négatives. Par conséquent, les valeurs d'opinions pour les domaines d'études dont les matières pré requises ont obtenues des valeurs positives avec une

forte certitude se verront augmentées. Alors que celles des domaines dépendants des matières ayant obtenues des valeurs d'opinion négatives avec une forte certitude verront leurs valeurs diminuées. De ce fait, le score d'orientation va favoriser d'avantage les domaines dont les matières ont obtenu des valeurs d'opinion positive avec une forte certitude.

III.8 Module de synthèse de l'orientation

Ce module du processus intervient au moment de l'orientation scolaire. Il est dédié à synthétiser les valeurs d'opinion calculées précédemment afin de les jumeler aux autres paramètres à savoir les notes et les aptitudes des élèves pour arriver à calculer le score d'orientation qui permet d'élaborer la carte d'aide à l'orientation des élèves. La figure 30 illustre l'architecture de ce module qui comporte trois sous modules détaillés ci-dessous.

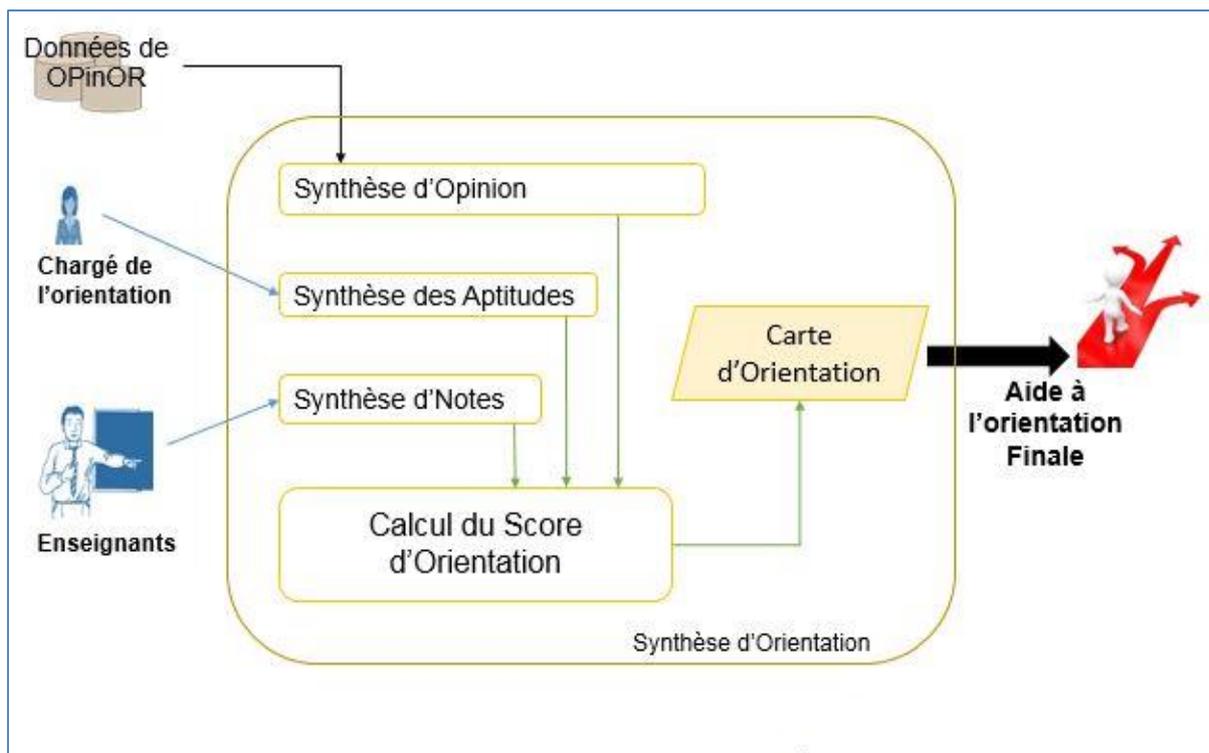


Figure 30 Architecture du module de synthèse d'orientation.

III.8.1 Module de synthèse d'opinion

Ce module reçoit en entrées les opinions sauvegardées durant les enquêtes d'opinion. Il a pour rôle de synthétiser, pour chaque intervenant, les opinions collectées afin de fournir des valeurs d'opinions de chaque intervenant, relatives à chaque domaine d'étude. Ces dernières seront aussi synthétisées afin de calculer l'opinion générale de tous les intervenants pour chaque domaine. Celle-ci servira au calcul du score d'orientation. Nous détaillons dans ce qui suit ces différentes étapes.

III.8.1.1 Moyenne périodique des opinions

La synthèse des opinions obtenues lors de toutes les enquêtes précédentes, commence par le calcul de la moyenne des opinions, relatives à chaque matière, pour toutes les périodes (nbT). Cette étape s'intéresse à identifier (prédire) la tendance de l'opinion des intervenants. Elle repose sur l'hypothèse que puisque l'opinion récente est plus représentative, alors il faut la favoriser lors du calcul de la moyenne des opinions. De ce fait, il est considéré que les opinions collectées lors des enquêtes constituent une série temporelle à laquelle est appliqué un lissage exponentielle souple (Dufour, 2002). Selon (Lagnoux, 1996), la moyenne des opinions sera alors donnée par la formule III.7, où b représente la constante de lissage fixé à 0.09 (Ostertagova & Ostertag, 2012).

$$\left\{ \begin{array}{l} \text{OMP} = (1 - b) \sum_{j=0}^{\text{nbT}-1} (b)^j \text{OGA}_{\text{nbT}-j} \\ \text{nbT} \text{ représente le nombre de périodes d'enquête.} \\ \text{b} \text{ représente la constante de lissage.} \end{array} \right. \quad (\text{III. 7})$$

Nous calculons ainsi :

- OMP_{EM} : L'opinion moyenne périodique de l'élève "E" au sujet de la matière "M" ;
- OMP_{SEM} : L'opinion moyenne périodique de l'enseignant "S" de la matière "M" au sujet de l'élève "E" ;
- OMP_{PEM} : L'opinion moyenne périodique du parent "P" de l'élève "E" au sujet de la matière "M".

III.8.1.2 Opinion par domaine :

L'opinion moyenne périodique de chaque intervenant sera synthétisée par domaine selon la formule III.9. En d'autres termes, les valeurs des opinions moyennes périodiques de chaque

matière serviront à calculer une valeur d'opinion pour chaque domaine à l'aide des coefficients de pondération y afférents (CF fig. 17), cela pour chacun des intervenants.

$$\left\{ \begin{array}{l} OD = \frac{(\sum_{m \in \text{dom}} \text{cof}_m * OMP_m)}{\sum_{m \in \text{dom}} \text{cof}_m} \\ \text{Cof}_m: \text{Coefficients de pondération de l'opinion des intervenants} \\ \text{par matière pour les domaines d'études } (f(\alpha_i), g(\alpha_i), h(\alpha_i)). \end{array} \right. \quad (\text{III. 8})$$

Nous calculons ainsi :

- OD_E : Opinion par domaine de l'élève ;
- OD_S : Opinion par domaine de l'enseignant ;
- OD_P : Opinion par domaine du parent.

III.8.1.3 Opinion générale :

La synthèse d'opinion aboutit enfin, au calcul d'une valeur d'opinion générale qui englobe l'opinion de l'élève, celle de ses parents ainsi que celle de ses enseignants, relative à chaque domaine selon la formule III.9.

$$\left\{ \begin{array}{l} OG_{EDOM} = \alpha * OD_E + \beta * OD_{SE} + \lambda * OD_{PE} \\ OD: \text{opinion par domaine} \\ (E: \text{élève}, SE: \text{enseignant}, PE: \text{parent}) \\ \alpha: \text{coefficients de pondération de } OD_E; \\ \beta: \text{coefficients de pondération de } OD_{SE}; \\ \lambda: \text{coefficients de pondération de } OD_{PE}. \end{array} \right. \quad (\text{III. 9})$$

Où :

- VOG_{EDOM} représente la valeur de l'opinion générale relative à l'élève « E » concernant le domaine « DOM ».

La figure 31 montre les différents niveaux de synthèse de l'opinion.

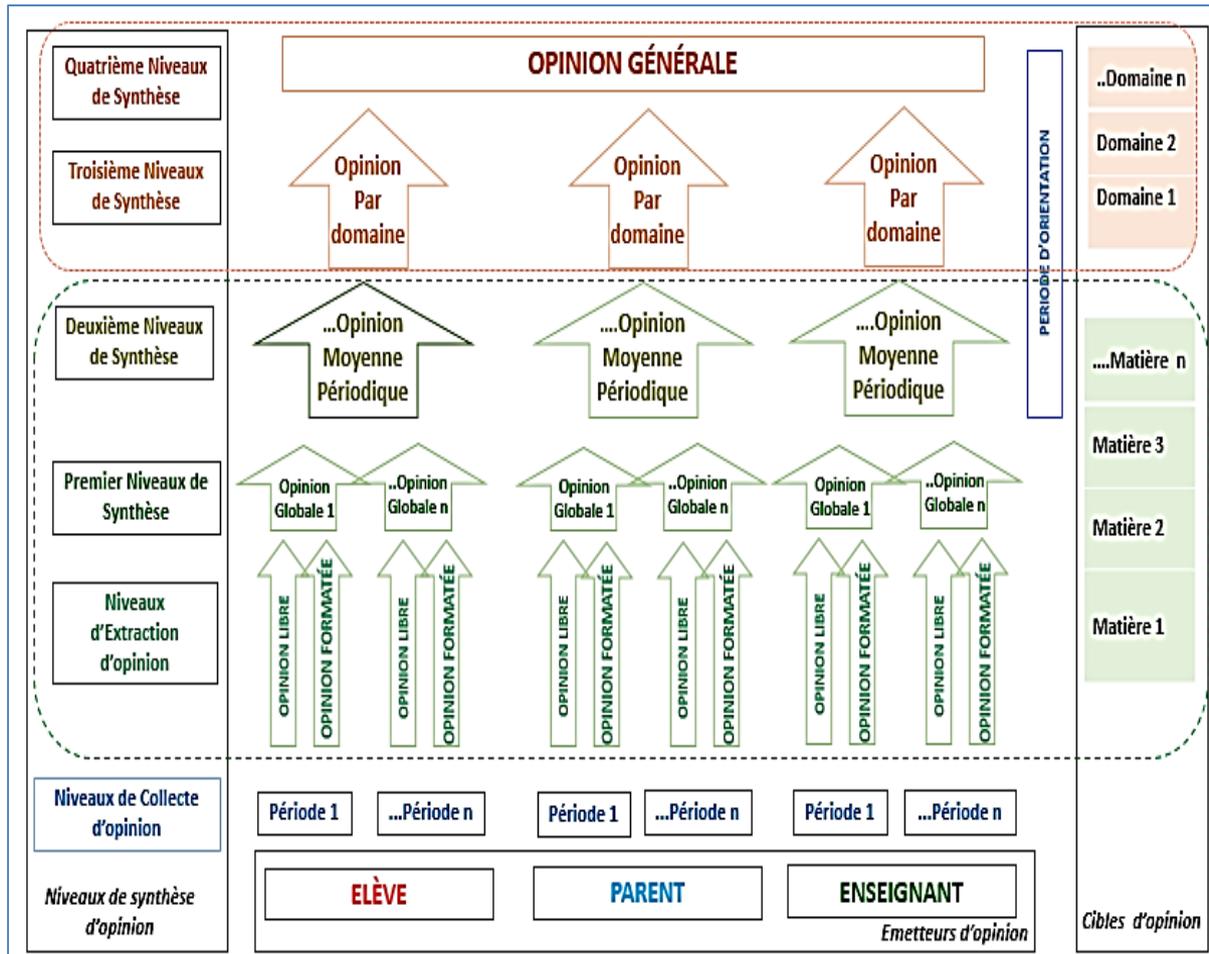


Figure 31 Niveaux de synthèse d'opinion.

III.8.2 Module de Calcul du Score d'Orientation

Ce module a pour rôle de calculer un score d'orientation pour chaque domaine d'étude. Comme montré dans la figure 30, il reçoit en entrée trois paramètres :

- **L'opinion générale** (VOG) obtenue selon la formule III.9.
- **La moyenne des notes** (VES) de chaque élève par domaine (AMV_{EDOM}) calculée à base de la moyenne des notes (AMV_{EM}) de l'élève (E) relative à chaque matière (M) comme donnée par la formule III.10.

$$AMV_{EDOM} = \frac{\sum_{M \in \text{dom}} AMV_{EM} * \alpha_M}{\sum_{M \in \text{dom}} \alpha_M} \quad (\text{III. 10})$$

Où :

- α_M représentent les coefficients de pondération des notes de chaque matière pour les domaines d'étude comme illustré dans la figure 17.

- **La valeur des aptitudes (VAP)** de chaque élève par domaine (DOM), est calculée par la pondération des aptitudes de l'élève selon la formule III.11.

$$VAP_{EDOM} = \frac{\sum_{ap \in dom} VAP_{eap} * \gamma_{ap}}{\sum_{ap \in dom} \gamma_{ap}} \quad (III.11)$$

Où :

- γ_{ap} représentent les coefficients de pondération des aptitudes (intelligence de Gardner) pour chaque domaine d'étude comme illustré par la figure 18.

Après cela, le score d'orientation pour chaque domaine d'étude (SOF_{EDOM}) est calculé selon la formule III.12.

$$\left\{ \begin{array}{l} SOF_{EDOM} = \sigma * AMV_{EDOM} + \gamma * VAP_{EDOM} + \omega * VOG_{EDOM} \\ \sigma: \text{Coefficient de pondération des moyennes des notes;} \\ \gamma: \text{Coefficient de pondération des aptitudes de l'élèves;} \\ \omega: \text{Coefficient de pondération des opinions générales.} \end{array} \right. \quad (III.12)$$

Après la synthèse du score d'orientation pour l'ensemble des élèves, la carte d'aide à l'orientation sera accessible aussi bien pour les élèves que pour les chargés de l'orientation. Elle les aidera à finaliser le choix du domaine d'études. La figure 32 illustre une carte détaillée d'aide à l'orientation d'un élève.

α	valeur de l'évaluation scolaire	β	valeur de l'aptitude	γ	score de l'orientation	recommandation
0.4	11.74	0.4	13.03	0.2	13.14	D4
0.4	10.23	0.4	12.32	0.2	13.97	D2
0.4	11.53	0.4	13.25	0.2	13.03	D6
0.4	10.21	0.4	15.24	0.2	14.71	D5
0.4	12.54	0.4	12.45	0.2	13.37	D1
0.4	13.21	0.4	12.65	0.2	13.96	D3
0.4	12.21	0.4	13.51	0.2	12.60	D7

Figure 32 Carte d'aide à l'orientation détaillée.

III.9 Structures de l'algorithme du processus proposé

Après avoir passé en revue les différents modules de OPinOR, nous présentons dans cette section son algorithme principal illustré dans la figure 33.

