

Urdu Sentiment Analysis: Future Extraction, Taxonomy, and Challenges

Mariam Mashooq, Shamyala Riaz, M.S Farooq

¹Department of Department of Computer Science, University of Management and Technology, Lahore, Punjab, Pakistan

*Corresponding author: shamyala.riaz@umt.edu.pk

ABSTRACT

By the newly gained attention from several research areas for the field of opinion mining, work in Sentiment Analysis (SA) has also been increased. Sentiment analysis is actually a natural language processing (NLP) method which is implemented to decide whether the data is negative, positive or neutral. This analysis can also utilized to provide most appropriate countermeasures for various issues that are connected with particular fields. It is a contextual extraction and arrangement of text which recognizes and pinpoints subjective information regarding source material and helps to understand the social sentiment of people while monitoring online conversations, comments, tweets, or information on blogs, etc. There is wide utilization of Urdu language in offering perspectives that's why the Urdu language also wants opinion mining as well. In this research, a systematic literature review on sentiment analysis of Urdu language has been performed. This SLR is focusing on explicit research questions and afterward contributions are described appropriately. The findings of the review present a taxonomy that is based on the techniques of sentiment classification. Furthermore, in this SLR, we have extracted all the preprocessing techniques that were used in these 24 papers, the most adopted algorithms by the researchers, the most implemented sentiment analysis approach, and the feature extraction techniques are also extricated. Eventually, a thorough survey is given on all these considerations. After a detailed and deep evaluation, we have computed their accuracy results for better understanding of future researchers.

KEY WORDS:

Sentiment analysis on Urdu language, Urdu sentiment analysis, machine learning, sentiment lexicon for Urdu language, sentiment analysis, opinion mining, prediction, classification

JOURNAL INFO

HISTORY: Received: March 19, 2022

Accepted: June 22, 2022

Published: June 30, 2022

I. INTRODUCTION

These days, Sentiment Analysis (SA) and opinion mining are commonly examined research areas. Sentiment analysis which is also called opinion mining is an application of NLP as it provides the way to use the natural language, and mining of text is performed to gather and know opinions or sentiments of people for an event. Usually, SA requires recognizing four components consisting entity, opinion holder, its aspect and his/her sentiment [1], [2]. The extricated opinions can be divided into subjective or objective text. In subjective text, text can also be sorted as positive or negative classes. Many researches implemented in SA have been conducted on human languages. NLP in Urdu is yet at the commencement period [3].

There is still need of new resources and tools. Consequently, the Urdu language yet faces difficulties in NLP tasks because of the complexity of its composition, history, and several customs. Tools and strategies in a large number, in the previous literature, are used to perform the task of sentiment analysis. Primarily, they are outlined to conduct SA in English [4]. The semantic approach collects the words containing sentiments and evaluates their polarities by the sentiment lexicon [5].

To create a novel model, Machine learning classifiers are implemented to train annotated data. After the conversion of annotated data to feature vectors, definite features are employed within a specific class. Eventually, the advanced model can be applied to speculate the class of the updated data. These techniques can be redesigned to other languages as well. The Urdu language has obtained lesser efforts as compared to other languages [6]. Though, many studies have been put forward for USA. Urdu sentiment analysis has become a common form of information extraction. These studies are helpful, such as displaying the product brand, distinguishing influencers, and fetching e-mail spam. USA has been examined in many domains, and a big number of articles have been published on the Urdu language [7].

This review proposes the SLR of the research efforts on USA. This SLR commences with describing the fundamental requirements and after reviewing different studies, this paper revealed that SVM classifier employing TFIDF with stemming within bigrams feature bettered the scenario execution in the Urdu sentiment analysis [8]. 45 papers have been originally observed to be consistent. Then the number of papers decreased to 42 articles after assessing the summaries of the studies. Ultimately, 24 papers are entirely reviewed. After a thorough survey of certain

articles, suitable knowledge has been collected. The research efforts on systematic literature review of USA are presented in this paper. Then an improved taxonomy of USA techniques is also introduced.

Section 2 of this paper is about review methodology. Findings of the SLR are given in Section 3. In section 4, the analysis and discussion are described. Section 5 contains the conclusion.

II. RESEARCH METHODOLOGY

A systematic literature review (SLR) follows an observed outlined system and the criterion is defined already before the review is carried including. An SLR distinguishes, chooses, and assesses the research in order to fulfill the formulated questions by following the research methodology [9], [10]. It is transparent research covering a lot of research done on a specific topic comprehensively within many databases and the literature that we can reproduce or recreate. This SLR is produced that works as an approach for identifying and reviewing the sentiment analysis in Urdu. It is implemented by following the systematic, explicit and precise standard that aims to summarize the current work. Moreover, also involves the element of analytical criticism. The SLR process model is shown in figure 1.

A. RESEARCH OBJECTIVES (RO)

The objectives of this review are mentioned below:

RO1: The major focus is to identify the most generally used feature extraction techniques.

RO2: To recognize the best preprocessing techniques in the most ranked articles until now in literature.

RO3: Process of prescreening of the selected papers that is showing the selection criteria after filtering them.

RO4: A generic graph for the distribution of USA reviewed articles over years.

RO5: Recognition of the common issues and unsolved challenges to acknowledge the future research prospects.

B. RESEARCH QUESTIONS (RQ)

A systematic literature review (SLR) follows an observed outlined system. The criteria are defined already before the SLR was carried out. An SLR distinguishes, chooses, and assesses the research in order to fulfill the formulated questions[11]. It is transparent research covering a lot of research done on a specific topic comprehensively within many databases and the literature that we can reproduce or recreate. This SLR is produced that works as an approach for identifying and reviewing the sentiment analysis in Urdu. It is implemented by following the systematic, explicit and precise standard that aims to summarize the current work. Moreover, also involves the element of analytical criticism. Table 1 shows the defined research question with their motivation for this systematic literature review.

1) SEARCH STRING

Practical and proper research has been performed by the usage of a string based on keywords to find and gather available literature in the domain of NLP for the Urdu language using different renowned research databases. To ensure the reliability of the search string concerning the relevancy of the articles, the main idea have been investigated by the help of research questions to achieve relevant keywords and terms used in the chosen area of study. As stated in the research questions, the search is performed with alterations of the search string. The keywords used for the search string were:

Urdu AND (Classification OR Prediction OR Polarity)
AND (Sentiment analysis OR Opinion Mining) AND
(Lexicon-based)

The articles included in this SLR vary in scope and type. It includes conference papers and journal articles, etc. The final keywords and their other similar terms (synonyms) needed to create a search string for the selection of the most suitable papers is specified in Table 2. In this table, the ‘+’ sign is used to express the inclusion of papers and the ‘-’ sign for exclusion of studies.

Table 1 RQ and major motivations

	Research Question	Major Motivation
RQ1	Which datasets are used in most of the articles?	To identify whether Urdu or Roman Urdu has been used for expressing the sentiments
RQ2	Which are the most common domains where Urdu sentiment analysis has been used?	To understand various databases and the domains that covers data relevant to Urdu language.
RQ3	What are the best preprocessing techniques used in the most ranked articles until now in literature?	To recognize the preprocessing techniques.
RQ4	What are the most involved feature extraction methods in USA?	To identify the best extraction techniques for the selection of features
RQ5	Which are the most noteworthy restrictions and gaps in the previous studies?	To understand the loopholes
RQ6	What are the objectives of future work on Urdu sentiment analysis?	To understand that gaps and loopholes and their best possible solution for best sentiment analysis.

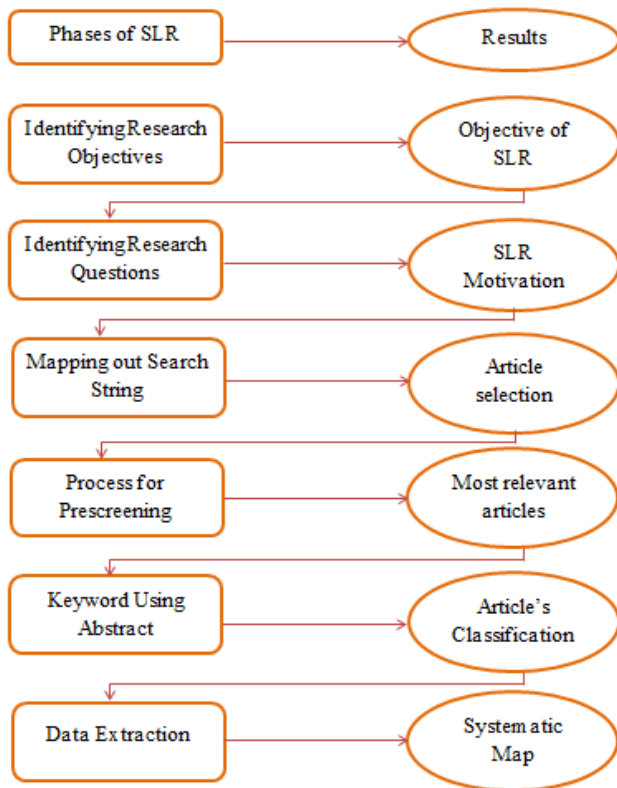


Figure 1 An SLR process model

Table 2 Terms and keywords used in search

Terms (Keywords)	Synonyms / Alternate Keywords
+ Sentiment Analysis on Urdu Language (USA)	Sentiment Analysis, Sentiment Analysis in Urdu language
+ Review on USA	Review or survey based papers on USA
- Opinion Mining or Sentiment Analysis in any other language	ASA, ESA etc.

