International Journal of Scientific Research & Engineering Trends
Volume 12, Issue 3, May-Jun-2026, ISSN (Online): 2395-566X

SMS Classifier with Encryption Decryption Using Machine Learning

Bhonde Vrushali Baban[1], Patil Renuka Rajendra[2], Wakle Anita Ashok[3], Kulkarni Mrunmayee Mangesh[4], Archana Sachin Gaikwad[5], Poornima Nandu Pathak[6]
Department of Computer Engineering KVN Naik Loknete Gopinathji Munde
College of Engineering And Research, Nashik

Abstract- — SMS spam is becoming more frequent as spammers use it to reach their targets. Although spam messages can be annoying, they can also be dangerous because they might try to steal personal information or direct users to harmful websites. This work explains how SMS spam is detected. It describes the different types of spam, the features used to find them, and the methods used to stop them. We also talk about some of the difficulties in identifying SMS spam and possible future methods to deal with them.

Keywords: SMS Classification, Machine Learning, Spam Detection, Encryption and Decryption, Data Security, Text Mining

I. INTRODUCTION

In today's digital age, Short Message Service (SMS) what left one of the most widely used ways for people to communicate, even though instant messaging apps are becoming increasingly popular. SMS is used for a lot of things, like banking alerts, promotions, personal conversations, and work updates. However, as the number of messages grows, it can become hard to manage inboxes. Users often struggle to find important messages and decide which ones to prioritize first.

Traditional SMS inboxes usually just show all messages in a simple list, without any sorting, filtering, or priority [7],[9]. This can make it difficult to spot important messages like transaction alerts, work updates, or service notifications, especially when they get mixed in with spam and promotional content. This lack of organization can slow down productivity and make the user experience worse. To fix these problems, we suggest an SMS Classifier System that automatically groups incoming messages into different categories such as Transactional, Personal, Promotional, Work, Service, Social, and Spam.

In addition, the app will have several useful features:

- **Favourite Messages:** This lets users mark important messages so they can find them quickly.
- **Restore Bin:** Similar to a recycle bin, this lets users bring back messages they accidentally deleted.
- **Encryption & Decryption:** This helps keep messages private by letting users secure sensitive messages and unlock them when needed.

II. LITERATURE REVIEW

SMS Guard: Enhancing Spam Detection using Multinomial naive Bayes for Secure Communication [1] Abayomi-Alli etl. aims to build a system that can quickly and accurately identify spam messages in real time. It uses the Multinomial Naive Bayes algorithm along with important steps such as data preprocessing and feature extraction techniques like TF-IDF to classify messages as either spam or genuine.

The system is designed to provide good accuracy while using minimal computational resources. The observe highlights that the use of TF-IDF features improves the machine's capability to apprehend essential words in messages. It additionally suggests that Multinomial Naive Bayes plays well despite large datasets due to its low computational value. (Manjula Devi C, 2025) [1]

Advanced SMS Spam Detection using Integrated Feature Extraction [2] A Qazi, etl. aims to create a framework that uses multiple types of features, including text-based, word-level, and context-based information, to improve spam detection. It applies methods such as TF-IDF, n-grams. This integrated method captures both the meaning and hidden patterns in SMS, resulting in higher accuracy, better detection of advanced spam, and fewer false alerts. It is widely used in mobile spam filtering, telecom security, banking fraud prevention. (G Subhashini, 2025) [2]

Spam SMS Classifier Using Machine Learning Algorithms [3] J. Chuma, etl. build a system that can Successfully discover and classify spam SMS messages. The research examines different machine learning methods, Support Vector Machine

© 2026 IJSRET

(SVM), Multinomial Naive Bayes, and Extra Tree Classifier, and concludes that Naive Bayes works the best. The look at compares the overall performance of various algorithms to recognize which one gives the maximum reliable outcomes. It highlights that Multinomial Naive Bayes performs better handles text-based capabilities effectively. (Bhagyashri Mankar, 2025) [3]

Combatting SMS Spam: A Machine Learning Approach for Accurate and Scalable Detection [4] Merugu, S. etl. seeks create machine learning can correctly and quickly detect spam messages. The system is built to determine whether a message is spam or a genuine one and uses techniques like tokenization and vectorization to improve detection. The observe emphasizes constructing a scalable a model that can deal with this big volume of SMS information efficiently. by making use of tokenization and vectorization, the gadget converts textual content meaningful numerical capabilities higher mastering. (Robin Britto V, 2025) [4]

