Sentiment Analysis of Social Media Content in Pashto Language using Deep Learning Algorithms

Saqib Iqbal[1], Farhad Khan[2], Hikmat Ullah Khan[2*], Tassawar Iqbal[2], Jamal Hussain Shah[2]

[1] College of Engineering, Al Ain University, United Arab Emirates
[2] Department of Computer Science, COMSATS University Islamabad, Wah Campus, Wah Cantt, Pakistan
saqib.iqbal@aau.ac.ae, farhadmohmand22@gmail.com, hikmat.ullah@ciitwah.edu.pk, tassawar@ciitwah.edu.pk, jhshah@ciitwah.edu.pk

Abstract

Sentiment Analysis (SA) has become an active research area due to introduction of social media as it provides content generation facility of its users. Thanks to social media platforms, common people can share their views, opinions and experiences. The main focus of the researchers has been to carry out sentiment analysis in the English language content. Minimal work has been done in the field of SA in content in Pashto language which is the national language of Afghanistan and widely spoken in Pakistan as well. In this research study, our aim is to perform SA in Pashto text of social media content using machine learning and state of the art deep learning algorithms. We exploit various text feature engineering techniques like Term Frequency-Inverse Document frequency, bag-of-words, n-gram, as well as deep features of word2vec, and GloVe. We perform three sets of test subjectivity analysis, binary and tertiary level sentiment classification. Being a pioneer work, we received satisfactory results on self-prepared datasets which is extracted from social media sources. The empirical analysis-based results are evaluated using standard performance evaluation measures such as accuracy, precision, recall and f-measure. Among numerous applied algorithms, Random Forest obtained better results as compared to other algorithms.

Keywords: Social media, Deep learning, Pashto language, Sentiment analysis

1 Introduction

The advancements in social media has attracted the researchers from data mining domain to carry out research tasks which are closely related to common people. Sentiment analysis, a trendy research area, is one of them. In recent years, due to the widespread internet and social media [1]. The users share their experiences through comments on social media sites such as Facebook, Twitter, Instagram, etc. These customer reviews are about various products and services are analyzed by carrying out sentiment analysis [2]. SA is the process of collecting text, processing text, and retrieving some valuable information to make business decisions [3]. Subjectivity, polarity, and emotion are the three basic categories of sentiment analysis [4]. Subjectivity analysis is to filter out the opinionated and natural facts [5]. A subjective sentence carries some opinion words, e.g., "the camera of this cell phone has excellent quality and I like its beautifu design" contains opinion, and objective doesn't have any opinion e.g., "The camera of this cell phone is 48 mega pixels" [6-7].The main foucs of SA is to analyze the text to determine what kind of feeling it expresses [8]; we assign negative, positive, or neutral to the feelings, which is called polarity [9-10]. Emotion includes a bunch of expressive, behavioral, phycological, and phenomenological characteristics [11-12]. The prediction of a person's emotional states by evaluating text written by them has many uses in computational linguistics, for example, e-learning environments or in the prevention of suicide [13].

SA is broadly utilized in organizations and industries for decision-making or future prediction or find out issues. In the past, if someone wanted to buy a product, they would ask friends or relatives about that product. Nowadays, people take help from the internet, seek reviews/feedback, and comments about particular products before buying [14]. The consumers are increasingly affected by online product review when choosing different products [15]. The direct source for sentiment analysis technology are online shopkeepers, movie reviewers, marketing managers, politicians, public relations agencies, and policymakers. One of the most popular websites for automatic product reviews is "Google Product Search" [16].

Pashto is the national language of Afghanistan and also mostly spoken in Pakistan. Internet access has spread worldwide; the data about any language on the web are increasing tremendously. The researchers have started to work on various languages. Regarding SA, mostly focus has been on English language. Pashto is a very resource-poor language also, the corpus of Pashto language is not available, it is a big challenge to create a corpus. Sentiment analysis of Pashto content presents research challenges due to the lack of plentiful linguistic tools and resources for performing various text analysis tasks such as stop words removal, stemming, and parsing. The main difference is the script difference, as the target content is from right to left. Stemming, Part of Speech (POS) tagging, tokenization, and semantics considerations have all been shown to enhance SA in the English language and should be evaluated for the Pashto language as well. In this research, our focus is to utilize state-of-the-art machine mearning and deep learning algorithms to address these

*Corresponding Author: Hikmat Ullah Khan; E-mail: hikmat.ullah@ciitwah.edu.pk
DOI: 10.53106/160792642022122307021

research challenges and then provide an effective model which may provide an accurate model for sentiment analysis of Pashto text. In this study, our aim is to develop the model for subjectivity and sentiment mining in Pashto text about any product services or political campaign. This study has following main research contributions:

- Data extraction, preprocessing and preparation for subjectivity and sentiment analysis.
- Definition of the list of stop-words which can be serve as a resource for future researchers working in the text data analysis for Pashto content.
- Feature engineering of text features such as TF-IDF, Bag-of-words, N-gram as well as deep features Word2vec, and GloVe.
- Application of state-of-the-art machine learning and deep learning algorithms for classification.
- Research evaluation using standard performance measures of accuracy, precision, recall and f-measure are to assess the performance of the classification model.