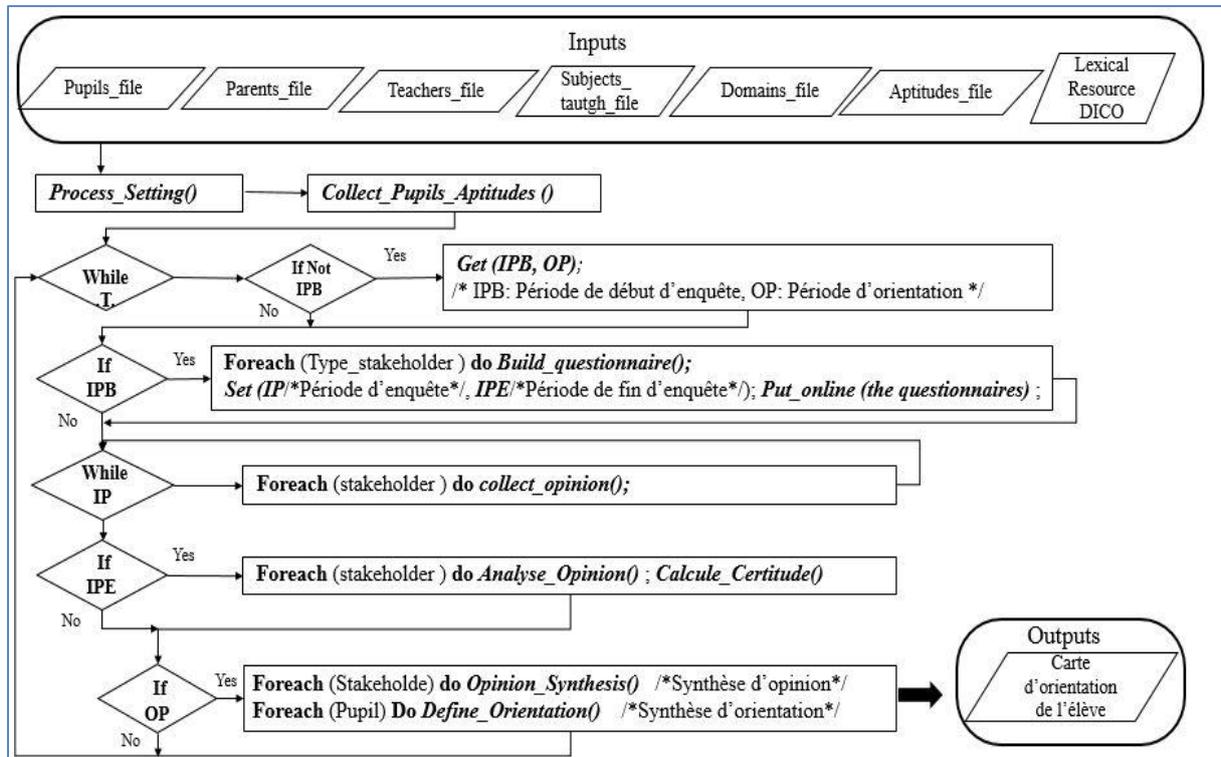


Figure 33 Organigramme principal d'OPinOR.

Conclusion

Dans ce chapitre, nous avons présenté nos propositions dans le cadre de l'amélioration de l'orientation scolaire. Pour ce faire, nous avons proposé un processus qui intègre de nouveaux paramètres tel que l'opinion des intervenants, les styles d'apprentissage de l'élève en plus de ses évaluations scolaires utilisées actuellement par les systèmes d'orientation. Ce processus s'articule sur cinq modules principaux : un module de paramétrage du système, un module dédiée à la construction d'une ressource lexicale exigée par l'analyse d'opinion, un module de traitement de l'opinion, un module de raffinement de l'opinion qui se traduit par la mesure de la certitude engagé par l'émetteur d'opinion et un module de calcul du score d'orientation qui synthétise tous les paramètres énoncés. Son objectif étant de préconiser à l'élève les domaines d'études favorables à ses attentes à travers la carte d'aide à l'orientation. Dans le prochain chapitre, nous allons présenter l'expérimentation et la validation de nos propositions.

Chapitre IV : Expérimentation et validation des propositions

Introduction

L'expérimentation de nos propositions a été menée au niveau local (Algérie). Elle s'est focalisée sur l'orientation des nouveaux bacheliers. La validation de nos propositions s'articule autour de trois parties. La première partie est consacrée à l'évaluation de l'extraction de l'opinion libre. La deuxième partie est dédiée à l'évaluation du système OPinOR sans la prise en compte de la mesure de certitude. La troisième partie est destinée à évaluer la proposition avec la prise en charge de la mesure de la certitude.

IV.2 Evaluation du module d'extraction de l'opinion libre

Pour évaluer l'extraction de l'opinion libre, nous recourons au calcul d'indicateurs de performance classiques à savoir : la justesse (accuracy), la précision, le rappel et le F1-score, sur un corpus annoté (Hossin & Sulaiman, 2015).

Le module d'extraction de l'opinion libre repose sur une approche lexicale qui utilise une ressource lexicale de langue française (DICO) destinée à l'analyse d'opinion dans le domaine de l'E-éducation. Pour rappel, cette ressource a été construite à partir de WOLF (WORDnet Libre Français) et SentiWordNet (Sentiment WordNet). Pour évaluer le module d'extraction de l'opinion libre, nous avons testé ce module sur un corpus de commentaires issus en partie de forums d'élèves et des réponses issues des enquêtes d'opinion menées. Ce corpus, regroupe 23768 commentaires annotés manuellement. Chaque commentaire détient une polarité (positive, négative ou neutre). Nous avons obtenu la matrice de confusion illustrée dans le tableau 10.

Tableau 10 Matrice de confusion Obtenue.

		Prédite		
		<i>Positive</i>	<i>Neutre</i>	<i>Négative</i>
Vraie	<i>Positive</i>	3100	2591	2344
	<i>Neutre</i>	1646	3789	1598
	<i>Négative</i>	2795	2720	3185

Les indicateurs de performance sont calculés selon les formules (I.6), (I.8), (I.9) et (I.10) respectivement (Hossin & Sulaiman, 2015). Les valeurs correspondantes à cette matrice sont :

- Justesse (accuracy) de 0.616 ;
- Précision de 0.425 ;
- Rappel de 0.43 ;
- F1-Score de 0.423.

L'implémentation actuelle de notre approche répond aux besoins initiaux de notre recherche, en effet, les indicateurs de performance obtenus sont proches de ceux donnés par les recherches similaires (Hemmatian & Sohrabi, 2017).

IV.3 Evaluation du système OPinOR sans prise en compte de la certitude

IV.3.1 Protocole d'évaluation

Dans le but de déterminer si le processus présenté est en mesure d'offrir aux élèves une orientation vers des domaines d'études qui répondent à leurs attentes, nous avons proposé de le tester sur un échantillon d'élèves en suivant les étapes suivantes :

- Conduite d'enquêtes d'opinion sur les matières dispensées, auprès d'un échantillon d'élève de terminal sur plusieurs périodes ;
- Recueil du choix explicite des domaines d'études souhaités par chaque élève durant la dernière enquête ;
- Calcul du score d'orientation pour les quatre cas :
 - Avec les notes obtenues à l'examen du Baccalauréat ;
 - Avec les valeurs des opinions recueillies durant les enquêtes ;
 - Avec les valeurs des aptitudes des élèves ;
 - En fin, avec tous les paramètres, notes, opinions et aptitudes.
- Comparaison pour chaque cas, du résultat donné par le score d'orientation avec le choix explicite de l'élève. L'objectif étant, de comparer l'ordre (hiérarchie) des domaines d'études suggéré par notre système à l'ordre fourni par l'élève dans son choix explicite. Nous examinons ensuite, le cas (parmi les quatre cités ci haut) qui offre une meilleure similarité entre les deux hiérarchies. Nous estimons cette similarité par le coefficient Kappa qui mesure le degré de concordance entre les

stades attribués par deux juges (Sim & Wright, 2005). L'interprétation du coefficient kappa est présentée dans le tableau 11 (Santos, 2015).

Tableau 11 Interprétation des valeurs du coefficient Kappa.

Kappa	Interprétation
<0	Grand désaccord
0-0,20	Accord très faible
0,21-0,40	Accord faible
0,41-0,60	Accord moyen (Modéré)
0,61-0,80	Accord satisfaisant (presque parfait)
0,81-1,00	Accord excellent (Parfait)

- Déduction à la fin, de la fiabilité du système proposé à partir des valeurs du coefficient Kappa obtenues pour chaque cas.

IV.3.2 Echantillon

Nous avons mené des enquêtes auprès d'un échantillon constitué de 300 élèves de classes de terminal de la filière sciences naturelles. Ces enquêtes sont menées aussi bien auprès des élèves, de leurs parents et de leurs enseignants durant les trois trimestres de l'année scolaire. Durant ces enquêtes :

- Les élèves ont exprimé leurs opinions au sujet des matières étudiées. Leurs choix explicites des domaines d'études universitaires selon leurs envies ont été recueillis durant l'enquête du dernier trimestre ;
- Les parents ont exprimé leurs opinions au sujet des matières étudiées par leurs enfants ;
- Les enseignants ont exprimé leurs opinions au sujet de leurs élèves.

Les domaines d'études universitaires auxquels on s'est intéressés sont donnés par la figure 34.

code	designation
D1	Sciences médicale
D2	Sciences et technologies
D3	Sciences de la matière
D4	Mathématique et informatique
D5	Sciences de la nature et vie
D6	Sciences de la terre et univers
D7	Sciences économique, management et commerciale

Figure 34 Domaines d'études universitaires considérés.

La liste des matières enseignées pour la filière des sciences naturelles, que nous avons considéré dans le cadre de cette étude, est illustrée dans la figure 35.

code	designation
M1	Mathématiques
M2	Physique
M3	sciences naturelles
M4	Anglais
M5	Arabe
M6	Français
M7	Histoire et géographie

Figure 35 Liste des matières enseignées considérées dans cette étude.

L'orientation des nouveaux bacheliers en Algérie se base sur la moyenne générale obtenue à l'examen du baccalauréat, combinée à celles de certaines matières selon l'importance de ces dernières pour les domaines d'études universitaires considérés. Cette importance se traduit par des coefficients de pondération (α_i) qui lient les matières enseignées aux domaines d'études

universitaires. La figure 36 illustre ces liaisons. Notons que, pour les besoins spécifiques de ce processus, nous avons ajouté une « matière fictive » représentant la moyenne générale obtenue au baccalauréat que nous avons codifié en m_0 à laquelle nous avons associé le coefficient de pondération α_0 dont la valeur est égale à 1.

α_i	Science medicale	Science et technologie	Science de la matière	Mathématique et informatique	Scince de la nature et de la vie	Science de la terre et de l'univers	Science économiques,management et commerciales
Note a l'examen du bac	1	1	1	1	1	1	1
Mathématique	4	6	3	6	4	4	1
Physique	4	5	5	5	4	4	2
sciences naturelles	6	3	6	3	5	5	2
Anglais	2	2	2	2	2	2	3
Français	1	1	1	1	1	1	3
Arabe	2	2	2	2	2	2	3
Histoire et géographie	3	1	1	1	1	1	3

Figure 36 Pondération des notes de matières (α_i) par domaines d'études.

Les fonctions affines qui calculent les coefficients de pondération des opinions relatives aux différents intervenants sont fixées comme suit :

- L'opinion des élèves avec la fonction $f(\alpha)=\alpha+3$;
- L'opinion des enseignants avec la fonction $h(\alpha)=\alpha+2$;
- L'opinion de leurs parents avec la fonction $g(\alpha)=\alpha+1$.

Les paramètres a et b de chaque fonction sont fixés selon l'importance de l'opinion de l'intervenant concerné, dans le processus d'orientation (Cf. Section III.4). De la même manière, nous avons fixé les coefficients de l'opinion générale comme suit :

- $\delta=0.5$ pour l'opinion globale des élèves ;
- $\beta=0.3$ pour l'opinion globale des enseignants ;
- $\lambda=0.2$ pour l'opinion globale des parents.

Comme mentionné précédemment (Cf. Section III.4), nous considérons que l'opinion des élèves est la plus importante, suivie par celle de leurs enseignants puis enfin par celle de leurs parents.

Les aptitudes des élèves représentées par les intelligences multiples de Gardner (Gardner, 2011) sont codifiées comme illustrées dans la figure 37. Les coefficients de pondération γ_i des aptitudes (intelligences multiples) de l'élève pour chaque domaine d'étude sont inspirés de l'étude menée par l'auteur de (Garceau, 2010).

Figure 37 Codification et Pondération des Aptitudes de l'élève.

Après avoir renseigné tous les paramètres nécessaires au lancement du processus OPinOR, nous présentons dans la prochaine section les résultats obtenus.

IV.3.3 Résultats

Les taux de réponse sont à hauteur de 60% pour les parents, 80% pour les enseignants et 90% pour les élèves. Nous avons constaté à travers les réponses aux questionnaires, que les élèves répondent, principalement, aux questions fermées. De ce fait ; lors du calcul de l'opinion globale des élèves par matière, une plus grande importance est accordée à l'opinion formatée. Les coefficients de pondération de l'opinion sont, alors, initialisés respectivement, à ($\theta=0,9$) pour l'opinion formatée et ($\varphi=0,1$) pour l'opinion libre. Les parents et les enseignants ont répondu aux deux types de questions, c'est pourquoi nous avons initialisé les coefficients d'opinion ($\theta=0,8$, $\varphi=0,2$) et ($\theta=0,7$, $\varphi=0,3$) respectivement.

Les valeurs obtenues pour les opinions sont ramenées à la même échelle que les évaluations scolaires.

Nous avons calculé le score d'orientation pour l'ensemble des élèves, en tenant compte uniquement des moyennes des notes ($\omega=0$, $\sigma=1$ et $\gamma=0$). Nous présentons la carte d'aide à l'orientation synthétisée d'un élève (Cf. Fig. 38) qui illustre l'ordre des domaines d'études obtenu suivant son choix explicite et l'ordre de ces mêmes domaines obtenu, selon la moyenne des notes obtenues lors de l'examen du baccalauréat.

Domaines d'études universitaires considérés		ordre par choix explicite	score de l'orientation $\alpha=0 \beta=1 \gamma=0$	recommandation
code	designation			
D1	Mathématiques et Informatique	1er	15.61	4eme
D2	Sciences de la Terre et de l'Univers	6eme	14.95	5eme
D3	Sciences Médicale	5eme	15.75	1er
D4	Science et Technologie	2eme	14.21	6eme
D5	Science de la matière	3eme	15.63	3eme
D6	Science de la nature et de la vie	4eme	15.65	2eme
D7	Sciences économique, management et commerciale	7eme	12.63	7eme

Figure 38 Carte d'aide à l'orientation synthétisée basée sur les notes du BAC.

Nous constatons à ce niveau que l'ordre des domaines d'études reposant sur le choix explicite de cet élève n'est pas similaire à celui qui est donné par la moyenne de ses notes du BAC.

La figure 39 illustre la carte d'aide à l'orientation du même élève, en tenant compte des trois critères : moyenne des notes du BAC, ses valeurs d'opinion et ses aptitudes. Nous remarquons que l'ordre des domaines recommandés par le score d'orientation est proche de celui de son choix explicite.

Domaines d'études universitaires considérés		ordre par choix explicite	score de l'orientation $\alpha=0,4 \beta=0,4 \gamma=0,2$	recommandation
code	designation			
D1	Mathématiques et Informatique	1er	14.95	1er
D2	Sciences de la Terre et de l'Univers	6eme	13.95	6eme
D3	Sciences Médicale	5eme	14.75	3eme
D4	Science et Technologie	2eme	14.83	2eme
D5	Science de la matière	3eme	14.64	5eme
D6	Science de la nature et de la vie	4eme	14.72	4eme
D7	Sciences économique, management et commerciale	7eme	12.63	7eme

Figure 39 Carte d'aide à l'orientation synthétisée avec les trois critères.

Pour généraliser ce constat, nous avons comparé ces deux paramètres pour l'échantillon des 300 élèves. Nous avons calculé le coefficient Kappa pour estimer la similarité entre l'ordre des domaines donné par le choix explicite à celui donné par le score d'orientation. Nous présentons ci-dessous les résultats obtenus pour chaque cas :

IV.3.3.1 Cas de la moyenne des notes seules

Nous avons calculé le score d'orientation pour l'ensemble des élèves de l'échantillon en tenant compte, uniquement, des notes du Bac ($\omega = 0$, $\sigma = 1$, et $\gamma = 0$). Ce cas correspond à une orientation ordinaire. Nous avons examiné la similarité des résultats donnés par ce score à ceux des choix explicites des élèves à travers le calcul du coefficient Kappa. Le tableau 12 présente la matrice de cette similitude des ordres par score d'orientation et choix explicite.

