International Journal of Artificial Intelligence, Data Science, and Machine Learning
Grace Horizon Publication | Volume 2, Issue 4, 60-69, 2021
ISSN: 3050-9262 | https://doi.org/10.63282/3050-9262.IJAIDSML-V2I4P107

Original Article

Predictive Modeling for Classification of SMS Spam Using NLP and ML Techniques

Aniruddha Arjun Singh Singh[1], Vetrivelan Tamilmani[2], Vaibhav Maniar[3], Rami Reddy Kothamaram[4], Dinesh Rajendran[5], Venkata Deepak Namburi[6]
[1]ADP, Sr. Implementation Project Manager.
[2]Principal Consultant (SAP), Infosys Ltd.
[3]Oklahoma City University, MBA / Product Management.
[4] California University of management and science, MS in Computer Information systems.
[5]Coimbatore Institute of Technology, MSC. Software Engineering.
[6]University of Central Missouri, Department of Computer Science.

Abstract - *Modern telecommunication systems have exposed users and service providers to complex forms of fraudulent communications via SMS spam, resulting in serious disruptions by sending unwanted messages, phishing attempts, and financial scams to millions of users worldwide. The SMS Spam Collection dataset (5,574 messages, 87.37% legitimate and 12.63% spam) is used to classify SMS spam in this study, which extensively evaluates NLP and ML techniques. It addresses the critical challenge of finding effective and precise detection methods for increasingly sophisticated spam. Conventional keyword-based filtering techniques struggle to manage linguistic variations and evolving spam profiles, necessitating more advanced computational approaches. An extensive ML model was developed, incorporating text preprocessing, systematic feature extraction through TF-IDF vectorization, and robust Support Vector Machine (SVM) classification trained on stratified 80-20 data splits with hyperparameter tuning. The system efficiently converts text input into numerical features by performing stemming, tokenization, and punctuation removal. The SVM model achieved 97.85% accuracy, outperforming Naive Bayes (93.9%), KNN (92.26%), and Random Forest (95.46%) in distinguishing spam from legitimate messages. These results demonstrate that SVM-based NLP techniques provide an accurate, scalable, and practical solution for improving telecommunications security and enhancing user experience in modern messaging systems.*

Keywords - *SMS Detection, Naïve Bayes, Spam Detection, Natural Language Processing, SMS Spam Collection Dataset, SVM, Machine Learning, KNN, Random Forest.*

1. Introduction

SMS is a very important communication means, personal, social, and business because of its simplicity, accessibility, and low cost. SMS is also used in SMS marketing which is a direct marketing [1]. SMS marketing can occasionally cause users to experience disturbances. However, the escalating reliance on SMS has also resulted in the proliferation of unwanted and malicious communications, which are mentioned to as SMS spam. Not only do these spam messages degrade the user experience, but they also pose significant risks by disseminating phishing links, fraudulent schemes, and malevolent content that can compromise user privacy and financial security [2]. SMS spam filtering has thus become a critical field of study to guarantee safe and effective mobile communication. Classification of SMS messages as spam or real (ham) communications has been extensively researched using predictive modelling approaches [3]. Rule-based filtering approaches struggle to keep up with the ever-evolving trend of spam communications, making data-driven prediction models more adaptable and accurate [4][5]. These models use labeled datasets to learn discriminative features of spam messages, and use them to filter unseen data, thus providing useful spam filtering in real time.

The use of ML predictive models for the detection of SMS spam has been a success. Ensemble learning, Random Forest models, Naïve Bayes, and Support Vector Machines (SVM) are all classifiers that have shown beneficial in enhancing the accuracy of classification [6]. These algorithms work well in high-dimensional text features and imbalanced data and hence applicable to real-world spam detection problems. Besides, hybrid and deep learning-based models have presented encouraging outcomes in accuracy and the low rate of FP and hence improve the trustworthiness of spam filters systems [7][8]. An important part of the predictive modelling pipeline is Natural Language Processing (NLP), which allows for efficient text pre-processing and feature extraction from the original SMS message. Unstructured SMS content is transformed into a structured input of ML models with the help of the methods of tokenization, stemming, lemmatization, removal of stop-words, and the methods of vectorization, e. g., TF-IDF or word embeddings [9]. More advanced NLP methods, like semantic embedding and deep contextual representation can reflect the linguistic nuances and subtle spam indications and achieve the classification even more

precise. Integration of NLP and ML thus comprise a comprehensive predictive modelling system that improves SMS spam detection systems.