2 Literature Review

In recent studies, subjectivity analysis is carried out using either machine learning algorithms or lexicons**.** Both methods are discussed in the following shows the approaches towards Sentiment Analysis.

2.1 Machine Learning

Machine learning is the analytical based model that focus on creating such type of automatic system that learns from experience and surrounding environments to extract pattern, take decision with minimal human interaction. These algorithms learn automatically without human help, so they become more brilliant. Furthermore, Machine learning is mainly categorized into three types: supervised learning, unsupervised learning, and reinforcement learning [17]. The ideas of using Logistic Regression, Naive Bayes, Random Forest, and K-Nearest Neighbor algorithms to automatically detect feelings expressed in English text for Amazon and Flipkart goods were explored and investigated [18]. The Product Comment Summarizer and Analyzer (PCSA) system was proposed in another study and utilized supervised learning algorithms including Random Forest, Naïve Bayes, SentiWordNet, and KNN [19]. In this work, they focused on to find polarities of positive, negative and neutral and used the Stanford Sentiment Treebank data set for movie review.

After application of Naïve Bayes classifier and some Deep Learning (DL) methods such as Deep Dense Network (DNN), Convolutional Neural Network (CNN), and Recurrent Neural Network (RNN); the performance is analyzed. Convolutional Neural network shows the best performance among all. The authors [20] predicted that factors are essential for movie success and profitability. They used Support vector Machine (SVM) and linear regression to analyze movie trailer views, Wikipedia page views, and the released time of the movie; the performance of SVM was 56.52 %. In this study [21], they predicted the election's actual outcomes using Twitter data in three countries of Asia: India, Pakistan, and Malaysia. They concluded that the strength of social media predictivity works perfectly in India and Pakistan but in the case of Malaysia is not. In another study [22], the prediction of stock market trends is attempted; two supervised machine learning models, one for daily prediction and the other for monthly prediction, have been built up. On a daily prediction model, supervised machine learning algorithms may achieve up to 70% accuracy. The monthly prediction model attempts to determine whether any two months' trends are comparable.

2.2 Lexicon-Based

Lexicon-based is another technique used for Sentiment analysis. From the semantic orientation of the lexicon, this methodology measures the sentiment polarity of all text or series of sentences. Both automatic and manual methods are used to create a lexicon dictionary. The most used dictionary is Wordnet. At the start, lexicons are created from the whole document; afterwards, WorldNet or another online thesaurus is used to find synonyms and antonyms to enhance the dictionary. However, it has the disadvantages of requiring human intervention in the text process.

These researchers [23] proposed the word-based approach for obtaining sentiment from text data in their paper. They expand the Semantic Orientation CALculator (SO-CAL) to additional elements of speech, building on prior research that used adjectives. In this study [24], the WKWSCI Sentiment Lexicon is a new sentiment lexicon created compared to five other lexicons: MPQA Subjectivity Lexicon, NRC Word-Sentiment Association Lexicon, Hu & Liu Opinion Lexicon, General Inquirer, and SO-CAL lexicon. WKWSCI, Hu & Liu, MPQA, and SO-CAL lexicons are similarly brilliant for product feedback sentiment categorization, with accuracy rates of 75 percent to 77 percent. The accuracy rate of WKWSCI was 69%. The Hu & Liu lexicons were suggested for product review papers, whereas the WKWSCI lexicon was recommended for non-review materials. In another study [25-26], the authors built a corpus from the Arabic tweet. First, they crawled the tweets and used lexicons for empirical analysis and achieved accuracy of 68.89%. When compared to the keyword-based method, the suggested methodology produced excellent results in terms of prediction accuracy.The authors of another study [27] used lexicons and dictionaries for sentiment classification. First, they capture the tweets and then identify subjective data applied to score the module after classifying it as either positive, negative, or neutral. In binary classification and multi-class classification, they got 92% and 87% accuracy respectively.

2.3 Urdu Sentiment Analysis

We discuss existing work related to Urdu data as it is somewhat similar to Pashto language. Because of the script, morphological and grammatical difference, resources of well-studied language like English is not adequate; a very partial work has been performed in Urdu sentiment as well [28]. The authors [29] perform sentiment analysis on blogs collected 14 different genres, apply ML algorithms Decision Tree (DT), KNN, SVM. They concluded that the performance of K-NN is better than other algorithms. The authors [30] used the bilingual method to develop Urdu's sentiment lexicons. The authors [31] conducted sentiment analysis on Roman Urdu using the Long-Short Time Memory Module (LSTM) which achieved good accuracy than lexicon-based method. In this study [15], initially, data is gathered from an online sources

and used five ML classifiers. Overall, IBK showed the best classifier for Urdu SA based on the results of the test conducted. This study [14] used three ML algorithms to analyze Roman Urdu Opinion Mining: Naïve Bayes (NB), DT, and KNN. NB outperformed other algorithms. In this study [24], the authors developed a Roman Urdu Opinion Mining System (RUOMiS) to help the Urdu native speakers as well as Non-native speakers to take benefits from the online review/comments posted in Roman Urdu. The experiment showed a satisfactory result as the recall of 100%; however, RUOMiS classified roughly 21.1 % of outcomes incorrectly.

