OPEN ACCESS
International Journal of Innovations in Science & Technology

Exploring Political Emotions: Sentiment Analysis of Urdu Tweets

Ehtisham Ur Rehman[1], Muhammad Nouman Khan[2], Najam Aziz[2]
[1]Department of Computer Science UET, Peshawar, Pakistan
[2]Department of Computer Systems Engineering UET Peshawar, Pakistan
***Correspondence**: ehtishamrehman@uetpeshawar.edu.pk, eng.noumankhann@gmail.com, najamkhattak@gmail.com

Citation | Rehman. E, Khan. M. N, Aziz. N, "Exploring Political Emotions: Sentiment Analysis of Urdu Tweets", IJIST, Special Issue pp 296-311, June 2024

Received | May 15, 2024 **Revised |** May 27, 2024 **Accepted |** May 31, 2024 **Published |** June 02, 2024.

This research is a multi-text categorization based on a collection of Pakistani political texts. The major goal of this research is to use Natural Language Processing (NLP) and Machine Learning classification models to categorize multi-text for Urdu. Political tweets from 13 different Pakistani famous leaders were collected for this research. These politicians make use of the platform to promote themselves and engage with their supporters. To analyze the model accuracy the desired dataset is divided into six categories which have been composed of their official Twitter account. We also collect top trends from Pakistan and around the world to examine current trends regularly. In the proposed research, the major political corpus data comprises 1300+ tweets in the Urdu language, encompassing political policies, campaigns, opinions, and so on. Sentiment analysis is an essential component of every deep learning approach. For that, we have used the deep learning approach i.e. sentiment analysis of the politician since it provides insight into their moods and views on a certain topic. Furthermore, text corpus pre-processing is conducted utilizing NLP techniques, such as data cleaning, data balancing, and stop word removal. TF-IDF is used as word filtering for feature extraction count vectors. Machine Learning classification algorithms such as SVM, Decision Tree, XGboost, and Random Forest, and for implementation of neural network we have used Word2vector.

Keywords: Political Emotions, Sentiment Analysis, Exploring, Urdu Tweets, Social Media, Political Discourse, Emotion Detection, Computational Linguistics, Urdu Language.

Introduction:

Sentiment analysis in politics uses Natural Language Processing (NLP) and linguistic computing to understand how people feel about a political term [1]. Textual data mining is used to identify linguistic patterns that reveal an individual's emotional, affective, cognitive state, attitude toward something or someone, personality traits, and other psychological constructs. Sentiment analysis can be used to gauge the response of the masses towards a politician's political attitude on any topic. For example, if people are satisfied with the stance of a political leader on a specific policy matter, the same is expressed in their views which may be propagated through various mediums such as interviews, Facebook posts, tweets, etc. Gathering and analyzing the political sentiments of the masses is important, as public opinion has a significant impact on elections and policy. It enables citizens' views to be analyzed. Anger and delight are the two most commonly expressed emotions in politics. These feelings can be found on both sides of the political aisle, but they are often aimed at politicians [1][2].

Looking at the classical methods that were used previously were mainly based on traditional surveys. The purpose of conducting political surveys is to understand registered voters' opinions and emotions. Many different groups, including political parties, Political action committees, consultants, council members, state departments, local school districts, and candidates, use questionnaires of this type. Political survey questions can be used to learn more about the base of support and the wants and requirements of the general populace. By asking voters these questions, political campaigns and activist groups can better understand their voters' priorities and develop policies that are more likely to win their support. In addition, such surveys can better understand the political environment and political campaign strategies and increase support from potential voters [3]. However, extensive surveys need considerable money and time. Political parties now emphasize the use of social media as a better and cheaper alternative for the collection and analysis of political sentiments. Social media provides a powerfutool for monitoring and analyzing political sentiments/opinions, resulting in an increased focus on social media platforms from political parties. Recent years have seen a dramatic increase in the use of Twitter as a platform from which political opinions may be shared and a conversation with ordinary citizens can occur.

In recent times, text mining has become more relevant due to the abundance of various data kinds from various sources, primarily in the form of semi-structured and un-structured data. The main objective of text mining is to enable users to extract data from diverse sources and then carry out a variety of activities, including data retrieval and classification (supervised, unsupervised, and semi-supervised), data mining, and NLP for automatic classification [4]. Political sentiment analysis has been the subject of multiple research projects in various languages, including English, German, Chinese, and others [5]. The 2009 German federal election was the primary focus as discussed by the Tumasjan et al. [6]. Twitter was used to monitor public opinion and predict the outcome of the election. In the neighborhood of a hundred thousand tweets, they looked for mentions of politicians or political parties. They analyzed the tweets for sentiment using the LIWC2007 tool LIWC is a reliable text analysis program designed to extract emotions, ideas, and personality from their actual text. They reasoned that the more tweets a candidate had, the better their chances of being elected. Online sentiment analysis was performed by [7] to predict the percentage of votes each contender will receive in the 2011 Singapore presidential election. Using Twitter data, [8] demonstrated a real-time sentiment application system for the 2012 US presidential election.

According to Ringsquandl et al. [5], the campaign on Twitter of the presidential candidates from the Republican Party was examined. The authors of the study came to the conclusion that politicians and the topics they address have a stable semantic link by combining the frequency of noun phrases with their PMI value and placing a constraint on aspect extraction. This new notion was presented in their study. Arabic text classification using the

WEKA software was utilized to center on the 2012 Egyptian presidential elections [9]. The scientists concluded that, of the various methods they tested, the Naive Bayesian approach was the most accurate and error-free. The purpose of studying Hindi tweets prior to India's 2016 general state election was to enable informed prediction-making [10]. They extracted 42,235 Hindi tweets from the Twitter Archiver, analyzed them using the SVM, dictionary-based, and Naive Bayes approach, and categorized them as "positive" "negative" or "neutral." Based on the results, the SVM predicted that the BJP had a 78.4 percent chance of winning additional seats in the general election as a result of the strong emotional response they received in their tweets. The Indian National Congress came in second with 26 out of 126 constituencies in the 2016 general election, while the Bharatiya Janata Party (BJP) won 60. This was significantly higher than any other political party [10].

