Investigation of Naive Bayes Combined with Multilayered Perceptron for Arabic Sentiment Analysis and Opinion Mining

Shakir Mrayyen, Mohammad Subhi Al-Batah, Malek Alzaqebah

Abstract—Sentiment analysis and opinion mining has recently become active research area, which associated with studying the people's opinions, emotions, and evaluation from a written text. The rapid growing of social networks increases the importance of sentiment analysis; in which people are always sharing their opinion about several subjects and topics over the internet. Therefore, it will be useful to classify people's reviews, discussions, and blogs to identify their opinion of specific product, movie or hotel ether is a positive or negative, to farther helps companies or other people for decision making. Sentiment analysis for English language has been well studied. Conversely, the work that has been carried out in terms of Arabic remains in its infancy; thus, more cooperation is required between research communities in order for them to offer a mature sentiment analysis system for Arabic. In this paper, the Naive Bayes (NB) algorithm and Multilayered Perceptron (MLP) network are combined with hybrid system for Arabic sentiment classification. The dataset are then classified into positive or negative polarities of sentiment using the proposed system. A comparison with many classifiers for the tweets dataset were tested. As a result, the performance of the combined MLP ranked first (99.5%) compared to other classifiers with recorded accuracy of 93.0%, 84.9%, 84.1%, 88.2%, 80.7%, and 78.1% for NB, Decision Table, JRip, OneR, PART, and ZeroR, respectively.

Keywords—Sentiment Analysis, Opinion Mining, Arabic Language Classification, Naive Bayes Algorithm, Multilayered Perceptron Network.

Shakir Mrayyen and Mohammad Subhi Al-Batah are with the Faculty of Science and Information Technology, Jadara University, Irbid, Jordan.
Malek Alzaqebah is with the Department of mathematics, Faculty of Science, Imam Abdulrahman Bin Faisal University, Dammam, Saudi Arabia.

I. INTRODUCTION

The sentiment classification is a task of classifying the people's text of reviews by performing Text Analysis techniques and Natural Language Processing to determine whether the people's attraction about some topic (product, hotel, or movie, etc.) is positive, negative, or neutral. Sentiment analysis and opinion mining recently become very popular and active research, that due to the importance of the sentiment analysis in various fields, such as in politic [1], marketing [2], education [3], healthcare [4] etc. In addition, sentiment analysis plays an important role in making decisions for individuals, organizations or governments; an appropriate decision can be made wisely by getting the opinion from others [5]. Many researchers have been studied the sentiment analysis in terms of English language, nonetheless the research work in terms of Arabic language still insufficient. Therefore, more research work of Arabic is required [6]. The authors Farra et al. in [7] used sentiment mining of Arabic text at both sentence level and document level. The study works in a grammatical approach and a semantic approach for sentence-level sentiment mining in Arabic text, and a document-level, also they construct semantic dictionary includes Arabic roots. The results showed high accuracy in grammatical approach. El-Beltagy et al. [8] studied the possibility of determining the semantic orientation in terms of Arabic Egyptian. The authors used two datasets from tweeter, first datasets contained 500 tweets, classified into 155 positive tweets, 310 negative tweets, and 35 neutral tweets; second dataset is the Dostour dataset which has 100 random comments. The comments were classified 40 positive, 38 negative, and 22 neutral.

Methodology proposed by Rushdi-Saleh et al. [9] proposes Arabic corpus for opinion mining, which collected from Arabic reviews of web pages related to movies and films, the authors used Support Vector Machines (SVM) and Naïve Bayes (NB) algorithms to classify the polarity of Arabic reviews. The authors have translated the Opinion Corpus for Arabic (OCA) corpus into English, generating the EVOCA corpus (English Version of OCA). The result has shown that the accuracy of EVOCA is low because of translation. Refaee et al. [10] used gold-standard annotated corpus. The corpus

contained 8,868 tweets collected manually at different time. Test set (1,365 tweets) development set (7,503 tweets). The result of the analysis showed the top of five frequent words. The researchers Al-Kabi et al. [11] used lexicon-based tool to analyze Arabic reviews and comments by using MSA language. This tool is able to identify the polarity, intensity and subjectivity, of Arabic reviews. The result of the study is a tool that more accurate for Arabic domain. The work of Cheong et al. [12] was done for a tweets collected from Twitter during the Australian floods in 2010-2011. The authors used the social network analysis techniques to generate and analyze the online networks traffic that arisen at that time. The SNA metrics is used to identify influential members on online communities. However, one limitation of it is the insufficient number of tweets collected for the purpose of generalizing the findings. Kouloumpis et al. [13] studied the effect of linguistic features for detecting the sentiment of Twitter messages and evaluated the usefulness of existing lexical resources. The researchers used three datasets (hashtag, emotion and iSieve) from twitter. Different features are used to classify data, use bigrams, unigrams and sentiment analysis, mainly features representing information from a sentiment lexicon and POS features.

