

THE IMPORTANCE OF NEUTRAL CLASS IN SENTIMENT ANALYSIS OF ARABIC TWEETS

Hamed AL-Rubaiee¹, Renxi Qiu², and Dayou Li³

Department of Computer Science and Technology
University of Bedfordshire
Bedfordshire, United Kingdom

ABSTRACT

Product reviews are becoming increasingly useful. In this paper, Twitter has been chosen as a platform for opinion mining in trading strategy with Mubasher products, which is a leading stock analysis software provider in the Gulf region. This experiment proposes a model for sentiment analysis of Saudi Arabic (standard and Arabian Gulf dialect) tweets to extract feedback from Mubasher products. A hybrid of natural language processing and machine learning approaches on building models are used to classify tweets according to their sentiment polarity into one of the classes positive, negative and neutral. In addition, Regarding to the comparison between SVM and Bayesian method, we have split the data into two independent subsets from different periods and the experiments were carried out for each subsets respectively in order to distinction between positive and negative examples by using neutral training examples in learning facilitates. Similar result has been given.

KEYWORDS

Sentiment analysis, Mubasher, Twitter, Saudi Arabia, machine learning, neutral class, Pre-Processing

1. INTRODUCTION

Social opinion has been analysed using sentiment analysis and some studies show that sentiment analysis of news, documents, quarterly reports, and blogs can be used as part of trading strategies [1-3]. Sentiment analysis has become the heart of social media research and many studies [4-6] have been conducted to obtain users' opinion in fields such as electronic commerce and trade, management and regarding political figures. Social media has recently become a rich resource in mining user sentiments. A retrieval system and sentiment mining was proposed. Firstly, document's Pre-processing are explored on the dataset. Secondly, Naive Bayes and Support Vector Machines (SVMs) are applied with different feature selection schemes like TF-IDF (Term Frequency–Inverse Document Frequency) and BTO (Binary-Term Occurrence). Thirdly, the proposed model for sentiment analysis is expanded to obtain the results for N-Grams term of tokens. Finally, human has labelled the data and this may involve some mistakes in the labelling process. At this moment, neutral class with generalisation of our classification will take results to different classification accuracy.

Moreover, same model found a clear relationship between Twitter messages and the Saudi Stock Market Index [7]. In order to explain the architectural design of Mubasher products, the Capital Market Authority (CMA) [8] was established in order to protect investors and the public in Saudi Arabia from unfair and fraudulent trading practices. Currently, Tadawul is the only Saudi Stock Exchange in Saudi Arabia. The CMA regulates Tadawul and within Tadawul, there are many

sectors such as Banks, Financial Services, Petrochemical Industries, retail, Telecommunication and Technology [9]. Tadawel authorises Mubasher [10] to access real-time price dissemination for a wide variety of products. Mubasher produces customized products for all type of investment needs Such as Mubasher Pro 10, Mubasher Pro 9, Mubasher NetPlus, Mubasher iPhone, Mubasher Android Phone, Mubasher iPad, Mubasher Touch for Android Tab.

The rest of the paper is organized as follows. In section 2, the related work presents the correlation between sentiment analysis and product review. Section 3 explains sentiment analysis and developments in this field. Section 4 shows the methodology and the process of sentiment analysis in Arabic tweet. Section 5 illustrates the experiment results, evaluations. The final section conclusions remarks and Future works.

2. RELATED WORK

Researchers tend to ignore the “neutral” class under the hypothesis that neutral is less to learn sentiment from neutral texts contrast to positive or negative classes. In the sentiment level, neutral usually means no opinion [11]. Koppel and Schler showed that it is conclusive to use neutral examples in learning polarity. Since, learning from negative and positive examples alone will not permit accurate classification of neutral examples. Besides, the use of neutral training examples in learning enables better distinction between positive and negative examples [12].

Indhuja and Raj P C [13] implemented an approach on sentiment classification using Fuzzy logic for product review documents. Their approach used novel fuzzy functions that emulate the effect of different linguistic hedges and consolidated them into the sentiment classification task. They achieved 85.58% accuracy and they found that some challenges affected the classification, such as entity recognition, co-reference resolution and domain dependency. Zhang and Varadarajan [14] build regression models by incorporating a diverse set of features; they tried to predict utility scores of product reviews, which is orthogonal to polarity classification. Their approach achieved high performance on three Amazon product review collections. Specially, the shallow syntactic features turned out to be the most effective predictors, which indicates that the realized utility of a product review highly and opinion extraction. Dehkharghani and Yilmaz [15] identified the quality of last Windows products against some important features such as security, internet connection and user interface. They used logistic classifier and the average accuracy was 66.6%. Bross and Ehrig [16] used the term aspect-based review mining to detect individual opinion expressions, determine their semantic orientation, then relate that to specific product aspects. Their experiments showed a high accuracy. Wang [17] proposed semi-supervised model called Co-LDA which is consists of two parts: the sentiment LDA part and the topic LDA part. He evaluated his model on two products review sets; the results showed that the Co-LDA is effective to topic sentiment analysis.

