VFAST Transactions on Software Engineering
http://vfast.org/journals/index.php/VTSE@ 2023, ISSN(e): 2309-3978, ISSN(p): 2411-6246
Volume 11, Number 1, January-March 2023 *pp:101-108*

Roman Urdu Sentiment Analysis of Songs‘ Reviews

Muhammad Aasim Qureshi[1], Muhammad Asif[2], Muhammad Farrukh Khan[3], Asad Kamal[4] and Bilal Shahid[4]

[1]Department of Computer Science, Bahria University Lahore, 54000, Pakistan
[2]Department of Computer Science, University of Luxembourg, Luxembourg City, 4365, Luxembourg
[3]Department of Computer Science, Minhaj University Lahore, Lahore, 54000, Pakistan
[4]Department of Computer Science, University of Central Punjab, Lahore, 54000, Pakistan
*Corresponding authors email address: maasimq@hotmail.com, asifhashmat255@gmail.com

ABSTRACT
To process Natural Language reviews using Machine Learning techniques is known as Sentiment Analysis. It is a way to categorize people's opinions, sentiments, and attitudes towards a specific entity. Due to easy access to the internet and smart devices, people are becoming habitual in posting reviews about any specific entity/product, they use. These reviews are very helpful for all types of users in decision-making. In the past, most of the work in Sentiment Analysis was carried out on resource-rich language but very little literature is witnessed on resource-poor languages. Very few efforts have been made to build language resources to process the Roman Urdu language. This research targets to perform Sentiment Analysis on Urdu (i.e. source-poor language) in Roman script. To perform sentiment analysis of roman urdu reviews on songs, the dataset is generated from the comments on songs. Three songs from the Sub-continent music industry opt from YouTube. After pre-processing the reviews, Roman Urdu reviews are analysed using Naïve Bayes, KNN, Decision Tree (ID3) and ANN. Naïve Bayes outperforms the other classifiers and achieved 82.41% results in terms of accuracy.

KEYWORDS
Sentiment analysis, Roman Urdu/Hinid, Reviews analysis, Machine learning, Natural language Processing

JOURNAL INFO
HISTORY: Received: December 28, 2022 Accepted: March 25, 2023 Published: March 31, 2023

INTRODUCTION

In recent years, the internet has become an important source of social interaction between people, due to easy access to computers, smartphones and high-speed internet [1]. Easy access to hand-held electronic devices with high-speed internet has popularized the use of these services for social interaction as well as for e-business [2]. Now, e-businesses have gained a lot of fame because of their easy use and time-saving process. Many organizations are working worldwide on e-business. They display their products and also provide a space for users to rate or review the product with the information the product. Users are also becoming habitual to give feedback about the specific product they used or accessed [3]. In the past, there exist some companies, which carries out a survey about the entity in hand and gather the reviews of the users about any specific entity. Their results may benefit organizations to improve their quality standards according to the users' needs or research [4]. This is an outdated way of data collection. Due to the increase in e-marketing and e-business, e-users are also increasing day by day [5]. New technological innovation in the field of the web has made this process simple through the provision of feedback mechanisms in the form of comments about the specific product [6]. This feedback helps new users to understand the object based on previous users' reviews. It can, also, helps the organization to improve its quality according to the customer's needs.

New technologies are reshaping the world. In almost every industry, this new technology is playing an important role. The entertainment industry is one of them. This industry of the sub-continent has evolved tremendously in the last few decades with the advancement of technology [7]. The Internet has an important role in its enhancement [8]. Using the internet, people are now accessing and watching plays, movies, and plays online. The attractive thing is that celebrities are directly connected to their fans through different social networks. Their fans give them feedback and comments on their social media pages as well as on websites or pages where they showcase their work.

YouTube is one of the most famous websites in this context. It is a very famous and useful tool for highlighting works through videos. Any video can be accessed easily. It also facilitates the users to give their reviews against the video [9]. These reviews can be in the form of text or/and emojis. The quality of the content provided in the video can easily be judged by reading these comments. It can be analysed manually if the number of comments is not large. However, if the number of review is too large then there is a need for an automated mechanism to analyze these reviews.

Information retrieval deals with the collection of such data from different blogs and sites where people leave their reviews [10]. Data analysis is an emerging field [11]. There exist different Machine Learning and Data Mining techniques which are used to analyse huge amounts of data [12]. It is required to analyse this data of reviews based on its polarity—positive and negative. It leads to a newly emerging field known as Sentiment Analysis.

Sentiment Analysis is the study of people's opinions, sentiments, attitudes and emotions about any specific entity [13]. It is also a way to categorize people's opinions as positive, neutral and negative towards any entity. It can be performed in different ways, e.g. Binary sentiment classification and Multi-class sentiment classification. In the Binary type of classification, sentiments could be analysed as positive/negative, like/dislike, good/bad, etc. Multi-class type of sentiment classification deals with more than two classes i.e Positive, Neutral, and Negative.