Urdu is categorized as an Indo-Aryan language and is spoken extensively throughout South Asia. In addition to being the official language of Pakistan, it is also one of 22 languages on India's schedule that enjoy de jure official status. Nepal too has its own regional dialect of Urdu. Over 70 million people use it as their first language and over 100 million people use it as a second language, mostly in Pakistan and India [11]. Even though English is the de facto social media language, people all around the globe still prefer to keep their words to themselves. Social media users ought to have the ability to communicate in a variety of languages, including Urdu, in addition to English. A major language in South Asia, Urdu is spoken by a huge number of native speakers. The absence of a conventional writing system in Urdu, however, makes it challenging to extract valuable information from written texts. Urdu sentiment analysis is still in its early stages, and the field of digital linguistics is booming, although studies on the language are few and far between [12][13]. This study lays the groundwork for future research into the feasibility of using Twitter to assess public sentiment on Urdu politics. Examining the sentiments expressed in the tweets collected for this study allowed us to identify their polarity, whether they were good or negative. Our contribution through this research is that we have created the "Urdu Sentiment Corpus (USC)" dataset and insights from Urdu tweets for sentiment analysis and polarity recognition. The dataset comprises tweets that cast a political shadow and present a competitive climate between political parties and the Pakistani government. We have gathered the tweets manually without using any application programming interface (API) or Python library. Approximately 1300 tweets were collected from the Urdu language from thirteen different Pakistani Politicians, with 702 entries being positive (P) and 613 being negative (n), as shown in Figure1.

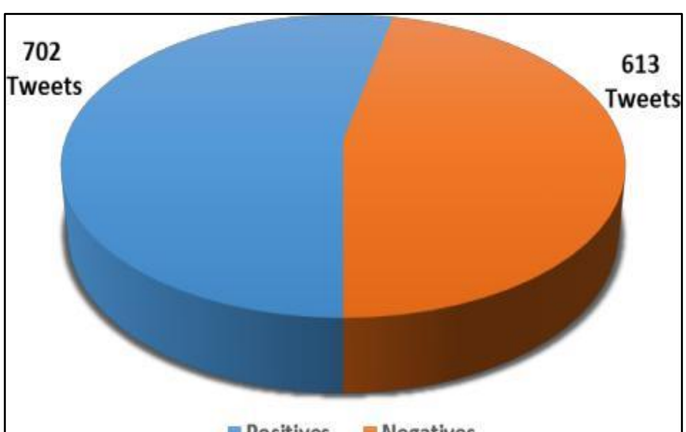

Figure 1: Total Number of Corpus (1300 Tweets)

This research describes visual insights into literary similarities, manifold learning, and other topics. In addition, this research proposed a Parts-of-Speech-wise analysis and applied Machine Learning (ML) algorithms, i.e., Logistic Regression, Decision Tree Classifier, Xgboost, Random Forest, and word2vector, on Urdu texts. We have investigated different approaches for developing Urdu-based sentiment analysis applications. Some of the factors that we considered and accounted for through this research are as follows: 1) Improving political sentiment classification accuracy. 2). Recognizing and classifying emotions in Urdu tweets is a challenging task. 3). Different machine learning and deep learning techniques were investigated for Urdu sentiment analysis.

Methodology:

The Dataset:

This section discusses the dataset's history, data gathering process, problems, and the dataset cleaning procedure, followed by the data labeling mechanism.

Background:

The tweets were gathered on February 17, 2022. Mr. Imran Khan, Chairman of Pakistan reek-e-Insaf (PTI), was the 22nd Prime Minister of the Islamic Republic of Pakistan during the period given. Meanwhile, other opposition groups such as the Pakistan Muslim League (PML-N), Pakistan People's Party (PPP), and Jamiat Ulema-e-Islam (F), among others, were criticizing the government and forming the Pakistan Democratic Organization (PDM), a political movement in Pakistan. It was formed in September.

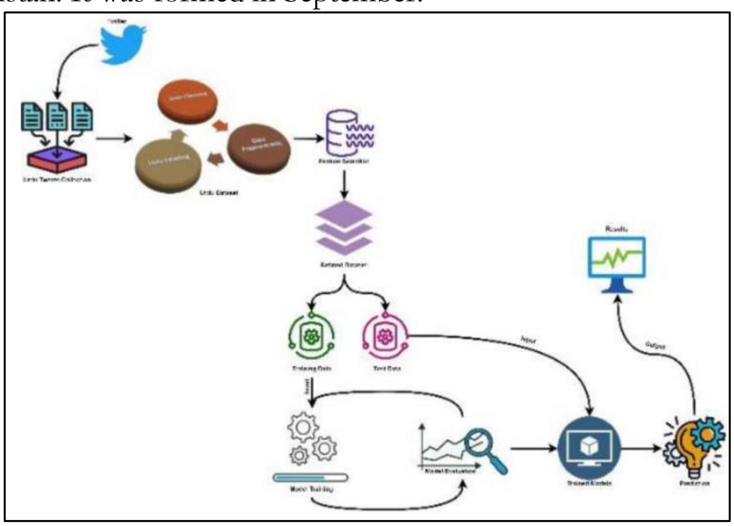

Figure 2: General Methodology

2020 as an opposition alliance against Prime Minister Imran Khan, accusing his administration of manipulating the 2018 Pakistani general election, bad governance, political persecution of opponents, and mismanagement of the economy and foreign affairs. Several dissidents from Khan's Pakistan Tehreek-e-Insaf (PTI) also joined the battle. The campaign successfully ousted Imran Khan in a no-confidence vote on April 10, 2022. As a result, the NNational Assembly of Pakistan chose Mr. Shehbaz Sharif as the 23rd Prime Minister of the Islamic Republic of Pakistan. Imran Khan said the U.S. was behind his ouster because he pursued an independent foreign policy and maintained close relations with China and Russia. His expulsion sparked outrage among his supporters across Pakistan. Imran Khan said that the PTI would begin the anti-government march. Although the data presented in the dataset exhibit a competitive political attire, we cannot say the political showcase dominates the entire dataset.

Twitter Top Trends:

Based on who you follow, your interests, and your location, an algorithm creates trends that are tailored to you. To help you find the most talked-about new topics on Twitter, this algorithm finds popular topics right now, not ones that have been popular for a long or even every day. One of the factors the algorithm takes into account while ranking and zrecognizing trends is the number of Tweets connected to them. Additionally, if trends and hashtags are connected to the same topic, an algorithm will group them [14]. Some trends may include a # indicator before the word or phrase. This is known as a hashtag, and it is used primarily in Tweets to identify them as related to a given topic so that others can follow the conversation in search. In our research, we have occasionally collected top trends in Pakistan and worldwide. Figure 3 shows some people's political interest, which changes with time. These trends are managed manually.

