WILEY | Hindawi

Research Article
Tanz-Indicator: A Novel Framework for Detection of Perso-Arabic-Scripted Urdu Sarcastic Opinions

Shabana Gul,[1] Rafi Ullah Khan [ID],[1] Mohib Ullah [ID],[1] Roman Aftab,[1] Abdul Waheed [ID],[2,3] and Tsu-Yang Wu [ID][4]

[1]Institute of Computer Science and Information Technology, The University of Agriculture, Peshawar 25150, Pakistan
[2]Department of Computer Science, Women University Swabi, Swabi 23430, Pakistan
[3]School of Electrical and Computer Engineering, Seoul National University, Seoul 08826, Republic of Korea
[4]College of Computer Science and Engineering, Shandong University of Science and Technology, Qingdao 266590, China

Correspondence should be addressed to Rafi Ullah Khan; rafiyz@gmail.com and Abdul Waheed; abdul@netlab.snu.ac.kr

Automatic sarcasm detection in textual data is a crucial task in sentiment analysis. This problem is complex because sarcastic comments usually carry the opposite meaning and are context-driven. The issue of sarcasm detection in comments written in Perso-Arabic-scripted Urdu text is even more challenging due to limited online linguistic resources. In this research, we proposed Tanz-Indicator, a lexicon-based framework to detect sarcasm in the user comments posted in Perso-Arabic Urdu language. We use a lexicon of over 3000 sarcastic tweets and 100 sarcastic features for experimentation. We also train two machine learning models with the same data to compare the performance of the lexicon-based model and machine learning-based model. The results show that the lexicon-based model correctly identified 48.5% sarcastic and 23.5% nonsarcastic tweets with the recall of 69.6% and 87.9% precision. The recall rate of Naïve Bayes and SVM-based machine learning models was 20.1% and 24.4%, respectively, with an overall accuracy of 65.2% and 60.1%, respectively.

1. Introduction

Sarcasm is defined as a combination of harsh, satirical words and gestural signs, which usually show the opposite meaning of the actual sentiment that may hurt the emotions of a receiver [1]. According to Spacey, sarcasm is an insincere statement to provoke someone. The motives of sarcastic statements are irony, flattery, insult, passive aggression, humour, satire, self-mocking, and others [2]. Sarcasm is usually expressed verbally and gesturally through facial expressions, loudness, pitch, word's vocal prolongation, and other notations in a normal conversation [1]. However, sarcasm detection in textual data is challenging due to the nonavailability of the gestures, tone, and other identification features [3]. A sarcastic sentence usually carries an opposite meaning in textual data, which humans understand due to their intellectual ability but not by the machine.

Although automatic sarcasm detection in textual data is beneficial for significant computing areas such as opinion mining, information retrieval, and market research, however, developing such a system is challenging, especially for native languages written in their native script. In this regard, many sarcasm detection models have been proposed for different languages such as Hindi [1, 4, 5], Indonesian [3], Dutch [6], English [7, 8], and Filipino [9]. However, very little attention is given to the Urdu language and especially the Perso-Arabic script of the Urdu language. It is challenging to detect sarcasm in the text written in Perso-Arabic-scripted Urdu. The nature of the sarcasm in a sentence is not fixed in most cases. Examples of sarcastic comments are shown in Figure 1 with English translation.

Urdu is the national language of Pakistan, and the majority of Pakistani users share their opinions and express their feelings on social media using either Roman Urdu

Sarcastic comments in Perso-Arabic scripted	Translation in english
ملکی ترقی کا راز مرغی کے انڈے میں چھپا ہے نئے پاکستان میں ۔	The secret of national development in new Pakistan is hidden in chicken eggs.
نئے پاکستان کا س آن ہے	The scene of new Pakistan is on.
ٹیم پہ پہرہ دے رہے	Guarding the dam.
انڈے_والی_سرکار#	Egg government
	Yes, they have gone fishing from Saqib Dam
جی وہ ثاقب ٹیم سے مچھلیاں پکڑنے گئے ہیں 🐳💧😂🤣🤣🤣	🐳💧😂🤣🤣🤣

Figure 1: Sarcastic comments in Perso-Arabic-scripted Urdu.

script or Perso-Arabic script [10]. Similarly, they reply in the same language that may contain sarcastic content. Akhter et al. proposed an offensive language detection system [11]. However, their proposed system detects offensive language in Urdu comments posted in Roman scripted Urdu. Haq et al. proposed USAD, a system to detect slang and abusive words in Perso-Arabic-scripted Urdu [10], but their proposed system was not designed to detect sarcasm in the comments. Therefore, there is an utmost need for a sarcasm detection system for comments posted in the Perso-Arabic Urdu language.

