Volume 9, Issue 12, December – 2024 International Journal of Innovative Science and Research Technology
ISSN No:-2456-2165 https://doi.org/10.5281/zenodo.14575886

Comparative Analysis of Large Language Models and Traditional Methods for Sentiment Analysis of Tweets Dataset

Suhag Pandya

Abstract:- Sentiment analysis is widely recognised as the most actively researched area in data mining. These days, a number of social media platforms have been created, and Twitter is a crucial instrument for exchanging and gathering people's thoughts, feelings, opinions, and attitudes about specific things. This paper presents a comparative analysis of Large Language Models (LLMs) and ML models for sentiment analysis on a Twitter dataset. The study evaluates a performance of XLNet, advance algorithms including KNN, RF, and XGBoost. The sentiment analysis methodology involves pre-processing the Twitter dataset through noise removal and tokenisation, followed by feature extraction using methods like Bag-of-Words and Word2Vec. Results show that XLNet is superior to the conventional models; It has 99% precision, recall, and F1-score values and an accuracy rate of 99.54%. In comparison, KNN achieves 78% accuracy, 85% precision, 88% recall, and 86% F1 score, while RF and XGBoost exhibit lower performance with accuracy rates of 69% and 60%, respectively. The performance comparison highlights the superior capabilities of XLNet for sentiment classification tasks, indicating its potential for enhancing text classification applications. Research in the future can look at ways to improve XLNet's performance on bigger and more complicated datasets by combining it with more sophisticated deep learning methods like attention mechanisms and transfer learning.

Keywords:- *Sentiment Analysis, LLM, Twitter Dataset, Natural Language Processing, Text Classification, Social Media Analytics.*

I. INTRODUCTION

The Internet has revolutionised the way individuals share their thoughts and ideas these days. These days, most people do it on websites that allow product reviews, social media, online forums, and blogs [1]. Facebook, Twitter, Google Plus, and many other social media platforms are attracting millions of users every day [2]. To share their thoughts and feelings as well as comments regarding their everyday lives [3]. Consumers may educate and persuade one another through online forums and online communities, which is an interactive kind of media [4]. Blog entries, reviews, comments, tweets, and status updates all contribute to the mountain of sentiment-rich data that is generated by social media [5].

Sentiment analysis (SA) informs consumers before they purchase a product whether the information is suitable [6][7]. This analytical data is used by marketers and businesses to better understand their goods and services so that they may be provided according to the user's requirements [8][9]. Processing, finding, or analysing the factual material available is the major emphasis of textual information retrieval strategies [10].

Worldwide, thanks to technological advancements, social media platforms like Instagram, Facebook, LinkedIn, YouTube, etc. [11][12]. people's ideas and thoughts on things, events, or objectives can be expressed through [13][14]. In today's world, people all over the world love to communicate their opinions and thoughts through short messages called tweets on the popular microblogging network known as Twitter [15][16]. An enormous amount of sentiment data derived from analyses of tweets is frequently generated [17]. Twitter is a great platform for sharing news and interacting with other users online [18][19]. The way people feel on Twitter has a big impact on many parts of our lives. The goal of SA and text classification is to extract information from texts and then classify a polarity as neutral (Ne), negative (N), or positive (P)[20].

Traditional machine learning (ML) approaches have long been utilised for SA tasks, relying on extensive preprocessing and feature extraction techniques [21][22]. However, the advent of LLMs such as GPT and BERT has transformed the landscape of NLP [23][24]. These models, pre-trained on vast corpora of text, demonstrate a superior ability to comprehend context, capture nuances, and generate accurate predictions, thereby outperforming conventional methods in many scenarios[25][26]. The main objective of this project is to use Twitter datasets to evaluate and contrast LLMs with more traditional ML techniques for sentiment analysis.[27]. By leveraging the strengths of both paradigms, this research seeks to highlight their efficacy, limitations, and practical applications.

- *Significance and Contribution of Study*

This study highlights a significance of leveraging both LLMs and traditional methods for sentiment analysis to enhance accuracy and contextual understanding. Finding out how well each method worked on a real-world Twitter dataset allows us to compare and contrast their merits and shortcomings. The findings contribute to advancing sentiment analysis techniques, enabling more effective applications in social media analytics and NLP. The contribution of this study

ISSN No:-2456-2165 https://doi.org/10.5281/zenodo.14575886

lies in its comprehensive evaluation of various Large language models for sentiment analysis using the Twitter dataset. The main contributions are:

- Analysed a substantial Twitter dataset with over 31,000 tweets, ensuring the study reflects practical applications.
- Captured semantic relationships by transforming words into vectors, enabling context-aware sentiment analysis.
- Identified key sentiment-related features, enhancing model efficiency and predictive accuracy.
- Provided a foundational baseline for sentiment analysis, enabling comparison with advanced techniques like Word2Vec.
- Compare the effectiveness of LLMs, such as XLNet, with traditional ML models like RF, XGBoost, and KNN for sentiment analysis of Twitter data.
- The models were evaluated using a number of relevant measures, including F1-score, recall, accuracy, and precision.

➢ *Structure of the Paper*

The research is organised as follows: In Section II, the existing literature on text classification is reviewed. In section III Methodology utilised to compile the data for this study. The findings and commentary on the text categorisation are presented in Section IV Section V concludes with a present finding.

II. LITERATURE REVIEW

A survey of the literature on sentiment is included in this section. Text categorisation using analysis of the Tweets dataset.

In this study, Ihnaini et al. (2024) fine-tuning process incorporates both supervised techniques and reinforcement learning from human feedback, specifically designed to align the models with the historical and cultural context of Song Ci. Notably, the ChatGLM-6B(8-bit) model achieved the best F1 Score of 0.840, demonstrating its exceptional ability to merge ancient literary analysis with modern computational technology, thereby broadening our understanding of the emotional spectrum of classical Chinese poetry [28].

In this study, Patrick et al. (2023) The model is developed and tested using a variety of learning algorithms, including Linear Regression, KNN, SVM, RF, Bagging, and Gradient Boosting. There are two sets of data: the test dataset and the training dataset. The test as well as training datasets were used to create and evaluate the model. It was determined that 1.1% of respondents were untrustworthy due to unanswered questions by the reliability check. The results demonstrate that out of the two algorithms, RF-supervised ML yields a superior accuracy of 0.711 compared to KNN's 0.515 [29].

In this study, Gore et al. (2023) conducted preprocessing using tokenisation, stemming, punctuation mark removal, and stop word removal. Word clouds may be used for both positive and negative word analysis. They got both good and bad comments on the Apple iPhone 14 Pro Max. Reviews that are favourable tend to be ones that are dynamic, good, improving, gorgeous, terrific, excellent, etc. Several machine learning techniques were used, including NB, LR, and SVM. LR achieved 90% accuracy, SVM 91% accuracy, and NB 94% accuracy [30].

In this study, Waspodo et al. (2022) The steps that makeup preprocessing include cleaning, lemmatisation, stemming, tokenising, and stop word removal. This study employs a variety of unsupervised learning techniques: Supervised learning based on Lexicon: SVM. The normalised Lexicon Sentiment with scaling is used to establish positive, neutral, and negative sentiment classes, and this data is then used to weight textual data. With 940 tweets analysed, they found 330 positive (35% of the total), 302 negative (32% of the total), and 308 neutral (33% of the total) tweets. The findings showed that the SVM achieved an average accuracy of 98.75% in its classification tests, with a precision of 0.98 and a recall of 0.98 [31].

