

Détection de polarité d'opinion dans les forums en langues arabe par fusion de plusieurs SVMs

Ziani Amel¹, Tlili Guiassa Yamina², Azizi Nabiha³

¹Département d'informatique, Université Badji Mokhtar Annaba (Algérie)

²Lri laboratory : laboratoire de recherche en informatique

³Labged laboratory : Laboratoire de gestion électronique des documents

Z_amel1911@live.fr, yamina.tlili@univ-annaba.org, nabiha.azizi@univ-annaba.org

RÉSUMÉ

Cet article décrit notre contribution sur la détection de polarité d'opinions en langue arabe par apprentissage supervisé. En effet le système proposé comprend trois phases: le prétraitement du corpus, l'extraction des caractéristiques et la classification. Pour la deuxième phase, nous utilisons vingt caractéristiques dont les principales sont l'émotivité, la réfexivité, l'adressage et la polarité. La phase de classification représente dans notre travail la combinaison des plusieurs classifieurs SVMs (Machine à Vecteur de Support) pour résoudre le problème multi classes. Nous avons donc analysés les deux stratégies de SVM multi classes qui sont : « un contre tous » et « un contre un » afin de comparer les résultats et améliorer la performance du système global.

ABSTRACT

Polarity Opinion Detection in Arabic Forums by Fusing Multiple SVMs

This article describes our contribution on the polarity's detection of opinions in Arabian language by supervised training. Indeed the proposed system consists of three phases: the pretreatment of the corpus, the extraction of the features and the classification. For the second phase, we use twenty features of which the main are emotionalism, the reflexivity, the adressage and the polarity. The phase of classification represents in our work the combination of the several SVMs (Support Vector Machine), to solve the multi class problem. We analyzed the two strategies of the SVMs multi class that are: "one against all" and "one against one" in order to compare the results and to improve the performance of the global system.

MOTS-CLÉS : Fouille d'opinions, apprentissage supervisé, Machine à Vecteur de Support (SVM), combinaison des classifieurs.

KEYWORDS : Opinion Mining, supervised training, Support Vector Machine (SVM), classifiers combination.

1 Introduction

De nos jours plusieurs applications et plateformes sur le web nous permettent de déposer des avis, de partager des sentiments et des opinions sur une variété de sujets. Vue l'importance de ces informations dans plusieurs domaines (politique, commercial ou individuel), il serait important de déterminer l'information subjective contenue dans les

textes.

Mais la détection automatique d'opinions et l'analyse des sentiments sont confrontées à des problèmes qui la distinguent de la recherche thématique traditionnelle, car le sentiment est exprimé de manière très variée et très subtile. La nécessité de traiter automatiquement les opinions se fait donc fortement ressentir.

La détection d'opinions est une tâche qui permet d'extraire les opinions d'un ensemble de documents pertinents pour un sujet donné. Elle est confrontée à des problèmes qui la distinguent de la recherche traditionnelle thématique dont les sujets sont souvent identifiés par des mots clés seulement. La classification du sentiment (polarité) est une sous-tâche de la détection d'opinions. Elle consiste de façon générale à déterminer si l'opinion du document sur le sujet est positive ou négative. Puisque nous avons besoin d'associer des notes à des textes, nous nous intéressons ici uniquement à la classification d'opinions. Deux grands types de méthodes sont utilisés pour cette tâche. Il y a tout d'abord les approches plutôt linguistiques qui consistent à répertorier le vocabulaire porteur d'opinions, puis à établir des règles de classification selon la présence ou l'absence des mots appartenant à ce vocabulaire. Il existe également les approches mettant en œuvre des outils issus du domaine de l'apprentissage automatique. Nous nous intéressons ici uniquement à celles de la deuxième famille, qui sur nos données se sont avérées nettement plus efficaces. Les méthodes utilisées dans ce cadre sont issues de la classification dite supervisée, où un classifieur est appris à l'aide d'exemples de données dont on connaît déjà la classe.

Ainsi, si aucune méthode de classification ne peut satisfaire entièrement aux exigences d'une application envisagée, l'utilisation simultanée de plusieurs méthodes en même temps peut éventuellement permettre d'en cumuler les avantages sans en cumuler les inconvénients. En effet, le comportement de chaque classifieur vis-à-vis de commentaire à classifier, est déterminé à partir des informations différentes représentant les caractéristiques extraites. L'exploitation des différents résultats générés par les classifieurs utilisant une des méthodes de combinaison de classifieurs, aboutit généralement à un commentaire bien classifié. Même si le classifieur est moins performant, la connaissance de son comportement apporte une certaine information utilisable à propos de la vraie classe pendant la combinaison. Donc, le but de la combinaison de classifieurs vise à réduire l'erreur et augmenter la fiabilité de la classification.

