JCBI
Journal of Computing & Biomedical Informatics Volume 08 Issue 01
ISSN: 2710-1606 2024

Research Article
https://doi.org/10.56979/801/2024

Urdu Language Text Summarization using Machine Learning

Imtiaz Khan[1], Muhammad Imran Khan Khalil[1*], Asif Nawaz[2], Izaz Ahmad Khan[3], Lubna Zafar[4], and Sheeraz Ahmed[5]

[1]University of Engineering & Technology, Peshawar, 25000, Pakistan.
[2]Higher Colleges of Technology, United Arab Emirates.
[3]Bacha Khan University, Charsadda, 24420, Pakistan.
[4]University of Poonch Rawalakot, 12350, Azad Kashmir, Pakistan.
[5]Iqra National University, Peshawar, 25000, Pakistan.
[*]Corresponding Author: Muhammad Imran Khan Khalil. Email: imrankhalil@uetpeshawar.edu.pk

Received: July 21, 2024 **Accepted:** September 29, 2024

Abstract: Text summarization involves creating a concise version of a text while preserving its essential details and core message. This technique allows for quick understanding of lengthy articles, documents, or books, saving time and enhancing comprehension. There are two primary approaches to summarization: extractive and abstractive. Extractive summarization involves selecting key sentences or phrases from the text, while abstractive summarization generates new sentences that convey the original meaning. In the context of Urdu news articles, summarization is particularly valuable, as it enables readers with limited time or attention to grasp the main points quickly. This study explores various methods for summarizing Urdu news articles, evaluating both extractive and abstractive approaches using various datasets. To evaluate the proposed model, we use ROUGE metric which shows the significant improvement and efficiency compared with existing models. The study also highlights challenges and future directions in this field, including the complexity of Urdu sentence structures, addressing biases in content of news, and incorporating the latest development natural language processing and deep learning.

Keywords: Urdu; Machine Learning; Abstractive; Extractive; Text Summarization; Deep Learning.

1. Introduction

The fast development of digital content and its processing has created a significant need for efficient content summarization methods. The key aspect, text summarization of NLP, aims to condense lengthy texts while preserving essential information and main ideas [1]. While extensive research has been conducted on summarization techniques for languages like English, Chinese, and French, there has been limited focus on Urdu, despite it being the language with the sixth highest number of speakers and the Pakistan's formally recognized language [2].

Urdu is widely used in government, education, and media, making the summarization of Urdu news articles particularly important. This is essential for Urdu speakers worldwide, enabling them to quickly and easily grasp the news [3].

News plays a vital role in keeping people informed about global events, major social issues, and developments across various sectors. However, with the overwhelming volume of news available online, individuals often struggle to stay updated [4]. The growing demand of technologies that can effectively con-dense news content while retaining key information. Text summarization addresses this challenge by extracting crucial details without altering the original meaning [5].

Summarizing Urdu news articles can have a positive impact across various domains, such as education, politics, and business. In education, text summarization can help students quickly comprehend complex news articles, keeping them informed about current events [1],[6]. In politics, summarizing news can enhance public awareness, promoting transparency and accountability in government. Businesses can al-so benefit, as text summarization enables scholars to in line with market trends.

Recent researches in Urdu news and it summarizing is a contemporary problem and it supports those people who speaks and understand Urdu language. The main objective of this research work is to define strategies and policies to summarize Urdu news and its related stories. The workflow for this research work is illustrated in Figure 1.