3. SENTIMENT ANALYSIS

Social media has recently become a rich resource in mining user sentiments for the purpose of customer relationship management (CRM) [18]. Social opinion has been analysed using sentiment analysis (SA). This is basically a natural language processing (NLP) application that uses computational linguistics and text mining to identify text sentiments as positive, negative and neutral. This technique is known in text mining field as emotional polarity analysis, opinion mining and review mining. In addition, to calculate sentiment score, the sentiment acquired from the text is compared to a dictionary in order to determine the strength of that sentiment. Studies on sentiment analysis focus on text written in English such as sentiment lexicons while applying this to other languages will cause domain adaptation problem [19]. On the other hand, using

semantic processes model called ontology is another technique for text classification. Ontology is a formal and expository representation that includes the vocabulary that refers to the terms in a particular topic domain and the logical statements that describe various expressions such as; how they are related to each other [20].

Traditional text classification is different from sentiment classification. Traditional text classification refers to pre-defined class to determine a document's category, and it gauges the theme of the text itself, while the main aim of the latter is to determine the attitudes and opinion through mining and analysing user interest or other subjective information [21]. Studies in sentiment analysis have found that Pre-processing the data is the procedure of cleaning and adapting. Haddi, Liu and Shi [22] defined the whole Pre-process in several steps: online text cleaning, white space removal, expanding abbreviation, stemming, stop words removal, negation handling and feature selection. The final step is called filtering while the rest are called transformations. Furthermore, the Pre-processing process is supposed to extract relevant contents from the text while leaving out the irrelevant ones.

Data mining discovery is applied to identify patterns in data. Likewise, text mining looks for patterns in text. Text mining can work and analyse unstructured data such as PDF files, emails, and XML files [23]. Many researchers have argued that the use of classifiers in tweets challenges sentiment analysis. Duncan and Zhang [24] claimed that Twitter sentiment analysis is different from traditional sentiment analysis Which should be addressed. The frequency of misspelling and slang in tweets is greater than in other sources due to the 140 character limit. They conducted their experiment using neural networks and the average accuracy was 74.15%. Pang, Lee, and Vaithyanathan [25] employed Naive Bayes, maximum entropy and SVM and their result indicates that SVM could outperform Naive Bayes and Maximum Entropy marginally. On the other hand, Khan, Baharudin, Baharum, Khan, Ullah [26] reported that the problem of domain dependency and a pre-knowledge base remains unsolved. In their study, Pang and lee [27] showed existing method and process for an opinion-oriented information retrieval, while not many researches in opinion mining considered blogs and even fewer addressed microblogging such as Twitter.

The nature of Arabic text is different from English and other languages such as German, Hindi and Chinese, not only orthographically, but also structurally. Arabic texts are written from right to left and the form of letters change depending on their location in the sentence and it consists of 28 letters. Moreover, Arabic words are precise to gender and number and the affixes can change accordingly. Capital letters are not used in Arabic text and special grammatical rules are required to detect nouns, acronyms and abbreviations [28]. Limited research has been conducted in the area of Arabic sentiment analysis and identifying this aspect in the Arabic language is still in its early stages in this field. Ibrahim, Abdou, and Gheith [29] presented MIKA a multi-genre tagged corpus of Modern Standard Arabic MSA and colloquial. Their data focused on MSA and Egyptian dialects. In addition, they mentioned the fact, which is natural language processing NLP approaches that has been applied to other languages, is not valid for process Arabic language directly.

Naive Bayes was used by El Kourdi, Bensaid, and Rachidi [30] to classify Arabic text web documents, which consisted of 300 Arabic documents from five categories from the Aljazeera website. The average accuracy over all the categories was 68.78%. Finally, El-Halees and Alaa Mustafa [31] combined approach depicts the extraction of opinion from Arabic text. His approach consisted of three methods: first, classifying documents using lexicon based methods. Second, the classified documents are used as a training set for the Maximum Entropy model; lastly, the author used K-Nearest Neighbour to classify the rest of the documents. The best classification accuracy was 80.40%. Al-Harbi, Almuhareb, Al-Thubaity, Khorsheed, and Al-Rajeh [32] studied the performance of two classification algorithms (SVM and C5.0) in order to predict the text polarity

of their seven Arabic corpora that contained 17.658 text documents. The result reported that SVM average accuracy is 68.65%, while the average accuracy of C5.0 is 78.42%.

4. METHODOLOGY

Figure 1, summarizes the sentiment analysis process of Arabic tweets. As one of the popular social media platforms, Twitter has been selected for this study as the base for the opinion mining. Since social media channel normally contain words with ambiguous means, the opinion mining based on this kind of social media is still under development, in particular for situations where spelling mistakes, use of emotions and other characters that express special meanings or use of English pronunciation but written it in Arabic characters occurs . In order to satisfy validation requirements, a mixture of Modern Standard Arabic (MSA) and local Saudi Arabian dialects will validate this study. These kinds of tweets consist of independent, semantics-oriented Arabic vocabularies, which complicate the study even further. In order to address the challenge, the ontology of a new keywords process model will be established to improve the text mining. The proposed approach would improve the extraction of hidden information. The machine learning approach is basically a supervised method in which a set of data labelled as positive or negative is represented by feature vectors. After that, the classifier uses those vectors as training data in order to conclude the combination of specific features that relate to a specific class.