In this research, we proposed Tanz-Indicator, a framework to detect sarcasm in the user comments posted in Perso-Arabic Urdu Language. The word Tanz (طنز) is an Urdu word used for the act of giving sarcastic judgment or opinion about the qualities of someone or something [12]. Our proposed framework is a lexicon-based system to detect sarcasm in the user comments posted in Perso-Arabic Urdu language. We use a lexicon of over 3000 sarcastic tweets and 100 sarcastic features for experimentation. We also train two machine learning models with the same data to compare the performance of the lexicon-based model and machine learning-based model. The results show that the lexicon-based model correctly identified 48.5% sarcastic and 23.5% nonsarcastic tweets with the recall of 69.6% and 87.9% precision.

In contrast, the recall rate of Naïve Bayes and SVM-based machine learning models was 20.1% and 24.4%, respectively, with an overall accuracy of 65.2% and 60.1%, respectively. This research can benefit many areas of interest in natural language processing applications such as information categorization and opinion mining market research. The contributions of this work are as follows:

(1) A lexicon-based framework (Tanz-Indicator) that identifies sarcastic comments posted in Perso-Arabic-scripted Urdu Tweets is proposed

(2) A dataset composed of hashtags, punctuation marks, emojis, patterns, and words used for sarcasm in Perso-Arabic-scripted Urdu is designed and annotated

Section 2 briefly discusses the related work; Section 3 explains the architecture and working of the proposed Tanz-Indicator model. In Section 4, we discuss the experi-mentation preliminaries. Sections 5 and 6 discussed the results and conclusions with future recommendations.

2. Related Work

Automatic sarcasm detection in textual data, especially in comments posted in native languages and scripts, is a challenging issue. In this regard, many sarcasm detection models have been proposed for different languages. This section discusses the models proposed for sarcasm detection in user comments in different languages.

González-Ibáñez et al. proposed a lexicon-based sarcasm detection mechanism for tweets [13]. They build a lexicon of positive and negative words and use a string comparison mechanism to detect sarcasm in user comments. Lunando and Purwarianti proposed a sarcasm detection mechanism for the Indonesian language [3]. They used a transformed SentiWordNet framework of the English language for sentiment classification using statistical machine translation of English to the Indonesian language. Rajadesingan et al. proposed Sarcasm Classification Using a Behavioral modelling Approach (SCUBA) for sarcasm detection [14]. They used the behavioral traits of the users to detect sarcasm in their tweets. Bamman and Smith used tweets, authors, audience, and responses to detect the sarcasm [5]. They claimed that a model trained with the above features offers better accuracy than the basic model. Kunneman et al. proposed a crosslingual sarcasm detection model for English and Dutch languages [6]. They identified that sarcasm in a text comment could be easily identified using hashtags (#) and punctuation marks. Mukherjee and Bala used a supervised and unsupervised learning method with some salient features such as content words, function words, parts of speech tags, and parts of speech n-grams to distinguish a sarcastic from a nonsarcastic tweet [15]. Desai and Dave proposed a pragmatic, lexical, and linguistic feature-based model to detect sarcasm in the Hindi language [4]. Their model used hashtags, emoticons, punctuation marks, and other features to identify sarcastic statements in Hindi comments. Filatova proposed a sentiment context identification model for sarcastic comment detection. Their model classifies the comments into sarcastic or nonsarcastic using salient and nonsalient meanings of the phrases in the given context [16]. Their results showed that sentiment flow shifts could

Sarcastic features	Examples
Hashtags	#مرغی #فراڈینے #چودھری_دلال #باز_رنگ #فراڈ #نامنظور_تبدیلی #گو_نیازی_عمران_گو #پٹواری_تلنگے #پیراگون #ڈیزل #شیدہ #تارزن #فسادرفیق #انڈے #پٹوارخانہ #ککڑی #لوہے_بی_دو_انڈے #اچھے_بی_دو_انڈے #بلورانی #گنجا #بھگوڑی #سکے_کھوتے #شرلی_عابد #کھوتا_خلائ #خنزیراعظم #شریف_نامہ #یوٹرن #فضائل #یوٹرن_کے #یوتھستان #لوہے_کے_چنے
Punctuation marks	؟؟؟ ؟؟؟؟ ؟؟؟؟؟ ؟؟؟؟؟؟ ؟؟؟؟؟؟؟ ؟؟؟؟؟؟؟؟ ؟؟؟؟؟؟؟؟؟ ....! ...!!! !!! !!!! !!!!! !!!!!! ...,,, ...
Emojis	😀 😃 😄 😆 😅 😂 🤣 😂 🙃 😉 😋 😜 😝 🤪 💲 😛 🤦 🤔 🤫 😢
Sarcastic terms	فساد پٹواری استعفی
Other patterns	واہ ہاہاہاہاہاہاہا ہاہاہاہا ہاہاہاہا ہاہاہاہا بابابابابابا بابابابابابابابا ہاہاہا ہاہاہا بابابابابابابابابا بابابابابابا بابابابا بابابا بابابا واہــ

FIGURE 2: Identified features in Perso-Arabic-scripted Urdu sarcastic comments.

