

Sentiment Analysis of Roman-Urdu Tweets about Covid-19 Using Machine Learning Approach: A Systematic Literature Review



Syed Muhammad Waqas Shah¹, Muhammad Nadeem², Muzamil Mehboob³

¹National University of Modern Languages, Pakistan, smwaqas@numl.edu.pk

²National University of Modern Languages, Pakistan, mnadeem@numl.edu.pk

³Bahauddin Zakariya University, Pakistan, muzamilmehboob@bzu.edu.pk

ABSTRACT

From the past decade, growth of social media encourages researchers to perform several academic studies on the user generated data. Sentiment Analysis is tool used to generate summary of public opinion about a topic, which helps in effective decision making. Most of the work in sentiment analysis has been performed on English Language Dataset. Many Languages such as Urdu, Hindi, and Italian faced lack of attention due to their complex morphological structure and unavailability of resources. COVID-19 has affected human life around the world. Many researchers have performed their studies using twitter data for knowing the sentiments of people throughout the time of pandemic. In this Paper we are giving a general overview about the process of Sentiment Analysis of Roman-Urdu Tweets. For the dataset query we have used the data generated by user about COVID-19.

Key words: Sentiment Analysis, Machine Learning, Roman-Urdu, COVID-19

1. INTRODUCTION

In today's world, the most usable platform used for communication between different people is internet. Due to the incredible grow of technology; people are using search engine for various aspects of life such as e-commerce, e-marketing, distance learning, bill payment, social media interaction etc. Different social media platform such as Facebook, YouTube and Twitter enable people to share their views about any product, event, service, political affair, and other viral news publicly [1][3][10]. People are spending more time on social media as compare to the time they spend on drinking, eating and on other social activities as analyzed by marketers. According to a survey of Smart Insights, within a minute, people manage to publish 3.3 million posts on Facebook and 4.5 million tweets on twitter [19]. There is too much user-generated data available on the internet used by the organization and companies to

know about the feedback of people. Through this available data, people can make their decision about purchase and companies can improve their product after knowing about the opinion of people [5]. As there is too much opinion available about everything on the internet, there is a need of some technique or algorithm in order to find accurate results. COVID-19 epidemic have affected millions of people around the globe. The virus is highly infectious because it transfers from person to person through any mean [24]. Many countries including Pakistan have taken several precautions to slow down the extent of COVID-19, such as ban of travelling, lockdown, closing of public places (e.g., gym, restaurants, schools), requiring people to practice good personal hygiene, keep social distancing of 1.5 meters, and work or study at home [35]. People in Pakistan have experienced many things first time due to the lockdown like wearing masks, work from home, delivery of items, online classes and exams. People have mixed views about them and they shared their opinion on social media platforms such as twitter. This massive personal post could be a valuable data for knowing the sentiments of people about incidents happened during COVID-19. Urdu is a national language of Pakistan and more than 64 million people worldwide speaks Urdu. Most of the people of Pakistan are very well aware of Roman script of Urdu and they use it to express their opinion on social-media platforms.

2. SENTIMENT ANALYSIS OF ROMAN URDU TWEETS

Sentiment Analysis (also known as opinion mining) is a tool used generate an overview of public opinion about a particular activities, events, objects, topics, services and products which helps in effective decision making [1][2][3]. It is a combination of Natural Language Processing and Data mining techniques. SA considered as a classification process, which performed at three levels named as document level, sentence level and aspect level. At document Level, entire document considered as one

unit of information for one topic and the whole document will classify as positive, neutral or negative on the basis of positive, neutral and negative sentences present in the document. At sentence level sentiment analysis, each sentence is classified as positive, negative or neutral. In aspect level Sentiment classification, different aspects (features) of the previously identified entities classified as positive or negative[1].It opens up opportunities for marketing teams, celebrities, politicians, and anyone concerned with opinions and moods of the general public. Many researchers have performed sentiment analysis during COVID-19 to find the reaction and opinion of people as mentioned in table 1. Wang et al. [18] collected posts from Weibo and used unsupervised BERT (Bidirectional Encoder Representations from Transformers) to classify sentiments of people into different categories (positive, neutral, and negative.