In recent years sentiment analysis can be performed in different languages and on different types of datasets i.e. movie reviews, hotel reviews, airline reviews and many more. There is not enough literature witnessed on the sentient analysis of Roman Urdu reviews on the music industry of the Sub-continent. The music industry of the sub-continent is now growing day by day. There is a huge collection of songs available on YouTube. After listening to the songs people post their reviews. Analysis of these reviews helps to judge the quality of the content in the song. Many songs have thousands and millions of reviews, so there is a need to collect (i.e. download) the first to analyse. Data collection is a fundamental component to perform any analysis [14]. Collecting these music reviews to make a reliably good dataset, and to perform Sentiment Analysis, is a challenging task. This paper is going to present a corpus of Roman Urdu of the Indo-Pak music industry with Sentiment Analysis using different Machine Learning Techniques. Reviews are collected from the comment section of YouTube using the scraper. Collected reviews consist of noisy data as well as reviews in other languages (other than Urdu like English, Hindi, Bengalese etc.) To separate the Roman Urdu reviews from the other languages' reviews, different filters are applied, details can be seen in section 4. Emojis are another way to show their feelings about anything. Some of the users post their reviews using emojis but this research only takes care of textual data only. So emojis are removed from the dataset, details can be seen in pre-processing section.

This research aims to perform roman urdu sentiment analysis on songs review. For the purpose, different Machine Learning Algorithms like Naïve Bayes, Decision Tree ID3, K-Nearest Neighbour, and Artificial Neural Networks are applied.

MOTIVATION

The music industry of Indo-Pak is growing day by day and it carries a good marketplace. Technology has a vital role in its enrichment. Reviews on the songs help to understand the quality of content provided in the song. Analysis of these reviews through a mechanism is called Sentiment Analysis. In the past, the study was carried on Sentiment Analysis in English and Chinese languages [15][16][17]. Not enough literature is witnessed in Sentiment Analysis on Roman Urdu Text [14]. Therefore there is a need to carry research on aspect level of Sentiment Analysis on Roman Urdu to improve the different business strategies according to users need. This research presented on roman urdu sentiment analysis of 3 sub-continet songs using 4 classification algorithms where Naïve Bayes outperoms the other classifiers.

LITRATURE REVIEW

In the past studies were carried out to perform sentiment analysis on resource-rich languages e.g., English, Chinese, Arabic etc. In resource-poor languages, not enough literature is witnessed on sentiment analysis i.e., Roman Urdu. Some of the related studies are presented in the following paragraphs.

In [1] sentiment analysis was performed on 10 English songs on the dataset of 369436 reviews. After performing preprocessing, the sentiment analysis was performed on aspect level. After that the sangs were rated based upom the aspects. In [18] Salsa music was used to perform an analysis of songs that contained 21651 song records. The paper focused to find the topology of the songs and songs that were classified based on their similarities whether they were part of salsa music or not.

In [19] analysis was performed on heterogeneous datasets that consist of information related to the gross domestic products of China and the United States. Dataset was extracted from different repositories that consist of currency exchange rates from Yuan to Dollar and from Dollar to Yuan. Currencies were converted by using generate values comparison and using different math rules. In [20] analyses were performed on a meta-dataset of de facto benchmark dataset for shot learning. The dataset consists of 1623 benchmark handwritten characters with 600 examples per class. KNN classifier is used to classify each query that was closest to its origin. K-NN model achieved maximum accuracy of 88.42%.

In [21] multimodal mood classification was performed on the MAC dataset consisting of audio (mp3) clips of western and Hindi songs of eight different moods. The 500 songs were included in the Hindi songs dataset which consists of 1753 audio clips. The 298 songs belonged to the Western music industry, consisting of 1111 audio clips. LibSVM achieved 58.9% in terms of precession on the Hindi music dataset and 70.5% precession on Western Music. In [22] Sentiment Analysis of Roman Urdu reviews was performed by using supervised Machine Learning algorithms on the hotel reviews. In [4] Sentiment Analysis was performed by using different Machine Learning algorithms for Roman Urdu mobile reviews.

Dimensionality Reduction of Sentiment Analysis performed was performed by using a dataset available on Kaggle free named “Bag of Words, Bag of Pop Corns” [23]. The dataset contained 25000 examples with an equal number of positive and negative examples. A survey was carried out on the sentiment detection of reviews [24]. The survey was carried out on Sentiment Detection of reviews in text. Text analysis has been a booming interest for the last

few decades back. Due to the number of re-uses, the number of reviews increasing day by day.

In [25] Semantics-based Roman Urdu Sentiment analysis was performed heterogeneous dataset collected from different blogs on different domains (movie, politics, mobile, drama, reviews, and music). In [26] RUOMIS, a system for opinion mining was proposed on a mobile dataset that was crawled from "what mobile". The reviews that are non in the Roman script were translated using different APIs. RUOMIS performed 0.427% F-score. RUOMIS did not perform well due to noisy data. In [27] Sentiment Analysis was performed and purposed a system for the Aspect level Detection of reviews. A spam detection model was used to avoid spam or noisy data. After pre-processing and annotating Support Vector Machine and Naïve Bayes were applied to check the rating of the product.

Products were classified into three categories (Low, Medium, & High). Product with 1- or 2-star ratings was classified as low product, products having 3 stars were classified as medium and a product having 4 and 5-star ratings were classified as a high product. For sentiment analysis proposed model classified the comments into two classes (Negative and Positive) quickly and correctly.