3 Proposed Methodology

In this research work, initially, we prepared Pashto language dataset from the social media site of Facebook by extracting comments of various users. Additionally, we extracted features, applied machine learning techniques to classify the positive, negative and neutral data, and evaluated the results using performance evaluation measures.

Figure 1 shows the framework of the proposed methodology.

3.1 Feature Engineering

We have considered the standard textual features such as TF-IDF, bag of words, nagram for Pashto language. In addition, we also take the deep features such as word2vec and GloVe. The details of these features are described as follows:

3.1.1 TF-IDF

TF-IDF is a feature extraction technique which is used to represent the importance of a phrase or any word in a given document.

3.1.2 Bags-of-Words

For each data case, Bag of Words(BoW) model builds a corpus of word counts. The count pure binary or sublinear counts are possible. In addition to word enrichment, a BoW model is needed, which may be used for predictive modelling. Three parameters are used for the BoWords model: Term Frequency, Document Frequency, and Regularization.

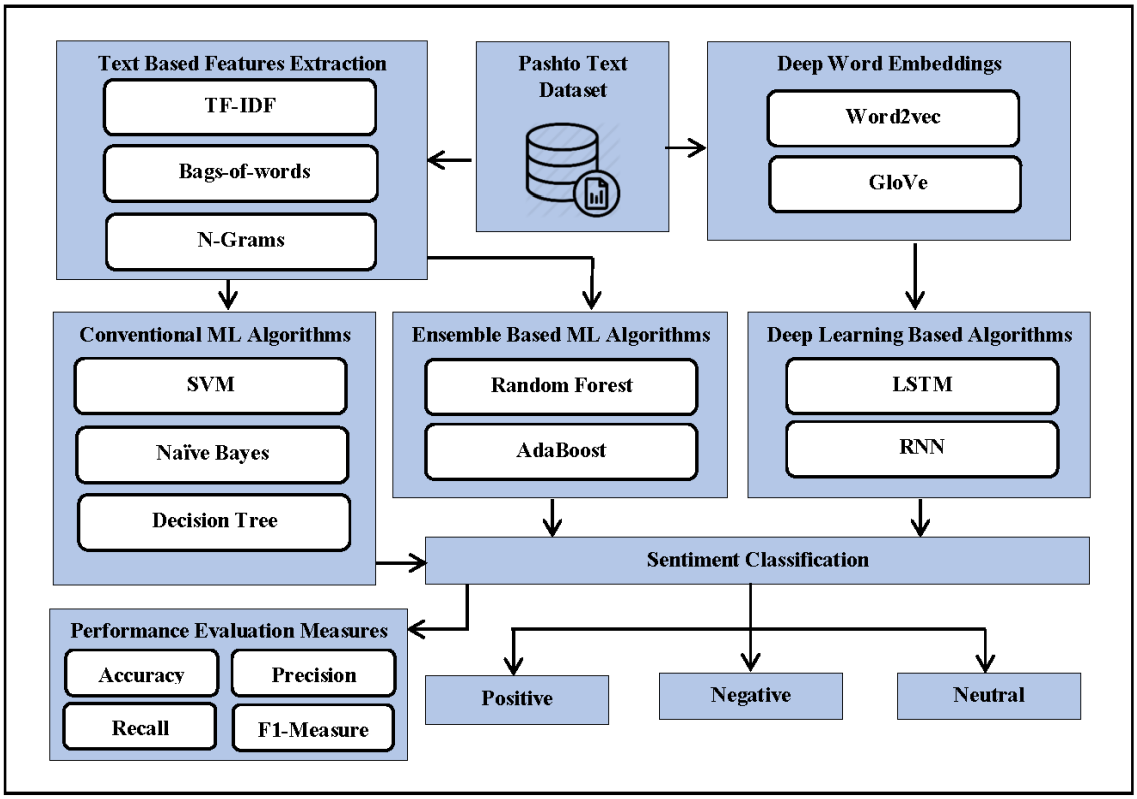

Figure 1. Framework showing steps of the proposed research methodology

3.1.3 N-gram

The N-gram technique is widely used in statistical NLP. Sequences of characters are modelled differently for different language identification. It can be analyzed using unigram, bigram and trigram.

3.1.4 Word2Vec

Word2vec uses a three-layered deep neural network for measuring the perspective of the document and relate parallel context phrases together. It has two variations Skip-Gram and Continues Bag of Words. For sensible word embedding and good feature generation, word2vec should be trained on a large and quality dataset.

