



A Comparative Analysis of Machine Learning and Deep Learning Models for Automated Fake News Detection

Student Name: Muhammad Haroon

Student ID: 2434025

Programme: MSc Computing

Module: ACFA7004 Research Methods and Professional Development

Supervisor: Dr Seena Joseph

Word Count: 3650

Date: Nov 10, 2025



Table of Contents

Abstract.....	2
1. Introduction	2
1.1. Research Problem.....	2
1.2. Motivation / Need for this Research	3
1.3. Shortcomings in Literature	3
2. Objectives.....	3
2.1. Objectives of The Study.....	3
2.2. What is to be Achieved.....	4
3. Importance of Research	4
4. Literature Review	4
4.1. Traditional Machine Learning Approaches: The Foundation	4
4.2. The Paradigm Shift to Deep Learning.....	4
4.3. Increasing Problems: Evaluation and The Benchmarking Gap	5
5. Hypothesis / Research Question	6
5.1. The hypothesis / research question can be answered analytically, logically, or statistically.....	6
5.2. The hypothesis / research question aligns with the objectives and the research problems.....	6
6. Research Design and Methodology	7
7. Research Limitations	8
8. Timeline.....	9
9. Expected Findings.....	9
10. References	10

Abstract

The challenge of fake news is a serious concern of social media and its impact on society. There has been research of using ML and Deep learning models dealing with automated detection of fake news. However, differentiating and analyzing the performance of these models on a common dataset has been absent. This research will involve systematic and empirical work in designing, deploying and assessing a fully automated fake news classification system built on a predefined ML and DL pipeline. The research will focus on publicly available datasets, perform exploratory data analysis, data cleansing, and formulate at least 4 machine learning models (Logistic Regression, Support Vector Machine (SVM), Random Forest, etc.) and 2 deep learning models (LSTM, CNN). The models will be exhaustively assessed and contrasted based on a set of evaluation metrics Accuracy, Precision, Recall, F1-Score. The intended outcome is a robust benchmark that will present the best performing model architectures along with the applicable trade-offs on the practical accuracy, resource consumption, and interpretability of the model.

1. Introduction

1.1. Research Problem

The rise of new technologies has created a thriving economy that has caught the interest of computational researchers who now see a multitude of frameworks that have been proposed to combat the phenomenon of fake news (Harris *et al.*, 2024). While separate divisions of current Machine Learning and Deep Learning paradigms have been gaining traction, the domain of fake news detection has still not received a distilled and rigorously constructed evaluative metric that is necessary for guide model selection (Bashaddadh *et al.*, 2025).

The main issue in the domain revolves around the practical shortcomings that stem from having no strategic model selection framework that rationalizes a real-world feasible set measure for detection optimized for detection of fake news. This issue can be elaborated further into a set of issues:

- **Disjointed and Non-comparable Assessments:** Existing narratives tend to construct a new technique and evaluate the performance of that new technique with a few selected baseline models on a closed world problem (Mishra, Shukla and Agarwal, 2022). Such a methodology creates a scenario where the conclusions are non-transferrable, and result in a non-informative body of work which ultimately results in the literature being more obscured than clarified (Garbacea, 2023).
- **The Lack of Evidence on the Balance of Simple Versus Complex:** The deep learning models (e.g., LSTM, CNN) are assumed to complex and data-hushed models, thereby rest, more able and efficient to deep learning than traditional ML models (e.g., SVM, Random Forest) (Ahmad *et al.*, 2018). The deep learning models are assumed to be more able rest complex, and more, efficient able LSTM, and CNN models while whose gain in ascertainable valued costly in data, computations and tethered loss in, rest, more, interpretation, outweigh is poorly justified. There is a gap in the comprehension of the complexity performance frontier on fake news detection (Bashaddadh *et al.*, 2025b).
- **The Literature Gap for Practical Application:** There is an absence of comprehensive arguments on the literature for social media practitioners, engines, and custom fact-checking extension developers concerning the questions: "When is a Random Forest model well optimized, and what are the boundary conditions for performance where an LSTM model must be used?" The literature

does not offer a like-for-like comparison. There is potential for underperformance coupled with inefficient resource utilization (Zhang *et al.*, 2023).