S.No	Date	Top Trends in Pakistan Trends	Top Trends in World Wide Trends/Tweets
1		#Petrol Prices	
2	17-Feb-2022	#Ukraine	
3		#Prime Minister ImranKhan	N/A
1		#FakeNews	
2	21-Feb-2022	#UkraineRussiaCrisis	N/A
1			##UkraineRussiaCrisis
2	22-Feb-2022	N/A	#Putin
1		#Ukraine	
2	24-Feb-2022	#World War 3	N/A
3		#kyiv	
1		#Ukraine 2859.3K tweets	#Moscow 45.8K
2	25-Feb-2022	#worldwar 3 26.1K tweets	#WWIII 572K
1		#PSLFinal 122K	#UkraineRussiaWar 497K
2	28-Feb-2022	#LahoreQalandars 54K	#Japan 460K
1		#pslfinal 12K	＃ファミマ春のおむすび祭 102K
2	01-March-2022	#ShabEMeraj 27K	＃無料引換クーポン 108.3K
1		#sunrisewithadeel	#SOTU 549K
2	02-March-2022	#UkraineRussiaWar 453K	#StateOfTheUnion 68K
1		N/A	Kansas 176K
2	5-April-2022	#surprise 30K	pt 25K
1		N/A	Lina 231K
2	11-April-2022	#ImranKhan 484K	Tadeu 48K
3		#BajwaSurrender 220K	Eliezer 54K

Figure 3: Top Trends on Twitter (Pakistan and Worldwide)

Data Collection and Challenges:

Most previous papers extract data, i.e., tweets, through API and some Python libraries. Our research gathers the tweets manually without using any API or Python library. Approximately 13K+ tweets were collected in the Urdu language from thirteen different Pakistani Politicians. Figure 4 shows the overall view of the dataset, which was collected manually from Twitter. Although it can be seen as overwhelming in size, there are significant challenges in dealing with the data, which are listed below:

Tweets	Sentiment	Tweet Date	Reply	Retweets	Likes
ٹانگ میں گولیاں لگیں لیکن اللہ نے بچا لیا۔ ڈاکٹرز کا کہنا ہے کہ شدید زخمی ہونے کے	p	31-Dec-21	1712	3939	16300
کر بہت زیادہ دکھ ہوا۔ اولاد کا غم شاید دنیا کا سب سے بڑا دکھ ہوتا ہے۔ اللہ تعالیٰ انہیں	p	29	1295	2798	15200
ہوئےہیں۔ غیر آئینی آرڈیننس کی متعلقہ شق فی الفور واپس لی جائے۔ ہمارے حال کو تو	n	23	826	3313	10200
٢٢ کروڑ کے ملک کو مہنگائی، لاقانونیت، نالائقی اور نااہلی کی دلدل میں دھکیل کر، بر	n	20	3040	6091	20900
ب اور مسلم لیگ ن کو بہت بہت مبارک! جن رہنماؤں اور کارکنوں نے محنت کی ، کام کی	p	16	997	2615	10800
الحمدللہ ربّ العالمین شیرو خانیوال کی شاندار فتح مبارک نواز شریف نام ہی کافی ہے	p	16	1775	3890	18400
اللہ کا شکر شیروں کو مبارک	p	5	52	875	3164
اک واری فیر ... شیررررررر الحمدللہ	p	5	2743	4325	19300
ہجان اور ہمارا اور آنے والی نسلوں کامستقبل ہے؟ کیا ہم ملک ایک محفوظ ملک تصور ہو	n	3	2504	4798	17400
۔ اور یہی واقعات وہ انمٹ ثبوت ہیں جن سے نا ثاقب نثار انکار کر سکتے ہیں نا ہی چھپا	n	27-Nov-21	2002	4102	9798
وہ ذلت جو چند روز کے اقتدار کے لیے عمران خان نے کمائی ہے۔ عبرت کا نشان۔ عمران	n	25	1179	2880	6160
، تکلیف اور بے بسی کی داستان ہے۔ میری دعا ہے انہی قطاروں میں عمران خان،	n	25	1238	3258	8721
کی انتہا ہے۔خلق خدا رُل رہی ہے، تڑپ رہی ہے مگر مجال ہے ظالم حکومت کو کوئی فر	n	25	2633	4239	14500
پھر بھی منہ کی کھانے ہیں	n	24	550	1382	4838
اللہ اکبر	p	21	2189	4065	19000
یاسی انتقام اور ظلم سہہ کر ایک بہتر پاکستان کی بنیاد رکھی۔ انشاءاللہ اس حقیقی ت	p	16	2586	4295	10500
ا شکریہ جو مشکل وقت میں ساتھ نبھاتے رہے۔ نواز شریف کے ساتھ آپ سب کو بھی اللہ	p	15	1041	2049	7284
ی اللہ نے ثابت کی۔ نا صرف سرخرو کیا بلکہ انکے خلاف جالی چلنے والوں کو رسوا اور نا	p	15	1901	4907	10900
ے ڈرنا چاہیے جو صبر کرے اور اپنا فیصلہ اللہ پہ چھوڑ دے۔ اس میں ظالموں کے لیے بہت	p	15	4098	5845	18500
اجلاس میں جہاد سمجھ کر ووٹ کریں تو کیا قوم یہ پوچھ سکتی ہے کہ جہاد اچانک ملتو	n	10	1392	3978	12100

Figure 4: Urdu Tweets Dataset

Word Segmentation:

One of the most challenging difficulties to solve is in Urdu (or any Perso-Arabic script language, such as Sindhi) or English words. It commonly happens when a user searches the Urdu corpus for English terms. The reader may read the word/words without difficulty, but the total corpus/words are treated as a single token for the computing task [15].

Duplicate Data:

While collecting the tweets, some of them were duplicated. Later, while training the desired dataset, these duplicated data were removed.

Has-Tags:

As discussed previously, a hashtag is a term or phrase that begins with the # symbol and is used to classify and associate any text/tweet with a trend/topic. Similarly, has-tags can produce a redundant set of tweets if these are added in replies.

Data Preparation and Cleaning:

Following the completion of the tweet collection activity, the procedure of data cleaning begins. While collecting tweets, we manually cleaned the data formally. Still, there were some regular expressions and URLs left. The issues are resolved in two steps. Thus, for a tweet, the first step involves the removal of URL anchors, Twitter handles, and has-tags therein, through the utilization of Regular Expressions, such that doing so will transform into the modified tweet [16]. Let Θ (β, Γ) be the function that takes a *RegEx* pattern B and text and gives you the $\Delta\tau$ in return. In the next step, $\Delta\tau$ is inducted into a hash-set (H) to gather the unique set of tweets. Algorithm1 defines the whole process applied to the entire collection of tweets T, such that $T = T1, T2\ Tn$.