Duwairi and Qarqaz [14] used crowd sourcing to collect a large dataset of tweets and developed a framework for analyzing Tweets to detect the positive, negative or neutral sentiments. However, there was a need to expand the dictionaries. El-Makky et al. [15] designed a novel subjective and sentiment analysis system for Arabic tweets by building a new Arabic lexicon and merging Modern standard Arabic lexica with two Egyptian Arabic lexica. The result of novel hybrid approach showed good result compared to the previous work. Shoukry and Rafea [16] used sentence level sentiment analysis for Arabic. 1000 tweets are collected from twitter and classified by using two algorithms in Weka program, Support Vector Machine (SVM) and naïve Bayes (NB) algorithms. The result showed that SVM accuracy was higher than NB. However, the study ignored neutral tweets. The study by Butgereit [17] aims to monitoring Twitter of weather status for a specific city. It used crowd sourced and collected almost 600 tweets from twitter and used the μ Model to analyze the tweets. The work by Rushdi-Saleh et al. [18] used verity web pages Arabic and blogs to build a small Arabic opinion corpus consists of 500 movies, 250 positive and 250 negative opinion using two machine learning; support vector machines algorithm and Naïve Bayes algorithm. The accuracy of the NB classifier was 84% and the accuracy of SVM classifier was 90%. However, there was a need to expand the corpus. The work by Siqi zhao [19] developed Sport Sense to detect sport events targeted for the National Football League (NFL) games. The results showed that major game events can be accurately and effectively extracted by using open access Twitter data. Sport Sense API is accessible for developers in order to create Twitter-enabled applications. The application was tested and used by the National Football League (NFL) games as a targeted domain.

Vieweg et al. in their research [20] analyzed the tweet posts in emergency events in North America during the Oklahoma grass fires and Red River Floods in spring 2009 by using Twitter Search API and geo-location. The aim of the study is to improve situational awareness in emergency situations. The authors Godfrey et al. [21] used k-means and the Non-Negative Matrix Factorization (NMF) to cluster the topics of number of tweets, the results demonstrate that k-means and NMF algorithms produces comparable results, but NMF showed to be faster. Duwairi in [22] introduced a framework to sentiment analysis on tweets written in Modern Standard Arabic (MSA) and Jordanian Dialectical Arabic (JDA). The researcher produced dialect lexicon which maps the words in JDA into their corresponding MSA words. The dataset consists of 22550 tweets, NB and SVM classifiers were used to determine the polarity of the tweets. Also, Korayem et al. in [23] presents a survey of different techniques for subjectivity and sentiment analysis (SSA) for Arabic. These surveys describe the main existing techniques and test corpora for Arabic SSA that has been introduced in the literature. The researchers Ibrahim and Salim [24] used 65 studies to determine features and methods used for twitter opinion mining. The results showed n-grams features is commonly used for twitter sentiments analysis and also used for Arabic tweets. The study also reported that the methods most common used is the Lexical based classification using Support Vector Machines (SVM) and Naive Bayes (NB). In addition, Al-Kabi et al. [25] used 4625 reviews/comments collect from Yahoo and Maktoob social media to determine length and likes/dislikes reviews. The dataset contains MSA language and various Arabic dialects. The result shows the accuracy of SVM (68.2%) and the accuracy of NB (61.43%).

II.   ARABIC SENTIMENT ANALYSIS

Sentiment analysis is a part of machine learning and natural language processing (NLP), which consists of extracting the sentiment, emotion or opinion of people about a specific topic from their writing. Many terms have been used throughout the literature to indicate the sentiment analysis, such as subjectivity analysis, opinion mining, review mining, emotion detection from text and appraisal extraction [26]. Furthermore, the sentiment analysis of the text can either be explicit or implicit aspects. In explicit aspect, a sentiment is given directly, such as (إنها مقالة جيدة / It is a good article), while in implicit aspect; the sentiment is implied in the text, such as (لها كثير من الاستشهادات / it has a lot of citations). Sentiment analysis can be defined formally as:

"Given a text t from a text set T, computationally assigning polarity labels p from a set of polarities P in such a way that p would reflect the actual polarity that is found in T" [27].

Sentiment analysis is an active research area, also known as text mining. Sentiment analysis can be powerful tool for decision making applied in online sentiments reviews or tweets [28]. Sentiment analysis of tweets is a powerful application of mining social media sites that can be used for a variety of social sensing tasks [29].

Tweets dataset are considered in our experiment. Twitter dataset for Arabic sentiment analysis were collected from tweeter on various topics such as: politics and arts, which include 2000 labelled tweets (1000 positive tweets and 1000 negative ones). These tweets include opinions written in both Modern Standard Arabic (MSA) and the Jordanian dialect [30].