Machine Learning-Based Solution for SMS Spam Detection Problem [5] S. M. Abdulhamid etl. develop a machine learning approach that correctly identify whether messages are spam. It evaluates several algorithms including Naive Bayes, SVM, Logistic Regression, at Random Forest, Decision and KNN to determine which one works best. The observe compares multiple algorithms to find the most correct and green classifier for SMS unsolicited mail detection. It highlights how one-of-a-kind fashions behave with various message styles and function units. The results display that some algorithms outperform others in terms of pace, precision, and reliability. (Ahmed Younes Shdefat, 2024) [5]

Natural Language Processing based classifier for identifying spam email and SMS message [6] Pham T.H. etl. Create a smart system that uses Natural Language Processing and machine learning to automatically recognize and remove spam. The classifier is trained on real data sets to accurately differentiate between harmful content and legitimate messages. The aim is to improve user experience and security by reducing unwanted spam and ensuring important messages are delivered. The machine blessings from NLP techniques that help it recognize the that means and shape of messages greater efficaciously. by using education on real-international datasets, the version becomes more study and adaptable to exceptional unsolicited mail formats. (Habibur Rahman, 2024) [6]

Enhancing SMS Classification with Ensemble Machine Learning Techniques [7] Gupta, P etl. focuses on building a system that can accurately and efficiently classify SMS messages as spam or genuine. It uses a group of models working together, including
Vector machine (SVM), multinomial Naïve Bayes (MNB) and Extra Classifier, to increase the overall Ability. The ensemble method enhances performance by bringing together the advantages of different algorithms rather than depending on just one. This allows the gadget handle different varieties of spam patterns extra successfully. (Sonam Tyagi, 2024) [7]

Spam SMS Detection Using Natural Language Processing [8] Nagare etl. create and put into use a system can automatically detect and remove spam SMS messages by using Processing and machine learning techniques. The system analyzes the text features of messages to classify them accurately, with the aim of making it better communication security and reducing exposure to unwanted messages.

by way of changing textual content into meaningful functions, the classifier can higher distinguish between and unsolicited mail content. (Dr. Satpalsing D. Rajput, 2024) [8]

Comparative Study of Machine Learning Algorithms for SMS Spam Detection [9] Chaturvedi etl. assess and compare how well different machine learning methods work in differentiating between spam and regular SMS messages. The research uses several models such as Naive Bayes. It presents insights into which models provide better accuracy, pace and reliability for real-time spam detection. The assessment allows pick out the strengths and weaknesses of each classifier. (Amani Alzahrani, 2019) [9]

Enhancing Spam Detection on Mobile Phone Short Message Service (SMS) Performance using FP- Growth and Naive Bayes Classifier [10] Shirani- Mehr etl. enhance the validity and speed of detecting spam messages by using Growth algorithm to extract features classifier to categorize message. .The FP-increase set of rules facilitates the system fast find frequent word styles which can be normally seen in unsolicited mail messages. by combining these extracted styles with the Naive Bayes classifier, the version will become more correct and efficient. (Dea Delvia Arifin, 2017) [10]

Research Gaps

1. **Binary Classification Only**: These systems classify SMS messages as either spam or ham, without further categorization like work, social, or promotional messages [3],[9].
2. **Keyword/Blacklist Based:** They mostly use fixed rules or keywords, so they cannot easily handle new types of data [1],[5].

© 2026 IJSRET

3. **High False Positives/Negatives:** Sometimes, real messages get incorrectly marked as spam, and spam messages might be mistaken for legitimate content [4],[6].
4. **Static Model:** Traditional ways of dealing with spam able to keep pace the new spammers [5],[9].
5. **No Security Layer:** While existing classifiers can detect spam, they do not protect the privacy of the SMS data itself [6],[8].

Solution

Our proposed system extends beyond just spam detection by integrating categorization + security + usability features:

Multi-Category Classification: Instead of only classifying messages as spam or ham, our system categorizes SMS into Transactional, Personal, Promotional, Service, Social, Work, and Spam. [7],[9].

Encryption & Decryption (AES/RSA): This ensures that sensitive messages, like banking details or OTPs, are kept private and secure. Unlike existing systems, our system protects data even after classification [6],[10].

Favourites Feature: Users can mark important messages for easy access. This adds a personal touch that is not available in the base system.

Recycle Bin (Restore Option): If a user from an accidentally deletes an SMS, it can be restored. This adds flexibility and helps prevent data loss.

Secure Storage Module: Messages are stored in an encrypted database, ensuring confidentiality data.

Objectives

Accurate SMS Classification: Automatically sort SMS messages into different categories like Transactional, Personal, Promotional, Work, Service, Social etc [3],[7].

Spam Detection and Reduction: Efficiently identify and filter out spam, phishing, and promotional messages, minimizing errors such as false positives or false negatives.

Data Privacy and Security: Use strong encryption and decryption methods like AES and RSA to ensure that SMS data is stored [6],[10].