FIGURE 3: Architecture of the proposed Tanz-Indicator model.

Tweet classification algorithm (Tanz-Indicator).
Input: C_T, S_D
Output: Tweet Polarity
1. for $C_{Ti} \in C_T$
2. $T_{Ci} \longleftarrow Tokenize\,(C_{Ti})$
3.   for $S_{Wi} \in S_D$
4.   if $T_{Ci} == S_{Wi}$
5.   return "Sarcastic"
6.   break ;
7.   end if
8.   end for
8. end for
9. return "Non – Sarcastic"

ALGORITHM 1:

be effectively used for sarcasm detection. Bharti et al. proposed a context-based sarcasm detection model for the Hindi language [1]. They used the Twitter platform to identify the tweet context according to its temporal information. Eke et al. conducted a systematic review on sarcasm detection in textual data from 2008 to 2019 [17]. They pointed out that content and context-based linguistics are used in most of the research. The most commonly used methods to detect sarcasm are PoS tagging and n-gram, while some researches also used well-known machine learning algorithms such as maximum entropy, NB, and SVM for sarcasm classification.

Mustafa et al. proposed a user's interest prediction mechanism based on tweets posted in scripted Urdu [18]. They used natural language processing and supervised

Table 1: Dataset properties.

Total tweets	3000
Sarcastic tweets	2092
Nonsarcastic tweets	908
Sarcastic hashtags	38
Sarcastic emojis	20
Sarcastic punctuation mark patterns	18
Other sarcastic patterns	16
Tweet time period	1st October 2018–31st March 2019
Machine learning model training data	2100 (70%)
Machine learning model testing data	900 (30%)

machine learning (SVM, K-NN, and NB) to propose a user's interest prediction mechanism. Their results showed that SVM performed better than the other selected classification methods. Kolchinski and Potts proposed a user tendency-based sarcasm detection model for textual data [19]. They proposed a hybrid model of dense embedding and simple Bayesian methods to model to discover the users' tendencies and their relationship comments.

Saha et al. proposed a polarity, polarity confidence, subjectivity, and subjectivity confidence-based sarcasm detection method for English tweets [20]. Ahuja et al. conducted sarcasm detection experiments using different machine learning algorithms [21]. They train their model on three tweets based on hashtags, i.e., positive, negative, and sarcastic. They reported that involving psychological and behavioral features is helpful for better sarcasm detection in textual data. Hazarika et al. proposed CASCADE, a content and context-driven approach-based method to detect sarcasm in social media posts [22]. Their results showed that CASCADE's performance is better than the other current neural network models, such as CUE-CNN and CNN-SVM. Chiragh proposed a lexicon and supervised machine learning technique-based sentiment analysis model for blogs in Urdu [23]. They used decision tree and K-NN algorithms for classification, while for the lexicon-based method, they used Urdu sentiment and a sentiment lexicon. Their experimental results showed that the lexicon-based technique performed better than the machine learning-based model. Kumar and Harish proposed a sarcastic text detection model using k-means clustering algorithms and feature selection techniques [24].

According to the available literature, many sarcasm detection models have been proposed for different languages such as Hindi [1, 4, 5], Indonesian [3], Dutch [6], English [7, 8], and Filipino [9]. However, very little attention is given to the Urdu language and especially the Perso-Arabic script of the Urdu language. Therefore, there is an utmost need for a sarcasm detection system for comments posted in the Perso-Arabic Urdu language, as Urdu is one of the world's most popular languages.

3. Tanz-Indicator

Automatic sarcasm detection in textual data is a crucial task in sentiment analysis. This problem is complex because sar-

castic comments usually carry the opposite meaning and are context-driven. The problem of sarcasm detection in comments written in Perso-Arabic-scripted Urdu text is even more challenging due to limited online linguistic resources [10, 25]. In this research, we proposed Tanz-Indicator, a framework to detect sarcasm in the user comments posted in Perso-Arabic Urdu language. In this section, we discuss the working of the proposed Tanz-Indicator model.