METHODOLOGY

This research is divided into six phases named below:

1. Data Collection
2. Data Pre-processing
3. Model Design/Training
4. Model Evaluation
5. Results and Evaluation

The sequence and dependencies can be seen in Figure 1.

Data Collection

Data collection is a trivial component of any analysis. For this research, text was scraped from the reviews of three songs by three different singers of the sub-continent, easily accessible on YouTube. Collected reviews are saved in CSV file format. The list of songs can be seen in Table 1.

Table 1 Artist and Song Names

Song Name	Singer	Country	Reviews
Eye to Eye	Tahir Shah	Pakistan	346
Tera Mukhra Haseen	Shahzad Roy	Pakistan	1593
Three Peg	Sharry Maan	India	49442

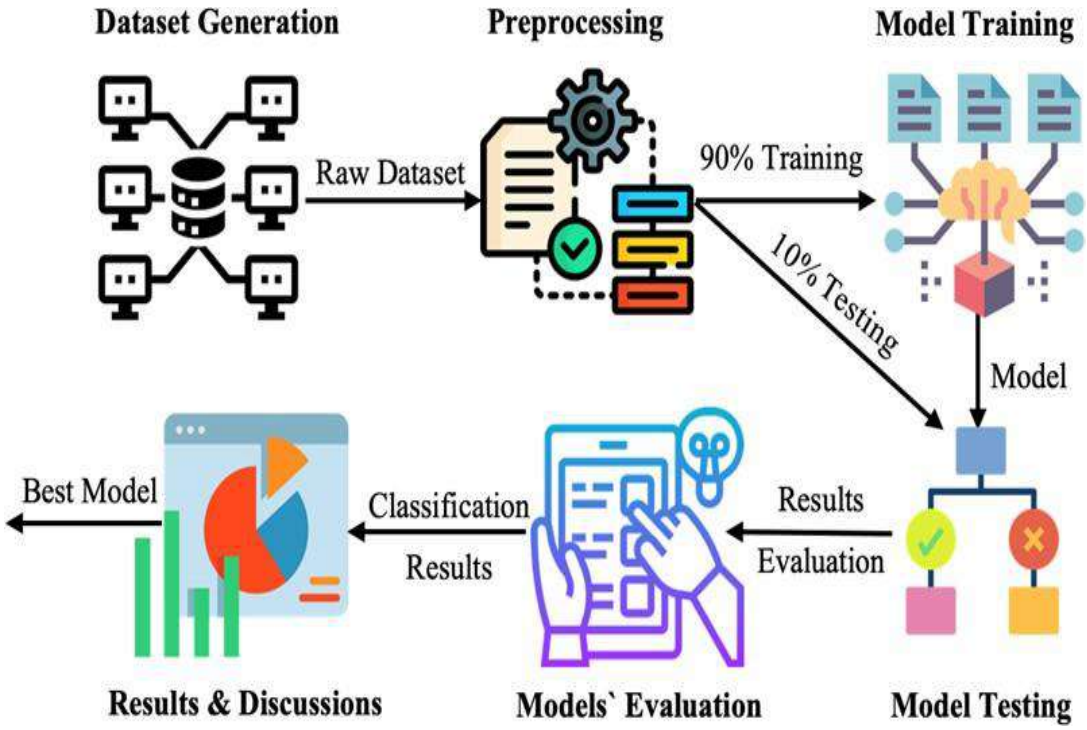

Figure 1 Proposed Methodology

Data Preprocessing

To refine data, different pre-processing techniques are applied. Details of each pre-processing technique are provided below.

Data Integration

Collected reviews from YouTube were in three different files. Handling and processing data in multiple files are more cumbersome than processing in a single file. For easy processing, data is integrated into one file [28].

Data Filtration

The reviews that are integrated into one file were in different scripts and of different languages like English, Tamil, Bengali and Punjabi as well. This study concentrated only on Roman Urdu/Hindi Sentiment Analysis, so there was a need to separate Roman Urdu/Hindi reviews from other languages' reviews. To separate Roman Urdu reviews different filters were applied to the dataset [29]. Microsoft Excel was used to apply different filters. Some sample data filters and their English translation are mentioned in Table 2. These filters filtered about 90% of the data.

Table 2. Sample data filters

Roman Urdu/Hindi	English Translation	Roman Urdu/Hindi	English Translation
jesa	As	Matlb	Means
lanat	Damn	kia	what
Bahut	Very	Mujy	Mine
wah	wow	Bura	Bad
pehly	First	Shab	Sir
zyada	More	Alla	Great
Iski	His/her	Sharm	shame

Remove Special Characters

The presence of non-alphabet characters and special characters affects the performance of the classifiers [30]. The filtered data that contains only Roman Urdu Reviews also had characters and special characters. This research is considering only text for the analysis so, all the special characters are removed from the data.

Remove Numbers

Machine learning algorithms simply classify the test examples in their training, they cannot make differences between the text data or numbers [31]. Many factors directly affect the classifier's performance. Different pieces of training lead to different outputs [32]. As discussed above this research is concentrating only on the text so all the numbers are removed to avoid accuracy issues. To avoid this issue all numeric, present in the dataset were removed.