In this study, Asmawati, Saikhu and Siahaan (2022) look at four different sentiment analysis approaches and evaluate their efficacy on text and image-based Indonesian memes. NB, SVM, DT, and CNNs were the supervised ML algorithms used to classify the retrieved text memes after their extraction. The study results showed that the most beneficial effects were obtained using sentiment analysis on meme text using the NB technique, with an accuracy rate 65.4% [32].

In this study, Gope et al. (2022) Sentiment analysis seeks to categorise an amount of good and negative emotions expressed in a certain text. ML methods such as LR, MNB, BNB, RF, and LSVM were used. They achieved a 91.90% success rate using the RFC. In our study, they also achieved maximum accuracy (97.52%) using RNN with LSTM as a DL strategy. Our model is well-suited to the RNN-LSTM technique [33].

In this paper, Al-Hagree and Al-Gaphari (2022) Data used for SA tasks were collected from evaluations left by users of banking mobile applications on the Google Play Store. The Arabic sentiment analysis is carried out by use of ML algorithms, namely the NB, KNN, DT, and SVM patterns. NB model has outperformed competing DT, KNN, and SVM algorithms in terms of evaluation quality. In comparison to the other models, the NB model had exceptional results according to recall(88.08%), accuracy(88.25%), and F-score(88.25%) [34].

Below, Table 1 provides a summary of a literature review with dataset approaches, results and limitations for text dataset classification.

Table 1 Summary of Literature Review of Sentiment Analysis of Tweets Dataset for Text Classification

Author	**Methods**	**Dataset**	**Accuracy**	**Limitation/gap**
Patrick et.al.	Linear Regression, KNN, SVM, RF, Bagging, GB.	Dataset of 884 employees was reduced to 802 after cleaning	Random Forest: 71.1%, KNN: 51.5%	Limited reliability control, only 1.1% deemed unreliable; lower accuracy with certain models
Gore et al.	Naive Bayes, SVM, Logistic Regression	Reviews on Apple iPhone 14 Pro Max	Logistic Regression: 90%, SVM: 91%, Naive Bayes: 94%	No mention of deep learning or ensemble techniques
Ihnaini et.al.	ChatGLM-6B (8-bit) model	Ancient Chinese poetry	F1 Score: 0.840	Focused on a very niche area, difficult to generalise results to modern language sentiment analysis
Waspodo et.al.	Lexicon-based (unsupervised) & SVM (supervised)	940 tweets	SVM: Accuracy: 98.75%, recall: 0.98, precision: 0.98	Narrow focus on tweets, results might not generalise to other text types.
Asmawati, Saikhu et.al.	Naive Bayes, SVM, Decision Tree, Convolutional Neural Networks	Indonesian memes (text and images)	Naive Bayes: 65.4%	Accuracy lower for complex text and image data combinations
Gope et.al.	Linear SVM, RF, Naive Bayes (Multinomial & Bernoulli), LR, RNN-LSTM	Amazon product reviews	Random Forest: 91.90%, RNN-LSTM: 97.52%	Limited focus on deep learning methods apart from RNN-LSTM
Al-Hagree and Al-Gaphari et al.	Naive Bayes, KNN, Decision Tree, SVM	Mobile app user reviews (Arabic)	Naive Bayes: 89.65%, recall: 88.08%, precision: 88.25%, F1: 88.25%	Focused only on Arabic language, other language models not explored.

III. METHODOLOGY

The methodology for the comparative analysis of Large Language Models (LLMs) and traditional methods for sentiment analysis of the Twitter dataset involves several key steps. Figure 1 shows a workflow of flowchart sentiment analysis for text classification employing ML models to analyse Twitter data. The sentiment analysis methodology involves analysing a Twitter dataset labelled with Positive and Negative sentiments. The dataset undergoes pre-processing to clean and standardise text, including noise removal and tokenisation. Traditional methods like Bag-of-Words and Word2Vec are used for feature extraction, capturing sentiment-relevant features such as word frequencies and vector similarities. The processed For model assessment, data is divided into training (80%) and testing (20%) sets. Both Large Language Models (LLMs) like XLNet and traditional ML algorithms like RF, XGBoost, and KNN are applied for classification. The models are trained and tested, and The F1-score, accuracy, precision, and recall are the typical measures used to assess their performance. Finally, the models forecast a sentiment of tweets as either Positive or negative, and their results are compared to determine the most effective approach for text classification.

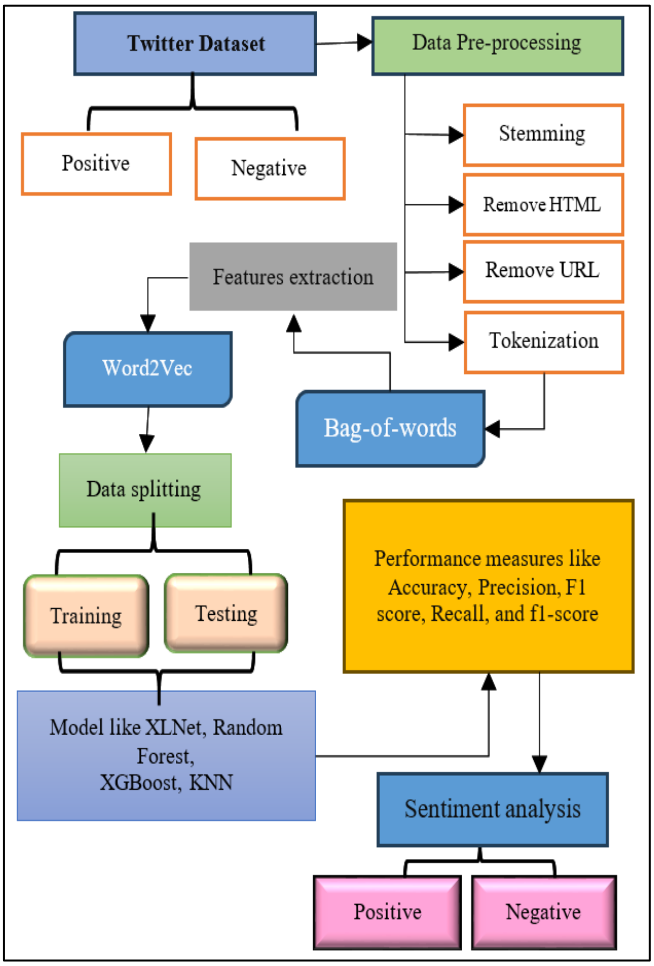

Fig 1 Flowchart for Sentiment Analysis

ISSN No:-2456-2165 https://doi.org/10.5281/zenodo.14575886

A. A following Steps of flowchart are Explained below:

➢ *Dataset Description*

Use the developer account to gather the Twitter API for this investigation. In training datasets, there are 31962 tweets in total, with "29959" representing positive values and "2003" representing negative values. There were four categories of perplexity in the data they gathered: joyful, sad, cheerful, and hateful. They concluded that the data had to be in binary form once they organised it [35]. For this, they made cheerful and cheery “Positive” and sad and hate “Negative”. In the realm of data management, this exists [36].

➢ *Data Preprocessing*

After the dataset has been collected, the following stage is to preprocess it in order to eliminate any junk values and extraneous information that may have been overlooked since it is a real-life dataset [37][38]. The following preprocessing steps are listed below:

- **Removing URLs:** Eliminate any hash marks (like #subject), URLs (like www.xyz.com), and targets (@username)[39].
- **Remove HTML:** As a text preparation step, HTML tag removal helps clean up HTML documents' text data [40][41].
- **Stemming:** The stemming technique is used to remove prefixes and suffixes from words [42]. The process of identifying and eliminating a word's root and stem is another possible definition [43].
- **Tokenization:** The process of tokenisation involves separating text into a set of discrete, meaningful components [44].