2) INCLUSION AND EXCLUSION CRITERIA

The analyzed papers were differentiated using separate criteria that include exclusion and inclusion, represented in Table 3. The exclusion criteria is used to maintain the quality of a research paper by listing the factors that should not be included while making a standardized paper. The included factors are listed in this process of inclusion criteria and it is called practical screening.

C. SEARCH CRITERIA

The arrangement of a search plan to sufficiently find and gather noteworthy articles in the selected field is the critical phase of an SLR. This cycle involves the format of an exploration string, databases that were explored to apply for search and the exclusion (inclusion/exclusion) stage to get the most relevant papers from the literature. Various parts of the selected papers were evaluated quality

wise to illustrate different viewpoints related to the research.

Table 3 Terms and keywords used in search

Inclusion	Exclusion
Research papers that were published from November 2010 to December 2020	Non-academic articles and gray literature comprising of technical reports
Prestige academic papers, survey papers and articles	Copied and audit articles introducing research in Arabic or some other language
The articles that focus on the classification of the sentiment analysis	Papers before November 2010
The research is connected with the Urdu language	Ignored papers with weak results and accuracy
Articles written in the English language	Articles with weak writing and analysis
Articles about NLP and the applications that comprise prediction of user's interest, sentiment analysis, opinion mining	
Articles about other domains as machine learning	
Articles about some other classifications, emotions classification and recognizing a dialect	
Articles creating corpus for Urdu language	

D. SELECTION OF RELEVANT PAPERS

The fixed approach of this article involves selecting resources, defining the population, acquiring the search strings, and adding the table of inclusion and exclusion phase. The literature search of this review includes inquiring widely known journals and conferences that deal with Urdu sentiment analysis including different databases, such as IEEE Xplore, Google scholar, Elsevier, Springer, MDPI shown in table 4. The articles that were published from November 2010 to December 2020 were included for the fixed range. Figure 2 shows the chosen venues of publication.

3) Data collection:

The data is gathered from every selected paper to manage and supervise the review of Urdu sentiment analysis and includes the following:

- Explained the beginning, is it a conference or a journal
- The author's detail including their name and institutions

- The title of the paper, publication year, and researcher
- Nature of the sentiment analysis tasks that were supervised
- Dataset along with its size and other details including source as well
- Level of sentiment analysis classification
- Technique used in Sentiment Analysis
- Sentiment Analysis algorithms and its accuracy
- Types of the Urdu language
- Flow for preprocessing
- The Feature selection and process of origination
- Related applications and tools to execute the SA process

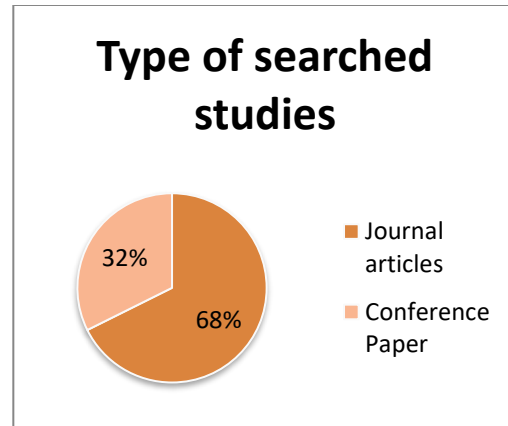


Figure 2 Type of searched studies

Table 4 Conferences, Databases and Journals explored for this supervised study

<i>Ref:</i>	<i>Conferences/journal</i>	<i>Databases</i>
[12]	International Conference on Identification, Information and Knowledge in the Internet of Things	Science Direct
[13]	International Journal of Pattern Recognition and Artificial Intelligence	World Scientific
[14]	Telematics and Informatics	Elsevier
[15]	The International Arab Journal of Information Technology, Volume 15	The International Arab Journal of Information Technology
[16]	Applied Modern Mathematics in Complex Networks	MDPI
[17]	(IJACSA) International Journal of Advanced Computer Science and Applications	Research Gate
[18]	IEEE 17th International Conference on SERA	IEEE Xplore
[19]	Mehran University Research Journal of Engineering and Technology	Mehran University Research Journal of Engineering and Technology
[20]	Artificial Intelligence Review	Springer Link
[21]	International Conference on Information Science and Communication Technology	IEEE Xplore
[22]	Information Processing & Management	Science Direct
[23]	International Symposium on Recent Advance in Electrical Engineering & Computer Sciences	IEEE Xplore
[24]	Egyptian Informatics Journal	Science Direct
[25]	Proceedings of the 1st Conference of the Asia-Pacific Chapter of the Association for Computational Linguistics and the 10th International Joint Conference on Natural Language Processing	ACL Anthology
Error! Reference source not found.	SN Computer Science: Article No-269	Springer Link

The SLR is implemented from the time period of November 2010 till December. Figure 3 illustrates the phases of the process of pre-selection of review articles in the SLR. Firstly, a study was conducted by studying and searching various journals and conferences. The reason for this study was to check the search variables. This phase proceeds to 42

research articles. Secondly, the next phase of articles was lessen to 30 by choosing the suitable articles. This stage included the procedure of diminishing similar papers and extracting non-similar articles by re-checking the titles and abstracts. Thirdly, the study was supervised one more time for some more papers at the end of December 2020

consisting 5 articles. At the end, extracting and choosing related articles were performed again after going through the full-text. 24 research articles were chosen.

4) *Data Extraction:*

This phase explains the plan for selecting data in the revised papers to explain the questions present in the research paper. Necessary data is chosen to supervise the review of Urdu sentiment analysis including the kind of SA activities, domain, source of the data, the dataset, classification level of Sentiment analysis, Approach used in SA, techniques, process for preprocessing and includes the generation process and feature selection.

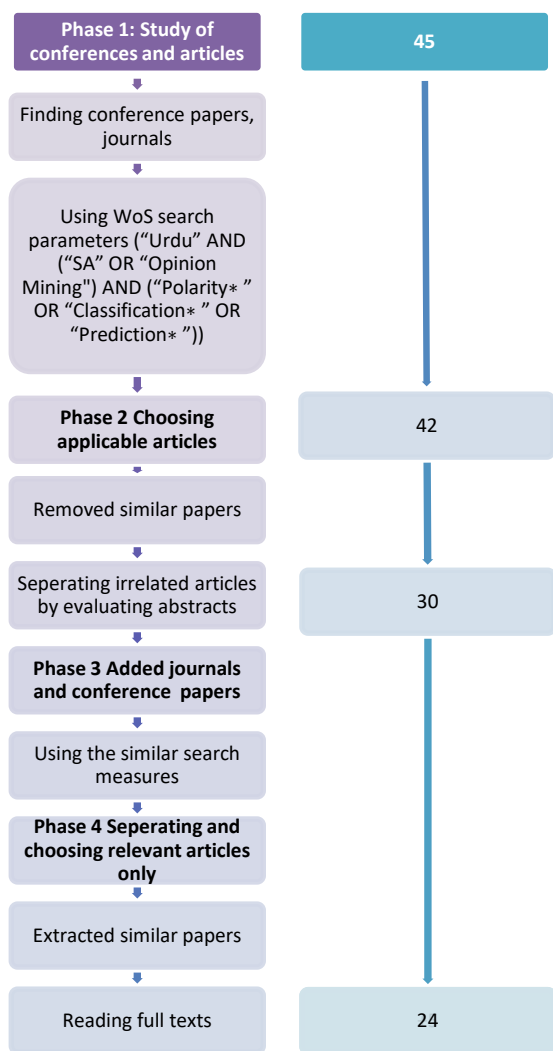


Figure 3 Prescreening of the paper

3. Related Work

In this paper, the researcher had performed lexicon-based sentiment analysis on the text of the Urdu language by the usage of SentiUnits. Along with that, the

creation of a sentiment annotated lexicon was also performed. The classification model was built for the processing and classification of text.