Lowercasing

Collected reviews having the text in uppercase as well as lowercase. The term lowercasing is a way to convert all the text into lowercase because classifiers take the uppercase as separate input and lowercase as another [33]. All the dataset text is converted into lowercase to overcome the performance of the classifiers.

Remove Emojis

Many users use emojis to express their sentiments. Analysis of sentiments in the form of emojis is another research problem but this research was carried out on text sentiments. So, all the emojis present in the dataset were removed to make data free from outliers.

Remove Multi Spaces

Many factors directly affect the performance of the classifiers. White spaces that include multiple spaces, feed-line characters, and carriage return characters are one of them [34]. To overcome the performance of the classifiers, all white spaces were removed. The final dataset only contained the reviews in the form of text data.

Data Annotation

After making the dataset free from all types of noises. The next step is to annotate the data. Data annotation is a static part, in which polarity is assigned to each text according to the subjectivity expressed in the text [35]. Data annotation is performed manually with the help of three annotators. All annotators are native speakers of the Urdu language and are graduates. After the preparation of annotation guidelines, they were presented to annotators followed by discussions. Each of the texts, i.e. review is assigned a single label according to its subjectivity. The reviews are labelled as positive if the sentiment is positive by the expression [36]. A review which is negative in terms annotated as negative. Firstly, two (out of three) annotators were involved, and 100 reviews were given to each annotator. After labelling these reviews, these labels were reviewed to find the level of conflict present in annotated data which was then resolved by involving a third annotator. The conflict-of-interest value is found with the help of "Kappa statistics". The common annotator score found with the help of kappa statistics is 0.87, which is a good score [37]. Reviews are annotated as positive or negative.

This research is carried out on 600-labelled Roman Urdu reviews collected from the above-mentioned songs. The sample labelled dataset is shown in Table 3.

MODEL DESIGN

Sentiment analysis has major two types of classification. Binary classification is in which two classes (Positive and Negative) are classified and the other is a polynomial classification that deals with three classes (Positive, Negative Neutral etc.) or more. This study performed a binary type of classification using a supervised Machine Learning algorithm on a dataset of Roman Urdu/Hindi reviews. Different machine learning algorithms

like Decision Tree (ID3), Naïve Bayesian, K-Nearest Neighbor and Artificial Neural Network are applied to design models to perform Sentiment Analysis on labelled data.

Table 3. Sample Label Dataset

Comments	Class
Ghanta Bey Eye Ka!	Negative
Lo is fakir Ka ek aur gaana aa gaya	Negative
Bhenchod gaana band kr	Negative
yrr tu sab kuchh kar gaana mat gaa	Negative
maa ki aankh bhul gaya paagal	Negative
hahahaha yeh kya bala hai?	Negative
Tu gaana band kar da	Negative
3 peg ke saat maza aagaya hit like	Positive
Full Shandaar Song	Positive
mast song	Positive
wah gi wah cha gye guru nice song	Positive
Jabardast gaana hai	Positive
gazab ka songs god bless u	Positive
diwana kar diya is gaine ne	Positive

Binary classification is performed on 1990 review datasets using the above-mentioned algorithms with an equal number of Positive and Negative (995 Positive, and 995 Negative reviews) reviews. Rapid-Miner is used to implement these classifiers.

The experimental setup is designed by splitting the data into training and testing datasets. 90% data of the dataset is used to train the model and the remaining 10% data is used to evaluate the performance of the model.

Decision Tree

The decision Tree algorithm is a tree-based classification algorithm. Data is classified by using a flowchart. The tree consists of leaves and nodes and each new branch, is chosen out of two. The structure of the Decision Tree is shown in Figure 2.

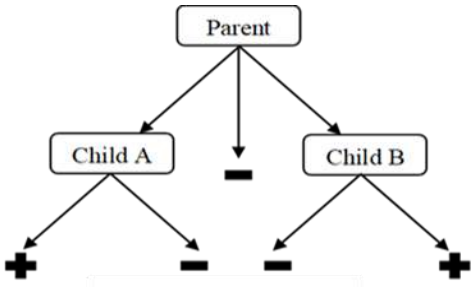

Figure 2: Decision Tree

It works on the top-down method. From the decision tree family, ID3 is used in this study, which performs well on text data. Noisy data causes over-fitting when the Decision Tree is used for the classification [38][39][40]. ID3 performs with 76.38% accuracy. The rest of the confusion matrix is presented in Table 4.

Table 4. Confusion Matrix of Decision Tree-ID3

	Actual Positive	Actual Negative
Predicted Positive	56	24
Predicted Negative	23	96

Naïve Bayes

Naïve Bayesian is a probability-based classification algorithm. It can simply work in a "Bayes theorem". It trains the model efficiently and at a fast speed. It can handle both discrete and continuous data. It calculates the posterior probability of the data and classifies the data according to the high posterior value [41][42][43]. The structure of Naïve Bayes is shown in Figure 3.

