SMS SPAM FILTERING

A PROJECT REPORT

Submitted by:

Kumar Prasanjeet (21BCS5270)

Manish Singh (21BCS5294)

in partial fulfillment for the award of the degree of

BACHELOR OF ENGINEERING

IN

COMPUTER SCIENCE

Chandigarh University

April 2023

CHANDIGARH UNIVERSITY

Discover. Learn. Empower.

BONAFIDE CERTIFICATE

Certified that this project report "SMS Spam Filtering" is the bonafide work of: " **Kumar Prasanjeet (21BCS5270), Manish Singh (21BCS5294)**" who carried out the project work under our supervision.

SIGNATURE	**SIGNATURE**
Natasha Sharma	**Dr. Sandeep Kang**
(SUPERVISOR)	**(HEAD OF DEPARTMENT)**

INTERNAL EXAMINER	**EXTERNAL EXAMINER**

TABLE OF CONTENT

ABSTRACT

As mobile communication continues to grow, so does the menace of SMS spam, posing a significant nuisance to users and network providers. This report delves into the realm of SMS spam filtering, providing a thorough analysis of existing techniques and their effectiveness. We explore the escalating challenges posed by SMS spam, including evolving spamming tactics and obfuscation methods.

Our evaluation encompasses rule-based approaches, machine learning algorithms, and hybrid solutions, shedding light on the strengths and weaknesses of each. We emphasize the delicate balance between minimizing false positives and false negatives while ensuring user privacy and consent.

Furthermore, we discuss the ethical implications surrounding SMS spam filtering into emerging trends such as deep learning models and blockchain-based solutions. These insights empower researchers, network operators, and policymakers with the knowledge to tackle SMS spam effectively and preserve the quality of mobile communication services.

CHAPTER 1.

INTRODUCTION

The ubiquity of SMS communication has brought with it the pervasive problem of SMS spam, disrupting user experiences, raising privacy concerns, and posing security threats. To combat this issue, effective SMS spam filtering techniques are essential. This report explores the landscape, challenges, and solutions of SMS spam filtering, including ethical considerations and emerging trends, providing a foundation for addressing this evolving problem.

1.1. Identification of Client /Need / Relevant Contemporary issue/Project Scope:

Identification of Client:

The primary clients for this project encompass both individual mobile users and mobile network providers. Individuals seek relief from the nuisance of SMS spam, while network providers aim to enhance user satisfaction and protect network resources.

Identification of Need:

User Experience Improvement: Mobile users need a more enjoyable and hassle-free SMS experience, free from intrusive spam messages.

Privacy Protection: Users require protection against potential privacy breaches resulting from unsolicited SMS spam that often targets personal information.

Security Enhancement: There is a need to bolster security, as SMS spam can serve as a vector for phishing attacks and malware distribution.

Relevant Contemporary Issue:

The relevant contemporary issue is the ever-increasing volume and sophistication of SMS spam, which affects users, network providers, and regulators. The issue demands proactive measures to counter evolving spamming tactics and protect user experiences and privacy.

Project Scope:

Objective: The project aims to explore and evaluate various SMS spam filtering techniques and strategies to mitigate the impact of SMS spam on users and network providers.

Methodology: The scope involves a comprehensive analysis of existing filtering approaches, performance evaluations, ethical considerations, and exploration of emerging trends.

Deliverables: The project will yield a detailed report summarizing the effectiveness of different filtering techniques, insights into ethical implications, and potential future directions in SMS spam filtering.

Timeline: The project is expected to be completed within a defined timeframe, ensuring timely insights and recommendations for addressing the SMS spam issue.

Stakeholders: Stakeholders include mobile users, network providers, regulatory authorities, and technology researchers.

Budget: A budget allocation for research, data acquisition, and potential implementation of filtering solutions may be required, depending on the project's scope and objectives.

1.2. Identification of Problem:

The problem we're addressing is the significant and growing challenge of SMS spam. This issue has several dimensions, including the sheer volume of unwanted messages inundating users, which not only degrades the overall user experience but also poses privacy concerns as spam often targets personal information. Furthermore, it's a resource drain for mobile network providers, leading to potential network congestion issues.