➢ *Bag-of-Words*

The information can be easily transformed into a bag-of-words format [45][46]. In this representation, some common words, like ‘go’ and ‘see,’ show a high distinction between classes, while others occur in similar amounts for every class.

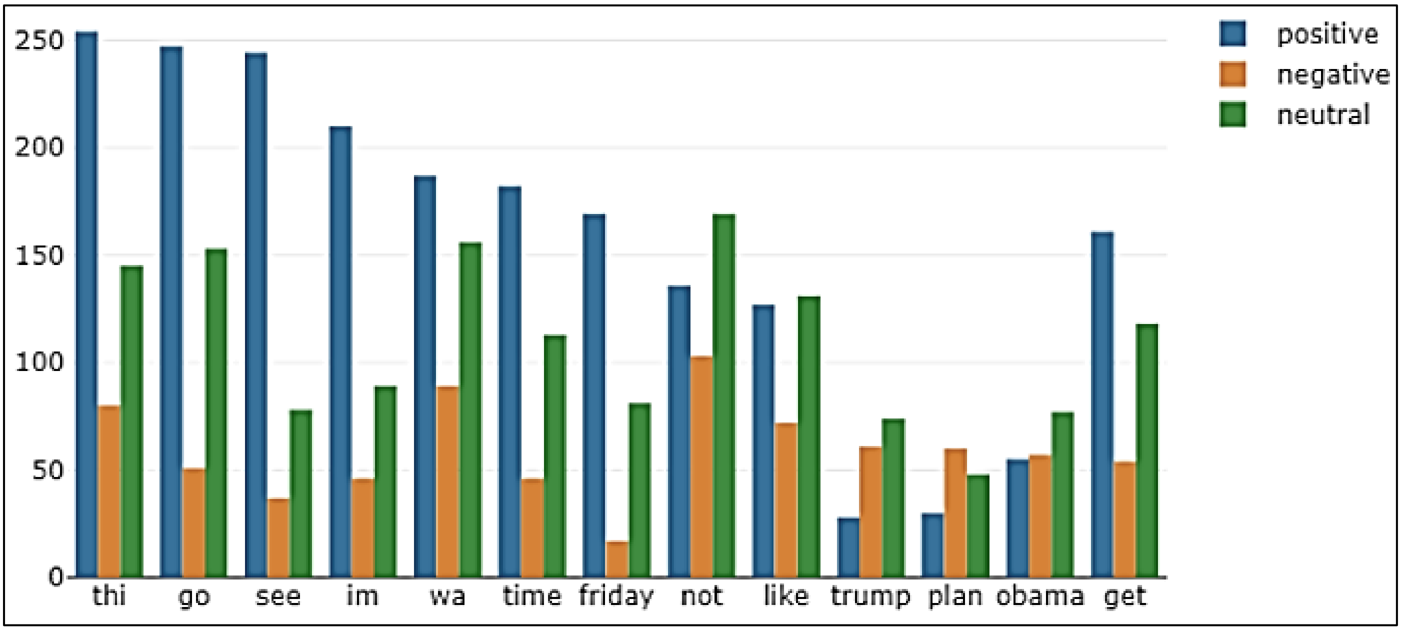

Fig 2 Most Common Words across Sentiments

Figure 2 illustrates sentiment analysis across common words, categorised into positive (blue), negative (orange), and neutral (green) sentiments. Words like "thi," "go," "time," "like," and names like "trump" and "obama" are analysed, with the y-axis showing frequency counts up to 250. This analysis, likely derived from social media or text data, reveals how word usage correlates with sentiment expression, offering insights into linguistic and emotional trends in the dataset.

➢ *Feature Extraction*

Feature extraction is a powerful tool for reducing resource requirements while preserving critical data [47][48]. The effectiveness and precision of ML models are greatly enhanced by feature extraction [49]. Emotion counts (both positive and negative), question marks, hashtags, and exclamation points are some of the key characteristics [50].

➢ *Word2Vec*

The process of word2vec involves converting words to vectors and then identifying words that are similar to each other. This enables it to identify words that reflect distinct emotions [51]. The construction of Word2Vec involves calculating the average similarity of a whole tweet to a given word, with the average values normalised to [0,1] to account for variations in word count across tweets [52][53]. Figure 3 displays the terms that are comparable.

ISSN No:-2456-2165 https://doi.org/10.5281/zenodo.14575886

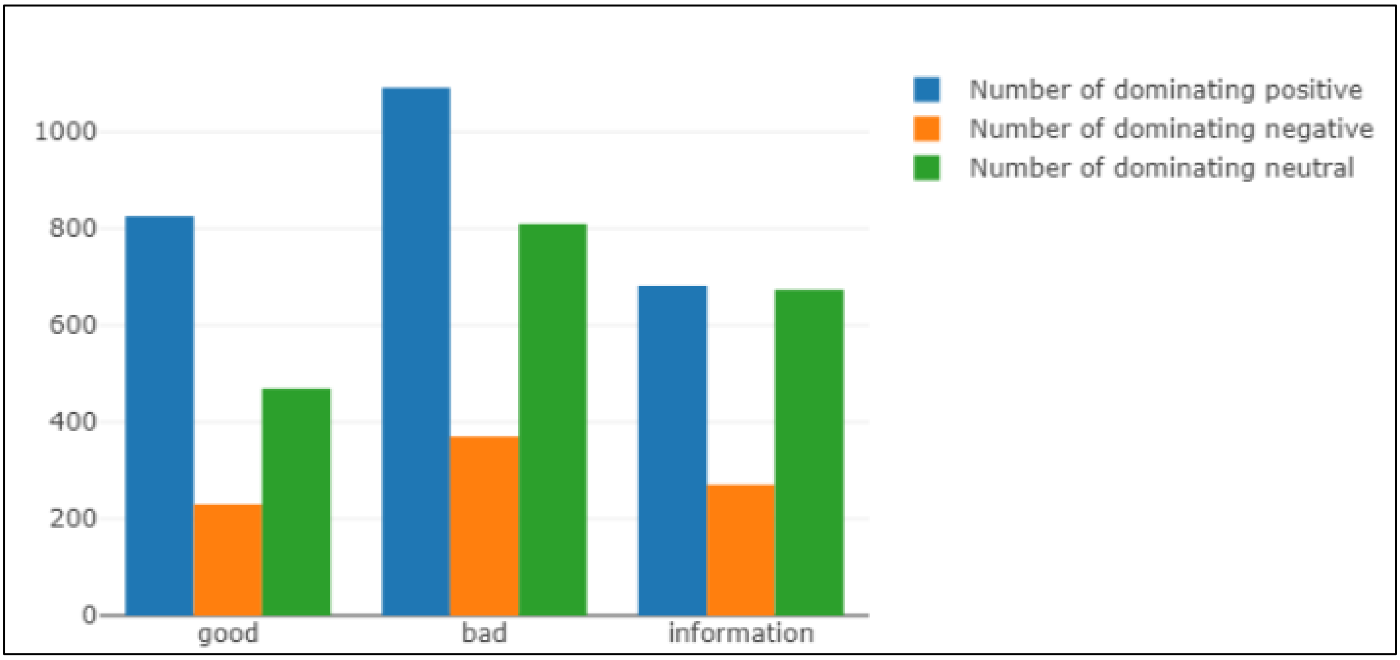

Fig 3 Words Similarity

Figure 3 visualises word similarity patterns across three keywords - "good," "bad," and "information" - showing their associations with different sentiment categories. For each word, the bars display varying heights indicating the frequency or strength of association, with "bad" showing the highest positive sentiment count (around 1000), followed by "good" (around 800), while "information" has more balanced sentiment distributions across all three categories. This visualisation effectively demonstrates how words that might seem inherently positive or negative can actually have complex sentiment associations in real-world usage.