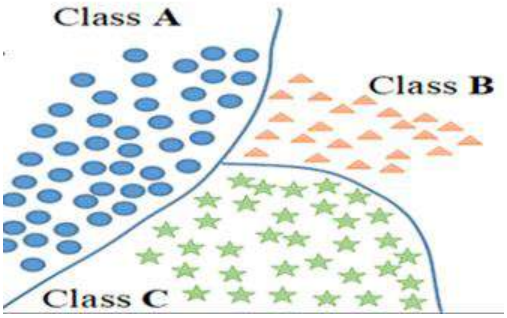

Figure 3. Naïve Bayes

Naïve Bayes correctly predicted 82.41% of reviews, with the 80.22% value of F-score. The resultant confusion matrix of Naïve Bayes is shown in Table 5.

Table 5. Confusion Matrix of Naïve Bayes

	Actual Positive	Actual Negative
Predicted Positive	71	19
Predicted Negative	16	93

KNN

K-Nearest Neighbors is used for both types of learning (Classification and Regression). It can assume the instances correspond to the points in the n-dimensional space. Make the clusters of the data near the n-dimensional space in one cluster. The structure of KNN is shown in Figure 4.

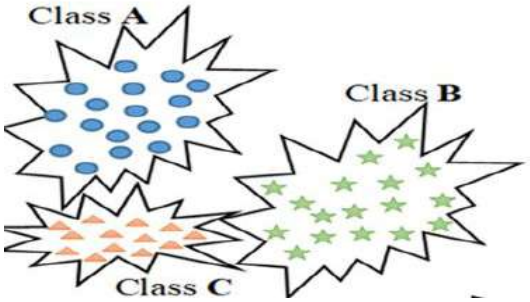

Figure 4. K-Nearest Neighbors

KNN is widely used in pattern recognition. It compares the given data instances with training data instances that are similar and are defined in terms of the standard Manhattan distance, Hamming distance, Minkowski distance and Euclidean distance. It does not perform well with missing data and it is a lazy learner because it memorizes the training data [44][45]. KNN produced results with 60.30% accuracy. The confusion matrix of KNN is shown in Table 6.

Table 6. Confusion Matrix of KNN

	Actual Positive	Actual Negative
Predicted Positive	29	32
Predicted Negative	47	91

ANN

ANN and Deep Learning algorithms are simply based on feed-forward neural networks. It can simply work on hierarchal or deep structured learning based on the multilayer feed-forward neural networks. In this study, Artificial Neural Network (ANN) is implemented because it works better on text classification. Deep Neural Networks are trained by using back-chain propagation with stochastic gradient descent that contains hidden layers. Deep Neural Nets consist of three main layers, the input layers, the hidden layers, and the output layers. Training of Neural family is a very challenging task and it is time-consuming but testing is easy instead of training [46][47][48]. The structure of ANN is shown in Figure 5.

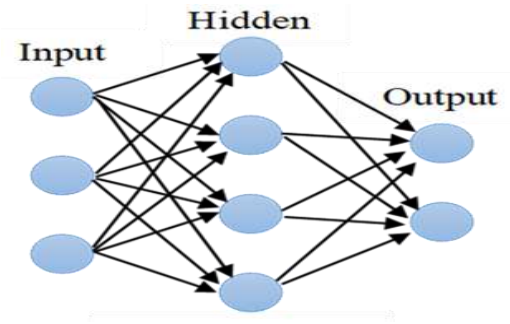

Figure 5. Artificial Neural Network

ANN correctly predicted 61.81% of reviews on the testing dataset. It predicted 123 reviews correctly out of 199 of the test datasets. The confusion matrix can be seen in Table 7.

Table 7. Confusion Matrix of ANN

	Actual Positive	Actual Negative
Predicted Positive	64	35
Predicted Negative	41	59

RESULTS AND DISCUSSION

In this study, Naïve Bayes, Decision Tree-ID3, KNN, and ANN are used to classify the data. RAPIDMINER is used to perform sentiment analysis on Roman Urdu Reviews. After testing the model on the testing dataset, the results are evaluated using 4 different evaluation measures i.e., Accuracy, Precision, Recall and F-score. In terms of accuracy, Naïve Bayes outperformed the other classifiers with 82.41% accuracy and the value of F-score is 80.23%. In precision and recall Naïve Bayes also outperforms the other classifiers with 78.79% precision and 81.60% recall. Whereas the closest competitor was Decision Tree-ID3 with 76.38% accuracy and 70.44% value of F-score. Detailed results can be seen in Table 8.

Table 8. Sample data filters

	Accuracy	Recall	Precision	Fscore
ID3	76.38	70.88	70.00	70.44
NB	**82.41**	**81.61**	**78.79**	**80.23**
KNN	60.30	38.16	47.54	42.34
ANN	61.81	60.95	64.65	62.75

A graphical analysis of the classifiers can be seen in Figure 6. Based on the accuracy and F-score Naïve Bayes outperforms the other classifiers implemented in this study.