Tableau 12 Matrice de similitudes - Cas des moyennes des notes.

		Ordre donné par le Score d'Orientation						
		D1	D2	D3	D4	D5	D6	D7
Ordre donné par le Choix Explicite	D1	30		8		7		3
	D2		24		9		11	
	D3	14		17		21		19
	D4		15		15			
	D5	19		6	9	9		11
	D6	5	9				17	9
	D7	1		3	5	7	2	16

Dans cette matrice, on peut lire par exemple qu'il y a trente (30) élèves pour lesquels les deux méthodes placent le domaine (D1) dans la même position. Aussi, on retrouve quatorze (14) élèves pour qui le score d'orientation place (D1) dans la même position qui est donnée par le choix explicite à (D3). La valeur du coefficient Kappa obtenue pour ce cas est égale à **0.2945**.

IV.3.3.2 Cas de la moyenne des opinions seules

Nous avons calculé le score d'orientation pour l'ensemble des élèves, en tenant compte cette fois ci uniquement de l'opinion des intervenants ($\omega=1$, $\sigma=0$ et $\gamma=0$). Nous avons, également, comparé les résultats obtenus à travers ce score et ceux obtenus à travers les choix explicites des élèves en vue de calculer le coefficient kappa. Le tableau 13 illustre la similitude des choix exprimés.

Tableau 13. Matrice de similitudes - Cas des opinions.

		Ordre donné par le Score d'Orientation						
		<i>D1</i>	<i>D2</i>	<i>D3</i>	<i>D4</i>	<i>D5</i>	<i>D6</i>	<i>D7</i>
Ordre donné par le Choix Explicite	<i>D1</i>	31		6	13		7	1
	<i>D2</i>		22		3	4		
	<i>D3</i>	4	10	39		2	7	
	<i>D4</i>	8		2	27			0
	<i>D5</i>			2		32		6
	<i>D6</i>	5	9				17	9
	<i>D7</i>	1		3	5	7	2	16

De la même manière, on peut lire dans cette matrice, à titre d'exemple, qu'il y a trente et un (31) élèves pour lesquels les deux méthodes placent le domaine (D1) dans la même position. Aussi, il y a huit (08) élèves pour qui le score d'orientation place (D1) dans la même position qui est donné par le choix explicite à (D4). La valeur du coefficient Kappa obtenue pour ce cas est égale à **0,5466**.

IV.3.3.3 Cas des aptitudes seules

Nous avons calculé le score d'orientations selon les aptitudes des élèves ($\omega=0$, $\sigma=0$ et $\gamma=1$) et de la même manière, nous avons comparés ces résultats aux choix explicites des élèves, nous avons obtenu les résultats figurants dans le tableau 14.

Tableau 14 Matrice de similitudes - Cas des aptitudes.

		Ordre donné par le Score d'Orientation						
		<i>D1</i>	<i>D2</i>	<i>D3</i>	<i>D4</i>	<i>D5</i>	<i>D6</i>	<i>D7</i>
Ordre donné par le Choix Explicite	<i>D1</i>	29		6	13		9	7
	<i>D2</i>		19		3	4		
	<i>D3</i>	4	10	26		2	7	
	<i>D4</i>	10		2	17			9
	<i>D5</i>			2		28		6
	<i>D6</i>	5	9				23	9
	<i>D7</i>	6		3	5	7	2	18

Aussi, on peut voir dans cette matrice, à titre d'exemple, qu'il y a dix-neuf (19) élèves pour qui les deux méthodes placent le domaine (D2) dans la même position. On voit aussi qu'il y a dix (10) élèves pour qui le score d'orientation place (D1) dans la même position que celle donnée par le choix explicite à (D4). La valeur du coefficient Kappa obtenue pour ce cas est égale à **0,4533**.

IV.3.3.4 Cas des trois paramètres combinés

En fin, nous avons calculé le score d'orientations selon les trois facteurs à savoir les moyennes des notes, les valeurs d'opinion et les valeurs des aptitudes des élèves. Nous considérons le cas où les notes et les opinions ont le même poids dans le calcul du score d'orientation. Nous fixons alors les coefficients de ce score à : $\omega=0.4$, $\sigma=0.4$ et $\gamma=0.2$.

Cependant, il est souhaitable que le choix des valeurs pour ces coefficients se fasse par les experts du domaine de l'éducation.

Aussi de la même manière, nous avons comparés ces résultats aux choix explicites des élèves, nous avons obtenus les résultats figurants dans le tableau 15.

Tableau 15 Matrice de similitudes - Cas des trois paramètres.

		Ordre donné par le Score d'Orientation						
		<i>D1</i>	<i>D2</i>	<i>D3</i>	<i>D4</i>	<i>D5</i>	<i>D6</i>	<i>D7</i>
Ordre donné par le Choix Explicite	<i>D1</i>	48		2	4	3	4	1
	<i>D2</i>		38		7		5	3
	<i>D3</i>	2		34		2		
	<i>D4</i>		8		29		2	2
	<i>D5</i>	5				25		
	<i>D6</i>	7	4	3	3		30	
	<i>D7</i>		3		2			24

Nous remarquons dans la diagonale de la matrice que les chiffres représentant la similitude des positions, entre le choix explicite et le score d'orientation, ont sensiblement

augmenté par rapport à ceux obtenus dans les cas précédents. Nous avons par exemple, quarante-huit (48) élèves pour qui le domaine (D1) est donné dans la même position aussi bien par le score d'orientation que par le choix explicite des élèves. La valeur du coefficient Kappa obtenue pour ce cas est égale à **0.7170**.

IV.3.4 Interprétation

Les valeurs que nous avons obtenues dans le cadre de cette étude sont résumées dans le tableau 16.

Tableau 16 Synthèse des valeurs du coefficient Kappa obtenues.

Score avec	Valeurs du coefficient Kappa obtenues	Interprétation
Les Notes du baccalauréat	0,2945	Accord faible
Opinion	0,5466	Accord moyen
Aptitudes	0,4533	Accord moyen
Les trois facteurs (Notes, Opinions, Aptitudes) combinées	0,7170	Accord satisfaisant

Ces résultats montrent que la similitude entre les choix explicites des élèves des domaines d'études universitaires et ceux donnés par les moyennes de leurs notes est en faible accord ce qui veut dire que si on ne tient compte que des moyennes des notes obtenues par les élèves, durant leurs orientations, ils se verront guidés vers des domaines d'études qui ne correspondent pas forcément à leurs choix.

Par ailleurs, lorsque nous introduisons des facteurs tels que l'opinion des intervenants dans ce processus et les aptitudes des élèves, l'orientation est nettement améliorée et les élèves se verront orientés vers les domaines qui concordent à leurs choix et à leurs attentes.

IV.4 Evaluation de la mesure de certitude et son impact sur l'orientation scolaire

IV.4.1 Protocole d'évaluation

L'objectif de la mesure de certitude de l'émetteur d'opinion est de définir le degré de crédibilité de cette opinion et de sa teneur en tant que paramètre de confiance lors de la prise de décision. L'insuffisance majeure que présente l'opinion en tant que paramètre de prise de

décision c'est sa versatilité. Elle peut en effet, être instable dans le temps et/ou être exprimée avec incertitude. Par conséquent, une opinion stable est considérée comme étant plus fiable. Nous nous basons sur cette hypothèse afin d'améliorer le paramètre opinion dans notre proposition.

Ainsi, pour évaluer la mesure de certitude que nous avons proposé, nous nous sommes intéressé à observer la corrélation entre la stabilité et la certitude exprimée dans les opinions d'un échantillon d'élèves. Ces opinions sont celles recueillies en partie lors des enquêtes menées dans le cadre de l'évaluation globale du système sans prise en compte de la mesure de certitude. Pour définir la stabilité de l'opinion, nous calculons d'abord sa variation (OV) correspondant à l'écart-type de l'ensemble des valeurs d'opinion considérées. Cette variation est calculée selon la formule VI.1.

$$\left\{ \begin{array}{l} OV = \text{Ecart - Type } (VO_1, VO_2, \dots, VO_{nbT}) \\ VO_i: \text{ Valeur de l'opinion obtenue pour l'enquête } i; \\ nbT: \text{ Nombre total d'enquêtes.} \end{array} \right. \quad (VI.1)$$

Etant donné que la variation de l'opinion est inversement proportionnelle à sa régularité nous proposons de considérer la stabilité(OS) de celle-ci, à titre illustratif, comme l'inverse de la variation. Ainsi, la stabilité de l'opinion va se calculer selon la formule IV.2.

$$OS = \frac{1}{OV} \quad (VI.2)$$

La corrélation de la stabilité et de la certitude des opinions est définie par le coefficient de corrélations (CC(CM,OS)) entre la stabilité des valeurs de l'opinion (OS) de l'échantillon considéré durant les périodes d'enquêtes (nbT) et la moyenne de la certitude(CM) exprimée. Ainsi, la corrélation est donnée par la formule IV.3.

$$CC(CM, OS) = \frac{\sum_{i=1}^{nbE} (CM_i + MCM)(OS_i + MOS)}{(nbE - 1)EC_{CM}EC_{OS}} \quad (IV.3)$$

Où :

- nbE : Taille de l'échantillon ;
- CM_i : Moyenne du coefficient d'amplification de certitude de toutes les périodes (nbT) ;
- MCM : Moyenne des (CM_i) pour l'échantillon considéré ;
- OS_i : Stabilité des opinions de l'élève i durant toutes les périodes (nbT) ;
- MOS : Moyenne des (OS_i) pour l'échantillon considéré ;
- EC_{CM} : Ecart-type(CM_i) ;
- EC_{OS} : Ecart-type(OS_i) ;

Afin d'évaluer la prise en compte du facteur certitude de l'opinion, nous calculons le score d'orientation avec amplification de l'opinion par le taux de certitude pour l'échantillon considéré. Nous comparons ensuite, les scores obtenus aux choix explicites des élèves comme détaillé plus haut. Nous considérons deux cas :

- Prise en compte des trois paramètres : Notes, aptitudes et opinion sans amplification de certitude ;
- Prise en compte des trois paramètres : Notes, aptitudes et opinion avec amplification de certitude ;

Nous calculons le coefficient Kappa pour chaque cas et nous comparons les résultats ainsi obtenus afin d'en déduire l'efficacité de la prise en compte de la mesure de certitude dans ce processus.

IV.4.2 Evaluation

Pour évaluer la mesure de certitude, nous avons utilisé les réponses aux questions libres d'un échantillon d'élèves de la branche science naturelle. Nous avons recueillis ces réponses pendant neuf (9) enquêtes. Certains de ces élèves font parties de l'échantillon étudié pour l'évaluation globale de OPinOR. Nous avons observé les opinions de cet échantillon d'élèves exprimées sur la matière mathématique. Nous avons calculé la stabilité de ces opinions. Aussi nous avons mesuré le taux de certitude contenu dans le texte des réponses libres avec un seuil de certitude (SC) arrêté à (0.5). Nous présentons, dans un premier temps, les résultats obtenus pour un élève. En effet, le tableau 17 montre les valeurs des opinions de cet élève, les taux de certitude correspondants à chaque texte des réponses libres, les valeurs du coefficient d'amplification (P) et les valeurs de l'opinion amplifiée (Cf. formule (III.5)).

Tableau 17 Les résultats obtenus pour un élève.

Période	Valeur d'Opinion	Taux de certitude	Coefficient d'amplification	Opinion Amplifiée	Observation
P1	4,25	0,98	0,98	5,23	
P2	3,8	0	0	3,8	Pas de certitude
P3	4,36	0,75	0,75	5,11	
P4	4,26	0,98	0,98	5,24	
P5	3,2	0,75	0,75	2,45	Opinion négatives
P6	3,89	0,36	0	3,89	Certitude inférieure au seuil
P7	4,72	0,8	0,8	5,52	
P8	4,86	0,89	0,89	5,75	
P9	4,95	0,9	0,9	5,85	

Nous avons calculé la stabilité des opinions avec et sans amplification avec le coefficient de certitude, Le tableau 18 illustre les résultats obtenus pour l'élève présenté précédemment. A ce niveau, nous remarquons que la stabilité de l'opinion est proportionnelle au degré de sa certitude.

Tableau 18 Valeur de la stabilité de l'opinion de l'élève.

Opinion		Moyenne du coefficient d'amplification par la Certitude (CM)
Variation (OV)	Stabilité (OS)	
0.56	1.77	0.67

De la même manière, nous avons calculé la stabilité de l'opinion pour tous les élèves considérés dans l'échantillon. Le tableau 19 illustre les résultats obtenus pour vingt (20) des cents (100) élèves.

Tableau 19 Stabilité et moyenne de certitude de l'opinion des élèves .

Elèves	Variation Opinion (OV)	Stabilité Opinion (OS)	Moyenne Certitude (CM)	Elèves	Variation Opinion (OV)	Stabilité Opinion (OS)	Moyenne Certitude (CM)
1	0,29	3,44	0,89	11	0,16	6,31	0,93
2	0,26	3,80	0,81	12	0,22	4,54	0,88
3	0,26	3,78	0,79	13	0,27	3,72	0,88
4	0,30	3,38	0,83	14	0,28	3,54	0,87
5	0,26	3,89	0,89	15	0,27	3,77	0,87
6	0,24	4,11	0,83	16	0,15	6,85	0,95
7	0,28	3,62	0,86	17	0,17	5,80	0,91
8	0,25	4,01	0,87	18	0,11	8,93	0,92
9	0,16	6,24	0,92	19	0,13	7,94	0,96
10	0,14	7,24	0,97	20	0,20	4,95	0,86

Afin d'illustrer la corrélation entre la stabilité de l'opinion et sa certitude, nous avons calculé le coefficient de corrélation (Cf. formule IV.3) et la valeur obtenue est de :

$$CC (CM_i, O_{Si}) = \mathbf{0,77}, i \text{ allant de } 1 \text{ à } 20 ;$$

Cette valeur correspond à une corrélation positive. Nous constatons, à ce niveau la corrélation entre la stabilité de l'opinion et la certitude contenu dans le texte qui l'exprime. Le graphe de la figure 40 confirme cette observation.

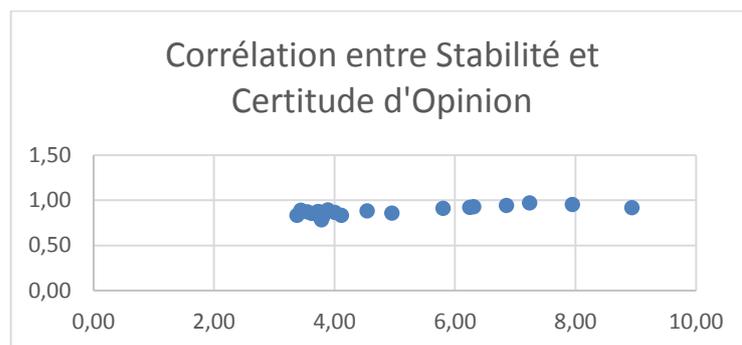


Figure 40 Stabilité et certitude de l'opinion.

Nous constatons ainsi une corrélation entre la stabilité et la certitude de l'opinion, ceci nous amène à dire que la prise en compte de ce facteur peut renforcer la crédibilité de l'opinion en tant que paramètre de prise de décision.

Nous avons aussi, examiné l'impact de la prise en compte de la certitude dans le processus d'orientation en reprenant les données relatives à l'échantillon des cent (100) élèves considéré dans cette section. Il y a lieu de signaler que parmi les 100 élèves considérés, il n'y a que quarante-huit (48) élèves pour qui nous disposons, aussi bien, des réponses de leurs parents que celles de leurs enseignants. De ce fait, nous avons restreint notre champ d'observation à ce sous échantillon. Le tableau 20 représente la couverture en certitude du sous échantillon considéré, en d'autres termes, le nombre d'élèves qui ont exprimé de la certitude dans leurs réponses pour chaque période et pour chacune des matières. Il y a lieu de noter que la certitude a été calculée uniquement pour les opinions des élèves.

