

# Sentiment Analysis of Arabic Tweets in e-Learning

<sup>1</sup>Hamed AL-Rubaiee, <sup>2</sup>Renxi Qiu, <sup>3</sup>Khalid Alomar and <sup>4</sup>Dayou Li

<sup>1</sup>Department of Computer Science and Technology, University of Bedfordshire, Bedfordshire, UK

<sup>2</sup>Department of Computer Science and Technology, University of Bedfordshire, Bedfordshire, UK

<sup>3</sup>Department of Information Systems, King Abdulaziz University, Jeddah, Kingdom of Saudi Arabia

<sup>4</sup>Department of Computer Science and Technology, University of Bedfordshire, Bedfordshire, UK

## Article history

Received: 24-10-2016

Revised: 26-10-2016

Accepted: 13-12-2016

## Corresponding Author:

Hamed Saad AL-Rubaiee

Department of Computer

Science and Technology,

University of Bedfordshire,

Bedfordshire, UK

Email: Rubaiee@study.beds.ac.uk

**Abstract:** In this study, we present the design and implementation of Arabic text classification in regard to university students' opinions through different algorithms such as Support Vector Machine (SVM) and Naive Bayes (NB). The aim of the study is to develop a framework to analyse Twitter "tweets" as having negative, positive or neutral sentiments in education or, in other words, to illustrate the relationship between the sentiments conveyed in Arabic tweets and the students' learning experiences at universities. Two experiments were carried out, one using negative and positive classes only and the other one with a neutral class. The results show that in Arabic, a sentiments SVM with an n-gram feature achieved higher accuracy than NB both with using negative and positive classes only and with the neutral class.

**Keywords:** Sentiment Analysis, Twitter, Arabic Tweets, Saudi Arabia, King Abdulaziz University, Machine Learning, Pre-Processing, e-Learning

## Introduction

The number of Arabic users on microblogging platforms such as Twitter has been rapidly increasing (Mourtada and Salem, 2014). In March 2014, the total number of active Arabic Twitter users reached 5,797,500 and the estimated number of tweets created was 533,165,900 tweets, with an average of 17,198,900 tweets per day (Mourtada and Salem, 2014). With the rapid growth of social media utilisation in the Arabic world, there also came a growth of Arabic reviews, comments, ratings, feedback and opinions.

Today, students use social media sites on a daily basis to express their opinions and describe their activities. Therefore, it follows that universities utilise social media in educational practice for the purpose of improving their teaching processes as well as for the purpose of analysing opinions and learning about experiences. This is especially relevant for students who mostly depend on the web, such as online distance education students. Therefore, this paper aims to understand distance education students' learning experiences by analysing students' opinions from the Twitter account of the Deanship of e-Learning and Distance Education (DELDE) at King Abdulaziz University in Saudi Arabia. Opinion mining and sentiment analysis identify how sentiments are expressed

in texts as negative, positive or neutral toward the subject (Nasukawa and Yi, 2003). Many studies (Pak and Paroubek, 2010; Agarwal *et al.*, 2011; Deng *et al.*, 2014) have been carried out to obtain users' opinions in different fields, such as e-commerce (Hu and Liu, 2004), the stock market (Bollen *et al.*, 2011; Martin, 2013; Zhang and Skiena, 2010) and politics (Tumasjan *et al.*, 2010; O'Connor *et al.*, 2010).

The rest of the paper is organised as follows: Section 2 covers related work. Section 3 presents the process of obtaining the Arabic sentiments from Twitter. Section 4 covers the classification methods used. The experiment and the results of the study are described in section 5. The final section covers the conclusion and plans for future works.

## Related Work

Tian *et al.* (2009) developed an e-learner's affective category model that applies a questionnaire. They first extract an emotion feature from the corpus on the Internet. They then compare emotion words with one another to measure the intensity of sentiment in each category. This is a comprehensive decision-tree approach that enables one to achieve positive results in dealing with classification challenges common to the analysis of interactive texts (Chinese, for example), which are characterised by the richness of emotions and short

abbreviations. Wang *et al.* (2010) in their studies described a monitoring and analysis system of Internet education for public sentiment. They worked on the basic function of public sentiment in education over the Internet. Wang (2010) established the Student Feedback Mining System (SFMS) on the basis of text analytics and an opinion-mining approach. In their studies, they performed an in-depth analysis of the qualitative student feedback and provided insight into teaching practices. As a result, they significantly improved student learning. Donovan *et al.* (2006) found out that at least 54% of online student comments were more detailed and informative than typical written feedback on papers. That emphasises the importance of paying more attention to online comments. However, the downside is that the analysis of online comments is more time-consuming and requires an automated feedback system to collect and analyse online comments. It also requires performing a visual analysis of the opinions-sentiments classification. In contrast to the traditional text classification, which is a pre-defined class for identifying a document's category, the sentiment classification defines the attitudes and opinions of users by mining and analysing their interests and other personal information (Fu and Wang, 2010).

### Sentiment Analysis

Sentiment Analysis (SA) is a form of natural language processing that utilises computational linguistics and data mining to identify text sentiments as negative, positive and neutral. Text mining has been used to detect patterns in text. The major advantage of the text-mining technique is that it is capable of analysing the unstructured data (Gharehchopogh and Khalifelu, 2011). However, Duncan and Zhang (2015) also highlighted that traditional sentiment analysis is fundamentally different from Twitter sentiment analysis and that difference should be considered. Because of the 140-character limit in tweets, there is a high level of misspellings and colloquial language on Twitter in comparison to any other source. Thus, the application of neural networks identified only 74.15% accuracy. In their study, Pang and Lee (2008) were two of the few researchers who attempted to use the sentiment analysis method for opinion-oriented information. Thus, there is a scarcity of studies on applying this method to blogs in general and even less research specifically in the area of microblogging platforms such as Twitter.