Nemes et al. [21] applied a Recurrent Neural Network (RNN)for the classification of tweets. He developed a model to analyze the emotional nature of

numerous tweets, using the recurrent neural network for emotional prediction, searching for connections between words, and marking them with positive or negative emotions. Samuel et al [22] collected data from twitter and found sentiment of people toward COVID-19. He further compared performances of different classification algorithms such as Naïve Bayer classifier, Logistic regression, and linear regression for the twitter data. Mostly the work in sentiment Analysis has been performed on English and there are many libraries and packages available in python for sentiment analysis of English language text such as Textblob [25].The table 1 below represents the different researches on roman Urdu data set.

Table 1: Survey on research types and methods using Roman Urdu datasets

| References | P. Channel | P. Year | Research Type | Empirical Type | Approach |
|------------|------------|---------|-------------------|----------------|----------|
| [12] | Conference | 2020 | Solution Proposal | Experiment | Method |
| [27] | Journal | 2016 | Solution Proposal | Experiment | Method |
| [15] | Journal | 2019 | Solution Proposal | Experiment | Modal |
| [12] | Conference | 2018 | Evaluation Paper | Experiment | Method |
| [1] | Conference | 2017 | Solution Proposal | Experiment | Method |
| [6] | Journal | 2019 | Solution Proposal | Experiment | Method |
| [7] | Journal | 2020 | Solution Proposal | Experiment | Modal |
| [10] | Journal | 2020 | Evaluation Paper | Survey | Method |
| [11] | Journal | 2020 | Solution Proposal | Experiment | Method |
| [13] | Journal | 2017 | Solution Proposal | Experiment | Method |
| [2] | Conference | 2019 | Evaluation Paper | Experiment | Method |
| [24] | Journal | 2020 | Evaluation Paper | Experiment | Method |
| [17] | Journal | 2020 | Evaluation Paper | Survey | Modal |
| [18] | Journal | 2017 | Solution Proposal | Experiment | Method |
| [23] | Journal | 2020 | Experience Paper | Experiment | Method |

| | | | | | |
|------|------------|------|-------------------|------------|--------|
| [22] | Journal | 2020 | Evaluation Paper | Survey | Method |
| [25] | Journal | 2020 | Evaluation Paper | Survey | Method |
| [22] | Journal | 2019 | Solution Proposal | Experiment | Method |
| [35] | Symposium | 2020 | Evaluation Paper | Experiment | Method |
| [40] | Conference | 2020 | Solution Proposal | Experiment | Method |

Currently limited research has been performed on other languages such as Urdu, Roman-Urdu, Hindi, Arabic and Italian. Sentiment Analysis of some language is difficult due to their complex morphological structure such as Urdu and Arabic. Sentiment Analysis of Roman-Urdu is as essential as of any other languages as majority of the people in subcontinent are not much well versed of English language and use Roman-Urdu to express their sentiments on social media platforms. Due to the COVID-19 a lot of research conducted on sentiment of people for knowing the opinion of people about outbreak virus, symptoms of virus, quarantine, lockdown etc.

There are two methods of finding sentiments of text data, Corpus-based and Lexicon-based. Corpus-based sentiment analysis used a machine learning classifier on labeled Dataset. The performance of the classifier depends on the quantity and quality of the training data. Lexicon-based sentiment analysis finds the polarity of every word or phrase in a text document with the use of a sentiment lexicon. Sentiment Analysis using Lexicon-based approach has two further types, dictionary based and corpus based. Dictionary based approach used an existing dictionary of words along with the sentiment strength. Corpus-based approach sentiment analysis approach searcher though vast amount of datasets to find the probability of a term occurrence [27].

Different researcher has made numerous efforts for solving the issue of Roman-Urdu script as mentioned in Table 1. But problems of Roman-Urdu text has not been fully solved yet. Javed et al. [31] conducted research on bilingual twitter dataset on a case study of general elections 2013 using lexicon based approach. Arif et al. [27] by using a dataset of English contain hotel reviews. He converted his dataset into Roman-Urdu using an online tool. He used different classifier such as SVM, KNN, Decision tree classifier, Ridge Classifier Naïve Bayes

etc. The results of this study show that Ridge Classifier with TF-IDF achieved maximum accuracy 98%. Bilal et al[1] used three classification models Naïve Bayes, KNN and decision tree in WEKA environment for the classification of opinion extracted from Blog. This study concluded that Naive Bayes performed better than KNN and Decision tree classifier. Rafae et al. [10] performed a study and concluded that several methods used for the sentiment analysis of other languages are not suitable for the Roman-Urdu script. Spelling variation and complex morphological structure is the main reason of this problem. Furthermore, the unavailability of language resources such as datasets and lexicons also makes it difficult to apply the current sentiment analysis techniques.