En effet, nous avons conçu et implémenté toutes les phases du processus de classification allant de la construction de la base de corpus jusqu'à la classification par combinaison d'expert, tout en passant par la phase d'extraction des caractéristiques. Concernant ces derniers, les caractéristiques que nous avons jugées importantes sont l'émotivité, la subjectivité, la réflexivité et l'adressage. Nous nous intéressons ensuite à la détection de polarité en déterminant si l'opinion est fortement positive, positive, fortement négative, négative ou neutre utilisant une des méthodes de classification. Toutefois, les méthodes les plus présentes dans la littérature, et qui semblent également être les plus performantes dans le domaine de fouille d'opinion sont les machines à vecteurs de supports. Comme notre système doit générer 5 classes, il est considéré comme un problème de classification multi classe, pour cela nous avons adopté les deux stratégies « un-contre-tous » et « un contre un » pour décider laquelle entre ces deux est meilleure dans le domaine de fouille d'opinions. Les SVMs sont jugés être très sensibles aux paramètres internes tels que la fonction noyau ; nous avons donc décidé d'analyser le changement de la fonction noyau en générant plusieurs classifieurs SVMs multi-classes associés chacun avec une fonction noyau différente.

Le résultat final sera la combinaison de ces classifieurs afin d'assurer la complémentarité existante entre les différentes fonctions. Pour cela on a utilisé les deux méthodes de fusion : Vote Majoritaire et Vote Pondéré pour générer les résultats finals.

Nous proposons dans cet article une approche de classification d'opinions dans les journaux arabes fondée sur la combinaison parallèle des classifieurs SVM (Support à Vecteur Machine) en utilisant les quatre fonctions (linéaire, polynomiale, gaussienne et tangente) afin de comparer les résultats. Dans un premier temps, nous introduisons le domaine de la fouille d'opinions avec un état de l'art relatif à notre problème. Ensuite, nous présentons les difficultés de la langue arabe et le processus général de la classification d'opinions. Puis, nous décrivons les différentes phases de l'approche proposée, incluant l'ensemble des caractéristiques adoptées. Après nous présentons et discutons les résultats obtenus à travers les expérimentations que nous avons menés. Nous achevons cet article par une conclusion et un ensemble de perspectives.

2 Etat de l'art

Le terme « fouille d'opinions » est utilisé pour évoquer le traitement automatique des opinions, des sentiments et de la subjectivité dans les textes. Ce domaine est connu sous les noms de : opinion mining (Pang et Lee, 2008), sentiment analysis (Liu, 2010), ou encore subjectivity analysis et est souvent associé à un problème de classification sur des textes évaluatifs comme ceux disponibles sur Amazon ou Epinions. Afin de décider de l'orientation d'un document (Turney, 2002), (Wilson et al., 2004) ou de la valeur positive/négative/neutre d'une opinion dans un document (Hatzivassiloglou et McKeown, 1997), (Yu et Hatzivassiloglou, 2003), (Kim et Hovy, 2004).

Le travail de Maurel et Dini en 2009 été caractérisé par une utilisation mixte d'une technologie symbolique fondée sur des règles et d'une technologie statistique reposant sur l'extraction d'une ontologie du domaine, approche dans laquelle la méthode symbolique a un poids plus important (Dini, 2002), (Dini et Mazzini, 2002), (Maurel et al., 2007), (Maurel et al., 2008), (Bosca et Dini, 2009).

Des travaux allant au-delà ont mis l'accent sur la force d'une opinion exprimée où chaque proposition dans une phrase peut avoir un fond neutre, faible, moyen ou élevé (Wilson et al., 2004). Des catégories grammaticales ont été utilisées pour l'analyse de sentiments dans (Bethard et al., 2004) où des syntagmes adjectivaux comme trop riche ont été utilisés afin d'extraire des opinions véhiculant des sentiments. (Bethard et al., 2004) utilisent une évaluation basée sur la somme des scores des adjectifs et des adverbes classés manuellement, tandis que (Chklovski, 2006) utilise des méthodes fondées sur un modèle pour représenter des expressions adverbiales de degré telles que parfois, beaucoup, assez ou très fort.

En 2002, Turney, Pang et coll. encouragent la recherche dans le domaine de sentiment analysis en classant des critiques de cinéma. (Ding et al., 2007) analysent les cooccurrences de mots à l'intérieur d'une phrase puis les cooccurrences entre les phrases. Ils combinent des règles issues des deux échelles pour obtenir un meilleur taux de bonne classification en classification de sentiments. Dans le même genre. (Wilson et al., 2004) ajoutent à la classification selon la polarité, la force de l'opinion exprimée.

En ce qui concerne les études sur la fouille d'opinions en langue arabe, elles n'en sont qu'à leur début, présentons ici quelques travaux : (Almas et Ahmed, 2007), (Abbasi et al., 2008), (Elhawary et Elfeky, 2010), (El halees, 2010), (Rushdi et al., 2011) et (Montassir et al., 2012).

3 La langue arabe

Par ses propriétés morphologiques et syntaxiques la langue arabe est considérée comme une langue difficile à maîtriser dans le domaine du traitement automatique de la langue. L'arabe doit sa formidable expansion à partir du 7ème siècle grâce à la propagation de l'islam et la diffusion du Coran. Les recherches pour le traitement automatique de l'arabe ont débuté vers les années 1970. Les premiers travaux concernaient notamment les lexiques et la morphologie.