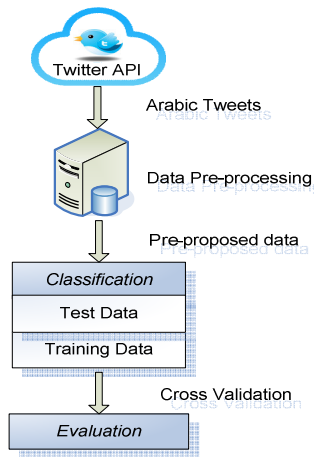


Figure 1. The Process of Sentiment Analysis in Arabic tweet.

To collect the corpus of data with Arabic tweets, a small desktop application (Twitter Data Grabber) was developed using C# with Twitter’s official Developers API. It allows us to label and save the relevant tweet and discard the irrelevant one. Moreover, the data was stored in our system and in normalised form with no hashtags, no duplicate tweets, no retweets, no URLs and no special characters. The tweets source data was obtained from Mubasher company website [10]in Saudi Arabia. The company is a leading stock analysis software provider in the Gulf region. A collection of tweets was compiled over a 57-day period between 15-7-2015 and 12-9-2015. The data set includes 2051 tweets. A selection of over 40 words and idioms in Arabic from the emotion corpus A selection of over 40 words and idioms in Arabic from the emotion corpus such as liberality, growth, Magnificent, Abruption, ineligious, Not accurate etc. was then formed into the following three classes: Positive, Negative, Neutral. For example, Table 1 shows the most common Arabic words and standard Saudi dialect represent the three classes.

Table 1. Common Arabic Words and Standard Saudi Dialect Represent the Classes

Arabic Positive Sentiment	English Translation	Arabic Negative Sentiment	English Translation
عطاء	liberality	انقطاع	Abruption
تطور	growth	مشكلة	Trouble
جميل	Magnificent	بطيئ	Slow
الأكثر ترشيحا	Most rated	الاتجاه أفضل	Etjah better
جيد	good	تكرشارت أفضل	Ticker Chart best
مكامل	Complementary	لم استفد	Uninteresting

The users' privacy issues were handled as follows. Tweets on Twitter are set to public and can be viewed by any person regardless of membership of Twitter. The tweets used in this study are based on data obtained from public timelines.

The Twitter API has several methods, such as GET statuses, user_timeline, GET statuses, home_timeline and GET search / tweets.

This study implemented two methods to download tweets:

- GET statuses/mentions_timeline

Returns the 20 most recent mentions (tweets containing a users's @screen_name) for the authenticating user. The timeline returned is the equivalent of the one seen when you view your mentions on twitter.com.

- GET statuses/user_timeline

Returns a collection of the most recent tweets posted by the user indicated by the screen_name or user_id parameters. User timelines belonging to protected users may only be requested when the authenticated user either "owns" the timeline or is an approved follower of the owner. The timeline returned is the equivalent of the one seen when you view a user's profile on twitter.com. Two Mubasher employees, who have experience in Mubasher software, labelled the data manually. Positive tweets were given the label "1", while negative tweets were given the label "-1". Neutral tweets were given the label "0" and the irrelevant tweets were deleted from the database. Table 2 shows the number of tweets in our database, 727 tweets were deleted in the labelling process.

Table 2. Collection of Tweets

Total	Positive	Negative	Neutral
1331	378	755	198

After the data labelling was completed, Rapidminer [33] was used for the next step: Pre-process-phase, in our experiment by loading the data and Pre-processing it with Rapidminer. The first step was to replace some words, such as company codes, percentages sign (%), Arabic words taking different shapes, and English words to standard Arabic equivalents. For example, Table 3 shows Replacing words in Pre-process-phase.

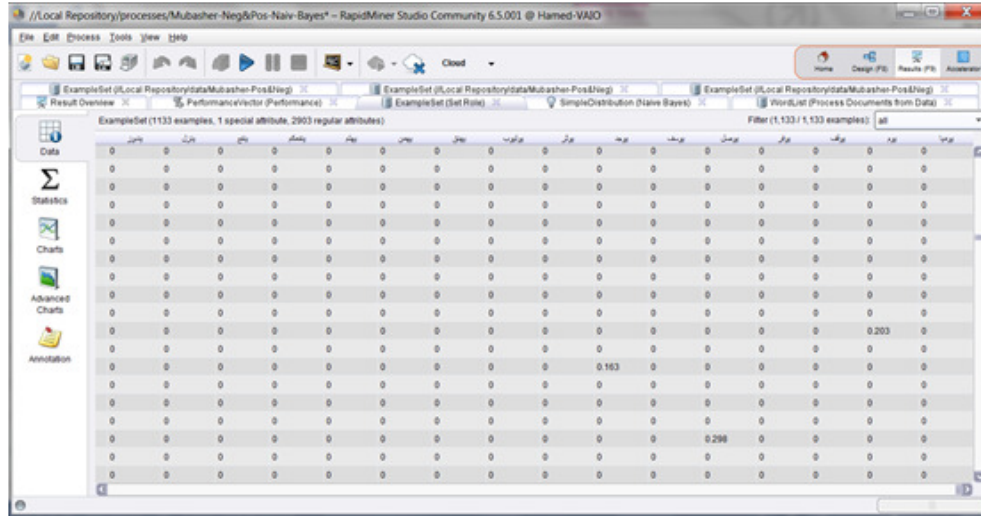


Figure 3: words vector representation.