Algorithm 1: Data Cleaning and Redundancy Removal

Result: Set of distinct tweets H

$T \hookleftarrow be\ the\ set\ of\ tweets$

$H \hookleftarrow be\ an\ empty\ hash-set$

$\mu \hookleftarrow be\ the\ RegEx\ pattern\ for\ URL's$

$\alpha \hookleftarrow be\ the\ RegEx\ pattern\ for\ Twitter\ handles$

$\lambda \hookleftarrow be\ the\ RegEx\ pattern\ for\ hash-tags$

$\boldsymbol{for}\ every\ \tau\ in\ T\ do$

| $\Delta\tau \longleftarrow \Theta\left(\lambda, \Theta(\alpha, \Theta(\mu, \tau))\right);$

| $H \longleftarrow H \cup \{\Delta\,\tau\};$

end

We have employed the has-set because, under the hash function, the redundant tweets collide to have the final result in the form of a set of separate tweets. Dictionaries or hash maps can be used as an alternative to hashsets. Lastly, we have tokenized the dataset and extracted the distinct words in the corpus. Then we manually performed the segmentation on the tokens, where space was not in-serted after non-joiners.

Data Labelling:

We created the 'Urdu Sentiment Corpus (USC)' dataset and insights from Urdu tweets for sentiment analysis and polarity recognition as part of this research. The dataset is composed of tweets that cast a political shadow and present a competitive climate between political parties and the Pakistani government. Overall, the dataset contains around 1300+ tokens, with 702 entries being positive and 613 being negative. In addition, this research describes visual insights into literary similarities, manifold learning, and other topics.

Each tweet in the dataset was manually labeled with two sentiments, i.e., Positive (P) and Negative (N). The label was considered positive, where the corpus shows the good aspects of a situation based on thought, feelings, and judgment. On the other hand, the negative sentiment shows the facts, concerns, or experiences based on unpleasant moments, and depressing and emotional idealism.

Sentiment Analysis on Individual Behavior:

They are starting from the assumption that the factors orienting political choices are always heterogeneous and multidimensional. Since Pakistan's inception, the improper functioning of political parties, ineffective leadership, and dismemberment of units have been significant contributors to a paralyzed political system, which has resulted in a slew of problems such as poverty, unemployment, crime, low women's status, child marriage, rape, and gender inequality. These are Pakistan's social difficulties, which, if addressed, may make a living more comfortable and society more productive.

This research study explores an individual's political tweets, allowing us to understand their sentiment regarding a specific theme. For that, we took two famous and well-known Pakistani politicians. Politician 1 and Politician 2. In this research, we haven't mentioned their names—why we are not targeting a specific political party or person. As we have already said, this research is entirely based on manually extracted data from their official Twitter accounts. We explored the political factors influencing politicians' attitudes and behavior through these tweets. According to their tweets, the latest development in Pakistani politics changes their behavior from the first phase. We have collected little tweets. But it shows the limits of a superficial inquiry, as established by looking at a lot of data. So, for future research, it might be useful to use a different technique to realize a comparison and to see if the size of the sample under consideration and the depth of the analytic technique can alter the results and the subsequent consideration. Based on political advancement and leading aspects, Figure 5 shows the tweets of both politicians' actions.

Machine Learning Classification Algorithms:

A dataset will often have significant particulars that may be utilized for generating decisions quickly. No system can make intelligent decisions in the absence of the classification of such datasets. Therefore, classification algorithms make the process easier by removing unnecessary steps, identifying important data categories, and generating valuable models. Each of the relevant papers can be classified into one of three different groups: supervised, unsupervised, or semi-supervised. Texts may be classified using a variety of different approaches, including SVM, KNN, Logistic Regression, and Decision Trees, among others. Word2vec was utilized for data evaluation in the present research; specifically, for the Support Vector Machine (SVM), Decision Trees, XGBoost, Random Forest, and neural network.

SVM Classifier:

A supervised machine learning approach for text categorization that was presented by Salton et al. [17] provides a framework for the SVM. For text classification, SVM has been selected by several researchers. for instance, The SVM has several benefits, including excellent accuracy and a lower risk of overfitting. SVM are great for text classification jobs due to their speed and ability to come up with solutions on the fly. It is also categorized as a technique for categorizing data as linear or non-linear, to put it another way. The SVM method employs a non-linear mapping methodology to transfer linear training data into a higher dimension and look for linear optimum separation hyperplanes [18].

Politician 1		Date	Politician 2		Date
Tweets	Sentiment		Tweets	Sentiment	
رات کی آغوش میں سکردو جگمگاتا ہے اور ہماری بچیاں گلگت بلتستان میں برف پر ہاکی کھیلتی ہیں۔ سکردو میں بین الاقوامی ہوائی اڈے کی تعمیر کے ساتھ یہ علاقہ اب سرما کی کھیلوں کا ایک عالمی مرکز بنے گا، انشاءاللہ۔	p	29-Jan-2022	ایک شیر، تین شکار مہنگائی، نااہلی، اور لوٹ مار خانیوال کا الیکشن جنوبی پنجاب میں پاکستان مسلم لیگ (ن) کی عوامی خدمت کا اعتراف ہے۔ میں پارٹی قائدین اور کارکنان کا شبانہ روز محنت پر شکریہ ادا کرتا ہوں اور شاباش ، دیتا ہوں۔	n	16-Dec-2021
یکجہتی کےاظہار+امریکی حمایت سےمقامی میرجعفروں کے بل پرحکومت بدلنے+ضمانتوں پررہا ڈاکوؤں کےبےحمیت ٹولےکواقتدار دلوانےکیخلاف جذبات سےبھرپوراحتجاج کیلئےبڑی تعدادمیں نکلنےپرمیں اہل پاکستان کامشکور ہوں۔اس سےثابت ہوتاہےکہ ملک وبیرونِ ملک مقیم پاکستانیوں نےاسےپوری شدت سےمستردکردیاہے	n	11-April-2022	معیشت کو درست سمت استوار کرکے مہنگائی کے ستائے عوام کی دادرسی کرنی ہے، روزگار کی فراہمی سے غریبوں کے چولہے پھر سے جلانے ہیں۔ عالمی سطح پر پاکستان کے مقام اور تعلقات کو قومی وقار اور مفادات کی بنیاد پر بحال کرنا ہے۔ قیادت کا امتحان ہی مشکل ترین حالات میں ہوتا ہے ... ۲/	p	8-April-2022
سوالات اٹھارہاہے۔معیشت پہلےہی تباہی کے گڑھے میں اتر چکی ہے۔ ڈاکوؤں اور ان کے آقاؤں کو میں خبردار کرنا چاہتا ہوں کہ یہ غیر جمہوری اور سفاکانہ اقدامات معاشی صورتحال میں مزید ابتری کے موجب بنیں گے اور ملک کو طوائف الملوکی کی جانب دھکیلیں گے۔	n	24-May-2022	آج یومِ عید ایک خاص دن سے بڑھ کر ایک خاص موقع ہے، بے مثال اتحاد و یگانگت کا۔ آج گلہ کرنے کا نہیں گلے ملنے کا دن ہے، قوم کے یک قالب ہونے کا دن ہے اور مل کر مسکرانے کا دن ہے۔ تمام پاکستانیوں اور عالمِ اسلام میں آباد ہر مسلمان کو میری طرف سے عید کی دلی مبارکباد	p	03-May-2022

Figure 5: Political Behavior on Individual Personality

Logistic Regression:

Logistic regression, often known as Linear Regression, is a classification issue solution. Its predictive analysis is based on probability. The sigmoid function, which is one of the more sophisticated algorithms, is the most commonly employed in Logistic Regression [19]. According to this study, Linear Regression fared well on text categorization, ranking first among all classification algorithms with an accuracy of 85%.