1.1. Motivation with Contribution

The logicality behind this effort was that SMS-based spam messages and the more sophisticated mechanisms of spamming are increasing exponentially and are posing a severe challenge to contemporary telecommunications systems and hence there is a need to develop sophisticated detection mechanisms to match the evolving spamming trends. Conventional key-word based filtering programs find it difficult to crack down on messages that are highly obfuscated and adopt techniques to escape the natural language processing. Conversely, Modern messaging systems that are based on rules are the requirement of SMS spam detection solutions that applied machine learning concepts to provide reliable and precise classification with the simultaneous enhancement of the user experience. The reason is that SMS datasets are not that simple, and they are inherently imbalanced in the class. The key contributions made by this SMS spam detecting environment are as follows:

- The paper introduces a fully implemented machine learning pipeline based on SMS Spam Collection dataset to identify spam in SMS in real-time. It involves text preprocessing, extraction of features, training of SVM model and evaluation.
- A lot of the statistics on SMS spam is wrong, with 87.37% of the messages being real and 12.63% being spam. This paper talks about the problem. It employs efficient data handling techniques to improve detection accuracy and trains its models rigorously.
- The paper uses SVM classification with optimized hyperparameter thus, performing better with an accuracy of 97.85% against the current methods such as NB, KNN, and RF models.
- In order to quickly transform textual SMS data into numerical features that may be utilized for ML categories, this study utilizes rigorous text preparation techniques such tokenization, stemming, punctuation removal, and TF-IDF text processing into vectors.
- This study is extremely efficient in computational time, its predictive ability is excellent, through systematic feature extraction, optimization of preprocessing, which offers a practical solution to the effectiveness of text classification as a time-saving spam detection tool.

1.2. Significance and Novelity

The research is significant because it introduces a powerful machine learning framework targeted at the serious issue of SMS spamming, which affects telecommunications networks. Its new thing is that it implements a lean Support Vector machine methodology that is superior to the existing ones in a comparative and systematic evaluation. Unlike the classical rule-based filtering systems, which operate in high-dimensional textual characteristics and imbalanced dataset normally present in spam detection contexts, this work uses text preprocessing and TF-IDF text vectorization and SVM classification. The suggested framework fulfills the pressing requirement of correct and real-time spam filtering of the latest messaging platforms, yet it does not sacrifice computational performance and high generalization rates, which makes it of great value when it comes to integration into telecommunications infrastructure, where detection accuracy and a low number of false positives are critical to the user experience.

1.3. Structure of Paper

The subsequent structure of the paper: Section II provide the literature review of SMS spam detection, Section III discussed the proposed methodology with each phase of this system design, Section IV evaluate the results of proposed models, comparison, discussion, last Section V provide the conclusion of this work with future work.

2. Literature of Review

The objective of this part is to review the literature on machine learning and natural language processing studies that attempted to classify SMS spam in network environments. The literature reviews that covered the following topics are summarized here: Table I:

Taloba and Ismail (2019) A ML approach that integrates evolutionary algorithms and decision trees is proposed as a means to achieve e-mail spam detection. Given its efficacy and precision, a genetic algorithm appears to be a reasonable candidate for improving decision trees' text categorization performance. The best way to prune the decision tree is to find the ideal value of a parameter called the confidence factor. A genetic algorithm can help with this optimization process. Principle Component Analysis (PCA) is an excellent option for addressing a crucial issue with any text categorization application, including spam detection. According to the findings, when contrasted with the conventional DT method, the hybrid GADT approach considerably improves the accuracy of spam e-mail identification. Furthermore, these outcomes demonstrate that GADT outperforms other conventional text classifiers following PCA [10].

Mansoor and Shaker (2019) System for detecting SMS spam that can adapt to the ever-changing nature of message services. Creating an Arabic and English spam filter is the focus of this project. Two classifiers are employed in the suggested system. A NN is employed as the secondary classifier, with NB serving as the primary classifier. A NB classifier is used to process the incoming messages. The message is relayed to the second classifier to test against spam, in case it was classified as ham;

otherwise it is not relayed. Using a dataset of 80% training data, and 95% accuracy, the proposed method obtained reasonable results with 97% accuracy on the English language [11].

Alzahrani and Rawat (2019) utilized ML techniques to effectively filter out spam in email. NB, LR, NN, and SVM are a few of the most popular and effective ML approaches. The study's primary objective is to identify the most effective method of spam filtering by analysing and contrasting several methodologies. Results show that when it comes to trained classifier models, neural networks perform the best when it comes to identifying ham and spam messages in received communications [12]. Navaney, Dubey and Rana (2018) compares and analyses the performance of several ML methods in identifying spam and legitimate communications, such as NB, SVM, and maximum entropy algorithms. Since more and more businesses are disclosing customers' private information online and more and more individuals are engaging in online activities, the number of spam texts sent by firms is rising can expect an SMS spam filter to perform similarly to how an email spam filter does. Find out that support vector machine gives the best results when compare it to other supervised learning techniques [13].