Finally, relevant data mining approaches were explored, e.g. Naive Bayes and SVM, for building models, which classify tweets according to their sentiment polarity into positive, and negative. The evaluation is carried out using precision and recall methods as shown in Figure 4.

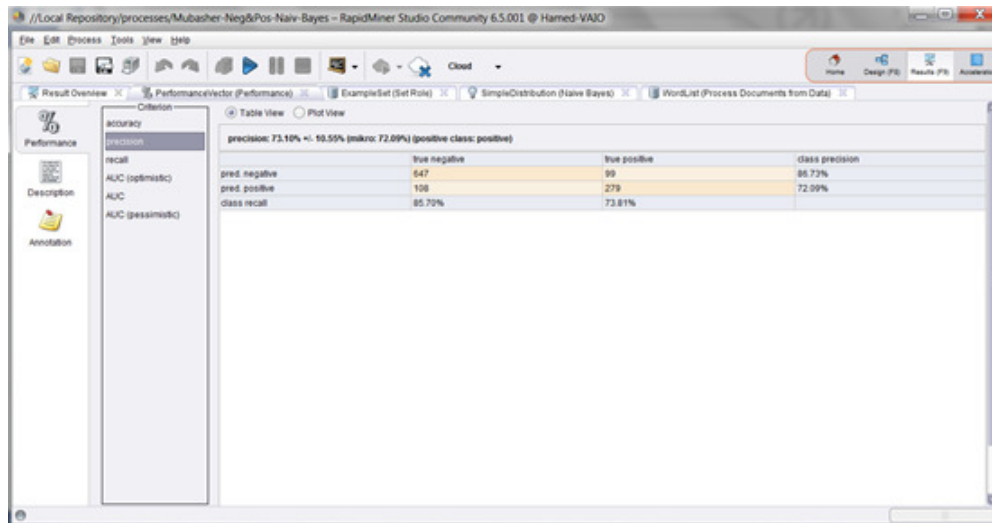


Figure 4: Precision and recall.

5. APPARATUS

The experiment was conducted on Sony laptop, windows 7 operating system (64 bit), Intel (R) Core (TM)-i5-3210M CPU @2.50 GHz, and memory RAM 4.00GB. On the other hand, to examine and validate the results of the study, the proposed data model has been established based on Raipdminer Platform which is a leading Data Mining Software.

6. PROCEDURE

Figure 5 shows the main process using rapidminer. First task is loading the dataset “Retrieve” operator was used, then “Process documents from data” Pre-processd the text using avalibel operators to create the word vecotr list. After that “Set role operator” used to change the role of the attributes which is in our case labeling the text into two class positive and negative. Finally, ”X-validation” operator was used to evaluate the accuracy.

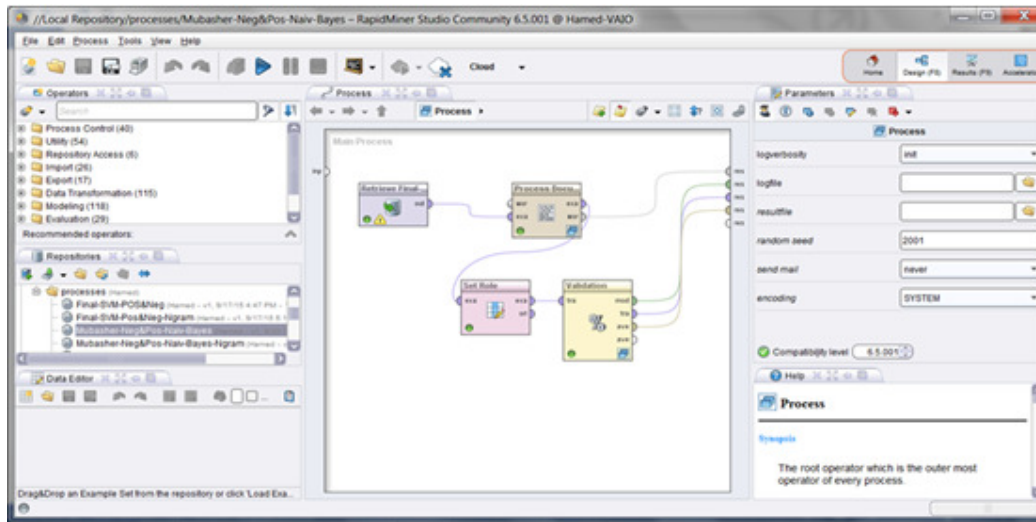


Figure 5: Main Process

In the documents Process, Rapidminer is applied for text processing. It has a list of text processing operators as shown in Figure 6. Those operators process Arabic text in order to create our vecotr list. Such as tokenization, removal stops word etc.