Avec la diffusion de la langue arabe sur le Web et la disponibilité des moyens de manipulation de textes arabes, les travaux de recherche ont abordé des problématiques plus variées comme la syntaxe, la traduction automatique, l'indexation automatique des documents, la recherche d'information, le résumé automatique, la fouille d'opinions etc.

Un des aspects complexes de la langue arabe est l'absence des voyelles dans le texte, qui risque de générer une certaine ambiguïté à deux niveaux :

- Sens du mot
- Difficulté à identifier sa fonction dans la phrase, (différencier entre le sujet et le complément).

Ceci peut influencer les fréquences des mots étant donné qu'elles sont calculées après la détection de la racine ou la lemmatisation des mots qui est basée sur la suppression de préfixes et suffixes.

Bien que la lemmatisation soit difficile pour les langues avec des morphologies complexes comme l'arabe, elle est particulièrement importante et utile en particulier dans les systèmes de recherche d'information. Il est suffisant de regrouper les mots qui se ressemblent le plus sans pour autant connaître la racine exacte.

Contrairement à l'anglais, la langue arabe possède un système dérivationnel très riche, et c'est dans cette caractéristique que réside la difficulté de traiter cette dernière.

4 Processus général de la classification d'opinions

Puisque nous avons besoin d'associer des notes à des textes, nous nous intéressons ici uniquement à la classification d'opinions. Deux grands types de méthodes sont utilisés pour cette tâche. Il y a tout d'abord les approches plutôt linguistiques qui consistent à répertorier le vocabulaire porteur d'opinions, puis à établir des règles de classification selon la présence ou l'absence des mots appartenant à ce vocabulaire.

Il existe également les approches mettant en œuvre des outils issus du domaine de l'apprentissage automatique. Dans notre cas, un ou plusieurs classifieurs sont utilisés pour apprendre la polarité des opinions traités, puis ils seront responsables de classer les opinions inconnues.

Nous nous intéressons dans ce travail à l'application de la deuxième approche avec un module d'extraction des caractéristiques pertinentes. Les méthodes utilisées dans ce cadre sont issues de la classification dite supervisée (ou apprentissage supervisé), où un classifieur

est appris à l’aide d’exemples de données (ici de textes de commentaires) dont on connaît déjà la classe (ici la polarité). Les mots des textes sont alors généralement considérés comme des données indépendantes et équivalentes les unes aux autres, leur sémantique n’étant pas explicitement prise en compte. On peut donner une définition un peu plus formelle du problème de la classification supervisée comme suit :

Définition : soit X un ensemble de données, Y un ensemble d’étiquettes (ou classes) et D un ensemble des représentations des données. Soit $d : X \rightarrow D$, une fonction connue qui associe à chaque donnée $x \in X$ une représentation $d(x) \in D$ et $S \subset D \times Y$ un ensemble de données étiquetées $(d(x), y)$. La classification supervisée consiste à construire en s’appuyant sur l’ensemble S un classifieur, c’est-à-dire une fonction de $D \rightarrow Y$ qui permette de prédire la classe de toute nouvelle donnée $x \in X$, représentée par $d(x) \in D$.

D’après cette définition, trois éléments distincts entrent en jeu dans la classification supervisée :

- Les étiquettes ou classes de prédiction (l’ensemble Y) ;
- Les exemples de données étiquetées, qui constituent le corpus d’apprentissage (l’ensemble S) ;
- Le classifieur ou prédicteur.

De plus, des prétraitements peuvent être appliqués sur les données avant la tâche de classification dans le but d’améliorer ses performances, que ce soit en termes de résultats ou de temps de calcul. (Poirier et al, 2011)

5 Structure générale du système AROPOL (ARabic Opinions POLarity) basée sur la classification supervisée

Nous pouvons résumer notre approche de classification d’opinions par combinaison des SVM dans les journaux en langue arabe par le schéma suivant :

