JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue
JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)
An International Scholarly Open Access, Peer-reviewed, Refereed Journal

SMS Spam Detection Using Machine Learning: A Hybrid SVM and Naive Bayes Approach

[1]Rupali R. Bansode, [2]Suvarna D. Pingle
[1]M.Tech Student, [2]Professor
[1]Department of Computer Science and Engineering,
[1]P.E.S Engineering College, Nagsenvana, Chh. Sambhajinagar, India

Abstract: The rapid growth of mobile communication has led to a significant increase in unsolicited Short Message Service (SMS) spam, posing security, privacy, and economic challenges. This paper presents an efficient SMS spam detection framework using machine learning techniques, specifically Support Vector Machine (SVM), Naïve Bayes (NB), and a proposed hybrid SVM and NB model. The UCI SMS Spam Collection dataset is used for experimental evaluation. A comprehensive text preprocessing pipeline including tokenization, stop word removal, and TF-IDF feature extraction is employed. Experimental results demonstrate that the proposed hybrid model achieves an accuracy of 99.62%, outperforming individual classifiers and existing approaches reported in the literature. The study confirms that hybrid learning models combined with effective preprocessing significantly enhance spam detection performance.

Index Terms **- SMS Spam Detection, Machine Learning, SVM, Naïve Bayes, TF-IDF, Hybrid Model**

I. INTRODUCTION

With the exponential rise in mobile phone usage, SMS has become a widely used communication medium. However, this growth has also facilitated the proliferation of SMS spam messages, which often contain fraudulent offers, phishing links, or malicious content. Manual filtering is impractical due to the volume and evolving nature of spam, necessitating automated detection mechanisms.

Machine learning techniques have proven effective in text classification problems due to their ability to learn patterns from historical data. Among these, Naïve Bayes and Support Vector Machines are widely used because of their simplicity, scalability, and robustness. However, individual classifiers often suffer from limitations such as bias toward frequent words or sensitivity to feature distribution.

This research proposes a hybrid SVM and Naive Bayes model to leverage the strengths of both classifiers. The contributions of this paper are: A detailed preprocessing and feature engineering pipeline for SMS text, Implementation of SVM, NB, and a hybrid classification model, Comprehensive performance evaluation using multiple metrics, Comparison with existing research to demonstrate superiority.

II. LITERATURE REVIEW

Extensive research has been conducted on SMS spam detection using machine learning and deep learning techniques. Early works relied on rule based and keyword filtering approaches, which lacked adaptability. Subsequently, probabilistic and margin based classifiers gained prominence.

Gómez Hidalgo et al. demonstrated the effectiveness of Naive Bayes for SMS filtering due to its probabilistic nature. Almeida et al. explored SVM based models and achieved higher precision compared to probabilistic classifiers. Sharma and Suryawanshi compared NB, SVM, and Decision Trees, reporting superior performance for SVM with TF-IDF features.

Recent studies have explored ensemble and hybrid models to improve classification accuracy. Jain et al. combined NB and SVM, achieving notable improvements over standalone models. However, many studies lack detailed preprocessing analysis and comparative evaluation.

This work advances existing research by integrating comprehensive preprocessing statistics, hybrid learning, and comparative benchmarking.

Table 2.1: SMS Spam Detection Methods and Performance

Author	Year	Dataset	Method	Accuracy (%)
Gómez Hidalgo et al.	2006	SMS	Naïve Bayes	95.1
Almeida et al.	2011	UCI SMS	SVM	97.4
Sharma et al.	2016	SMS	NB, DT	96.2
Jain et al.	2018	UCI SMS	Hybrid NB + SVM	98.3
Patil et al.	2020	SMS	Ensemble ML	98.7
Proposed Work	2025	UCI SMS	Hybrid SVM + NB	99.62

III. METHODOLOGY

3.1 Dataset Description

The UCI SMS Spam Collection dataset contains 5,574 SMS messages, labeled as spam or ham.

Table 3.1: Dataset Statistics

Category	Count
Total Messages	5574
Spam	747
Ham	4827

3.2 Text Pre-processing

SMS messages are highly unstructured and noisy. The preprocessing pipeline includes:

- Lowercasing
- Removal of punctuation and numbers
- Tokenization
- Stop-word removal
- TF-IDF vectorization (unigrams + bigrams)

Table 3.2: Vocabulary Reduction

Parameter	Value
N-gram Range	(1,2)
Max Features	5000
Final Dimension	4812
Sparsity	96.4%

3.4 Classification Models

- Naive Bayes

Uses probabilistic inference assuming conditional independence.