III. NAIVE BAYES CLASSIFIER

Naive Bayes (NB) classifier is a classifier based probabilistic method that can be categorized as supervised classification method. NB applies Bayes' theorem with strong independence rules, which approved to be fast, accurate, effective and simple, in text classification [31]. The NB classifier is an algorithm that can be applied to problems that associating an object with a discrete category.

To perform text classification method, it's required to extract features from given text. Features are stored in a vector to be analyzed by the classifier, which called the feature vector or Term vector. The vector is generated based on vocabulary from the training dataset with unique words. The methods used by NB text classification can be categorized to two as presented in [32] by Mccallum and Nigam, the multinomial model and the multivariate Bernoulli model. Both methods are performed by applying the Bayes rule as follow [33].

$$P(c_i \mid d_j) = \frac{P(c_i) * P(d_j \mid c_i)}{P(d_j)} \qquad (1)$$

where, c_i specific class that might be positive or negative

  d_j is the text to be classified
  $P(c_i)$ and $P(d_j)$ are the Prior probabilities
  $P(c_i \mid d_j)$ and $P(d_j \mid c_i)$ are the Posterior probabilities

The multinomial model is used in this study. That due to in text classification the multinomial model outperforms the multivariate Bernoulli model [34].

NB multinomial model calculates the word frequency in a given text [35]. The Maximum Likelihood Estimate (MLE) estimates the parameters of given training data based on relative frequency [36]. MLE used the likelihood function to find the parameter values that maximize the most likely value. For the prior probability, the estimation is given in Equation 2.

$$P(c_i) = \frac{Nc}{N} \qquad (2)$$

where Nc is the number of documents in the class c_i
  and N is the total number of documents.

Multinomial model considers that all attributes are independent for a given context of class. The drawback of MLE occurs if the number of words in the class c_i is zero, which means the word did not occurred in the training data. In other words the training data are insufficient to represent the

frequency of rare occurrence words. In order to overcome this drawback and avoid zero in frequency of words, add-one smoothing method is used (or Laplace smoothing), this method simply adds one to each word count. The probability $P(W_k/c_i)$ is a multinomial distribution presented in Equation 3, which shows the relative frequency of the word W in class c_i.

$$P(W_k \mid c_i) = \frac{n_k(c_i) + 1}{n(c_i) + |V|} \qquad (3)$$

Where: $n(c_i)$ is the number of words belonging to a class c_i (which can be pos or neg), $n_k(c_i)$ is the number of occurrences of word k in class c_i. V is the vocabulary which stores the unique words |V|.

Then the classification of a new given text will be based on the value of naïve-bayes classifier V_{NB}, which shown in Equation 4.

$$V_{NB} = \text{argmax}_{c_j \in v} \ P(c_j) \prod_{w \in words} P(w_i \mid c_j) \qquad (4)$$

Where, c_j is the jth class and w_i is the ith word.

To show how Naive Bayes classifier works, can best be explained by providing an illustrative example. Our target is to classify whether a new tweet has positive attitude or negative attitude. Let's consider we have number of classified tweets on a topic as training set as in table I.

Table I: Set of classified tweets

No	tweets	Translation	class
1	فكرة جيده جداً	good idea	pos
2	انا مهتم جداً	I'm very interested	pos
3	انه سيئ انا غيرمهتم	It is bad and I'm not interested	neg
4	تدعو للعنصريه	Reinforce Racist	neg

Now we will use this example to estimate the probability of the positive and negative attitudes and the probability of words in each tweet for the positive and negative attitudes as follow:

- P(pos)
- P(neg)
- P(word in doc|pos) positive probability of word
- P(word in doc|neg) negative probability of word

There are 10 unique Arabic words to be stored in a vector of words. This stores all unique words for each tweet.
 Note that: - The words "انا" and "جداً" occurred twice and each will be counted as one word.
    - The word "غير مهتم" is the antonym of "مهتم" will be considered one word because it contains "غير" ["Not"].

Training Part:
Step 1: calculate priori probability of pos and neg by using Equation 2 as follow:
  p (pos) = 2/4=0.5 (two positive out of four)

p (neg) = 2/4 =0.5 (two negative out of four)

Step 2: calculate the individual probability of all possible words with positive attitude and negative attitude based on likelihood smoothing NB estimate using Equation 3 as follow:

For positive attitude: as in table I the word "فكرة" occurred once in pos class, and 1 for likelihood smoothing, the value will be (1+1). The number of unique words is 5 in pos class and the number of all unique words is 10, so the value will be (5+10), do the same for all unique words as follow:

P (فكرة| pos) = (1+1)/ (5+10) = 0.13
P (جيده| pos) = (1+1)/ (5+10) = 0. 13
P (جداً| pos) = (2+1)/ (5+10) = 0.2 (occurred twice)
P (انا| pos) = (1+1)/ (5+10) = 0. 13
P (مهتم| pos) = (1+1)/ (5+10) = 0. 13
P (انه| pos) = (0+1)/ (5+10) = 0.06 (no occurrence)
P (سيئ| pos) = (0+1)/ (5+10) = 0.06
P (غير مهتم| pos) = (0+1)/ (5+10) = 0.06
P (تدعو| pos) = (0+1)/ (5+10) = 0.06
P (عنصريه| pos) = (0+1)/ (5+10) = 0.06