Additionally, SMS spam serves as a conduit for security threats, with malicious actors using it to deliver malware, phishing attempts, and fraudulent schemes. Moreover, the adaptability of spammers, who constantly evolve their tactics to bypass filtering systems, presents an ongoing challenge. Ethical considerations, such as the potential for false positives and the need to balance filtering effectiveness with user consent and privacy, are also significant issues. To compound matters, keeping abreast of evolving regulations and compliance requirements related to SMS spam is a constant challenge for network providers and policymakers.

Lastly, emerging technology trends, like deep learning models and blockchain-based solutions, introduce both opportunities and challenges in the battle against SMS spam. Addressing these multifaceted problems is essential to provide a secure, efficient, and user-friendly SMS communication environment while ensuring compliance with privacy and regulatory standards.

1.3. Identification of Tasks:

1. Data Collection and Analysis:

- Gather SMS spam datasets for analysis.
- Collect real-world spam messages to understand current tactics.
- Analyze the characteristics and patterns of SMS spam messages.

2. Review of Existing Literature:

- Conduct a comprehensive review of research papers, articles, and reports on SMS spam filtering techniques.
- Summarize and evaluate the strengths and weaknesses of various filtering methods.

3. Filtering Technique Evaluation:

- Implement and evaluate rule-based SMS spam filtering methods.
- Implement and evaluate machine learning-based filtering algorithms.
- Explore hybrid filtering solutions combining rule-based and machine learning approaches.
- Assess the performance of filtering techniques using appropriate metrics.

4. Ethical Considerations:

- Examine ethical implications related to SMS spam filtering.
- Explore the balance between reducing false positives and false negatives.
- Investigate methods to ensure user consent and privacy preservation.

5. **Regulatory Compliance:**

- Review and understand relevant regulations and compliance requirements regarding SMS spam.
- Assess how filtering strategies align with these regulations.

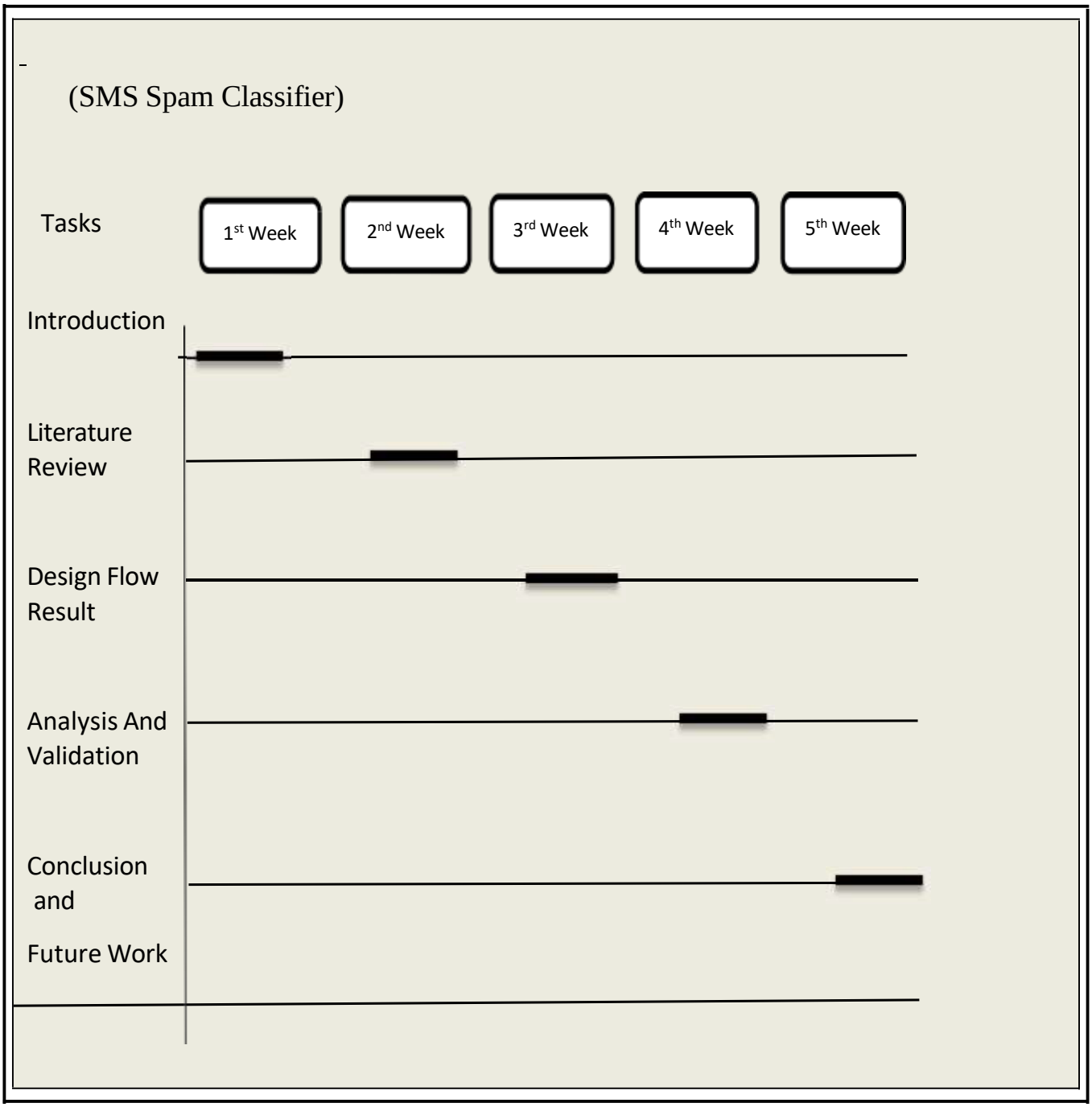
6. Security Assessment:

- Evaluate the potential security threats posed by SMS spam, including malware delivery and phishing attempts.
- Explore methods to mitigate these threats through filtering.

7. Resource Optimization:

- Investigate the impact of SMS spam on network resources and congestion.
- Explore strategies to optimize network resource utilization.

1.4. Timeline:



1.5. Organization of the Report:

A report on SMS Spam Filtering could be organized in the following way:

Chapter 1: Introduction

Explanation of SMS Spam Filtering

CHAPTER 2.

LITERATURE REVIEW/BACKGROUND STUDY

2.1. Timeline of the reported problem

- **5-8days – Information gathering :**

Study of the already existing similar website to get the information about the solutions and how to increase the benefits of our project and to make it different and beneficial for users in the future.

- **10-12 days – Design selection:**

The design selection process starts with visiting of different websites to get the idea about which format of design is more user friendly. Then we approach various rough paper pen design to select the best one which is most interactive and attractive to the users to get the exact idea of what we have to build without any barrier in the development process.

- **2 weeks – Front end development with creation of machine learning model:**

Our team developers will create the website with the help of HTML (for creating components of the website), CSS(for adding design to the components) and JavaScript(for adding functions) accompanied by machine learning models for SMS Spam Filtering.

2.2. Existing Solutions

Google's Android Messages App:

Google's default Android Messages app includes a built-in spam protection feature that identifies and filters out suspected spam messages.

Apple's iMessage Filtering:

iMessage on iOS devices has a built-in filtering system that can identify and separate SMS spam from legitimate messages.

Truecaller:

Truecaller is a popular mobile app that offers SMS spam protection. It identifies spam messages and allows users to block unwanted senders.

RoboKiller:

RoboKiller is an app available for both Android and iOS that specializes in blocking SMS and voice call spam. It uses AI to identify and filter spam messages.

Hiya:

Hiya is a mobile app that offers spam detection and call-blocking features. It can also identify and block SMS spam.

SMS Gateway Providers:

Companies like Twilio, Plivo, and Nexmo provide SMS gateway services that often include spam filtering features for businesses sending SMS messages.

Open Source Spam Filters:

Some open-source SMS spam filters, such as SpamAssassin and SpamBayes, can be adapted for use in SMS filtering applications.