Models are not the issue, but the absence of a conclusive comparative framework remains a challenge. There is a gap with an absence of large consecutive empirical studies that takes a wide array of traditional and contemporary models and pitted them against one another on the same dataset, same preprocessing pipeline and the same evaluation protocol. The insights garnered should speak unambiguous and actionable on the competitive stance, operational moderation, and robustness. This study attempts to remedy this gap.

1.2. Motivation / Need for this Research

The blurring of the truth and the rapid dissemination of fake news erodes public trust in institutions, alters public perception, and even poses risks to the social order. The detection of fake news using automated techniques is imperative in order to scale the identification of nefarious content (Sharma *et al.*, 2019). Nonetheless, developers and users of the applications face a model selection problem that is a trade-off between model performance and explainability, availability of computational resources (Mori and Uchihiro, 2019). This problem drives the research with motivation to model selection for a real-world problem using data-driven approaches.

1.3. Shortcomings in Literature

Most of the works in this field focus on either proposing intricate, novel models, or exploring a narrow selection of machine learning models. Very few works compare, and critique tuned classical ML models with contemporary DL models, thus making the benefits of computationally expensive approaches unclear. This gap does not serve the interests of the practitioners (Barbierato and Gatti, 2024).

2. Objectives

- Conduct an Exploratory Data Analysis on the selected Fake News Dataset by studying its notable characteristics including the class distribution, text length, and prevalent keywords.
- Preprocess the text data following the defined steps in the process flow which includes tokenization, stop-word filtering, and vectorization (TF-IDF and word embeddings for ML and Deep Learning models respectively).
- Train and evaluate four models (Logistic Regression, Naïve Bayes, Support Vector Machine, Random Forest) and document the baseline performance metrics for all.
- Train and evaluate two Deep Learning models (an LSTM and a CNN) for sequence classification.
- Assess the models using a separate test set and document performance using Accuracy, Precision, Recall, F1-Score, and Confusion Matrix.
- Document the findings in a performance report structured for model evaluation and comparison discussing the trade-offs for each model outcome.

2.1. Objectives of The Study

Moving beyond theoretical comparisons and actionable evidence on the model's efficacy, this study aims to set a defined benchmark, deeply rooted in evidence, for the accuracy of the detection models for fake news.

2.2. What is to be Achieved

This research is expected to deliver an ordered performance evaluation of several Machine Learning and Deep Learning models, a collection of validated and thoroughly documented code for the model of ease of use, and model-agnostic prescriptive recommendations based on the target task to be accomplished (e.g., detection of fake news and the operational trade-off to be made, tactical precision as opposed to the provided computational capability).

3. Importance of Research

- **Scientific Value:** This research adds a thorough benchmark study to the field of NLP and computational social science. It furthers understanding of the strengths of various algorithmic paradigms in text classification and, in particular, as a counterpoint to black box, complex DL models, delineates the circumstances under which one ought to deploy less complex, more transparent ML models (Allam *et al.*, 2025).
- **Practical Value:** These findings can certainly aid technology companies and social media platforms in their efforts to design and deploy content moderation systems. In particular, the findings can help design fact-checking browser extensions for greater accuracy and efficiency and aid AI teams in understanding the performance-cost ratios in different models to better target their efforts (Ridwan, 2025).

4. Literature Review

Automatic fake news detection literature contains many gaps. This review identifies sandwiched models of feature-engineered to deep learning neural networks to the present-day issues related to evaluation and generalization. This understanding is important for the context of this research.

4.1. Traditional Machine Learning Approaches: The Foundation

Pioneer works in fake news detection concentrated on Machine Learning (ML) models with custom-built features. The main focus was to uncover linguistic and stylistic markers to differentiate fake news from authentic content.