### Sentiment Analysis Approach

Most studies on sentiment analysis are focused on text written in English using the sentiment lexicons technique. However, applying the lexicons technique to other languages will cause a domain adaptation problem (Cambria *et al.*, 2013). Also, there is no lexicon for Arabic sentiment words and thus the machine-learning

or corpus-based technique is usually used. Both techniques are supervised methods, where a set of data is first categorised as negative, positive or neutral and appears by feature vectors. Afterwards, the classifier uses those vectors as training data to identify similar features and to group the data them in a certain class (Shoukry and Rafea, 2012).

### The Uniqueness of the Arabic Language

This paper discusses the structural complexity of the Arabic language and, as a result, the challenges associated with Arabic text analysis. These challenges are also coupled with the recent spread of social media and the use of colloquial Arabic language on social platforms. This work highlights the lack of research in the field and the limitations of the previous studies. As an alternative, it also proposes a different method for text analysis, called sentiment analysis, which is further discussed in detail in the below sections.

In comparison to other language families, including German, Hindi and Chinese, the structure of the Arabic language is fundamentally different. First, the writing rule is based on the principle of 'right to left', while the form of letters depends on their place in a sentence. Second, gender plays a key role in determining the affixes of words. Third, capital letters do not subsist in the Arabic language and there are grammatical rules that aim to detect entities, acronyms and abbreviations (Fu and Wang, 2010). All these complexities result in difficulties in Arabic text analysis, which are also coupled with the lack of research in that field. A few authors have attempted to analyse Arabic texts by applying different methods. However, their approaches have a number of limitations. For instance, Arabic text web documents from the *Al Jazeera* website were classified into five main categories using Naïve Bayes (Gharehchopogh and Khalifelu, 2011; El Kourdi *et al.*, 2004). Their dataset contained 300 Arabic articles using NB. The average accuracy in those identified categories did not exceed 68.78%. El-Halees' applied a different method based on depicting the extraction of opinion from Arabic texts (Duncan and Zhang, 2015; El-Halees, 2007). That method encompassed three stages. First, the documents were classified according to lexicon; second, those documents were utilised to train the maximum entropy model; and finally, the rest of them were categorised by the k-Nearest Neighbour. Such approaches enabled the achievement of a higher level of accuracy - 80.40%. Ibrahim *et al.*, (2015) in their studies investigated the application of Modern Standard Arabic (MSA) and colloquial Arabic to Egyptian dialects. Even though they emphasised that it is a Natural Language Processing approach (NLP) that works relatively well for other languages, the main limitation of this approach is that it is not applicable to the Arabic language. The main

reason for that is the morphological richness of Arabic as a language (Pang and Lee, 2008; Cambria *et al.*, 2013; Duwairi *et al.*, 2014; Farghaly and Shaalan, 2009; Farra *et al.*, 2010). In fact, the Arabic informal (colloquial) language lacks structure and is difficult to standardise. There is limited research on how to analyse the Arabic colloquial language. This limitation is also coupled with the complexity of the social media phenomenon, which tends to include a large number of words with ambiguous meanings. Applying the opinion-mining technique to the sentiment analysis in social media is still in the process of its development, especially in the area of spelling mistakes, emoticons or any other special characters. To deal with the discussed complexities of the Arabic language, this study will do data mining, applying the sentiment analysis technique to modern standard Arabic with the colloquial Saudi Arabian dialects. The main complexity of this research design is in the analysis of the isolated semantics common to the structure of the Arabic language. To deal with this complexity and to facilitate the process of text mining, the ontology of the incipient keyword-processing model will be constructed. The suggested method will enable the withdrawal of unobserved information, applying the results for higher-order n-grams. The n-gram feature is a subsequence of N items from a given text. Higher order N-grams are more accurate in detecting the context, as they give a better understanding of the word location. The strategy is to use tokens such as trigrams ( $N = 3$ ) in the feature space rather than just unigrams ( $N = 1$ ) (Tripathi and Naganna, 2015). This model demonstrates the essentiality of a neutral class in sentiment analysis in the case of the Arabic language. A neutral class with generalisation of the proposed classification methods will lead to a higher classification accuracy (Koppel and Schler, 2006; Hamed *et al.*, 2016).

## Methodology

Figure 1 presents the process of obtaining the Arabic sentiments from Twitter.

### Data Collection

To accumulate the Arabic tweets, an application was developed to grab tweets from a famous social site called Twitter. Two thousand tweets by different students on the DELDE Twitter account (K.A. University, 2016) were downloaded by this application. The application was developed in C# and Twitter's official Developers API was utilised to download tweets.

The tweets were preserved in a database. The tweets could then be tagged in the grid view based on the text. The tweet could be manually marked as negative (-1), positive (1) or neutral (0). Marked tweets are updated in

the database. The marked tweets can then be exported to text files, Excel sheets and .csv files.

The REST APIs were habituated to read and write Twitter data, for instance, GET statuses/mentions\_timeline and GET statuses/user\_timeline.

### Labelling Process

The labelling process started with all the Arabic tweets saved in the main dataset. Then, all the saved tweets were filtered. The filtering process was performed by applying the following criteria:

- Tweets should not contain hashtags
- Tweets should not contain links/URLs
- Duplicate tweets are eliminated
- Tweets should not contain special characters

Then, all filtered tweets were saved in a dataset called the 'ToBeMarked' dataset. The user culled the sentiment of the tweet to be rated and then preserved it in the database. After the tweet was preserved, it was moved back to the 'ToBeMarked' dataset. The final dataset holds only tweets that are certain to be marked correctly. Once the label assigned by the owner matches the label assigned by the supervisor, then the tweet is considered as marked correctly. In case the anterior two labels are mismatched, a third one (manager) comes to avail and all three decide the final outcome. If these precedent conditions are met, the tweet peregrinates to the final stage and is then utilised in the subsequent phases. The supervisor is one of the authors and the owner is the employee who deals with the Twitter account. The manager is the vice dean of the Deanship of e-Learning and Distance Education and is responsible for the Twitter account and also has experience dealing with students' quandaries. Table 1 displays the number of collected tweets (2000 Arabic tweets); 1121 tweets were marked for the training dataset and a total of 879 unrelated tweets were deleted.

