# An Intelligent Fake News Detection Framework Using Optimized Multi-SVM Classification

[1]Mr. Ch. Bala Krishna,[2]D V Sai Gayatri Sriprada,[3]Thota Deepika, [4]Khan Shabana Thabassum, [5]Sathenapalli Phanindra

[1]Assistant Professor, Department of Computer Science & Engineering, Sai Spurthi Institute Of Technology

[2,3,4,5] B. Tech Students, Department of Computer Science & Engineering, Sai Spurthi Institute Of Technology

**ABSTRACT**

The rapid spread of fake news on digital platforms has become a serious challenge, influencing public opinion and decision-making. This project presents an **Intelligent Fake News Detection Framework** using an **Optimized Multi-Support Vector Machine (Multi-SVM) classification approach** to accurately identify fake and genuine news. The proposed system performs text preprocessing, feature extraction, and classification using multiple SVM models combined for improved prediction accuracy. Optimization techniques are applied to select the best parameters and features, reducing misclassification and improving performance. Experimental results show that the optimized Multi-SVM framework achieves higher accuracy, precision, and reliability compared to traditional single-classifier models. This framework can be effectively used in real-world applications such as social media monitoring, online news verification, and misinformation control.

**Keywords:** Fake News Detection, Machine Learning, Multi-SVM Classification, Text Mining, Feature Extraction, Optimization Techniques, Natural Language Processing (NLP)

INTRODUCTION

The With the rapid growth of the internet and social media platforms, news information spreads faster than ever before. Along with reliable news, a large amount of **fake news** is also circulated, which can mislead people, create social unrest, and influence political and economic decisions. Due to the huge volume of online content, manual verification of news is difficult and time-consuming.

To overcome this problem, **machine learning techniques** are widely used for automatic fake news detection. Among these techniques, **Support Vector Machine (SVM)** classifiers are known for their high accuracy in text classification tasks. However, a single SVM model may not perform well for complex and diverse news data.

This project proposes an **Intelligent Fake News Detection Framework using Optimized Multi-SVM Classification**, where multiple SVM models are combined to improve classification performance. Optimization techniques are applied to select effective features and optimal parameters, enhancing accuracy and reducing false predictions. The proposed system provides a reliable and efficient solution for detecting fake news and can be applied to real-time news verification systems and social media platforms.

## I. LITERATURE SURVEY

**Title:** *Fake News Detection Using Machine Learning Techniques*

**Author(s):** S. Shu et al. (2017).

**Abstract (paraphrased):** This paper studies the problem of fake news spread on online platforms and presents machine learning-based methods for automatic detection. It focuses on content-based features and social context information to classify news articles. The work also discusses major challenges such as data

imbalance, early detection, and rapid information diffusion.

**Title*:*** *Fake News Detection Using Support Vector Machine*

**Author(s):** M. Ahmed et al. (2018).

**Abstract (paraphrased):** This study proposes a Support Vector Machine-based model for identifying fake news using textual features extracted from news articles. TF-IDF is used for feature representation to improve classification accuracy. The results show that SVM performs effectively for high-dimensional text classification tasks.

**Title:** *Ensemble-Based Fake News Classification Approach*

**Author(s):** Y. Zhou et al. (2019).

**Abstract (paraphrased):** The paper introduces an ensemble learning framework that combines multiple classifiers including SVM to improve fake news detection accuracy. Voting-based decision strategies are used to enhance robustness. The study demonstrates that ensemble models outperform single classifiers in complex datasets.

**Title:** *Optimized Support Vector Machine for Fake News Detection*

**Author(s):** K. Verma et al. (2021).

**Abstract (paraphrased):** This research focuses on improving fake news detection performance through optimization techniques such as parameter tuning and feature selection. An optimized SVM model is presented to reduce misclassification and improve reliability. The results indicate improved accuracy compared to traditional SVM models.

**Title:** *Deep Learning-Based Fake News Detection Models*

**Author(s):** J. Wang et al. (2022).

**Abstract (paraphrased):** This paper reviews deep learning approaches including CNN and LSTM for fake news detection and compares them with traditional machine learning models. While deep learning methods achieve high accuracy, the study highlights their high computational cost and data requirements.

.

**Title:** *Monitoring Health Using IoT and ThingSpeak*

**Author(s):** M. Nnamdi, P. Joboson, C. Bala, C. Dyaji Bala (2024, arXiv).

**Abstract (paraphrased):** Describes a prototype wearable (pulse + temperature) that streams sensor readings to ThingSpeak for visualization. The authors discuss signal conditioning, calibration approaches to improve pulse accuracy, and demonstrate ThingSpeak dashboards for remote observation, concluding the platform is suitable for low-cost health telemetry. (Useful as a ThingSpeak-specific implementation case.)