➢ *Data Splitting*

Data splitting in ML models is a crucial step for evaluating the performance of models. The basic premise is to partition the dataset into separate parts for testing and training. They use 80% of the data for training and just 20% for testing.

➢ *Large Language Model (XLNET)*

XLNet[54] is one of Google AI's 2019 Transfer Learning models; it's similar to BERT but uses an AR pre-training strategy to generalise its features, making it perform better than BERT on many benchmark datasets [12][55]. They demonstrate below how XLNet has used PLM to overcome AE models' shortcomings, namely the issue of obtaining bidirectional context [56]. In comparison to BERT, XLNet's convergence time is much longer due to the fact that it trains through all possible word sequences utilising permutations of occurrences for a particular word [57][58].

The core concept of XLNet is to enhance PLM with additional capabilities in order to capture situations that are bidirectional [59][60]. It is possible to accomplish AR factorisation in T! by thinking about every possible location on each side of a token. separate sequences x tokens of length T in a phrase [61]. Consider ZT to be the collection of every combination of T-length sequences that could exist.

$$\max E_z \sim Z_T\{\sum_{t=1}^{T} \log p(x_{z_t} | x_z < t)\} \quad (1)$$

Where z_t and z < t represent a permutation's t-th and t-1 elements, respectively. Z_T

A likelihood of token x_{z_t} given previous tokens $x_{z_t<t}$ is computed using the XLNet auto-regressive permutation approach, as displayed in Equation 1. Although XLNetonly modifies the factorisation order and not a sequence_order during training, it preserves an original sequence_order and utilises Transformers to match the original sequence's positional encoding[62][63]. This quality is helpful for fortuning as it takes into account just the sequence's natural order [64]. They thus use XLNet in our research since its design differs from that of BERT [65][66].

➢ *Performance Metrics*

To evaluate the performance of phishing email detection, used a set of evaluation metrics, also known as performance metrics [67][68]. A confusion matrix may be used to compare the predicted and actual values of a model in order to evaluate its accuracy [69][70]. Five measures were utilised to evaluate the final models: accuracy, precision, recall, and score. This evaluation process begins with confusion matrices that use the following criteria to rank the models: with TP, FP, TN, and FN, where:

- TP indicates the amount of occurrences when the actual class was accurately predicted.
- FN, the picture shows how many times the real class was wrongly assumed to be some other kind.
- TN displays the quantity of correctly classified records as normal.
- FP, it is the sum of all the times a different class was wrongly assumed to be the real class.

ISSN No:-2456-2165 https://doi.org/10.5281/zenodo.14575886

- **Accuracy:** The ratio of properly categorised undesirable and normal events to total oil well events in the dataset is known as accuracy. It is stated as follows Eq.(2):

$$Accuracy = \frac{TP+TN}{TP+FP+FN+TN} \quad (2)$$

- **Precision:** According to formula (3), precision is a percentage of positive observations that accurately forecast a total number of positive predictions.

$$Precision = \frac{TP}{TP+FR} \quad (3)$$

- **Recall:** Recall is a percentage of properly detected positive observations, as computed using Eq. (4):

$$Recall = \frac{TP}{TP+FN} \quad (4)$$

- **F1 Score:** The following Eq. (5) is the formula for the F1 Score, which is an all-encompassing assessment that strikes a balance between recall and precision values:

$$F1 = \frac{2 \times recall \times precision}{recall+precision} \quad (5)$$

The purpose of these matrices is to facilitate the evaluation of various ML models on the Twitter dataset.

IV. RESULT ANALYSIS AND DISCUSSION

The experiment is done on the twitter dataset. On this twitter data apply Large Language Model (XLNET), and compare (see in Table 3) with KNN[71], Random Forest[72], XGBoost [73] models. Table II displays the outcomes of a sentiment analysis using a LLM model.

Table 2 XLNet model Performance on Twitter Dataset

Measure	XL Net
Accuracy	99.54
Precision	99
Recall	99
F1-score	99

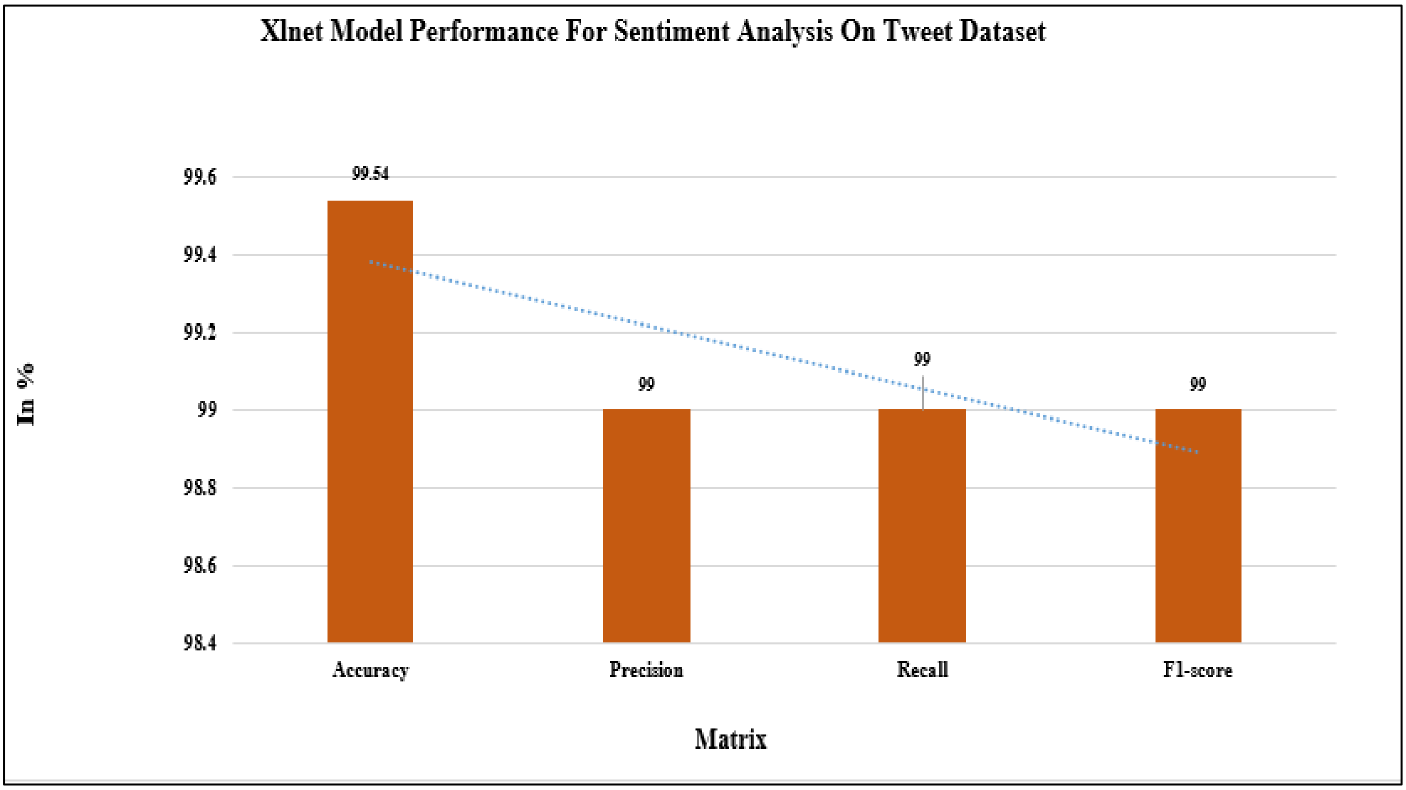

Fig 4 Performance of XL-Net Model

The Figure 4 bar graph displayed a performance of a XLNet model for Twitter data. In this graph, the XLNet model's performance for sentiment analysis achieves an accuracy of 99.54, precision of 99%, recall of 99%, and F1-Score of 99%, reflecting its excellent overall results in classification tasks.