### Data Distribution

Figure 2 shows the Twitter data and their daily distributions. The figure shows the average numbers of total, negative, positive and neutral tweets. The data distribution depends on the Twitter API methods that have been used and the number of tweets that have been posted on the DELDE twitter account (K.A. University, 2016).

### Data Pre-Processing

Many studies in sentiment analysis corroborated that pre-processing consists of several steps: Online text cleaning, stop words abstraction, white space abstraction, stemming, expanding abbreviation, negation handling and feature culling (Gokulakrishnan *et al.*, 2012; Haddi *et al.*, 2013).

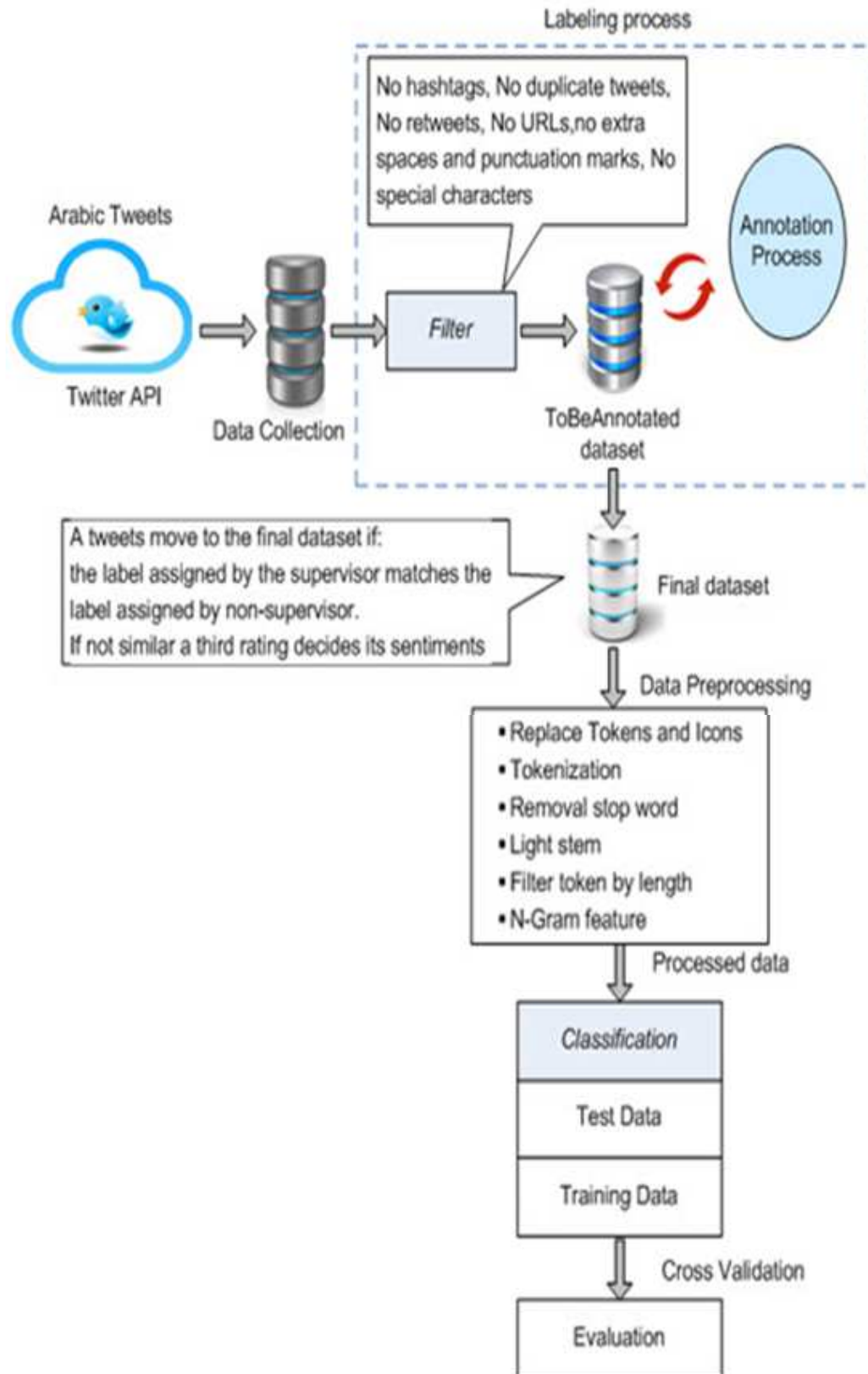


Fig. 1. The framework for obtaining the Arabic sentiments from Twitter

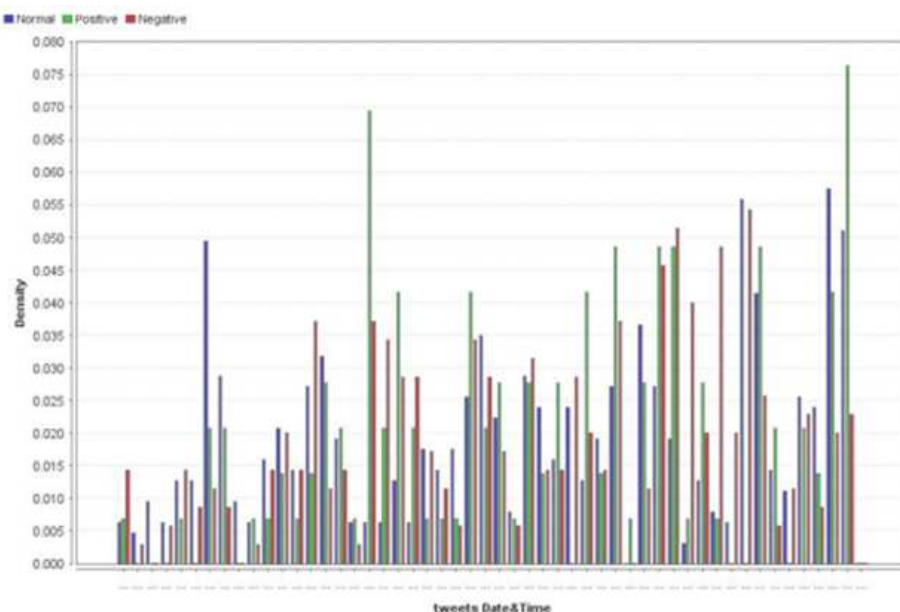


Fig. 2. Tweets distribution

Table 1. The number of obtained tweets

Total	Positive	Negative	Neutral
1121	144	350	627

Gokulakrishnan *et al.* (2012) carried out essential steps in the comprehensive classification process. The steps are as follows:

#### Replacing Emoticons

Emoticons are used in microblogging to express users' emotions in a brief way. There are about 30 icons, such as “☹”, “😊”, “: D”, “=”, “=”, etc., which can be replaced with frown keywords.

#### Upper Case and Lower Case

The goal is to make the cases consistent for the classifier and to map tokens to the corresponding feature irrespective of casing.

#### URL Extracting

URLs are used to share more explanation than what can be given in a short post. Some studies replace that with an equivalence class <URL> in order to reduce feature size, while others simply remove URLs from the text.

#### Detection of Pointers (Usernames and Hashtags)

In Twitter, messages can point to another utiliser with the utilisation of the @ token in front of the username. Moreover, a hashtag is utilized to categorise the post. Supersession of @ and the hashtag with the fine-tuned symbols <utilizer> and <hashtag> will reduce the feature size.

#### Identification of Punctuation

Punctuation is used in microblogging to avoid proper grammar and to communicate emotions more easily. Removing punctuation from the text will reduce the redundant feature in the training set.

#### Removal of Stop Words

It is a common method to remove unvalued substantial words such as “a”, “an” and “the” in the classification process.

#### Compression of Words

Twitter users often express strong emotions in an informal way, for example, writing “happyyyyyyy” and “cooooooooool”. Reducing these expressions to three characters will make a difference between the regular usage and emphasised usage of these words. Rapidminer (2016) was used to pre-process the collected Arabic tweets and then replace any Arabic words that use different styles. For example, Table 2 shows the replacement of some words and icons that use different styles in the pre-process phase.

The following steps were used to pre-process tweets using Rapidminer:

- Tokenization: Splits the text of a tweet into a sequence of tokens
- Stop-word process: Filters Arabic stop words
- Light stem: Stems Arabic tweets by removing the suffixes and prefixes
- Filter token by length: Filters tokens based on their character numbers

Table 2. Replacing words and icons

Words and Icons	English	Replaced
👋	Bye	مع السلامة
😊, 🙄, 🙇	Wonder	تعجب
🌺, 🍀, 🌻, ❤️, ♥️, 🍃, 🍷, 🎧	Positive	ايجابي
😞	Negative	سلبي
إ-أ-آ-أ-إ-أ-ة-ي	Arabic letters	ا-لا-ه-و-ى
ممتاااااااااا	Perfect	ممتاااااااااا
أخوووووي	Brother	اخوووى
الاداب	Arts	الاداب
تكفووووونننن	Please	تكفونون

Then, the 'Process Documents from Data' operator engenders a word vector from the dataset utilising the TF-IDF method that can be calculated with a multiple of the value of TF and IDF for a particular word. This scheme is often utilised in information retrieval and text mining (Hamed *et al.*, 2016; Tripathi and Naganna, 2015). NB and SVM were used to build the classification models used to classify tweets as negative, positive and neutral. Finally, precision and recall methods were used to evaluate the classification results.

## Classification Methods

The NB model has many advantages: It treats each feature independently, it is relatively expeditious to compute, it is easy to construct with no desideratum for any intricate reiterated parameter estimation schemes and it has less over-fitting compared to other models. On the other hand, the strength assumption of conditional independence between features reduces the potential of NB (Alsaedi *et al.*, 2014). In addition, SVM were applied prosperously in many sentiment analysis tasks. SVM has outperformed other machine-learning techniques because of its primary advantages. For instance, generality of text categorisation dilemmas are linearly separable, it's robust in high-dimensional spaces and powerful when there is a sparse set of samples and any feature is pertinent (Rushdi-Saleh *et al.*, 2011). The fundamental conception of SVM is to find a hyper-plane represented by a vector that does not only dissever the document vectors into a different class from those in other documents.

## Evaluation

The precision and recall metrics were obtained to evaluate the classification results (Khan *et al.*, 2014):

- Precision =  $tp/(tp + fp)$
- Recall =  $tp/(tp + fn)$

## Experiment

This experiment explores Arabic text classification in e-learning as an education pattern with different

algorithms such as NB and SVM. It would be necessary to pre-process the MSA Arabic language with the dialects and analyse students' feedback (comments) to test the performance of the sentiment classification method, which can help improve the e-learning system process. The classifiers, which were used to explore the polarity of 2000 tweets with the DELDE, were NB and SVM. Table 3 shows the classification results using NB and SVM.

Table 4 shows the classification results using NB and SVM with an n-gram equal to three.