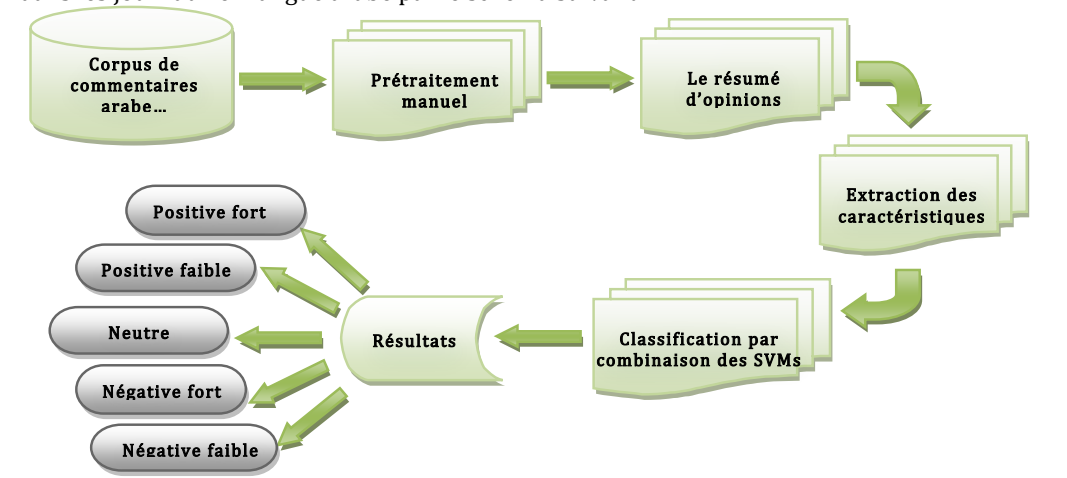


FIGURE 1 –Le Processus de classification d’opinions par l’approche AROPOL

5.1 Le corpus d’apprentissage

La classification supervisée nécessite des exemples (données étiquetées) afin de construire le «corpus d’apprentissage». Ce corpus ayant un impact direct sur l’apprentissage des règles, et par conséquent sur la classification, il est nécessaire que les exemples soient représentatifs de l’ensemble des données.

Cette hypothèse est généralement difficile à vérifier. En classification d’opinions, ou plus généralement en classification de textes, les corpus étudiés sont assez restreints car l’étiquetage des exemples est souvent effectué à la main. Ceci entraîne un coût élevé et ne permet donc pas l’obtention d’un gros corpus d’apprentissage. (Poirier et al, 2011)

Pendant notre travail, nous avons utilisés un corpus de cent cinquante commentaires sur des articles, recueillis à partir des journaux arabes algériens disponibles sur le net (Eshorouk الشروق, Akher saa أخرساعةetc.). L’ensemble touchait plusieurs thèmes différents (économiques, politiques, etc.). Notre but était d’effectuer une classification supervisée pour pouvoir déterminer la polarité des commentaires: positive fort, positive faible, neutre, négative fort ou négative faible. C’est pour cela on a appliqué un ensemble des prétraitements manuels sur ce corpus.

Exemple de commentaire :

Nous présentons un exemple de chaque type d’opinion que nous devons classifier.

	Polarité	Les commentaires en français	Les commentaires en arabe
1	Positif faible	Merci pour ce que tu as écrit Mr amine zaoui celui qui dirige ministère de la culture en Algérie est celui que le dépense sans dignité.	شكراعلماكتب سيد الأمين الزاوي إن وزارة الثقافة بالجزائر يستلمها من يصر فالمال العام بغير وجه حق.
2	Positif fort	Félicitation à nous tous, vous avez fait le bon choix. La journaliste leila bouzidi est réellement compétente et d’une personnalité professionnelle lui permettant de bien gérer sa carrière que dieu soit avec elle.	مبروك علينا وعليكم لقد أحسنت الاختيار. الصحفية القديرة ليلي بوزيدي حقا هي متمكنة وذات شخصية مهنية متحكمة في إدارة مهنتها أعانها الله ووفقها.
4	Négatif faible	Franchement, c’est Article non professionnel de la part d’un expert connu voulant la réussite de l’équipe national.	مقال غير احترافي صراحة من إعلامي كبير يحرص على نجاح المنتخب وقاسي كذلك.
6	Négatif fort	Franchement, je n’ai pas aimé le style de l’écrivain, et aussi son point de vus envers ce sujet et j’ai haï ses moqueries pour l’art.	صراحة لم أحب أبدا أسلوب الكاتب ولا رأيه في الموضوع وكرهت استهزاءه بالفن.
8	Neutre	Non, c’est son point de vue personnel	كلا انه رأيه الشخصي.

TABLE 1 – Exemple de commentaires

5.2 Extraction des caractéristiques

Nous commençons la phase d’extraction de caractéristiques par la construction manuelle des tables des marqueurs d’opinions du corpus utilisé définies comme suit :

❖ **Marqueur :**

La table marqueur contient tous les prédicats, les adjectifs et les adverbes construits à partir du corpus avec leurs polarités et intensités.

Exemples:

- **Prédicat:** أحب aime, كره détester, ظن penser.
- **Adjectif:** جميلة bien fait, رائعة magnifique, ركيكة lâche.
- **Adverbe:** غنية riche, مضجرة fatigante, مفيدة intéressante.
- ❖ **Intensité:** كثيرا beaucoup, جدا très, مئة بالمئة cent pour cent.
- ❖ **Négation:** لا (non, ni, pas), لن لم ليس pas.
- ❖ **Adressage:** أنت tu, يا سيدي الكاتب Mr l’écrivain.
- ❖ **Réflexivité:** أنا moi, رأيي mon opinion.

5.3 Les caractéristiques de représentation d’opinion

Nous avons aussi adoptés un ensemble des caractéristiques inspirées de celle des travaux de (Boughanem et al, 2010) et qui ont montrés leur efficacité dans la représentation d’un commentaire.