3.1.5 GLoVe:

The GloVe is an unsupervised learning methodology for generating a word vector representation. It concentrates on term co-occurrence throughout the whole corpus. Its embedding is related to the changes of two terms being together.

3.2 Machine Learning Techniques

Here we have used conventional ML approaches and ensemble-based ML approach.

3.2.1 Conventional ML Methods

SVM*:* SVM is the supervised type of ML algorithm. SVM is capable of classifying the data into two groups or classes. It needs some training data to train and learn the model; then, it predicts the new data based on training samples.

NB**:** NB is also known as a "probabilistic classifier." NB predicts the classes of new data based on training from the attributes of the input data. The attributes are treated independently in NB. It is easy to implement and more suitable for extensive data.

Decision Tree**:** It is used for both classifications as well as for regression problems. The key idea is to create a training model using a tree representation which can be used to predict the class by learning concluded decision rules. The internal nodes of the tree represent attributes, and leaf nodes represent the specific class label.

3.2.2 Ensemble Methods

Ensemble methods use multiple ML techniques to achieve better results by decreasing variance (called bagging), bias (called boosting), or improving predictions.

Random Forest**:** RF consists of an abundant distinct Decision Tree. RF uses the crowd's wisdom for model prediction such that every Decision Tree gives a class prediction and the class with the most votes become the model's prediction. It can be used for both regression and classification.

AdaBoost**:** AdaBoost combines a number of different "weak classifiers" into a single "strong classifier." AdaBoost's weak learners are referred to as decision stumps. Signal splits exist in these decision stumps. AdaBoost works by giving more weight to cases that are difficult to categorize and less weight to classes that are already well-handled.

3.3 Deep Learning Methods

RNN**:** Recurrent neural network (RNN) is derived from feedforward network. It can solve the problem of prediction in sequential data. It takes intellectual decisions and works like a human's brain. It has a short-term memory. So, its structure is like a circle that uses your current value for the next round.

LSTM**:** Long short-term memory (LSTM) is derived from RNN. It has solved the issue of RNN. It is used for a long and short memory. It can use the whole data for performing experiments rather than a single data point. The data sample has been divided into small sets of data. There are various activation functions such as ReLu and sigmoid.

4 Experimental Setup

In this section, we discuss about prepared dataset and performance evaluation measures which are used for results analysis.

4.1 Dataset

We prepared the corpus of Pashto language by extracting Facebook using FacePager. The corpus contains 600 text documents in Pashto language. The annotators classified the sentences manually from the annotators in three classes: positive negative and neutral. The accuracy of the data annotation was carried out using Kappa Cohen statistics and the overall result falls in very good raing which is highly acceptable for research task.

4.2 Performance Evaluation Measures

Following measures are used for the research evaluations:

Accuracy**:** The percentage of successfully categorized test sets is known as accuracy.

$$\text{Accuracy} = \frac{\text{TP+TN}}{\text{TP+TN+FP+FN}}. \quad \textbf{(1)}$$

where *TP* is original True as well as predicted True by the classifier, *TN* is originally True but predicted negatively by the classifier, *FP* is the Original false but predicted positively by the classifier, and *FN* is the original false and predicted false by the classifier.

Precision**:** Precision is exactness, i.e., what % of entries that the classifier labelled as positive are positive.

$$\text{Precision} = \frac{\text{TP}}{\text{TP+FP}}. \quad \textbf{(2)}$$

Recall**:** The recall is completeness, i.e., what % of the positive tuples did the classifier label as positive?

$$\text{Recall} = \frac{\text{TP}}{\text{TP+ FN}}. \quad \textbf{(3)}$$

F1-measure**:** It is the Harmonic mean of recall along with precision.

$$\text{F} = \frac{\text{2 X precision X Recall}}{\text{precision+ Recall}}. \quad \textbf{(4)}$$

5 Results and Discussions

We summarize the findings of proposed study and demonstrate the usefulness of our proposed model. The stop-words removal was also a challenging task because there is not any standard of stop-words of Pashto text available. We defined 180 stop-words and filtered out. Let us discuss here first the experimental results of conventional-based,

ensemble-based and deep learning algorithms for subjectivity analysis, sentiment polarity, and sentiment analysis.

5.1 Results of Conventional and Ensemble-based ML Algorithms

5.1.1 Subjectivity Analysis

To analyze the opinionated sentences and facts, we applied the Machine Learning algorithms to our Pashto text. We showed the results of subjectivity analysis of conventional-based and ensemble-based ML algorithms using features extraction algorithms such TF-IDF, Bag-of-words, and N-gram. The Table 1 shows the results of the tf-idf model. According to the results, the RF outperformed SVM, Naïve Bayes, Decision Tree, AdaBoost and attained 87% accuracy. According to the results, using Bag-of-words results for Subjectivity Analysis, SVM performs better over other techniques and attains 88.33% accuracy. According results, using TF-IDF and N-gram for Subjectivity Analysis Random Forest performed well than the other algorithms in terms of accuracy.