For negative attitude:
P (فكرة| neg) = (0+1)/ (6+10) = 0.06
P (جيده| neg) = (0+1)/ (6+10) = 0.06
P (جداً| neg) = (0+1)/ (6+10) = 0.06 (occurred twice)
P (انا| neg) = (1+1)/ (6+10) = 0.12
P (مهتم| neg) = (0+1)/ (6+10) = 0.06
P (انه| neg) = (1+1)/ (6+10) = 0.12 (no occurrence)
P (سيئ| neg) = (1+1)/ (6+10) = 0.12
P (غير مهتم| neg) = (1+1)/ (6+10) = 0.12
P (تدعو| neg) = (1+1)/ (6+10) = 0.12
P (عنصريه| neg) = (1+1)/ (6+10) = 0.12

Testing Part:
Consider the testing sentence (TS) is:
" مهتم جداً هذه فعاليات جيده "
[Note that, it's Positive sentence]

Table II shows the probability for each word in pos and neg classes taken from **Step 2**:

Table II: probability for each word

	جيده	فعاليات	هذه	جداً	مهتم
Pos probability for each word	0. 13	-	-	0.2	0. 13
Neg probability for each word	0.06	-	-	0.06	0.06

"-" means not exist in the vector of words

Calculate the V_{NB} for the testing sentence (TS) using Equation 4.
p (pos | TS)= [p (pos) calculated in step 1 * Pos probability for each word in TS]
p (pos | TS) = 0.5 * 0.13 * 0.2 * 0.13 = 0.00169

p (neg | TS)= [p (neg) calculated in step 1* Neg probability for each word in TS]
p (neg | TS) = 0.5 *0.06* 0.06* 0.06= 0.000108

Using Bayes' rule, we label the new tweet TS with pos class that achieves the highest probability, since the value of p (pos | TS) is greater than the value of p (neg | TS), so the testing sentence has positive attitude.

IV. CLASSIFICATION

In this paper, the MLP, Decision Table, JRip, OneR, PART, and ZeroR classifiers are proposed for tweets classification. Since the inputs of tweets are text, it was not possible for these classifiers to recognize and categorize the text data. Only, the NB has the ability for classifying the polarity of the text dataset. Thus, a combined method for NB with the other classifiers is proposed. In the combined method, the weight of the tweets and the predicted output using NB is automatically feed to next classifier.
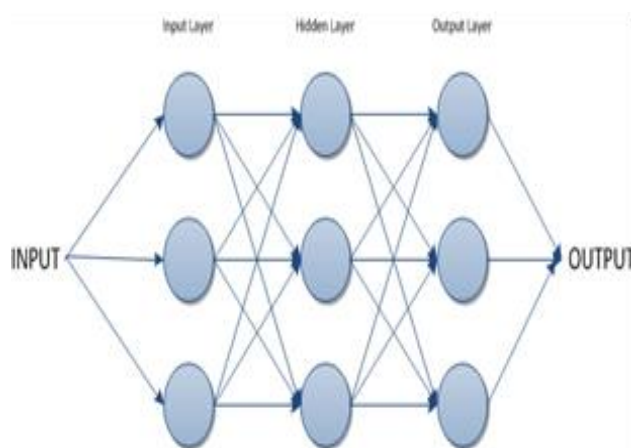

Figure I: A Multilayer Perceptron

- Decision Table algorithm builds and using a simple decision table majority classifier as proposed by Kohavi [37]. Decision Table employs the wrapper method to find a good subset of attributes for inclusion in the table.
- Java Repeated Incremental Pruning (JRip) this implements a propositional rule learner. Repeated Incremental Pruning to Produce Error Reduction (RIPPER) [38].
- One Rule (OneR) this algorithm uses the minimum-error attribute for prediction, and discretizing numeric attributes [39].
- Partial C4.5 (PART) is a separate and conquer rule learner. Builds a partial C4.5 decision tree in each iteration and makes the "best" leaf into a rule [40].
- 0-R classifier (ZeroR) this algorithm predicts the mean (for a numeric class) or the mode (for a nominal class) [41].

- Multilayer perceptron (MLP) uses a feedforward architecture and can have multiple hidden layers [42-44]. MLP uses dot products between inputs and weights and sigmoidal activation functions (or other monotonic functions such as rectified linear unit (ReLU)). The training is usually done through back-propagation for all layers [45-47]. This type of neural network is used in deep learning with the help of many techniques such as dropout or batch normalization [48-50]. The structure of the three layered back-propagation neural network is shown in Figure I.