Choudhary and Jain (2017) lot of people use SMS as a way to communicate online. On the other hand, that has led to an upsurge in assaults targeting mobile devices, such as Introducing SMS Spam, a revolutionary approach to spam filtering that leverages machine learning categorisation algorithms. There are ten telltale signs of spam SMS messages that have been established after a thorough investigation into this topic. True positives are at 96.5% and false positives are at 1.02% according to the Random Forest classification method [14]. Suleiman and Al-naymat (2017) used algorithms for comparisons in ML are NB, DL, and RF. Use them as classifiers in DL and RF, but they are also useful for figuring out which characteristics are most important to feed into these classifiers, as well as NB and RF. A URL's existence in the SMS text and the number of digits are the two most critical criteria that can effect the identification of SMS spam, according to the results of the research. With a recall of 86%, accuracy of 91%, precision of 96%, and f-measure of 96%, the dataset suggested by UCI ML Repositories is utilized in the experiment [15].

ML has recently been investigated for its potential use in identifying spam SMS in contemporary communication settings; first results show promising improvements in accuracy and generalizability. Ensemble models such as XGBoost, RF and hybrid stacking have become useful when working with skewed data sets and modeling diverse spam patterns. Furthermore, have used clustering-based and graph-driven techniques to detect anomalies and latent links in spam messages. Some of them also point out that Artificial Neural Networks and logistic regression are significantly more accurate in classification, and others suggest optimization-based options to increase real-time filtering behavior in dynamic messaging systems. Even with these developments, several critical issues remain, such as the ongoing changes in the content of spam, the lack of interpretability in black-box models, and the challenge of making sure that it is scalable over heterogeneous and multilingual messaging platforms. This has led to an increased research focus on bypassing explainable AI methods by using semantic feature analysis to make SMS spam detection systems more transparent, more trustworthy and more effective.

Table 1. Comparative Analysis of Recent Studies on SMS Spam Detection Using Machine Learning

Author (Year)	Methodology	Key Findings	Advantages	Limitations	Future Work
Taloba and Ismail (2019)	Method that combines PCA with a hybrid of decision trees and genetic algorithms	GADT improves decision tree performance by optimizing the confidence factor; PCA enhances accuracy	Effective optimization of parameters; improved accuracy in spam detection.	Computationally expensive; limited to e-mail spam dataset.	Extend GADT to SMS and multilingual spam datasets.
Mansoor and Shaker (2019)	Two-stage classifier: Naïve Bayes followed by Neural Network for Arabic & English SMS	Achieved 97% accuracy for English and 95% for Arabic with selected features.	Handles multilingual SMS; high accuracy for English spam filtering.	Performance drop for Arabic; feature dependency.	Enhance feature extraction for Arabic; apply deep learning models.
Alzahrani and Rawat (2019)	Classifiers from ML (e.g., NN, LR, NB) used for e-mail spam filtering	Neural Network outperformed other classifiers in accuracy for filtering messages.	Demonstrated effectiveness of NN in spam detection; comparative analysis provided.	Focused on e-mail spam only; limited SMS context.	Extend models to SMS datasets; apply hybrid classifiers.
Navaney, Dubey, and Rana (2018)	Contrasting NB, SVM, and Maximum Entropy for the purpose of	SVM achieved highest accuracy compared to other classifiers.	SVM robustness and high classification accuracy.	Limited feature exploration; evaluation restricted to small dataset.	Incorporate advanced features; apply ensemble

	analysing SMS spam.				approaches.
Anonymous (2018)	RF for SMS spam with 10 extracted features	Got a 96.5% success rate with only a 1.02% false positive.	High detection accuracy with low false positives.	Feature selection limited to 10; dataset not specified in detail.	Expand feature engineering; test across benchmark datasets.
Suleiman and Al-Naymat (2017)	Application of RF, DL, and NB on the UCI SMS dataset	Most significant features are number of digits and URLs; accuracy of 96%, precision 96%, recall 86%, F1-score 91%.	Identified key discriminative features; demonstrated deep learning benefits.	Imbalanced precision-recall trade-off; dataset limitation.	Apply feature selection on larger dataset