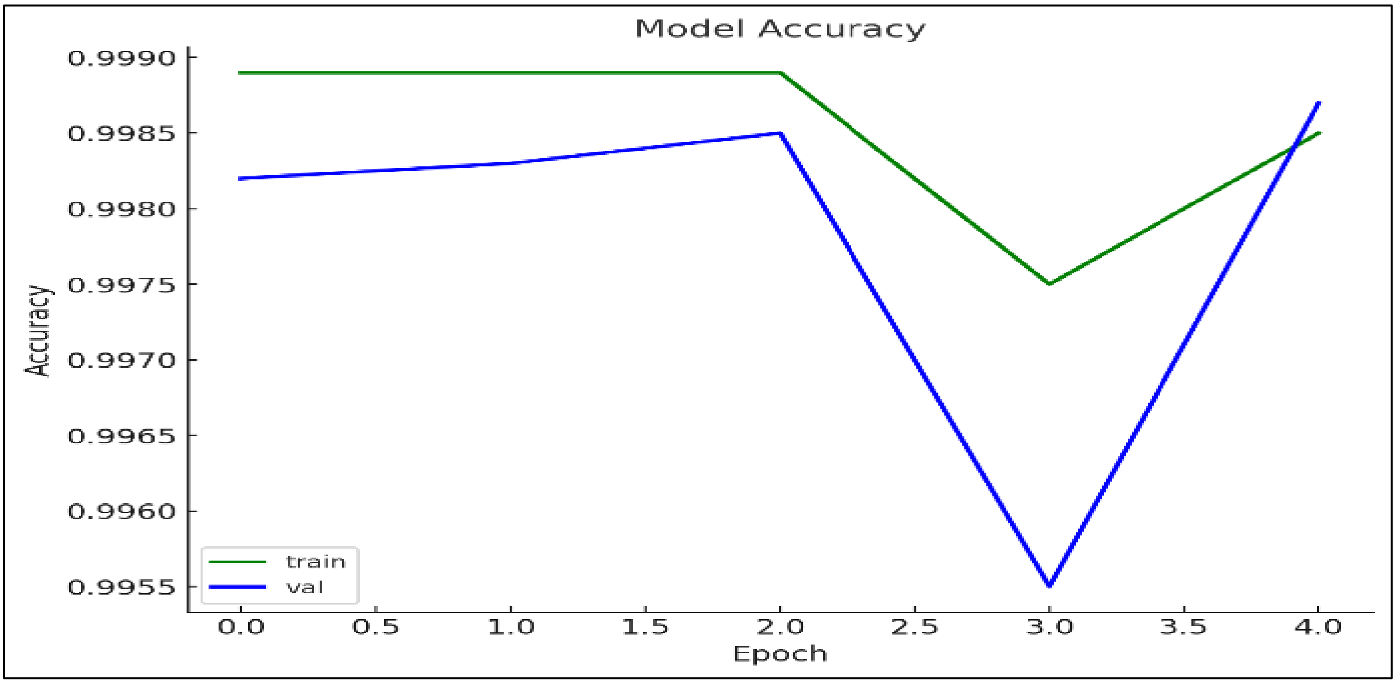

Fig 5 Accuracy Plot for XLNet Model

Figure 5 is a line plot showing the model accuracy of XLNet during training and validation across epochs. The accuracy is displays on a y-axis, while a number of epochs is displayed on an x-axis. A plot shows a fluctuation in validation accuracy around epochs 2-3, with a notable dip, while training accuracy remains relatively stable.

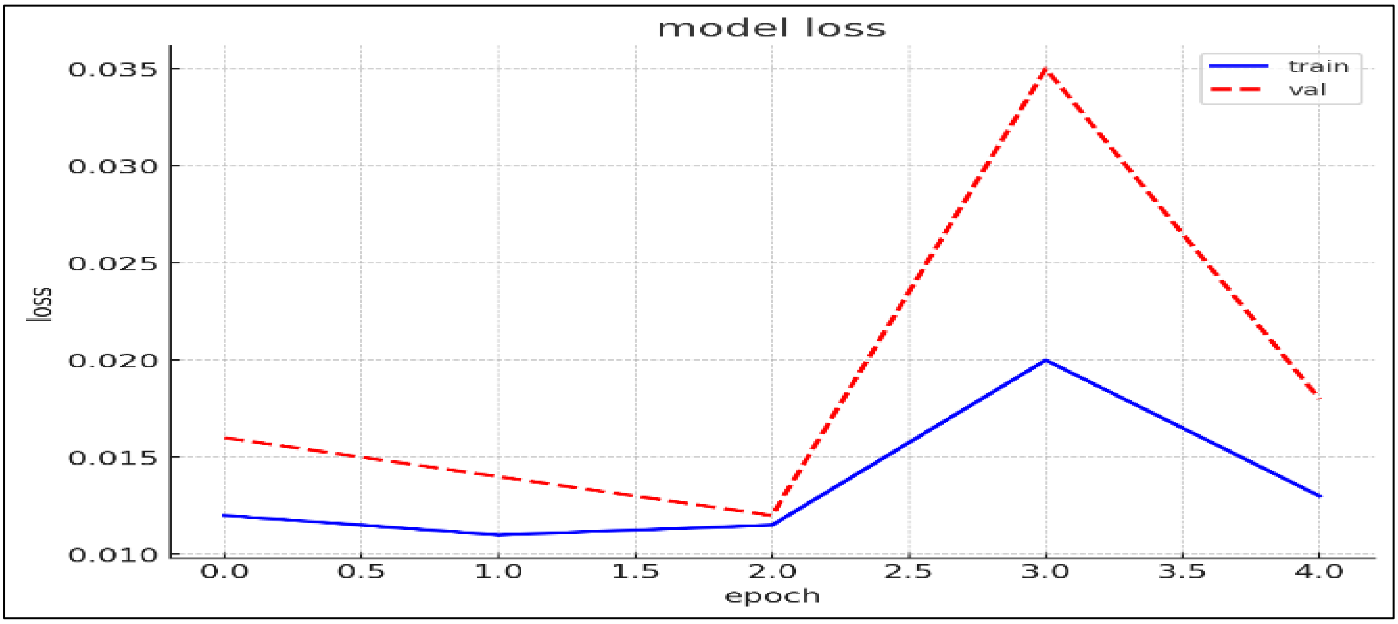

Fig 6 Loss Plot for XLNet Model

The following Figure 6 shows the loss plot for XLNet during training. An x-axis spans epochs between 0 and 4, and the y-axis represents loss values from approximately 0.01 to 0.03. This graph is useful for tracking the model's learning progress and seeing signs of under- or overfitting.