5.1.2 Sentiment Polarity

To carry out sentiment polarity, our main focus is to carry out binary classification. We applied ML algorithms with features extraction methods such as, TF-IDF, Bag-of-words, and N-gram on Pashto corpus. According to Table 2, Random Forest outperformed SVM, Decision Tree and Naïve Bayes, and AdaBoost in terms of and achieved 86% accuracy. However, in terms of precision value, Naïve Bayes and Random Forest both algorithms achieved 99% precision. The results of conventional machine learning and ensemble-based ML algorithms using Bag-of-words to analyze the sentiment polarity are presented in Table 2. According to results NB and Random Forest showed better accuracy than others. We used the TF-IDF and N-gram with ML algorithms for Polarity. In which SVM and RF outperformed over techniques.

Table 1. Sentiment polarity using conventional and ensemble-based method with TF-IDF, Bags-of-words and N-grams

Features extraction	Model	Accuracy %	Precision %	Recall %	F1-measure %
TF-IDF	SVM	87	97	90	93
	Naïve Bayes	88	99	88	94
	Decision Tree	83	92	89	91
	Random Forest	**89**	**99**	**89**	**94**
	AdaBoost	87	85	90	93
Bags of words	SVM	88	96	91	94
	Naïve Bayes	88	99	89	94
	Decision Tree	86	96	89	92
	Random Forest	**88**	**99**	**88**	**94**
	AdaBoost	85	95	89	92
TF-IDF & N-Grams	SVM	88	98	90	94
	Naïve Bayes	88	99	88	94
	Decision Tree	78	85	89	87
	Random Forest	**89**	**99**	**89**	**94**
	AdaBoost	86	95	89	92

Table 2. Subjective analysis using conventional and ensemble-based method with TF-IDF, Bags-of-words and N-grams

Features Extraction	Model	Accuracy %	Precision %	Recall %	F1-measure %
TF-IDF	SVM	85	95	88	92
	Naïve Bayes	85	99	85	92
	Decision Tree	82	93	87	90
	Random Forest	**86**	**99**	**86**	**92**
	AdaBoost	84	85	87	91
Bags of words	SVM	83	94	87	91
	Naïve Bayes	**87**	97	87	03
	Decision Tree	82	94	86	90
	Random Forest	**87**	**98**	**86**	**93**
	AdaBoost	83	96	86	91
TF-IDF & N-Grams	SVM	**86**	98	87	92
	Naïve Bayes	85	99	85	92
	Decision Tree	81	93	86	89
	Random Forest	**86**	**99**	**86**	**92**
	AdaBoost	83	94	87	91

5.1.3 Sentiment Analysis

Here we performed tertiary classification considering neutral as third class. We applied the conventional ML and ensemble-based ML algorithms with feature extraction, such TF-IDF, Bag-of-word and N-gram on Pashto corpus. Table 3 shows the results of conventional ML and ensemble-based ML algorithms with TF-IDF, Bag-of-words and N-grams for

feature extraction. According to results SVM obtained the 66% accuracy and outperformed the other algorithms. In results of ML algorithms with Bag-of-words, Naïve Bayes showed better accuracy than the other algorithms. We used the TF-IDF and N-gram for feature extraction for sentiment analysis of Pashto text in which SVM attains better accuracy than the others algorithms.

Table 3. Sentiment analysis using conventional and ensemble-based methods using diverse features

Features Extraction	Model	Accuracy %	Precision %	Recall %	F1-measure %
TF-IDF	SVM	**66**	91	**71**	**79**
	Naïve Bayes	60	**99**	60	75
	Decision Tree	60	92	60	73
	Random Forest	59	94	61	71
	AdaBoost	57	91	62	74
Bags of words	SVM	65	91	67	77
	Naïve Bayes	**66**	92	69	**79**
	Decision Tree	56	91	57	70
	Random Forest	59	**97**	58	73
	AdaBoost	57	95	**80**	72
TF-IDF & N-Grams	SVM	**86**	98	**87**	**92**
	Naïve Bayes	85	**99**	85	**92**
	Decision Tree	81	93	86	89
	Random Forest	**86**	99	86	92
	AdaBoost	83	94	**87**	91

5.2 Deep Learning Algorithms Results

In this research work, in addition to machine learning, we also applied two deep learning algorithms including LSTM and RNN with deep features like Word2vec and GloVe for subjectivity analysis and sentiment polarity of our Pashto text. In Table 4, the results of sentiment polarity and subjectivity analysis using deep learning algorithm along with word2vec and GloVe are presented. According to results for sentiment polarity LSTM with word2vec achieved 88% accuracy and outperformed over other models. Moreover, for subjective analysis RNN and LSTM with word2vec obtained 85% accuracy outperformed GloVe model.