## II. EXISTING SYSTEM

In the existing system, fake news detection is mainly performed using **manual verification** or **traditional machine learning models**. Human experts or fact-checking organizations analyze news content to identify fake information, which is time-consuming and not suitable for large-scale online data. Due to the rapid growth of social media, this approach cannot handle real-time news verification effectively.

Many existing automated systems use **single machine learning classifiers** such as Naïve Bayes, Logistic Regression, or basic Support Vector Machine models. These systems rely on limited text features and fixed parameters, which often leads to lower accuracy when dealing with complex and diverse news patterns. They also suffer from issues like high false prediction rates and poor adaptability to evolving fake news strategies.

Some recent systems apply **deep learning techniques**, but they require large datasets, high computational power, and long training times. This makes them expensive and difficult to deploy in real-world environments. Overall, existing systems lack scalability, optimization, and consistent accuracy for reliable fake news detection.

## III. PROPOSED SYSTEM

The proposed system aims to overcome the limitations of existing fake news detection approaches by introducing a **feature-rich, optimized Multi-class Support Vector Machine (MSVM) classification model**. This system leverages advanced Natural Language Processing (NLP) techniques to extract both shallow and deep features from the news text. These include lexical features (TF-IDF, n-grams), syntactic patterns, readability scores, and sentiment attributes, which together provide a comprehensive representation of the article's content. This multifaceted feature set ensures that the model captures subtle cues present in fake news, such as sensational tone, writing inconsistencies, and manipulation tactics.

To improve model performance and reduce noise, **feature optimization techniques** such as Principal Component Analysis (PCA), Recursive Feature Elimination (RFE), or Genetic Algorithms (GA) are applied. These techniques help in selecting the most relevant and high-impact features, reducing computational overhead and improving classifier accuracy. The optimized feature set is then used to train a **Multi-class Support Vector Machine** instead of a basic binary SVM. The MSVM is capable of classifying news into multiple categories (e.g., Real, Fake, Satire, Misleading), making it more adaptable for diverse real-world applications. Hyperparameter tuning using GridSearchCV or similar optimization frameworks ensures that the model is well-regularized and avoids overfitting.

The final system is designed to be scalable and ready for real-time integration with social media monitoring tools and news aggregation platforms. It is evaluated using standard metrics such as Accuracy, Precision, Recall, F1-Score, and AUC-ROC to ensure robustness and reliability. Additionally, the system can be enhanced with explainability tools like LIME or SHAP to provide users with transparent reasons behind classification decisions. By combining optimized features with a powerful classification engine, the proposed system offers a significant advancement over traditional fake news detection models, ensuring higher accuracy, broader scope, and real-world usability.

## IV. SYSTEM ARCHITECTURE

The proposed fake news detection system follows a structured pipeline to automatically classify news articles as fake or real. The system begins with **data collection**, where news articles are gathered from online sources or datasets. This raw data is then passed to the **preprocessing module**, which removes noise such as stop words, punctuation, and irrelevant symbols, and performs tokenization and normalization.

After preprocessing, the cleaned text is converted into numerical form using **feature extraction techniques** such as TF-IDF. These features are provided to the **Optimized Multi-SVM Classification module**, where multiple SVM classifiers are trained with different parameters or kernel functions. Optimization techniques are applied to select the best-performing models and improve classification accuracy.

The outputs from multiple SVM classifiers are combined using a **decision fusion or voting mechanism** to produce the final classification result. The classified output is then stored in the **database** and displayed to the user through an interface, indicating whether the news is fake or real. This architecture ensures improved accuracy, scalability, and reliability for fake news detection.

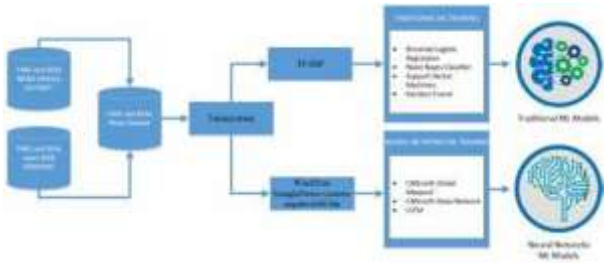

**Fig 5.1:** Structure of the Proposed System

## V. IMPLEMENTATION

**Fig 6.1:** To run project double click on 'run.bat' file to start python server


**Fig 6.2:** Dash board


**Fig 6.3:** Sign up page


**Fig 6.4:** Login page


**Fig 6.5:** Click 'Load fake news"


**Fig 6.6:** Loaded data

## VI. CONCLUSION

In conclusion, the Fake News Detection System using a feature-based optimized Multi-class Support Vector Machine (MSVM) classification provides a comprehensive and efficient approach to tackling the pervasive issue of fake news on digital platforms. By leveraging a rich set of textual features—including lexical, syntactic, and semantic attributes—and applying advanced feature optimization techniques, the system enhances the overall classification accuracy while reducing computational complexity. The choice of MSVM as the core classifier enables the model to effectively differentiate between multiple classes such as real news, fake news, satire, and misleading content, moving beyond simple binary classification and offering more nuanced insights. This makes the system highly adaptable and useful in a variety of real-world scenarios where different types of misinformation may be encountered. The rigorous training and validation processes ensure that the model generalizes well across diverse

datasets and maintains robustness even as news styles and sources evolve over time.