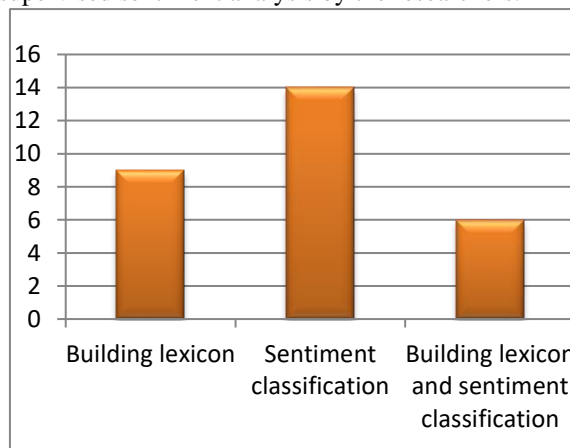
Normalization and segmentation were used as preprocessing techniques [26]. Normalization, Diacritic Omission, Tokenization, and Segmentation preprocessing techniques were implemented for extricating SentiUnits. In this paper, 82.5% of the accuracy was obtained as shallow parsing has been used for the target expressions.

This system proposed a novel framework operated on associating polarity with the sentences of the Urdu language. A sentiment classifier has been used in this research work. After reforming and improving the lexicon, 66% of the accuracy has been achieved after overall architecture in this work [27]. POS tagging and data cleaning were employed as the preprocessing techniques in this sentence-level sentiment analysis on the text of Urdu nouns and after utilizing analyzer on two different corpora. In the domain of movie review, the analyzer attained 72% of accuracy, and 78% of accuracy has been attained in the product review domain.

The sentiment analysis of the news relevant tweets in the Urdu language was focused in this paper covering the known news sources in our country. Firstly, the collection of tweets data was performed, and then a sentiment lexicon was built and used for the SA. After the transformation of text into unigram tokens, Filtration was performed. Then after the segmentation of the data, the sentiment score was calculated and achieved 77% accuracy. This lexicon will help upcoming researchers for future research comprising major knowledge. A comparison with different machine learning (ML) approaches was given and created a local dataset for the use of the public for sentiment analysis. It is an

Figure 4 Number of studies approaching SA Tasks

unsupervised sentiment analysis by the researchers.



After the detection of the sentences by identifying the lexicons, the researchers also built four sentiment lexicons having separate lexemes and roots [28].

In this paper, five classifiers were employed after the collection of data using WEKA. 10-fold cross-validation has been used after the three top classifiers including SVM, J48, and IBK. According to the results, SVM is much better than J48 for good predictions results. In this research work, the researchers used the model of Deep Neural LSTM for the training of the model of Roman Urdu sentiment analysis. In this paper, it is proved that the LSTM network is systematic and beneficial for the sequential data models. As opposed to Lexicon-based approach, it accomplished the most increased accuracy on sentiment binary classification on Roman Urdu language [29].

In this paper, [13] Urdu sentiment analysis has been performed by using a supervised machine learning approach. Firstly, Weka was used for the categorization of data. After that, three classifiers were used including Lib SVM, J48, and K-NN. The accuracy achieved in this paper was below 50% [30]. In this paper [14], Stop words removal were implemented in the preprocessing techniques. A lexicon-based approach was performed in different domains for sentiment analysis. Data was collected from the Urdu blogs and it is sentence-level sentiment analysis. 89.03% of accuracy was achieved by the usage of a lexicon-based approach. It concludes that the lexicon-based approach works well and gives the highest results instead of the supervised machine learning approach. It does not perform well in the terms of accuracy and precision.

A segmentation approach was used in this paper for the Urdu sentiment analysis. A technique carried out on sub-opinions in the text to indicate and quantify the general polarity of the sentiments in the text of Urdu language. POS Tagging, Stop words removal, and Tokenization were implemented in the preprocessing methods. 67.07% of the precision was attained in this paper [15], [31].

As explained in the previous studies, tasks were divided into sentiment classification, building lexicon.

Furthermore, it is clear from this figure 4 that the activity of sentiment classification and creating resources dealt with all the other activities as well. But, the Urdu language yet requires tools so that they can be used to help classifying sentiment.

RQ1: Which datasets were used in most of the articles?

After the step of preprocessing, sentiment analysis approaches of the literature are discussed. These preprocessing techniques cover all important phases including normalization, stemming, tokenization, named entity recognition etc. Next column displays the name of datasets such as sentiment corpora, Urdu lexicon, sentence-level lexicon, and RU Dataset. Dataset size and the number of total words, positive data, negative data is provided accordingly [32]. The dataset ranged from the size of 318 reviews up to 11000 reviews/sentences including their source as well. The data is divided into three basic types. Figure 5 shows the frequency of publications of each year. We have identified two major types of Urdu language as shown in figure 6.

Afterwards, seventh table presents the sources of the datasets involving different areas. Then the second last column describes the validation/test field. The last column lists the information about the data, whether it was publicly available or not. Table 6 shows the datasets used in different articles. The sources of data set are shown in figure 7.

In table 7, sentiment analysis approaches are presented in the third column. All details are mentioned. These approaches include semi-supervised, supervised, lexicon-based, unsupervised, and hybrid approach. Algorithms implemented in the papers are listed below. Survey of the extracted papers have been performed in below. The accuracy of each paper have been mentioned properly in percentages. The studies who had used the SVM, KNN and decision tree had achieved more accuracy.

Table 6 Datasets in the articles

Ref:	Dataset used	Dataset size	+VE	-VE	Neutral	Source	For validation/test data	Publicly available Yes/No
[12]	Urdu word embedding	Not provided			-	FastText		Yes
[13]	Sentence-level sentiment classification (Annotator 1)	1661 sentences	1493 sentences	111 sentences	57 sentences	Urdu blogs	Two annotators are used for the results of 5175 sentences	Yes
[14]	This dataset is for sentiment Analyzer	1800 sentences	600 sentences	600 sentences	600 sentences	Urdu blogs	6025 sentences are used in this research and every time a different sample is taken.	Yes
[15]	D1	443	194	249	-	Cosmetic products – Social media platforms	Overall performance was calculated by D1 and D2	Yes

[16]	D2	401	197	204	-	Electronic devices – Social media platforms	Overall performance was calculated by D1 and D2	Yes
[17]	Dataset for sentiment analyzer	2073 tweets	1245 tweets	33 tweets	795 tweets	Public accounts of Twitter	These tweets were analyzed from Text Blob, SentiWordNet and W-WSD	Yes
[18]	Urdu Lexicon	Not provided				Social networks, online forums, and newspapers		Yes
[33]	Training dataset	500 tweets	272 tweets	168 tweets	58 tweets	Twitter	500 tweets have been used by the training dataset in three categories.	Yes
[19]	RU-Dataset	11000 reviews	5686 reviews	5314 reviews	-	To gather realistic data, various online blogs, social media were used.	-	Yes
[20]	Urdu annotated corpus	6025 sentences	1876 sentences	2753 sentences	1388 sentences and 8 were discarded	Online blogs	Collected these sentences from 151 online blogs and two human annotators were hired for their labeling	Yes
[21]	Urdu sentiment lexicon	21317 words	9578 words	11739 words	Not defined	Online sources	-	Yes
[22]	Urdu lexicon	17, 185 words	8, 184 words	8, 748 words	253 words	Twitter	-	Yes
[23]	Urdu dataset	11000	5686 reviews	5314 reviews	-	Data is collected from 6 different domains		Yes
[24]	Training dataset	6000 tweets	Not provided	Not provided	-	Data is collected from the official accounts of Geo and Urdu news	Consists of five classes, each class has 1200 tweets.	Yes
[34]	Test dataset	1800 tweets	Not provided	Not provided	-	Data is collected from the official accounts of Geo and Urdu news	Consists five classes, each class has 1200 tweets	Yes
[25]	Roman Urdu dataset	454 reviews	51,61,75	29,36,93	106 Noisy reviews	Samsung, Nokia, and QMobile	This paper aims to find the polarity about these products	Yes
[35]	Urdu tweets	7000 tweets, 3000 tweets	2510 comments	2654 comments	1508 comments	Cricket domain, Football domain	In this dataset, tweets were collected and evaluation is performed on the type of their comments	Yes
Error! Reference source not found.	Urdu lexicon – In seed set	1400 tweets	213	618	569	Twitter		Yes