In this paper we are following two aspects, first to deliver an overview about performing the sentiment analysis of twitter data. The second aspect of our paper is to give reader an idea of how Roman-Urdu tweets can be used in the sentiment analysis.

3. METHODOLOGY

In this section we will study different methods of sentiment analysis of Roman- Urdu tweets. The Methodology of this research consists of five steps as shown in figure 1. In the first phase data will be collected using twitter API. After the complete collection of dataset, preprocessing will be performed to clean the data because clean and labeled data gives good results with machine learning algorithms [30]. Roman-Urdu does not have any labeled dataset, so in order to conduct research; data will be collected and manually labeled by the annotator. After performing the preprocessing several features will be extracted from the dataset. Many researchers have used TF-IDF with unigram and bigram for their research. These two features works well with Roman-Urdu Script. After preprocessing, features will be selected using chi-squared. For the classification, we have presented several studies which have used different algorithms for solving classification problems. At the end model will be evaluated with different methods like Accuracy, precision and recall. Figure 1 below shows the flow of events in the system to analyze the Roman Urdu tweets.

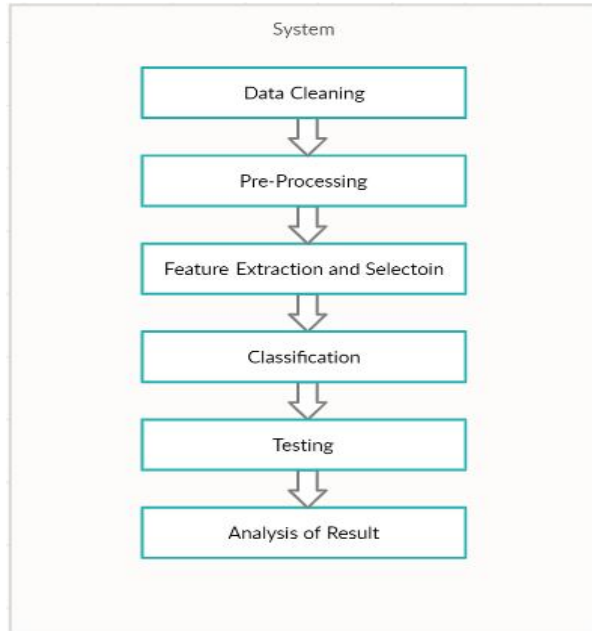


Figure 1: Sentiment Analysis of Roman Urdu Tweets

3.1 Dataset

Launched in 2006, the microblogging platform Twitter became one of the most popular microblogging providers. Initially, the users were given 140 characters space to tweet but recently, this has been extended to 280 characters. Twitter received more than 330 million tweets daily, which makes it a great platform worldwide. It is the treasure of sentiment of people around the globe; since People use twitter to share their thoughts, as well as to spread the information.

The rapid sharing of the opinions of the user on Twitter has encouraged researchers to find the sentiments on almost everything, including sentiments towards products, movies, politics and digital technology. Twitter data has been widely used for crises situation exploration and recording people reaction toward it [18][24][23][22].De Choudhury et al. [32] performed a study for behavioral changes and mood of new mothers suffering from postnatal. He collected data from twitter and evaluate it for postnatal changes like social engagement, emotion, social network, and linguistic style. With this study, they concluded that twitter data can be very useful in identifying mothers at risk of postnatal depression.

In sentiment analysis of twitter Tweepy python library is used for data extraction from twitter API. Twitter API recognized in the 10 world's most powerful APIs. It offers a developer account to the user through which they can access twitter data.

Tweepy is a powerful library of python which enable data retrieval by searching through keywords, hashtags, timelines, trends or Geo-Location. Python is the fastest growing language for developers due to its easy to understand syntax and fast compilation. It has many libraries such as numpy, pandas, nlkt and frameworks like TensorFlow which helps in deep learning. For this paperwe have targeted only those research papers which have used COVID-19 tweets dataset. They all have stored Data in CSV file and applied further analysis on it.

3.2 Preprocessing

Preprocessing is an important step that keeps the important words and removes the unnecessary words. Python programming language mostly used for the data analysis. The steps usually followed for the preprocessing are mentioned below

- 1: Remove hash tag such as (#COVID and @user) and other URLS as they do not make any contribution in Sentiment Analysis.
- 2: Remove all the English characters and emoticons.
- 3: Remove the special characters, Arabic characters, punctuation and numbers from the dataset.
4. Conversion of all tweets into lower case.
- 5: Removing stop words. The common words which do not give meaning to the sentiment are called stop words.
- 6: Remove repetitive words like convert too worried to worry.