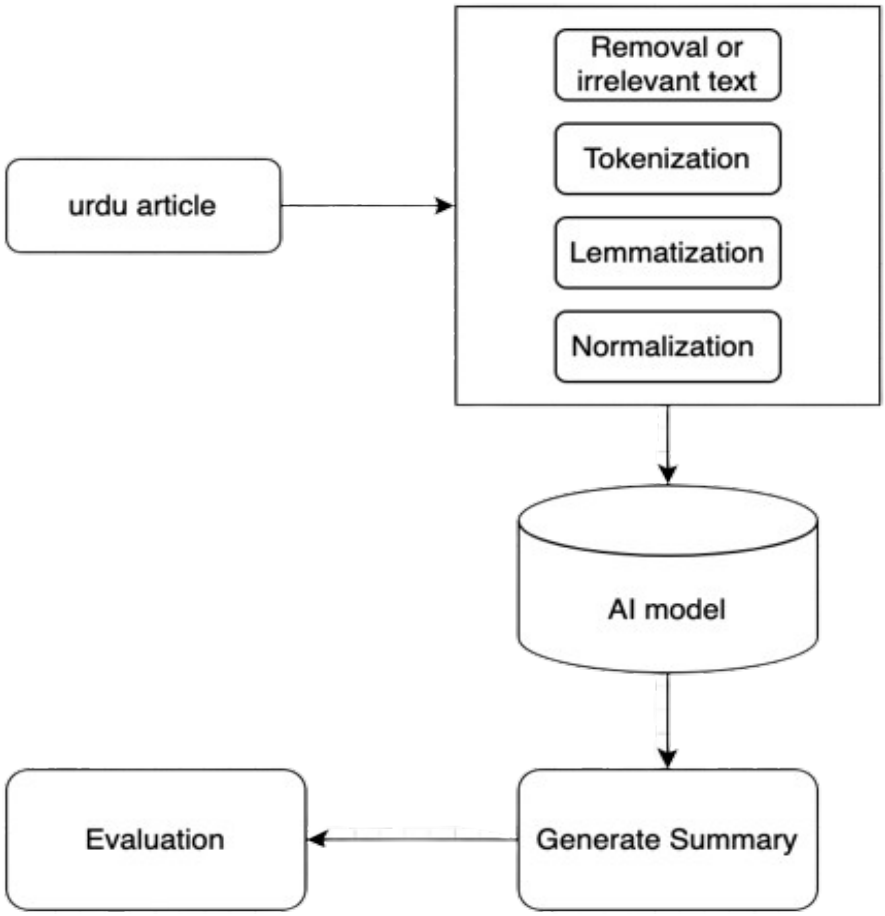

Figure 1. Workflow of the research work

2. Literature Review

In natural language processing the text summarization is a difficult job. In this regard, various techniques have been suggested to handle this issue. In recent years, there has been growing interest in summarizing Urdu news articles, driven by the widespread use of Urdu in Pakistan and other countries. This paper reviews key research efforts aimed at summarizing Urdu news.

One of the pioneering works in this area was presented by Khan and Fraz [5], who proposed a rule-based method for selecting important sentences from Urdu text to create summaries. Their approach involves identifying and extracting significant sentences from Urdu documents using linguistic rules. This method enhances the efficiency of summarization by simplifying the process of extracting and understanding information in Urdu. [6] Introduced a different approach which graph based and utilizes both semantic and lexical parameters for sentence similarity graph construction. The most important sentences are identified based on this graph, achieving a summarization ratio of 30%, outperforming previous methods.

[7] Focused upon Urdu language text summarization extractive models. The authors evaluated various techniques, aiming to improve summarization for Urdu content. Their work offers insights into adapting summarization methods for specific languages, leading to more effective information. [8] Proposed the model based on DL for the summarization of Urdu news. To identify the relevant phrases in Urdu news they used CNN and RNN machine learning models. They evaluate their proposed models using a data set of more than hundred articles and achieved thirty percent summarization ratio. [9] Designed a model based on recurrent neural network hybrid to summarize Urdu newspapers and use

statistical techniques to evaluate the model. For the evaluate of model they used a data set of twenty five percent urdu language ar-ticles and shows a significant improvement. [10] designed a contemporary technique using unsupervised machine learning model and NLP. They attain thirty five percent ratio of summarization and achieve out-standing results as compared with other models.