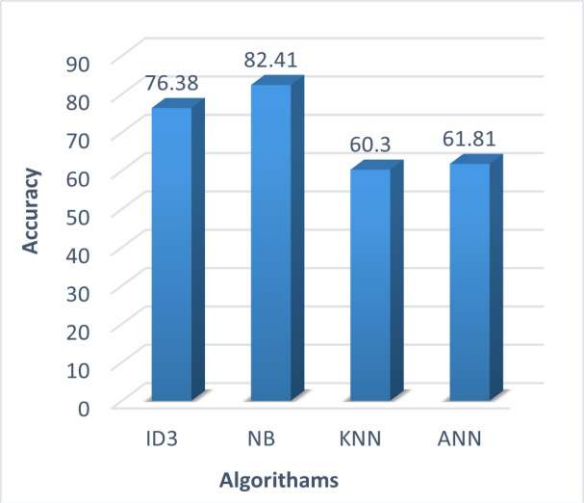

Figure 6. Comparison of Classification Accuracy's.

CONCLUSION

Sentiment analysis has attracted the researcher's attention for more than a decade. Maximum work is done in rich languages like English, Arabic, and Chinese but a few in resource-poor languages like Urdu or Hindi. To overcome the gap, this study has carried out a sentiment analysis on Roman Urdu Reviews. A review of three songs from the Sub-continent music industry is scraped. These reviews were in different languages. A Dataset of Roman Urdu/Hindi was created using the filtration technique which is followed by different pre-processings and labelling. For the classification, different machine learning algorithms—NB, ID3, KNN, and ANN are applied. NB outperformed all with 82.41% accuracy, 81.61% recall, 78.79% precision and 80.23% F-score.

FUTURE DIRECTIONS

A big dataset of roman Urdu/Hindi, either combined or separated, is required for better model design. At the same time, full-length research can be taken into account for the standardization of roman Urdu/Hindi script.

CREDIT AUTHOR STATEMENT

Muhammad Aasim Qureshi: Supervision, Writing- Reviewing and Editing, Conceptualization**: Muhammad Asif**: Conceptualization, Methodology, Data curation, Writing- Original draft preparation**: Muhammad Farrukh Khan**: Reviewing and Editing, Data curation, Visualization, Investigation**: *Asad Kamal*:** Reviewing and Editing, Data curation, Visualization, Data Annotation**: Bilal shahid**: Reviewing and Editing, Data curation, Data Annotation, Visualization

COMPLIANCE WITH ETHICAL STANDARDS

It is declared that all authors don't have any conflict of interest. Furthermore, informed consent was obtained from all individual participants included in the study.

REFERENCES

[1] M. A. Qureshi *et al.*, "Aspect Level Songs Rating Based Upon Reviews in English," *Comput. Mater. Contin.*, vol. 74, no. 2, pp. 2589–2605, 2023, doi: 10.32604/cmc.2023.032173.

[2] M. Mhatre, D. Phondekar, P. Kadam, A. Chawathe, and K. Ghag, "Dimensionality reduction for sentiment analysis using pre-processing techniques," in *Proceedings of the International Conference on Computing Methodologies and Communication, ICCMC 2017*, 2018, vol. 2018-Janua, pp. 16–21. doi: 10.1109/ICCMC.2017.8282676.

[3] Z. Papacharissi, "The Virtual Sphere 2.0: The Internet, the Public Sphere and beyond," in *Handbook of Internet Politics*, Routledge, 2009, pp. 1–35. doi: 10.1111/1478-9302.12016_66.

[4] M. Bilal, H. Israr, M. Shahid, and A. Khan, "Sentiment classification of Roman-Urdu opinions using Naïve Bayesian, Decision Tree and KNN classification techniques," *J. King Saud Univ. - Comput. Inf. Sci.*, vol. 28, no. 3, pp. 330–344, Jul. 2016, doi: 10.1016/j.jksuci.2015.11.003.

[5] R. G. Curty and P. Zhang, "Social commerce: Looking back and forward," *Proc. ASIST Annu. Meet.*, vol. 48, no. 1, pp. 1–10, 2011, doi: 10.1002/meet.2011.14504801096.

[6] J. P. Verma, B. Patel, and A. Patel, "Big data analysis: Recommendation system with hadoop framework," in *Proceedings - 2015 IEEE International Conference on Computational Intelligence and Communication Technology, CICT 2015*, Feb. 2015, pp. 92–97. doi: 10.1109/CICT.2015.86.

[7] H. L. Vogel, *Entertainment industry economics: A guide for financial analysis, ninth edition*. Cambridge University Press, 2015. doi: 10.1017/CBO9781139871679.

[8] A. Abid *et al.*, "A survey on search results diversification techniques," *Neural Comput. Appl.*, vol. 27, no. 5, pp. 1207–1229, 2016, doi: 10.1007/s00521-015-1945-5.

[9] A. Madden, I. Ruthven, and D. McMenemy, "A classification scheme for content analyses of YouTube video comments," in *Journal of Documentation*, Sep. 2013, vol. 69, no. 5, pp. 693–714. doi: 10.1108/JD-06-2012-0078.

[10] P. Chiranjeevi, D. Teja Santosh, and B. Vishnuvardhan, "Survey on sentiment analysis methods for reputation evaluation," in *Proceeding of Cognitive Informatics and Soft Computing 2017*, 2019, vol. 768, pp. 53–66. doi: 10.1007/978-981-13-0617-4_6.