Tableau 20 Couverture en certitude.

		Matières Objets des Enquêtes							Taux de Couverture en Certitude
		M1	M2	M3	M4	M5	M6	M7	
Périodes d'Enquête Considérées	P1	43	36	37	19	2	10	4	45
	P2	41	36	38	23	4	14	6	48
	P3	42	36	39	25	1	14	2	48
	P4	36	37	34	16	2	12	7	43
	P5	34	39	39	14	2	16	4	44
	P6	39	39	41	12	4	12	5	45
	P7	38	41	40	9	0	18	2	44
	P8	41	38	41	20	2	15	5	48
	P9	40	40	43	12	5	12	7	47

Ainsi, le système a calculé le taux de certitude pour les textes des réponses libres des élèves relatifs à chaque matière. Ensuite, il a pondéré, avec les coefficients d'amplification obtenus, les opinions globales correspondantes à chacune des matières. Il a fourni les scores d'orientation pour l'ensemble des élèves, en tenant compte :

- Des trois paramètres (Notes, opinion et aptitude selon les coefficients : $\omega=0.4$, $\sigma=0.4$ et $\gamma=0.2$.) sans amplification des opinions par la certitude. Notons que pour ce faire, on fixe le seuil de certitude (SC) à 1 (Cf. Fig. 29) ;
- Des trois paramètres (Notes, opinion et aptitude selon les coefficients : $\omega=0.4$, $\sigma=0.4$ et $\gamma=0.2$.) avec amplification des opinions par la certitude. Le seuil de certitude étant été fixé à 0.5.

Le tableau 21 illustre les matrices de similitude ainsi obtenues. Les coefficients kappa relatif à chacun des cas sont **0,6029** et **0,7777**.

Tableau 21 Matrices de similitude avec et sans amplification par la certitude.

		Ordre donné par le Score d'Orientation - Sans amplification par la certitude							Ordre donné par le Score d'Orientation - Avec amplification par la certitude						
		<i>D1</i>	<i>D2</i>	<i>D3</i>	<i>D4</i>	<i>D5</i>	<i>D6</i>	<i>D7</i>	<i>D1</i>	<i>D2</i>	<i>D3</i>	<i>D4</i>	<i>D5</i>	<i>D6</i>	<i>D7</i>
		<i>D1</i>	<i>D2</i>	<i>D3</i>	<i>D4</i>	<i>D5</i>	<i>D6</i>	<i>D7</i>	<i>D1</i>	<i>D2</i>	<i>D3</i>	<i>D4</i>	<i>D5</i>	<i>D6</i>	<i>D7</i>
Ordre donné par le Choix Explicite	<i>D1</i>	8		2	1		1	1	9		1	1		1	
	<i>D2</i>		5				2			5				1	
	<i>D3</i>	1		3		1					4		1		
	<i>D4</i>		1		7			2		1		8			1
	<i>D5</i>					4							5		
	<i>D6</i>	1	1		1		2					1		3	
	<i>D7</i>		1					3		1					5

Nous résumons les résultats obtenus ainsi que leurs interprétations dans le tableau 22.

Tableau 22 Interprétation des valeurs du coefficient Kappa obtenues.

Score	Valeurs du coefficient Kappa obtenues	Interprétation
Sans amplification par la certitude	0,6029	Accord satisfaisant
Avec amplification par la certitude	0,7777	Accord satisfaisant

Nous remarquons à ce niveau une augmentation du coefficient kappa lors de l'amplification de l'opinion par la certitude. Ceci explique que l'introduction du facteur certitude de l'opinion rapproche les résultats de l'orientation préconisés par le système aux choix explicites des élèves, d'où l'importance de la prise en compte de la certitude. Elle renforce la crédibilité de l'opinion d'une part et elle rapproche l'orientation automatique du choix explicite de l'élève.

Conclusion

Nous avons présenté dans ce chapitre les expérimentations que nous avons menées afin de valider nos propositions. Nous avons commencé par évaluer le module d'extraction de l'opinion libre avec un corpus de commentaires annotés issus du domaine de l'éducation. Nous avons, ensuite, expérimenté le système OPinOR sur un échantillon de 300 élèves en menant des enquêtes auprès d'eux, de leurs parents et de leurs enseignants. Nous avons traité les opinions

sans prendre en compte la mesure de certitude. Nous avons calculé le score d'orientation avec seulement les notes du Baccalauréat, ensuite avec les opinions et les aptitudes, nous avons mesuré la similarité du score donné par OPinOr et le choix explicite des élèves. Enfin, nous avons présenté la validation du module de mesure de certitude. Nous avons utilisé les données collectés précédemment pour observer la corrélation entre la stabilité de l'opinion ainsi que sa certitude d'une part, et l'impact de la prise en compte de la certitude sur le score d'orientation.

Dans le chapitre suivant nous synthétisons et discutons les résultats obtenus. Nous présentons aussi, les impacts de notre étude.

Chapitre V : Synthèse et discussion de nos propositions

Introduction

Nous nous intéressons dans ce chapitre à discuter et à synthétiser les résultats des expérimentations conduites dans le cadre de la validation de nos propositions. Nous mettons l'accent sur les points forts et les limites de nos propositions. Nous présentons enfin l'impact de ces recherches sur le domaine socioéconomique, d'une part, et celui de la recherche en E-éducation, d'autre part.

V.2. Synthèse et discussion

Dans cette thèse, nous avons proposé un système d'aide à l'orientation scolaire basé sur trois dimensions (la dimension opinion, la dimension personnelle, la dimension capitalisation). Nous avons concentré nos travaux sur la dimension opinion. Nous avons conçu une ressource lexicale pour permettre l'extraction de l'opinion à partir de texte, nous avons proposé une fiabilisation de cette opinion par la mesure de certitude de son émetteur, nous avons enfin intégré cette opinion dans le processus d'orientation scolaire.

Dans le but de valider nos propositions, nous avons mené des enquêtes auprès des intervenants dans le processus d'orientation. Nous avons exploité les données recueillies pour évaluer l'approche proposée pour l'extraction de l'opinion libre, la synthèse d'orientation et enfin la mesure de certitude.

Les indicateurs de performance enregistrés pour l'approche lexicale proposée (Justesse (accuracy) de 0.616, Précision de 0.425, rappel de 0.43 et F1-Score de 0.423) sont acceptables comparés à ceux des approches similaires qui donne une justesse plutôt faible comparé à des approches statistiques. Mais les approches lexicales ont l'avantage de ne pas exiger des corpus annotés, et offre des accès faciles aux mots et à leurs orientations (polarités).

De ce fait nous considérons que notre approche constitue une base pour une plateforme d'analyse d'opinion pour le domaine de l'éducation. En effet, notre ressource lexicale peut servir à d'autres applications de l'analyse d'opinion dans le domaine de l'éducation. Elle peut offrir d'autres pistes de recherche locale portant notamment sur l'intégration du multilinguisme.

Les attributs de l'opinion les plus exploités sont la polarité, la cible avec parfois ses caractéristiques, son émetteur et le temps où elle est émise. Dans notre étude, nous nous sommes intéressés à un autre attribut qui est l'engagement de son émetteur à travers le taux de certitude exprimé dans les textes d'opinion. Ce concept est intéressant dans le sens où il permet de distinguer les opinions bien fondées et d'en renforcer ainsi leur crédibilité.

Les applications de l'analyse d'opinion dans le domaine de l'éducation se focalisent généralement sur la détection de l'opinion des élèves dans le but d'améliorer la qualité de l'enseignement en ligne ou en présentiel (Oramas Bustillos, et al., 2019; Balahadia, et al., 2016; Nitin, et al., 2015). Elles s'intéressent aussi à la capture de l'opinion publique sur les services de l'éducation (Invento, et al., 2017; Hamzah & Widyastuti, 2016; Khomonenko, et al., 2017; Omar, et al., 2017). Notre étude s'est focalisée sur l'application de l'analyse d'opinion à l'orientation scolaire. Les recherches citées précédemment, recourent à différentes approches de l'analyse d'opinion nommément : les approches symboliques, les approches statistiques ou encore les approches hybrides. Notre proposition repose, comme déjà dit, sur la construction d'une ressource lexicale dédiée au domaine de l'éducation basé sur SentiWordNet et Wolf.

Les logiciels d'aide à l'orientation (LAO) sont généralement axés sur des informations relatives aux domaines d'études supérieurs et aux métiers offerts par les secteurs de la formation professionnelle. Aussi ils se concentrent sur les informations personnelles des élèves telles la découverte de l'image de soi. Le système que nous proposons est basé sur trois facteurs : les notes des évaluations continues obtenues par les élèves, un test psychologique qui détermine les intelligences multiples des élèves selon l'approche de Gardner (Gardner, 2011). Ajouté à cela, notre système incorpore un nouvel élément qui est l'opinion des intervenants comme paramètre de prise de décision. Nous rappelons précisément que pour l'analyse d'opinion dédiée à l'orientation scolaire, nous n'avons pas rencontré de systèmes ou de recherches qui combinent tous les aspects des profils des intervenants tel que proposé dans cette thèse.

V.3. Points forts et limites de nos propositions

Tout au long de notre étude, nous avons eu à constater un ensemble de points forts de nos propositions. Aussi, nous avons relevé un certain nombre de limites. Nous allons étaler tous ces points dans ce qui suit.

V.3.1 Points forts de notre proposition

- Le système OPinOR renforce l'éducation au choix car la prise en compte de l'opinion des intervenants en général et celle des élèves en particulier au sujet des matières étudiées renforce la formation de ces derniers à choisir et non à subir une orientation préfabriquée.
- L'association des différents intervenants dans le système permet à ces derniers de jouer le rôle d'acteurs et non de spectateurs subissant une orientation imposée par l'administration.
- Les systèmes d'orientation scolaire existants juxtaposent les évaluations continues des élèves à leur image de soi. Ces systèmes, bien que performants et très répandus (dans les pays développés notamment), ils restent concentrés sur l'aspect exploration et découverte. Ils incitent l'élève à découvrir ses capitalisations, à se découvrir soi-même à travers des questionnaires psycho pédagogiques, à découvrir le monde du travail par une panoplie d'informations. Tout cela, Afin que l'élève fasse son choix. Notre étude s'inscrit aussi dans cette démarche, elle vient compléter ces aspects par la possibilité offerte aux intervenants d'introduire, en amont, leurs opinions sur des éléments de base que sont les matières étudiées. Ceci rendu possible grâce aux outils de l'analyse d'opinion. Ceci, infèrent de manière automatique les domaines d'études les plus propices. En d'autres termes, notre système aide les élèves à prendre leur décision lors du choix grâce à la synthèse de leurs opinions et celles d'intervenants non moins important dans ce processus.
- Les systèmes d'orientations scolaires offrent des informations riches sur les domaines d'études ainsi que les métiers afin de faciliter le choix aux élèves. Ils définissent aussi les matières pré-requises pour chacun des domaines et des métiers. Ils incitent, ainsi les élèves à exceller dans les matières pré-requises pour le domaine ou le métier souhaité sans tenir compte du sentiment éprouvé par l'élève envers chacune de ces matières. L'élève peut alors se trouver frustré car il met tous ses efforts dans des matières alors qu'il ne les apprécie pas forcément. Notre système permet d'aider les élèves à choisir des domaines d'études parmi ceux auxquels ils éprouvent des sentiments positifs, avec une capitalisation de connaissance acceptable et dont ils disposent des aptitudes nécessaires. Grace à cela, il sera dans une positive attitude dans son cycle universitaire. Sachant que la

positive attitude est recherchée actuellement car elle constitue un élément clé de succès.

- La positive attitude est octroyée dans notre cas par un sentiment de bien-être qui s'exprime par une opinion positive. L'intégration de l'opinion des intervenants dans un processus permet de renforcer l'acceptabilité par ces derniers du résultat obtenu. Nous l'avons vérifié dans notre étude, car notre système permet de rapprocher les résultats donnés par les scores d'orientation à ceux des choix explicites des élèves (Kappa = **0,7170**). Alors que l'orientation ordinaire donne un faible accord entre ces deux paramètres (Kappa= 0,2945).
- Le système que nous proposons implique un traitement automatisé d'un volume important de données. Cela n'implique pas, pour autant, de contraintes et ne surchargera pas les administrations.
- Dans notre étude, nous nous sommes penchées également sur la crédibilité du facteur opinion à travers l'engagement de son émetteur. Cet engagement qui se traduit par le taux de certitude contenu dans les textes d'opinion permet de distinguer les opinions les plus crédibles. Nous avons vérifié la corrélation entre la stabilité de l'opinion et le taux de certitude qui y est exprimé. La prise en compte de la certitude d'opinion dans l'orientation scolaire améliore sensiblement les résultats obtenus dans notre étude (Kappa = **0,6029** sans prise en compte de la certitude, Kappa = **0,7777** avec prise en compte de la certitude). Ce concept renforce les appuis de l'orientation scolaire intégrant l'analyse d'opinion, en se basant sur des paramètres crédibles.

VII.3.2 Points faibles de nos propositions

Les limites que nous avons constatées lors de notre étude et que nous comptons travailler dans nos recherches futures se résument à :

- **Concernant le module d'extraction de l'opinion libre**
 - Bien que l'implémentation actuelle de notre approche réponde aux besoins préliminaires de notre étude, elle requiert, tout au moins, un certain nombre de perfectionnements, nous citons l'amélioration de la pertinence des polarités des Synsets de DICO.
 - La deuxième limite est relative à la complexité d'analyse de texte avec une mixture de langues. Nous avons constaté, lors des enquêtes menées dans le cadre de la validation de notre approche, que certaines réponses données par

les élèves, sont un mélange d'arabe et de français, parfois d'anglais et de tamazight.

Cette contrainte a besoin, à notre sens, d'outils de traitement de langage naturel assez sophistiqués et spécifiques répondants à la prise en charge du multilingue et l'analyse de textes « mixte ».

- Nous avons aussi rencontré des cas où les réponses aux questions sont sarcastiques et parfois ironiques. Dans ces cas, la classification du texte ne se fait pas correctement.

- **Concernant le système d'orientation proposé**

Le système OPinOR a besoin d'être complété par un module d'information sur les domaines d'études et les métiers offerts ainsi que les débouchés de travail pour chacun d'eux. Ceci peut être pris en charge aisément par l'intégration d'un module reposant sur une base de données des métiers et des domaines d'études existants.

- **Concernant la mesure de certitude**

La limite de notre proposition à ce niveau réside dans la classification des verbes d'opinion, où une généralisation de cette classification à l'ensemble des verbes d'attitude contenus dans la ressource DICO est nécessaire. Aussi élargir cette classification à d'autres catégories de mots tels les adverbes afin de capter le maximum de certitude contenue dans les textes.

- **Concernant l'intégration de la certitude de l'opinion dans l'orientation scolaire**

Notre proposition à ce stade s'est axée sur l'amplification de l'opinion des élèves par le coefficient de certitude lors du calcul du score d'orientation. Bien que les résultats préliminaires soient assez intéressants, il est toutefois important de mener une étude plus approfondie en vue d'intégrer plus d'aspects de certitude d'opinion dans l'orientation scolaire et aussi pour d'autres applications.

V.4 Impacts de notre proposition

L'orientation scolaire est une problématique épineuse dans le domaine de l'éducation. Les statistiques parlent d'elles-mêmes lorsque nous observons les taux d'échec en première année universitaire. Ceci, est causé principalement par une mauvaise orientation des élèves notamment après l'obtention du baccalauréat. Les élèves se retrouvent à étudier des domaines

qui parfois ne les intéressent nullement. Ils sont alors complètement démotivés et parfois renoncent carrément à leur scolarité. Les problèmes de la jeunesse sans formation sont connus de tous. Les jeunes virent rapidement vers toutes sortes de fléaux sociaux.