Table 7 Survey of the extracted data

Ref:	Lang	SA Approach	Algorithm	SA level	Domain	Accuracy (%)
------	------	-------------	-----------	----------	--------	--------------

[36]	Urdu	Sentiment annotated lexicon based approach	No algorithm is used in this paper	Senti units	Movies and main brands in Pakistan	80-85%
[37]	Urdu	Lexicon-based	Entropy Weighted Genetic Algorithm	Senti units	Movies and main brands in Pakistan	82.5% of accuracy was obtained
[38]	Urdu	Lexicon-based	Sentiment classifier	Document, sentence, aspect level	BBC News, Dawn news	66% Accuracy
[39]	Urdu	Lexicon-based	Sentiment classifier	Sentence level	Urdu news headlines	86.8% Accuracy
[40]	Urdu	Lexicon-based	Algorithm contains three stages	Tweets wise	Twitter accounts of Daily Jung, BBC Urdu, Dawn news	77% accuracy
[41]	Urdu	Sentiment lexicons	Polarity Detection, Naïve Bayes, Logistic Regression	Sentence Level	ANEW, AFINN, SenticNet, NRC	60.54% of accuracy
[42]	Urdu	Sentiment lexicon	PART, decision tree (J48), (KNN), NB multinomial text, Lib SVM	Sentence Level	Urdu blogs	Not provided
[12]	Roman urdu	Lexicon and Machine Learning approaches	RNN, LSTM	Not described	FastText	Random forest 0.88 Naive Bayes 0.77
[13]	Urdu	Lexicon-based	SVM, KNN and Decision tree	Sentence level	Urdu blogs	Below 50%
[14]	Urdu	Lexicon-based	SVM, KNN and Decision tree	Sentence level	Urdu blogs	89.03% Accuracy
[15]	Urdu	Lexicon-based	Algorithm involves polarity score calculation, sentiment segmentation and sentiment polarity identification.	Sentence level	Media forums	67.07% Accuracy
[16]	English	Machine Learning-Based Sentiment Analysis	SVM	Sentence level	Twitter	Achieved 70% accuracy after using SVM classifier
[17]	Urdu	Annotated Lexicon of Urdu	SVM	Sentence level	social network, online forums, and newspapers	Not mentioned
[18]	Urdu	Urdu sentiment	Decision Tree	Sentence level	Twitter	90%
[33]	Urdu	Urdu sentiment	DFST	Word level, character level	Hamari web, Twitter, Masala tv, etc	Results are provided in the paper
[19]	Urdu	Lexicon based	Algorithms are not used	Multiple negations	Urdu blogs	78.32%
[20]	Urdu	Lexicon based	An efficient algorithm was implemented -- using Java JDK 6	Sentence level	Urdu blogs	89.03% accuracy
[21]	Urdu	Urdu corpus	3 algorithms are defined	Document level – word level	Twitter	Not provided in percentage

[22]	Urdu and Hindi	Urdu sentiment	KNN, Decision Tree, RF, NB, ANN, SVM, AdaBoost,		BBC Urdu, Jang, Express news papers	Results are described in the paper
[23]	Urdu	Supervised and NLP approach	NB, SVM, KNN	Short length sentence	Twitter	91%, 94%, 65%
[24]	Urdu	Survey has been performed	-	-	Online forums	Performance is gauged by using two kinds of data sets in the quantitative comparison
[34]	Urdu	Lexicon based	Two algorithms are performed in this paper	Sentence level	Whatmobile.com	60%
[25]	Urdu	Aspect based	naive Bayes, Random Forest, and KNN	Aspect level	Twitter	Not defined
[35]	Urdu	Lexicon based, machine learning based	XGBoost, SVM, Naïve bayes	Sentence level	Twitter	Results are defined in the paper

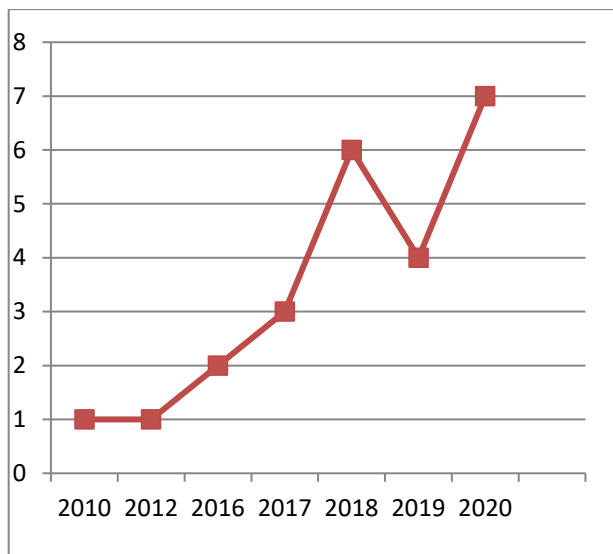


Figure 5 Distribution of Urdu Sentiment Analysis reviewed articles over years

4. Findings

This phase provides data to answer the first two questions. The phase under this section is related to the achieved outcomes from SLR. Urdu sentiment analysis is required since the percentage of Urdu audience raised a lot. 45 papers were found from November 2010 till the end of December 2020. Urdu sentiment Analysis is an interesting topic for writing a research article. It has improved a lot. As known, The Urdu language is commonly varied in Roman Urdu. Figure 5 displays the number of research articles listing the types.

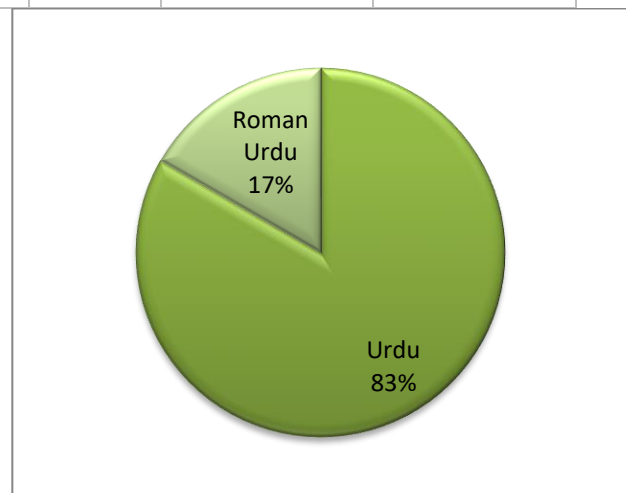


Figure 6 Urdu language types

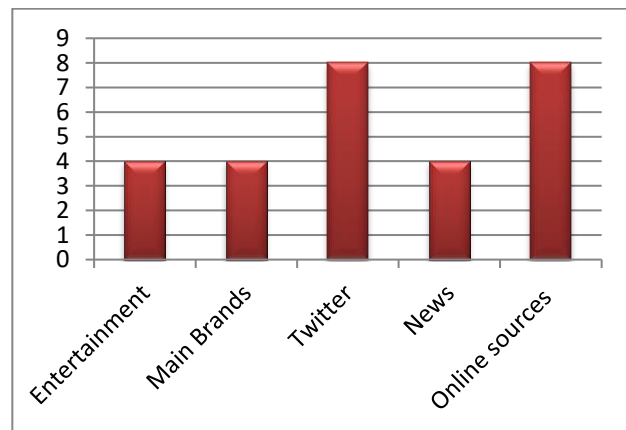


Figure 7 Sources of dataset

RQ2: Which are the best preprocessing techniques used in the most ranked articles until now in literature?

It represents the sources of datasets utilized for the USA. It is examined that most of the data in articles is used from Twitter as it is the most common application of social media used in the previous articles. But it is limited to the size of a very short message named as tweets. Machine learning (ML) approaches for the USA comprises the below phases, comprising feature generation, data preprocessing, selection etc. In the previous papers, different methods were recommended in the phase of SA to increase the results. The Urdu research conducted in all of these stages is described in the below phases. Figure 8 shows the preprocessing techniques.

1) *Preprocessing of Urdu Text:*

Urdu is a different language and a difficult challenging semantic area in Natural Language Processing. Urdu language has morphological difficulties that require advance preprocessing. Preprocessing of Urdu raw data immensely lessens the noise and enhances capability. Regrettably, most of the utmost studies focused on preprocessing of the English text. Still, some studies

concentrated on Urdu text [44]. The figure 7 explains the knowledge about the usual preprocessing procedures practiced in USA. It is observed from the figure 7 that most of the maximum studies do the preprocessing phase, including text cleaning, normalization, stemming etc. It is discovered that a big number of the studies examined stop words removal as an essential step. Meantime, text cleaning, normalization, Parts of speech tagging, and tokenization were also used by most of the articles [45]. Table 8 shows the preprocessing approaches used in this study.