[11] Y. Yao, N. Zhong, and Y. Zeng, "Information retrieval support systems," in *Understanding Information Retrieval Systems: Management, Types, and Standards*, vol. 2, Auerbach Publications, 2011, pp. 363–371. doi: 10.1201/b11499-30.

[12] M. A. Qureshi *et al.*, "A Novel Auto-Annotation Technique for Aspect Level Sentiment Analysis," *C. Mater. \& Contin.*, vol. 70, no. 3, pp. 4987–5004, 2022.

[13] K. Mehmood, D. Essam, K. Shafi, and M. K. Malik, "An unsupervised lexical normalization for Roman Hindi and Urdu sentiment analysis," *Inf. Process. Manag.*, vol. 57, no. 6, p. 102368, 2020, doi: 10.1016/j.ipm.2020.102368.

[14] M. Asif, M. A. Qureshi, A. Abid, and A. Kamal, "A Dataset for the Sentiment Analysis of Indo-Pak Music Industry," in *2019 International Conference on Innovative Computing (ICIC)*, 2019, pp. 1–6.

[15] C. Zhang, D. Zeng, J. Li, F. Y. Wang, and W. Zuo, "Sentiment analysis of chinese documents: From sentence to document level," *J. Am. Soc. Inf. Sci. Technol.*, vol. 60, no. 12, pp. 2474–2487, Dec. 2009, doi: 10.1002/asi.21206.

[16] Z. Zhang, "Sentiment Analysis of Chinese Commodity Reviews Based on Deep Learning," in *International Conference on Modern Educational Technology and Innovation and Entrepreneurship (ICMETIE 2020)*, 2020, pp. 22–28.

[17] A. Percovich, A. Tosi, L. Chiruzzo, and A. Rosa, "Ludic Applications for Language Teaching Support using Natural Language Processing," *Proc. - Int. Conf. Chil. Comput. Sci. Soc. SCCC*, vol. 2019-Novem, 2019, doi: 10.1109/SCCC49216.2019.8966429.

[18] G. M. M. Sarria, J. Diaz, and C. Arce-Lopera, "Analyzing and Extending the Salsa Music Dataset," in *2019 22nd*

Symposium on Image, Signal Processing and Artificial Vision, STSIVA 2019 - Conference Proceedings, Apr. 2019, pp. 1–5. doi: 10.1109/STSIVA.2019.8730229.

[19] J. Hendler, "Data integration for heterogenous datasets," *Big Data*, vol. 2, no. 4, pp. 205–215, Dec. 2014, doi: 10.1089/big.2014.0068.

[20] E. Triantafillou *et al.*, "Meta-Dataset: A Dataset of Datasets for Learning to Learn from Few Examples," *arXiv Prepr. arXiv1903.03096*, Mar. 2019, [Online]. Available: http://arxiv.org/abs/1903.03096

[21] B. G. Patra, D. Das, and S. Bandyopadhyay, "Multimodal mood classification of Hindi and Western songs," *J. Intell. Inf. Syst.*, vol. 51, no. 3, pp. 579–596, Dec. 2018, doi: 10.1007/s10844-018-0497-4.

[22] Z. Papacharissi *et al.*, "Sentiment Analysis of Roman Urdu/Hindi using supervised methods," *Ain Shams Eng. J.*, vol. 2, no. 3, pp. 1093–1113, 2013.

[23] M. Mhatre, D. Phondekar, P. Kadam, A. Chawathe, and K. Ghag, "Dimensionality reduction for sentiment analysis using pre-processing techniques," in *Proceedings of the International Conference on Computing Methodologies and Communication, ICCMC 2017*, Jul. 2018, vol. 2018-Janua, no. Iccmc, pp. 16–21. doi: 10.1109/ICCMC.2017.8282676.

[24] A. Yousif, Z. Niu, J. K. Tarus, and A. Ahmad, "A survey on sentiment analysis of scientific citations," *Artif. Intell. Rev.*, vol. 52, no. 3, pp. 1–34, Oct. 2017, doi: 10.1007/s10462-017-9597-8.

[25] Z. Sharf, D. Saif, and U. Rahman, "Performing Natural Language Processing on Roman Urdu Datasets," *IJCSNS Int. J. Comput. Sci. Netw. Secur.*, vol. 18, no. 1, pp. 141–148, 2018.

[26] M. Daud, R. Khan, M. Mohibullah, and A. Daud, "Roman Urdu Opinion Mining System (Ruomis)," *Comput. Sci. Eng. An Int. J.*, vol. 4, no. 6, pp. 1–9, Dec. 2014, doi: 10.5121/cseij.2014.4601.

[27] A. Bilal, A. Rextin, A. Kakakhel, and M. Nasim, "Analyzing Emergent Users' Text Messages Data and Exploring Its Benefits," *IEEE Access*, vol. 7, pp. 2870–2879, 2019, doi: 10.1109/ACCESS.2018.2885332.

[28] S. Vijayarani, M. J. Ilamathi, M. Nithya, A. Professor, and M. P. Research Scholar, "Preprocessing Techniques for Text Mining -An Overview," *Int. J. Comput. Sci. Commun. Networks*, vol. 5, no. 1, pp. 7–16, 2015.