Table 4. Sentiment polarity and subjective analysis using deep learning with Word2Vec and GloVe

Analysis	Features extraction	Model	Accuracy %	Precision %	Recall %	F1-measure %
Sentiment polarity	Word 2 Vec	LSTM	**88**	99	88	93
		RNN	87	99	88	93
	GloVe	LSTM	84	97	84	91
		RNN	82	97	82	90
Subjective analysis	Word 2 Vec	LSTM	**85**	99	85	92
		RNN	**85**	98	85	92
	GloVe	LSTM	83	99	84	91
		RNN	84	97	84	91

5.3 Comparison of ML and DL

Figure 2 and Figure 3 show the comparison of ML and DL for subjectivity analysis and sentiment polarity respectively. According to the results, Random Forest outperformed machine learning as well as deep learning algorithms for both subjective analysis and sentiment analysis.

6 Conclusion

In this research study, conventional, ensemble and deep learning-based framework is presented for Pashto text sentiment analysis. The proposed framework includes several texts and deep learning based feature extraction techniques such as TF-IDF, bags-of-words, N-grams, word2vec and GloVe respectively. The features are computed form Pashto text dataset. The detail experiments are carried out to validate the proposed model. The results of performance evaluation measures verify that ensemble-based methods with text-based features played an important role in the Pashto sentiment analysis. Random forest with TF-IDF and N-Gram outperformed all other classifiers by achieving 89% accuracy. Moreover, deep learning model with word2vec also performs better and achieve 88% accuracy. The role of text-based features are more effective in classification of Pashto sentiment into positive, negative and neutral.

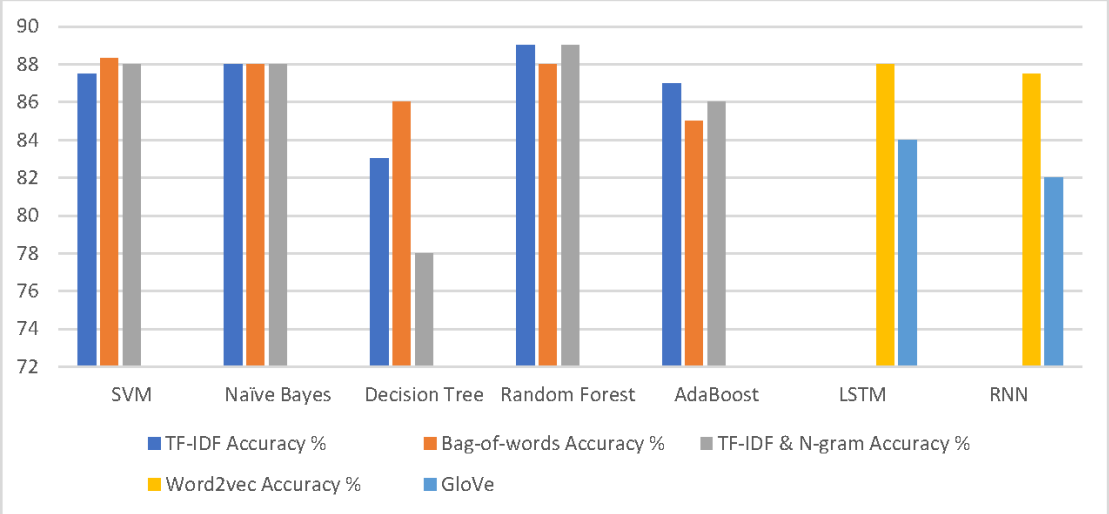

Figure 2. Performance comparison of ML and DL for subjectivity analysis

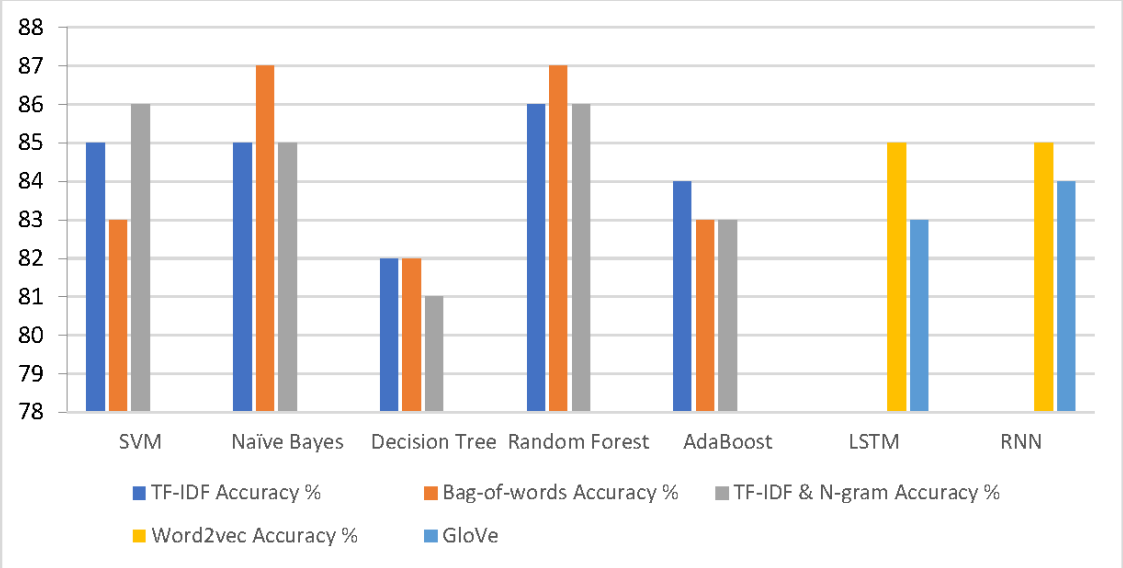

Figure 3. Performance comparison of ML and DL algorithms for sentiment polarity

References

[1] N. X. Bach, T. M. Phuong, Leveraging User Ratings for Resource-poor Sentiment Classification, *Procedia Computer Science*, Vol. 60, pp. 322-331, 2015.