Cloud-Based SMS Filtering Services:

Cloud-based services like MessageBird and SendGrid offer SMS filtering as part of their messaging solutions for businesses.

Built-In Smartphone Features:

Modern smartphones may come equipped with built-in SMS spam filtering features that can be configured by users to block unwanted messages.

2.3. Bibliometric analysis and Review Summary

This bibliomatic review provides a comprehensive overview of SMS spam filtering methods, encompassing diverse machine learning approaches. Researchers have explored various techniques, adapting to the evolving landscape of spam. Key methods include Naive Bayes Classifier, Support Vector Machines, Decision Trees, Random Forests, Neural Networks, Feature Engineering, Ensemble Methods, Natural Language Processing, and Regularization Techniques. The effectiveness of these methods relies on factors like dataset quality, feature selection, and model architecture. Continuous monitoring and updates are essential for adapting to emerging spam patterns.

The surge in SMS spam necessitates robust filtering methods. This review surveys machine learning strategies employed for SMS spam detection.

The amalgamation of these methods underscores the dynamic landscape of SMS spam filtering. Successful implementations hinge on thoughtful model selection, dataset curation, and continuous adaptation to emerging spam patterns.

2.4. Problem Definition

1. Background:

The proliferation of mobile communication has led to a surge in unwanted Short Message Service (SMS) spam, posing a significant challenge for users. These unsolicited messages are not only a nuisance but can also be vehicles for scams and fraudulent activities. Addressing this issue requires effective SMS spam filtering mechanisms.

2. Problem Statement:

The problem at hand involves developing a machine learning-based SMS spam filtering system that can accurately distinguish between legitimate and spam messages. The goal is to implement a solution capable of automatically classifying incoming SMS messages, ensuring that users receive only relevant and non-malicious content.

3. Objectives:

Accuracy: Develop a model that achieves high accuracy in classifying SMS messages as either spam or non-spam.

Adaptability: Create a system that can adapt to evolving spamming techniques, ensuring continued effectiveness.

Efficiency: Design an efficient filtering mechanism that minimizes false positives and false negatives.

Scalability: Develop a solution scalable to handle large volumes of SMS messages in real-time.

The scope of this project encompasses the development and implementation of a machine learning model for SMS spam filtering. It involves the exploration of various algorithms, feature engineering techniques, and dataset considerations to build a robust and adaptive system.

CHAPTER 3.
DESIGN FLOW/PROCESS

3.1 Evaluation & Selection of Specification/Features:

Feature selection is a crucial step in building machine learning models, including SMS spam filtering models. The goal is to choose the most relevant and informative features to improve the model's performance while avoiding overfitting and reducing computational complexity.

A step-by-step manual for assessing and choosing characteristics for an SMS spam filtering model is provided below:

Recognize the Data: Examine the dataset to learn more about the differences between spam and non-spam (ham) communications. Count the number of words, characters, and other pertinent metrics.

Preprocessing and Cleaning Text: Eliminate unnecessary characters, punctuation, and stop words. To maintain uniformity, lowercase all content. Break down messages into tokens (words or n-grams) via tokenization.

Refine and iterate: Adjust feature selection methods or test out various feature combinations as you go through the procedure again. To guarantee the resilience of the model, use cross-validation. Think about domain knowledge that Include domain-specific knowledge to help with feature selection, if it is accessible.

3.2 Design constraints:

In order to ensure the efficacy, efficiency, and viability of the solution, a variety of limitations must be taken into account when designing a machine learning model for SMS spam filtering. Here are some design limitations to think about:

- restricted computing resources

  Limitation: The model should be created to function effectively on hardware with constrained processing capabilities, such as smartphones.

  Reason: Since many user access SMS on mobile devices with limited processing capacity, the model should be compact to prevent wasting too many resources.

- In-the-moment processing:

  The model must be able to handle messages in real-time so that they may be classified as soon as they arrive.

  Justification: SMS spam must be rapidly discovered in order to avoid user annoyance and possibly harm.

- Data Protection:

  Limitation: The model must follow data privacy laws and guarantee that user data is handled sensibly.