I. RESULT AND DISSCUSSION

The proposed approach combined the NB with many classifier for Arabic sentiment classification. Six classifiers were tested; MLP, Decision Table, JRip, OneR, PART, and ZeroR. This study employs many methods to arrange the number of the data for training and testing sets [51]. The methods for test options are full training set, cross-validation (2 folds to 10 folds), and percentage split (10%-99%) [52-54]. From the experiments, the best accuracy is considered for each classifier.
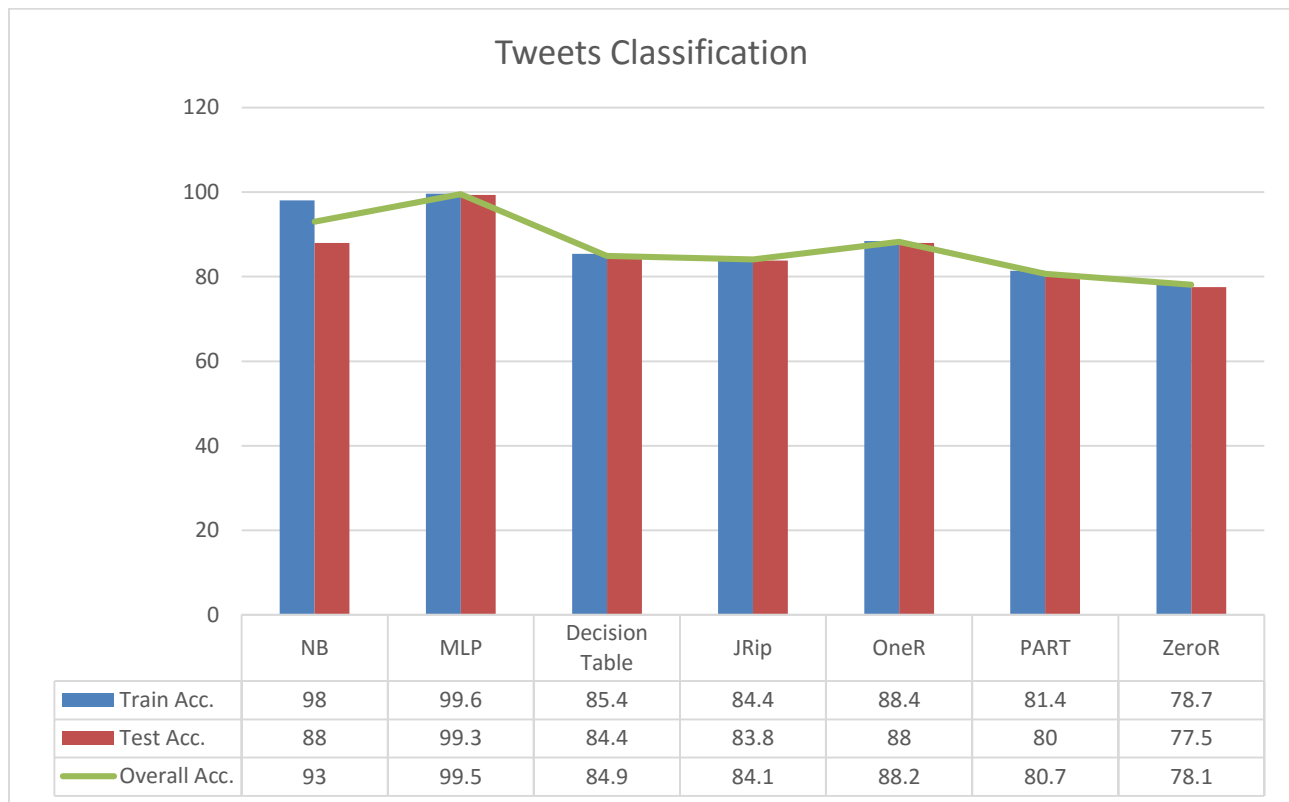

Tweets Classification

	NB	MLP	Decision Table	JRip	OneR	PART	ZeroR
Train Acc.	98	99.6	85.4	84.4	88.4	81.4	78.7
Test Acc.	88	99.3	84.4	83.8	88	80	77.5
Overall Acc.	93	99.5	84.9	84.1	88.2	80.7	78.1

Figure II: Accuracy for tweets different classifiers

Figure II shows the classification accuracy of tweets using the NB and the combined MLP, Decision Table, JRip, OneR, PART, and ZeroR. Based on the results obtained, the combined MLP produces better training accuracy compared to other classifiers. The training accuracy are 98.0%, 99.6%, 85.4%, 84.4%, 88.4%, 81.4%, and 78.7% for NB, MLP, Decision Table, JRip, OneR, PART, and ZeroR, respectively. While the test accuracy are 88.0%, 99.3%, 84.4%, 83.8%, 88.0%, 80.0%, and 77.5% for NB, MLP, Decision Table, JRip, OneR, PART, and ZeroR, respectively. Also, the result shows that the combined MLP provides better overall accuracy 99.5%, as compared with NB (93.0%), Decision Table (84.9%), JRip (84.1%), OneR (88.2%), PART (80.7%), and ZeroR (78.1%).