3.1. Working of Tanz-Indicator. The proposed Tanz-Indicator model is divided into two significant lexicon building and testing steps. We initially collected user tweets posted in the Perso-Arabic script and performed data preprocessing. We removed URLs, stop words, mentions, and other language characters in data preprocessing. Then, we replace all the emojis with their unicode value to make the machine understandable and then tokenize the tweets as a single entity. After preprocessing the data, sarcastic features are extracted from the tweets to build the sarcastic lexicon. The sarcastic features identified in the tweets are hashtags, punctuation marks, emojis, and other patterns, as shown in Figure 2. In the data testing step, clean and processed tweets are provided to the classification module for testing. The classification module tested input tweets against the sarcastic lexicon for classifying tweets as sarcastic or nonsarcastic. The architecture of the proposed Tanz-Indicator model is shown in Figure 3. The algorithm of the classification module is shown in Algorithm 1.

C_T represents the set of training and testing comments, S_D represents the dictionary of sarcastic words, and T_{Ci} represents the tokenized comment. The contents of the training set and sarcastic dictionary are shown in Equations (1) and (2), respectively. The running time complexity of the proposed Tanz-Indicator algorithm is $O(n^2)$.

$$C_T = (C_1, C_2, C_3, C_4, \cdots . . C_n), \tag{1}$$

$$S_D = (S_{W1}, S_{W2}, S_{W3}, S_{W4}, \cdots . . S_{Wn}). \tag{2}$$

We also conducted experiments with the machine learning-based classification model for sarcasm detection in Perso-Arabic-script tweets. For testing purposes, we used two famous classification algorithms Naïve Bayes [26] and support vector machine (SVM) [27]. The testing data is also provided to the machine learning-based sarcasm detection

TABLE 2: Confusion matrix value comparison.

Confusion matrix values	Lexicon	Naïve Bayes	SVM
True positive	48.5%	8.3%	9.5%
False positive	6.7%	1.7%	2.7%
True negative	23.5%	56.9%	50.6%
False negative	21.2%	33.1%	37.2%

TABLE 3: Performance comparison of lexicon model with machine learning models.

Performance matrices	Lexicon	Naïve Bayes	SVM
Precision	0.879	0.828	0.777
Recall	0.696	0.201	0.204
F-measure	0.776	0.324	0.323
Accuracy	0.720	0.652	0.601

model for classifying tweets as sarcastic or nonsarcastic (as shown in Figure 3).

4. Experimentation Preliminaries

A Python-based lexicon building and the testing tool is developed to implement the proposed Tanz-Indicator model for sarcastic Urdu tweet detection. A sarcastic lexicon is developed using the sarcastic features that appear in user tweets, which are then used to classify tweets. The experiments are conducted on a workstation with 8 GB memory and a 2.8 GHz Intel Core i5 processor. This section discusses the dataset and lexicon generation methods, machine learning algorithms used in machine learning-based experiments, and performance evaluation parameters.

4.1. Dataset and Sarcastic Lexicon Creation. We crawled more than 3000 tweets posted in Perso-Arabic script from October 2018 to March 2019 (6 months) for sarcastic lexicon creation. After collecting the raw tweets, we processed them and manually annotated them as sarcastic and nonsarcastic. Then, we extract sarcastic features from the sarcastic tweets. The sarcastic features identified in the tweets are hashtags, punctuation marks, emojis, and other patterns, as shown in Figure 2. Similarly, the details of the dataset are given in Table 1. The same data is also used for a machine learning-based sarcasm detection model with a 70 : 30 training and testing ratio.

4.2. Algorithms Used in Machine Learning Model. We also train a machine learning-based classification model for sarcasm detection in Perso-Arabic-script tweets, apart from lexicon-based experiments. We train the machine learning model using two famous classification algorithms, Naïve Bayes and support vector machine (SVM). The Naïve Bayes algorithm is chosen because it is easy and fast to predict the class of test data set and is suitable for multiclass predictions. SVM algorithm is selected as it is not prone to catastrophic

failures and can correlate with other elements within the corpus. The Naïve Bayes model is based on the Bayes theorem, which works on conditional probability. In contrast, the support vector machine (SVM) algorithm uses a statistical learning method for data classification.

4.3. Performance Matrices. The performance of the proposed Tanz-indicator model is evaluated using standard machine learning metrics, i.e., precision, recall, F-measure, and accuracy. The mathematical representation of all the metrics is shown in the equations below.

$$\text{Precision} = \frac{\text{T.P.cases}}{\text{TP cases} + \text{FP cases}},$$

$$\text{Recall} = \frac{\text{T.P.cases}}{\text{TP cases} + \text{FN cases}},$$

$$F\text{-measure} = 2 \times \frac{\text{precision} \times \text{recall}}{\text{precision} + \text{recall}},$$

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

(3)

where T.P., TN, F.P., and F.N. stands for true positive, true negative, false positive, and false negative, respectively.