[2] U. Ishfaq, H. U. Khan, K. Iqbal, Identifying the influential bloggers: a modular approach based on sentiment analysis, *Journal of Web Engineering*, Vol. 16, No. 5-6, pp. 505-523, September, 2017.

[3] A. Mahmood, H. U. Khan, M. Ramzan, On Modelling for Bias-Aware Sentiment Analysis and Its Impact in Twitter, *Journal of Web Engineering*, Vol. 19, No. 1, pp. 1-28, March, 2020.

[4] A. Zamir, H. U. Khan, W. Mehmood, T. Iqbal, A. U. Akram, A feature-centric spam email detection model using diverse supervised machine learning algorithms, *The Electronic Library*, Vol. 38, No. 3, pp. 633-657, July, 2020.

[5] M. Alghobiri, U. Ishfaq, H. U. Khan, T. A. Malik, Exploring the role of sentiments in identification of active and influential bloggers, *International Journal of Business and Technology*, Vol. 4, No. 2, pp. 48-55, Spring, 2016.

[6] A. Khattak, M. Z. Asghar, A. Saeed, I. A. Hameed, S. A. Hassan, S. Ahmad, A survey on sentiment analysis in Urdu: A resource-poor language, *Egyptian Informatics Journal*, Vol. 22, No. 1, pp. 53-74, March, 2021.

[7] I. Chaturvedi, E. Cambria, R. E. Welsch, F. Herrera, Distinguishing between facts and opinions for sentiment analysis: Survey and challenges, *Information Fusion*, Vol. 44, pp. 65-77, November, 2018.

[8] H. U. Khan, A. Daud, Using Machine Learning Techniques for Subjectivity Analysis based on Lexical and Non-lexical Features, *International Arab Journal of Information Technology (IAJIT)*, Vol. 14, No. 5, pp. 481-487, July, 2017.

[9] A. V. Bidwaikar, M. H. Dakhore, A Survey on Polarity Checking of the Text, *International Journal of Engineering Trends and Technology*, Vol. 42, No. 7, pp. 375-376, December, 2016.