➤ **Adverbe**

Total-adverbes : tot(adv) = Nombre total des adverbes du document

Moy-adverbes = tot(adv) / (tot(adj) + tot(adv) + tot(pred))(1)

➤ **Adjectif**

Total-adjectives : tot(adj) = Nombre total des adjectifs du document

Moy-adjectives = tot(adj) / (tot(adj) + tot(adv) + tot(pred))(2)

➤ **Émotivité**

Les chercheurs ont exploité la présence des adverbes et des adjectifs dans un document comme un indicateur qui permet de déterminer les opinions. Nous calculons l’émotivité d’un document en comptant le nombre des adverbes et des adjectifs dans ce document.

$$Emot(d) = \frac{|\{\omega \in d \setminus type(\omega) \in \{adjectif, adverbe\}\}|}{|\{\omega \in d \setminus type(\omega) \in \{predicat\}\}|} \quad (3)$$

➤ **Adressage**

La plupart des phrases trouvées dans les blogs et les forums contiennent des mots comme suit « أنت toi, vousمَنْ, toi-même نفسك, vous-même أنفسكم, هو il, elle هي, ils هم, lui-même نفسه, elle-même نفسها » car les utilisateurs écrivent des commentaires sur un sujet, en s’adressant aux autres. De ce fait l’utilisation de ces pronoms d’adressage est très fréquente. Par conséquent, nous considérons que la composante d’adressage dans le cadre de notre détection d’opinions, est comme suit :

$$Add(d) = \frac{|\{\omega \cap \omega' \setminus \omega \in d, \omega' \in A\}|}{|A| + |R|} \quad (4)$$

- $|\{\omega \cap \omega' \backslash \omega \in d, \omega' \in A\}|$ Représente le nombre d’occurrences des termes d’adressages dans le document d qui appartiennent à la liste d’adressage $|A|$ que nous avons préparé.
- $|A|$ est égal au nombre total de pronoms dans la liste d’adressage A .
- $|R|$ est le nombre total de pronoms dans la liste de réflexivité.

➤ **Réflexivité**

Les blogueurs utilisent beaucoup de pronoms réflexifs comme « أنا شخصيا، أنا » « moi, moi-même » lors de l’écriture. Par exemple, l’utilisation de « ي » dans « رأيي » « Je pense que », « من وجهة نظري », « mon point de vu est que », etc. Toutes ces phrases font référence à une opinion d’opinion, et par conséquent, nous incluons la mesure de la réflexivité. L’idée est que tout document avec un plus grand nombre de ces mots sera plus subjectif par rapport à celui qui a moins de nombre de ces mots. Cette mesure est exprimée par la réflexivité $Ref(d)$.

$$Ref(d) = \frac{|\{\omega \cap \omega' \backslash \omega \in d, \omega' \in R\}|}{|R| + |A|} \tag{5}$$

- $|\{\omega \cap \omega' \backslash \omega \in d, \omega' \in R\}|$ est le nombre de pronoms réfléchis dans le document d qui appartiennent à la liste de réflexivité R que nous avons construit.
- $|R|$ est le nombre total de pronoms dans la liste de réflexivité R .
- $|A|$ est l’ensemble de nombre total de pronoms dans la liste d’adressage.

En plus il y a un ensemble de caractéristiques que nous avons proposées qui sont décrites dans le tableau suivant :

Phrase	Nombre de phrases	Σ phrases
Positif fort	Nombre des mots positifs forts	Σ mot (posFo)
Positif faible	Nombre des mots positifs faibles	Σ mot (posFa)
Négatif fort	Nombre de mots négatifs forts	Σ mot (negFo)
Négatif faible	Nombre de mots négatifs faibles	Σ mot (negFa)
Neutre	Nombre de mots neutres	Σ mot (ntr)
Prédicat	Total-prédictats	Nbr(pred)=Nombre total des prédicats du document
	Moy- prédictats	$Nbr(pred)/(\text{tot}(\text{adj}) + \text{tot}(\text{adv}) + \text{tot}(\text{pred}))$
Polarité	Somme polarité	$SomPolarite = \Sigma \text{mot}(\text{posFo}) + \Sigma \text{mot}(\text{posFa}) + \Sigma \text{mot}(\text{negFo}) + \Sigma \text{mot}(\text{negFa}) + \Sigma \text{mot}(\text{ntr})$
	Mots positifs forts	$\Sigma \text{mot}(\text{posFo}) / SomPolarite$
	Mots positifs faibles	$\Sigma \text{mot}(\text{posFa}) / SomPolarite$
	Mots négatifs forts	$\Sigma \text{mot}(\text{negFo}) / SomPolarite$
	Mots négatifs faibles	$\Sigma \text{mot}(\text{negFa}) / SomPolarite$
	Mots neutres	$\Sigma \text{mot}(\text{ntr}) / SomPolarite$