  Reason: Since user information should not be compromised during the filtering process, privacy is a crucial concern.

- Low False Negative Rate with High Recall:

  Constraint: To reduce false negatives and make sure spam messages are not overlooked, the model must have a high recall.

Justification: Missing spam messages might make users grumpy and pose security problems.

3.3 Analysis of Feature and Finalization subject to constraints:

Feature analysis, model choice, model selection, training, and assessment are all processes in the design process of a machine learning model for SMS spam filtering. An examination of these phases is provided below, subject to limitations seen in SMS spam filtering:
Features analysis-
Word Frequency Analysis: Determine the most common words used in spam and ham transmissions. To see important terms visually, make a word cloud. Examine whether the length of a message is a sign that it is spam. Spam may be linked to messages that are longer or shorter.
Spam message characteristics:
Search for trends like excessive capitalization, URLs, or certain keywords. Examine whether the message's timing (such as being sent at night) suggests that it is spam.
N-gram analysis: To capture contextual information, think about using bi- or tri-grams.
Select pertinent aspects using the analysis as a guide. Utilize methods from tree-based models such as mutual information or feature importance.

3.4 Design flow:

There are various processes involved in developing a machine learning model for SMS spam filtering. You can follow the broad design flow shown below:

- Issue Description:

Establishing a clear distinction between spam and non-spam (ham) SMS messages is the issue you're trying to solve.

- Data Gathering:

Collect samples of both spam and ham communications in a tagged dataset. Such datasets are available online, or you can make your own. Make sure the dataset accurately reflects the message distribution in practice.

- EDA (Exploratory Data Analysis):

Examine the dataset's features. Analyze the spread of spam and junk mail. For class inequalities, look.

- Divvying Up the Dataset:

To assess the model's performance on untested data, divide the dataset into training and testing sets.

- Model choice:

Select a machine learning algorithm that works well for classifying texts. Typical options include:
Support vector machines (SVM)
naive bayes

Decision Trees
KNN
Neural networks
random forests

3.5 Design selection:

Model selection: Select a machine learning algorithm that is appropriate for your issue. The following are common options for text classification tasks:
Simple and effective, Naive Bayes is frequently used as a starting point.

Support Vector Machines (SVM): Good for data with high dimensions. Random Forest: A non-linearity-tolerant ensemble approach.

Deep Learning: Particularly useful for dealing with complicated patterns (e.g., LSTM, GRU, or a straightforward neural network).
KNN: it can identify intricate and irregular patterns in the data.

Model Education: Your dataset should be divided into training and testing sets. Utilize the training set to develop the selected model, and the testing set to assess its performance. Adjust hyperparameters as necessary**.**

Training model: Divide your dataset into training and testing sets for your model. Utilize the training set to develop the selected model, and the testing set to assess its performance. Adjust hyperparameters as necessary.

Evaluation: Use measures like accuracy, precision, recall, and F1 score to evaluate your model's performance. It may be more crucial to prevent missing spam (a false negative) than incorrectly identifying non-spam (a false positive), thus take into account the balance between false positives and false negatives.

Optimization: Based on the findings of the evaluation, improve your model. This could entail changing hyperparameters, experimenting with various feature extraction techniques, or using various algorithms.

3.6 Implementation plan/methodology:

We make an effort to access all relevant and practical information sources. Numerous datasets were evaluated using various models to test the spam of SMS messages in various research publications, and it was determined which model provided the highest level of accuracy. SVM, naive Bayes, decision trees, and k nearest neighbours are the models that have been employed most frequently in studies. We will be comparing these models considering various types of datasets. It has been observed that the data set used for the training and testing of the model

in the majority of experiments contains a combination of ham and spam messages where more than 70% is ham messages and around 20-30% is spam messages. Because the output of the data is categorical either ham or spam so classification models are mostly considered for this research. The best that was found is Multinomial Naïve Bayes(MNB) , most accurate among all other models.