- An example of such a pioneer is (Onan, 2018) and the study on linguistic features. He used SVM and Logistic Regression. He increased the model competitiveness by demonstrating that a set of n-grams and punctuation psycholinguistic and LIWC syntactic features can work in a model.
- (Pattun and Kumar, 2025) in their broad survey explained the problem's complexity by discussing not only linguistic aspects but also the social context features. They, however, in the survey emphasized the reliance on manual feature engineering as the major bottleneck of the approach as it needed domain knowledge and could miss more sophisticated and hidden patterns in the text.

4.2. The Paradigm Shift to Deep Learning

The arrival of Deep Learning (DL) saw the removal of manual feature engineering as the major bottleneck and replaced it with models that were able to learn features in a hierarchy and in layers, even from unprocessed and raw text.

- **Convolutional Neural Networks (CNN)** She is one of the first to note that CNN can be used in the classification of sentences and, as in other works, uses the successful application with vision as a starting point (Kim, 2014). And in the case of the classification of fake news, CNN are able to locate

and identify the very informative and critical phrases within a text, e.g. the sensationalist or the inconsistent word combinations, regardless of the distance between words or their position within the text.

- Recurrent Neural Networks (RNN) and LSTMs: RNNs focus on sequential information, contrary to CNNs. For rumor detection on microblogs, (Mienye, Swart and Obaido, 2024) applied LSTM, a subset of RNN. LSTMs are good at modeling contextual flow and long-range dependencies of language, such as the narrative structure of a news article. Subsequently, (Liu *et al.*, 2016) introduced a model that fused CNNs with RNNs to exploit local and global temporal contextual features.

4.3. Increasing Problems: Evaluation and The Benchmarking Gap

The rise in the number of proposed models led to an important problem of how to evaluate models in a comprehensive and systematic manner. A large portion of the literature concentrates on proposing new model(s) and benchmarking against a handful, often one or two, of baselines to establish dominance.

- The Incremental Improvement Question: Research focused on large-scale controlled benchmarking studies is sorely lacking and poses a critical pragmatic query. In cross a standard public benchmark, do modern DL models, with their added complexity and DL-specific computational costs, provide a consistently better advantage over a well-tuned traditional ML model (Chab, Li and Setia, 2025).

Table 1: Previous Related Work

References	Method	Aim/Objectives	Techniques/Algorithms	Limitation	Programming Languages/Tools	Comments on the Method or Techniques
(Gurram and Nalam, 2024)	Traditional ML	To identify fake news using linguistic features and style markers.	Logistic Regression, SVM with features like n-grams, punctuation, syntax.	Relies entirely on manual feature engineering; may not capture deep semantic patterns.	Python (Scikit-learn), LIWC	Established a strong, interpretable baseline. Performance is often capped by the quality of the engineered features.
(Talele, Jangale and Vijayalakshmi, 2021)	Deep Learning	To detect rumors on Twitter by modeling the temporal dynamics of text.	LSTM (Recurrent Neural Network).	Computationally intensive; model performance can be sensitive to hyperparameter tuning.	Not specified (typically Python with Keras/TensorFlow).	Pioneering use of RNNs for sequence modeling in rumor detection. Effective for context but may be less sensitive to key phrases than CNNs.

(Ahamed, 2018)	Deep Learning	To classify sentences using a CNN model with pre-trained word vectors.	CNN with pre-trained word embeddings (e.g., Word2Vec, GloVe).	Fixed window size may miss long-range dependencies in longer texts.	Not specified (typically Python with Keras/TensorFlow).	Demonstrated the power of CNNs for extracting local features in text. Highly influential but designed for shorter text sequences.
(Abdullah <i>et al.</i> , 2025)	Hybrid Deep Learning	To capture both local features and global context for improved fake news detection.	Combined CNN and RNN (LSTM) architectures.	Increased model complexity and computational cost; requires more data to train effectively.	Python, Keras/TensorFlow	A sophisticated approach that aims to get the "best of both worlds," but the added complexity may not always justify the performance gain for all datasets.
(Hu, Mao and Zhang, 2025)	Data Mining & Survey	To provide a holistic overview of fake news detection, including data mining perspectives.	N/A (Survey paper).	Does not provide empirical comparisons between a wide range of models.	N/A	A highly valuable resource that outlines the landscape and explicitly identifies the need for more robust, data-driven solutions and benchmarks.