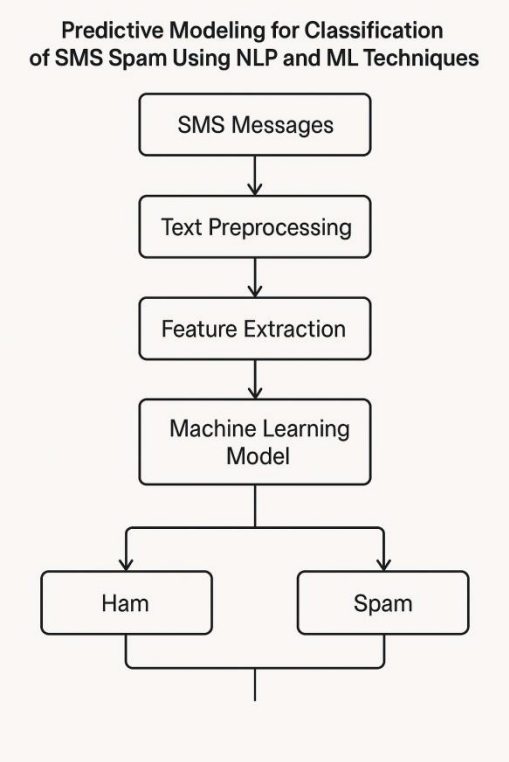

Figure 1: Flowchart for SMS Spam Detection Using Machine Learning Models

3. Methodology

The suggested SMS spam detection methodology is based on a systematic pipeline system starting with the SMS spam collection data as the input into the complete text analysis. Data preprocessing stage consists of several steps such as data cleaning to isolate irrelevant data, removing punctuation marks to exclude non-textual characters and normalizing capitalization to maintain a uniform text format. After that, messages are tokenized and the stemming process is done to reduce words to root words to have better feature consistency. In order for ML algorithms to function with textual input, feature extraction must be performed, and TF-IDF is used to do this task. Half of the processed data should be used for training purposes and half for testing purposes to ensure fair testing. There is a split of 80/20. Secondly, train the support vector machine classifier to differentiate between legitimate and spam messages using the preprocessed attributes. Finally, use traditional classification measures like F1-score, accuracy, precision, and recall to assess the prospective SVM-based method for automated SMS spam detection in telecom systems. This analysis is a component of the larger flow diagram shown in Figure 1.

3.1. Data Collection

The data utilized to conduct the various investigations described here came from the SMS Spam Collection Project. All told, there are 5574 SMS messages here, with 4827 being real and 747 being spam. Researchers have had access to this dataset for the purpose of testing solutions for spam messages. As of right now, it's the most used dataset for studies that aim to categories and detect spam communications. highlighting its relevance for detecting SMS spam some of the visualization are given below:

Fig 2: Occurrence Frequencies of Words in the SMS Spam Collection Dataset

This word cloud visualization represents the frequency distribution of terms commonly found in SMS spam messages, with larger text indicating higher occurrence rate in Figure 2. Prominent spam-related keywords include "free," "call," "txt," "claim," "win," and "prize," which are typical indicators used by ML models for automated spam detection and classification in telecommunications filtering systems.

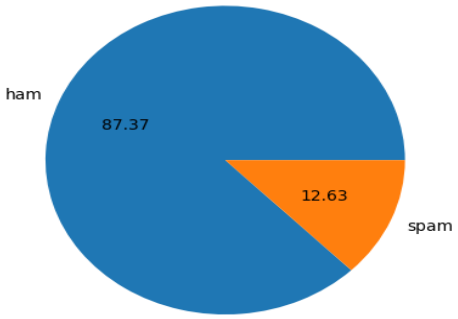

Fig 3: Data Distributions Graph Pie Chart

The distribution of classes in the SMS dataset used for spam detection analysis is shown in Figure 3 via this pie chart. The dataset contains 87.37% legitimate messages (ham) represented in blue and 12.63% spam messages shown in orange, indicating a significantly imbalanced dataset typical in telecommunications spam filtering applications where legitimate messages substantially outnumber spam instances.

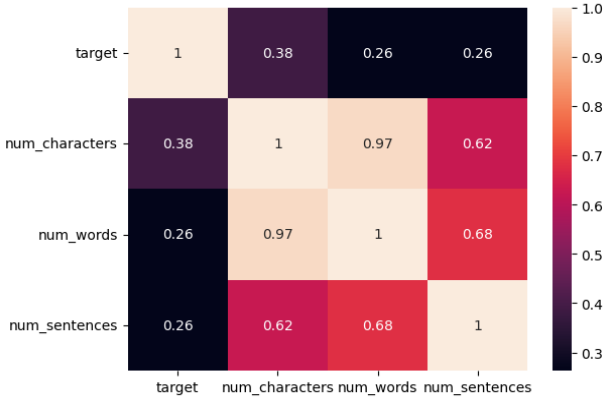

Fig 4: Correlation Heatmap of Different Feature Dataset

This correlation heatmap displays the Pearson correlation coefficients between SMS dataset features for spam detection analysis in Figure 4. Strong positive correlations (0.97) exist between num_characters and num_words, indicating text length relationships. The target variable shows moderate correlations (0.26-0.38) with textual features, suggesting these metrics are useful discriminators for SMS spam classification tasks.

3.2. Data Preprocessing

The initial procedure in data preprocessing was to clean the input text thoroughly by eliminating punctuation and changing it to lowercase. Missing values and duplicate entries were identified and handled to maintain dataset integrity. Tokenization and stemming were used to normalize the text by reducing words to their base forms so that they could be represented consistently. To make text data more machine-learning-friendly, feature extraction was applied using TF-IDF vectorization. Finally, in order to maintain a balanced distribution of classes in the dataset, stratified sampling was employed. The dataset was then divided into

two parts: a training set, comprising 80% of the total, and a testing set, including 20%. Important processes in data preparation involve:

- **Data cleaning:** The suggested model begins with data cleaning. Streamlining the text messages by removing extraneous words and symbols help the machine learning model perform better.
- **Punctuation:** edit out any extraneous symbols and punctuation from the article.
- **Capitalization:** lowercase all words to get rid of the capital letters.