2) *Urdu Sentiment Analysis Applications:*

In recent times, the Urdu language has earned significant recognition and the applications of this language are expanding in practically every feasible domain. Figure 10 presents the multiple popular domains that were targeted to implement sentiment classification. It can be observed that most researchers were preferred to apply sentiment analysis in business, political and domains as Twitter. On the other hand, the field that was used in very few domains for Urdu sentiment analysis was the entertainment domain numbering in just four articles for this domain [45].

Table 1 Steps for preprocessing conducted in the previous studies

Reference	Authors	Pre-processing Techniques
[36]	Afraz Z. Syed	Normalization, Segmentation
[37]	Afraz Z. Syed	Normalization, Diacritic Omission, Tokenization, Segmentation
[38]	Zia Ul Rehman	Tokenization, Calculate Polarity, Polarity Identification
[39]	Faiza Hashim	POS tagging, Data cleaning
[40]	Kamran Amjad	Tokenization, POS Tagging
[41]	Muhammad Yaseen Khan	Normalization, Tokenization, Stop words removal, Special characters removal
[42]	Neelam Mukhtar	Preprocessing is not performed in this paper
[12]	Hussain Ghulam	Preprocessing is not performed in this paper
[13]	Neelam Mukhtar	Stop words removal
[14]	Neelam Mukhtar	Stop words removal
[15]	Muhammad Hassan	POS Tagging, Stop words removal, Tokenization,
[16]	Ali Hasan	Stop words removal
[17]	Khairullah Khan	Noise removal, Tokenization, POS Tagging, Sentence boundary detection
[18]	Raheela Bibi	Stop words removal, POS Tagging
[33]	Khawar Mehmood	Preprocessing is not performed in this paper
[19]	Neelam Mukhtar	Preprocessing is not performed in this paper
[20]	Neelam Mukhtar	POS Tagging
[21]	Muhammad Yaseen Khan	Data cleaning, Tokenization, POS Tagging
[22]	Daryl Essam	Normalization
[23]	Faizan ul Mustafa	Tokenization, Text cleaning, Stop words removal, Stemming
[24]	Asad Khattak	Preprocessing is not performed in this paper
[34]	Halima Sadia	Remove Noise, Tokenization, Stop words removal
[25]	Sadaf Rani	Stop words removal
[35]	Zarmeen Nasim	Stop words removal

RQ3: What are the best preprocessing techniques used in the most ranked articles until now in literature?

Previous studies have proposed a broad range of techniques and methods to resolve the USA problem. This figure demonstrates the maximum techniques used in USA [46]. It is observed that SVM, NB, KNN are the extremely

employed methods in papers, where logistic regression and boosting are the lowest. SVM has been used in various earlier works appearing in 9 papers out of total papers and NB was practiced in 7 articles. It is critical to take note of that utilizing SVM classifier in the previous studies has been preceding different classifiers [32], [47].

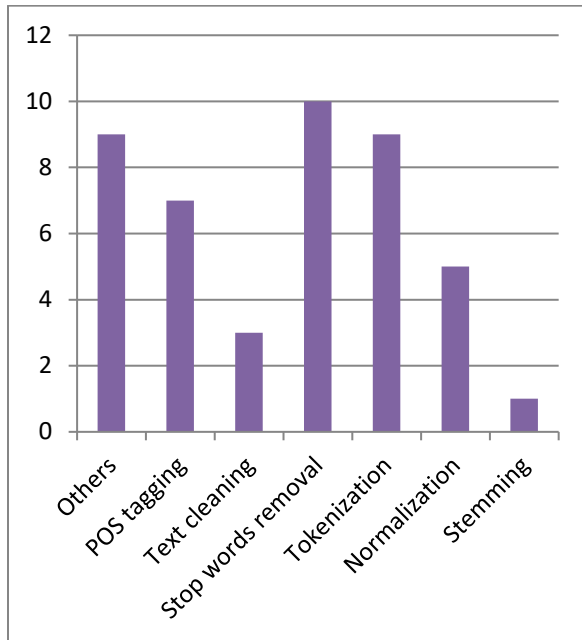


Figure 1: Preprocessing approach used in USA

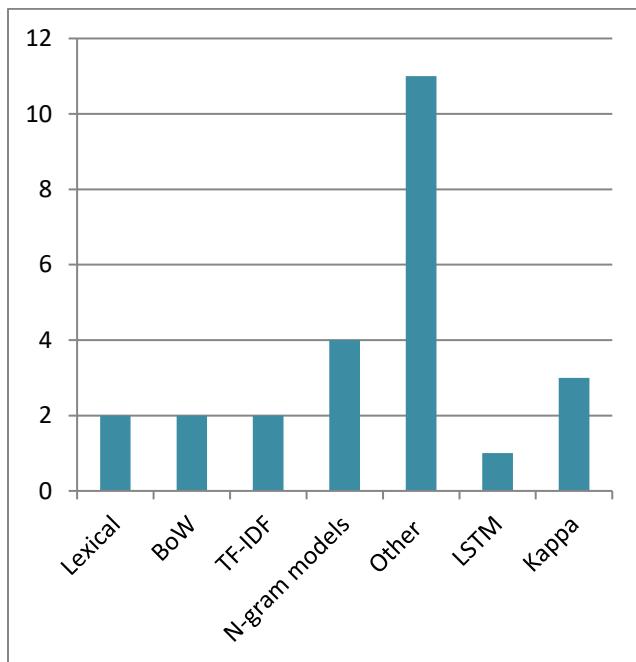


Figure 2: Frequent features in USA

RQ4: What are the most involved feature extraction methods in USA?

Machine learning offers many methods for the classification of sentiments. Nonetheless, the difficulty of obtaining sentiment in the form of written text is to choose the most useful features to be applied. Features produce extensive summarization of the results with detailed analysis. Figure 9 discloses the common features that are practiced in USA. It reveals that n-gram models are good in

the examined studies after applying the features[48] . Figure 10 shows utilized method for classification.

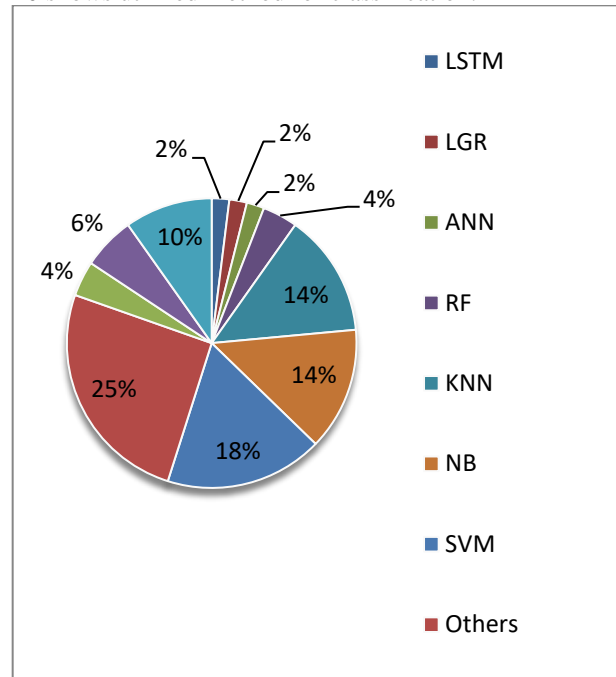


Figure 3: Methods utilized in classification

For the most part, concentrates on focused on USA applications in a fix set of areas, for example, Urdu sites, entertainment and movies, Twitter. Numerous papers were published to examine USA for different objectives such as making an Urdu sentiment dictionary, creating a structure for USA, and balancing sentiment analysis tools that support the Urdu language. These studies included handling small datasets with a capacity of less than 500 tweets that are associated to multiple domains such as politics, and news. The targeted domains for dataset are shown in figure 11.

In [18], a set of 600 tweets were obtained from the official Twitter account of Urdu News for opinion mining. The paper performed sentence-level sentiment analysis using a decision tree that divides the tweets into three classes. After the phase of preprocessing, feature vector was designed. It was formulated through type of words and presence of negation. After creating the feature vector, results of the proposed methodology explains significant.

RQ5: Which are the most noteworthy restrictions and gaps in the previous studies?

Indeed, the preprocessing step is a notable track in SA, it is however minimized and not widely included in the literature. Also, the most suitable preprocessing methods, that perform a decisive role and effects the results in USA, are still an extensive plot to read and explore [49]. A real preprocessing directs to choose a proper characteristic. Article description involves a grammatical description that is yet a tough job in this field. In this way, unsettling word semantic is likely with associative semantic plans.