5. Hypothesis / Research Question

5.1. The hypothesis / research question can be answered analytically, logically, or statistically.

- Primary Research Question: "Among a selected set of Machine Learning and Deep Learning classifiers, which architecture obtains the best performance for fake news detection after a thorough multi-metric evaluation?"
- It will be answered through the statistical analysis of performance metrics, e.g., F1-scores, along with the logical deduction gained from F1 confusion matrix evaluation.

5.2. The hypothesis / research question aligns with the objectives and the research problems.

- This research question is exactly with the fundamental research dilemma of the absence of comprehensive systematic comparison within the presented research. It is addressed through the achievement of the Objectives 3, 4, and 5, which pertain to the implementation of the models

with their performance evaluation. The question presents a vital quality which interlinks all the experimental objectives.

6. Research Design and Methodology

The research implements provided seems to cover the objectives as stated.

The approach taken is Quantitative Research with Experimental method in a Design Science Research Framework. This is warranted as the objectives stipulate the construction of IT artifacts (trained models) and their subsequent quantification (performance metrics) in a controlled setting.

The objective delineated procedural steps have been rationalized above.

These procedural steps are in a logical sequence.

- Objective 1 involves Data Acquisition and EDA (Exploratory Data Analysis) covering the preliminary analysis.
- For Objective 2 the Data is Preprocessed.
- For Objective 3 the models are trained for the constructed ML models.
- Objective 4 involves the training of DL models.
- For Objective 5 the models are Evaluated and the pertinent Performance Metrics are calculated.
- For Objective 6 the models are subjected to a Comparative Analysis and Synthesis.

The set of selected techniques is deemed to be Highly Appropriate for the problem at hand.

- For ML models TF-IDF Vectorization is used as it draws text into a numerical feature set.
- For DL models Word Embedding in particular GloVe is the default and proper input representation used in NLP.
- The metrics Performance Metrics in particular Precision, Recall and F1-Score are the metrics used in classification problems for instance when there is class imbalance and a tradeoff between false positives and false negatives is warranted.
- The evaluation process is considered realistic in nature and it is aimed at measuring model generalizability in context of unseen data. This is done through the process of Train Test Split.