Receiver Operating Characteristics (ROC) graphs are a utilisable technique to evaluate and visualise the predictive power of a binary classifier. ROC graphs represent the difference of likelihood level between the True Positive Rate (TPR) that is plotted on the y-axis and the Mendacious Positive Rate (FPR), which is plotted on the x-axis (Fawcett, 2004). The best point in the ROC space is located in the upper left corner (0,1). Figure 3-5 show the ROC curve for the SVM and Naive-Bayes classifiers with different parameters determining how the ROC are evaluated, i.e., correct predictions are counted first, last or alternately (Rapidminer, 2016):

- **Optimistic:** The correct classified examples are taken into account before visually examining the erroneous classification
- **Pessimistic:** Erroneous classifications are taken into account before optically canvassing the correct classifications
- **Neutral:** A combination of both optimistic and pessimistic calculation methods. Here, correct true and erroneous classifications are taken into account reciprocally

The conclusion of the first experiment is that the best precision was achieved by SVM without an n-gram feature. On the other hand, the worst precision was consummated by NB when the n-gram feature was involved. In addition, researchers tend to ignore the neutral class under the hypothesis that there is less learning sentiment from neutral texts in contrast to positive or negative classes. At the sentiment level, neutral usually means no opinion (Liu, 2012). Koppel and Schler (2006) showed that it is important to use neutral examples in learning polarity. Table 5 shows the classification results using NB and SVM with all classes.

Lastly, to show the difference in the accuracy and the class precision and recall, the n-gram feature added to this experiment. Table 6 shows the classification results using NB and SVM with n-gram for all classes.

A lift chart is a quantification of the efficacy of a predictive model calculated as the ratio between the results obtained with and without the predictive model



(Jaffery and Liu, 2009). The lift chart is easier to read and illustrate than the ROC chart, which is a discrete version used to represent and visualise the classifier performance. The highest confidence numbers are

shown first and, as can be seen, the confidences numbers decrease at some point. For example, Fig. 6 shows the lift charts for the neutral class with the NB classifier.