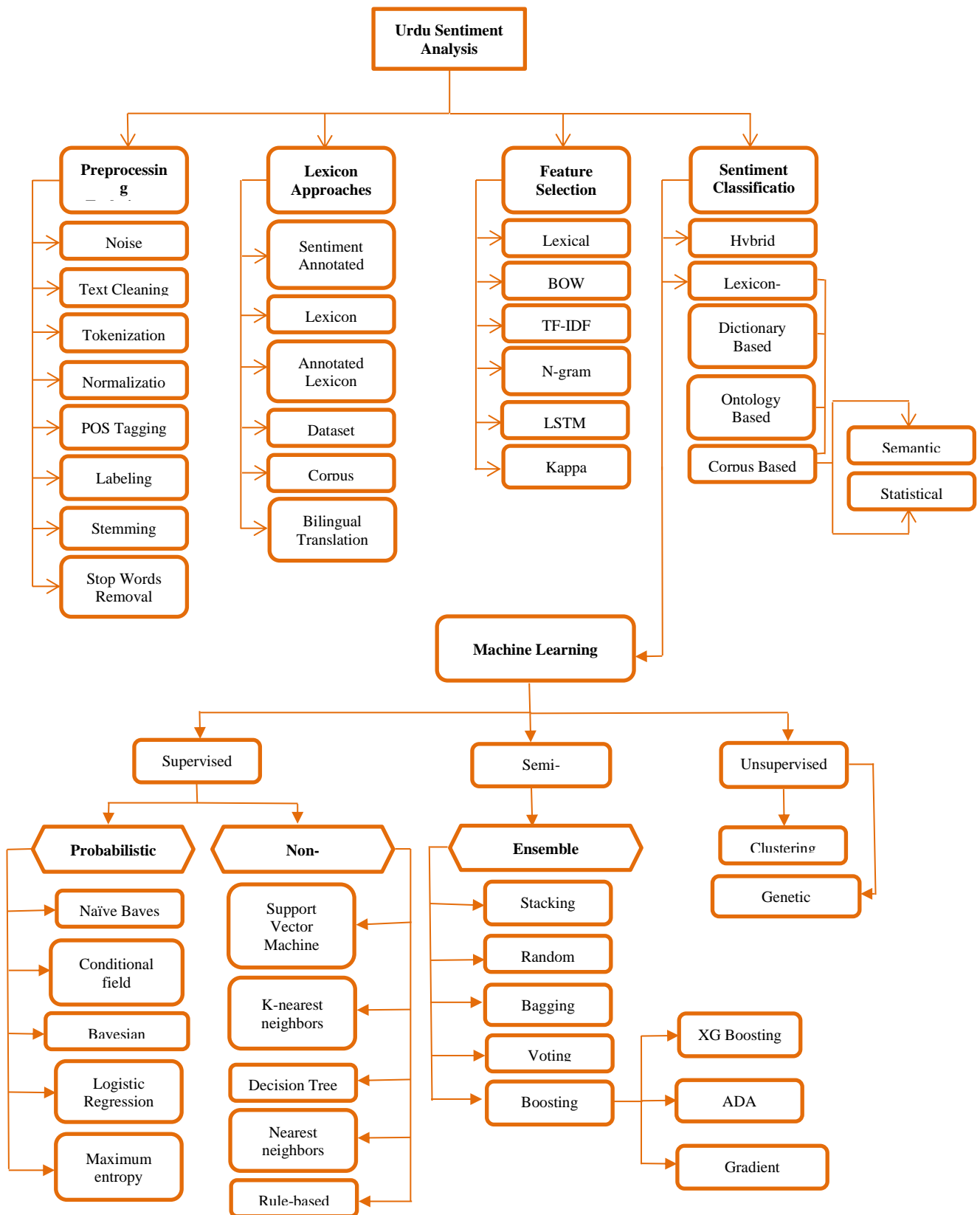


Figure 12 Taxonomy

Appending embedding of word with a class of N-gram models will foster the consequences of SA [50].

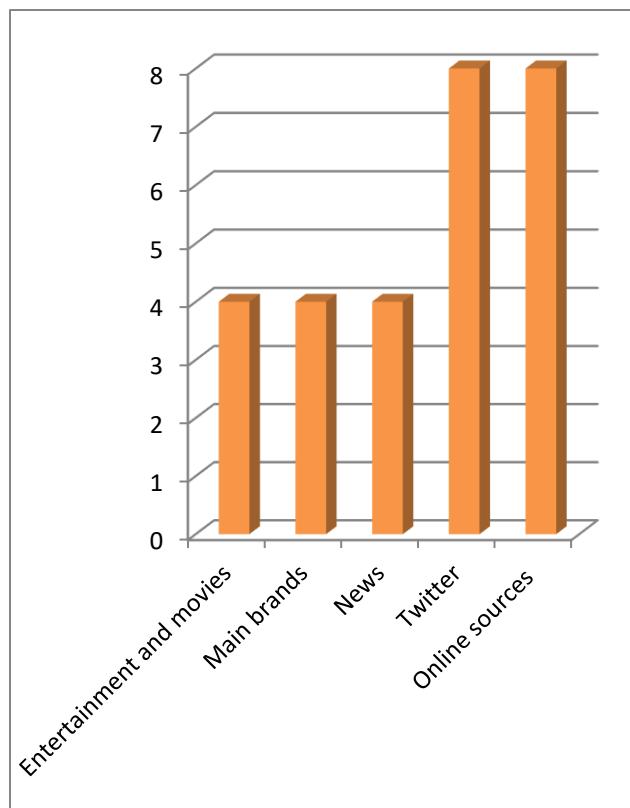


Figure 4: Targeted Domains for dataset

The motivation behind deep learning methods is to execute the algorithms of deep neural network to obtain the extreme features that are extracted a colossal measure of raw data. They need a major amount of data to implement well so that the feature extraction and inherence of assets can build the performance of deep learning models [3].

Different methodologies were proposed with the Urdu sentiment classification. Still, the efficiency of these techniques is different due to huge datasets, and articles. Applying the DL model in USA will cooperate exceedingly to resolve difficulties involved in additional well-known techniques, such as SVM and NB. Deep neural system has been fortunately chosen to obtain articles. It has the larger advantage over other Machine learning method [51].

4. DISCUSSIONS:

This section is required to consider the achieved outcomes from SLR and to present answers to the research questions. 24 papers have been interpreted in USA to achieve the present situation and obtain the purposes. Plans were compiling the efficient methods applied in USA, showing the gaps and flaws in the examined papers, and showing the goals of planned research on USA. USA has been examined by three major perspectives [52]. The first is preprocessing approaches that completely affect the

achieved outcomes of USA classification. Second process is the characteristic formation to improve the results. The next method is the sorting process that gets the vectors by characteristics formation to arrange the thoughts. Urdu data mostly include noisy and inappropriate data [53]. It is necessary to preprocess that data to enhance the results of sentiment analysis. Implementing a broad collection of preprocessing approaches, such as normalization, tokenization will enhance the classification results [54], [55].

Additionally, various papers analyzed certain preprocessing approaches, including a few weighting plans controlling individual circumstances for every individual [56]. In depth, N-gram models were applied with and without stemming categories. Further, the outcomes showed that SVM classifier utilizing TFIDF with stemming inside bigrams features highlight bettered the situation execution [57].

RQ6: What are the objectives of future work on Urdu sentiment analysis?

The survey drives that there is deficient research on making regulated datasets and executing promising classification techniques in the Urdu language. Besides, the survey additionally uncovered an absence of research from originating another feature representation that suits the characteristics of Urdu language.

3.5: Taxonomy of Urdu Sentiment Analysis Methods

To accomplish Urdu sentiment analysis, different techniques of sentiment classification were introduced. As shown in figure 12 the categorization on the classification of the Urdu sentiment analysis techniques is explained already. These implemented techniques can be categorized depending upon the strategies using machine learning [58]. ML strategies are of different categories. The unsupervised ML approaches are not convenient as it is tough to search labeled documents. All methods can be classified into semi-supervised methods. Furthermore, USA techniques vary [59]. These techniques can be based on a dictionary or corpus.

5. CONCLUSION:

In this Systematic Literature Review, the research papers on Urdu Sentiment Analysis were reviewed thoroughly. The contributions were examined concerning particular research questions. It displays a precise survey of current research in USA. 24 published research papers comprising journals, conference proceedings, and articles, are analyzed thoroughly in this SLR. USA became an important concern in the situation for preprocessing procedure, features selection, and methods for classification. The state of the literature of USA has revealed the different and popular works from diverse perspectives. This SLR focuses on the common approaches and the techniques applied in the phase of feature selection. Moreover, it describes taxonomy on the techniques of sentiment classification. It was created to explain the

research questions. By SLR, it is clear that USA still requires deeper research.