[10] M. Darwich, S. A. M. Noah, N. Omar, Deriving the sentiment polarity of term senses using dual-step context-aware in-gloss matching, *Information Processing & Management*, Vol. 57, No. 6, Article No. 102273, November, 2020.
[11] D. M. E.-D. M. Hussein, A survey on sentiment analysis challenges, *Journal of King Saud University - Engineering Sciences*, Vol. 30, No. 4, pp. 330-338, October, 2018.
[12] A. K. Nassirtoussi, S. Aghabozorgi, T. Y. Wah, D. C. L. Ngo, Text mining for market prediction: A systematic review, *Expert Systems with Applications*, Vol. 41, No. 16, pp. 7653-7670, November, 2014.
[13] K. Sailunaz, M. Dhaliwal, J. Rokne, R. Alhajj, Emotion detection from text and speech: a survey, *Social Network Analysis and Mining*, Vol. 8, No. 1, pp. 1-26, April, 2018.
[14] A. Abirami, V. Gayathri, A survey on sentiment analysis methods and approach, *2016 Eighth International Conference on Advanced Computing*, Chennai, India, 2017, pp. 72-76.
[15] P. Racherla, M. Mandviwalla, D. J. Connolly, Factors affecting consumers' trust in online product reviews, *Journal of Consumer Behaviour*, Vol. 11, No. 2, pp. 94-104, March/ April, 2012.
[16] R. Feldman, Techniques and applications for sentiment analysis, *Communications of the ACM*, Vol. 56, No. 4, pp. 82-89, April, 2013.
[17] A. Rafique, M. K. Malik, Z. Nawaz, F. Bukhari, A. H. Jalbani, Sentiment analysis for roman urdu, *Mehran University Research Journal of Engineering & Technology*, Vol. 38, No. 2, pp. 463-470, April, 2019.
[18] A. Dadhich, B. Thankachan, Social & Juristic challenges of AI for Opinion Mining Approaches on Amazon & Flipkart Product Reviews Using Machine Learning Algorithms, *SN Computer Science*, Vol. 2, No. 3, Article No. 180, May, 2021.
[19] L. Zhang, S. Wang, B. Liu, Deep learning for sentiment analysis: A survey, *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, Vol. 8, No. 4, Article No. e1253, July/ August, 2018.
[20] V. Subramaniyaswamy, M. V. Vaibhav, R. V. Prasad, R. Logesh, Predicting movie box office success using multiple regression and SVM, *2017 international conference on intelligent sustainable systems*, Palladam, India, 2017, pp. 182-186.
[21] K. Jaidka, S. Ahmed, M. Skoric, M. Hilbert, Predicting elections from social media: a three-country, three-method comparative study, *Asian Journal of Communication*, Vol. 29, No. 3, pp. 252-273, 2019.
[22] A. Nayak, M. M. Pai, R. M. Pai, Prediction models for Indian stock market, *Procedia Computer Science*, Vol. 89, pp. 441-449, 2016.
[23] M. Taboada, J. Brooke, M. Tofiloski, K. Voll, M. Stede, Lexicon-based methods for sentiment analysis, *Computational linguistics*, Vol. 37, No. 2, pp. 267-307, June, 2011.
[24] C. S. Khoo, S. B. Johnkhan, Lexicon-based sentiment analysis: Comparative evaluation of six sentiment lexicons, *Journal of Information Science*, Vol. 44, No. 4, pp. 491-511, August, 2018.
[25] N. Abdulla, N. A. Ahmed, M. Shehab, M. Al-Ayyoub, Arabic sentiment analysis: Lexicon-based and corpus-based, *IEEE Jordan Conference on Applied Electrical Engineering and Computing Technologies*, Amman, Jordan, 2013, pp. 1-6.
[26] M. Al-Ayyoub, S. B. Essa, I. Alsmadi, Lexicon-based sentiment analysis of arabic tweets, *International Journal of Social Network Mining*, Vol. 2, No. 2, pp. 101-114, October, 2015.
[27] F. M. Kundi, A. Khan, S. Ahmad, M. Z. Asghar, Lexicon-based sentiment analysis in the social web, *Journal of Basic and Applied Scientific Research*, Vol. 4, No. 6, pp. 238-248, June, 2014.
[28] A. Z. Syed, M. Aslam, A. M. Martinez-Enriquez, Lexicon based sentiment analysis of Urdu text using SentiUnits, *Mexican International Conference on Artificial Intelligence*, Pachuca, Mexico, 2010, pp. 32-43.
[29] N. Mukhtar, M. A. Khan, Urdu sentiment analysis using supervised machine learning approach, *International Journal of Pattern Recognition and Artificial Intelligence*, Vol. 32, No. 2, Article No. 1851001, Februray, 2018.
[30] M. Z. Asghar, A. Sattar, A. Khan, A. Ali, F. M. Kundi, S. Ahmad, Creating sentiment lexicon for sentiment analysis in Urdu: The case of a resource-poor language, *Expert Systems*, Vol. 36, No. 3, Article No. e12397, June, 2019.
[31] H. Ghulam, F. Zeng, W. Li, Y. Xiao, Deep learning-based sentiment analysis for roman urdu text, *Procedia computer science*, Vol. 147, pp. 131-135, 2019.

Biographies

Saqib Iqbal received his doctoral in Software Engineering from the University of Huddersfield in 2013. Prior to this, he completed his MSc in Software Engineering from the Queen Mary University of London, the UK in 2007. He has taught in various postgraduate universities for over 10 years and has worked in the industry as a Software Engineer for more than 3 years. He is actively involved in research in areas particularly related to requirements engineering, software design, model-transformation, and software testing.

Farhad Khan is research scholar and doing her master degree in computer science from Department of Computer Science, COMSATS University Islamabad, Wah Cantt, Pakistan. His research interests include Machine Learning, Deep Learning, Sentiment Analysis and Data mining.

Hikmat Ullah Khan received the master's degree in computer science and the Ph.D. degree in computer science from International Islamic University, Islamabad. He has been an Active Researcher for the last ten years. He is currently an Assistant Professor with the Department of Computer Science, COMSATS University Islamabad, Wah Cantt, Pakistan. He has authored 50+ articles in top peer-reviewed journals and international conferences. His research interests include social Web mining, semantic Web, data

science, information retrieval, and scientometrics. He is an Editorial Board Member of a number of prestigious impact factor journals.

Tassawar Iqbal is presently serving as Assistant Professor (TTS) at Department of Computer Science in COMSATS University Islamabad, Wah Campus, Pakistan. He has completed his PhD degree from Vienna University of Technology in 2012. He received his MS (CS) degree from COMSATS University Abbottabad Campus in year 2007. His current research interests include Studying impact of technologies in society and Human Computer Interaction. He has 29 articles published in reputed journals and conferences.

Jamal Hussain Shah received the Ph.D. degree in computer science from University of Science and Technology of China (USTC), China. He is currently an Assistant Professor with the Department of Computer Science, COMSATS University Islamabad, Wah Cantt, Pakistan. His areas of interest are Digital image Processing and Networking. He has authored several research articles published in well reputed peer-reviewed journals.