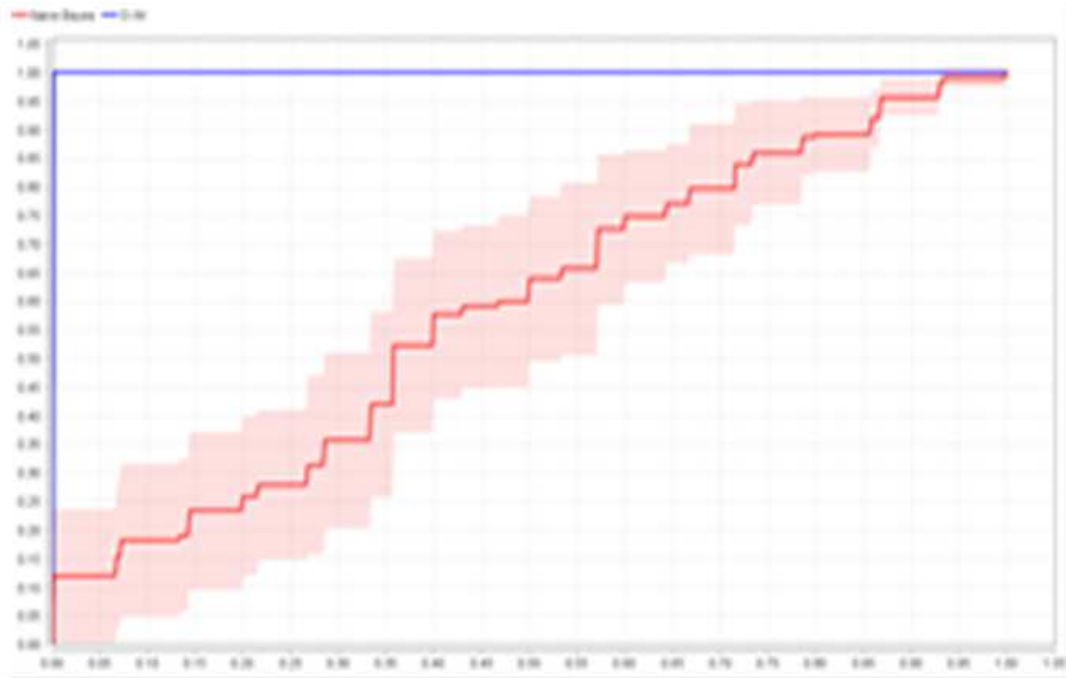


Fig. 3. Optimistic ROC curve for the classifiers

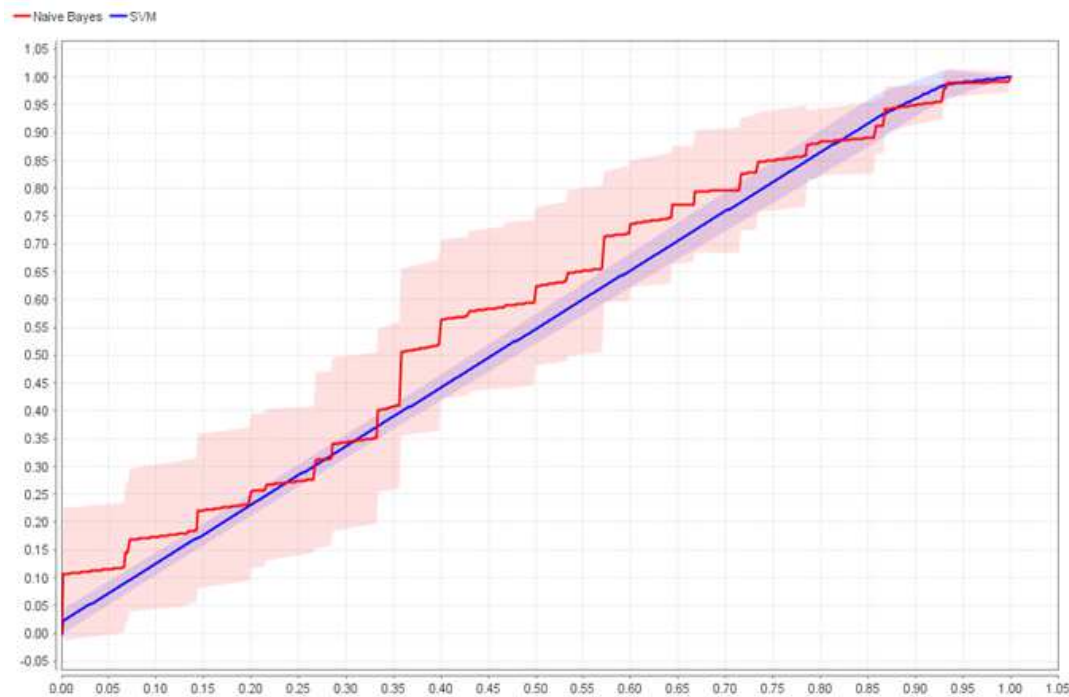


Fig. 4. Pessimistic ROC curve for the classifiers

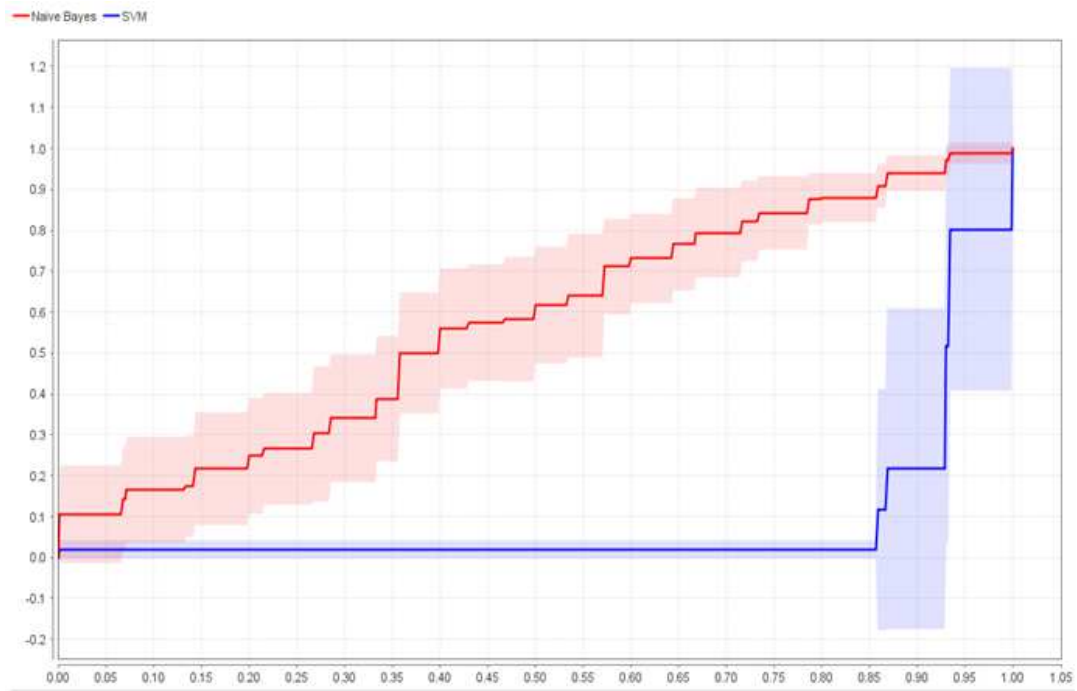


Fig. 5. Neutral ROC curve for the classifiers

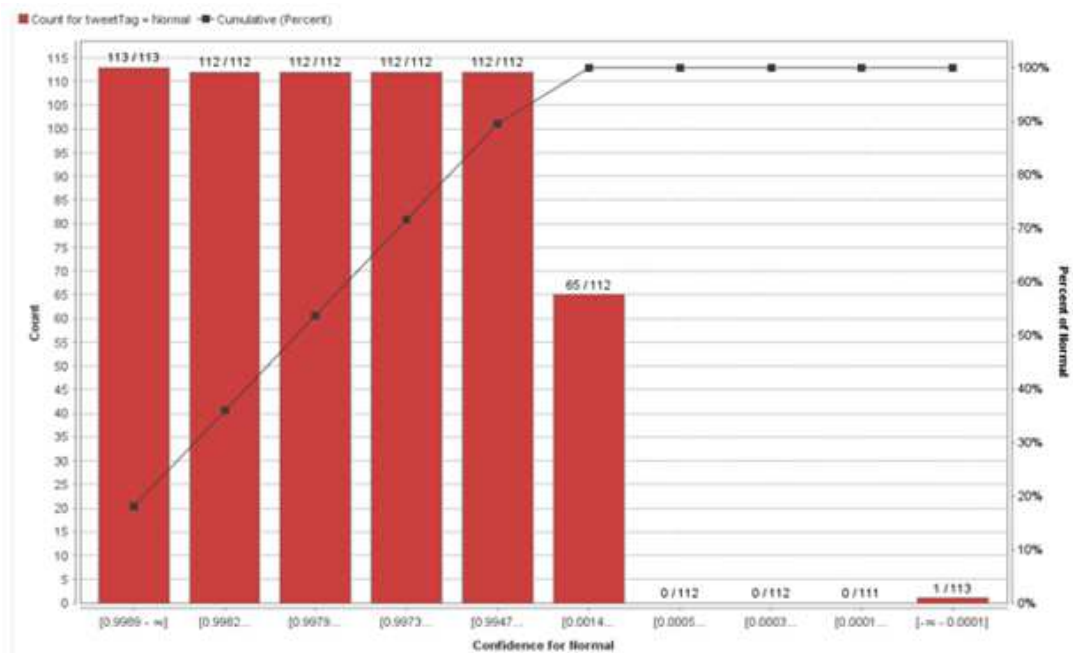


Fig. 6. NB lift chart for the neutral class

Table 3. Classification results for the positive and negative classes

Classifier	Accuracy	Recall	Precision	Weighting schemes
Naïve-Bayes	79.73	77.02	75.96	TF-IDF
SVM	84.62	99.14	82.62	TF-IDF



Table 4. Classification results for the positive and negative classes with n-gram equal to three

Classifier	Accuracy	Recall	Precision	Weighting Schemes
Naïve-Bayes	65.36	65.48	63.26	TF-IDF
SVM	84.84	98.00	85.72	TF-IDF

Table 5. Classification results for all classes

Classifier	Accuracy	Recall	Precision	Weighting schemes
Naïve-Bayes	57.89	57.44	55.43	TF-IDF
SVM	72.08	61.43	75.59	TF-IDF

Table 6. Classification results with n-gram for all classes

Classifier	Accuracy	Recall	Precision	Weighting schemes
Naïve-Bayes	51.56	52.70	51.29	TF-IDF
SVM	73.15	62.28	78.91	TF-IDF

In conclusion, the first run shows the NB performance drops when we utilised the n-gram feature. On the other hand, there was minor amelioration in the precision of the result when SVM utilised the same feature.