[11] Proposed a new technique which based on neural networks for Urdu news summarization. To identify key terminologies in Urdu language they used domain specific linguistic method. They achieved a significant improvement in results compared with other models. [12] Designed a significantly different technique for Urdu language news summarization using CNN and evolutionary efficient algorithms. The result shows the tremendous improvement in results considering ROUGE scores and other statistical evaluation methods.

[13] Proposed a novel technique which is based on graph theory for summarizing Urdu news. Their method utilizes both lexical and semantic features to identify important sentences and shows the significant good results compare with other algorithms. [14], a prominent figure in text summarization research specializing in Human Language Technologies (HLT), offers a comprehensive literature review that ex-amines advancements in text summarization. The review provides a thorough analysis of the technological developments and algorithms used in creating summarization systems, with a focus on HLT.

[15] Introduced a novel approach to extract meaningful information in Urdu language for summarization. The algorithm used in this research work shows the significant improved results compared with other techniques. This algorithm guarantees the conservation of sentence structure whereas minimize the test length. The main objective of this research work is to maximize the accuracy of summarization of Urdu language.

[16] Conducted an in-depth investigation into automatic text summarization, considering machine learning models. This research work highlights the development is techniques of summarization for different languages focusing on the integration of machine learning and Long Short-Term Memory (LSTM) networks [17]. The article provides an overview of recent advancements in both extractive and abstractive summarization techniques, highlighting the role of LSTM networks in improving summarization effectiveness. The literature review offers valuable insights into the field, guiding future research.

[18] Presents a unique approach to extractive summarization using topic modelling techniques. The study explores the strategic implementation of topic-specific modelling, which group's text based on subject matter to aid in summarization. This method results in precise and clear extractive summaries, demonstrating the integration of topic modelling into the summarization process. Mancha [19] investigates artificial text summarization, emphasizing the extraction of information from large volumes of written content on the internet. The research utilizes advanced Natural Language Processing (NLP) techniques and deep learning methods to evaluate the different models.

3. Experimental work and its Evaluations

The problem setting of this research work is summarized in the following steps

3.1. Pre-processing and Data Collection

To improve the summarization output, preprocessing of the collected data involves removing stop words, stemming, and normalizing the text. The preprocessing steps include:

1. **Normalization:** This involves standardizing the text by correcting spelling, simplifying complex structures, and ensuring uniformity throughout the text.
2. **Removing Punctuation:** This step entails eliminating punctuation marks such as commas, periods, and question marks from the Urdu text.
3. **Removing Accents:** In Urdu, accents such as 'zabar,' 'zair,' and 'pesh' indicate consonant pronunciation. Removing these accents enhances text clarity.
4. **Replacing URLs:** Web addresses are replaced with placeholders like "URL" or symbols such as "link" to simplify the text.
5. **Replacing Numbers:** Numerical digits are substituted with corresponding words or symbols representing numbers in the text.
6. **Replacing Currency and Symbols:** Currency symbols (e.g., "$" for dollars or "Rp" for Pakistani rupees) are replaced with words or blanks in the text.
7. **Removing English Words:** English words or sentences are identified and appropriately modified or removed from the Urdu text.

8. **Normalizing White Spaces:** This step ensures consistent use of tabs, spaces, and line breaks by removing excess spacing and consolidating multiple spaces into one.
9. **Text Tokenization:** The text is split into individual words or tokens for further analysis.
10. **Lemmatization/Stemming:** Words are either reduced to their base form (lemmatization) or truncated to their root form (stemming) for better analysis.

3.2. Data Collection

A large dataset of Urdu news articles is compiled from various publications and online sources, covering a range of genres and topics, including politics, entertainment, and sports. The dataset consists of 10,000 articles categorized into four main groups: Business and Economics (3,500 articles), Entertainment (1,500 articles), Science and Technology (2,700 articles), and Sports (2,300 articles). The dataset's structure is visualized to provide a clear overview of the categories. The Pre-Processing steps are below in the figure.