After performing preprocessing, we have clean data for the analysis.

3.3 Data Annotation

In most of research paper we have read, annotation is performed manually by human annotators. For manual annotation, every researcher has its own rules according to the suitability of dataset. Mostly annotators are native Urdu speaker, familiar with Roman-Urdu Script. Some of the rules are mentioned below

- 1: How to annotate (sentences contain positive terms or negative terms)
- 2: What to annotate and what not to annotate?
- 3: How to handle Sarcastic and neutral cases?

If the sentence has both negative and positive sentiments, then annotators will classify it as neutral.

3.4 Tokenization

After the Accomplishment of Preprocessing and manual annotation of Dataset, Tokenization is performed. It is a process of converting a random string of characters into Token. They serve as a reference to the original data. Tokenization is used to identify noun, pronoun, verb adjective etc. Combining words with same meaning is one of the processes of NLP. Mostly Tokenization for Roman-Urdu script is usually performed by Spacy Toolkit.

3.5 Feature Extraction

After cleaning and preprocessing data, Feature extraction is performed on the dataset in order to gather useful feature for sentiment Analysis. The most common feature set for NLP of Roman-Urdu script includes n-gram, Bigram and TF-IDF. Earlier researchers have used six algorithms for sentiment classification such as Naïve Bayes, Logistic Regression, Linear Regression, Decision Tree etc. They found that TF-IDF with Bigram worked excellent with most of the Dataset.

TF-IDF: stands for Term Frequency-Inverse Document Frequency. It is use to assign weights to the terms which have relative importance in the dataset [8][9][10]. The TF-IDF value of a words increases when it appears frequently in a tweets. In research, sometime TF-IDF is used to eliminate repeated words and some researcher used to extract frequent words for topic modeling [24].The words such as “the” have higher frequency and these kinds of terms needs to weigh down. Researcher set parameters of TF-IDF for effective results. For example if they set minimum value 2 in term frequency, than the words appear less than 2 times will not be considered. Mostly feature selection is based on word density. Word density is determined by TF-IDF method. Word density helps to know about the polarity of the tweet. The words which do not appear frequently in the tweets remain present in the corpus. TF-IDF is also used for calculating the polarity of the text. The positive or negative status of the word is calculated through the number of time the word arises in given tweet dataset. Dataset executed with TF-IDF, show the rank of each word occur frequently in the dataset. The higher rank means that the word is relevant and it can contribute in the polarity of tweet. TF-IDF basically tells the important of a term in the dataset and helps in extracting relevant terms.

BAG OF WORDS: In natural language processing, Bag of words is used for the simple representation of document or dataset. The document or a sentence represented as bag (multiset) of word by discarding

grammar and the order of occurrence of word but keeping the multiplicity. Bag of words often used for the classification of document in sentiment analysis and topic modeling. It is also used for capturing the emoticon in a text [9]. This model is very simple and can be used many ways for the extraction of appropriate features from the dataset. The main drawback of Bag of word is the absence of information about sentence’s grammar.

UNIGRAM AND BIGRAM: Unigram is define as the series of word in a tweet with a fixed window size such as N=1 which is use to extract valuable information from the dataset. They are often used for the clustering of large dataset. The size of bigram is 2 and it is extensively used in text mining tasks. A bigram takes a sentence as an input and return two consecutive words from the sentence. The benefits of unigram and bigram models are simplicity and scalability. As tweets contain less than 300 words so bigram + Unigram works best in many cases.

3.6 Feature Selection

There is no such method in research for helping in selecting feature set for feeding model. Generally researchers used TF-IDF with N-gram model for the sentiment classification of Roman-Urdu Tweets. Feature selection helps in identifying precise set of feature from the dataset to contribute in the accuracy of the predictive modal. It simplifies the model by eliminating unnecessary attributes. It helps in improving the performance of algorithm by increasing classification accuracy, reducing training time and decreases overfitting. Generally there are four approaches used for the feature selection which are hybrid, Wrapper, Filter and Embedded. Arif et al [27] used filter method for the selection of decent feature set. Some methods used in Filter approach are defined below

CHI-SQUARED TEST: It is used for calculating the occurrence of two procedures, the occurrence of term and the occurrence of document. Feature with high Chi-squared helps in the selection of good features [27] [12] [16].