Decision Tree Classifier:

The decision tree algorithm is a supervised learning technique used in machine learning. The If-then rule structure on which it is based makes it easy to use. The ability to comprehend results and deal with relationships between features are two of the decision tree's strongest suits. Not only does it function well with numerical data, but also with text and other types of data. Avoiding overfitting in decision trees is possible through a process called "tree pruning" [20].

XGBoost Classification:

The supervised learning algorithm XGBoost (Extreme Gradient Boosting) is widely used. On a predefined Urdu dataset, we have utilized it for regression and classification. It uses a training method that is extremely scalable and avoids overfitting by building short decision trees consecutively [21]. Wrapper classes are provided by XGBoost so that models may be used in the Scikit-learn environment in the same way as classifiers and regressors are. Therefore, the entire Scikit-learn package may be used with XGBoost models.

Random Forest Classifier:

It can classify data and run regressions. A decision tree is built on the training side. From each tree, it predicts the means of regression and classification and assigns classes. The idea behind Random Forest is that features are selected at random during induction and that the classification it generates is based on a random selection of data points/samples from the training data. As far as text classification algorithms go, this study is the third. In order to create different decision trees, it uses random subspace and tree bagging methods, as well as randomly selected data samples from the training data [18]. Here, separate trees are built using random samples, and the classification decision is anticipated to be made by each tree in the random forest. The tagged data is permitted to traverse the trees once the entire forest has been built. Now we get to the proximities; two occurrences are considered closer together by one if they both occur on the same leaf node. The total number of trees in the forest is used to adjust the proximities [22].

Word2vec and TF-IDF:

The algorithm has difficulty processing raw text and can only make sense of numerical data. Therefore, as a first step, we must transform the plain text into numerical form. NLP can help change the raw text into vectors [23]. The feature generated by word embedding features dense and low-dimensional, while the part caused by TF-IDF is sparse and high-dimensional. So, it's also an outstanding example of semantic meaning.

Word2vec embedding, developed by Google in 2013, is one of the significant advances. Word2vec's embedding superiority over TF-IDF is the determining factor in many situations [24][25]. In 2017, the Transformer network was introduced, and after extensive study, the Research, Bidirectional Encoder Representation from Transformer, was published (BERT).

Word2vec:

As a single-hidden-layer neural network model, word2vec is essential. Each word in the sentences or corpus has its neighboring words predicted. The model's hidden layer's learning weights can act as word embeddings. Thus, we'll need to get them. Word2vec, in its simplest form, transforms a word or phrase into a vector in D-dimensional space [26].

The purpose of pre-processing data is to organize it into a collection of word arrays, which enables additional pre-processing. Case folding, tokenization, stemming, stop-word removal, and padding are all pre-processing components. As a processing advantage, case folding converts all words to lowercase, making them uniform and suitable to be represented in the Word2Vec method in the same way they were before. The next step is tokenization [27], in which the text is split into smaller, more manageable chunks called tokens. Tokens might be symbols, phrases, or other discrete items with a clear meaning. The purpose of eliminating filler words is to narrow attention to only the most relevant terms in preparation for further processing. Lemmatization is a method for standardizing text that transforms each word into its stem form. The lemmatizer function from the Urdu Hack library is used in the research. During the preparation phase, padding is the last step. This is a crucial stage since it ensures that all LST training input papers are of the same length. Finally, all the documents in the data collection are under the minimum required word count. This occurs when the document's size exceeds the maximum allowed length when the token" < *pad* >" is inserted [28].

Training of Word2vec:

The author [29] recommends using Word2Vec. By using a neural network architecture, the Word2Vec training process allows the system to grab vector representations of words. The pre-processed Urdu political tweet data is used as the input for training the Word2Vec algorithm. Word2Vec takes as input a set of pre-processed Urdu political tweets and produces vector representations of each word. The creation of a vocabulary from input data is the first step in the process of developing a Word2Vec model.

After that, we learn a vector representation of the word. Both negative and positive sampling are utilized to evaluate the performance of the Word2Vec model embedded in this study. Figure 6 provides a visual breakdown of the various Word2Vec layouts. The original proposal for the hierarchical sigmoid function can be found in Morin et. al. [30], which proposes a binary tree structure of the vocabulary set with a word count (W), with each word represented by a leaf node and each node having a determined probability for its child nodes. The probabilities of the words are assigned using a random walk. Generative Adversarial Networks [31], presented negative sampling, which is predicated on the estimate of noise contrast; that is, a decent model should be able to distinguish between the genuine signal and the fake signal using logistic regression. The dimensions of word2vec are chosen to be in the range of 100 to 300 since this is the range most frequently used in existing research. The features of high-dimensional spaces are also lost if we choose dimensions smaller than 100.

```
Model: "sequential"
_________________________________________________________________
 Layer (type)                Output Shape              Param #
=================================================================
 embedding (Embedding)       (None, 52, 128)           1058560

 conv1d (Conv1D)             (None, 48, 512)           328192

 max_pooling1d (MaxPooling1D  (None, 9, 512)           0
 )

 conv1d_1 (Conv1D)           (None, 5, 256)            655616

 max_pooling1d_1 (MaxPooling  (None, 1, 256)           0
 1D)

 flatten (Flatten)           (None, 256)               0

 dense (Dense)               (None, 32)                8224

 dropout (Dropout)           (None, 32)                0

 dense_1 (Dense)             (None, 1)                 33

=================================================================
Total params: 2,050,625
Trainable params: 992,065
Non-trainable params: 1,058,560
_________________________________________________________________
```

Figure 6: Word2vec Model Layers

Word2vec Using LSTM:

The LSTM is trained to develop a model for the sentiment classification model. On the other hand, testing aims to evaluate how well the classification model functioned after training [32]. By "word embedding," we mean the technique of modeling a document's vocabulary with a set of vectors. Word embeddings may be learned more easily with the help of the embedding layer. Glove and Word2Vec are two methods used to study word embeddings. The Word2Vec word embedding method [33].