TABLE 2 – Les mesures proposées

5.4 La classification

Ils existent plusieurs méthodes de classification supervisée et beaucoup d'entre elles ont été testées pour la classification d'opinions. On peut citer les arbres de décision, les réseaux de neurones, la régression logistique, les règles de décision ainsi que des méthodes combinant différents classifieurs comme les systèmes de votes ou les algorithmes de Boosting. Toutefois, les méthodes les plus présentes dans la littérature, et qui semblent également être les plus performantes sur les textes, sont les machines à support de vecteurs (Pang et Lee, 2004), (Wilson et al, 2004), (Nigam et Hurst, 2006), (Généreux et Santini, 2007), (Trinh, 2007), (Crestan et al, 2007), (Planté et al, 2008), (Kaiser et al, 2010), (Poirier et al, 2011), (Abbasi et al, 2008), (Montassir et al, 2012).

Au premier temps, les SVM sont conçus pour traiter le problème de classification binaire. Leur extension aux problèmes de classification multi-classes demeure un sujet de recherche très actif. En général, les méthodes multi-classes proposées dans la littérature suivent l'une des deux approches suivantes :

- ✓ La combinaison de plusieurs classifieurs binaires classiques.
- ✓ La conception d'un seul classifieur SVM en résolvant un seul problème d'optimisation.

Dans notre travail nous avons adopté ici les deux stratégies de la première approche « un-contre-tous » et « un contre un » pour décider laquelle entre ces deux est meilleure pour la classification d'opinions multi classes.

5.5 La combinaison des classifieurs

Plutôt que de chercher à optimiser un seul classifieur en choisissant les meilleures caractéristiques pour un problème donné, les chercheurs ont trouvé plus intéressant de combiner plusieurs méthodes de classification.

La multiplication des travaux sur la combinaison de classifieurs a entraîné la mise au point de nombreux schémas traitant les données de manières différentes. Nous utilisons dans cette partie la combinaison parallèle. L'approche parallèle laisse les différents classifieurs opérer indépendamment les uns des autres puis fusionne leurs réponses respectives.

Notre objectif est d'analyser le comportement d'un système combinant plusieurs classifieurs afin d'augmenter les performances de classification. Donc, pour construire un système multi classifieurs, il y a une méthode qui se base sur l'utilisation d'un même classifieur en modifiant à chaque fois ses paramètres internes ; ce qui va générer des classifieurs différents. Cette différence sera traduite par une complémentarité durant le processus de classification d'un commentaire inconnu.

En effet, Vu que chaque classifieurs SVM selon la fonction noyau génère un résultat différent, et vu que dans la littérature, on n'a pas pu prouver la supériorité d'une fonction par rapport aux autres dans le cas général, on a utilisé quatre classifieurs, chacun d'eux à une fonction noyau différente qui sont les suivantes (linéaire, polynomiale, gaussienne et tangente). En combinant ces classifieurs, chacun d'eux va assurer un certain niveau de complémentarité au système global.

Le système généré est construit de 4 classifieurs, où chacun d'eux offre 5 sorties. Afin de fusionner les résultats des 4 classifieurs, nous avons appliqué deux méthodes de

combinaison les plus connus (le vote majoritaire et le vote pondéré).

6 L'évaluation du système

Au niveau du test on applique notre système de classification sur un nouvel ensemble différent de celui de l'apprentissage, contient 30 commentaires étiquetés manuellement.

Pour évaluer les résultats obtenus il faut calculer le taux de chaque classe pour chaque stratégie (un contre tous et un contre un) et pour chaque méthode de fusion (le vote et le vote pondéré).

6.1 La stratégie un contre tous

	Linéaire	Polynomiale	Gaussienne	Tangente	Vote	Vote pondéré
Neutre	0.19	0.27	0.31	0.09	0.30	0.39
Positif fort	0.69	0.75	0.82	0.20	0.81	0.91
Positif faible	0.40	0.49	0.55	0.11	0.55	0.59
Négatif fort	0.89	0.90	1.00	0.31	1.00	1.00
Négatif faible	0.45	0.51	0.59	0.20	0.59	0.60

TABLE 3 – Variation du taux de classification de chaque classe pour les quatre fonctions noyaux et avec les deux méthodes de fusion

A partir du tableau ci-dessus on peut constater que les classifieurs SVMs conçus avec les fonctions noyaux polynomiale et gaussienne donne le meilleur taux de classification par comparaison aux autres SVMs.

On constate notamment que la classe de type « négatif fort » offre le meilleur taux de classification ; en revanche la classe neutre est celle ayant le taux le plus faible. Cela est dû à notre avis aux poids des marqueurs représentant ces deux classes.

On remarque que les résultats de la combinaison (soit pour la méthode de vote ou de vote pondéré) sont meilleurs qu'en utilisant un seul classifieur; en effet, malgré que la fonction tangente génère un taux de reconnaissance très faible mais sa présence dans le processus de combinaison enrichie la classification; d'où l'intérêt majeur de la combinaison de classifieurs.