In addition, the results in Figure II show that the combined MLP is able to achieve better classification performance than other classifiers. The combined MLP outperformed the others

in terms of the percentage of training accuracy by more than 1.6% for NB, 14.2% for Decision Table, 15.2% for JRip, 11.2% for OneR, 18.2% for PART and 20.9% for ZeroR. Also, the combined MLP outperformed the other classifiers with difference of testing accuracy percentage equal to 11.3%, 14.9%, 15.5%, 11.3%, 19.3% and 21.8% for NB, Decision Table, JRip, OneR, PART, and ZeroR, respectively. The outcomes consistently demonstrate the effectiveness of the combined MLP for tackling Arabic sentiment classification tasks.

II. CONCLUSION

Sentiment analysis plays vital role in making decisions. In this study, an experiment is conducted for the tweets using the NB and the combined MLP, Decision Table, JRip, OneR, PART, and ZeroR. The tweets are classified into positive or negative attitudes. The accuracy of sentiment analysis is

improved using the combined MLP. Over the whole classifiers, the combined MLP outperformed the others in terms of the percentage of overall accuracy by more than 6.5% for NB, 14.6% for Decision Table, 15.4% for JRip, 11.3% for OneR, 18.8% for PART and 21.4% for ZeroR. These results show the effectiveness of the combined MLP as compared with other classifiers. In our future work, we targeted to apply this method on other domains with feature selection schemes to select features that carry sentiment from those that do not.

REFERENCES

[1] Tumasjan, A., Sprenger, T. O., Sandner, P. G., & Welpe, I. M. (2010). Predicting elections with twitter: What 140 characters reveal about political sentiment. Icwsm, 10(1), 178-185.

[2] Rambocas, M., & Gama, J. (2013). Marketing research: The role of sentiment analysis (No. 489). Universidade do Porto, Faculdade de Economia do Porto.

[3] Pang, B., & Lee, L. (2004, July). A sentimental education: Sentiment analysis using subjectivity summarization based on minimum cuts. In Proceedings of the 42nd annual meeting on Association for Computational Linguistics (p. 271). Association for Computational Linguistics.

[4] Greaves, F., Ramirez-Cano, D., Millett, C., Darzi, A., & Donaldson, L. (2013). Use of sentiment analysis for capturing patient experience from free-text comments posted online. Journal of medical Internet research, 15(11).

[5] Pawar, A. B., Jawale, M. A., & Kyatanavar, D. N. (2016). Fundamentals of Sentiment Analysis: Concepts and Methodology. In Sentiment Analysis and Ontology Engineering (pp. 25-48). Springer International Publishing.

[6] Alotaibi, S. S. (2015). Sentiment analysis in the Arabic language using machine learning (Doctoral dissertation, Colorado State University).

[7] Farra, Noura, et al. "Sentence-level and document-level sentiment mining for Arabic texts." Data Mining Workshops (ICDMW), 2010 IEEE International Conference on. IEEE, 2010.

[8] El-Beltagy, Samhaa R., and Ahmad Ali. "Open issues in the sentiment analysis of Arabic social media: A case study." Innovations in information technology (iit), 2013 9th international conference on. IEEE, 2013.

[9] Rushdi-Saleh, Mohammed, et al. "Bilingual experiments with an arabic-english corpus for opinion mining." (2011).

[10] Refaee, Eshrag, and Verena Rieser. "An Arabic twitter corpus for subjectivity and sentiment analysis." Proceedings of the Ninth International Conference on Language Resources and Evaluation (LREC'14), Reykjavik, Iceland, may. European Language Resources Association (ELRA). 2014.

[11] Al-Kabi, M. N., Alsmadi, I. M., Gigieh, A. H., Wahsheh, H. A., & Haidar, M. M. (2014). Opinion Mining and Analysis for Arabic Language. IJACSA) International Journal of Advanced Computer Science and Applications, 5(5), 181-195.

[12] Cheong, France, and Christopher Cheong. "Social Media Data Mining: A Social Network Analysis Of Tweets During The 2010- 2011 Australian Floods." PACIS11 (2011): 46-46.

[13] Kouloumpis, Efthymios, Theresa Wilson, and Johanna Moore. "Twitter sentiment analysis: The good the bad and the omg!." Icwsm 11 (2011): 538-541.

[14] Duwairi, R. M., & Qarqaz, I. (2014, August). Arabic sentiment analysis using supervised classification. In Future Internet of Things and Cloud (FiCloud), 2014 International Conference on (pp. 579-583). IEEE.