The LSTM technique is a subset of the RNN series designed for processing models containing long input sequences. The RNN method builds on the foundation of regular feed-forward neural networks. However, RNN models have problems with explosions and gradient vanishing. The LSTM model was developed to overcome these challenges. This capacity arises from the LSTM model cell memory, which helps maintain the system's current state [34].

In addition to the cell memory, the LSTM model has three gates: the input gate, the forget gate, and the output gate. The gate input specifies which bits of information are modified and delivered into the memory cell. The forget gate's function is to determine whether or not the data being input and output is safe to proceed on. Forgetting happens if the forget gate's output is near zero; retaining occurs if it's close to one. As a result of this forget-fate operation, LSTM can deal with the exploding problem and the vanishing gradient problem. The state of the cell is unaffected by the gate output; however, the date differentiates between the cell state and the valid information.

```
Model: "sequential_1"
_________________________________________________________________
 Layer (type)                Output Shape              Param #
=================================================================
 embedding_1 (Embedding)     (None, 52, 128)           1058560

 bidirectional (Bidirectiona  (None, 512)              788480
 l)

 dense_2 (Dense)             (None, 128)               65664

 dropout_1 (Dropout)         (None, 128)               0

 dense_3 (Dense)             (None, 64)                8256

 dropout_2 (Dropout)         (None, 64)                0

 dense_4 (Dense)             (None, 1)                 65

=================================================================
Total params: 1,921,025
Trainable params: 862,465
Non-trainable params: 1,058,560
_________________________________________________________________
```

Figure 7: Word2vec Model using LSTM

Pooling Layer:

The pooling layers compress the spatial input, which in turn decreases the number of network parameters and accelerates computation while also controlling overfitting. It is common practice for the pooling layer to use both maximum and average pooling. Their names suggest that maximum and average values are used in the respective methods.

Fully Connected Layer:

Given that the fully connected layer requires a vector as an input, a transformation from the pooling layer's multidimensional array output is needed.

Evaluation Metrics:

An assessment metric quantifies a predictive model's effectiveness. This process typically entails training a model on a data set, using the model to make predictions on a holdout data set that was not utilized during training, and then comparing the predicted values to the expected values in the holdout dataset [35]. In this research, we have different evaluation metrics as mentioned below;

Precision:

Precision is a measure of performance that is used in pattern recognition, information retrieval, and classification. (machine learning) to evaluate information that has been obtained from a dataset. In our case, we have used Urdu Political tweets. The purpose of the performance is to determine how many positive observations were predicted accurately as a proportion of all positive occurrences [36].

Recall:

The recall is the ratio of correctly predicted positive observations to all observations in the actual class [37].

F1-Score:

Precision and Recall are weighted to produce the F1 Score. True positives and false negatives are both factored into this score. F1 is typically more useful than accuracy, especially if your class distribution is not uniform, despite the fact that it is not as intuitively appealing. When the costs of false positives and false negatives are roughly equal, accuracy works best. It is more prudent to examine both Precision and Recall if the cost of false positives and false negatives is substantial [38].

Accuracy:

The proportion of the total set of data that was accurately predicted is how prediction accuracy is measured. More precisely, it is the ratio of the total number of positive and negative results divided by the total number of positive and negative results, which includes both true and false positives and negatives [39].

Results and Discussion:

Several machine learning classifiers, selected for training text data classification, have been covered in this section. In this research, SVC classifiers are the most widely used supervised learning classifiers [18], Decision Tree [20], XG-Boost [21], and Random Forest. For the neural network, we have used Word2vec. Table 1 shows the Accuracy and other evaluation metrics results based on Precision, Recall, and F1-score, obtained through Urdu text classification from the corpus (Tweets).

It is worthwhile to mention that the Logistic regression classifier attained the highest accuracy (overall) among the rest of the tested algorithms. The neural network i.e. word2vec has the lowest accuracy of all classifiers. Most of the models predicted well during the test on Urdu Tweets text data to analyze the efficiency and accuracy of each classification model.

Sentiment analysis on Twitter can be viewed as a classification problem in which positive and negative tweets must be sorted into separate categories. As a result, we have used a variety of classifiers for the task, and we have checked their performance by measuring their Accuracy of prediction. Precision, Recall, and F1 Score are also calculated to measure performance. Parameters are derived from the confusion matrix.

Table 1: Urdu Text Classification Result

Models	Sentiment	Precision	Recall	F1-Score	Support	Accuracy
SVM Classifier	Negative	0.85	0.80	0.82	186	
on TF-IDF Vectorizer	Positive	0.83	0.87	0.85	209	0.84
Logistic Regression Classifier	Negative	0.85	0.82	0.83	182	
on TF-IDF Vectorizer	Positive	0.85	0.87	0.86	213	0.85
Decision Tree Classifier	Negative	0.58	0.82	0.68	125	
on TF-IDF Vectorizer	Positive	0.89	0.73	0.80	270	0.75
Boost Classifier	Negative	0.52	0.82	0.63	111	
on TF-IDF Vectorizer	Positive	0.91	0.70	0.79	284	0.73
Random Forest Classifier on TF-IDF Vectorizer	Negative	0.67	0.80	0.73	147	
	Positive	0.87	0.77	0.81	248	0.78
	Negative	0.61	0.56	0.58	140	
word2vec	Positive	0.54	0.60	0.57	123	0.58
word2vec	Negative	0.12	0.56	0.19	27	
using LSTM Layer	Positive	0.91	0.53	0.67	236	0.52

Table 2: Confusion Matrix

Predicted Class	Positive Prediction	Negative Prediction
True Positive	TP (True Positive)	FN (False Negative)
True Negative	FP (False Positive)	TN (True Negative)

Although there is no scarcity of classifiers (or classification models), we have focused on the most suitable for text analysis. Accuracy detection is a key tool for evaluating the classifier's performance. For this purpose, accuracy alone is not enough for evaluating classifier performance; other metrics such as precision, recall, and F1-score measure are required. Table 2 shows an example of a confusion matrix, which compares the true results with those predicted by a classifier. The F1-score metric represents a harmonic mean of the Precision and Recall measures. Determining a decision based on Precision and Recall alone can be challenging, especially when one of the scores becomes extremely high or extremely low. Then we can decide upon the validity of the result by considering the F1-score measure.