Tools Pipeline	Query (Urdu)
Query	لاہور کی قدیم عمارتیں
Tokenization	لاہور، کی، قدیم، عمارتیں
Stop Word Removal	لاہور، قدیم، عمارتیں
Lemmatization	لاہور، قدیم، عمارت
Query Expansion	لاہور، قدیم، عمارت/عمارتیں/عمارات/عمارتوں

Figure 2. Pre-processing steps

S. No	Total Articles	Articles Domain
1	10000	Urdu Articles
2	3500	Business and Economics
3	1500	Entertainment
4	2700	Science and Technology
5	2,300	Sports

Figure 3. Collected dataset

4. Sentence Extraction

Sentence extraction (SE) involves selecting the most critical sentences from a longer text, such as an essay or report. This approach effectively reduces the content while preserving key information, allowing researchers to quickly identify the relevant details without reading the entire document [20]. To extract the most important phrases from the pre-processed text and capture the core ideas, various methods can be employed. These include machine learning models, graph-based algorithms, and term frequency-inverse document frequency (TF-IDF) techniques [21].

SE is particularly beneficial for Urdu language learning for several reasons. It enables the condensation of lengthy texts, allowing readers to grasp the main ideas quickly without investing significant time in reading the full content. Additionally, it simplifies the evaluation of different types of articles, helping experts efficiently locate specific information within large datasets [22]. This method relies on language-aware algorithms and automated tools to identify sentences with the most crucial information. Several methodologies are applied, including contextual evaluation, keyword analysis, and trend identification. By refining the extraction process, these techniques ensure that the selected terms align with the research objectives [23].

5. Sentence Ranking

Sentence placement involves systematically analyzing and organizing sentences within a text based on their significance and relevance to the study's objectives. The goal is to arrange sentences in a logical sequence that reflects their importance [24]. This technique is useful for reducing and evaluating information, allowing us to quickly identify key phrases in a lengthy text. To assess the effectiveness of the selected sentences, factors such as sentence length and position within the text are considered. Techniques like Text Rank, centroid-based summarization, and other clustering algorithms are employed for this purpose [25]. Sentence ranking has various applications in Urdu language studies. By prioritizing essential terms, researchers can extract critical information from a document, making the reading process more efficient. Sentence evaluation algorithms help identify the most important lines.

To enable the creation of brief summaries of Urdu papers automated sentence ranking is useful for the summarization of content [26]. Additionally, it plays a crucial role in developing search engines and text analysis tools that allow users to quickly access vital information. By prioritizing and categorizing terms, developers and experts working with Urdu can enhance the clarity and accessibility of their content [27].

6. Generation of Summary

To make summaries meaningful from long Urdu articles summary generation is a useful tool [28]. The process involves selecting the most relevant sentences or phrases based on specific summarization criteria or a predetermined length. The resulting summary should be brief, easy to understand, and accurately convey the main ideas of the original text [29]. Summarization is an effective tool for managing and interpreting vast amounts of information in Urdu language studies [30]. It leverages natural language processing techniques to analyze the text, identify key sentences, and organize them logically. This ensures that the summary captures the essence of the full text while remaining coherent and faithful to the original content.

In essence, generating summaries for Urdu language content facilitates academic work, information sharing, and reading comprehension. It simplifies complex Urdu texts into more accessible formats, enhancing transfer of knowledge and allowing the broader public to take advantage from the detail understanding of research.

7. Selection of Features

The selected features from the previous step are used to train the models. The data is divided into training and testing sets, with 70% allocated for training and 30% reserved for testing. This is accomplished using the "*sklearn.model_selection.train_test_split*" package from scikit-learn, which randomly splits the data into two groups: "train" and "test." [31] The training set, comprising 70% of the data, is used to build the models, while the remaining 30% is utilized to evaluate their performance. To ensure a balanced distribution of accurate and fake articles, the data is shuffled during the splitting process.