V.4.1 Impact sur l'apprentissage et l'enseignement

- Le système que nous avons proposé intervient pour améliorer la qualité de l'orientation scolaire en aidant d'une part les élèves à choisir les domaines d'études favorables à leurs attentes. Et d'autre part, il va assister l'administration à garantir une orientation adéquate aux élèves tout en respectant les délais qui lui sont impartis.
- L'intégration de l'opinion dans le processus d'orientation permet aux élèves d'assumer pleinement leurs choix et de renforcer ainsi leurs motivations que ce soit durant leur cursus de formation ou dans leur vie active. Nous pensons que cela impactera de manière positive le rendement scolaire en diminuant notamment le taux de décrochage.
- Les enseignants aussi s'impliqueront plus dans leurs tâches quotidiennes sachant que leurs opinions auront un impact sur le devenir de chacun de leurs élèves.
- Le traitement automatique des données, notamment celle relatives aux opinions, offert par notre système assure une meilleure transparence et donc plus d'équité, dans le processus de l'orientation scolaire.

V.4.2 Impact la citoyenneté

- L'amélioration du processus d'orientation scolaire a un impact direct sur la vie sociale. En effet elle contribue à la formation de citoyens responsables et émancipés. Ce qui permet d'éloigner les citoyens notamment les jeunes des différents fléaux sociaux.
- Sur le plan économique, nos propositions contribueront, à travers l'amélioration de l'orientation scolaire, à offrir au marché du travail, une main d'œuvre plus qualifiée et plus impliquée. Car cette main œuvre aura à exercer des métiers qui lui sont favorables et qui renforcent sa positive attitude. Ces facteurs réunis offrent toutes les chances de réussite que ce soit pour le citoyen où pour les entreprises.

V.4.3 Impact sur la recherche en E-Education versus E-orientation

- Notre étude est une contribution aux applications de l'analyse d'opinion au domaine de l'éducation en générale et l'e-orientation en particulier. Nous avons proposé une application de l'analyse d'opinion qui aide à améliorer la prise de décision et offre une meilleure satisfaction aux acteurs concernés.
- Ce processus contribue, avec un système d'aide intégrant les outils d'intelligence artificielle, à accélérer le traitement des données massives générées à cet effet. Cette étude peut être améliorée en intégrant d'autres profils des intervenants dans le processus d'orientation notamment, des

profils de l'élève.

- Cette étude présente un système de prise de décision offrant une synthèse multiniveaux d'opinion. Cette technique peut être généralisée à d'autres processus relevant de l'éducation ou autres.
- Elle ouvre une ébauche pour la fiabilisation de l'opinion par la mesure de certitude. Cette proposition peut impacter toutes les applications de l'analyse d'opinion qui intègre ce paramètre dans les systèmes de prise de décision et ou de recommandation.

Conclusion

Nous avons présenté dans ce dernier chapitre de notre thèse, une synthèse et une discussion des résultats des expérimentations conduites pour la validation de nos propositions. Nous avons mis en évidence les points forts et les impacts de nos contributions. Nous avons aussi mis l'accent sur les insuffisances de notre étude, qui feront l'objet de perspectives d'améliorations de nos travaux.

Conclusion Générale Et Perspectives De Recherche

L'orientation constitue un pivot dans la vie des élèves d'aujourd'hui, adultes de demain. Elle détermine, de manière inéluctable, leurs chances de réussite. Elle ne doit, de ce fait, rester à l'écart de l'évolution technologique. Elle pose un nombre de questions qui ne doivent pas laisser la communauté scientifique indifférente. Parmi ces questions, l'enrichissement de ses critères par des éléments fiables regroupant tous les intervenants dans ce processus.

Nous avons proposé un système d'E-orientation basé sur l'analyse d'opinion (OPinOR). Il a pour objectif d'améliorer l'orientation des élèves du secondaire vers les domaines d'études universitaires ou professionnelles. Il enrichit les critères d'orientation par l'intégration de l'opinion des intervenants dans ce processus d'une part et les aptitudes des élèves représentés par leurs intelligences multiples de Gardner d'autre part. Il conjugue ces deux éléments aux évaluations scolaires, critère de base des systèmes classiques d'orientation. Il permet, en effet de faire collaborer tous les acteurs concernés par l'orientation scolaire à savoir les élèves, leurs parents ainsi que leurs enseignants. Les opinions de tous ces intervenants sont prises en compte pour proposer aux élèves des cartes d'orientation leur permettant de faire le choix du domaine d'étude propice à leurs attentes.

Pour réaliser ce système, nous avons conçu une ressource lexicale de la langue française basée à la fois sur SentiWordNet et Wolf. Cette ressource est utilisée pour l'extraction de l'opinion à partir de textes.

Le système OPinOR propose de recueillir, périodiquement, les opinions des intervenants à travers des enquêtes menées avec des questionnaires en ligne. Ces questionnaires comportent des questions fermées dédiées à recueillir des opinions formatées, des questions ouvertes destinées à recueillir les opinions libres sous forme de textes. Les questionnaires relatifs aux élèves servent à recueillir l'opinion de ces derniers sur les matières étudiées. Ceux des parents portent aussi sur les matières étudiées par leurs enfants. Les questionnaires des enseignants, par contre servent à recueillir les opinions de ces derniers au sujet de leurs élèves.

Ces questionnaires sont traités afin d'extraire les opinions. Celles-ci sont amplifiées par des coefficients de certitude calculés à partir du texte des réponses libres. Ces coefficients permettent de distinguer les opinions fiables.

En fin de cycle, le système procède à la synthèse de toutes les opinions calculées. Il les combine aux valeurs des aptitudes et celles des évaluations scolaires des élèves. Il leur octroie ainsi, une carte d'orientation permettant de les aider dans le choix du domaine d'étude.

Nous avons testé nos propositions grâce à deux prototypes. Le premier, nous l'avons effectué avec un échantillon d'élèves de terminal. Le deuxième, nous l'avons mis en œuvre sur un échantillon de 100 élèves lycéens. Les résultats obtenus sont intéressants.

Nos propositions pour la e-orientation en éducation peuvent être généralisées pour avoir une plate-forme de génération de cartes d'orientations pour d'autres domaines comme la médecine, le commerce, ... où il y a lieu d'intégrer l'analyse d'opinion.

En dépit que les objectifs fixés pour notre thèse soient atteints, le système OPinOR peut être amélioré et complété. Nos perspectives de recherches dans ce sens sont :

- **A court terme**
 - Nous envisageons de mettre en œuvre un processus de recueil implicite d'opinion libre plus approprié à travers des forums dédiés en remplacement aux questionnaires.
 - Améliorer les polarités de la ressource lexicale proposée DICO via un processus d'apprentissage automatique sur un corpus de commentaires issus du domaine de l'éducation.
- **À moyen terme**
 - Le système OPinOR peut réutiliser les travaux sur les logiciels d'aide à l'orientation (LAO), tels ceux proposés dans (Feral, 2013) pour intégrer des modules d'informations et de communication sur les domaines d'études et les métiers.
 - Il peut être greffé comme module d'orientation automatique dans la plateforme MOOC proposé dans (Haddadi, et al., 2018 -a).
 - Pour améliorer la mesure de certitude, élargir la classification des verbes d'attitudes et étendre cette proposition aux adverbes.
 - Pour améliorer la crédibilité de l'opinion, mener une étude approfondie sur l'intégration de la certitude dans le processus de E-orientation basé sur l'analyse d'opinion.

- Dans le même contexte, approfondir la classification d'opinion par la détection de l'ironie et le sarcasme, fréquents dans les réponses des élèves.
- **A long terme**
 - Dans les travaux futurs et dans l'optique de valoriser les résultats obtenus, nous souhaiterons présenter notre recherche à l'observatoire nationale de l'éducation et de la formation afin de bénéficier d'un échantillon beaucoup plus significatif pour tester notre proposition sur le terrain en vue de son adoption.

Références bibliographiques Et Webographies

Références bibliographiques

- (Alfaro , et al., 2016) Alfaro, C., Cano-Montero, J., Gómez, J., Moguerza, J. M., & Ortega, F. 2016. A multi-stage method for content classification and opinion mining on weblog comments. *Annals of Operations Research*, 236(1), pp. 197-213.
- (Altrabsheh, et al., 2013) Altrabsheh, N., Gaber, M. & Cocea, M., 2013. *SA-E: sentiment analysis for education*. Sorrento Italy, International Conference on Intelligent Decision Technologies, Springer, pp. 353-362.
- (Baccianella, et al., 2010) Baccianella, S., Esuli, A. & Sebastiani, F., 2010. Sentiwordnet 3.0: an enhanced lexical resource for sentiment analysis and opinion mining. *LREC*, 10(2010), pp. 2200-2204.
- (Baeza-Yates & Ribeiro-Neto , 1999) Baeza-Yates , R. & Ribeiro-Neto , B., 1999. *Modern information retrieval*. New York, ACM Press.
- (Balahadia, et al., 2016) Balahadia, F. F., Fernando, G. M. C. & Juana, C. I., 2016. *Teacher's performance evaluation tool using opinion mining with sentiment analysis*. Sanur Indonésie, 2016 IEEE Region 10 Symposium (TENSYP), IEEE, pp. 95-98.
- (Balahur, et al., 2012) Balahur, A., Hermida, J. M. & Montoyo, A., 2012. Detecting implicit expressions of emotion in text: A comparative analysis. *Decision Support Systems*, Volume 53, p. 742–753.
- (Bandura, 1986) Bandura, A., 1986. Social foundations of thought and action. *Englewood Cliffs*, 5(1), 23-28.
- (Banea, et al., 2014) Banea, C., Mihalcea, R. & Wiebe, J., 2014. Sense-level Subjectivity in a Multilingual Setting. *Computer Speech & Language*, 28(1), pp. 7-19.
- (Baucom, et al., 2013) Baucom, E., Sanjari, A., Liu, X. & Chen, M., 2013. *Mirroring the real world in social media: Twitter, geolocation, and sentiment analysis*. San Francisco, California, USA, Proceedings of the 2013 international workshop on Mining unstructured big data using natural language processing, ACM, pp. 61-68.

- (Ben Flisse, 2014) Ben Flisse, K., 2014. *La référence dans l'orientation scolaire et professionnelle*. Alger : Office des Publications Universitaires (OPU).
- (Berou, 2010) Berou, M., 2010. *Impact de l'orientation scolaire sur la réussite scolaire au secondaire*. Tizi-Ouzou : El Amel.
- (Bing, et al., 2014) Bing, L., Chan, K. C. & Ou, C., 2014. *Public Sentiment Analysis in Twitter Data for Prediction of a Company's Stock Price Movements*. Guangzhou -Chine, 2014 IEEE 11th International Conference on e-Business Engineering, pp. 232-239.
- (Bloch, et al., 1997) Bloch , H, Dépret, E, Gallo, A, Garnier , Ph, Gineste , M D, Leconte, P, Le Ny , J F, Postel , J Reuchlin, M, Casalis, D., 1997. *Dictionnaire fondamental de la psychologie*. Larousse.
- (Bollegala, et al., 2013) Bollegala, D., Weir, D. & Carroll, J., 2013. Cross-domain sentiment classification using a sentiment sensitive thesaurus, *IEEE Transactions on Knowledge and Data Engineering*, 25(8), pp. 1719-1731.
- (Bollen, et al., 2011) Bollen, J., Mao, H. & Zeng, X., 2011. Twitter mood predicts the stock market. *Journal of Computational Science*, 2(1), pp. 1-8.
- (Bond & Paik, 2012) Bond, F. & Paik, K., 2012. A Survey of WordNets and their Licenses. *Small*, 8(4), pp. 5-12.
- (Boone & Boone, 2012) Boone, H. N. & Boone, D. A., 2012. Analyzing Likert data. *Journal of extension*, 50(2), pp. 1-5.
- (Boullier & Lohard, 2012) Boullier, D. & Lohard, A., 2012. *Opinion mining et Sentiment analysis: Méthodes et outils*. Marseille, France: Open Edition Press.
- (Bravo-Marquez, et al., 2014) Bravo-Marquez, F., Mendoza, M. & Poblete, B., 2014. Meta-Level Sentiment Models for Big Social Data Analysis. *Knowledge-Based Systems*, Volume 69, pp. 86-99.
- (Cambria, et al., 2010) Cambria, E., Speer, R., Havasi, C. & Hussain, A., 2010. *SenticNet: A Publicly Available Semantic Resource for Opinion Mining*. Arlington, Virginia, In: Commonsense Knowledge: Papers from the AAAI Fall Symposium, pp. 14-18.

- (Chaumartin, 2007)** Chaumartin, F.-R., 2007. *WordNet et son écosystème : un ensemble de ressources linguistiques de large couverture*. Montréal, Canada, Colloque BD lexicales.
- (Chen, et al., 2014)** Chen, L., Wang, F., Qi, L. & Liang, F., 2014. Experiment on sentiment embedded comparison interface. *Knowledge-Based Systems*, Volume 64, pp. 44-58.
- (Cho, et al., 2014)** Cho, H., Kim, S., Lee, J. & Lee, J.-S., 2014. Data-driven integration of multiple sentiment dictionaries for lexicon-based sentiment classification of product reviews. *Knowledge-Based Systems*, Volume 71, pp. 61-71.
- (Chvanova, et al., 2016)** Chvanova, M. S., Hramov, A. E., Khramova, M. V. & Pitsik, E. N., 2016. *Is it possible to improve the university education with social networks: The opinion of students and teachers?* St. Petersburg , Russia, Conference on Quality Management, Transport and Information Security, Information Technologies (IT&MQ&IS), IEEE, pp. 33-38.
- (Colace, et al., 2014)** Colace, F., De Santo, M., Greco, L. & Guerriero, G., 2014. *Sentiment Analysis and E-Learning: a Proposal*. Riga, Latvia, ICEE/ICIT International Conference on Engineering Education & International Conference on information Technology, pp. 268-277.
- (Collins & Quillian, 1969)** Collins, A. M. & Quillian, M. R., 1969. Retrieval time from semantic memory. *Journal of Verbal Learning and Verbal Behavior*, 8(2), pp. 240-247.
- (Demeuse & Lafontaine, 2005)** Demeuse, M. & Lafontaine, D., 2005. L'orientation scolaire en communauté française de Belgique. *International Journal of Education Sèvres*, Volume 38, pp. 35-52.
- (Di Caro & Grella , 2013)** Di Caro, L. & Grella, M., 2013. Sentiment analysis via dependency parsing. *Computer Standards & Interfaces*, 35(5), pp. 442-453.
- (Dragos & Jaulent, 2010)** Dragos, V. & Jaulent, M.-C., 2010. *Apprentissage de patrons lexico-syntaxiques à partir de texte*. Hammamet, Tunisie., Association Internationale Francophone d'Extraction et de Gestion des Connaissances, pp. 615-620.