3.3. Tokenization

The TensorFlow Keras Tokenizer API makes text vectorization easier by turning words into integers and making integer or vector sequences. It is an important part of natural language processing jobs that does this. When text is tokenized, punctuation is automatically deleted, which reduces the number of unique words and makes the text representation cleaner, making it more suited for machine learning models. The tokenizer maintains a word index dictionary that maps each unique word to a corresponding integer value, enabling consistent numerical encoding across training and testing datasets.

3.4. Stemming

Stemming is a text preparation method that is used to normalize the various forms of a word by reducing it to its base or root form by removing suffixes. This technique enhances computational efficiency, reduces the number of the vocabulary and enhances generalization because it focuses on the meaning of words.

3.5. Feature Extraction Using TF-IDF

The most significant process is featuring extraction after eliminating the unnecessary words, tokenizing, and converting the information. Text characteristics are more dimensional in nature and contain noisy features, which are suited to this approach. Text mining makes use of the TF-IDF, a statistical metric, mostly to ascertain the significance and relevance of documents. This method uses a word frequency in a given document (TF), as well as, frequency of that word across all papers (IDF). Using the TF-IDF method, one can determine the relative importance of words or tokens within a corpus of documents. The TF-IDF value is often decreased when the word appears more frequently in the corpus, making up for the fact that few words appear more often generally. On the other hand, it climbs proportionally with the number of token occurrences in the document [16]. When compared to a basic count method, TF-IDF yields superior results, making it one of the most widely used term-weighting techniques nowadays. In mathematical terms, TF-IDF is given by Equation (1),

$$TF(t,d) = \frac{f_{t,d}}{\sum_{t'} f_t', d}$$

The phrase fb represents the frequency of term t in document d and $\sum b' f Z'$ stands for the total number of terms in document d, as stated in Equation (2).

$$IDF(t,D) = log\,\frac{N}{|d \epsilon D : t \epsilon d|}$$

The number of documents that contain the word t is represented by the equation $|d \epsilon D : t \epsilon d|$, where N is the total number of documents.

3.6. Data Splitting

The standard data used for SMS spam detection typically consists of 80% training data and 20% testing data. In this manner, the model's efficacy can be suitably assessed. The SVM model learns from the training set to identify particular textual features and linguistic attributes that allow it to distinguish between valid and spam communications.

3.7. Proposed Models of SVM Model in SMS Spam Detection

Supervised learning strategies are often made use of in the field of fraudulent review detection by using classification techniques. This learning technique requires two datasets; test data and training data. In SVM classification, a good hyperplane that successfully separates data points into various classes is the main goal. This improves generalization and decreases misclassification. The basic classification role is referred to as Equation (3):

$$f(x) = sign(w.x + b)$$

x is an input feature vector, www is a weight vector obtained in the course of training, and b is a bias. This linear decision boundary is useful when the data is linearly separable and in real life scenarios like SMS spam detection message feature may be non-linearly related. To address this problem, SVM uses helper functions of the kernel to map the input data into the higher dimensional spaces so as to separate them using a nonlinear hyperplane. Generalized decision formulated based on kernels takes the form of as Equation (4):

$$f(x) = sign\left(\sum_{i=1}^{N} \alpha_i\, y_i k(x_i, x) + b\right)$$

Where α_i are Lagrange multipliers, $y_i \epsilon \{+1, -1\}$ are class labels representing spam and ham and K (x_i, x) is either a linear, polynomially, or radial basis function. Applying this mathematical form to SMS spam detection, the SVM classifier is able to discern accurately between legitimate and spam messages by extracting complex patterns in textual information. Spam: SVM is an effective option in managing high-dimensional feature because messages with a value of f(x) above the specified threshold are spam and ham otherwise.

3.8. Performance Matrix

The capability of the proposed explainable machine to detect mobile SMS spam was evaluated using a number of performance evaluation indicators [4].

- **True Positive (TP) -** The quantity of correctly labelled samples.
- **True Negative (TN) -** the total number of samples that should be immediately removed from consideration.
- **False Positive (FP) -** the sum of all incorrectly rejected class samples.
- **False Negative (FN) -** sum up all the samples that were wrongly placed in the right category.

3.8.1. Accuracy

Precision, represented by A, is the fraction of all SMS that accurately identify spam messages, as shown in Equation (5).