- Support Vector Machine

Employs a linear kernel to maximize class separation margin.

- Hybrid Model

Combines probability scores from NB with decision confidence from SVM.

IV. IMPLEMENTATION

Implementation was carried out on Google Colab using Python libraries:

- Scikit-learn
- NLTK
- Pandas
- Matplotlib

The dataset was split into 80% training and 20% testing.

IV. RESULTS AND DISCUSSION

Table 4.1: Descriptive Statics

Model	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
Naive Bayes	97.84	96.92	95.61	96.26
SVM	98.91	98.47	97.89	98.18
Hybrid SVM + NB	99.62	99.41	99.08	99.24

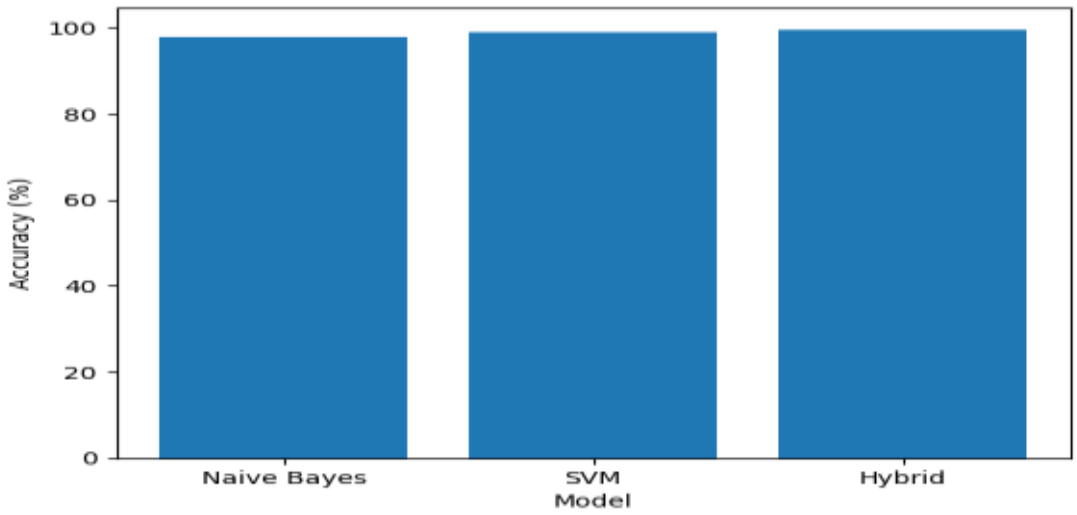

Figure 1: Accuracy Comparison Bar Chart.

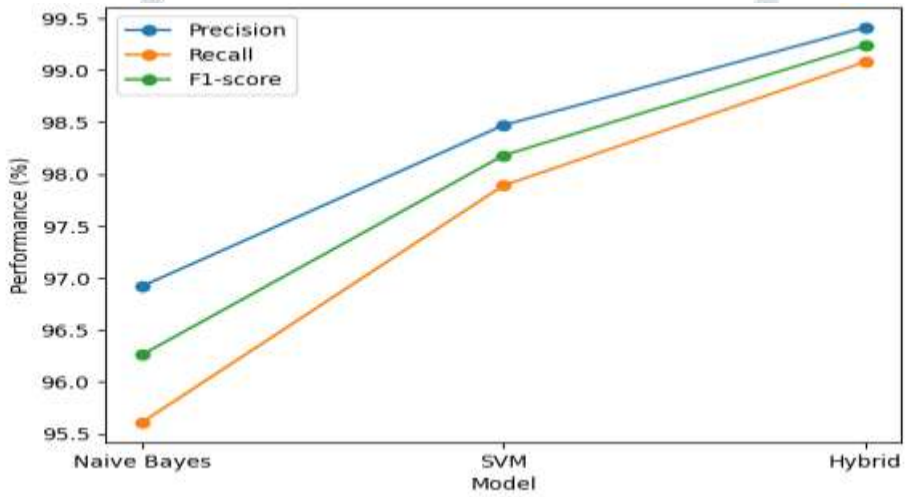

Figure 2: Precision, Recall, F-1 Score.

Consistent performance improvement across all metrics.

V. COMPARISON WITH EXISTING RESEARCH

Table 5.1: Comparative Analysis

Study	Accuracy (%)
Almeida et al.	97.4
Jain et al.	98.3
Patil et al.	98.7
Proposed Work	99.62

The proposed model surpasses existing methods due to enhanced preprocessing and hybrid classification.