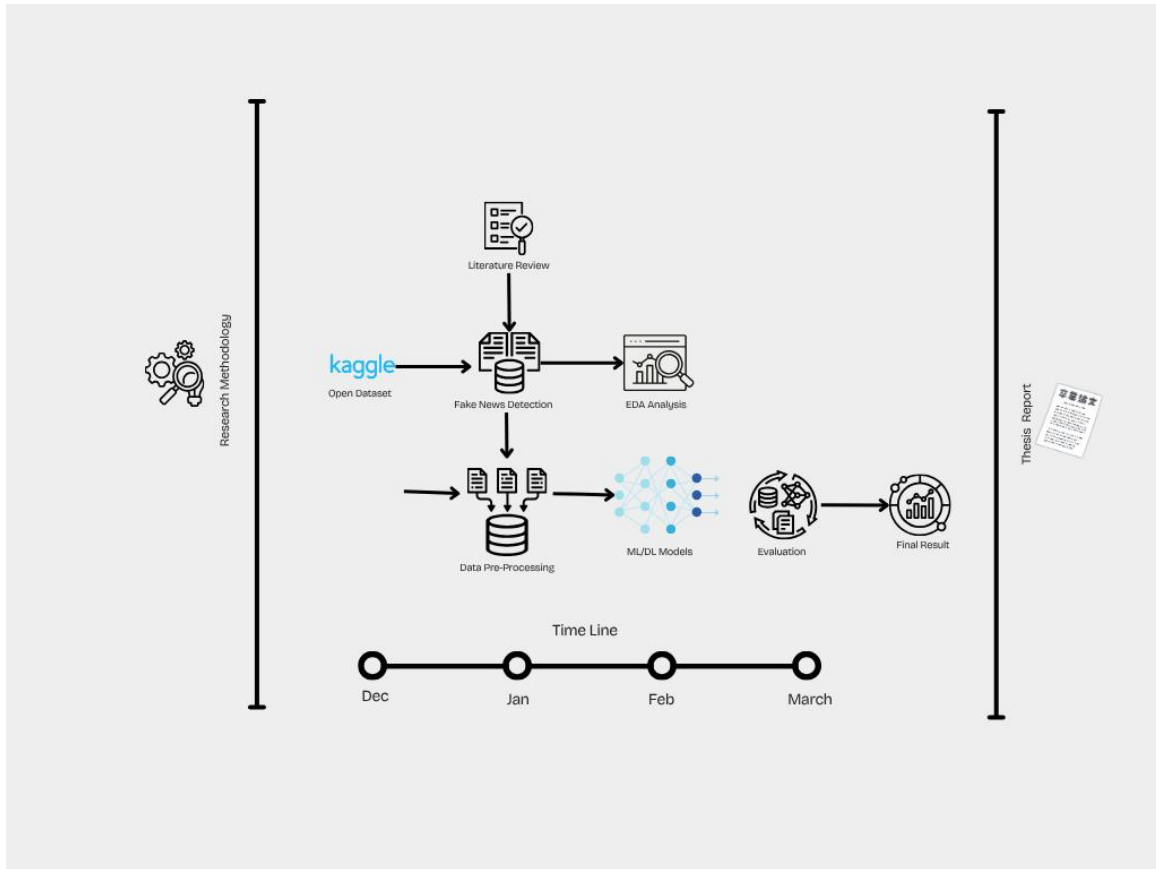


Figure 1: Research Methodology

7. Research Limitations

The effect of design constraints on the research has been stated.

- Limitation: The study is oversimplified to a single dataset based on a single modality of data. It excludes combined data (for example, images) and relevant social context features (for example, networks of propagation, user data) (Lahat, Adali and Jutten, 2015).
- Impact: This does not enhance the ability to extrapolate the study's conclusion. The model which has the highest efficacy on this textual dataset is unlikely to perform better on datasets, which are primarily visual, or on datasets which follow certain propagation structures. The conclusions are relevant primarily to systems of text-focused detection (Oluwagbemi, Mesioye and Akinbo, 2025).

The effect of design constraints on the research has been stated.

- Limitation: The research approach approaches the dataset in question as a static, immutable system and thus can, at best, superficially grasp the dynamic, adversarial, and ever-changing system of fake news generation and propagation.
- Impact: The findings demonstrate performance at a particular snapshot in time. The models are significantly vulnerable to concept drift, which is the new fake news styles, not captured in the training datasets, which can emerge, therefore, making the models ineffective on a long-term scale, if not retrained continuously.