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \times 100 \quad (5)$$

3.8.2. Precision

Precision of SMS spam detection, defined as the accuracy of affirmative predictions, provides an accurate estimate of the number of messages that are truly spam (as indicated in Equation (6)):

$$Precision = \frac{TP}{TP+FR} \times 100 \quad (6)$$

3.8.3. Recall

Recall (R), sometimes called sensitivity, is defined as the ratio of the number of relevant instances retrieved to the number of relevant examples actually present in the sample. It is explained in Equation (7):

$$Recall = \frac{TP}{TP+FN} \times 100 \quad (7)$$

3.8.4. F1 Score

The F-measure (F1) is the harmonic mean of recall and accuracy. The following equations show the trade-off between recall and accuracy shown in Equation (8):

$$F1 - score = \frac{2 \times recall \times precision}{recall + precision} \quad (8)$$

4. Result and discussion

This section presents a comprehensive experimental comparison of the SVM in the collection of SMS spam. As standard binary classification metrics, the model's ability to differentiate between valid and spam SMS messages was assessed using the accuracy, precision, recall, and F1-score. These tests were run in a Jupiter Notebook on Google Colab with Python 3.8. The important libraries used were scikit-learn to run support vector machine, pandas and NumPy to prepare and visualize data and seaborn and matplotlib to visualize data. The compute service was based on NVIDIA RTX 3070 with 32 GB of RAM that were used to train SVMs and test their effectiveness on the SMS dataset. The text-based SMS data was transformed into numerical characteristics that SVMs in classification using text preparation methods such as tokenization, stop words removal, and TF-IDF vectorization. By utilizing confusion matrices, ROC curves, and classification reports, analysis provides comprehensive details regarding SVM's behaviour when it comes to distinguishing between ham and spam messages.

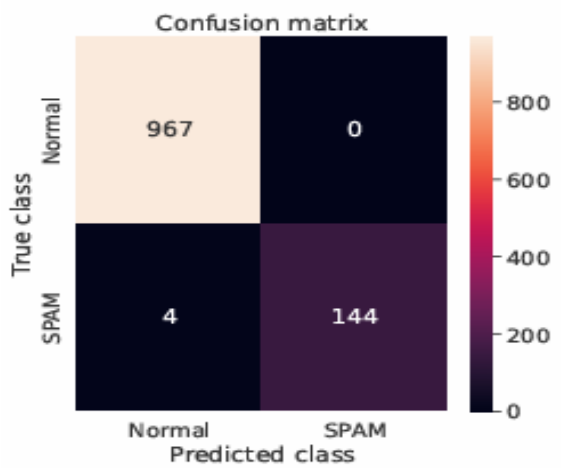

Fig 5: Confusion matrix of SVM model

This confusion analyses the SVM performance over spam in Figure 5. The model accurately identified 967 normal emails and 144 spam emails and only 4 false positives (normal was considered spam) and 0 false negatives (spam was considered normal). The heat map visualization represents the accuracy of prediction between the true and predicted class labels.

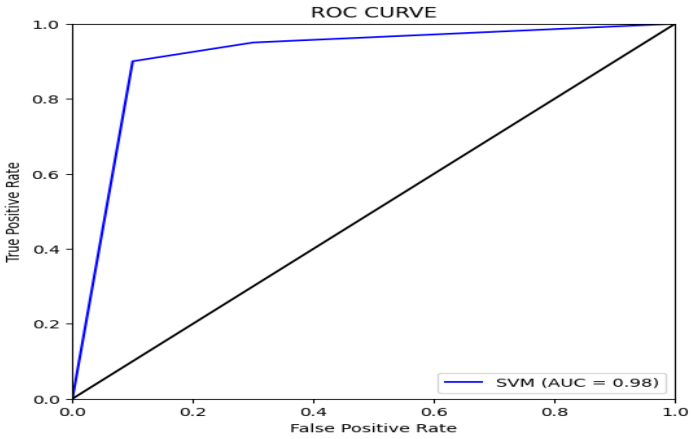

Fig 6: Roc Curve of SVM Technique

Figure 6 demonstrates that an SVM model achieves an AUC of 0.98 on this ROC curve. The blue curve increases steeply out of the top left corner and remains close to it, which represents high true positive rates with a low amount of FP. The black line is the diagonal line which shows random chance performance.