8. Research Model to Establish Abstractive Summary

The research model in this research work are twofold: the encoder and the decoder. The encoder processes the input sequence and transforms it into a fixed-length context vector, which encapsulates the essential information of the input [32]. The decoder then uses this context vector to generate the output sequence. Both the encoder and decoder incorporate an embedding layer, which converts raw tokens into meaningful representations. Long Short-Term Memory (LSTM) layers are used in both the encoder and decoder, effectively managing sequential data by capturing the relationships between input and output patterns. The LSTM layer generates a sequence of output lines for the decoder, which are then passed through a dense layer with SoftMax activation [33]. This layer produces a probability distribution over the vocabulary, allowing the model to generate sequences of varying lengths. This architecture is particularly well-suited for tasks such as language translation and text generation.

9. Extractive Summary Generation Model

To categorize and extract the important Urdu phrases and sentences to create summary of the text, the model of extractive summarization uses different techniques based on machine learning and NLP. Most commonly used approaches and techniques are SVM and TF-IDF which plays an important role to achieve the research objectives [34].

In this process, sentences are assigned scores based on their TF-IDF values, with higher scores indicating greater importance. The algorithm then ranks the sentences and selects the top N phrases based on their prediction scores, which are subsequently used to generate the extracted summary.

10. Algorithm Used

The following learning algorithms are discussed:

a. Confusion Matrix

To compare the properly identified outcomes with improperly identified outcomes we use confusion matrix. The result show in table 3.

Table 3. Confusion Matrix

	Predicated True	**Predicated False**
Actual true	True Positive (TP)	False Negative (FN)
Actual false	False Positive (FP)	True Negative (TN)

b. Results Accuracy

The Table 3 provides a basic view, a more comprehensive measure of classifier performance is needed. Accuracy is one such metric, calculated as follows:

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

c. Recall (Sensitivity)

Recall evaluates the method's ability to identify true positives among all actual positives, computed by:

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

d. Precision

Precision measures the proportion of true positives among all instances predicted as positive. It is calculated using:

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

e. F-1 Score

The F1 Score is a metric that combines precision and recall into a single measure to evaluate the overall performance of a summarization method. It is given by:

$$F1 - Score = 2\frac{Precision\ x\ Recall}{Precision + Recall}$$

f. ROUGE Score

ROUGE (Recall-Oriented Understudy for Gisting Evaluation) scores assess the quality of automatic summaries by comparing them to reference summaries created by humans. These scores include precision, recall, and F1 measures for various n-grams and word overlaps.

11. Abstractive Summary

Our abstractive summarization model was evaluated using the ROUGE metric to analyze its performance. ROUGE is a widely recognized method for assessing the quality of summaries by comparing them with reference summaries [35]. The model we developed is based on a sequence-to-sequence architecture that employs transformers to create both an encoder and a decoder. To enhance the quality of the summaries, we implemented a beam search technique.

The evaluation results of our proposed model were promising. It achieved ROUGE scores with an average F1 score of 0.44 for ROUGE-1, 0.26 for ROUGE-2, and 0.24 for ROUGE-L. These results highlight the effectiveness of our model in natural language processing and summarization tasks, showing its ability to generate accurate and relevant summaries from a variety of input texts.

Table 4. Abstractive model and evaluation metric

Model Type	Evaluation Metric	Average F1 Score	Accuracy
Abstractive Model (Transformer)	ROUGE-1	0.44	-
	ROUGE-2	0.26	-
	ROUGE-L	0.24	-

12. Extractive Summary

Our extractive summarization model employs the Support Vector Machine (SVM) technique to identify key sentences for summarization. SVM is well-suited for binary classification problems, and we have tailored this approach by treating each sentence as a data point to decide its inclusion in the summary.