Looking at Table 1, Word2vec shows worse results compared to other classifier models. The reason behind this is that Word2Vec always struggles with words that are not in its dictionary. Out-of-vocabulary (OOV) words are given a potentially unsatisfactory random vector representation. Local lexical knowledge is used for this purpose. Words' semantic representations are based entirely on their surrounding contexts, which isn't always ideal. It is not possible to share parameters used in the training of new languages. We have to start from scratch if we want to train word2vec in a new language like Urdu. It is incapable of dealing with ambiguity. If a word has multiple meanings, and there are many of these terms in the real world, embedding will reflect the average of these senses in vector space. More data is needed for the network to converge (particularly when employing skip-gram). The dataset size of 1300 tweets is not large enough to train a word2vec neural network.

Limitation:

In the field of NLP, sentiment analysis is among the most successful and widespread uses. There are still many issues with it, despite all the hype it has garnered since it was founded. One of the most difficult aspects of creating datasets is ignoring potentially biased perspectives of individuals due to the cultural backgrounds of those individuals or their own personal opinions. In addition to this, sentiment analysis of tweets based on biased views can choose just those results and data that support their major arguments, which, needless to say, may have an effect on our research. We have paid close attention to the explanation of the primary problem we tried to solve, and we make sure that we have collected accurate data.

Systems for sentiment analysis trained on data from political tweets often perform far worse when applied to data from other domains. Politicians do not always communicate sentiment in the same way that a critic does on social media, which is different from how a social media poster does. Therefore, it is not uncommon for machine learning systems trained on political tweet data to fail miserably when faced with the task of predicting sentiment in a more general domain. Contextual understanding is another facet of NLP. This area is now experiencing a surge in research activity. Memory networks are a type of model that may derive answers to queries from context. Similar to how a person's brain processes separate words, attention networks can process different sections of a phrase independently. The larger context of a statement can be understood with the help of RNN, like bidirectional LSTM. Our research has led to the beginning of incorporating these into sentiment analysis.

By adjusting the parameter values based on their learning rate, epochs, batch size, etc., we have previously tried numerous fine-tuning strategies, as seen by the results. The results were still inappropriate. We can try to train more models for this specific domain, ideally using a more sophisticated technique (like deep neural networks). In terms of both time and money, this is the worst possible option. Yet, if we are able to accomplish this with our available resources, we may end up with a valuable and exclusive piece of intellectual property.

Conclusion and Future Work:

This study employed a variety of machine learning classification algorithms—including SVM, Logistic Regression, Decision Tree, XGBoost, and Random Forest—using the TF-IDF measure—to classify Urdu text (Tweets). In order to extract the idea of relatedness across words, the neural network makes use of Word2vec. We used data cleaning, feature extraction, and feature selection as NLP pre-processing approaches before we used Urdu-Text. Nevertheless, when compared to other classifiers tested on a comparable dataset for Urdu tweets, the results demonstrated that Logistic Regression and SVM achieved 85% and 84% accuracy, respectively.

The results demonstrate that the TF-IDF model performs the Word2Vec model because of the imbalanced data points in each emotion class and a large number of classes with an insufficient number of data points. Contrasted with the abundance of other emotions, the number of surprise feelings constitutes a small fraction of the total data set. The results of the research may be impacted by data limitations. We will require a sizable dataset in the future that can handle these constraints.

References:

[1] G. L. Yovellia Londo, D. H. Kartawijaya, H. T. Ivariyani, P. W. P. Yohanes Sigit, A. P. Muhammad Rafi, and D. Ariyandi, "A Study of Text Classification for Indonesian News Article," Proceeding - 2019 Int. Conf. Artif. Intell. Inf. Technol. ICAIIT 2019, pp. 205–208, Mar. 2019, doi: 10.1109/ICAIIT.2019.8834611.

[2] Y. Zheng, "An exploration on text classification with classical machine learning algorithm," Proc. - 2019 Int. Conf. Mach. Learn. Big Data Bus. Intell. MLBDBI 2019, pp. 81–85, Nov. 2019, doi: 10.1109/MLBDBI48998.2019.00023.

[3] S. K. Dwivedi and C. Arya, "Automatic text classification in information retrieval: A Survey," ACM Int. Conf. Proceeding Ser., vol. 04-05-March-2016, Mar. 2016, doi: 10.1145/2905055.2905191.

[4] V. Korde, "Text Classification and Classifiers:A Survey," Int. J. Artif. Intell. Appl., vol. 3, no. 2, pp. 85–99, Mar. 2012, doi: 10.5121/IJAIA.2012.3208.

[5] "Analyzing political sentiment on twitter", [Online]. Available: https://cdn.aaai.org/ocs/5702/5702-24478-1-PB.pdf

[6] A. Tumasjan, T. O. Sprenger, P. G. Sandner, and I. M. Welpe, "Predicting Elections with Twitter: What 140 Characters Reveal about Political Sentiment," Proc. Int. AAAI Conf. Web Soc. Media, vol. 4, no. 1, pp. 178–185, May 2010, doi: 10.1609/ICWSM.V4I1.14009.

[7] M. Choy, M. L. F. Cheong, M. N. Laik, and K. P. Shung, "A sentiment analysis of Singapore Presidential Election 2011 using Twitter data with census correction," Aug. 2011, Accessed: Jun. 06, 2024. [Online]. Available: https://arxiv.org/abs/1108.5520v1

[8] S. N. Hao Wang, Dogan Can, Abe Kazemzadeh, François Bar, "A System for Real-time Twitter Sentiment Analysis of 2012 U.S. Presidential Election Cycle", [Online]. Available: https://aclanthology.org/P12-3020/

[9] T. Elghazaly, A. Mahmoud, and H. A. Hefny, "Political sentiment analysis using twitter data," ACM Int. Conf. Proceeding Ser., vol. 22-23-March-2016, Mar. 2016, doi: 10.1145/2896387.2896396.

[10] P. Sharma and T. S. Moh, "Prediction of Indian election using sentiment analysis on Hindi Twitter," Proc. - 2016 IEEE Int. Conf. Big Data, Big Data 2016, pp. 1966–1971, 2016, doi: 10.1109/BIGDATA.2016.7840818.

[11] A. Daud, W. Khan, and D. Che, "Urdu language processing: a survey," Artif. Intell. Rev., vol. 47, no. 3, pp. 279–311, Mar. 2017, doi: 10.1007/S10462-016-9482-X/METRICS.

[12] D. P. Rose and D. M. Greeley, "Education in fragile states: Capturing lessons and identifying good practice", [Online]. Available: https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=d0d3c719eced2e039700dc583919188d60773d34

[13] S. W. Kazmi, "Role of education in globalization: A case for pakistan", [Online]. Available: http://www.developyst.jellyfish.com.pk/files/article/39/Role of Education in Globalization.pdf

[14] R. Lu and Q. Yang, "Trend Analysis of News Topics on Twitter," Int. J. Mach. Learn. Comput., pp. 327–332, 2012, doi: 10.7763/IJMLC.2012.V2.139.