- (Dufour, 2002)** Dufour, J. M., 2002. *Lissage exponentiel*. Montréal: Université de Montréal.
- (Eckert, 2006)** Eckert, P., 2006. Communities of practice. Encyclopedia of Language and Linguistics. Elsevier. *Perspective, USA Empowerment in Organizations*, 6(7), pp. 177-186.
- (ElKateb, et al., 2006)** ElKateb, S., Black, W., Rodríguez, H., Alkhalifa, M., Vossen, P., Pease, A., & Fellbaum, C. 2006. *Building a WordNet for Arabic*. Genoa –Italy, LREC, pp. 29-34.
- (Esuli & Sebastiani, 2006)** Esuli, A. & Sebastiani, F., 2006. Sentiwordnet: A publicly available lexical resource for opinion mining. *LREC*, Volume 6, pp. 417-422.
- (Farek, 2009)** Farek, L., 2009. *Identification d'opinions dans les journaux arabes*. Annaba, Algérie: Thèse de Doctorat, Université Badji Mokhtar.
- (Fellbaum, 1999)** Fellbaum, C., 1999. La représentation des verbes dans le réseau sémantique "WORDNET". *Langages*, Volume 33(136), pp. 27-40.
- (Fellbaum, 1998)** Fellbaum, C. 1998. A semantic network of English verbs. *WordNet: An electronic lexical database*, Volume 3, pp. 153-178.
- (Gardin, 2009)** Gardin, P., 2009. *Application de la théorie de l'Appraisal à l'analyse d'opinion*. Avignon France, Acte de MANifestation des JEunes Chercheurs en Sciences et Technologies de l'Information et de la Communication (MajecSTIC).
- (Gardner, 2006)** Gardner, H., 2006. *multiples Intelligences – New Horizons*. New York : Basic books.
- (Gardner, 2011)** Gardner, H., 2011. *Frames of mind: The theory of multiple intelligences*. New York : Basic books.
- (Greenwood, et al., 2005)** Greenwood, M. M. S., Guo, Y., Harkema, H. & Roberts, A., 2005. *Automatically acquiring a linguistically motivated genic interaction extraction system*. Bonn, Germany, Proceedings of the 4th Learning Language in Logic Workshop, pp. 46-52.

- (Gruber, 1995)** Gruber, T. R., 1995. Toward principles for the design of ontologies used for knowledge sharing? *International journal of human-computer studies*, 43(5), pp. 907-928.
- (Guerry, et al., 1993)** Guerry, M., Catelain, A. & Caron, J., 1993. La compréhension de marqueurs modaux : verbes d'attitude propositionnelle et adverbs. *L'Année psychologique*, 2(93), pp. 201-225.
- (Guichard & Huteau, 2001)** Guichard, J. & Huteau, M., 2001. *Psychologie de l'Orientalion*. Paris : Dunod.
- (Haddadi, et al., 2018 -a)** Haddadi, L., Bouarab-Dahmani, F., Guin, N., Berkane, T., & Lazib, S. 2018. Peer assessment and groups formation in massive open online courses. *Computer Applications in Engineering Education*, 26(5), pp. 1873-1887.
- (Haddadi, et al., 2018 -b)** Haddadi, L., 2018. *Modélisation et évaluation des apprenants dans les MOOCs basés sur l'approche ODALA*. Tizi-Ouzou : Thèse de doctorat, Université Mouloud Mammeri Tizi-Ouzou.
- (Haddadi, et al., 2018-c)** Haddadi, L., Bouarab-Dahmani, F. & Guin, N., 2018. Multi-Level Learner Assessment in Massive Open Online Courses. *International Journal of Knowledge and Learning*, 12(4), pp. 325-351.
- (Hamzah & Widyastuti, 2016)** Hamzah, A. & Widyastuti, N., 2016. *Opinion extracting and classification from questionnaire comments using HMM-POS Tagger and machine learning techniques*. Denpasar, Indonesia, Data and Software Engineering (ICoDSE), International Conference on. IEEE, pp. 1-6.
- (Haseena, 2014)** Haseena, R. P., 2014. Opinion Mining and Sentiment Analysis-Challenges and Applications. *International Journal of Application or Innovation in Engineering & Management (IJAIEM)*, 3(5) pp.401-403.
- (Hatzivassiloglou & McKeown, 1997)** Hatzivassiloglou, V. & McKeown, K. R., 1997. *Predicting the semantic orientation of adjectives*. Madrid, Spain . In 35th annual meeting of the association for computational linguistics and 8th conference of the european chapter of the association for computational linguistics, pp. 174–181.

- (Hawk & Ami, 2007)** Hawk, T. F. & Amit, J. S., 2007. Using learning style instruments to enhance student learning. *Decision Sciences Journal of Innovative Education*, 5(1), pp. 1-19.
- (Hemmatian & Sohrabi, 2017)** Hemmatian, F. & Sohrabi, M. K., 2017. A survey on classification techniques for opinion mining and sentiment analysis. *Artificial Intelligence Review*, pp. 1-51.
- (Hermann, 2010)** Hermann, C., 2010. *Entre Web 2.0 et 3.0 : opinion mining*. Thèse de Doctorat, Haute Ecole de Gestion & Tourisme.
- (Heydari, et al., 2015)** Heydari, A., Tavakoli, M. A., Salim, N. & Heydari, Z., 2015. Detection of review spam: A survey. *Expert Systems with Applications*, 42(7), pp. 3634-3642.
- (Hossin & Sulaiman, 2015)** Hossin, M. & Sulaiman, M. N., 2015. A review on evaluation metrics for data classification evaluations. *International Journal of Data Mining & Knowledge Management Process (IJDKP)*, 5(2), pp.1-11.
- (Huang & Yen, 2013)** Huang, H. & Yen, D. C., 2013. Predicting the Helpfulness of Online Reviews—A Replication. *International Journal of Human-Computer Interaction*, 29(2), pp. 129-138.
- (Hubert, 1970)** Hubert, R., 1970. *Traité de pédagogie générale*. Paris : Presses Universitaires de France.
- (Hung & Lin, 2013)** Hung, C. & Lin, H.-K., 2013. Using Objective Words in SentiWordNet to Improve Word-of Mouth Sentiment Classification. *IEEE Intelligent Systems*, (2), pp. 47-54.
- (Invento, et al., 2017)** Invento, A. C. B., Lerias, T. J. & Ceniza, A. M., 2017. *Sentiment Analysis on the Impact of K-12 Program in the Philippines using Naïve Bayes and Lexicon Approach with Code Switching*. Singapore, Singapore, The 2017 International Conference on Information Technology, ACM, pp. 103-106.

- (Jeyapriya & Selvi, 2015)** Jeyapriya, A. & Selvi, C. S. K., 2015. *Extracting aspects and mining opinions in product reviews using supervised learning algorithm*. Coimbatore, India, 2nd International Conference on Electronics and Communication Systems (ICECS), IEEE, pp. 548-552.
- (Kechaou, et al., 2011)** Kechaou, Z., Ben Ammar, M. & Alimi, A. M., 2011. *Improving e-learning with sentiment analysis of users' opinions*. Amman, Jordan, 2011 IEEE global engineering education conference (EDUCON), IEEE, pp. 1032-1038.
- (Khan, et al., 2016)** Khan, F. H., Qamar, U. & Bashir, S., 2016. SentiMI : Introducing point-wise mutual information with SentiWordNet to improve sentiment polarity detection. *Applied Soft Computing ACM*, Volume 39, pp. 140-153.
- (Khomonenko, et al., 2017)** Khomonenko, A. D., Dashonok, V. L., Ivanova, K. A. & Kassymova, D. T., 2017. *Approach to processing of data from social networks for detecting public opinion on quality of educational services*. St. Petersburg, Russia, 2017 XX IEEE International Conference on Soft Computing and Measurements (SCM). IEEE, pp. 737-739.
- (Koehler, et al., 2015)** Koehler, M. J., Greenhalgh, S. & Zellner, A., 2015. *Potential Applications of Sentiment Analysis in Educational Research and Practice—Is SITE the Friendliest Conference?* Las Vegas USA, Association for the Advancement of Computing in Education (AACE), pp. 1348-1354.
- (Kreutzer & Witte, 2013)** Kreutzer, J. & Witte, N., 2013. *Opinion Mining Using SentiWordNet*. Uppsala : Université Uppsala.
- (Krishnamoorthy, 2015)** Krishnamoorthy, S. 2015. Linguistic features for review helpfulness prediction. *Expert Systems with Applications*, 42(7), pp. 3751-3759.
- (Lagnoux, 1996)** Lagnoux, A., 1996. *Séries chronologiques*. Toulouse France : Université de Toulouse Le Mirail.

- (Lark, et al., 2015) Lark, J., Morin, E. & Saldarriaga, S. P., 2015. *CANÉPHORE : un corpus français pour la fouille d'opinion ciblée*. Caen France, 22e conférence sur le Traitement Automatique des Langues Naturelles(TALN).
- (Lazib & Bouarab-Dahmani, 2017) Lazib, S. & Bouarab-Dahmani, F., 2017. *Integrating of opinion holder engagement in an educational guidance help system*. Adrar, Algérie, In International Conference on Mathematics and Information Technology (ICMIT), pp. 247-254.
- (Lazib, et al., 2019) Lazib, S., Bouarab-Dahmani, F. & Haddadi, L., 2019. Educational Guidance Process Integrating Opinion Mining. *Computer Applications in Engineering Education*, 27(5), pp. 1139-1155.
- (Lazib, et al., 2016 -a) Lazib, S., Bouarab-Dahmani, F. & Haddadi, L., Juil. 2016. *Orientation process for secondary school pupils based on opinion analysis*. Barcelonne, Espagne, the 8th annual International Conference on Education and New Learning Technologies, pp. 3229- 3239.
- (Lazib, et al., 2016 -b) Lazib, S., Bouarab-Dahmani, F. & Haddadi, L., Nov. 2016. *Analyse d'opinions pour l'orientation scolaire*. Msila Algérie, the Second International Symposium on Informatics and its Applications, pp. 87.
- (Lesk, 1986) Lesk, M., 1986. *Automatic sense disambiguation using machine readable dictionaries: how to tell a pine cone from an ice cream cone*. Proceedings of the 5th annual international conference on Systems documentation, Citeseer, pp. 24-26.
- (Li & Liu, 2012) Li, G. & Liu, F., 2012. Application of a clustering method on sentiment analysis. *Journal of Information Science*, 38(2), pp. 127–139.
- (Li, et al., 2012) Li, G., Chang, K. & Hoi, S. C. H., 2012. Multiview Semi-Supervised Learning with Consensus. *IEEE Transactions on Knowledge and Data Engineering*, 24(11), pp. 2040 - 2051.
- (Liao, et al., 2016) Liao, C., Feng, C., Yang, S. & Huang, H., 2016. A hybrid method of domain lexicon construction for opinion targets extraction using syntax and semantics. *Journal Comput Sci Technol*, Volume 31, pp. 595–603.

- (Littré, 1984)** Littré, E., 1984. *Dictionnaire de la langue française*. Hachette & cie.
- (Liu & Zhang, 2012)** Liu, B. & Zhang, L., 2012. *A Survey of Opinion Mining and Sentiment Analysis*. Boston, MA., Mining text data. Springer, pp.115-163.
- (Liu, 2010)** Liu, B., 2010. Sentiment analysis and subjectivity. *Handbook of Natural Language Processing, Second Edition*, 2(2010), pp. 627-666.
- (Liu, 2012)** Liu, B., 2012. Sentiment analysis and opinion mining. *Synthesis lectures on human language technologies*, 5(1), pp. 1-167.
- (Liu, 2015)** Liu, B., 2015. *Sentiment analysis: mining opinions, sentiments, and emotions*. Cambridge: Cambridge University Press.
- (Luo, et al., 2015)** Luo, W. et al., 2015. QPLSA: Utilizing quad-tuples for aspect identification and rating. *Information Processing & Management*, 51(1), pp. 25-41.
- (Lv, et al., 2017)** Lv, Y., Liu, J., Chen, H., Mi, J., Liu, M., & Zheng, Q., 2017. Opinioned post detection in Sina Weibo. *IEEE Access*, Volume 5, pp. 7263-7271.
- (Maccario, 1987)** Maccario, B., 1987. Cardinet (Jean). Evaluation scolaire et mesure. Cardinet (Jean). Evaluation scolaire et pratique. Cardinet (Jean). Pour apprécier le travail des élèves. *Revue française de pédagogie*, 81(1), pp. 103-108.
- (Marchand, 2015)** Marchand, M., 2015. *Domaines et fouille d'opinion*. Paris France : Université Paris-Sud.
- (Marrese-Taylor, et al., 2014)** Marrese-Taylor, E., Velásquez, J. D. & Bravo, F., 2014. A novel deterministic approach for aspect-based opinion mining in tourism products reviews. *Expert Systems with Applications*, 41(17), pp. 7764-7775.
- (Martin & White, 2003)** Martin, J. R. & White, P. R., 2003. *The language of evaluation*. Macmillan: Springer.

- (Martineau, et al., 2011) Martineau, C., Voyatzi, S., Varga, L., Brizard, S., & Migeotte, A. 2011. *Détection fine d'opinion et sentiments : attribution fine de la polarité et calcul incrémental de l'intensité*. Nicosia, Cyprus, 30th International Conference on Lexis, pp. 319-334.
- (Masson, 1997) Masson, P. 1997. Elèves, parents d'élèves et agents scolaires dans le processus d'orientation. *Revue française de sociologie*, 38(1), pp. 119-142.
- (Matmati, 2001) Matmati, M. 2001. Les enjeux de l'intranet dans la GRH. *Les Cahiers du Management Technologique*, 11(2), pp. 53-68.
- (Maurel, et al., 2007) Maurel, S., Curtoni, P., & Dini, L. 2007. *Classification d'opinions par méthodes symbolique, statistique et hybride*. Actes du troisième DÉFI Fouille de Textes, pp. 121-127.
- (Maurel, et al., 2008) Maurel, S., Curtoni, P. & Dini, L., 2008. *L'analyse des sentiments dans les forums*. Fontainebleau France, INFORSID 2008, Atelier Fouille des Données d'Opinion, pp. 9-22.
- (Medagoda, et al., 2015) Medagoda, N., Shanmuganathan, S. & Whalley, J. L., 2015. *Sentiment lexicon construction using SentiWordNet 3.0*. Zhangjiajie, China, 11th International Conference on Natural Computation (ICNC) IEEE, pp. 802-807.
- (Medhat, et al., 2014) Medhat, W., Hassan, A. & Korashy, H., 2014. Sentiment analysis algorithms and applications: A survey. *Ain Shams engineering journal*, 5(4), pp. 1093-1113.
- (Meunier, 2008) Meunier, O., 2008. *Orientation scolaire et insertion professionnelle, Approches sociologiques*. National Institute for Educational Research.
- (Mihalcea & Banea, 2007) Mihalcea, R., Banea, C., & Wiebe, J. 2007. *Learning multilingual subjective language via cross-lingual projections*. In Proceedings of the 45th annual meeting of the association of computational linguistics, pp. 976-983.
- (Miller & Fellbaum, 1991) Miller, G. A. & Fellbaum, C., 1991. Semantic networks of English. *Cognition*, 41(1-3), pp. 197-229.

- (Miller, et al., 1990)** Miller, G. A. et al. 1990. Introduction to WordNet : An On-line Lexical Database*. *Special issue of the International Journal of Lexicography*, 3(4), pp. 235–244.
- (Miller, 1995)** Miller, G. A., 1995. WordNet: a lexical database for English. *Communications of the ACM*, 38(11), pp. 39-41.
- (Min & Park, 2012)** Min, H. J., & Park, J. C. 2012. Identifying helpful reviews based on customer’s mentions about experiences. *Expert Systems with Applications*, 39(15), pp. 11830-11838.
- (Molina-González, et al., 2013)** Molina-González, M. D., Martínez-Cámara, E., Martín-Valdivia, M.-T. & Perea-Ortega, J. M., 2013. Semantic orientation for polarity classification in Spanish reviews. *Expert Systems with Applications*, 40(18), pp. 7250-7257.
- (Molina-González, et al., 2015)** Molina-González, M. D., Martínez-Cámara, E., Martín-Valdivia, M. T., & Ureña-López, L. A. 2015. A Spanish semantic orientation approach to domain adaptation for polarity classification. *Information Processing & Management*, 51(4), pp. 520-531.
- (Montoyo, et al., 2012)** Montoya, A., Martínez-Barco, P. & Balahur, A., 2012. Subjectivity and sentiment analysis: An overview of the current state of the area and envisaged developments. *Decision Support Systems*, Volume 53, pp. 675-679.
- (Moreo, et al., 2012)** Moreo, A., Romero, M., Castro, J. & Zurita, J. M., 2012. Lexicon-based comments-oriented news sentiment analyzer. *Expert Systems with Applications*, 39(10), pp. 9166-9180.
- (Ngo-Ye & Sinha, 2014)** Ngo-Ye, T. L., & Sinha, A. P. 2014. The influence of reviewer engagement characteristics on online review helpfulness: A text regression model. *Decision Support Systems Elsevier*, Volume 61, pp. 47-58.
- (Nitin, et al., 2015)** Nitin, G. I., Swapna, G. & Shankararaman, V., 2015. *Analyzing educational comments for topics and sentiments: A text analytics approach*. El Paso, Texas, USA, 2015 IEEE Frontiers in Education Conference (FIE), IEEE, pp. 1-9.