The evaluation of the model demonstrated promising results, with an F1 score of approximately 0.75 and an accuracy rate of around 86%. This SVM-based extractive summarization algorithm effectively identified and selected relevant sentences, resulting in concise and informative summaries across various texts.

Table 5. Extractive model and evaluation metric

Model Type	Evaluation Metric	Average F1 Score	Accuracy
Extractive model (SVM)	F1 Score	0.75	0.86
	Accuracy	-	0.86

13. Abstractive and Extractive Summary

The results indicate that both the abstractive and extractive summarization methods are effective. The performance of the abstractive model is assessed using the ROUGE evaluation metric, while the effectiveness of the extractive model is measured by its F1 score and accuracy.

The abstractive model, despite its complexity, achieved a high ROUGE score, demonstrating its ability to produce valuable summaries. Similarly, the extractive model showed high F1 scores and accuracy, reflecting its capability to identify and select important sentences for concise summarization.

Table 6. Abstractive and extractive model and evaluation metric

Model Type	Evaluation Metric	Average F1 Score	Accuracy
Abstractive Model (Transformer)	ROUGE-1	0.44	-
	ROUGE-2	0.26	-
	ROUGE-L	0.24	-
	F1 Score	0.75	0.86

Extractive model (SVM)	Accuracy	-	0.86

14. Conclusion and Future Directions

This study explores the benefits of text summarization techniques applied to lengthy Urdu news articles, focusing on both extractive and abstractive methods. Given the complexities of the Urdu language, including its intricate word meanings and morphological challenges, the study underscores the importance of adapting summarization strategies to address these difficulties. It highlights how these techniques improve sentence structure and semantic understanding through advanced technologies such as deep learning and natural language generation. The findings provide valuable insights into the effectiveness of summary models. The abstractive summarization model achieved a ROUGE score of 0.44, while the extractive model performed better with an F1 score and precision rate of 0.86. The research underscores the value of summarization in enhancing the accessibility and engagement of Urdu speakers with news content. It also identifies potential areas for further investigation, including the impact of bias on news reporting and advancements in summarization technology.

Future research should explore integrating visual elements into Urdu text summarization. This approach could enhance the completeness and in formativeness of summaries by incorporating visual data alongside textual information. Investigating the interplay between written and visual content may lead to more effective summarization methods that capture the full context provided by images and videos. Understanding the role of visual features in summarization could lead to the development of more advanced and efficient models, tailored to the unique aspects of the Urdu language. This aligns with the evolving field of natural language processing and information retrieval, offering new opportunities to improve and enrich Urdu text summarization in a multimedia-driven environment.