5. Results and Discussions

This research proposes Tanz-Indicator, a framework to detect sarcasm in the user comments posted in Perso-Arabic Urdu language. We build a Python-based testing environment for both lexicon and machine learning model-based testing for experimentation. We crawled more than 3000 tweets for lexicon building and built a lexicon of 2092 sarcastic tweets, 908 nonsarcastic tweets, and over 100 sarcastic features. For both lexicon and machine learning-based sarcasm detection models, 70% of data is used for training and 30% for testing. This section discusses the results of lexicon-based Tanz-Indicator and machine learning models.

The results show that the lexicon-based model correctly identified 48.5% sarcastic and 23.5% nonsarcastic tweets with the recall of 69.6% and 87.9% precision. Similarly, the Naïve Bayes-based machine learning model correctly identified 8.3% sarcastic tweets and 56.9% nonsarcastic tweets with a recall of 20.1% and 82.8% precision. While support vector machine- (SVM-) based machine learning model correctly identified 9.5% sarcastic tweets and 50.6% nonsarcastic tweets with a recall of 20.4% and 77.7% precision. The models' performance comparison results in terms of confusion matrix values are shown in Table 2, while the results of the models' performance comparison in terms of precision, recall, F-measure, and accuracy are shown in Table 3 and plotted in Figure 4.

The results showed that the lexicon-based Tanz-Indicator model performed better than the machine learning-based models. The precision rate of both Naïve Bayes- and SVM-based models are comparable to the
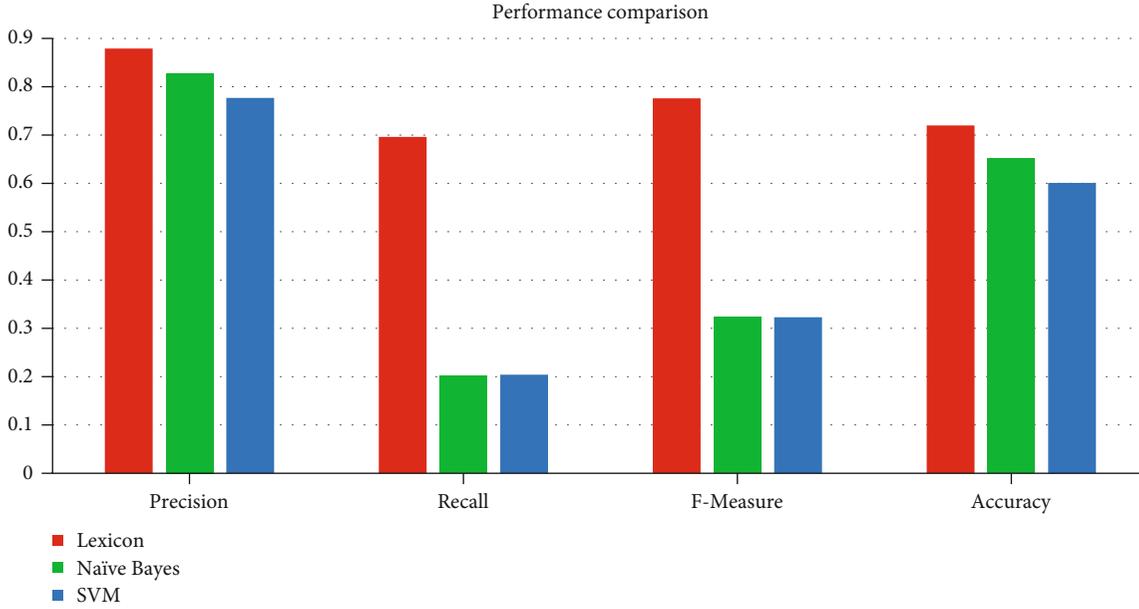
FIGURE 4: Performance comparison of lexicon model with machine learning models.

lexicon-based model. Moreover, the recall rate of both the machine learning-based models is meagre. Similarly, the accuracy of both the machine learning-based models is very low compared to that of the lexicon-based model. In both machine learning and lexicon models, the false-negative rate is high, due to which the accuracy of the models is affected. Upon investigation of misclassified tweets by both machine learning and lexicon models, it was found that the tweets were misclassified due to the limited information about sarcastic context, limited sarcastic terms in the lexicon, misspelt sarcastic words, and the limited number of sarcastic features such as sarcastic emojis, sarcastic hashtags, and sarcastic punctuation marks. The findings mentioned above are discussed in this section with examples in Figure 5.