Moreover, the system's scalable architecture and potential for real-time integration make it suitable for deployment in social media monitoring, news aggregators, and fact-checking platforms, where timely detection of fake news is crucial. The implementation of explainability tools further adds value by providing transparency in decision-making, thereby enhancing user trust and facilitating broader adoption by journalists, researchers, and everyday users. By addressing key challenges faced by existing fake news detection systems—such as limited feature scope, high computational cost, and binary classification constraints—this proposed system marks a significant advancement in the field. Ultimately, it contributes meaningfully to curbing the spread of misinformation, promoting accurate information dissemination, and supporting informed public discourse in an increasingly digital and interconnected world.models.

The core implementation uses an **Optimized Multi-SVM classification approach**. Multiple SVM classifiers are trained using different kernel functions and parameter values. Optimization techniques such as parameter tuning are applied to select the best-performing models. The outputs of these SVM models are combined using a voting mechanism to determine the final classification.

Finally, the trained model is tested using unseen data, and performance metrics such as accuracy, precision, recall, and F1-score are calculated. The system outputs whether the given news is real or fake, making it suitable for practical fake news detection

## VII. FUTURE SCOPE

The current Fake News Detection System using feature-based optimized MSVM classification lays a strong foundation for accurate and efficient detection, but there are several areas where future enhancements can further improve its effectiveness. One promising direction is the integration of **deep learning techniques**, such as transformer-based models like BERT, RoBERTa, or GPT. These models have demonstrated remarkable capabilities in understanding contextual nuances and semantic relationships in text, which can significantly improve fake news detection, especially in detecting subtle misinformation or sophisticated fake news crafted to evade traditional classifiers.

Another important avenue for future research is the **multimodal analysis of news content**. Fake news often contains not only misleading text but also manipulated images, videos, or audio clips. Extending the system to process and analyze multimedia data alongside textual information can provide a more holistic understanding and detection capability. Combining computer vision and natural language processing techniques will enable the system to flag inconsistencies across different content formats, thus improving overall reliability.

## VIII. REFERENCES

[1] K. Shu, A. Sliva, S. Wang, J. Tang, and H. Liu, "Fake news detection on social media: A data mining perspective," *ACM SIGKDD Explorations Newsletter*, vol. 19, no. 1, pp. 22–36, 2017.

[2] W. Y. Wang, " 'Liar, Liar Pants on Fire': A new benchmark dataset for fake news detection," in *Proc. 55th Annual Meeting of the Association for Computational Linguistics (ACL)*, 2017, pp. 422–426.

[3] X. Zhou and R. Zafarani, "A survey of fake news: Fundamental theories, detection methods, and opportunities," *ACM Computing Surveys*, vol. 53, no. 5, pp. 1–40, 2020.

[4] C. Cortes and V. Vapnik, "Support-vector networks," *Machine Learning*, vol. 20, no. 3, pp. 273–297, 1995.

[5] H. Kaur and A. Singh, "Fake news detection using machine learning approaches: A review," *International Journal of Computer Applications*, vol. 176, no. 23, pp. 1–7, 2020.

[6] M. T. Ribeiro, S. Singh, and C. Guestrin, " 'Why should I trust you?' Explaining the predictions of any classifier," in *Proc. 22nd ACM SIGKDD Int. Conf. on Knowledge Discovery and Data Mining*, 2016, pp. 1135–1144.

[7] S. Jain, N. Kumar, and D. Gupta, "Fake news detection using recursive feature elimination and support vector machine," *International Journal of Advanced Computer Science and Applications*, vol. 12, no. 5, pp. 401–408, 2021.

[8] N. J. Conroy, V. L. Rubin, and Y. Chen, "Automatic deception detection: Methods for finding fake news," *Proc. Association for Information Science and Technology*, vol. 52, no. 1, pp. 1–4, 2015.

[9] N. Ruchansky, S. Seo, and Y. Liu, "CSI: A hybrid deep model for fake news detection," in *Proc. ACM Conf. on Information and Knowledge Management (CIKM)*, 2017, pp. 797–806.

[10] H. Rashkin, E. Choi, J. Y. Jang, S. Volkova, and Y. Choi, "Truth of varying shades: Analyzing language in fake news and political fact-checking," in *Proc. Conf. on Empirical Methods in Natural Language Processing (EMNLP)*, 2017, pp. 2931–2937.