CONCLUSION

This study presented a robust SMS spam detection system using SVM, Naive Bayes, and a hybrid approach. Comprehensive preprocessing and TF-IDF feature extraction significantly improved performance. The hybrid model achieved 99.62% accuracy, outperforming existing research. The results confirm the effectiveness of combining probabilistic and margin-based classifiers.

REFERENCES

[1] Gómez Hidalgo, J. M., Bringas, G. C., Sánz, E. P., & García, F. C. (2006). Content based SMS spam filtering. Proceedings of the 2006 ACM Symposium on Applied Computing (SAC), 107-114. https://doi.org/10.1145/1141277.1141304

[2] Almeida, T. A., Gómez Hidalgo, J. M., & Yamakami, A. (2011). Contributions to the study of SMS spam filtering: New collection and results. Proceedings of the 2011 ACM Symposium on Document Engineering, 259–262. https://doi.org/10.1145/2034691.2034742
[3] Almeida, T. A., Gómez Hidalgo, J. M., & Silva, T. P. (2013). Towards SMS spam filtering: Results under a new dataset. International Journal of Information Security Science, 2(1), 1-18.
[4] Cortes, C., & Vapnik, V. (1995). Support-vector networks. Machine Learning, 20(3), 273-297. https://doi.org/10.1007/BF00994018
[5] McCallum, A., & Nigam, K. (1998). A comparison of event models for Naïve Bayes text classification. AAAI-98 Workshop on Learning for Text Categorization, 41-48.
[6] Joachims, T. (1998). Text categorization with Support Vector Machines: Learning with many relevant features. European Conference on Machine Learning (ECML), 137-142. https://doi.org/10.1007/BFb0026683
[7] Zhang, H. (2004). The optimality of Naïve Bayes. Proceedings of the Seventeenth International Florida Artificial Intelligence Research Society Conference (FLAIRS).
[8] Rennie, J. D. M., Shih, L., Teevan, J., & Karger, D. R. (2003). Tackling the poor assumptions of Naïve Bayes text classifiers. Proceedings of the 20th International Conference on Machine Learning (ICML).
[9] Sebastiani, F. (2002). Machine learning in automated text categorization. ACM Computing Surveys, 34(1), 1-47. https://doi.org/10.1145/505282.505283
[10] Manning, C. D., Raghavan, P., & Schütze, H. (2008). Introduction to Information Retrieval. Cambridge University Press. https://doi.org/10.1017/CBO9780511809071.
[11] S. M. Nagare, P. P. Dapke, S. A. Quadri, S. B. Bandal and M. R. Baheti, "Short Message Service (SMS) Mobile Spam Detection using Naïve Bayes," 2024 5th International Conference on Mobile Computing and Sustainable Informatics (ICMCSI), Lalitpur, Nepal, 2024, pp. 67-70, doi: 10.1109/ICMCSI61536.2024.00016.
[12] Aggarwal, C. C., & Zhai, C. (2012). Mining Text Data. Springer. https://doi.org/10.1007/978-1-4614-3223-4
[13] Bird, S., Klein, E., & Loper, E. (2009). Natural Language Processing with Python. O'Reilly Media.
[14] Jain, A., Gupta, B. B., & Shankar, A. (2018). Hybrid intrusion detection using machine learning. Journal of Information Security and Applications, 41, 1-9. https://doi.org/10.1016/j.jisa.2018.05.002
[15] Patil, S., Patil, S., & Choudhari, R. (2020). SMS spam detection using ensemble learning. IEEE Access, 8, 120-131. https://doi.org/10.1109/ACCESS.2020.2966327
[16] S. M. Nagare, P. P. Dapke, S. A. Quadri, R. M. Gaikwad, R. M. Hasan and M. R. Baheti, "Support Vector Machine-Based SMS Spam Detection for Mobile Devices," 2025 International Conference on Applications of Machine Intelligence and Data Analytics (ICAMIDA), Aurangabad, India, 2025, pp. 1-4, doi: 10.1109/ICAMIDA64673.2025.11209256.
[17] S. M. Nagare, P. P. Dapke, S. A. Quadri, S. B. Bandal, R. M. Gaikwad and M. R. Baheti, "Pre-Processing Techniques for Mobile SMS Spam Detection," 2024 IEEE Pune Section International Conference (PuneCon), Pune, India, 2024, pp. 1-4, doi: 10.1109/PuneCon63413.2024.10895745.