REFERENCES

- [1] A. Abbasi, H. Chen, and A. Salem, "Sentiment analysis in multiple languages: Feature selection for opinion classification in web forums," *ACM Trans. Inf. Syst.*, vol. 26, no. 3, pp. 1–34, 2008.
- [2] M. S. Farooq, M. Khan, and A. Abid, "A framework to make charity collection transparent and auditable using blockchain technology," *Comput. Electr. Eng.*, vol. 83, p. 106588, 2020, doi: 10.1016/j.compeleceng.2020.106588.
- [3] R. Tehseen, M. S. Farooq, and A. Abid, "Fuzzy expert system for earthquake prediction in western Himalayan range," *Elektron. ir Elektrotehnika*, vol. 26, no. 3, pp. 4–10, 2020, doi: 10.5755/j01.eie.26.3.25744.
- [4] O. Aziz, M. S. Farooq, A. Abid, R. Saher, and N. Aslam, "Research Trends in Enterprise Service Bus (ESB) Applications: A Systematic Mapping Study," *IEEE Access*, vol. 8, pp. 31180–31197, 2020, doi: 10.1109/ACCESS.2020.2972195.
- [5] U. Omer, M. S. Farooq, and A. Abid, "Cognitive learning analytics using assessment data and concept map: A framework-based approach for sustainability of programming courses," *Sustain.*, vol. 12, no. 17, 2020, doi: 10.3390/su12176990.
- [6] E. Mehmood, A. Abid, M. S. Farooq, and N. A. Nawaz, "Curriculum, Teaching and Learning, and Assessments for Introductory Programming Course," *IEEE Access*, vol. 8, pp. 125961–125981, 2020, doi: 10.1109/ACCESS.2020.3008321.
- [7] N. S. Khan, A. Abid, K. Abid, U. Farooq, M. S. Farooq, and H. Jameel, "Speak Pakistan: Challenges in Developing Pakistan Sign," *South Asian Stud. A Res. J. South Asian Stud.*, vol. 30, no. 2, pp. 367–379, 2015, [Online]. Available: <http://journals.pu.edu.pk/journals/index.php/IJSAS/article/view/3027>
- [8] M. S. Farooq *et al.*, "a Qualitative Framework for Introducing Programming Language At High School," *J. Qual. Technol. Manag.*, vol. VIII, no. Ii, pp. 135–151, 2012.
- [9] I. Obaid, M. S. Farooq, and A. Abid, "Gamification for Recruitment and Job Training: Model, Taxonomy, and Challenges," *IEEE Access*, vol. 8, pp. 65164–65178, 2020, doi: 10.1109/ACCESS.2020.2984178.
- [10] A. Arooj, M. S. Farooq, A. Akram, R. Iqbal, A. Sharma, and G. Dhiman, *Big Data Processing and Analysis in Internet of Vehicles: Architecture, Taxonomy, and Open Research Challenges*, vol. 29, no. 2. Springer Netherlands, 2022. doi: 10.1007/s11831-021-09590-x.
- [11] A. Dewey and A. Drahota, "Introduction to systematic reviews: online learning module Cochrane Training," Retrieved from, 2016.
- [12] H. Ghulam, F. Zeng, W. Li, and Y. Xiao, "Deep learning-based sentiment analysis for roman urdu text," *Procedia Comput. Sci.*, vol. 147, pp. 131–135, 2019.
- [13] N. Mukhtar and M. A. Khan, "Urdu sentiment analysis using supervised machine learning approach," *Int. J. Pattern Recognit. Artif. Intell.*, vol. 32, no. 02, p. 1851001, 2018.
- [14] N. Mukhtar, M. A. Khan, and N. Chiragh, "Lexicon-based approach outperforms Supervised Machine Learning approach for Urdu Sentiment Analysis in multiple domains," *Telemat. Informatics*, vol. 35, no. 8, pp. 2173–2183, 2018.
- [15] M. Hassan and M. Shoaib, "Opinion within opinion: segmentation approach for urdu sentiment analysis.," *Int. Arab J. Inf. Technol.*, vol. 15, no. 1, pp. 21–28, 2018.
- [16] A. Hasan, S. Moin, A. Karim, and S. Shamshirband, "Machine learning-based sentiment analysis for twitter accounts," *Math. Comput. Appl.*, vol. 23, no. 1, p. 11, 2018.
- [17] K. Khan, W. Khan, A. Rehman, A. Khan, and A. Khan, "Urdu sentiment analysis," *Int. J. Adv. Comput. Sci. Appl.*, vol. 9, no. 9, 2018.
- [18] R. Bibi, U. Qamar, M. Ansar, and A. Shaheen, "Sentiment Analysis for Urdu News Tweets Using Decision Tree," in *2019 IEEE 17th International Conference on Software Engineering Research, Management and Applications (SERA)*, 2019, pp. 66–70.
- [19] N. Mukhtar, M. A. Khan, N. Chiragh, A. U. Jan, and S. Nazir, "Recognition and Effective Handling of Negations in Enhancing the Accuracy of Urdu Sentiment Analyzer," *Mehran Univ. Res. J. Eng. Technol.*, vol. 39, no. 4, pp. 759–771, 2020.
- [20] N. Mukhtar and M. A. Khan, "Effective lexicon-based approach for Urdu sentiment analysis," *Artif. Intell. Rev.*, pp. 1–28, 2019.
- [21] M. Y. Khan and M. S. Nizami, "Urdu sentiment corpus (v1.0): Linguistic exploration and visualization of labeled dataset for urdu sentiment analysis," in *2020 International Conference on Information Science and Communication Technology (ICISCT)*, 2020, pp. 1–15.
- [22] 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.
- [23] F. ul Mustafa, I. Ashraf, A. Baqir, U. Ahmad, S. Malik, and S. Mehmood, "Prediction of User's Interest Based on Urdu Tweets," in *2020 International Symposium on Recent Advances in Electrical Engineering & Computer Sciences (RAEE & CS)*, 2020, vol. 5, pp. 1–6.
- [24] A. Khattak, M. Z. Asghar, A. Saeed, I. A. Hameed, S. A. Hassan, and S. Ahmad, "A survey on sentiment analysis in Urdu: A resource-poor language," *Egypt. Informatics J.*, 2020.
- [25] S. Rani and W. Anwar, "Resource Creation and Evaluation of Aspect Based Sentiment Analysis in Urdu," in *Proceedings of the 1st Conference of the Asia-Pacific Chapter of the Association for Computational Linguistics and the 10th International Joint Conference on Natural Language Processing: Student Research Workshop*, 2020, pp. 79–84.
- [26] M. S. Farooq, S. A. Khan, and A. Abid, "A Framework for the Assessment of First Programming Language," *J. Basic Appl. Acientific Res.*, vol. 2, no. 8, pp. 8144–8149, 2012.
- [27] A. Arooj, M. S. Farooq, T. Umer, G. Rasool, and B. Wang, "Cyber Physical and Social Networks in IoV (CPSN-IoV): A Multimodal Architecture in Edge-Based Networks for Optimal Route Selection Using 5G Technologies," *IEEE Access*, vol. 8, pp. 33609–33630, 2020, doi: 10.1109/ACCESS.2020.2973461.
- [28] B. Hassan, M. S. Farooq, A. Abid, and N. Sabir, "Pakistan Sign Language: Computer Vision Analysis & Recommendations," *VFAST Trans. Softw. Eng.*, vol. 9, no. 1, p. 1, 2015, doi: 10.21015/vtse.v9i1.386.