Document Frequency (DF): It is define as the number of documents containing a particular term. It can be used to eliminate several unimportant words from the document

INFORMATION GAIN (IG): It can be used for feature selection by estimating the gain of each variable in the context of target variable.

3.7 Classification

Sentiment Analysis has been performed using three approaches; machine learning approach, semi-

supervised approach, List based approach (SentiWordNet) and grouping of all of the above. Machine learning approach has two type of classification.

Supervised Classification: In this kind of classification, labeled dataset is used for training model. Classification is an example of supervised learning.

Unsupervised Classification: In this type of classification, Labels of dataset are not known in advance. Once classification is performed; class labels are assigned to the records on the basis of some natural similarities. Clustering is an example of unsupervised learning.

The two types of sentiment classifications are binary classification and multi-class sentiment classification. In binary classification, a feature, sentence or document is classified according to the two polarities, positive or negative, favorable or unfavorable whereas multi-class sentiment classification can label the review into more than two classes. There are some classifiers which have been widely used for the sentiment analysis of Roman Urdu Tweets. Some of them are explained below.

NAÏVE BAYES CLASSIFIER: It is one of the significant classifier of natural language processing applications. It has been widely used for the text classification since 1950's. It is theoretically based on Bayes theorem. The two approaches for Naïve Bayes classifier are Multinomial Naïve Bayes (used for binary representation of the features) and Bernoulli Naïve Bayes (used for the feature representation in frequency). Many researchers have used NBC for the sentiment analysis. Samuel et al [22] used Naïve Bayes classifier for classification of public sentiments toward COVID-19. They observed an accuracy of 91% with Naïve Bayes which is incredible. Mahmood et al [12] performed sentiment analysis of roman Urdu tweets with Naïve Bayes and observed highest accuracy as compare to the other classifiers which is 75%. Bilal at al [1] performed an experiment using WEKA environment for the sentiment analysis of Roman-Urdu opinions. The result of this experiment shows that NBS perform better than KNN and Decision tree. It shows the accuracy of 97%.

LOGISTIC REGRESSION: Logistic regression has been widely used in classification problems. It was first developed by David Cox in 1958 [36]. In Logistic regression, the probability of the outcome of the single trial is modeled by using a logistic function. In logistic function, the probability of the outcomes is converted into binary values. Thus, the

logistic function always predicts answer between 0 and 1. LR classifier helps in the appropriate classification of categorical outcomes. However, this prediction needs each data point to be independent to each other [36]. Samuel et al [22] used Logistic Regression classifier for classification of public sentiments toward COVID-19. They observed an accuracy of 74% with LR which is good. Mahmood et al [16] did a study on sentiment analysis of labeled Roman-Urdu dataset of six different domains. They concluded after applying several algorithms for improving results that logistic regression with Unigram performs better than any classification algorithm.

K-NEAREST NEIGHBORS (KNN): Bilal at al [1] performed an experiment using WEKA environment with different classifiers for the sentiment analysis of Roman-Urdu opinions. The accuracy of KNN in this study was 95%. Arif et al [16] conducted a study with a dataset of hotel reviews in Roman-Urdu and observed 79% accuracy with KNN.

DECISION TREE (DT): Decision tree have been widely used for the sentiment analysis of different languages and it shows good results. Arif et al [16] conducted an experiment for the sentiment analysis of Hotel reviews in Roman-Urdu Script. Decision tree have shown 77% accuracy in this experiment. Another experiment was performed by Bilal et al [1] in the Weka environment and Decision tree has shown 92% accuracy in this experiment. Many researchers have used Naïve Bayes for the Sentiment analysis and for solving Classification problems[1] [12] [22] [4] [16]. Deep learning algorithm such as RCNN has also been used for the sentiment Analysis. Zainabat al [7] conducted a study to evaluate the sentiment of Roman- Urdu text using Rule-based, N-gram with Recurrent Convolutional Neural Network. They used a huge dataset of more than 10,000 sentences. They performed both binary and multiclass classification. The result of this study shows 65% accuracy for binary classification and 57% accuracy for multiclass classification.