Table 3 Comparison between LLM and Another ML models Performance

Measures	**XLNet**	**KNN**	**RF**	**XGBoost**
Accuracy	99.54	78	69	60
Precision	99	85	44	56.3
Recall	99	88	88	51.3
F1-score	99	86	59	52.3

ISSN No:-2456-2165 https://doi.org/10.5281/zenodo.14575886

The following Table 3 presents the performance of experiments that used ML for text classification on the Twitter dataset. A number of metrics were used to assess an efficacy of a RF, XGBoost, XLNet, and KNN models, including recall, accuracy, precision, and F1-score. In this comparison, KNN achieving superior results with accuracy78%, precision85%, recall88%, and F1-score86%, Random Forest achieves 69% accuracy, with 44% precision, 88% recall and a 59% F1-score and XGBoost achieves 60% accuracy with 56.3% of precision, 51.3 of recall and 52.3% of F1-score indicating relatively moderate performance and XLNet is performed strongly, with an accuracy 99.54%, precision 99%, recall 99%, and F1-score 99%. Overall, the XLNet achieves the highest performance in contrast with other models.

V. CONCLUSION AND FUTURE WORK

As a result of technological advancements, social media has become a vital component of people's daily lives. A person's innermost feelings, beliefs, and opinions may be easily shared via social media. Sentiment analysis's objective is to find out how a piece of writing is generally felt by analysing whether it's good, negative, neutral, or has some other subjective feeling. The significance of sentiment analysis on strategic decision-making makes it crucial for both business and society. In summary, XLNet performs noticeably better than KNN, RF, and XGBoost when compared to traditional ML models for sentiment analysis on the Twitter dataset. It achieves an impressive 99.54% accuracy rate along with high precision, recall, and F1-score. This research does have some limitations, however. One of them is that it only used one dataset, which could not be representative of the complexity of actual sentiment analysis jobs in business. Future work can focus on applying XLNet to larger, more diverse datasets, including multilingual text, to assess its generalizability. Improving the model's performance and its suitability for large-scale sentiment analysis applications in real time might be achieved by integrating Twitter data and investigating advanced approaches like hyperparameter optimisation and transfer learning.

REFERENCES

[1]. B. Patel, V. K. Yarlagadda, N. Dhameliya, K. Mullangi, and S. C. R. Vennapusa, “Advancements in 5G Technology: Enhancing Connectivity and Performance in Communication Engineering,” Eng. Int., vol. 10, no. 2, pp. 117–130, 2022, doi: 10.18034/ei.v10i2.715.

[2]. S. Bauskar and S. Clarita, “AN ANALYSIS: EARLY DIAGNOSIS AND CLASSIFICATION OF PARKINSON’S DISEASE USING MACHINE LEARNING TECHNIQUES,” Int. J. Comput. Eng. Technol., vol. 12, no. 01, pp. 54-66., 2021, doi: 10.5281/zenodo.13836264.

[3]. V. V. Kumar, A. Sahoo, S. K. Balasubramanian, and S. Gholston, “Mitigating healthcare supply chain challenges under disaster conditions: a holistic AI-based analysis of social media data,” Int. J. Prod. Res., 2024, doi: 10.1080/00207543.2024.2316884.

[4]. J. Thomas, “A Survey of E-Commerce Integration in Supply Chain Management for Retail and Consumer Goods in Emerging Markets,” J. Emerg. Technol. Innov. Res., vol. 10, no. 12, 2023.

[5]. V. A. and S. S. Sonawane, “Sentiment Analysis of Twitter Data: A Survey of Techniques,” Int. J. Comput. Appl., 2016, doi: 10.5120/ijca2016908625.

[6]. Sahil Arora and Pranav Khare, “AI/ML-Enabled Optimization of Edge Infrastructure: Enhancing Performance and Security,” Int. J. Adv. Res. Sci. Commun. Technol., vol. 6, no. 1, pp. 1046–1053, 2024, doi: 10.48175/568.

[7]. Pranav Khare and Shristi Srivastava, “Data-driven product marketing strategies: An in-depth analysis of machine learning applications,” Int. J. Sci. Res. Arch., vol. 10, no. 2, pp. 1185–1197, Dec. 2023, doi: 10.30574/ijsra.2023.10.2.0933.

[8]. K. Patel, “AN ANALYSIS OF QUALITY ASSURANCE FOR BUSINESS INTELLIGENCE PROCESS IN EDUCATION SECTOR,” IJNRD - Int. J. Nov. Res. Dev., vol. 9, no. 9, pp. a884–a896, 2024.

[9]. V. S. Thokala, “Improving Data Security and Privacy in Web Applications : A Study of Serverless Architecture,” Int. Res. J., vol. 11, no. 12, pp. 74–82, 2024.

[10]. S. Elmitwalli and J. Mehegan, “Sentiment analysis of COP9-related tweets: a comparative study of pre-trained models and traditional techniques,” Front. Big Data, vol. 7, 2024, doi: 10.3389/fdata.2024.1357926.

[11]. S. G. Thomas Jubin, Kirti Vinod Vedi, “Enhancing Supply Chain Resilience Through Cloud-Based SCM and Advanced Machine Learning: A Case Study of Logistics,” J. Emerg. Technol. Innov. Res., vol. 8, no. 9, 2021.

[12]. N. Abid, “A Climbing Artificial Intelligence for Threat Identification in Critical Infrastructure Cyber Security,” Int. J. Res. Anal. Rev., vol. 9, no. 4, 2022.

[13]. S. S. Pranav Khare, “The Impact of AI on Product Management : A Systematic Review and Future Trends,” Int. J. Res. Anal. Rev. (IJRAR, vol. 9, no. 4, 2022.

[14]. A. P. A. Singh and N. Gameti, “Leveraging Digital Twins for Predictive Maintenance : Techniques , Challenges , and Application,” IJSART, vol. 10, no. 9, 2024.

[15]. V. V. Kumar, F. W. Liou, S. N. Balakrishnan, and V. Kumar, “Economical impact of RFID implementation in remanufacturing: a Chaos-based Interactive Artificial Bee Colony approach,” J. Intell. Manuf., vol. 26, no. 4, pp. 815–830, Aug. 2015, doi: 10.1007/s10845-013-0836-9.

[16]. S. G. Jubin Thomas, Kirti Vinod Vedi, “Effects of supply chain management strategies on the overall performance of the organization,” Int. J. Sci. Res. Arch., vol. 13, no. 01, pp. 709–719, 2024.

[17]. S. Bauskar, “A Review on Database Security Challenges in Cloud Computing Environment,” Int. J. Comput. Eng. Technol., vol. 15, pp. 842–852, 2024, doi: 10.5281/zenodo.13922361.