## Conclusion and Future Works

This paper presents the design and implementation of Arabic text classification in regard to King Abdul-Aziz University students' opinions using different algorithms such as SVM and NB. The analysis shows promising results with online student comments. We found that their text was more traditional and that local Saudi Arabian dialects are exercised in feedback more often than Modern Standard Arabic (MSA). To summarise, in the two experiments, which used negative and positive classes only, the best accuracy was achieved by SVM with the n-gram feature. On the other hand, the best accuracy was completed by SVM when the neutral class was added and the n-gram feature was involved.

In future work, we plan to gather blackboard structured data from King Abdulaziz University and then label the student rows into three classes, negative, positive and neutral, based on behaviour study patterns. Then, semantic schema will be established utilising the exact same approach as this paper. The extracted features can help to understand the comportment of the students during their study.

## Acknowledgment

The first author would like to thank the Deanship of e-Learning and Distance Education at King Abdulaziz University for their helpful comments and support. He would also like to thank the Saudi Arabian Cultural Bureau in London for its support in collecting the required data.

## Authors' Contributions

All authors equally contributed to this work.

## Ethics

All corresponding authors have made the same effort on this article and all agree that no ethical issues are involved in this manuscript. In addition, all authors agreed that this article contains unpublished material.

## References

- Agarwal, A., B. Xie, I. Vovsha, O. Rambow and R. Passonneau, 2011. Sentiment analysis of twitter data. Proceedings of the Workshop on Languages in Social Media, Jun. 23-23, Association for Computational Linguistics, Portland, Oregon, pp: 30-38.
- Alsaedi, N., P. Burnap and O.F. Rana, 2014. A combined classification-clustering framework for identifying disruptive events. Proceedings of the ASE SocialCom Conference, May 27-21, Stanford University, CA., USA.
- Bollen, J., H. Mao and X. Zeng, 2011. Twitter mood predicts the stock market. J. Comput. Sci., 2: 1-8. DOI: 10.1016/j.jocs.2010.12.007
- Cambria, E., B. Schuller, Y. Xia and C. Havasi, 2013. New avenues in opinion mining and sentiment analysis. IEEE Intell. Syst., 28: 15-21. DOI: 10.1109/MIS.2013.30
- Deng, Z.H., K.H. Luo and H.L. Yu, 2014. A study of supervised term weighting scheme for sentiment analysis. Expert Syst. Applic., 41: 3506-3513. DOI: 10.1016/j.eswa.2013.10.056
- Donovan, J., C.E. Mader and J. Shinsky, 2006. Constructive student feedback: Online Vs. traditional course evaluations. J. Interactive Online Learn., 5: 283-296.
- Duncan, B. and Y. Zhang, 2015. Neural networks for sentiment analysis on Twitter. Proceedings of the IEEE 14th International Conference on Cognitive Informatics and Cognitive Computing, Jul. 6-8, IEEE Xplore Press, pp: 275-278. DOI: 10.1109/ICCI-CC.2015.7259397