3.8 System Evaluation

After applying sentiment classification model, system is evaluated for finding the effectiveness of model. Several studies have used different approaches for the evaluation of modal according to their problem statement. Some researcher used test data for the validation of their model. They spilt dataset into a ratio of 75:25 or 80:20 as training and testing dataset. Some other methods used for the evaluation of model are defined below.

ACCURACY: It is express as the overall number of correctly identified sentences from the entire dataset of sentences.

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN}$$

PRECISION AND RECALL: It is another performance matrix used in the classification problem. Precision is defined as the measurer of result relevancy whereas recall is known as the ability of a classifier to find all the relevant data in a dataset.

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

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

F-MEASURE: It is defined as the Harmonic measure of precision and recall.

$$\text{F1 score} = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$$

4 RESULTS

In this research we have covered a general approach toward sentiment analysis of Roman-Urdu with COVID-19 tweets. With the advancement of social media, a lot of user generated data is available for research. This valuable data can solve a lot of problems. We have defined the possible method of sentiment analysis of Roman Urdu tweets. For problem statement we have targeted Dataset of tweets with COVID-19 related hashtags. For expressing the importance of our targeted dataset, we have also outlined papers which used dataset on COVID-19 and give valuable results. In this research, we have find out that overall performance of the machine learning algorithms depends on the dataset used in training. TF-IDF with N-gram models are the most usable features set of NLP for Sentiment Analysis. At the end for classification, Naïve Bayes is the most used classification algorithm. The possible reason of better performance of NB algorithm works on assumption that all features are independent. Therefore it is utilized by many researchers. The scope of SA analysis of Roman-Urdu tweets can be increased with quality and size of Dataset.

5 FUTURE WORK AND DISCUSSION

The importance of Text analysis has increased over the past few years due the availability of huge Labeled and Unlabeled dataset. Twitter API has contributed a lot in motivating the researchers to perform several studies on Natural language processing. With the outbreak of COVID-19 several studies have been performed for finding the reaction of people toward different problems associated with COVID-19. Yin et al [24] performed a study to detect

dynamic topic from a huge dataset of COVID-19. They Applied Latent Dirichlet Allocation with TF-IDF to find dynamic topic within the dataset. These kinds of studies for the dynamic topic modeling do not exist with Roman-Urdu tweets. People in the subcontinent mostly use Urdu but due to the unavailability of Urdu keyboard, Roman script is used to write Urdu language which is known as Roman Urdu. With this kind of study more interesting topic and sentiment can be discovered from the dataset of Roman-Urdu Tweets. Due to the complexity of Roman-Urdu script and unavailability of labeled dataset, many problems related to Roman-Urdu remain unsolved. Availability of Roman-Urdu Stop word file has made the preprocessing phase easy but other advancements are not available in Roman-Urdu as compare to the English. It is observed in several studies that machine learning algorithms are simpler and often give better results than the lexicon-based approach.

6 CONCLUSION

In this paper we have presented a systematic literature review of Sentiment Analysis of Roman-Urdu with Covid-19 Tweets. Due to the advancement of social-media, people are using it to share their opinion literally about everything. This scenario generates a lot of user data for the research and effective decision. Roman-Urdu script has lack of resources due to its morphological structure. Every researcher has developed its own dataset for the sentiment analysis according to the problem statement. During the study we have observed that TF-IDF with Unigram and Bigram are the most selected features for the sentiment analysis of Roman-Urdu text. Hybrid approaches are also making progress with Roman-Urdu, but a lot of work is required to achieve good results. Many studies have used Naïve Bayes for the Roman-Urdu and it give good results even with the large dataset. The overall performance of machine learning algorithms depends on the quality of dataset. Labeled dataset gives more accurate results as compare to the unlabeled dataset.

REFERENCES

- [1] Bilal, Muhammad & Israr, Huma & Shahid, Muhammad & Khan, Amin. (2016). Sentiment classification of Roman-Urdu opinions using Navie Bayesian, Decision Tree and KNN classification techniques. Journal of King Saud University - Computer and Information Sciences. 28. 10.1016/j.jksuci.2015.11.003.
- [2] Noor, Faiza & Bakhtyar, Maheen & Baber, Junaid. (2019). Sentiment Analysis in E-commerce

Using SVM on Roman Urdu Text. 10.1007/978-3-030-23943-5_16.

[3] Daud, Misbah & Khan, Rafiullah & Mohibullah, & Daud, Aitazaz. (2015). Roman Urdu Opinion Mining System (Ruomis). Computer Science & Engineering: An International Journal. 4. 10.5121/cseij.2014.4601.