8. Timeline

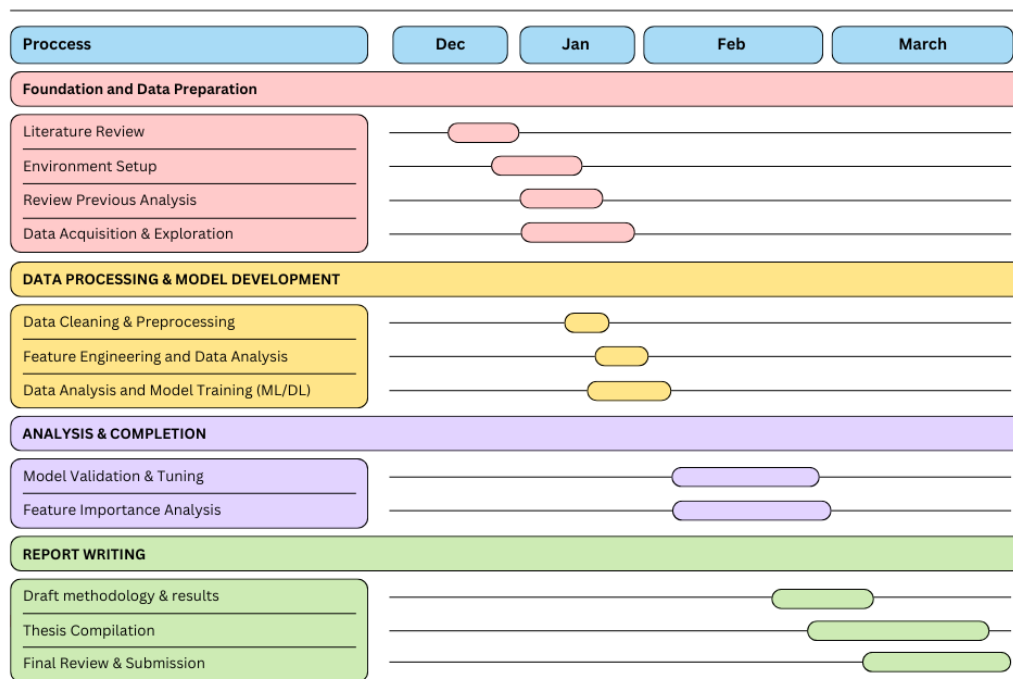


Figure 2: Timeline and Deliverables

9. Expected Findings

It is possible that the outcomes from this research will mirror the primary outcomes from the rigorously defined objectives of undertaking a systematic benchmark and comparative analysis. The research is expected to reveal a number of outcomes that would add to the body of knowledge as well as the practice of automated fake news detection.

From the previous works done on the application of machine learning on text classification, it is reasonable to expect that Deep Learning models would outperform other Machine Learning models on the relevant metrics of Accuracy, Support, Precision and F1-Score. Such a conclusion is justified by the capacity of deep learning models to capture the requisite hierarchical structures and nuances of high dimensional problems as well as to manage very large raw text databanks where diverse and fine nuances are intricately intertwined. Such a conclusion would certainly justify the claims made regarding deep learning models and the features associated with high-stakes classification of data. It would also satisfy a part of the comparison goals of this research.

In addition, there is expected to be a rank order of performance within the class of traditional Machine Learning models. It is expected that base estimators which employ ensemble techniques will outperform simpler estimators which rely on linear as well as probabilistic models.

10. References

- Abdullah, A.A. *et al.* (2025) "A Hybrid Approach for Fake News Detection on Cloud Computing," *International Journal of Theoretical and Applied Research*, 0(0), pp. 600–608. Available at: <https://doi.org/10.21608/ijtar.2025.402732.1136>.
- Ahamed, S. (2018) *Text Classification Using CNN, LSTM and Pre-trained Glove Word Embeddings: Part-3* | by Sabber Ahamed | Medium. Available at: <https://sabber.medium.com/classifying-yelp-review-comments-using-cnn-lstm-and-pre-trained-glove-word-embeddings-part-3-53fcea9a17fa> (Accessed: November 6, 2025).
- Ahmad, I. *et al.* (2018) "Performance comparison of support vector machine, random forest, and extreme learning machine for intrusion detection," *IEEE access*, 6, pp. 33789–33795.
- Allam, H. *et al.* (2025) "Text classification: How machine learning is revolutionizing text categorization," *Information*, 16(2), p. 130.
- Barbierato, E. and Gatti, A. (2024) "The challenges of machine learning: A critical review," *Electronics*, 13(2), p. 416.
- Bashaddadh, O. *et al.* (2025a) "Machine Learning and Deep Learning Approaches for Fake News Detection: A Systematic Review of Techniques, Challenges, and Advancements," *IEEE Access* [Preprint]. Available at: <https://ieeexplore.ieee.org/abstract/document/11008608/> (Accessed: November 6, 2025).
- Bashaddadh, O. *et al.* (2025b) "Machine Learning and Deep Learning Approaches for Fake News Detection: A Systematic Review of Techniques, Challenges, and Advancements," *IEEE Access* [Preprint]. Available at: <https://ieeexplore.ieee.org/abstract/document/11008608/> (Accessed: November 6, 2025).
- Chab, R., Li, F. and Setia, S. (2025) "Algorithmic Techniques for GPU Scheduling: A Comprehensive Survey," *Algorithms*, 18(7), p. 385.
- Garbacea, G.-C. (2023) *Neural language generation for content adaptation: Explainable, efficient low-resource text simplification and evaluation*. PhD Thesis. Available at: <https://deepblue.lib.umich.edu/handle/2027.42/178028> (Accessed: November 6, 2025).
- Gurram, V.V.V.R. and Nalam, J.M. (2024) "Automated Detection of Fake News in Natural Language Processing: A Comparative Study of TF-IDF and Lexical-Based Stance Detection with Logistic Regression." Available at: <https://www.diva-portal.org/smash/record.jsf?pid=diva2:1883087> (Accessed: November 6, 2025).
- Harris, S. *et al.* (2024) "Fake news detection revisited: An extensive review of theoretical frameworks, dataset assessments, model constraints, and forward-looking research agendas," *Technologies*, 12(11), p. 222.
- Hu, B., Mao, Z. and Zhang, Y. (2025) "An overview of fake news detection: From a new perspective," *Fundamental Research*, 5(1), pp. 332–346. Available at: <https://doi.org/10.1016/j.fmre.2024.01.017>.
- Kim, Y. (2014) "Convolutional Neural Networks for Sentence Classification." arXiv. Available at: <https://doi.org/10.48550/arXiv.1408.5882>.