[15] J. R. Saffran, E. L. Newport, and R. N. Aslin, "Word Segmentation: The Role of Distributional Cues," J. Mem. Lang., vol. 35, no. 4, pp. 606–621, Aug. 1996, doi: 10.1006/JMLA.1996.0032.

[16] S. Zhang, C. Zhang, and Q. Yang, "Data preparation for data mining," Appl. Artif.

Intell., vol. 17, no. 5–6, pp. 375–381, May 2003, doi: 10.1080/713827180.
[17] G. D. Salton, R. J. Ross, and J. D. Kelleher, "Idiom Token Classification using Sentential Distributed Semantics," 54th Annu. Meet. Assoc. Comput. Linguist. ACL 2016 - Long Pap., vol. 1, pp. 194–204, 2016, doi: 10.18653/V1/P16-1019.
[18] Y. Al Amrani, M. Lazaar, and K. E. El Kadirp, "Random Forest and Support Vector Machine based Hybrid Approach to Sentiment Analysis," Procedia Comput. Sci., vol. 127, pp. 511–520, Jan. 2018, doi: 10.1016/J.PROCS.2018.01.150.
[19] "Kleinbaum, D.G., Dietz, K., Gail, M., et al. (2002) Logistic Regression. Springer-Verlag, New York, 43-53. - References - Scientific Research Publishing." Accessed: Jun. 06, 2024. [Online]. Available: https://www.scirp.org/reference/referencespapers?referenceid=2959305
[20] S. R. Safavian and D. Landgrebe, "A Survey of Decision Tree Classifier Methodology," IEEE Trans. Syst. Man Cybern., vol. 21, no. 3, pp. 660–674, 1991, doi: 10.1109/21.97458.
[21] E. Podasca, "Predicting the Movement Direction of OMXS30 Stock Index Using XGBoost and Sentiment Analysis", Accessed: Jun. 06, 2024. [Online]. Available: www.bth.se
[22] J. Ali, R. Khan, N. Ahmad, and I. Maqsood, "Random Forests and Decision Trees," 2012, Accessed: Jun. 06, 2024. [Online]. Available: www.IJCSI.org
[23] W. Zhu, W. Zhang, G. Z. Li, C. He, and L. Zhang, "A study of damp-heat syndrome classification using Word2vec and TF-IDF," Proc. - 2016 IEEE Int. Conf. Bioinforma. Biomed. BIBM 2016, pp. 1415–1420, Jan. 2017, doi: 10.1109/BIBM.2016.7822730.
[24] O. B. Deho, W. A. Agangiba, F. L. Aryeh, and J. A. Ansah, "Sentiment analysis with word embedding," IEEE Int. Conf. Adapt. Sci. Technol. ICAST, vol. 2018-August, Oct. 2018, doi: 10.1109/ICASTECH.2018.8506717.
[25] C. Esteves, C. Allen-Blanchette, X. Zhou, and K. Daniilidis, "Polar Transformer Networks," 6th Int. Conf. Learn. Represent. ICLR 2018 - Conf. Track Proc., Sep. 2017, Accessed: Jun. 06, 2024. [Online]. Available: https://arxiv.org/abs/1709.01889v3
[26] K. W. CHURCH, "Word2vec," Nat. Lang. Eng., [Online]. Available: https://www.cambridge.org/core/journals/natural-language-engineering/article/word2vec/B84AE4446BD47F48847B4904F0B36E0B
[27] G. Grefenstette, "Tokenization," pp. 117–133, 1999, doi: 10.1007/978-94-015-9273-4_9.
[28] X. Rong, "word2vec Parameter Learning Explained," Nov. 2014, Accessed: Jun. 06, 2024. [Online]. Available: https://arxiv.org/abs/1411.2738v4
[29] Y. Goldberg et al., "word2vec Explained: deriving Mikolov et al.'s negative-sampling word-embedding method," Feb. 2014, Accessed: Jun. 06, 2024. [Online]. Available: https://arxiv.org/abs/1402.3722v1
[30] F. Morin and Y. Bengio, "Hierarchical Probabilistic Neural Network Language Model." PMLR, pp. 246–252, Jan. 06, 2005. Accessed: Jun. 06, 2024. [Online]. Available: https://proceedings.mlr.press/r5/morin05a.html
[31] A. Creswell, T. White, V. Dumoulin, K. Arulkumaran, B. Sengupta, and A. A. Bharath, "Generative Adversarial Networks: An Overview," IEEE Signal Process. Mag., vol. 35, no. 1, pp. 53–65, Jan. 2018, doi: 10.1109/MSP.2017.2765202.
[32] C. Steinruecken, "Compressing sets and multisets of sequences," IEEE Trans. Inf. Theory, vol. 61, no. 3, pp. 1485–1490, Mar. 2015, doi: 10.1109/TIT.2015.2392093.
[33] P. F. Muhammad, R. Kusumaningrum, and A. Wibowo, "Sentiment Analysis Using Word2vec And Long Short-Term Memory (LSTM) For Indonesian Hotel Reviews," Procedia Comput. Sci., vol. 179, pp. 728–735, Jan. 2021, doi: 10.1016/J.PROCS.2021.01.061.

[34] R. Huang et al., "Well performance prediction based on Long Short-Term Memory (LSTM) neural network," J. Pet. Sci. Eng., vol. 208, p. 109686, Jan. 2022, doi: 10.1016/J.PETROL.2021.109686.

[35] M. Hossin and Sulaiman, "A REVIEW ON EVALUATION METRICS FOR DATA CLASSIFICATION EVALUATIONS," Int. J. Data Min. Knowl. Manag. Process, vol. 5, no. 2, 2015, doi: 10.5121/ijdkp.2015.5201.

[36] M. Kane, "The precision of measurements," Appl. Meas. Educ., vol. 9, no. 4, pp. 355–379, 1996.

[37] F. Buckland, M., and Gey, ""The relationship between recall and precision.,"" J. Am. Soc. Inf. Sci., vol. 45, no. 1, pp. 12–19, 1994.

[38] O. M. Sulea, M. Zampieri, S. Malmasi, M. Vela, L. P. Dinu, and J. Van Genabith, "Exploring the Use of Text Classification in the Legal Domain," CEUR Workshop Proc., vol. 2143, Oct. 2017, Accessed: Jun. 06, 2024. [Online]. Available: https://arxiv.org/abs/1710.09306v1

[39] F. Freese, "Testing Accuracy," For. Sci., vol. 6, no. 2, pp. 139–145, Jun. 1960, doi: 10.1093/FORESTSCIENCE/6.2.139.