User-Friendly Inbox Management: Include helpful tools such as Favorites for easy access to important messages, and a Recycle Bin or Restore option to recover accidentally deleted messages.

Deployment and Real-Time Use: Create a web or mobile interface that allows real-time SMS classification and secure management.

Research and Innovation: Improve existing spam detection systems by transitioning from a simple Spam/Ham classification to a more detailed multi category.

III. METHODOLOGY

SMS Input Module: Collect incoming SMS messages from the user's device or a server.

Pre-processing: Clean and prepare the text data by breaking it into words, removing unnecessary words, and reducing words to their base form, making it easier for the next steps [4],[6].

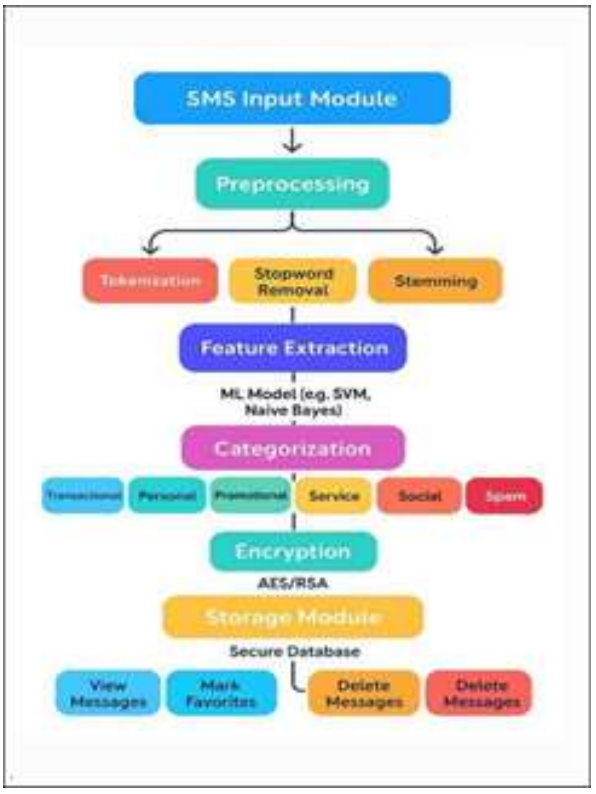

Fig: System Architecture

Feature Extraction: Turn the text into numbers using methods like TF-IDF Word2Vec, which assist machine learning models grasping the meaning the messages [2],[8].

Classification: Use machine learning techniques like Support Vector Machine or Naive Bayes to classify messages into categories such Transactional, Personal, Promotional, Work, Service, Social, or Spam [3],[5].

© 2026 IJSRET

After pre-processing, every SMS is converted into numbers the usage of function extraction techniques like TF-IDF or Word2Vec.

These numerical functions are given to a machine studying version such as:

- Naive Bayes
- Vector gadget (SVM)

The version studies styles from the training dataset, like:

- Words used in promotional SMS
- Sentence structure in personal SMS
- Common features in transactional SMS

Using those learned styles, the version predicts which category the brand new SMS belongs to.

1. **Categorization:** Group messages based on their classifications, helping users filter and manage their SMS more effectively.
2. Type offers the predicted magnificence (for instance:
3. "Promotional").
4. Categorization takes this elegance and locations the SMS into the perfect category in the UI.

It creates prepared message sections like:

- Inbox
- Spam
- Favorites
- Promotional
- Work
- Social
- Personal
- Recycle Bin

1. **Encryption Module:** Protect messages by encrypting them using algorithms like AES or RSA before they are stored [6],[10].
2. **Storage Module**: Keep encrypted messages in a secure database to maintain their integrity and keep them private.
3. **User Interface:** Offer users an easy-to-use platform where they can view SMS, mark favorites, and delete messages they don't need.
4. **Recycle Bin:** Temporarily store deleted messages so users can choose to restore them or permanently delete them, improving overall data handling.

IV. MATHEMATICAL MODEL

The mathematical model represents how an SMS message is classified into a specific category using machine learning.

Let:

$M = m_1, m_2, m_3 \ldots, m_n$ be the group of SMS messages.

$C = c_1, c_2, c_3 \ldots, C_k$ be the set of categories (Transactional, Personal, Promotional, Work, Social, Spam).

Each message is converted into a numerical feature vector using TF-IDF.

TF-IDF Formula:

$$TF - IDF(t,d) = TF(t,d) \times \log\left(\frac{N}{dF(t)}\right)$$

Where:

$TF(t,d)$ = frequency of term t in document d

$DF(t)$ = number of documents holding back term t

N = total number of documents

For classification, Naive Bayes theorem is applied:

$$P(C|X) = \frac{P(X|C) \times P(C)}{P(X)}$$

Where:

X = input SMS message

C = predicted class

The output is chosen to be the class with the highest probability.