Table 2: Proposed Models Performance on Sms Spam Detection on SMS Spam Collection Dataset

Measure	SVM
Accuracy	97.85
Precision	98.70
Recall	97.97
F1-score	97.31

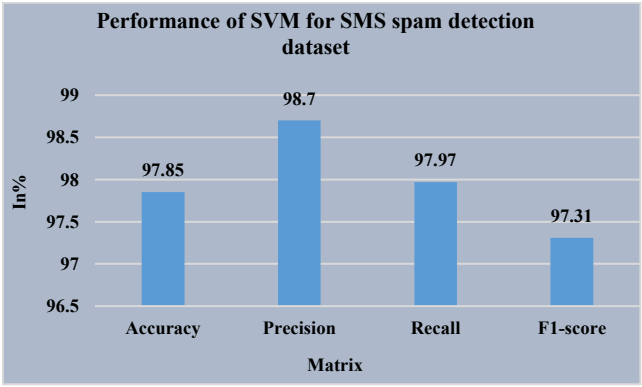

Fig 7: Comparison of Model Performance Metrics

Figure 7 and Table II show the results of the suggested SVM model's performance analysis on the SMS Spam Collection dataset, correspondingly. The collected results show that the SVM model has an F1-score of 97.31%, a recall of 97.97%, a precision of 98.70%, and an accuracy of 97.85%. In addition to demonstrating the SVM classifier's high efficiency in minimising the amount of FP and FN, these data also indicate to its high efficiency in differentiating between spam and valid SMS messages, indicating a high overall efficiency in spam detection activities.

4.1. Discussion

Table III displays the results of a comparison study of SMS spam detectors using data from the SMS Spam Collection. The research looked at four different ML approaches and how well they performed in text categorization circumstances by measuring accuracy. SVM model suggested a highest accuracy of 97.85 percent that virtually surpassed the already existing models and showed that it was more useful in detection of a spam pattern within the SMS textual data. RF showed a high performance of 95.46% accuracy, KNN performance was 92.26% and Naive Bayes performance was 93.9%. The results clearly show that the suggested SVM model does indeed outperform the existing machine learning frameworks on SMS spam detection within the textual communication information, which is presumably due to its ability to adequately address the high-dimensional

feature space created by the TF-IDF vectorization procedure and the multidimensional textual use patterns that are frequently present during spam classification problems. Its high-performance rate, the great adaptability towards solving text classification problems, and good generalization features make the method most suitable in telecommunication systems where the most important considerations are accuracy and reliability in automated spam filtering and enhancing user experience.

Table 3: Comparison between All Proposed Model And Existing Models for Sms Spam Detection

Measure	Accuracy
SVM	97.85
Naïve Bayes[17]	93.9%
KNN[18]	92.26
RF[19]	95.46

The SVM model proposed has an excellent performance in SMS spam classification and the high accuracy of 97.85 is significantly higher than the existing SMS classification models in the text classification task. The model, with the help of the powerful kernel approach and separation of the hyperplane that is best with SVM, is also proficient in identifying complex patterns in texts and working with high-dimensional feature space created by TF-IDF vectorization processing to correctly classify spam in telecommunications contexts. Its high accuracy of the SVM algorithm demonstrates the efficiency of the approach with the imbalanced SMS collections and multi-dimensional linguistic diversities and provides constant predictions to the automated spam filtering systems. However, adaptation to changing spam techniques and computation complexity in large scale messaging systems are some of the challenges possible. The excellent generalization performance of the SVM model, together with its ability to operate with non-linear decision functions through the aid of the kernel functions, makes it particularly applicable in the text classification scenarios where either the dimension of features, as well as the complexity of the pattern, have significant influence. Generally, this SVM-based system provides telecommunication companies and security gurus with an effective, reliable and accurate aspect of detecting spam messages and maintaining high-detection rates and enabling successful spam blocking in modern communications system.

5. Conclusion and Future work

The current telecommunications environment demands a high security level because of the rapid development of technologies and the growth of spam. Identifying spam messages in communication systems is crucial because threats of privacy infiltration and service interruption are enormous. Dealing with SMS spam collection files becomes more challenging when junk messages outnumber real ones. Existing studies on telecommunications system spam detection and management using NLP and ML have shown limitations. Findings in this research support previous work, demonstrating that SVM classification is an effective spam classifier for real-time SMS. With a high accuracy of 97.85%, the SVM model was found to be reliable in detecting spam messages with low error rates. It also efficiently manages high-dimensional textual features using kernels and TF-IDF vectorization. The results indicate that the model can be applied in real-time telecommunications systems where quick, precise, and effective detection is essential. This research provides a practical, data-driven solution for SMS spam classification and emphasizes the importance of advanced text processing methods. Future work can expand the framework to deep learning algorithms like neural networks, RNNs, and LSTMs, improving accuracy, efficiency, and adaptability against evolving spam strategies.

References

[1] L. N. Lota and B. M. M. Hossain, “A Systematic Literature Review on SMS Spam Detection Techniques,” *Int. J. Inf. Technol. Comput. Sci.*, vol. 9, no. 7, pp. 42–50, Jul. 2017, doi: 10.5815/ijitcs.2017.07.05.

[2] T. A. Almeida, J. M. G. Hidalgo, and A. Yamakami, “Contributions to the study of SMS spam filtering: new collection and results,” in *Proceedings of the 11th ACM Symposium on Document Engineering*, in DocEng ’11. New York, NY, USA: Association for Computing Machinery, 2011, pp. 259–262. doi: 10.1145/2034691.2034742.