- Lahat, D., Adali, T. and Jutten, C. (2015) "Multimodal data fusion: an overview of methods, challenges, and prospects," *Proceedings of the IEEE*, 103(9), pp. 1449–1477.
- Liu, Q. *et al.* (2016) "Predicting the next location: A recurrent model with spatial and temporal contexts," in *Proceedings of the AAAI conference on artificial intelligence*. Available at: <https://ojs.aaai.org/index.php/AAAI/article/view/9971> (Accessed: November 6, 2025).
- Mienye, I.D., Swart, T.G. and Obaido, G. (2024) "Recurrent neural networks: A comprehensive review of architectures, variants, and applications," *Information*, 15(9), p. 517.
- Mishra, S., Shukla, P. and Agarwal, R. (2022) "Analyzing Machine Learning Enabled Fake News Detection Techniques for Diversified Datasets," *Wireless Communications and Mobile Computing*. Edited by M.F. Hashmi, 2022(1), p. 1575365. Available at: <https://doi.org/10.1155/2022/1575365>.
- Mori, T. and Uchihira, N. (2019) "Balancing the trade-off between accuracy and interpretability in software defect prediction," *Empirical Software Engineering*, 24(2), pp. 779–825. Available at: <https://doi.org/10.1007/s10664-018-9638-1>.
- Oluwagbemi, J.B., Mesioye, A.E. and Akinbo, R.S. (2025) "Depress-HybridNet: A Linguistic-Behavioral Hybrid Framework for Early and Accurate Depression Detection on Social Media," *Journal of Future Artificial Intelligence and Technologies*, 2(3), pp. 432–444.
- Onan, A. (2018) "Sentiment analysis on Twitter based on ensemble of psychological and linguistic feature sets," *Balkan Journal of Electrical and Computer Engineering*, 6(2), pp. 69–77.
- Pattun, G. and Kumar, P. (2025) "Feature Engineering Trends in Text-Based Affective Computing: Rules to Advance Deep Learning Models," *International Research Journal of Multidisciplinary Technovation*, 7(2), pp. 87–107.
- Ridwan, I.B. (2025) "Optimizing enterprise decision-making through causal machine learning models and real-time business intelligence integration," *Int J Adv Res Publ Rev*, 2(5), pp. 67–88.
- Sharma, K. *et al.* (2019) "Combating Fake News: A Survey on Identification and Mitigation Techniques," *ACM Transactions on Intelligent Systems and Technology*, 10(3), pp. 1–42. Available at: <https://doi.org/10.1145/3305260>.
- Talele, M., Jangale, S. and Vijayalakshmi, M. (2021) "Rumor Detection Using Various Deep Learning Approaches."
- Zhang, Y. *et al.* (2023) "Comparison of random forest, support vector regression, and long short term memory for performance prediction and optimization of a cryogenic organic rankine cycle (ORC)," *Energy*, 280, p. 128146.