[29] S. Yordanova and D. Kabakchieva, "Sentiment Classification of Hotel Reviews in Social Media with Decision Tree Learning," *Int. J. Comput. Appl.*, vol. 158, no. 5, pp. 1–7, Jan. 2017, doi: 10.5120/ijca2017912806.

[30] M. Syahrul and M. Dwi, "Aspect-based Sentiment Analysis to Review Products Using Naïve Bayes," vol. 020060, 2017, doi: 10.1063/1.4994463.

[31] D. Kalita, "Supervised and Unsupervised Document Classification-A survey," *Int. J. Comput. Sci. Inf. Technol.*, vol. 6, no. 2, pp. 1971–1974, 2015, [Online]. Available: www.ijcsit.com

[32] G. Qi, Z. Zhu, K. Erqinhu, Y. Chen, Y. Chai, and J. Sun, "Fault-diagnosis for reciprocating compressors using big data and machine learning," *Simul. Model. Pract. Theory*, vol. 80, pp. 104–127, Jan. 2018, doi: 10.1016/j.simpat.2017.10.005.

[33] D. Patel, S. Shah, and H. Chhinkaniwala, "Fuzzy logic based multi document summarization with improved sentence scoring and redundancy removal technique," *Expert Syst. Appl.*, vol. 134, pp. 167–177, Nov. 2019, doi: 10.1016/j.eswa.2019.05.045.

[34] D. Shubham, P. Mithil, M. Shobharani, and S. Sumathy, "Aspect level sentiment analysis using machine learning," *IOP Conf. Ser. Mater. Sci. Eng.*, vol. 263, no. 4, 2017, doi: 10.1088/1757-899X/263/4/042009.

[35] M. A. Qureshi *et al.*, "Sentiment Analysis of Reviews in Natural Language: Roman Urdu as a Case Study," *IEEE Access*, vol. 10, no. 1, pp. 24945–24954, 2022, doi: 10.1109/ACCESS.2022.3150172.

[36] M. Asif, M. Bashir, M. A. Qureshi, H. M. Zain, and M. Shoaib, "Roman Urdu Sentiment Analysis of Reviews on PSL Anthems," vol. 06, no. 03, pp. 4–11, 2022.

[37] N. Mukhtar and M. A. Khan, "Effective lexicon-based approach for Urdu sentiment analysis," *Artif. Intell. Rev.*, 2019, doi: 10.1007/s10462-019-09740-5.

[38] N. Mukhtar and M. A. Khan, "Urdu Sentiment Analysis Using Supervised Machine Learning Approach," *Int. J. Pattern Recognit. Artif. Intell.*, vol. 32, no. 02, pp. 1851001–1851007, Feb. 2017, doi: 10.1142/S0218001418510011.

[39] M. Kubat, *An Introduction to Machine Learning*, vol. 35. Cham: Springer International Publishing, 2017. doi: 10.1007/978-3-319-63913-0.

[40] M. Khan and K. Malik, *Sentiment Classification of Customer's Reviews About Automobiles in Roman Urdu BT - Advances in Information and Communication Networks*. Springer International Publishing, 2019. doi: 10.1007/978-3-030-03405-4.

[41] L. Jiang, C. Li, S. Wang, and L. Zhang, "Deep feature weighting for naive Bayes and its application to text classification," *Eng. Appl. Artif. Intell.*, vol. 52, no. 01, pp. 26–39, 2016, doi: 10.1016/j.engappai.2016.02.002.

[42] N. Ben Amor, S. Benferhat, and Z. Elouedi, "Naive Bayes vs decision trees in intrusion detection systems," in *Proceedings of the ACM Symposium on Applied Computing*, 2004, vol. 1, pp. 420–424. doi: 10.1145/967900.967989.

[43] V. Priya and K. Umamaheswari, "Ensemble based parallel k means using map reduce for aspect based summarization," *ACM Int. Conf. Proceeding Ser.*, vol. 25-26-Augu, 2016, doi: 10.1145/2980258.2980308.

[44] M. L. Zhang and Z. H. Zhou, "ML-KNN: A lazy learning approach to multi-label learning," *Pattern Recognit.*, vol. 40, no. 7, pp. 2038–2048, Jul. 2007, doi: 10.1016/j.patcog.2006.12.019.

[45] B. G. Priya, "EMOJI BASED SENTIMENT ANALYSIS USING KNN," vol. 07, no. 04, pp. 859–865, 2019.

[46] A. P. Ben Veyseh, F. Dernoncourt, D. Dou, and T. H. Nguyen, "A joint model for definition extraction with syntactic connection and semantic consistency," *AAAI 2020 - 34th AAAI Conf. Artif. Intell.*, pp. 9098–9105, 2020, doi: 10.1609/aaai.v34i05.6444.

[47] O. Obafemi, A. Stephen, O. Ajayi, and M. Nkosinathi, "A survey of artificial neural network-based prediction models for thermal properties of biomass," *Procedia Manuf.*, vol. 33, pp. 184–191, 2019.

[48] S. Timotheou, "The random neural network: A survey," *Comput. J.*, vol. 53, no. 3, pp. 251–267, Mar. 2010, doi: 10.1093/comjnl/bxp032.