References

1. Andhale N. and Bewoor, L. A. "An overview of Text Summarization techniques," International Conference on Computing Communication Control and automation (ICCUBEA), Pune, India, 2016, pp. 1-7.
2. Allahyari, M., Pouriyeh, S., Assefi, M., Safaei, S., Trippe, E. D., Gutierrez, J. B., & Kochut, K. "Text summarization techniques: a brief survey," 2017, arXiv preprint arXiv:1707.02268
3. Nenkova, A., and McKeown, K. "A survey of text summarization techniques", 2016, Mining text data, pp. 43-76.
4. Khan, F., and Fraz, F. "Rule-based sentence extraction for Urdu text summarization". In 2010 International Conference on Frontiers of Information Technology IEEE, 2010, pp. 155-160.
5. Tariq, M. A., and Raza, S. "Urdu news summarization using graph-based approach". International Journal of Advanced Computer Science and Applications, 2017, vol. 8, no. 6, pp. 102-108.
6. Nawaz, A., Bakhtyar, M., Baber, J., Ullah, I., Noor, W., & Basit, A. "Extractive text summarization models for Urdu language", Information Processing & Management, 2020, vol. 57, no. 6, p. 102383.
7. Mahmood, T., Ahmed, N., Ahmed, M., & Islam, N. "A Deep Learning Based Approach for Urdu News Summarization", 2021, arXiv preprint arXiv:2101.12550.
8. Ali, A., and Khan, F. "A hybrid approach for Urdu text summarization". In 2012 11th International Conference on Information Sciences, Signal Processing and their Applications (ISSPA) IEEE, 2019, pp. 296-301.
9. Hassan, S., Khan, F., & Hafeez, U. "Clustering-Based Urdu Text Summarization". In the International Conference on Innovative Computing and Communication Springer, Cham, 2021, pp. 276-286.
10. Mir, M. F., Mushtaq, U., & Qureshi, I. M. "A rule-based approach for Urdu news summarization", Journal of King Saud University-Computer and Information Sciences, 2022, vol. 31, no. 1, pp. 77-85.
11. Ali, S. S., and Shah, S. A. "Genetic algorithm-based summarization model for Urdu language using artificial neural network", Multimedia Tools and Applications, 2021, vol. 79, no. 23, pp.16211-16234.
12. Rashid, T., Muhammad, F., & Hussain, M. "Graph-Based Urdu News Summarization", IEEE Access, vol. 9, pp. 49215-49225.
13. Hafeez, R., Anwar, M. W., Jamal, M. H., Fatima, T., Espinosa, J. C. M., López, L. A. D., & Ashraf, I. "Contextual Urdu Lemmatization Using Recurrent Neural Network Models", Mathematics, 2021, vol. 11, no. 2, p. 435.
14. Lin, C. Y. "ROUGE: A Package for Automatic Evaluation of Summaries. In Text Summarization Branches Out", Proceedings of the ACL-04 Workshop, 2010, pp. 74-81.
15. Khalil, M. I. K.; Afsheen, A.; Taj, A.; Nawaz, A.; Jan, N.; and S. Ahmad, "Enhancing Security Testing Through Evolutionary Techniques: A Novel Model" Journal of Computing & Biomedical Informatics, 2023, vol. 6, no. 1, pp. 375 – 393.
16. Khalil, M. I. K. and Taj, T. "Factors Affecting the Efficacy of Software Development: A Study of Software Houses in Peshawar, Pakistan," International Review of Basic and Applied Sciences, 2021, vol. 9, no. 3, pp. 385-393.
17. Khalil, M. I. K. Ullah, A.; Taj, A.; Khan, I.A.; Ullah, F.; Taj, F.; and Shah, S. "Analysis of Critical Risk Factors Affecting Software Quality: A Study of KPK, Pakistan Software Industry," International Review of Basic and Applied Sciences, 2022, vol. 10, no. 2, pp. 338-348.
18. Saqib, A.; Ullah, M.; Hyder, S.; Khatoon, S.; and Khalil, M. I. K. "Creative Decision Making in Leaders: A Case of Beer Game Simulation," Abasyn Journal of Social Sciences, 2020, vol. 12, no. 2, pp. 379-387.
19. Khan, I.A; Ullah, F.; Abrar, M.; Shah, S.; Taj, F. and Khalil, M.I.K. "Ransomware Early Detection Model using API-Calls at Runtime by Random Decision Forests," International Review of Basic and Applied Sciences, 2022, vol. 10, no. 2, pp. 349-359.
20. Jan, S.; Maqsood, I.; Ahmad, I.; Ashraf, M.; Khan, F. Q.; and Khalil, M. I. K. "A Systematic Feasibility Analysis of User Interfaces for Illiterate Users," Proceedings of the Pakistan Academy of Sciences, 2020, vol. 56, no. 4, pp. 2518-4253.
21. Khalil, M. I. K.; Shah, S. A. A.; Khan, I. A.; Hijji, M.; Shiraz, M.; and Shaheen, Q. "Energy cost minimization using string matching algorithm in geo-distributed data centers," Computers, Materials, and Continua, 2023, vol. 75, no. 3, pp. 6305-6322.
22. Ahmad, I.; Khalil, M. I. K.; and Shah, S. A. A. "Optimization-based workload distribution in geographically distributed data centers: A survey," International Journal of Communication Systems, 2020, vol. 33, no. 12, p. e4453.
23. Khalil, M. I. K.; Ahmad, I.; Shah, S. A. A.; Jan, S.; and Khan, F. Q. "Energy cost minimization for sustainable cloud computing using option pricing," Sustainable Cities and Society, 2020, vol. 63, p. 102440.