[3] H. Sajedi, G. Z. Parast, and F. Akbari, “SMS Spam Filtering Using Machine Learning Techniques : A Survey,” *Mach. Learn. Res.*, vol. 1, no. 1, pp. 1–14, 2016, doi: 10.11648/j.mlr.20160101.11.

[4] S. M. Abdulhamid *et al.*, “A Review on Mobile SMS Spam Filtering Techniques,” *IEEE Access*, vol. 5, pp. 15650–15666, 2017, doi: 10.1109/ACCESS.2017.2666785.

[5] P. Pathak, A. Shrivastava, and S. Gupta, “A survey on various security issues in delay tolerant networks,” *J Adv Shell Program.*, vol. 2, no. 2, pp. 12–18, 2015.

[6] S. S. S. Neeli, “Serverless Databases : A Cost-Effective and Scalable Solution,” *IJIRMPS*, vol. 7, no. 6, 2019.

[7] S. J. Delany, M. Buckley, and D. Greene, “SMS spam filtering: Methods and data,” *Expert Syst. Appl.*, vol. 39, no. 10, pp. 9899–9908, Aug. 2012, doi: 10.1016/j.eswa.2012.02.053.

[8] J. M. G. Hidalgo, G. C. Bringas, E. P. Sánz, and F. C. García, “Content based SMS Spam Filtering,” in *roceedings of the 2006 ACM symposium on Document Engineering*, in DocEng ’06. New York, NY, USA, NY, USA: ACM, Oct. 2006, pp. 107–114. doi: 10.1145/1166160.1166191.

[9] P. Sethi, V. Bhandari, and B. Kohli, “SMS spam detection and comparison of various machine learning algorithms,” in *2017 International Conference on Computing and Communication Technologies for Smart Nation (IC3TSN)*, 2017, pp.

28–31. doi: 10.1109/IC3TSN.2017.8284445.
[10] A. I. Taloba and S. S. I. Ismail, "An Intelligent Hybrid Technique of Decision Tree and Genetic Algorithm for E-Mail Spam Detection," in *2019 Ninth International Conference on Intelligent Computing and Information Systems (ICICIS)*, IEEE, Dec. 2019, pp. 99–104. doi: 10.1109/ICICIS46948.2019.9014756.
[11] H. H. Mansoor and S. H. Shaker, "Using classification techniques to SMS spam filter," *Int. J. Innov. Technol. Explor. Eng.*, vol. 8, no. 12, pp. 1734–1739, 2019, doi: 10.35940/ijitee.L3206.1081219.
[12] A. Alzahrani and D. B. Rawat, "Comparative Study of Machine Learning Algorithms for SMS Spam Detection," in *2019 SoutheastCon*, 2019, pp. 1–6. doi: 10.1109/SoutheastCon42311.2019.9020530.
[13] P. Navaney, G. Dubey, and A. Rana, "SMS Spam Filtering Using Supervised Machine Learning Algorithms," in *2018 8th International Conference on Cloud Computing, Data Science & Engineering (Confluence)*, 2018, pp. 43–48. doi: 10.1109/CONFLUENCE.2018.8442564.
[14] N. Choudhary and A. K. Jain, "Towards Filtering of SMS Spam Messages Using Machine Learning Based Technique," in *Advanced Informatics for Computing Research*, D. Singh, B. Raman, A. K. Luhach, and P. Lingras, Eds., Singapore: Springer Singapore, 2017, pp. 18–30.
[15] D. Suleiman and G. Al-naymat, "SMS Spam Detection using H2O Framework," *Procedia Comput. Sci.*, vol. 113, pp. 154–161, 2017, doi: 10.1016/j.procs.2017.08.335.
[16] N. Hussain, H. Turab Mirza, G. Rasool, I. Hussain, and M. Kaleem, "Spam Review Detection Techniques: A Systematic Literature Review," *Appl. Sci.*, vol. 9, no. 5, 2019, doi: 10.3390/app9050987.
[17] S. Gheewala and R. Patel, "Machine Learning Based Twitter Spam Account Detection: A Review," in *2018 Second International Conference on Computing Methodologies and Communication (ICCMC)*, 2018, pp. 79–84. doi: 10.1109/ICCMC.2018.8487992.
[18] H. Raj, Y. Weihong, S. K. Banbhrani, and S. P. Dino, "LSTM Based Short Message Service (SMS) Modeling for Spam Classification," in *Proceedings of the 2018 International Conference on Machine Learning Technologies*, New York, NY, USA: ACM, May 2018, pp. 76–80. doi: 10.1145/3231884.3231895.
[19] A. Tekerek, "Support Vector Machine Based Spam SMS Detection," *Politek. Derg.*, vol. 22, no. 3, pp. 779–784, Sep. 2019, doi: 10.2339/politeknik.429707.