ISSN No:-2456-2165 https://doi.org/10.5281/zenodo.14575886

[18]. K. Patel, "The Impact of Data Quality Assurance Practices in Internet of Things (IoT) Technology," Int. J. Tech. Innov. Mod. Eng. Sci., vol. 10, no. 10, pp. 1–8, 2024.
[19]. V. S. Thokala, "Integrating Machine Learning into Web Applications for Personalized Content Delivery using Python," Int. J. Curr. Eng. Technol., vol. 11, no. 6, pp. 652–660, 2021, doi: https://doi.org/10.14741/ijcet/v.11.6.9.
[20]. N. Parveen, P. Chakrabarti, B. T. Hung, and A. Shaik, "Twitter sentiment analysis using hybrid gated attention recurrent network," J. Big Data, 2023, doi: 10.1186/s40537-023-00726-3.
[21]. S. K. R. Anumandla, V. K. Yarlagadda, S. C. R. Vennapusa, and K. R. V Kothapalli, "Unveiling the Influence of Artificial Intelligence on Resource Management and Sustainable Development: A Comprehensive Investigation," Technol. \& Manag. Rev., vol. 5, no. 1, pp. 45–65, 2020.
[22]. Abhishek Goyal, "Driving Continuous Improvement in Engineering Projects with AI-Enhanced Agile Testing and Machine Learning," Int. J. Adv. Res. Sci. Commun. Technol., vol. 3, no. 3, pp. 1320–1331, Jul. 2023, doi: 10.48175/IJARSCT-14000T.
[23]. H. Sinha, "ANALYZING MOVIE REVIEW SENTIMENTS ADVANCED MACHINE LEARNING AND NATURAL LANGUAGE PROCESSING METHODS," Int. Res. J. Mod. Eng. Technol. Sci. (, vol. 06, no. 08, pp. 1326–1337, 2024, doi: https://www.doi.org/10.56726/IRJMETS60895.
[24]. R. Tandon, "An Analysis Of COVID-19 Tweets Sentiments Based On Large Language Models (Llms)," Int. J. Res. Anal. Rev., vol. 11, no. 3, 2024.
[25]. N. Yadav, O. Kudale, A. Rao, S. Gupta, and A. Shitole, "Twitter Sentiment Analysis Using Supervised Machine Learning," in Lecture Notes on Data Engineering and Communications Technologies, 2021. doi: 10.1007/978-981-15-9509-7_51.
[26]. A. and P. Khare, "Cloud Security Challenges : Implementing Best Practices for Secure SaaS Application Development," Int. J. Curr. Eng. Technol., vol. 11, no. 6, pp. 669–676, 2021, doi: https://doi.org/10.14741/ijcet/v.11.6.11.
[27]. R. Tandon, "Social Media Texts Classification and Analysis Based on Large Language Models (LLMs)," J. Emerg. Technol. Innov. Res., vol. 11, no. 9, pp. a751–a757, 2024.
[28]. B. Ihnaini, W. Sun, Y. Cai, Z. Xu, and R. Sangi, "Sentiment Analysis of Song Dynasty Classical Poetry Using Fine-Tuned Large Language Models: A Study with LLMs," in 2024 7th International Conference on Artificial Intelligence and Big Data (ICAIBD), 2024, pp. 590–597. doi: 10.1109/ICAIBD62003.2024.10604440.
[29]. H. A. Patrick, J. Praveen Gujjar, M. H. Sharief, and U. Mukherjee, "Sentiment Analysis Perspective using Supervised Machine Learning Method," in 2023 5th International Conference on Electrical, Computer and Communication Technologies, ICECCT 2023, 2023. doi: 10.1109/ICECCT56650.2023.10179807.
[30]. R. D. Gore, V. Y. Borole, R. S. Gupta, V. Sonaje, P. Chaudhari, and B. W. Gawali, "Sentiments Analysis on Amazon Product Reviews Using Supervised Machine Learning Algorithms," in 2023 International Conference on Integration of Computational Intelligent System, ICICIS 2023, 2023. doi: 10.1109/ICICIS56802.2023.10430298.
[31]. B. Waspodo, Q. Aini, F. R. Singgih, R. H. Kusumaningtyas, and E. Fetrina, "Support Vector Machine and Lexicon based Sentiment Analysis on Kartu Prakerja (Indonesia Pre-Employment Cards Government Initiatives)," in 2022 10th International Conference on Cyber and IT Service Management, CITSM 2022, 2022. doi: 10.1109/CITSM56380.2022.9935990.
[32]. E. Asmawati, A. Saikhu, and D. Siahaan, "Sentiment Analysis of Text Memes: A Comparison among Supervised Machine Learning Methods," in International Conference on Electrical Engineering, Computer Science and Informatics (EECSI), 2022. doi: 10.23919/EECSI56542.2022.9946506.
[33]. J. C. Gope, T. Tabassum, M. M. Mabrur, K. Yu, and M. Arifuzzaman, "Sentiment Analysis of Amazon Product Reviews Using Machine Learning and Deep Learning Models," in 2022 International Conference on Advancement in Electrical and Electronic Engineering, ICAEEE 2022, 2022. doi: 10.1109/ICAEEE54957.2022.9836420.
[34]. S. Al-Hagree and G. Al-Gaphari, "Arabic Sentiment Analysis Based Machine Learning for Measuring User Satisfaction with Banking Services' Mobile Applications : Comparative Study," in 2022 2nd International Conference on Emerging Smart Technologies and Applications, eSmarTA 2022, 2022. doi: 10.1109/eSmarTA56775.2022.9935486.
[35]. P. K. Vishwakarma and M. R. Suyambu, "An Analysis of Engineering , Procurement And Construction (EPC) -Contracts Based on Renewable Energy," IJSART, vol. 10, no. 10, 2024.
[36]. R. Goyal, "EXPLORING THE PERFORMANCE OF MACHINE LEARNING MODELS FOR CLASSIFICATION AND IDENTIFICATION OF FRAUDULENT INSURANCE CLAIMS," Int. J. Core Eng. Manag., vol. 7, no. 10, 2024.
[37]. P. Khare and S. Srivastava, "AI-Powered Fraud Prevention: A Comprehensive Analysis of Machine Learning Applications in Online Transactions," J. Emerg. Technol. Innov. Res., vol. 10, pp. f518–f525, 2023.
[38]. R. K. Arora, Mohd.Muqeem, and M. Saxena, "Developing a Comprehensive Security Framework for Detecting and Mitigating IoT device Attack," Sep. 30, 2024. doi: 10.21203/rs.3.rs-5165811/v1.
[39]. N. G. Abhinav Parashar, "Asset Master Data Management: Ensuring Accuracy and Consistency in Industrial Operations," IJNRD - Int. J. Nov. Res. Dev., vol. 9, no. 9, pp. 861-a867, 2024.
[40]. S. Bauskar, "Enhancing System Observability with Machine Learning Techniques for Anomaly Detection," Int. J. Manag. IT Eng., vol. 14, no. 10, pp. 64–70, 2024.

ISSN No:-2456-2165 https://doi.org/10.5281/zenodo.14575886

[41]. R. Arora, S. Gera, and M. Saxena, "Mitigating security risks on privacy of sensitive data used in cloud-based ERP applications," in Proceedings of the 2021 8th International Conference on Computing for Sustainable Global Development, INDIACom 2021, 2021. doi: 10.1109/INDIACom51348.2021.00081.

[42]. S. Bauskar, "BUSINESS ANALYTICS IN ENTERPRISE SYSTEM BASED ON APPLICATION OF ARTIFICIAL INTELLIGENCE," Int. Res. J. Mod. Eng. Technol. Sci., vol. 04, no. 01, pp. 1861–1870, 2022, doi: DOI : https://www.doi.org/10.56726/IRJMETS18127.

[43]. V. V. Kumar, A. Sahoo, and F. W. Liou, "Cyber-enabled Product Lifecycle Management: A Multi-agent Framework," Procedia Manuf., vol. 39, pp. 123–131, 2019, doi: 10.1016/j.promfg.2020.01.247.

[44]. Rajesh Goyal, "THE ROLE OF BUSINESS ANALYSTS IN INFORMATION MANAGEMENT PROJECTS," Int. J. Core Eng. Manag., vol. 6, no. 9, 2020.

[45]. N. G. Singh, Abhinav Parashar A, "Streamlining Purchase Requisitions and Orders : A Guide to Effective Goods Receipt Management," J. Emerg. Technol. Innov. Res., vol. 8, no. 5, 2021.