- (Nofer & Hinz, 2015)** Nofer, M. & Hinz, O., 2015. Using Twitter to predict the stock market. *Business & Information Systems Engineering*, 57(4), pp. 229–242.
- (O'Connor, et al., 2010)** O'Connor, B., Balasubramanyan, R., Routledge, B. R. & Smith, N. A., 2010. *From Tweets to Polls: Linking Text Sentiment to Public Opinion Time Series*. Washington USA, Fourth International AAAI Conference on Weblogs and Social Media, pp. 122-129.
- (Olson & Delen, 2008)** Olson, D. L. & Delen, D., 2008. *Advanced data mining techniques*. Berlin, Springer Science & Business Media.
- (Omar, et al., 2017)** Omar, M. S., Njeru, A., Paracha, S., Wannous, M., Yi, S., 2017. *Mining tweets for education reforms*. In 2017 International Conference on Applied System Innovation (ICASI), IEEE, pp. 416-419.
- (Oramas Bustillos, et al., 2019)** Oramas Bustillos, R., Zatarain Cabada, R., Barrón Estrada, M. L. & Hernández Pérez, Y., 2019. Opinion mining and emotion recognition in an intelligent learning environment. *Computer Applications in Engineering Education*, 27(1), pp. 90-101.
- (Ortigosa, et al., 2014)** Ortigosa, A., Martín, J. M. & Carro, R. M., 2014. Sentiment analysis in Facebook and its application to e-learning. *Computers in Human Behavior*, Volume 31, pp. 527-541.
- (Ostertagova & Ostertag, 2012)** Ostertagova, E. & Ostertag, O., 2012. Forecasting using simple exponential smoothing method. *Acta Electrotechnica et Informatica*, 12(3), pp. 62-66.
- (Pandarachalil, et al., 2015)** Pandarachalil, R., Sendhilkumar, S. & Mahalakshmi, G., 2015. Twitter sentiment analysis for large-scale data: an unsupervised approach. *Cognitive computation*, 7(2), pp. 254-262.
- (Pang & Lee, 2004)** Pang, B. & Lee, L., 2004. *A sentiment education: Sentiment analysis using subjectivity summarization based on minimum cuts*. Barcelona, Spain, Proceedings of the 42nd Annual Meeting of the Association for Computational Linguistics (ACL-04), Association for Computational Linguistics, pp. 271-278.

- (Pang & Lee, 2008)** Pang, B. & Lee, L., 2008. Opinion Mining and Sentiment Analysis. *Foundations and Trends® in Information Retrieval*, 2(1-2), pp. 1-135.
- (Pang, et al., 2002)** Pang, B., Lee, L. & Vaithyanathan, S., 2002. *Thumbs up? : Sentiment classification using machine-learning techniques*. In Proceedings of the 2002 Conference on Empirical Methods in Natural Language Processing, Association for Computational Linguistics, pp. 79-86.
- (Pedersen, et al., 2004)** Pedersen, T., Patwardhan, S. & Michelizzi, J., 2004. *WordNet::Similarity: measuring the relatedness of concepts*. Boston, Massachusetts, *Demonstration Papers at HLT-NAACL*, pp. 38-41.
- (Pieron & Salinas, 1969)** Pieron, H. & Salinas, M. S., 1969. *Examens et docimologie - La Evaluación de los Aprendizajes en la Universidad, Facultad de Educación*. Presses universitaires de France - Universidad de Antioquia.
- (Polanco, 2001)** Polanco, X., 2001. *Textmining et intelligence économique : aujourd'hui et demain*. Colloque Veille technologique, Intelligence économique et Bibliométrie.
- (Pugin, 2008)** Pugin, V., 2008. *L'orientation scolaire et professionnelle*, Lyon, France, Millénium : The Prospective Resource Center of Greater Lyon.
- (Purnawirawan, et al., 2012)** Purnawirawan, N., De Pelsmacker, P. & Dens, N., 2012. Balance and Sequence in Online Reviews: How Perceived Usefulness Affects Attitudes and Intentions. *Journal of Interactive Marketing*, Volume 26, pp. 244–255.
- (Qiu, et al., 2011)** Qiu, G., Liu, B., Bu, J. & Chen, C., 2011. Opinion word expansion and target extraction through double propagation. *Computer Linguistics*, 37(1), pp. 9-27.
- (Rajput & Solanki, 2016)** Rajput, R. & Solanki, A. K., 2016. Review of Sentimental Analysis methods using lexicon based approach. *IJCSMC*, 5(2), pp. 159-166.
- (Rana & Cheah, 2016)** Rana, T. A. & Cheah, Y.-N., 2016. Aspect extraction in sentiment analysis: comparative analysis and survey. *Artificial Intelligence Review*, 46(4), pp. 459-483.

- (Ravi & Ravi, 2015)** Ravi, K. & Ravi, V., 2015. A survey on opinion mining and sentiment analysis: tasks, approaches and applications. *Knowledge-Based Systems*, Volume 89, pp. 14-46.
- (Regragui, et al., 2016)** Regragui, Y. et al., 2016. *Arabic wordnet: New content and new applications*. In Proceedings of the Eighth Global WordNet Conference, pp. 330-338.
- (Richit, 2014)** Richit, N., 2104. L'éducation à l'orientation dans le cadre d'un enseignement d'exploration en seconde. *Recherches en didactiques*, 2(18), pp. 101-118.
- (Riloff & Wiebe, 2003)** Riloff, E. & Wiebe, J., 2003. *Learning extraction patterns for subjective expressions*. Morristown, USA, Association for Computational Linguistics, pp. 105–112.
- (Rudnicka, et al., 2012)** Rudnicka, E., Maziarz, M., Piasecki, M. & Szpakowicz, S., 2012. *A Strategy of Mapping Polish WordNet onto Princeton WordNet*. Mumbai, India, In Proceedings of COLING 2012, pp. 1039–1048.
- (Sagot, 2017)** Sagot, B., 2017. Représentation de l'information sémantique lexicale : le modèle wordnet et son application au français. *Revue française de linguistique appliquée*, 22(1), pp. 131-146.
- (Sebastiani, 2002)** Sebastiani, F., 2002. Machine learning in automated text categorization. *ACM Computer Surv.*, 34(1), pp. 1-47.
- (Sim & Wright, 2005)** Sim, J. & Wright, C. C., 2005. The kappa statistic in reliability studies: use, interpretation, and sample size requirements. *Physical therapy*, 85(3), pp. 257-268.
- (Tan, et al., 2009)** Tan, S., Cheng, X., Wang, Y. & Xu, H., 2009. *Adapting Naive Bayes to Domain Adaptation for Sentiment Analysis*. Berlin, Heidelberg. *European Conference on Information Retrieval*. Springer, pp. 337–349.
- (Tang, et al., 2009)** Tang, H., Tan, S. & Cheng, X., 2009. A survey on sentiment detection of reviews. *Expert Systems with Applications*, 36(7), pp. 10760-10773.
- (Tarde, 1989)** Tarde, G., 1989. *L'opinion et la foule [original edition 1901]*. Paris, France: PUF.

- (Trainor, et al., 2014)** Trainor, K. J., Andzulis, J. M., Rapp, A. & Agnihotri, R., 2014. Social media technology usage and customer relationship performance: A capabilities-based examination of social CRM. *Journal of Business Research*, 67(6), pp. 1201-1208.
- (Tsytsarau & Palpanas, 2012)** Tsytsarau, M., & Palpanas, T. 2012. Survey on mining subjective data on the web. *Data Mining and Knowledge Discovery*, 24(3), 478-514.
- (Turney, 2002)** Turney, P., 2002. *Thumbs up or thumbs down? Semantic orientation applied to unsupervised classification of reviews*. Philadelphia, Pennsylvania., Association for Computational Linguistics, pp. 417-424.
- (Vossen, 1998)** Vossen, P. 1998 EuroWordNet: A Multilingual Database with Lexical Semantic Networks. *Computational Linguistics*, 25(4).
- (W Lent, 2008)** W Lent, R., 2008. Une conception sociale cognitive de l'orientation scolaire et professionnelle : considérations théoriques et pratiques. *L'orientation scolaire et professionnelle*, 37(1), pp. 57-90.
- (Wang, et al., 2011)** Wang, S. et al., 2011. A feature selection method based on improved fisher's discriminant ratio for text sentiment classification. *Expert Systems with Applications*, 38(7), pp. 8696-8702.
- (Wattarujeekrit, et al., 2004)** Wattarujeekrit, T., Shah, P. K. & Collier, N., 2004. PASBio: predicate-argument structures for event extraction in molecular biology. *BMC Bioinformatics*, 5(1), p. 155.
- (Weichselbraun, et al., 2013)** Weichselbraun, A., Gindl, S. & Scharl, A., 2013. Extracting and grounding contextualized sentiment lexicons, Knowledge - based Approaches to Concept - Level Sentiment Analysis. *IEEE Intelligent Systems*. 28(2), pp. 39-46.
- (Wen, et al., 2014)** Wen, M., Yang, D. & Rosé, C. P., 2014. *Sentiment analysis in MOOC discussion forums: What does it tell us*. Londre, RU, Educational data mining. pp.130-137.

- (Wilson, et al., 2005)** Wilson, T., Hoffmann, P., Somasundaran, S., Kessler, J., Wiebe, J. Choi, Y. Cardie, C. Riloff, E. Patwardhan, S., 2005. *Opinion Finder : A System for Subjectivity Analysis*. Vancouver HLT/EMNLP 2005 Interactive Demonstrations, pp. 34-35.
- (Xia, et al., 2015)** Xia, Y., Cambria, E. & Hussain, A., 2015. AspNet : aspect extraction by bootstrapping generalization and propagation using an aspect network. *Cognitive Computation*, 7(2), pp. 241-253.
- (Xuan, et al., 2012)** Xuan, H. N. T., Le, A. C. & Nguyen, L. M., 2012. *Linguistic features for subjectivity classification*. In 2012 International Conference on Asian Language Processing, IEEE, pp. 17-20.
- (Yi, et al., 2003)** Yi, J., Nasukawa, T., Bunescu, R. & Niblack, W., 2003. *Sentiment analyzer extracting sentiments about a given topic using natural language processing techniques*. Melbourne, USA, The IEEE International Conference on Data Mining (ICDM), pp. 427–432.
- (Zhang & Ferrari, 2014)** Zhang, L. & Ferrari, S., 2014. Intensité et polarité : un modèle opératoire articulant plusieurs travaux linguistiques. *Langue française*, 4(184), pp. 35-54.
- (Zhang, 2008)** Zhang, Z., 2008. Weighing stars: Aggregating online product reviews for intelligent e-commerce applications. *IEEE Intelligent Systems*, 23(5), pp. 42-49.

Références webographies

- (CSE, 2000) Conseil Supérieur de l'Éducation. (2000). Éducation et nouvelles technologies. Rapport annuel 1999-2000 sur l'état et les besoins de l'éducation. (Gouvernement du Québec). Consulté le 25/07/2019.
Disponible sur : <https://www.cse.gouv.qc.ca/fichiers/documents/publications/RapportsAnnuel/rapann00.pdf>
- (Devauchelle, 2007) Devauchelle, B., 2007. L'orientation scolaire et l'insertion professionnelle : Quel usage peut-on faire des Technologies de l'Information et de la Communication ? Consulté le 22/07/2019.
Disponible sur : <http://www.brunodevauchelle.com/orientation.htm>.
- (Ettouahria, 2015) Ettouahria, S., 2015. Éducation nationale : Journées portes ouvertes sur « l'orientation scolaire ». El Moudjahid, 14/04/2015. Consulté le 22/07/2019.
Disponible sur : <http://www.elmoudjahid.com/fr/actualites/76077>.
- (Feral, 2013) Feral, G., 2013. Les outils de L'orientation. Consulté le 13/03/2017.
Disponible sur : <http://www.pluriactivite.org/IMG/pdf/outils-orientation-centreinfo-janvier2013.pdf>.
- (Garceau, 2010) Garceau, O., 2010. Questionnaire quelles sont vos intelligences multiples ? Consulté le 25/02/2016.
Disponible sur : <http://www2.cegep-ste-Foy.qc.ca/freesite/fileadmin/groups/158/documents/Intelligences multiples.pdf>.
- (Imès, 2018) Imès, N., 2018. Orientation scolaire : La fin du bricolage ? *le soir d'Algérie*, Consulté le 25/06/2019.
Disponible sur : <https://www.lesoirdalgerie.com/actualites/la-fin-du-bricolage-16337>.
- (JO 04, 2008) JO 04, 2008. Loi no 08-04 du 23 janvier 2008 portant loi d'orientation sur l'éducation nationale. Consulté le 20/03/2019.
Disponible sur : <https://www.joradp.dz/FTP/JO-FRANCAIS/2008/F2008004.pdf?znjo=04>.

- (Lacoste, et al., 2005)** Lacoste, S., Esparbès-Pistre, S. & Tap, P., 2005. « L'orientation scolaire et professionnelle comme source de stress chez les collégiens et les lycéens », *L'orientation scolaire et professionnelle* 34/3.
Consulté le 30/06/2019.
Disponible sur : <https://journals.openedition.org/osp/617#quotation>
- (Leblanc, 2017)** Leblanc, C., 2017. Les logiciels d'aide à l'orientation (L.A.O.), 2017. *SAIO Rectorat de Versailles*.
Consulté le 30/04/2018.
Disponible sur : <https://docplayer.fr/23659610-Les-logiciels-d-aide-a-l-orientation-l-a-o.html>.
- (MEN, 2016)** MEN, 2016. Système Éducatif Algérien - Ministère de l'Éducation Nationale.
Consulté le 27/07/2019.
Disponible sur : <http://www.education.gov.dz/fr/systeme-educatif-algerien/>.
- (MESRS, 2018)** MESRS, 2018. GUIDE BACHELIER - Ministère de l'Enseignement Supérieur et de la Recherche Scientifique.
Consulté le 25/07/2019.
Disponible sur : <https://www.univ-boumerdes.dz/Portail%20nouveaux%20bacheliers%202019/GUIDE%202019.pdf>.
- (Santos, 2015)** Santos, F., 2015. Le kappa de Cohen : un outil de mesure de l'accord inter-juges sur des caractères qualitatifs.
Consulté le 06/02/2017.
Disponible sur : http://www.Paceau-Bordeaux1.en/IMG/pdf/Kappa_Cohen.pdf.
- (Sikali, 2009)** Sikali, L. A., 2009. Mémoire Online - Ouverture aux technologies de l'information et de la communication et performance professionnelle des conseillers d'orientation au Cameroun : le cas des lycées camerounais.
Consulté le 30/04/2018.
Disponible sur : https://www.memoireonline.com/08/10/3823/m_Ouverture-aux-technologies-de-linformation-et-de-la-communication-et-performance-professionnelle-de3.html.
- (UNESCO, 2001)** UNESCO, 2001. Définition de l'UNESCO en 1992 amendée par le Conseil de l'Éducation et de la Formation en 2001.
Consulté le 17/03/2017.
Disponible sur : <http://www.brunette.brucity.be/ferrer/>.

(Vasilescu, 2003)

Vasilescu, F., 2003. Désambiguïsation de corpus monolingues par des approches de type Lesk.
Consulté le 04/05/2017.
Disponible sur :
<http://rali.iro.umontreal.ca/rali/sites/default/files/publis/tina.pdf>.