Concernant les méthodes de combinaison, on a testé les deux méthodes les plus utilisées pour les classifieurs de type classe qui sont « vote majoritaire et vote pondéré » afin de maintenir celle qui offre le taux de classification le plus fort. Dans notre application, on constate que le taux du vote pondéré est supérieur de celui de vote majoritaire.

6.2 La stratégie un contre un

	Linéaire	Polynomiale	Gaussienne	Tangente	Vote	Vote pondéré
Neutre	0.50	0.57	0.61	0.29	0.62	0.65
Positif fort	0.81	0.91	1.00	0.31	1.00	1.00
Positif faible	0.88	0.88	0.88	1.00	0.88	0.95
Négatif fort	0.88	0.98	1.00	0.23	1.00	1.00
Négatif faible	0.49	0.50	0.62	0.12	0.62	0.68

TABLE 4 – Variation du taux de classification de chaque classe pour les quatre fonctions noyaux et avec les deux méthodes de fusion

D'après le tableau ci dessus on constate que la fonction gaussienne donne de bon taux de classification pour les quatre classes (neutre, positif faible, négatif fort et négatif faible), par contre la fonction tangente donne le meilleur taux pour la classe positif faible.

En générale les deux méthodes de fusion produisent des bons résultats, mais c'est la méthode de vote pondéré qui a le meilleur taux pour toutes les classes.

Nous concluons alors, que pour améliorer le taux de classification, il est préférable d'utiliser la méthode de vote pondéré.

Après la mise en œuvre des stratégies, il s'avère que les résultats obtenus par la stratégie « un contre un » sont meilleurs que ceux de la stratégie « un contre tous ». En effet comme le montre le tableau précédant (Table 4), le taux de toutes les classes est augmenté pour tous les fonctions, ce qui signifie que la classification est excellente.

7 Conclusion

Au terme de ce travail, nous procédons dans les lignes qui se suivent à un récapitulatif du travail effectué. Rappelons que notre travail consistait essentiellement à développer une application qui permet la détection de polarité d'opinions dans les forums en langue arabe par combinaison de plusieurs SVM. Ce système a pour rôle d'extraire les caractéristiques représentant les commentaires du corpus et de les classer en catégories par la coopération de plusieurs classifieurs SVMs.

Donc notre système opère en trois phases, la première consiste à la construction et le prétraitement manuel du corpus recueillis à partir des journaux arabes algériens. La seconde phase est une extraction des caractéristiques afin de détecter et représenter les commentaires. Le choix de ces caractéristiques est fait par une recherche approfondie sur les plus importantes caractéristiques pouvant représenter de mieux le commentaire tout en évitant la redondance et la confusion des données d'entrée, et nous avons conclu qu'il est difficile de faire le choix des bons caractéristiques. Enfin la troisième phase est la réalisation du module de classification et dans le but de bénéficier des avantages des systèmes multi-classifieurs, on a proposé un système combinant 04 classifieurs SVMs multi-classes représentant chacun par une fonction noyau différente pour les deux stratégies (un contre tous et un contre un). Les résultats issus de la combinaison sont très encourageants et nous ont permis de mieux s'investir dans cet axe tout en analysant le rôle que peut jouer chaque paramètre des caractéristiques.

Les tests sur les commentaires en langue arabe avec les deux stratégies « un contre tous » et « un contre un », nous ont permis de prouver que la stratégie un contre un donne de meilleurs résultats avec les commentaires des journaux en langue arabe.

Cette expérience, nous a permis de faire une première exploration du vaste domaine qui est l'opinion mining, et nous incite à aller plus loin dans ce domaine dans le cadre de travaux futurs, car l'opinion mining représente un axe de recherche très prometteur et très passionnant, ce qui explique l'intérêt recrudescant que portent les chercheurs pour ce domaine.

Comme perspective par rapport à ce travail, il serait intéressant de généraliser notre système avec n'importe quelle base de commentaires.

Une autre perspective consiste à utiliser d'autres caractéristiques afin d'enrichir le vecteur des caractéristiques pour une meilleure représentation des commentaires, et optez pour l'ajout d'un module de sélection de caractéristiques.

Il serait aussi intéressant de combiner plusieurs algorithmes de classification tels que les SVM (machine à vecteur de support) et les réseaux bayesiens, pour pouvoir évaluer et améliorer les performances.

Références

ABBASI, A., CHEN, H. et SALEM, A. (2008). Sentiment analysis in multiple languages: Feature selection for opinion classification in Web forums. *ACM Transactions on Information Systems Volume 26 Issue 3, Jun. 2008*.

ALMAS, Y. et AHMAD, K. (2007). A Note on Extracting 'Sentiments' in Financial News in English, Arabic & Urdu. *In Proceedings of the 2nd Workshop on Computational Approaches to Arabic Script-based Languages Linguistic Institute, Stanford, California, USA, July 21-22, 2007*.

Bethard, S., Yu, H., Thornton, A., Hatzivassiloglou, V. et Jurafsky, D. (2004). Automatic extraction of opinion propositions and their holders. *Actes d'AAAI'04, 2004*.