[46]. V. K. Yarlagadda, "Cutting-edge developments in Robotics for Smart Warehousing and Logistics Optimization," Robot. Xplore USA Tech Dig., vol. 1, no. 1, pp. 61–79, 2024.

[47]. A. P. A. Singh, "Best Practices for Creating and Maintaining Material Master Data in Industrial Systems," Int. J. Res. Anal. Rev., vol. 10, no. 1, pp. 112–119, 2023.

[48]. S. Pandya, "Predictive Analytics in Smart Grids : Leveraging Machine Learning for Renewable Energy Sources," Int. J. Curr. Eng. Technol., vol. 11, no. 6, pp. 677–683, 2021, doi: https://doi.org/10.14741/ijcet/v.11.6.12.

[49]. P. K. and S. Arora, "the Impact of Machine Learning and Ai on Enhancing Risk-Based Identity Verification Processes," Int. Res. J. Mod. Eng. Technol. Sci., vol. 06, no. 05, pp. 8246–8255, 2024, doi: 10.56726/IRJMETS57063.

[50]. B. Boddu, "Serverless Databases Are the Future of Database Management," https://jsaer.com/download/vol-6-iss-1-2019/JSAER2019-6-1-277-282.pdf, vol. 6, no. 1, p. 5, 2020.

[51]. A. P. A. Singh, "STRATEGIC APPROACHES TO MATERIALS DATA COLLECTION AND INVENTORY MANAGEMENT," Int. J. Bus. Quant. Econ. Appl. Manag. Res., vol. 7, no. 5, 2022.

[52]. P. Khare, S. Arora, and S. Gupta, "Integration of Artificial Intelligence (AI) and Machine Learning (ML) into Product Roadmap Planning," in 2024 First International Conference on Electronics, Communication and Signal Processing (ICECSP), IEEE, Aug. 2024, pp. 1–6. doi: 10.1109/ICECSP61809.2024.10698502.

[53]. Muthuvel Raj Suyambu and Pawan Kumar Vishwakarma, "Improving grid reliability with grid-scale Battery Energy Storage Systems (BESS)," Int. J. Sci. Res. Arch., vol. 13, no. 1, pp. 776–789, Sep. 2024, doi: 10.30574/ijsra.2024.13.1.1694.

[54]. P. Rajapaksha, R. Farahbakhsh, and N. Crespi, "BERT, XLNet or RoBERTa: The Best Transfer Learning Model to Detect Clickbaits," IEEE Access, 2021, doi: 10.1109/ACCESS.2021.3128742.

[55]. R. Tandon, "Face mask detection model based on deep CNN techniques using AWS," Int. J. Eng. Res. Appl., vol. 13, no. 5, pp. 12–19, 2023, doi: 10.9790/9622-13051219.

[56]. N. Abid, "Enhanced IoT Network Security with Machine Learning Techniques for Anomaly Detection and Classification," Int. J. Curr. Eng. Technol., vol. 13, no. 6, pp. 536–544, 2023, doi: https://doi.org/10.14741/ijcet/v.13.6.5.

[57]. S. Arora and A. Tewari, "AI-Driven Resilience : Enhancing Critical Infrastructure with Edge Computing," Int. J. Curr. Eng. Technol., vol. 12, no. 2, pp. 151–157, 2022, doi: https://doi.org/10.14741/ijcet/v.12.2.9.

[58]. M. Z. Hasan, R. Fink, M. R. Suyambu, M. K. Baskaran, D. James, and J. Gamboa, "Performance evaluation of energy efficient intelligent elevator controllers," in IEEE International Conference on Electro Information Technology, 2015. doi: 10.1109/EIT.2015.7293320.

[59]. R. Goyal, "THE ROLE OF REQUIREMENT GATHERING IN AGILE SOFTWARE DEVELOPMENT: STRATEGIES FOR SUCCESS AND CHALLENGES," Int. J. Core Eng. Manag., vol. 6, no. 12, pp. 142–152, 2021.

[60]. Prity Choudhary, Rahul Choudhary and S. Garaga, "Enhancing Training by Incorporating ChatGPT in Learning Modules: An Exploration of Benefits, Challenges, and Best Practices," Int. J. Innov. Sci. Res. Technol., vol. 9, no. 11, 2024.

[61]. K. Ullah et al., "Short-Term Load Forecasting: A Comprehensive Review and Simulation Study With CNN-LSTM Hybrids Approach," IEEE Access, vol. 12, pp. 111858–111881, 2024, doi: 10.1109/ACCESS.2024.3440631.

[62]. R. Tandon, "Prediction of Stock Market Trends Based on Large Language Models," J. Emerg. Technol. Innov. Res., vol. 11, no. 9, pp. a615–a622, 2024.

[63]. M. H. Hashmi, M. Affan, and R. Tandon, "A Customize Battery Management Approach for Satellite," in Proceedings of the 2023 24th International Carpathian Control Conference, ICCC 2023, 2023. doi: 10.1109/ICCC57093.2023.10178893.

[64]. Ramesh Bishukarma, "Privacy-preserving based encryption techniques for securing data in cloud computing environments," Int. J. Sci. Res. Arch., vol. 9, no. 2, pp. 1014–1025, Aug. 2023, doi: 10.30574/ijsra.2023.9.2.0441.

[65]. Sahil Arora and Apoorva Tewari, “Fortifying Critical Infrastructures: Secure Data Management with Edge Computing,” Int. J. Adv. Res. Sci. Commun. Technol., vol. 3, no. 2, pp. 946–955, Aug. 2023, doi: 10.48175/IJARSCT-12743E.

[66]. S. Bauskar, “ADVANCED ENCRYPTION TECHNIQUES FOR ENHANCING DATA SECURITY IN CLOUD COMPUTING ENVIRONMENT,” Int. Res. J. Mod. Eng. Technol. Sci., vol. 05, no. 10, 2023, doi: 10.56726/IRJMETS45283.

[67]. H. Sinha, “Benchmarking Predictive Performance Of Machine Learning Approaches For Accurate Prediction Of Boston House Prices : An In-Depth Analysis,” iternational J. Res. Anal. Rev., vol. 11, no. 3, 2024.

[68]. M. R. S. and P. K. Vishwakarma, “THE ASSESSMENTS OF FINANCIAL RISK BASED ON RENEWABLE ENERGY INDUSTRY,” Int. Res. J. Mod. Eng. Technol. Sci., vol. 06, no. 09, pp. 758–770, 2024.

[69]. H. Sinha, “Predicting Employee Performance in Business Environments Using Effective Machine Learning Models,” IJNRD - Int. J. Nov. Res. Dev., vol. 9, no. 9, pp. a875–a881, 2024.

[70]. R. Bishukarma, “Optimising Cloud Security in Multi-Cloud Environments : A Study of Best Practices,” Int. Res. J., vol. 11, no. 11, pp. 590–598, 2024.

[71]. H. Cam, A. V. Cam, U. Demirel, and S. Ahmed, “Sentiment analysis of financial Twitter posts on Twitter with the machine learning classifiers,” Heliyon, 2024, doi: 10.1016/j.heliyon.2023.e23784.

[72]. T. Zahra and I. Hussain, “Sentiment Analysis of Twitter Dataset Using Lle and Classification Methods,” no. 01, pp. 1151–1164, 2021.

[73]. K. Deepa, H. Sangita, and H. Shruthi, “Sentiment Analysis of Twitter Data Using Machine Learning,” Lect. Notes Electr. Eng., vol. 905, no. June, pp. 259–267, 2022, doi: 10.1007/978-981-19-2177-3_26.