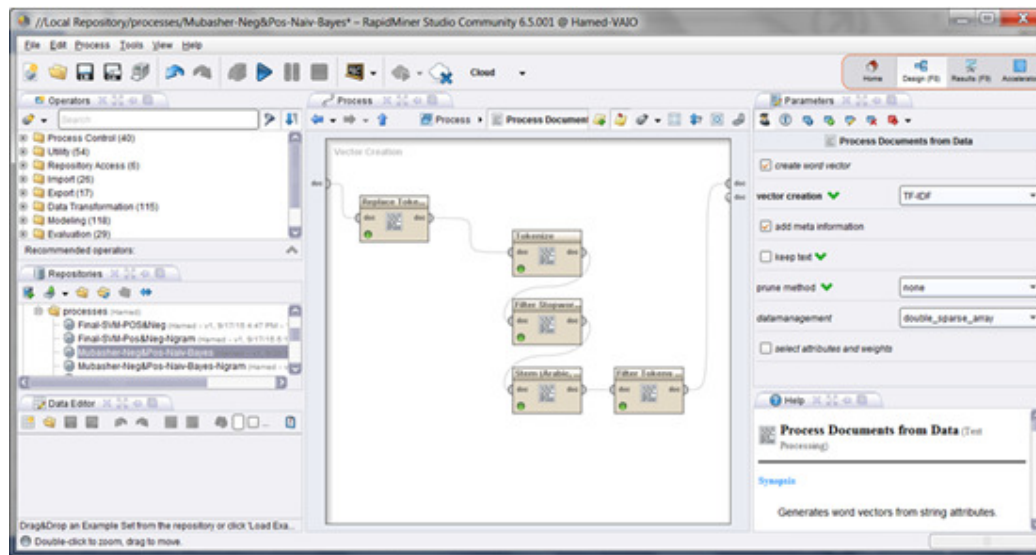


Figure 6: Documents Process

Finally, the Validation and performance Process, which Cross-validation was used and repeated 10 times in order to estimate the accuracy of the statistical model by divided the dataset into two partions, a training data and test data. The model is trained using the training data only and the accuracy is evaluated using the test data, as shown in Figure 7.

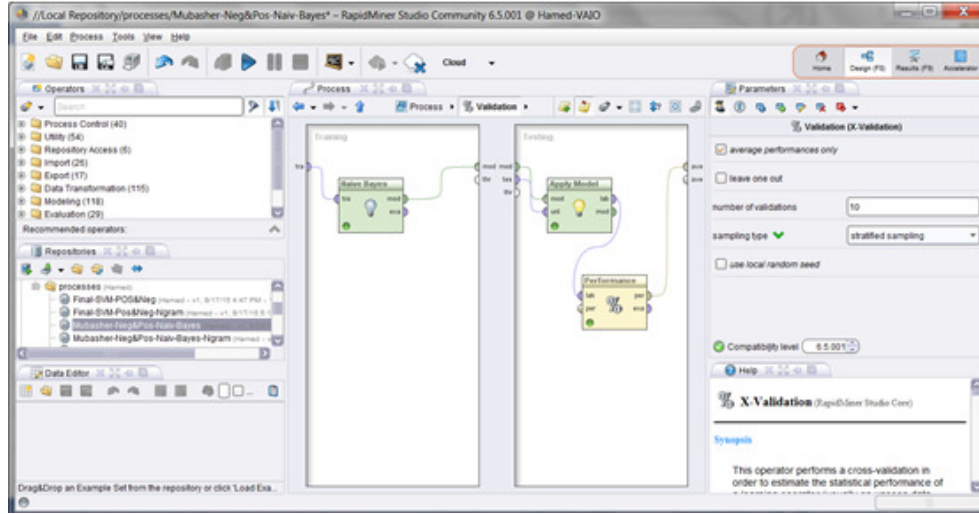


Figure 7: Validation and Performance Process

The performance metrics that were widely used to evaluate the classification results were precision and recall [34].

$$\text{Precision} = \frac{TP}{TP + FP}$$

Higher precision means less false positives.

$$\text{Recall} = \frac{TP}{TP + FN}$$

Higher recall means less false negatives.

Naive Bayes and Support Vector Machines are applied with different weighting schemes like TF-IDF (Term Frequency–Inverse Document Frequency) and BTO (Binary-Term Occurrence). These schemes are often used in information retrieval and text mining [35], [36].

- TF-IDF is the multiple of the value of TF and IDF for a particular word. The value of TF-IDF increases with the number of occurrences within a document and with infrequency of the term across the corps.

Term frequency (TF) is the Product of two statistics, term frequency and inverse document frequency where, TF (term, document) is the frequency of that term in a document i.e. the number of times that term t occurs in document d.

$$\text{TF (term, document)} = \text{Frequency of term} / \text{Number of Document}$$

Inverse document frequency (IDF) calculates whether the word is rare or common in all documents. IDF (term, document) is obtained by dividing total number of Documents by the number of documents containing that term and taking log of that.

$$\text{IDF (term, document)} = \log (\text{Total Number of Document} / \text{Number of Document containing term})$$

- BTO (Binary-term occurrence): word or term receives one if present in document, 0 otherwise.