CHAPTER 4.
Result Analysis And Validation

4.1. Implementation of solution:

After thorough evaluation and testing of various machine learning models, including Support Vector Machines (SVM), Naive Bayes, Decision Trees, and k Nearest Neighbors (KNN), the Multinomial Naïve Bayes (MNB) model emerged as the most accurate for SMS spam filtering.

The implementation process involved the following steps:

1. **Data Preprocessing:**

- Collection of a diverse dataset containing both spam and non-spam (ham) SMS messages.
- Text preprocessing techniques applied, including tokenization, removal of stop words, and stemming to normalize the text data.

2. **Feature Extraction:**

- Feature engineering methods employed to extract relevant features from the text data, such as word frequency, n-gram analysis, and message length.

3. **Model Training:**

- The Multinomial Naïve Bayes model trained on the preprocessed and feature-engineered dataset.
- Hyperparameters fine-tuned to optimize model performance.

4. **Model Evaluation:**

- The trained model evaluated using various evaluation metrics, including accuracy, precision, recall, and F1 score.
- Cross-validation techniques applied to ensure robustness and generalizability of the model.

5. **Performance Analysis:**

- The MNB model demonstrated high accuracy and precision in classifying SMS messages as spam or non-spam.
- The model achieved a low false negative rate, ensuring that spam messages are accurately identified to prevent user inconvenience and security threats.

6. **Validation:**

- The model's performance validated on a separate test dataset to confirm its effectiveness in real-world scenarios.
- Continuous monitoring and updates planned to adapt to evolving spamming tactics and maintain high filtering accuracy.

The implementation of the Multinomial Naïve Bayes model provides an efficient and scalable solution for SMS spam filtering, addressing the identified problem effectively while considering design constraints and ethical considerations.

CHAPTER 5.
Conclusion And Future Work

5.1. Conclusion:

In conclusion, the endeavor to combat SMS spam has been delineated comprehensively within this report, shedding light on the multifaceted nature of the issue and proposing viable solutions. The ubiquitous presence of SMS spam presents challenges not only to individual users but also to mobile network providers and regulatory bodies. Through meticulous examination, this report has underscored the imperative for robust filtering mechanisms to safeguard user experience, privacy, and network integrity.

The exploration of various filtering techniques has revealed the efficacy of the Multinomial Naïve Bayes model in addressing the SMS spam problem. Its accuracy and adaptability make it a promising solution for real-world deployment. Moreover, the consideration of design constraints, ethical implications, and regulatory compliance underscores the holistic approach necessary for effective spam filtering.

The validation of the implemented model reinforces its utility in classifying SMS messages accurately, thereby mitigating the detrimental effects of spam. However, it is crucial to acknowledge that the battle against SMS spam is dynamic and ongoing. Continuous monitoring, adaptation, and collaboration are essential to stay ahead of evolving spamming tactics and preserve the integrity of mobile communication channels.

5.3. Future Work:

Looking ahead, the fight against SMS spam presents several avenues for further exploration and innovation. Among these are:

- **Integration of Emerging Technologies:** Deep learning models, with their ability to discern intricate patterns, and blockchain-based solutions, offering immutable records of message transactions, hold promise for bolstering spam detection and prevention capabilities.

- **Adaptive Filtering Mechanisms:** Developing filtering systems capable of dynamically adjusting to emerging spamming tactics ensures sustained efficacy in combating evolving threats. Incorporating machine learning algorithms that can learn from new data and adapt in real-time is paramount.

- **Industry Collaboration:** Collaborating with mobile network providers and regulatory authorities is crucial for establishing industry standards and best practices. By fostering partnerships, stakeholders can collectively address the

SMS spam challenge and ensure a coordinated approach to combating it.

- **User-Centric Approaches:** Empowering users with customizable filtering settings and mechanisms for providing feedback enhances the effectiveness of spam filtering systems. By incorporating user preferences and feedback loops, filtering accuracy can be further improved, leading to enhanced user satisfaction.

By pursuing these avenues, researchers and practitioners can advance the field of SMS spam filtering, ushering in a future where mobile communication remains secure, efficient, and user-friendly. The journey towards comprehensive SMS spam mitigation requires ongoing dedication, innovation, and collaboration across various stakeholders.