5.1. Limited Information about Sarcastic Context.
In both types of the proposed sarcasm detection models (i.e., lexicon- and machine learning-based), one of the significant limitations is the nonavailability of the context of the sarcastic tweet. Sarcasm is usually evolved from some event, action, or conversation words. These events, actions, or conversation words provide the context for sarcasm. In this research, we aimed to identify the sarcasm in the tweet using different features. We used the hashtag feature as a sarcastic context in our lexicon-based approach, due to which the accuracy of the lexicon-based approach is far better than the machine learning-based approach. Our lexicon was built on data gathered from the specific period; therefore, the lexicon-based model could not classify tweets having hashtags developed before and after that period. In machine learning-based models, however, hashtag features were not explicitly defined. Therefore, the accuracy of the machine learning model is poorly affected due to which false-negative rate is high in those models.

5.2. Limited Number of Sarcastic Terms in the Lexicon.
Another feature for sarcasm detection is sarcastic terms closely connected with sarcastic context. There are very few specific sarcastic terms available in Urdu literature, and most of them are slang or abusive. Similarly, people also use opposite meaning words/terms in their sarcastic replies. Those words are harmless or nonabusive, but the whole sentence gave a sarcastic sense. Detecting sarcasm in these types of sentences is very much tricky without context. Therefore, another reason for the misclassification of sarcastic tweets is the limited number of sarcastic words in the lexicon and training data in this research work.

5.3. Misspelled Sarcastic Terms.
Misspelling words is a prevalent practice on social media in almost all languages. Most of the users on social media are usually careless about the spelling mistakes in their posts. It is a challenging task for a classification model to understand misspelt words. Therefore, another reason for the low accuracy of the model is misspelling terms in tweets posted by users.

5.4. Sarcastic Emojis.
One of the significant sarcasm identification features in the proposed model is emoji. Emojis are small digital images used to express an emotion or idea in textual data. Emoji is one of the powerful features through which one can easily understand the tone of the message/post. The common practice of the users to use emojis in the sarcastic post is either positive text with negative emoji or negative text with positive emoji. In this research, we only use emojis to classify the tweet as sarcastic, not the associated textual data, due to which many nonsarcastic tweets are classified as sarcastic.

5.5. Sarcastic Punctuation Marks.
Punctuation marks are another important sarcasm detection feature. The common

Tweet	Actual	Classified as	Reason
اب ڈیزل کا وزن ہوگا ترازو میں۔ سمجھدار کیلنے اشارہ کافی #متحدہ_مجلس_بے‌عمل #کھوٹے_سکے	Sarcastic	Non-sarcastic	Lack of context
قومیں #میٹرو یا #موٹروے بنانے سے نہیں #آنڈے #مرغیاں اور #کٹے پالنے سے بنتی ہیں	Sarcastic	Non-sarcastic	The sentence mage of Hashtags
سنا ہے بھاگنے کی روایت کو برقرار رکھتے ہوئے نیب سے نواز کا وکیل بھاگ گیا	Sarcastic	Non-sarcastic	Lack of any feature
ریڈیو پاکستان کی بلڈنگ لیز پر دینے کا فیصلہ واپس۔ #یو ٹرن تو اچھے ہوتے ہیں	Sarcastic	Non-sarcastic	Non-availability of hashtag in lexicon
پولیس نے نامعلوم افراد کے خلاف مقدمہ درج کرلیا	Sarcastic	Non-sarcastic	Lack of context
اگر ملک کو بچھا نہی سکتے تو استیفا دیدو۔	Sarcastic	Non-sarcastic	Misspell
جیسے اقرار بھائی سے آپ نے شادی کی ہے	Sarcastic	Non-sarcastic	Limited of words
جس سیاسی لیڈر نے پاک دھرتی کو لوٹا ہے اللہ انکا حشر فرعون اور ابوجہل کے "ساتھ فرمانے" #جیالوں #پٹواریوں #انصافیوں آؤ مل کر سب کہو "آمین	Non-sarcastic	Sarcastic	Sarcastic hashtag
یبی تھرڈ کلاس ڈویژن سے پاس ہونے والا بندہ بریڈفورڈ … کا چانسلر ربا۔۔۔۔ کسے بیوقوف بنا رہے؟ ہر کوئی پٹواری نہیں ہوتا !!!	Non-sarcastic	Sarcastic	Sarcastic punctuation marks
کہا بھی تھا 😃	Non-sarcastic	Sarcastic	Sarcastic emoji
جلیبی کی طرح سیدھا	Sarcastic	Non-sarcastic	Sarcastic proverb

FIGURE 5: Confusion tweets with their types and reasons.

practice of the users to use punctuation in sarcastic posts is to use a pattern of punctuation marks (question marks, exclamation marks, periods, and others) in the text message. This research only used punctuation mark patterns to detect sarcasm; therefore, some nonsarcastic tweets are classified as sarcastic due to a lack of information on associated text data.