BOSCA, A. et DINI, L. (2009). User Logs as a Means to Enrich and Refine Translation Dictionaries. *Workshop of Cross-Language Evaluation Forum. CLEF (1) 2009*.

BOUGHANEM, M. et BELBACHIR, F. (2010). Expérimentation de fonctions pour la détection d'opinions dans les blogs (mémoire de Master).

CHKLOVSKI, T. (2006). Deriving quantitative overviews of free text assessments on the web. *Actes d'IUI'06*.

CRESTAN, E. et ACUNA-AGOST, R. (2007). Quel modèle pour détecter une opinion ? Trois propositions pour généraliser l'extraction d'une idée dans un corpus. *In Actes de l'atelier de clôture du 3ème défi fouille de textes, Grenoble, France*.

DINI, L. (2002). Compréhension multilingue et extraction de l'information. *In, F. Segond (ed.), Multilinguisme et traitement de l'information (Traité des sciences et techniques de l'information), Editions Hermes Science, 2002*.

DINI, L. et MAZZINI, G. (2002). Opinion classification through information extraction. *In A. Zanas, C. A. Brebbia, N. F. F. Ebecken, P. Melli (Eds), Data Mining III, WIT Press, 2002*.

EL-HALEES, A. (2011). Arabic Opinion Mining Using Combined Classification Approach. *In Proceeding The International Arab Conference On Information Technology, Azraq, Jordan 2011*.

ELHAWARY, M. et ELFEKY, M. (2010). Mining Arabic Business Reviews. *IEEE International Conference on Data Mining Workshops. 2010*.

GÉNÉREUX, M. et SANTINI, M. (2007). Défi : Classification de textes français subjectifs. *Actes de l'atelier de clôture du 3e Défi Fouille de Textes, AFIA, Grenoble, France, 2007*.

HATZIVASSILOGLLOU, V. et MCKEOWN, K. R. (1997). Predicting the semantic orientation of adjectives. *Actes d'ACL'97, 1997*.

KIM, S.-M. et HOVY, E. (2004). Determining the sentiment of opinions. *Actes de COLING'04, 2004*.

- KAISER, C., KROCKEL, J. et BODENDORF, F. (2010). Swarm Intelligence for Analyzing Opinions in Online Communities.
- LIU, B. (2010). Sentiment Analysis. *Invited talk at the 5th Annual Text Analytics Summit*.
- MAUREL, S., CURTONI, P. et DINI, L. (2007). L'analyse des sentiments dans les forums.
- MAUREL, S., CURTONI, P. et DINI, L. (2008). L'analyse des sentiments dans les forums. *Actes de l'atelier FOuille des Données d'OPinions, 2008*.
- MILGRAM, J. (2007). Contribution à l'intégration des machines à vecteurs de support au sein des systèmes de reconnaissance de formes: application à la lecture automatique de l'écriture manuscrite.
- Mountassir, A., Benbrahim, H. et Berrada, I. (2012). A cross-study of Sentiment Classification on Arabic corpora. *In Research and Development in Intelligent Systems XXIX. Springer London, 2012*.
- NIGAM, K. ET HURST, M. (2006). Towards a Robust Metric of Polarity, Computing Attitude and Affect in Text: Theory and Applications. *Springer, Dordrecht, The Netherlands, 2006*.
- PANGB, et Lee L. (2004). A Sentimental Education: Sentiment Analysis Using Subjectivity Summarization Based on Minimum Cuts.
- PANG, B. et LEE, L. (2008). Opinion Mining and Sentiment Analysis.
- PLANTIÉ, M., ROCHE, M., DRAY, G. et PONCELET, P. (2008). Is a Voting Approach Accurate for Opinion Mining ?.
- POIRIER, D., FESSANT, F. et TELLIER, I. (2011). De la classification d'opinions à la recommandation : l'apport des textes communautaires. *In TALN 2011 (Traitement automatique des langues naturelles)*.
- RUSHDI, S., MOHAMMED, M., MARTÍN-VALDIVIA, T., UREÑA-LÓPEZ, A. L., et JOSÉ M. (2011). OCA: Opinion corpus for Arabic. *Journal of the American Society for Information Science and Technology* 62, no. 10 (2011): 2045-2054.
- TRINH, A.-P. (2007). Classification de texte et estimation probabiliste par machine à vecteur de support. *In Actes de l'atelier de clôture du 3ème défi fouille de textes, Grenoble, France*.
- TURNER, P. (2002). Thumbs Up or Thumbs Down? Semantic Orientation Applied to Unsupervised Classification of Reviews. *Proceedings of the Association for Computational Linguistics, ACL'02, 2002*.
- WILSON, T. WIEBE, J. et HWA, R. (2004). Just How Mad Are You? Finding Strong and Weak Opinion Clauses.
- YU, H. et HATZIVASSILOGLU, V. (2003). Towards answering opinion questions: Separating facts from opinions and identifying the polarity of opinion sentences. *Actes d'EMNLP'03, 2003*.