(Vignoli, 2012)

Vignoli, E., 2012. « Introduction », L'orientation scolaire et professionnelle, 41/3.
Consulté le 30/06/2019.
Disponible sur : <https://journals.openedition.org/osp/3835>.

Annexes

Annexes 1 – Questionnaire pour recueillir les aptitudes de l'élève.

QUESTIONNAIRE QUELLES SONT VOS INTELLIGENCES MULTIPLES ?

Ce questionnaire a pour objet de vous aider à cerner les formes d'intelligence qui sont particulièrement développées chez vous, selon la théorie du psychologue américain Howard Gardner.

Pour chacun des énoncés, cochez la case appropriée selon l'échelle suivante :

1. Pas du tout comme moi
2. Un peu comme moi
3. Moyennement comme moi
4. Beaucoup comme moi
5. Tout à fait comme moi

Puis, faites le total de vos points pour chaque catégorie. Vous trouverez ainsi vos quatre formes d'intelligence principales.

INTELLIGENCE VERBALE/LINGUISTIQUE	1	2	3	4	5
1- J'aime les jeux de mots et autres blagues du genre.					
2- Je me sens à l'aise dans le monde de la langue et des mots et j'en tire un renforcement positif.					
3- J'aime faire des mots croisés et jouer à des jeux comme ScrabbleMC.					
4- Je me rappelle mot pour mot ce que les gens me disent.					
5- J'aime participer à des débats ou à des discussions.					
6- Je préfère les questions à développer que les questions à choix multiples.					
7- J'aime tenir un journal ou écrire des histoires ou des articles.					
8- J'aime beaucoup lire.					
Mon intelligence verbale/linguistique – Total					

INTELLIGENCE LOGIQUE/MATHÉMATIQUE	1	2	3	4	5
1- Je travaille mieux quand mon plan de travail est bien organisé.					
2- J'aime les mathématiques et les sciences.					
3- Je garde une liste des choses à faire.					
4- J'aime jouer à des jeux de réflexion ou qui font appel à la pensée logique comme JeopardyMC et ClueMC.					
5- J'aime savoir le pourquoi des choses et chercher des éclaircissements aux questions qui m'intéressent.					
6- Je travaille mieux avec un agenda ou un calendrier.					
7- Je saisis rapidement les relations de cause à effet.					
8- Mes estimations sont souvent bonnes.					
Mon intelligence logique/mathématique – Total					

INTELLIGENCE VISUELLE/SPATIALE	1	2	3	4	5
1- Je comprends les combinaisons de couleurs et vois quelles couleurs vont bien ensemble.					
2- J'aime faire des puzzles, des labyrinthes ou des jeux de patience.					
3- Je n'ai pas de mal à lire les cartes.					
4- J'ai un bon sens de l'orientation					
5- Dans les films, je m'intéresse particulièrement aux scènes et aux activités.					
6- Quand je dors, mes rêves me paraissent très réels.					
7- Je peux prévoir les mouvements et les conséquences dans un plan de jeu (par exemple au hockey, aux échecs).					
8- J'ai une mémoire surtout visuelle.					
Mon intelligence visuelle/spatiale – Total					

INTELLIGENCE INTERPERSONNELLE	1	2	3	4	5
1- Je travaille mieux quand j'interagis avec les autres.					
2- Je préfère les sports d'équipe aux sports individuels.					
3- La présence des autres me donne de l'énergie.					
4- Je préfère faire des choses en groupe plutôt qu'être seul.					
5- J'aime apprendre à connaître les autres cultures.					
6- J'ai l'habitude de parler de mes problèmes personnels avec mes amis.					
7- J'aime partager mes idées et mes sentiments avec d'autres.					
8- Je travaille mieux au sein d'un groupe où je peux discuter avec les autres de diverses questions.					
Mon intelligence interpersonnelle – Total					

INTELLIGENCE INTRAPERSONNELLE	1	2	3	4	5
1- J'aime ma solitude et avoir un monde à moi.					
2- J'ai quelques amis proches.					
3- J'ai des opinions bien arrêtées sur des questions controversées.					
4- Je travaille mieux quand je peux le faire à mon rythme.					
5- Je ne me laisse pas influencer facilement par les autres.					
6- Je comprends bien ce que je ressens et comment je réagis aux circonstances.					
7- Je soulève souvent des questions sur les valeurs et les croyances.					
8- Je me sais responsable de mon comportement.					
Mon intelligence intra-personnelle – Total					

INTELLIGENCE CORPORELLE/KINESTHÉSIQUE	1	2	3	4	5
1- J'aime bouger, taper du pied ou me remuer quand je suis assis.					
2- J'aime les sports aux sensations extrêmes (comme le kayak de mer, la planche à neige, le vélo de montagne).					
3- Je suis curieux sur le plan des sensations et j'aime toucher les objets pour en sentir la texture.					
4- J'ai une bonne coordination.					
5- J'aime le travail manuel.					
6- Je préfère participer physiquement plutôt que de rester assis et de regarder.					
7- Je comprends mieux les choses si je les fais (les touche, les bouge ou interagis avec elles).					
8- J'aime créer, faire des choses avec mes mains.					
Mon intelligence corporelle/kinesthésique – Total					

INTELLIGENCE MUSICALE	1	2	3	4	5
1- Je joue de la musique dans ma tête.					
2- Pour me rappeler des choses, je compose des comptines.					
3- Il m'est facile de suivre le rythme d'une musique.					
4- J'aime mettre une musique à des chansons ou des poèmes.					
5- Quand une musique joue, je bats la cadence.					
6- Je perçois les fausses notes.					
7- Je trouve facile de m'engager dans des activités musicales.					
8- Je suis fier de mes réalisations musicales.					
Mon intelligence musicale – Total					

INTELLIGENCE NATURALISTE	1	2	3	4	5
1- J'ai une collection de coquillages, de tasses, de pierres, de cartes de hockey, etc.					
2- Je remarque les similarités et les différences dans les arbres, les fleurs et autres objets de la nature.					
3- Je contribue activement à la protection de l'environnement.					
4- J'aime faire des fouilles et découvrir des objets et d'autres choses inhabituelles.					
5- Je préfère être à l'extérieur qu'à l'intérieur.					
6- J'aime faire des plantations et prendre soin d'un jardin.					
7- J'aime pêcher et suivre des animaux à la piste.					
8- La meilleure façon d'apprendre pour moi est d'aller en excursion, de voir des expositions sur la nature, etc.					
Mon intelligence naturaliste – Total					

Annexes 1-bis Les intelligences multiples et les Professions correspondantes

LES HUIT FORMES DE L'INTELLIGENCE MULTIPLE	QUELQUES PROFESSIONS À TITRE INDICATIF
1- Verbale/linguistique : capacité d'utiliser les mots efficacement, oralement ou par écrit. Aimer écrire, lire, parler, expliquer, comprendre les consignes, convaincre.	Enseignant, rédacteur, avocat, historien, animateur d'émissions-débats, journaliste, écrivain, politicien, interprète, secrétaire, bibliothécaire, comédien, poète, etc.
2- Logique/mathématique : capacité de déduction, d'utiliser les nombres efficacement et de bien raisonner. Aimer classifier, catégoriser, observer, résoudre des problèmes, calculer, quantifier.	Scientifique, comptable, économiste, médecin, mathématicien, programmeur-analyste, actuaire, informaticien, ingénieur, etc.
3- Visuelle/spatiale : capacité de bien percevoir le monde spatiovisuel, de se faire une image mentale juste des choses, de penser en trois dimensions. Aimer travailler avec des objets, construire, assembler, fabriquer, imaginer.	Artiste, architecte, pilote, guide de plein air, inventeur, designer, marin, ingénieur, sculpteur, chirurgien, topographe urbaniste, cartographe, peintre, joueur d'échecs, etc.
4- Interpersonnelle : capacité de comprendre les motivations, les sentiments des autres et d'entrer en relation avec eux; empathie, sensibilité aux expressions verbales et non verbales d'autrui. Aimer aider, enseigner, influencer.	Psychologue, acteur, travailleur social, conseiller touristique, enseignant, thérapeute, vendeur, guide spirituel, politicien, infirmier, sociologue, administrateur, médiateur, récréologue, etc.
5- Intrapersonnelle : capacité de se comprendre et d'utiliser cette compréhension pour mieux vivre. Aimer résoudre ses problèmes personnels, aider à résoudre ceux de l'humanité.	Philosophe, chercheur, conseillers, consultants, thérapeute, psychologue, théologien, planificateur, entrepreneur, etc.
6- Corporelle/kinesthésique : capacité d'apprendre par la pratique, de bien manipuler les objets, d'utiliser le corps pour exprimer ses émotions comme en danse ou en sport; habileté à manipuler et bonne coordination de tout le corps (visuomotrice, dextérité manuelle, motricité fine, etc.).	Inventeur, athlète, mécanicien, danseur, menuisier, acteur, mime, chirurgien, sculpteur, artisan, chorégraphe, massothérapeute, bijoutier, entraîneur, etc.
7- Musicale : capacité de comprendre la musique et de s'exprimer par elle, c'est-à-dire une aptitude à produire et à apprécier un rythme, une tonalité et un timbre; appréciation des formes d'expression musicale.	Musicothérapeute, présentateur de disques, critique musical, musicien, parolier, chef d'orchestre, ingénieur du son, etc.
8- Naturaliste : capacité de comprendre, de classer et d'expliquer la nature (plantes, animaux, phénomènes naturels); être capable de reconnaître, classifier les individus, les espèces; intérêt pour l'écologie et l'environnement.	Botaniste, vétérinaire, anthropologue, météorologue, physicien, biologiste, explorateur, géologue, etc.

Annexe 2 : Liste des communications et publications personnelles

- **Publications dans des revues internationales**

Lazib, S. ,Bouarab-DahmaniF. &Haddadi, L.(2019).Educational Guidance Process Integrating Opinion Mining. In *Computer Applications in Engineering Education*.⁹

(Article publié le 04/09/2019, versé pour la soutenance)

Haddadi, L., Bouarab-Dahmani, F., Guin, N., Berkane, T. &Lazib, S. Peer assessment and groups formation in massive open online courses. In *Computer Applications in Engineering Education*, 2018; 1–15, <https://doi.org/10.1002/cae.22005>.

- **Communications internationales**

Lazib, S., Bouarab-Dahmani, F., (2017, December). Integrating of opinion holder engagement in an educational guidance help system. In *International Conference on Mathematics and Information Technology (ICMIT), 2017* (pp. 247-254). IEEE. Adrar, Algeria.

Lazib, S., Bouarab-Dahmani, F., & Haddadi, L. (2016, July). Orientation process for secondary school pupils based on opinion analysis. *Proceeding of the 8th annual International Conference on Education and New Learning Technologies*(pp 3229- 3239), Barcelona, Spain.

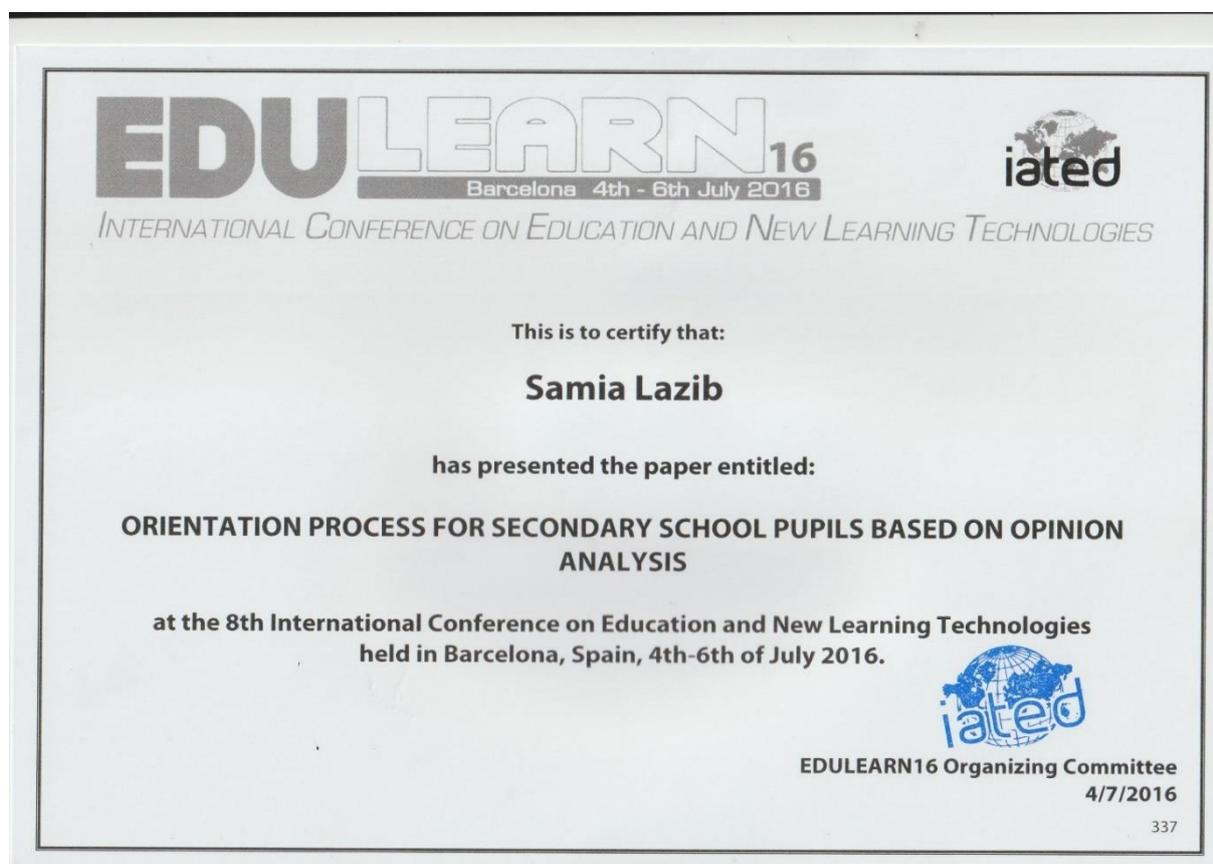
Lazib, S., Bouarab-Dahmani, F.,& Haddadi, L. (2016, November). Analyse d'opinions pour l'orientation scolaire.In *Proceeding of the Second International Symposium on Informatics and its Applications*, (p. 87),M'sila, Algeria..

Haddadi, L., Bouarab-Dahmani, F., Berkane, T., & Lazib, S. (2017, May). Peer assessment in MOOCs based on learners' profiles clustering. In *8th International Conference in Information Technology(ICIT), 2017 on* (pp. 532-536). IEEE, Amman, Jordan.

Haddadi, L., Bouarab-Dahmani, F. Tali, M. A., & Lazib, S. (2017, April). MOOC classification centered on learners' needs. In *Actes du colloque international IPAPE '2017* (p. 92), Djerba, Tunisia.

⁹<https://onlinelibrary.wiley.com/page/journal/10990542/homepage/productinformation.html>

Annexe 3 : Certifications de participation aux conférences



The Second International Symposium of Informatics and its Applications November 8-9, 2016

Computer Science Department, University of M'Sila, Algeria



Certificate of Participation

We herewith certify that

Samia Lazib

Has participated in the symposium with an Oral communication

Entitled: Analyse d'opinions pour l'orientation scolaire

Co-authors: Farida Bouarab-Dahmani and Lynda Haddadi

President of the Conference

Head of Department



The University Ahmed Draia of Adrar
 Faculty of Science and Technology
 Department of Mathematics and Computer Science



Attendance Certificate

We hereby notify that

Samia LAZIB

has attended the 2017 International Conference on Mathematics and Information Technology (ICMIT 2017) that was held at the University of Adrar, Algeria, on the 4th-5th of December 2017, and presented the paper entitled:

“Integrating of opinion holder engagement in a educational guidance help system”

Co-Authors: **Farida BOUARAB-DAHMANI**