7. THE EXPERIMENT RESULTS AND EVALUATIONS

This paper explores Arabic text classification in product review within different algorithms such as SVM and Naive Bayes. The analysis shows promising results in SVM algorithms. On the other hand, this phase demonstrates that the text Pre-processing is a key factor in sentiment analysis classification. In order to prove this, when the tweets have different ways of Pre-processing the result will show different levels of accuracy. By creates N-Grams term of tokens, which is a subsequence of n items from a given text [37]. The accuracy in sentiment analysis will change according to the polarity of the text and the classifiers that were used.

7.1 First Experiment

The following tables show the type of tuning in order to provide the best accuracy in sentiment analysis for Arabic text in relation to the Mubasher products using all data, which are 1331 tweets. The classifiers, which were used to explore the polarity of a given data, were Naive-Bayes and SVM. Two different weighting schemes (TF-IDF and BTO) were used. Table 4 shows the class Precision and recall for the Naive-Bayes and SVM classifiers: Crosse-validation=10, sampling type=stratified sampling, prune=none.

Table 4. Precision and Recall for the Naive-bayes and SVM Classifiers

classifier	Accuracy	Recall	Precision	weighting schemes
Naïve-Bayes	81.74	73.91	73.10	TF-IDF
	82.70	79.16	72.71	BTO
SVM	89.68	72.04	96.00	TF-IDF
	88.18	66.46	97.35	BTO

Table 5 shows the class Precision and recall for the Naive-Bayes and SVM classifiers with N-Gram feature that set to two.

Table 5.. Naive-Bayes and SVM Precision and Recall for the Positive and Negative Classes Lasses with N-Gram

classifier	Accuracy	Recall	Precision	weighting schemes
Naive-Bayes	77.53	67.92	82.84	TF-IDF
	76.26	66.62	74.65	BTO
SVM	84.82	66.44	58.91	TF-IDF
	84.89	66.49	59.00	BTO

To sum up the first experiment, the best accuracy was achieved by SVM without N-Gram feature. On the other hand, the best accuracy was completed by Naive-Bayes when the N-Gram feature is involved. Moreover, human has labelled the data and this may involve some mistakes in the labelling process, only for this reason, generalisation of our classification with neutral class will show the importance of neutral class in sentiment analysis and different results of classification accuracy. The reason results are different is that the Pre-processing for the text data consists of all the words that belong to positive and negative classes as well. Table 6 shows the class Precision and recall for the Naive-Bayes and SVM classifiers: Crosse-validation=10, sampling type=stratified sampling, prune=none.

Table 6. Precision and Recall for the Naive-Bayes and SVM Classifiers

classifier	Accuracy	Recall	Precision	weighting schemes
Naïve-Bayes	76.86	69.16	82.16	TF-IDF
	76.26	69.17	76.35	BTO
SVM	84.97	66.58	58.56	TF-IDF
	84.97	66.58	58.77	BTO

Lastly, adding N-Gram feature to conclude this experiment was accomplished to show the different of the accuracy and the class Precision and recall. Table 7 shows the class Precision and recall for the Naive- Bayes and SVM classifiers with N-gram: Crosse-validation=10, sampling type=stratified sampling, prune=none.

Table 7. Precision and Recall for the Naive-Bayes and SVM Classifiers with N-gram

classifier	Accuracy	Recall	Precision	weighting schemes
Naïve-Bayes	83.67	72.06	78.25	TF-IDF
	84.20	74.41	77.94	BTO
SVM	79.62	39.78	98.07	TF-IDF
	78.65	36.62	98.62	BTO

In conclusion, the first experiment show Naive-Bayes performance is better when we used the N-Gram feature. On the other hand, there was performance drop when SVM used the same feature.

7.2 SECOND EXPERIMENT

For the second experiment, we have split the data into two independents part for different periods of date and the experiment has been reproduce again for each part. The details of the second experiment first part are shown in Table 8, 9, 10, 11, and 12, whereas the details of the second experiment second part are shown in Table 13, 14, 15, 16, and 17. Moreover, the experiment followed the same setup, which has been mentioned above. The main purpose of the existing experiment is to demonstrate the success of the proposed method on sentiment analysis of Arabic language. In addition, the polarity problems might be best handled as three classes because we believe that neutral documents are involved mixed of sentiment.

Table 8. Number of Tweets

Total	Positive	Negative	Neutral
665	189	377	99

Table 9. Naïve-Bayes & SVM for Positive & Negative Classes

classifier	Accuracy	Recall	Precision	weighting schemes
Naïve-Bayes	64.58	64.38	64.58	TF-IDF
	68.34	65.21	64.86	BTO
SVM	66.73	50.00	33.36	TF-IDF
	66.73	50.00	33.36	BTO

Table 10. Naive-Bayes & SVM with N-gram for Positive & Negative Classes

classifier	Accuracy	Recall	Precision	weighting schemes
Naïve-Bayes	67.99	64.05	64.08	TF-IDF
	69.22	65.75	65.73	BTO
SVM	66.73	50.00	33.36	TF-IDF
	66.73	50.00	33.36	BTO

Table 11. Naive-Bayes & SVM for All Classes

classifier	Accuracy	Recall	Precision	weighting schemes
Naïve-Bayes	59.87	49.63	51.14	TF-IDF
	60.31	50.83	52.77	BTO
SVM	56.70	33.33	18.90	TF-IDF
	56.70	33.33	18.90	BTO