[15] El-Makky, N., Nagi, K., El-Ebshihy, A., Apady, E., Hafez, O., Mostafa, S., & Ibrahim, S. (2015). Sentiment Analysis of Colloquial Arabic Tweets.

[16] Shoukry, Amira, and Ahmed Rafea. "Sentence-level Arabic sentiment analysis." Collaboration Technologies and Systems (CTS), 2012 International Conference on. IEEE, 2012.

[17] Butgereit, Laurie. "Crowd sourced weather reports: an implementation of the µ model for spotting weather information in twitter." ist-africa conference proceedings, 2014. ieee, 2014.

[18] Rushdi_Saleh, Mohammed et al. "OCA: Opinion corpus for Arabic." Journal of the American Society for Information Science and Technology 62.10 (2011): 2045-2054.

[19] Siqi zhao, Detecting events from twitter in Real time, 2013.

[20] Vieweg, Sarah, et al. "Microblogging during two natural hazards events: what twitter may contribute to situational awareness." Proceedings of the SIGCHI conference on human factors in computing systems. ACM, 2010.

[21] Godfrey, Daniel, et al. "A Case Study in Text Mining: Interpreting Twitter Data From World Cup Tweets." arXiv preprint arXiv:1408.5427 (2014).

[22] Duwairi, Rehab M. "Sentiment analysis for dialectical Arabic." Information and Communication Systems (ICICS), 2015 6th International Conference on. IEEE, 2015.

[23] Korayem, Mohammed, David Crandall, and Muhammad AbdulMageed. "Subjectivity and sentiment analysis of arabic: A survey." Advanced Machine Learning Technologies and Applications. Springer Berlin Heidelberg, 2012. 128-139.

[24] Ibrahim, M A, and Salim N. "OPINION ANALYSIS FOR TWITTER AND ARABIC TWEETS: A SYSTEMATIC LITERATURE REVIEW."Journal of Theoretical & Applied Information Technology 56.3 (2013).

[25] Al-Kabi, Mohammed N., Nawaf A. Abdulla, and Mahmoud AlAyyoub. "An analytical study of arabic sentiments: Maktoob case study." Internet Technology and Secured Transactions (ICITST), 2013 8th International Conference for IEEE, 2013.

[26] B. Pang and L. Lee. Opinion mining and sentiment analysis. Foundations and trends in information retrieval, volume 2(1-2): pages 1–135, Jan. 2008. ISSN 1554-0669.

[27] ElSahar, H., & El-Beltagy, S. R. (2015, April). Building Large Arabic Multi-domain Resources for Sentiment Analysis. In CICLing (2) (pp. 23-34).

[28] Nabil, M., et al.: LABR: A Large Scale Arabic Book Reviews Dataset. arXiv Prepr. arXiv1411.6718 (2014).

[29] Pang, B., et al.: Thumbs up? Sentiment Classification using Machine Learning Techniques. In: Conf. Empir. Methods Nat. Lang. Process. (EMNLP 2002), pp. 79–86 (2002).

[30] Abdulla, N., Mahyoub, N., Shehab, M., & Al-Ayyoub, M. (2013). Arabic sentiment analysis: Corpus-based and lexicon-based. In Proceedings of The IEEE conference on Applied Electrical Engineering and Computing Technologies (AEECT).

[31] C. Tseng, N. Patel, H. Paranjape, T. Y. Lin, S. Teoh, "Classifying twitter data with naive bayes classifier," in IEEE International Conference on Granular Computing, 2012.

[32] Andrew Mccallum, Kamal Nigam. "A comparison of event models for naive bayes text classification". In AAAI-98 Workshop on Learning for Text Categorization. Pages 41-48. 1998.

[33] K. M. Leung, "Naive Bayesian classifier," [Online]. Available: http://www.sharepdf.com/81fb247fa7c54680a94dc0f3a253fd85/naiveBayesianClassifier.pdf, [Accessed: January 2018].

[34] H. Shimodaira, "Text classification using naive bayes," pp. 1-5, [Online].Available: http://www.inf.ed.ac.uk/teaching/courses/inf2b/learn notes/inf2blearn-note07-2up.pdf, [Accessed: January 2018].

[35] M.Radovanovic, M. Ivanovic, "Text mining: Approaches and applications," vol. 38, no. 3, pp. 227-234, [Online]. Available:http://www.emis.de/journals/NSJOM/Paper s/383/NSJOM 3[Accessed: January 2018].

[36] Lina L. Dhande and Girish K. Patnaik. Analyzing Sentiment of Movie Review Data using Naïve Bayes Neural Classifier. International Journal of Emerging Trends & Technology in Computer Science (IJETTCS). Volume 3, Issue 4 July-August 2014, Pages 313-320.

[37] Ron Kohavi: The Power of Decision Tables. In: 8th European Conference on Machine Learning, 174-189, 1995.

[38] R.C. Holte (1993). Very simple classification rules perform well on most commonly used datasets. Machine Learning. 11:63-91.