Algorithm

Input: SMS message

Output: Classified and encrypted SMS

1. Read SMS message.
2. Perform preprocessing (remove symbols, convert to lowercase).
3. Extract features using TF-IDF.
4. Apply Naive Bayes or SVM classifier.
5. Predict category of SMS.
6. Encrypt the message using AES or RSA.
7. Store encrypted message in database.
8. Decrypt when user requests to view message.

Performance Evaluation

To assess the performance of the system following metrics are used:

© 2026 IJSRET

Accuracy:

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN}$$

Precision:

$$Precision = \frac{TP}{TP+FP}$$

Recall:

$$Recall = \frac{TP}{TP+FN}$$

F1score:

$$F1 = \frac{2\times Precision\times Recall}{precision+Recall}$$

Where:

- TP = True Positives
- TN = True Negatives
- FP = False Positives
- FN = False Negatives

These measures how correctly the SMS messages are classified.

Scope

1. To automatically classify incoming SMS messages into categories like non-public, Promotional, Transactional, work, Social, and junk mail.
2. To provide secure message handling the use of AES and RSA encryption to guard sensitive statistics.
3. Providing features to enhance the user experience like Favorites, Recycle Bin, and organized message folders.
4. To lessen spam and undesirable messages the use of system mastering algorithms inclusive of Naive Bayes or SVM.
5. To develop a smart, user-friendly SMS management system that can be integrated into mobile apps.
6. To mix message class with security to make certain both accuracy and privacy in verbal exchange.

Application

1. **Mobile Messaging Apps:** For automatic SMS sorting and secure message storage.
2. **Banking and Financial Services:** To classify transactional alerts and protect OTPs and account- related messages.
3. **Telecommunication industry:** For unsolicited mail detection, SMS filtering, and at ease communication offerings.
4. **Customer Support systems:** To categorize customer messages and secure sensitive user data.
5. **Business Communication Platform systems:** To separate work-related messages from personal ones and prevent data leaks.
6. **E-commerce & Marketing:** To clear out promotional SMS and reduce user annoyance with ads.
7. **Cybersecurity applications:** For encrypted data storage and secure message transmission.

V. CONCLUSION

The project "SMS Classifier with Encryption & Decryption using Machine Learning" helps organize SMS communication, makes it more secure, and improves the user experience. It sorts messages into different categories, keeps sensitive information safe with encryption, and includes useful features like Favorites and Recycle Bin. These features lead to more accurate message handling, better privacy and a more convenient experience for users compared to other system.

REFERENCES

1. Abayomi-Alli, O., Misra, S., Abayomi-Alli, A., Odusami, M. (2019). “A review of soft techniques for SMS spam classification: Methods, approaches, and applications”. Engineering Applications of Artificial Intelligence, 86, 197-212.
2. Qazi, N. Hasan, R. Mao, M. E. M. Abo, S. K.Dey and G.Hardaker, “Machine Learning-Based Opinion Spam Detection: A Systematic Literature Review,”inIEEEAccess,doi:10.1109/ACCESS.2024.3 3 99264.
3. J. Chuma, “Short message service (SMS) spam detection on mobile phones (Doctoral dissertation BUSE)”, 2018.
4. Merugu, S., Reddy, M. C. S., Goyal, E., Piplani, L. (2019). “Text message classification using supervised machine learning algorithms.” In ICCCE 2018: Proceedings of the International Conference on Communications and Cyber Physical Engineering 2018 (pp. 141-150). Springer Singapore.
5. S. M. Abdulhamid et al., “A review on mobile SMS spam filtering techniques,” IEEE Access, vol. 5, pp. 15 650–15 666, 2017.
6. Pham, T.H. and Le-Hong, P., 2016, November. “Content-based approach for Vietnamese spam SMS filtering”. In

© 2026 IJSRET

2016 International Conference on Asian Language Processing (IALP) (pp. 41-44). IEEE
7. Gupta, P., Sahu, A., Singh, A., Awasthi, A. and Singh, R., 2022, March. “SMS Spam Classifier by Naive Bayes Classifier”. In Proceedings of the International Conference on Innovative Computing Communication (ICICC).
8. Nagare, Samadhan. (2021). “Mobile SMS Spam Detection using Machine Learning Techniques”. 7. 331-334.
9. Chaturvedi, S.A. and Purohit, L., 2023. “Feature selection-based spam system in SMS and email domain. In Sentiment Analysis and Deep Learning”: Proceedings of ICSADL 2022 (pp. 37-52). Singapore: Springer Nature Singapore.
10. Shirani-Mehr, Houshmand. “SMS spam detection using machine learning approach” (2013):

© 2026 IJSRET