5.6. Lack of Sarcastic Proverb Lexicon. The Urdu language has some proverbs used explicitly for sarcasm, unlike sarcastic terms. However, these proverbs were not very common in our collected data. As in this research, we build our lexicon and train our model based on users' data; therefore, we did not consider sarcastic proverbs, due to which some sarcastic tweets are classified as nonsarcastic.

6. Conclusions

Automatic sarcasm detection in textual data is a crucial task in sentiment analysis. In this research, we proposed Tanz-Indicator, a lexicon-based framework to detect sarcasm in the user comments posted in Perso-Arabic Urdu language. We use a lexicon of over 3000 sarcastic tweets and 100 sarcastic features (Table 1). We also train two machine learning models with the same data to compare the performance of the lexicon-based model and machine learning-based model. The results show that the lexicon-based model correctly identified 48.5% sarcastic and 23.5% nonsarcastic tweets with the recall of 69.6% and 87.9% precision.

In contrast, the recall rate of the Naïve-Bayes and SVM-based machine learning models was 20.1% and 24.4%, respectively, with an overall accuracy of 65.2% and 60.1%, respectively. It is concluded that the proposed lexicon-based Tanz-Indicator model performed better than the machine learning-based models. Despite this, the precision rate of both the Naïve Bayes and SVM-based models is com-

parable to that of the lexicon-based model. However, the recall rate of both the machine learning-based models is very low.

It was further noticed that in both machine learning and lexicon models, the false-negative rate is high, due to which the accuracy of the models is affected. Upon investigation, the tweets were misclassified due to the limited information about sarcastic context, limited sarcastic terms in the lexicon, misspelt sarcastic words, sarcastic proverbs, and the limited number of sarcastic features, sarcastic emojis, sarcastic hashtags, and sarcastic punctuation marks.

6.1. Major Findings. From the results of this research, we have drawn the following significant conclusions and findings:

(1) The hashtag is a handy feature to find contextual sarcasm and understand the tweets' context

(2) People usually use harmless or nonabusive words to pass a sarcastic comment, and detecting sarcasm in these sentences is difficult without context

(3) Including all possible misspelt sarcastic terms in a lexicon will significantly improve the model's performance

(4) The common practice of the users to use emojis in a sarcastic post is either positive text with negative emoji or negative text with positive emoji. Therefore, analysis of labelled emoji with associated textual data will also improve the model's performance

(5) A lexicon of sarcastic proverbs will improve the accuracy and performance of the proposed model

6.2. Future Work. In the future, we aim to enhance the sarcastic lexicon based on the finding mentioned in the previous section, such as the inclusion of sarcastic proverbs, misspelt sarcastic terms, and trending sarcastic hashtags. Similarly, context-based sarcasm detection in user tweets is a challenging task. Our next primary aim is to find a connection between the user profile and sarcastic tweets and develop a user profile-based sarcasm detection system for the Urdu language. Furthermore, deep neural networks produce promising results in many classification problems. Therefore, our other aim is to improve the performance of the proposed framework using deep neural network algorithms in the future.

Data Availability

All the relevant data is available in the manuscript.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

References

[1] S. K. Bharti, K. S. Babu, and R. Raman, "Context-based sarcasm detection in hindi tweets," in *In 2017 Ninth International Conference on Advances in Pattern Recognition (ICAPR)*, pp. 1–6, Bangalore, India, 2017.

[2] J. Spacey, "11 types of sarcasm," 2018, https://simplicable.com/new/sarcasm.

[3] E. Lunando and A. Purwarianti, "Indonesian social media sentiment analysis with sarcasm detection," in *In 2013 International Conference on Advanced Computer Science and Information Systems (ICACSIS)*, pp. 195–198, Sanur Bali, Indonesia, 2013.

[4] A. D. Dave and N. P. Desai, "A comprehensive study of classification techniques for sarcasm detection on textual data," in *In 2016 International Conference on Electrical, Electronics, and Optimization Techniques (ICEEOT)*, pp. 1985–1991, Chennai, India, 2016.

[5] D. Bamman and N. Smith, "Contextualized sarcasm detection on Twitter," in *In Proceedings of the International AAAI Conference on Web and Social Media*, pp. 574–577, Cologne, Germany, 2015.

[6] F. Kunneman, C. Liebrecht, M. Van Mulken, and A. Van den Bosch, "Signaling sarcasm: from hyperbole to hashtag," *Information Processing & Management*, vol. 51, no. 4, pp. 500–509, 2015.

[7] M. Bedi, S. Kumar, M. S. Akhtar, and T. Chakraborty, "Multimodal sarcasm detection and humor classification in code-mixed conversations," *IEEE Transactions on Affective Computing*, 2021.