- Duwairi, R., R. Marji, N. Sha'ban and S. Rushaidat, 2014. Sentiment analysis in Arabic tweets. Proceedings of the 5th International Conference on Information and Communication Systems, Apr. 1-3, IEEE Xplore Press, pp: 1-6.  
 DOI: 10.1109/IACS.2014.6841964
- El Kourdi, M., A. Bensaid and T.E. Rachidi, 2004. Automatic Arabic document categorization based on the Naïve Bayes algorithm. Proceedings of the Workshop on Computational Approaches to Arabic Script-based Languages, Aug. 28-28, Geneva, Switzerland, pp: 51-58.
- El-Halees, A.M., 2007. Arabic text classification using maximum entropy. IUG J. Nat. Stud., 15: 157-167.
- Farghaly, A. and K. Shaalan, 2009. Arabic natural language processing: Challenges and solutions. ACM Trans. Asian Lang. Inform. Process., 8: 14-14.  
 DOI: 10.1145/1644879.1644881
- Farra, N., E. Challita, R.A. Assi and H. Hajj, 2010. Sentence-level and document-level sentiment mining for Arabic texts. Proceedings of the IEEE International Conference on Data Mining Workshops, Dec. 13-13, IEEE Xplore Press, pp: 1114-1119. DOI: 10.1109/ICDMW.2010.95
- Fawcett, T., 2004. ROC graphs: Notes and practical considerations for researchers. Mach. Learn., 31: 1-38.
- Fu, G. and X. Wang, 2010. Chinese sentence-level sentiment classification based on fuzzy sets. Proceedings of the 23rd International Conference on Computational Linguistics: Posters, Aug. 23-27, Beijing, China, pp: 312-319.
- Gharehchopogh, F. and Z. Khalifelu, 2011. Analysis and evaluation of unstructured data: Text mining versus natural language processing. Proceedings of the 5th International Conference on Application of Information and Communication Technologies, Oct. 12-14, IEEE Xplore Press, pp: 1-4.  
 DOI: 10.1109/ICAICT.2011.6111017
- Gokulakrishnan, B., P. Priyathan, T. Ragavan, N. Prasath and A. Perera, 2012. Opinion mining and sentiment analysis on a twitter data stream. Proceedings of the International Conference on Advances in ICT for Emerging Regions, Dec. 12-15, IEEE Xplore Press, pp: 182-188.  
 DOI: 10.1109/ICTer.2012.6423033
- Haddi, E., X. Liu and Y. Shi, 2013. The role of text pre-processing in sentiment analysis. Proc. Comput. Sci., 17: 26-32.  
 DOI: 10.1016/j.procs.2013.05.005
- Hamed, A.R., R. Qiu and D. Li, 2016. The importance of neutral class in sentiment analysis of Arabic tweets. Int. J. Comput. Sci. Inform. Technol., 8: 17-31.
- Hu, M. and B. Liu, 2004. Mining and summarizing customer reviews. Proceedings of the 10th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Aug. 22-25, ACM, Seattle, WA, USA., pp: 168-177.  
 DOI: 10.1145/1014052.1014073
- Ibrahim, H.S., S.M. Abdou and M. Gheith, 2015. MIKA: A tagged corpus for modern standard Arabic and colloquial sentiment analysis. Proceedings of the IEEE 2nd International Conference on Recent Trends in Information Systems, Jul. 9-11, IEEE Xplore Press, pp: 353-358. DOI: 10.1109/ReTIS.2015.7232904
- Jaffery, T. and S. Liu, 2009. Measuring campaign performance by using cumulative gain and lift chart. SAS Global Forum.
- K.A. University, 2016. Deanship of e-learning and distance education twitter account.
- Khan, F.H. S. Bashir and U. Qamar, 2014. TOM: Twitter opinion mining framework using hybrid classification scheme. Decision Support Syst., 57: 245-257. DOI: 10.1016/j.dss.2013.09.004
- Koppel, M. and J. Schler, 2006. The importance of neutral examples for learning sentiment. Comput. Intell., 22: 100-109.  
 DOI: 10.1111/j.1467-8640.2006.00276.x
- Liu, B., 2012. Sentiment analysis and opinion mining. Synthesis Lect. Human Lang. Technol., 5: 1-167.
- Martin, V., 2013. Predicting the French stock market using social media analysis. Proceedings of the 8th International Workshop on Semantic and Social Media Adaptation and Personalization, Dec. 12-13, IEEE Xplore Press, pp: 3-7.  
 DOI: 10.1109/SMAP.2013.22
- Mourtada, R. and F. Salem, 2014. Citizen Engagement and Public Services in the Arab World: The Potential of Social Media. 6th Edn., Arab Social Media Report Series.
- Nasukawa, T. and J. Yi, 2003. Sentiment analysis: Capturing favorability using natural language processing. Proceedings of the 2nd International Conference on Knowledge Capture, Oct. 23-25, ACM, Sanibel Island, FL, USA, pp: 70-77.  
 DOI: 10.1145/945645.945658
- O'Connor, B., R. Balasubramanian, B.R. Routledge and N.A. Smith, 2010. From tweets to polls: Linking text sentiment to public opinion time series. Proceedings of the 4th International Conference on Weblogs and Social Media, May 23-26, Washington, DC, USA., pp: 1-2.
- Pak, A. and P. Paroubek, 2010. Twitter as a corpus for sentiment analysis and opinion mining. Proceedings of the 7th conference on International Language Resources and Evaluation, (LRE' 10), pp: 1320-1326.

- Pang, B. and L. Lee, 2008. Opinion mining and sentiment analysis. *Foundat. Trends Informat. Retrieval*, 2: 1-135. DOI: 10.1561/15000000011
- Rapidminer, 2016. Available: <https://rapidminer.com/>
- Rushdi-Saleh, M., M.T. Martín-Valdivia, L.A. Ureña-López and J.M. Perea-Ortega, 2011. OCA: Opinion corpus for Arabic. *J. Am. Society Inform. Sci. Technol.*, 62: 2045-2054. DOI: 10.1002/asi.21598
- Shoukry, A. and A. Rafea, 2012. Sentence-level Arabic sentiment analysis. *Proceedings of the International Conference on Collaboration Technologies and Systems*, May 21-25, IEEE Xplore Press, pp: 546-550. DOI: 10.1109/CTS.2012.6261103
- Tian, F., Q. Zheng, R. Zhao, T. Chen and X. Jia, 2009. Can e-Learner's emotion be recognized from interactive Chinese texts? *Proceedings of the 13th International Conference on Computer Supported Cooperative Work in Design*, Apr. 22-24, IEEE Xplore Press, pp: 546-551. DOI: 10.1109/CSCWD.2009.4968116
- Tripathi, G. and S. Naganna, 2015. Feature selection and classification approach for sentiment analysis. *Mach. Learn. Applic.: Int. J.*, 2: 1-16. DOI: 10.5121/mlaij.2015.2201
- Tumasjan, A., T.O. Sprenger, P.G. Sandner and I.M. Welp, 2010. Predicting elections with twitter: What 140 characters reveal about political sentiment. *Proceedings of the 4th International Conference on Weblogs and Social Media*, May 23-26, Washington, DC, USA, pp: 178-185.
- Wang, P., M. Zuo and L. Diao, 2010. Design of the monitoring and analysis system of the Internet Education Public Sentiment. *Proceedings of the 2nd International Conference on Future Computer and Communication*, May 21-24, IEEE Xplore Press, pp: V2-571-V2-574. DOI: 10.1109/ICFCC.2010.5497529
- Wang, W., 2010. Sentiment analysis of online product reviews with Semi-supervised topic sentiment mixture model. *Proceedings of the 7th International Conference on Fuzzy Systems and Knowledge Discovery*, Aug. 10-12, IEEE Xplore Press, pp: 2385-2389. DOI: 10.1109/FSKD.2010.5569528
- Zhang, W. and S. Skiena, 2010. Trading strategies to exploit blog and news sentiment. *Proceedings of the 4th International Conference on Weblogs and Social Media*, May 23-26, Washington, DC, USA.