[39] Eibe Frank, Ian H. Witten: Generating Accurate Rule Sets Without Global Optimization. In: Fifteenth International Conference on Machine Learning, 144-151, 1998.

[40] Inamdar S., Narangale S., and Shinde G., (2011) Preprocessor agent approach to knowledge discovery using Zero-R algorithm, International Journal 2, 2011.

[41] B.R. Gaines and P. Compton. Induction of ripple-down rules applied to modeling large databases.

[42] Klaib, M. F. J., Al-batah, M. S., and Rasrasc, R. J., 3-way Interaction Testing using the Tree Strategy, Journal of Procedia Computer Science, International Conference on Communication, Management and Information Technology (ICCMIT 2015), Procedia Computer Science 65(2015), 845–852.

[43] Al-batah, M. S. (2014) Testing the Probability of Heart Disease Using Classification and Regression Tree Model, Annual Research & Review in Biology, ISSN: 2231-4776, 4(11): 1713-1725, 2014, SCIENCEDOMAIN international, GURGAON, INDIA and USA.

[44] Alkhasawneh, M. Sh., Ngah, U., Tay, L. T., Al-batah, M. S., and Mat Isa, N. A. (2014) Intelligent Landslide System Based on Discriminant Analysis and Cascade-Forward Back-Propagation Network, Arabian Journal for Science and Engineering, ISSN: 1319-8025, Springer, HEIDELBERG, GERMANY.

[45] Baareh A.K., Sheta A.F., Al-Batah, M. S. (2012) Feature based 3D Object Recognition using Artificial Neural Networks, International Journal of Computer Applications, ISSN: 0975-8887, 44(5):1-7, April 2012. USA.

[46] Mat Isa, N. A., Sani, Z. M., Al-Batah, M. S. (2011) Automated Intelligent real-time system for aggregate classification, International Journal of Mineral Processing, ISSN: 0301-7516, 100(1-2): 41-50, 2001, Science direct, Elsevier, AMSTERDAM, NETHERLANDS.

[47] Al-batah, M. S., Alkhasawneh, M. Sh., Tay, L. T., Ngah U., Lateh, H. H., and Mat Isa, N. A., Landslide Occurrence Prediction Using Trainable Cascade Forward Network and Multilayer Perceptron, Hindawi Publishing Corporation, Mathematical Problems in Engineering, Volume 2015, Article ID 512158, 9 pages.

[48] Alkhasawneh, M. Sh., Ngah, U., Tay, L. T., Al-batah, M. S., and Mat Isa, N. A. (2013) Determination of Important Topographic Factors for Landslide Mapping Analysis Using MLP Network, Scientific World Journal, ISSN: 1537-744X, Volume 2013, Article ID 415023, 12 pages, Hindawi Publishing Corporation, NEW YORK, USA.

[49] Mat Isa, N. A., Al-Batah, M. S., Zamli, K. Z., Azizli, K. A., Joret, A., Mat Noor, N. R. (2008) Suitable features selection for the HMLP and MLP networks to identify the shape of aggregate, Construction and Building Materials, ISSN: 0950-0618, Volume 22, Issues 3, Pages 402-410.

[50] Mat-Isa, N.A., Joret, A., Al-Batah, M.S., Ali A.N., Zamli, K.Z. and Azizli, K.A. (2006) Microcontroller Based HMLP Realization for Aggregate Classification System, International Journal of Factory Automation, Robotics and Soft Computing, ISSN: 1828-6984, Issue 2, Pages 19-26.

[51] Quteishat, A., al-batah, M., al-mofleh, A., Alnabelsi, S. H. (2013) Cervical Cancer Diagnostic System Using Adaptive Fuzzy Moving K-means Algorithm and Fuzzy MIN-MAX Neural Network, Journal of Theoretical and Applied Information Technology, ISSN: 1992-8645, 57(1):48-53, 2013, Little Lion Scientific Islamabad Pakistan, Pakistan and USA.

[52] Mat Noor, N. R., Mat Isa, N. A., Al-Batah, M. S., Automatic glass-slide capturing system for cervical cancer pre-screening program, American Journal of Applied Sciences, Volume 5, Issue 5, Pages 461-467.

[53] Md-Sani, Z., Mat-Isa, N.A., Suandi, S.A., Al-Batah, M.S., Zamli, K.Z. and Azizli, K.A. (2007) Intelligent Rock Vertical Shaft Impact Crusher Local Database System, International Journal of Computer Science and Network Security, ISSN : 1738-7906, Volume 7, Issues 6, Pages 57-62.

[54] Al-Batah, M.S., Mrayyen S., Alzaqebah M. (2018) Arabic Sentiment Classification using MLP Network Hybrid with Naive Bayes Algorithm, Journal of Computer Science, Volume 14, Issue 8, Pages 1104-1114.