[8] C. Techentin, D. R. Cann, M. Lupton, and D. Phung, "Sarcasm detection in native English and English as a second language speakers," *Canadian Journal of Experimental Psychology/ Revue Canadienne De Psychologie Expérimentale*, vol. 75, no. 2, pp. 133–138, 2021.

[9] M. J. C. Samonte, C. J. T. Dollete, P. M. M. Capanas, M. L. C. Flores, and C. B. Soriano, "Sentence-level sarcasm detection in English and Filipino tweets," in *In Proceedings of the 4th international conference on industrial and business engineering*, pp. 181–186, New York, 2018.

[10] N. U. Haq, M. Ullah, R. Khan et al., "USAD: an intelligent system for slang and abusive text detection in Perso-Arabic-scripted Urdu," *Complexity*, vol. 2020, 7 pages, 2020.

[11] M. P. Akhter, Z. Jiangbin, I. R. Naqvi, M. Abdelmajeed, and M. T. Sadiq, "Automatic detection of offensive language for Urdu and Roman Urdu," *IEEE Access*, vol. 8, pp. 91213–91226, 2020.

[12] K. Khullar, "Influence of October revolution on Urdu literature," *Indian Literature*, vol. 24, pp. 124–139, 1981.

[13] R. González-Ibánez, S. Muresan, and N. Wacholder, "Identifying sarcasm in twitter: a closer look," in *In Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies*, pp. 581–586, Portland Oregon, USA, 2011.

[14] A. Rajadesingan, R. Zafarani, and H. Liu, "Sarcasm detection on twitter: a behavioral modeling approach," in *In Proceedings of the eighth ACM international conference on web search and data mining*, pp. 97–106, New York, 2015.

[15] S. Mukherjee and P. K. Bala, "Sarcasm detection in microblogs using Naive Bayes and fuzzy clustering," *Technology in Society*, vol. 48, pp. 19–27, 2017.

[16] E. Filatova, "Sarcasm detection using sentiment flow shifts," in *In the Thirtieth International Flairs Conference*, Marco Island, Florida, USA, 2017.

[17] C. I. Eke, A. A. Norman, L. Shuib, and H. F. Nweke, "Sarcasm identification in textual data: systematic review, research challenges and open directions," *Artificial Intelligence Review*, vol. 53, no. 6, pp. 4215–4258, 2020.

[18] F. U. Mustafa, I. Ashraf, A. Baqir, U. Ahmad, S. Malik, and S. Mehmood, "Prediction of user's interest based on urdu tweets," in *In 2020 International Symposium on Recent Advances in Electrical Engineering & Computer Sciences (RAEE & CS)*, pp. 1–6, Islamabad, Pakistan, 2020.

[19] Y. A. Kolchinski and C. Potts, "Representing social media users for sarcasm detection," in *In Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing*, pp. 1115–1121, Brussels, Belgium,, 2018.

[20] S. Saha, J. Yadav, and P. Ranjan, "Proposed approach for sarcasm detection in twitter," *Indian Journal of Science and Technology*, vol. 10, no. 25, pp. 1–8, 2017.

[21] R. Ahuja, S. Bansal, S. Prakash, K. Venkataraman, and A. Banga, "Comparative study of different sarcasm detection algorithms based on behavioral approach," *Procedia Computer Science*, vol. 143, pp. 411–418, 2018.

[22] D. Hazarika, S. Poria, S. Gorantla, E. Cambria, R. Zimmermann, and R. Mihalcea, "CASCADE: contextual sarcasm detection in online discussion forums," in *In Proceedings of the 27th International Conference on Computational Linguistics*, pp. 1837–1848, Santa Fe, New Mexico, USA, 2018.

[23] N. Chiragh, *Enhancing accuracy of Urdu sentiments analysis, using lexicon-based approach*, University of Peshawar, Peshawar, Department of Computer Science, 2018.

[24] H. K. Kumar and B. Harish, "Sarcasm classification: a novel approach by using content based feature selection method," *Procedia Computer Science*, vol. 143, pp. 378–386, 2018.

[25] M. Daud, R. Khan, and A. Daud, "Roman Urdu opinion mining system (RUOMiS)," 2015, http://arxiv.org/abs/1501.01386.

[26] R. Khan, A. Ahmad, A. O. Alsayed, M. Binsawad, M. A. Islam, and M. Ullah, "QuPiD attack: machine learning-based privacy quantification mechanism for PIR protocols in health-related web search," *Scientific Programming*, vol. 2020, 11 pages, 2020.

[27] F. Noor, M. Bakhtyar, and J. Baber, "Sentiment analysis in e-commerce using SVM on Roman Urdu text," in *In International Conference for Emerging Technologies in Computing*, pp. 213–222, Springer, Cham, 2019.