[4]K Sreelakshmi, B Premjith, K.P. Soman, Detection of Hate Speech Text in Hindi-English Code-mixed Data, Procedia Computer Science, Volume 171, 2020, Pages 737-744,

[5] Sharf, Z., & Rahman, S. U. (2017). Lexical normalization of roman Urdu text. International Journal of Computer Science and Network Security, 17(12), 213–221.

[6] Sharf, Z., & Rahman, S. U. (2018). Performing natural language processing on Roman Urdu datasets. International Journal Of Computer Science And Network Security.

[7] Zainab Mahmood, Iqra Safder, Rao Muhammad Adeel Nawab, Faisal Bukhari, Raheel Nawaz, Ahmed S. Alfakeeh, Naif Radi Aljohani, Saeed-Ul Hassan, Deep sentiments in Roman Urdu text using Recurrent Convolutional Neural Network model, Information Processing & Management, Volume 57, Issue 4, 2020, 102233, ISSN 0306-4573, <https://doi.org/10.1016/j.ipm.2020.102233>.

[8] Khan, S. N., Khan, K., & Khan, W. (2018).

[9] M. U. Arshad, M. F. Bashir, A. Majeed, W. Shahzad and M. O. Beg, "Corpus for Emotion Detection on Roman Urdu," 2019 22nd International Multitopic Conference (INMIC), Islamabad, Pakistan, 2019, pp. 1-6, doi: 10.1109/INMIC48123.2019.9022782.

[10] Alamoodi, A. & Albahri, O.s & Albahri, A.s & Bahaa, Bilal & Malik, Rami & Zaidan, A. & Alaa, Musaab & Mohammed, K. & Hammed, Hamsa & Al-Qays, Z & Alqaysi, M.. (2020). Sentiment analysis and its applications in fighting COVID-19 and infectious diseases: A systematic review. Expert Systems with Applications. 10.1016/j.eswa.2020.114155.

[11] F. Mehmood, M. U. Ghani, M. A. Ibrahim, R. Shahzadi, W. Mahmood and M. N. Asim, "A Precisely Xtreme-Multi Channel Hybrid Approach for Roman Urdu Sentiment Analysis," in IEEE Access, vol. 8, pp. 192740-192759, 2020, doi: 10.1109/ACCESS.2020.3030885.

[12] Mehmood, Khawar & Essam, Daryl & Shafi, Kamran. (2020). Sentiment Analysis System for Roman Urdu.

[13] Pokharel, Bishwo Prakash. (2020). Twitter Sentiment Analysis During Covid-19 Outbreak in Nepal. SSRN Electronic Journal. 10.2139/ssrn.3624719.

[14] Mukhtar, Neelam & Khan, Mohammad Abid. (2017). Urdu Sentiment Analysis Using Supervised Machine Learning Approach. International Journal of Pattern Recognition and Artificial Intelligence

[15] Ghulam, Hussain & Zeng, Feng & Li, Wenjia & Xiao, Yutong. (2019). Deep Learning-Based Sentiment Analysis for Roman Urdu Text. Procedia Computer Science. 147. 131-135. 10.1016/j.procs.2019.01.202.

[16]Mehmood, Khawar & Essam, Daryl & Shafi, Kamran. (2018). Sentiment Analysis for a Resource Poor Language—Roman Urdu

[17] Asad Khattak, Muhammad Zubair Asghar, Anam Saeed, Ibrahim A. Hameed, Syed Asif Hassan, Shakeel Ahmad, A survey on sentiment analysis in Urdu: A resource-poor language, Egyptian Informatics Journal, 2020,

[18] T. Wang, K. Lu, K. P. Chow and Q. Zhu, "COVID-19 Sensing: Negative Sentiment Analysis on Social Media in China via BERT Model," in IEEE Access, vol. 8, pp. 138162-138169, 2020

[19] Xue, Jia & Chen, Junxiang & Chen, Chen & Zheng, Chengda & Li, Sijia & Zhu, Tingshao. (2020). Public discourse and sentiment during the COVID 19 pandemic: Using Latent Dirichlet Allocation for topic modeling on Twitter. PloS one. 15. e0239441. 10.1371/journal.pone.0239441.

[20] Jang, Hyeju & Rempel, Emily & Carenini, Giuseppe & Janjua, Naveed. (2020). Exploratory Analysis of COVID-19 Related Tweets in North America to Inform Public Health Institutes.