24. Khalil, M. I. K.. "Improve quality of service and secure communication in mobile adhoc networks (MANETS) through group key management," International Review of Basic and Applied Sciences, 2013, vol. 1, no. 3, pp. 107-115.
25. Muhammad, D.; Ahmad, I.; Khalil, M. I. K.; Khalil, W.; and Ahmad, O. A. "A generalized deep learning approach to seismic activity prediction," Applied Sciences, MDPI, 2023, vol. 13, p. 1698.
26. Khalil, M. I. K. "Job satisfaction and work morale among Ph.D's: A study of public and private sector universities of Peshawar, Pakistan," International Review of Management and Business Research, 2013, vol. 02, no. 2, p. 362.
27. Ahmad, I.; Ahmad, M. O.; Alqarni, M. A.; Almazroi, A. A.; and Khalil, M. I. K. "Using algorithmic trading to analyze short-term profitability of Bitcoin," PeerJ Computer Science, 2021, vol. 7, p. e337.
28. Khalil, M. I. K.; Shah, S. A. A.; Taj, A.; Shiraz, M.; Alamri, B.; Murawat, S.; and Hafeez, G. "Renewable aware geographical load balancing using option pricing for energy cost minimization in data centers," Processes, MDPI, 2022, vol. 10, no. 10, p. 1983.
29. Khalil, M. I. K.; Ahmad, A.; Almazroi, A.A. "Energy Efficient Workload Distribution in Geographically Distributed Data Centers," IEEE Access, 2019, vol. 7, no. 1, pp. 82672-82680.
30. Khalil, M. I. K..; Mubeen, A.; Taj, A.; Jan, N.; Ahmad, S. "Renewable and Temperature Aware Load Balancing for Energy Cost Minimization in Data Centers: A Study of BRT, Peshawar," Journal of Computing & Biomedical Informatics, 2023, vol. 06, no. 3, pp. 183-194.
31. Anum, H.; Khalil, M. I. K..; Nawaz, A.; Jan, N.; and Ahmad, S. "Enhancing rumor detection on social media using machine learning and empath features," Journal of Computing & Biomedical Informatics, 2023, vol. 06, no. 2, pp. 272-281.
32. Khalil, M. I. K.; Khan, I.A; Nawaz, A.; Latif, S.; Ahmad, S. and Ahmad, S. "Unveiling the security Maze: A comprehensive review of challenges in Internet of things," Special Issue on Intelligent Computing of Applied Sciences and Emerging Trends (ICASET), Journal of Computing & Biomedical Informatics, 2024, pp. 10-19.
33. Ahmad, I.; Ahmad, M. O.; Alqarni, M. A.; Almazroi, A. A.; and Khalil, M. I. K. "Using algorithmic trading to analyze short-term profitability of Bitcoin," PeerJ Computer Science, 2021, vol. 7, p. e337.
34. Khalil, M. I. K.; Shah, S. A. A.; Taj, A.; Shiraz, M.; Alamri, B.; Murawat, S.; and Hafeez, G. "Renewable aware geographical load balancing using option pricing for energy cost minimization in data centers," Processes, MDPI, 2022, vol. 10, no. 10, p. 1983.
35. Rahman, Z.U..; Khalil, M.I.K..; Nawaz, A.; Khan, I.A..; Jan, N..; and Sheeraz, A. "Analysis and clustering of Pakistani music by lyrics: A study of CokeStudio Pakistan," Journal of Computing & Biomedical Informatics, 2024, vol. 7, no. 1, pp. 281-296.