- [29] A. Arooj, M. S. Farooq, T. Umer, and R. U. Shan, "Cognitive Internet of Vehicles and disaster management: A proposed architecture and future direction," *Trans. Emerg. Telecommun. Technol.*, no. May, 2019, doi: 10.1002/ett.3625.
- [30] M. S. Farooq *et al.*, "Taxonomy and design considerations for comments in programming languages: a quality perspective," *J. Qual. Technol. Manag.*, vol. 10, no. 2, pp. 167–182, 2015.
- [31] N. Khan, A. Shahzada, S. Ata, and A. Abid, "A Vision Based Approach for Pakistan Sign Language Alphabets Recognition," *Pensee*, vol. 76, no. 3, pp. 274–285, 2014, [Online]. Available: http://www.researchgate.net/profile/Tahir_Mushtaq2/publication/260762401_A_Vision_Based_Approach_for_Pakistan_Sign_Language_alphabets_Recognition/links/0f3175349003529e8a000000.pdf
- [32] M. Ramzan, M. S. Farooq, A. Zamir, W. Akhtar, M. Ilyas, and H. U. Khan, "An Analysis of Issues for Adoption of Cloud Computing in Telecom Industries," *Eng. Technol. Appl. Sci. Res.*, vol. 8, no. 4, pp. 3157–3161, 2018, doi: 10.48084/etasr.2101.
- [33] K. Mehmood, D. Essam, K. Shafi, and M. K. Malik, "Discriminative feature spamming technique for roman Urdu sentiment analysis," *IEEE Access*, vol. 7, pp. 47991–48002, 2019.
- [34] H. Sadia *et al.*, "An efficient way of finding polarity of roman urdu reviews by using Boolean rules," *Scalable Comput. Pract. Exp.*, vol. 21, no. 2, pp. 277–289, 2020.
- [35] Z. Nasim and S. Ghani, "Sentiment Analysis on Urdu Tweets Using Markov Chains," *SN Comput. Sci.*, vol. 1, no. 5, pp. 1–13, 2020.
- [36] A. Z. Syed, M. Aslam, and A. M. Martinez-Enriquez, "Lexicon based sentiment analysis of Urdu text using SentiUnits," in *Mexican International Conference on Artificial Intelligence*, 2010, pp. 32–43.
- [37] A. Z. Syed, M. Aslam, and A. M. Martinez-Enriquez, "Associating targets with SentiUnits: a step forward in sentiment analysis of Urdu text," *Artif. Intell. Rev.*, vol. 41, no. 4, pp. 535–561, 2014.
- [38] Z. U. Rehman and I. S. Bajwa, "Lexicon-based sentiment analysis for Urdu language," in *2016 sixth international conference on innovative computing technology (INTECH)*, 2016, pp. 497–501.
- [39] F. Hashim and M. Khan, "Sentence level sentiment analysis using urdu nouns," *Dep. Comput. Sci. Univ. Peshawar, Pakistan*, pp. 101–108, 2016.
- [40] K. Amjad, M. Ishtiaq, S. Firdous, and M. A. Mehmood, "Exploring Twitter news biases using urdu-based sentiment lexicon," in *2017 International Conference on Open Source Systems & Technologies (ICOSST)*, 2017, pp. 48–53.
- [41] M. Y. Khan, S. M. Emaduddin, and K. N. Junejo, "Harnessing english sentiment lexicons for polarity detection in urdu tweets: A baseline approach," in *2017 IEEE 11th International Conference on Semantic Computing (ICSC)*, 2017, pp. 242–249.
- [42] N. Mukhtar, M. A. Khan, and N. Chiragh, "Effective use of evaluation measures for the validation of best classifier in Urdu sentiment analysis," *Cognit. Comput.*, vol. 9, no. 4, pp. 446–456, 2017.
- [43] A. Abid adnan.abid@umt.edu.pk, M. S. Farooq, I. Raza, U. Farooq, and K. Abid, "Variants of Teaching First Course in Database Systems.," *Bull. Educ. Res.*, vol. 37, no. 2, pp. 1–17, 2015, [Online]. Available: <https://login.proxy006.nclive.org/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=eue&AN=113573080&site=ehost-live&scope=site>
- [44] A. Abid, M. S. Farooq, U. Farooq, K. Abid, and M. Shafiq, "A Strategy for the Design of Introductory Computer Programming Course in High School," *J. Elem. Educ.*, vol. 25, no. 1, pp. 145–165.
- [45] R. Tehseen, M. S. Farooq, and A. Abid, "A Framework for the Prediction of Earthquake Using Federated Learning," *PeerJ Comput. Sci.*, vol. 7, pp. 1–20, 2021, doi: 10.7717/PEERJ-CS.540.
- [46] M. Attique, M. S. Farooq, A. Khelifi, and A. Abid, "Prediction of Therapeutic Peptides Using Machine Learning: Computational Models, Datasets, and Feature Encodings," *IEEE Access*, vol. 8, pp. 148570–148594, 2020, doi: 10.1109/ACCESS.2020.3015792.
- [47] A. Ali Shah, M. Khurram Ehsan, K. Ishaq, Z. Ali, and M. Shoab Farooq, "An Efficient Hybrid Classifier Model for Anomaly Intrusion Detection System," *IJCSNS Int. J. Comput. Sci. Netw. Secur.*, vol. 18, no. 11, pp. 127–136, 2018, [Online]. Available: <https://www.researchgate.net/publication/330039285>
- [48] R. Tehseen, U. Omer, and S. Farooq, "Ethical Guidelines for Artificial Intelligence: A Systematic Literature Review," *FAST Trans. Softw. Eng.*, vol. 9, no. 3, pp. 33–47, 2021, [Online]. Available: <http://vfast.org/journals/index.php/VTSE@>
- [49] U. Omer, M. S. Farooq, and A. Abid, "Introductory programming course: review and future implications," *PeerJ Comput. Sci.*, vol. 7, pp. 1–33, 2021, doi: 10.7717/PEERJ-CS.647.
- [50] K. Abid, A. Abid, M. S. Farooq, U. Farooq, and A. Abbas, "A Strategy for the Promotion of Computer Programming Using Urdu Language in Pakistan.," *South Asian Stud.*, vol. 31, no. 1, pp. 305–318, 2016, [Online]. Available: <http://search.ebscohost.com/login.aspx?direct=true&db=bth&AN=116697081&site=ehost-live&scope=site>
- [51] A. Abid, W. Ali, M. S. Farooq, U. Farooq, N. S. Khan, and K. Abid, "Semi-automatic classification and duplicate detection from human loss news corpus," *IEEE Access*, vol. 8, pp. 97737–97747, 2020, doi: 10.1109/ACCESS.2020.2995789.
- [52] L. A. Haafza, M. J. Awan, A. Abid, A. Yasin, H. Nobanee, and M. S. Farooq, "An empirical study of requirements-based test generation on an automobile control system," *Electron.*, vol. 10, no. 24, 2021, doi: 10.3390/electronics10243125.
- [53] A. Rashid, M. S. Farooq, A. Abid, T. Umer, A. K. Bashir, and Y. Bin Zikria, "Social media intention mining for sustainable information systems: categories, taxonomy, datasets and challenges," *Complex Intell. Syst.*, 2021, doi: 10.1007/s40747-021-00342-9.
- [54] M. F. Manzoor, A. Abid, M. S. Farooq, N. A. Nawaz, and U. Farooq, "Resource allocation techniques in cloud computing: A review and future directions," *Elektron. ir Elektrotehnika*, vol. 26, no. 6, pp. 40–51, 2020, doi: 10.5755/j01.eie.26.6.25865.
- [55] A. Khelifi, O. Aziz, M. S. Farooq, A. Abid, and F. Bukhari, "Social and Economic Contribution of 5G and Blockchain with Green Computing: Taxonomy, Challenges, and

- Opportunities,” *IEEE Access*, vol. 9, no. Mm, pp. 69082–69099, 2021, doi: 10.1109/ACCESS.2021.3075642.
- [56] I. A. Khawaja, A. Abid, M. S. Farooq, A. Shahzada, U. Farooq, and K. Abid, “Ad-Hoc Collaboration Space for Distributed Cross Device Mobile Application Development,” *IEEE Access*, vol. 8, pp. 62800–62814, 2020, doi: 10.1109/ACCESS.2020.2980319.
- [57] K. Ishaq, K. Ishaq, A. Abid, S. Farooq, U. Farooq, and M. Ijaz, “Use of Cloud Computing In Higher Education of Pakistan,” *Int. J. Innov. Technol. Explor. Eng.*, vol. 9, no. 2, pp. 3221–3225, 2019, doi: 10.35940/ijitee.b7730.129219.
- [58] M. Ahmed, A. Nawaz, M. Ahmed, and M. S. Farooq, “Decentralized Power Control Strategy in Microgrid for Smart Homes,” *Smart Grid Renew. Energy*, vol. 10, no. 03, pp. 43–53, 2019, doi: 10.4236/sgre.2019.103004.
- [59] D. M. Vistro, M. S. Farooq, A. U. Rehman, and S. Malik, “Smart Application Based Blockchain Consensus Protocols: A Systematic Mapping Study,” *Proc. 3rd Int. Conf. Integr. Intell. Comput. Commun. Secur. (ICIIC 2021)*, vol. 4, no. Iciiic, pp. 573–581, 2021, doi: 10.2991/ahis.k.210913.073.