Table 121. Naive-Bayes & SVM with N-gram For All Classes

classifier	Accuracy	Recall	Precision	weighting schemes
Naïve-Bayes	60.17	49.94	52.08	TF-IDF
	60.76	50.53	54.05	BTO
SVM with N-gram	56.70	33.33	18.90	TF-IDF
	56.70	33.33	18.90	BTO

Table 132. Number of Tweets

Total	Positive	Negative	Neutral
666	189	378	99

Table 14. Naive-Bayes & SVM for Positive & Negative Classes

classifier	Accuracy	Recall	Precision	weighting schemes
Naïve-Bayes	88.35	89.68	92.86	TF-IDF
	88	88.09	93.79	BTO
SVM	90.48	99.74	87.79	TF-IDF
	86.76	99.74	83.88	BTO

Table 15. Naive-Bayes & SVM with N-gram for Positive & Negative Classes

classifier	Accuracy	Recall	Precision	weighting schemes
Naïve-Bayes	90.47	92.33	93.47	TF-IDF
	90.65	91.00	94.92	BTO
SVM	82.19	100.00	79.07	TF-IDF
	81.12	100.00	78.22	BTO

Table 163. Naive-Bayes & SVM for All Classes

classifier	Accuracy	Recall	Precision	weighting schemes
Naïve-Bayes	74.17	65.74	65.91	TF-IDF
	74.32	66.34	65.93	BTO
SVM	56.70	33.33	18.90	TF-IDF
	56.70	33.33	18.90	BTO

Table 174. Naive-Bayes & SVM with N-gram for All Classes

classifier	Accuracy	Recall	Precision	weighting schemes
Naïve-Bayes	75.07	66.10	66.58	TF-IDF
	75.52	66.82	66.81	BTO
SVM	56.76	33.33	18.90	TF-IDF
	56.76	33.33	18.90	BTO

To sum up the second experiment, using negative and positive classes only, the best accuracy was achieved by SVM without N-Gram feature. On the other hand, the best accuracy was completed by Naive-Bayes when the neutral class is added and N-Gram feature is involved. Moreover, the analysis shows same results with SVM for each part of data, which means accurate classification when neutral examples involved.

8. CONCLUSION AND FUTURE WORKS

The main target of this paper is to introduce neutral class with generalisation of our classification, which will take results to the right direction. The analysis shows similar result when we split the data, adding neutral class shows different results of classification accuracy. The reason results are different is that the pre-processing for the text data consists of all the words that belong to positive and negative classes as well.

For future research, we intend to extract technical features aspects of e-learning systems for King Abdulaziz University in Saudi Arabia. It is important to know students' opinions. The extracted features can help developers know the advantages and the disadvantages of their systems.

ACKNOWLEDGMENT

Mubasher Company supported this work. We would like to thank the Managing Director of Mubasher KSA Dr. Walid Rasheed AL-Ballah, Dr. Khalid Al-Omar for their helpful comments and support.

REFERENCES

- [1] J.Bollen, H. Mao, and X. Zeng, "Twitter mood predicts the stock market," *Journal of Computational Science*, vol. 2, pp. 1-8, 2011.
- [2] V.Martin, "Predicting the French Stock Market Using Social Media Analysis," in *Semantic and Social Media Adaptation and Personalization (SMAP)*, 2013 8th International Workshop on, 2013, pp. 3-7.
- [3] W.Zhang and S. Skiena, "Trading Strategies to Exploit Blog and News Sentiment," in *ICWSM*, 2010.
- [4] A.Pak and P. Paroubek, "Twitter as a Corpus for Sentiment Analysis and Opinion Mining," in *LREC*, 2010, pp. 1320-1326.
- [5] A.Agarwal, B. Xie, I. Vovsha, O. Rambow, and R. Passonneau, "Sentiment analysis of twitter data," in *Proceedings of the Workshop on Languages in Social Media*, 2011, pp. 30-38.
- [6] Z.-H.Deng, K.-H. Luo, and H.-L. Yu, "A study of supervised term weighting scheme for sentiment analysis," *Expert Systems with Applications*, vol. 41, pp. 3506-3513, 2014.
- [7] H.AL-Rubaiee, R. Qiu, and D. Li, "Analysis of the relationship between Saudi twitter posts and the Saudi stock market," in *2015 IEEE Seventh International Conference on Intelligent Computing and Information Systems (ICICIS)*, 2015, pp. 660-665.
- [8] (2016, 28 April). Capital Market Authority (CMA), . Available: <http://www.cma.org.sa/en/Pages/home.aspx>
- [9] Tadawul. (2016, 28 April). Available: <http://www.tadawul.com.sa>
- [10] Mubasher. (2016, 28 April). Mubasher Main Website. Available: <http://global.mubashertrade.com>
- [11] B.Liu, "Sentiment analysis and opinion mining," *Synthesis lectures on human language technologies*, vol. 5, pp. 1-167, 2012.
- [12] M.Koppel and J. Schler, "The importance of neutral examples for learning sentiment," *Computational Intelligence*, vol. 22, pp. 100-109, 2006.
- [13] K.Indhuja and R. P. Reghu, "Fuzzy logic based sentiment analysis of product review documents," in *Computational Systems and Communications (ICCSC)*, 2014 First International Conference on, 2014, pp. 18-22.
- [14] Z.Zhang and B. Varadarajan, "Utility scoring of product reviews," in *Proceedings of the 15th ACM international conference on Information and knowledge management*, 2006, pp. 51-57.
- [15] R.Dezhkharhani and C. Yilmaz, "Automatically identifying a software product's quality attributes through sentiment analysis of tweets," in *Natural Language Analysis in Software Engineering (NaturaLiSE)*, 2013 1st International Workshop on, 2013, pp. 25-30.
- [16] J.Broß and H. Ehrig, "Generating a context-aware sentiment lexicon for aspect-based product review mining," in *Web Intelligence and Intelligent Agent Technology (WI-IAT)*, 2010 IEEE/WIC/ACM International Conference on, 2010, pp. 435-439.
- [17] W.Wang, "Sentiment analysis of online product reviews with Semi-supervised topic sentiment mixture model," in *Fuzzy Systems and Knowledge Discovery (FSKD)*, 2010 Seventh International Conference on, 2010, pp. 2385-2389.
- [18] M.M.Mostafa, "More than words: Social networks' text mining for consumer brand sentiments," *Expert Systems with Applications*, vol. 40, pp. 4241-4251, 2013.
- [19] E.Cambria, B. Schuller, Y. Xia, and C. Havasi, "New avenues in opinion mining and sentiment analysis," *IEEE Intelligent Systems*, pp. 15-21, 2013.
- [20] K.Sam and C. Chatwin, "Ontology-based text-mining model for social network analysis," in *Management of Innovation and Technology (ICMIT)*, 2012 IEEE International Conference on, 2012, pp. 226-231.