[21]Nemes, László & Kiss, Attila. (2020). Social media sentiment analysis based on COVID-19. 1-15. 10.1080/24751839.2020.1790793.

[22] Samuel, Jim & Ali, G. G. Md. Nawaz & Rahman, Md. Mokhlesur & Esawi, Ek & Samuel, Yana. (2020). COVID-19 Public Sentiment Insights and Machine Learning for Tweets Classification. Information (Switzerland). 11. 1-22. 10.3390/info11060314.

[23] Bhat, Muzafar & Qadiri, Monisa & Beg, Noor-ul-Asrar & Kundroo, Majid & Ahanger, Abdul Naffi & Agarwal, Basant. (2020). Sentiment analysis of

Social Media Response on the Covid19 outbreak. Brain, Behavior, and Immunity. 87. 10.1016/j.bbi.2020.05.006.

[24] Yin, Hui & Yang, Shuiqiao & Li, Jianxin. (2020). Detecting Topic and Sentiment Dynamics Due to COVID-19 Pandemic Using Social Media.

[25] Drias, Habiba & Drias, Yassine. (2020). Mining Twitter Data on COVID-19 for Sentiment analysis and frequent patterns Discovery (Preprint). 10.2196/preprints.19856.

[26] Hassan, S. U., Imran, M., Iftikhar, T., Safder, I., & Shabbir, M. (2017). Deep stylometry and lexical & syntactic features based author attribution on plot digital repository. International conference on Asian digital libraries (pp. 119–127).

[27] Arif, H., Munir, K., & Salman, A. (2016). Sentiment Analysis of Roman Urdu/Hindi using supervised methods.

[28] Masroor, H., Saeed, M., Feroz, M., Ahsan, K., & Islam, K. (2019). Transtech: Development of a novel translator for Roman Urdu to English. Heliyon, 5(5), e01780.

[29] Luo, Z., Huang, S., & Zhu, K. Q. (2019). Knowledge empowered prominent aspect extraction from product reviews. Information Processing & Management, 56(3), 408–423.

[30] Lytos, A., Lagkas, T., Sarigiannidis, P., & Bontcheva, K. (2019). The evolution of argumentation mining: From models to social media and emerging tools. Information Processing & Management, 56(6), 102055.

[31] Javed, I., Afzal, H., Majeed, A., & Khan, B. (2014). Towards the creation of linguistic resources for bilingual sentiment analysis of twitter data. International Conference on Applications of Natural Language to Data Bases/Information Systems (pp. 232–236). Springer.

[32] De Choudhury, M.; Counts, S.; Horvitz, E. Predicting Postpartum Changes in Emotion and

Behavior via Social Media. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, Paris, France, 27 April–2 May 2013; pp. 3267–3276.

[33] Zhang, L., Wang, S., & Liu, B. (2018). Deep learning for sentiment analysis: A survey. Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, 8(4), e1253.

[34] Vuong, T., Saastamoinen, M., Jacucci, G., & Ruotsalo, T. (2019). Understanding user behavior in naturalistic information search tasks. Journal of the Association for Information Science and Technology.

[35] Tuarob, S., & Mitranont, J. L. (2017). Automatic discovery of abusive Thai language usages in social networks. November International Conference on Asian Digital Libraries 267–278.

[36] Shardlow, M., Batista-Navarro, R., Thompson, P., Nawaz, R., McNaught, J., & Ananiadou, S. (2018). Identification of research hypotheses and new knowledge from the scientific literature. BMC medical informatics and decision making, 18(1), 46.

[37] Sailunaz, K., & Alhaji, R. (2019). Emotion and sentiment analysis from Twitter text. Journal of Computational Science.

[38] Safder, I., Hassan, S. U., & Aljohani, N. R. (2018). AI cognition in searching for relevant knowledge from scholarly big data, using a multi-layer perceptron and recurrent convolutional neural network model. Companion Proceedings of The Web Conference 2018 (pp. 251–258).

[39] Qadir, H., Khalid, O., Khan, M. U., Khan, A. U. R., & Nawaz, R. (2018). An optimal ride sharing recommendation framework for carpooling services. IEEE access : practical innovations, open solutions, 6, 62296–62313.

[40] Nakov, P., Ritter, A., Rosenthal, S., Sebastiani, F., & Stoyanov, V. (2016). SemEval-2016 task 4: Sentiment analysis on Twitter.