- [21] G.Fu and X. Wang, "Chinese sentence-level sentiment classification based on fuzzy sets," in Proceedings of the 23rd International Conference on Computational Linguistics: Posters, 2010, pp. 312-319.
- [22] E.Haddi, X. Liu, and Y. Shi, "The role of text pre-processing in sentiment analysis," *Procedia Computer Science*, vol. 17, pp. 26-32, 2013.
- [23] F.Gharehchopogh and Z. Khalifelu, "Analysis and evaluation of unstructured data: text mining versus natural language processing," in *Application of Information and Communication Technologies (AICT)*, 2011 5th International Conference on, 2011, pp. 1-4.
- [24] N.F.da Silva, E. R. Hruschka, and E. R. Hruschka, "Tweet sentiment analysis with classifier ensembles," *Decision Support Systems*, vol. 66, pp. 170-179, 2014.
- [25] B.Pang, L.Lee, and S. Vaithyanathan, "Thumbs up?: sentiment classification using machine learning techniques," in Proceedings of the ACL-02 conference on Empirical methods in natural language processing-Volume 10, 2002, pp. 79-86.
- [26] K. Khan, B. Baharudin, A. Khan, and A. Ullah, "Mining opinion components from unstructured reviews: A review," *Journal of King Saud University-Computer and Information Sciences*, vol. 26, pp. 258-275, 2014.
- [27] B.Pang and L.Lee, "Opinion mining and sentiment analysis," *Foundations and trends in information retrieval*, vol. 2, pp. 1-135, 2008.
- [28] K.Ahmad and Y.Almas, "Visualising sentiments in financial texts?," in *Information Visualisation, 2005. Proceedings. Ninth International Conference on*, 2005, pp. 363-368.
- [29] H.S.Ibrahim, S.M.Abdou, and M.Gheith, "MIKA: A tagged corpus for modern standard Arabic and colloquial sentiment analysis," in *Recent Trends in Information Systems (ReTIS)*, 2015 IEEE 2nd International Conference on, 2015, pp. 353-358.
- [30] M.El Kourdi, A. Bensaid, and T.-e. Rachidi, "Automatic Arabic document categorization based on the Naïve Bayes algorithm," in Proceedings of the Workshop on Computational Approaches to Arabic Script-based Languages, 2004, pp. 51-58.
- [31] A. M. El-Halees, "Arabic text classification using maximum entropy," 2007.
- [32] S. Al-Harbi, A. Almuhareb, A. Al-Thubaity, M. Khorsheed, and A. Al-Rajeh, "Automatic Arabic text classification," 2008.
- [33] Rapidminer. (2016, 28 April). Available: <https://rapidminer.com/>
- [34] F.H.Khan, S.Bashir, and U. Qamar, "TOM: Twitter opinion mining framework using hybrid classification scheme," *Decision Support Systems*, vol. 57, pp. 245-257, 2014.
- [35] G.T.a.N.S, "Feature Selection and Classification Approach for Sentiment Analysis," *Machine Learning and Applications: An International Journal (MLAIJ)*, vol. Vol.2, pp. 1-16, June 2015.
- [36] (2016, 28 April). Available: https://www.researchgate.net/post/how_can_we_find_the_tf-idf_value_of_a_word_in_the_corpus
- [37] G.T.a.N.S, "Feature Selection and Classification Approach for Sentiment Analysis," *Machine Learning and Applications: An International Journal (MLAIJ)*, vol. 2, 2015.