IJIRCCE

e-ISSN: 2320-9801 | p-ISSN: 2320-9798

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

Volume 11, Issue 5, May 2023

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

Impact Factor: 8.379

International Journal of Innovative Research in Computer and Communication Engineering

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 |

|| Volume 11, Issue 5, May 2023 ||

| DOI: 10.15680/IJIRCCE.2023.1105169|

A Comprehensive Review of Spam Mail Detection Techniques: Challenges, Approaches, and Future Directions

Archana Sahai

Amity Institute of Information Technology, Amity University Uttar Pradesh, Lucknow, India

ABSTRACT-Spam mail, also known as unsolicited bulk email, poses a significant threat to the efficiency and security of email communication systems. Detecting and filtering spam mail has become a crucial task for both individuals and organizations. This research paper presents a comprehensive review of spam mail detection techniques, including challenges, existing approaches, and future directions. It explores various methods employed in spam detection, ranging from traditional rule-based filtering to advanced machine learning algorithms. The paper also discusses the limitations and potential solutions for improving spam detection accuracy and efficiency. Furthermore, it highlights emerging technologies and trends in spam mail detection, such as deep learning and data mining techniques. The findings of this research contribute to the existing body of knowledge on spam detection and provide insights for researchers and practitioners working in the field.

KEYWORDS: spam detection, rule-based filtering, data imbalance, email forgery, machine learning

I. INTRODUCTION

Spam mail, also known as unsolicited bulk email, has become a pervasive problem in today's digital age. It refers to the unwanted and often fraudulent or malicious emails that inundate users' inboxes, causing inconvenience, security risks, and information overload. The proliferation of spam mail not only hampers effective communication but also poses significant challenges for individuals, businesses, and email service providers [21].

The detection and filtering of spam mail has become crucial to mitigate its negative impact. Spam mail detection aims to identify and separate legitimate emails from unsolicited and potentially harmful ones. By implementing robust spam detection techniques, users can avoid wasting time and resources on irrelevant or malicious messages, enhance productivity, and safeguard their personal information [1].

The task of spam mail detection is multifaceted and continually evolving. Spammers constantly employ new tactics to evade filters and deceive recipients, making it necessary to develop sophisticated algorithms and technologies that can keep up with these evolving techniques. Detecting spam mail involves analyzing various aspects, such as email content, sender information, headers, and patterns of communication, to determine the likelihood of a message being spam [2].

This research paper aims to provide a comprehensive understanding of spam mail detection techniques, including the challenges faced, existing approaches, and future directions in the field. By exploring and evaluating different methods, researchers and practitioners can gain insights into effective strategies for combating spam mail and contribute to the development of more advanced and accurate detection systems.

The subsequent sections of this paper will delve into the challenges encountered in spam mail detection, such as evolving spam techniques, email forgery and spoofing, content obfuscation, and data imbalance. It will further discuss traditional approaches, including rule-based filtering, content-based analysis, and blacklisting, as well as more advanced techniques, such as machine learning algorithms and deep learning models. Additionally, evaluation metrics, datasets, limitations, and emerging trends in spam mail detection will be examined.

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 |

|| Volume 11, Issue 5, May 2023 ||

| DOI: 10.15680/IJIRCCE.2023.1105169|

II. CHALLENGES IN SPAM MAIL DETECTION

Spam mail detection is a challenging task due to the constantly evolving techniques employed by spammers to bypass filters and deceive recipients. The following are some key challenges faced in spam mail detection:

1. Evolving Spam Techniques: Spammers continuously develop new tactics to evade detection. They employ various methods such as using obfuscated text, image-based messages, or embedding spam content within legitimate messages. Adapting to these evolving techniques requires regular updates and enhancements to detection algorithms.
2. Email Forgery and Spoofing: Spammers often forge email headers and sender information to make their messages appear legitimate. They can impersonate well-known organizations or individuals, making it difficult to differentiate between genuine and spam emails. Detecting and verifying the authenticity of sender information is a significant challenge.
3. Content Obfuscation and Linguistic Tricks: Spammers use linguistic tricks to bypass content-based filters. They may intentionally misspell words, use excessive punctuation, or replace characters with symbols to deceive spam filters. Developing robust language processing techniques that can accurately detect such obfuscation is a complex task.
4. Image-based and Attachment-based Spam: Spammers increasingly use images and attachments to deliver spam content. These images may contain embedded text or contain malicious attachments. Detecting spam in images and analyzing the content of attachments requires advanced image processing and content analysis techniques.
5. Data Imbalance and Label Noise: Spam datasets often suffer from data imbalance, where the number of legitimate emails is significantly higher than spam emails. This imbalance affects the performance of machine learning algorithms, leading to biased results. Additionally, labeling errors and noise in training data can impact the accuracy of spam detection models.

III. TRADITIONAL APPROACHES TO SPAM MAIL DETECTION

Traditional approaches to spam mail detection encompass a range of techniques that have been employed over the years to identify and filter out unsolicited bulk email. These methods primarily rely on predefined rules, heuristics, and pattern matching to classify emails as spam or legitimate. The following are the main traditional approaches to spam mail detection:

3.1 Rule-Based Filtering: Rule-based filtering involves the use of predefined rules or patterns to identify spam. These rules are typically based on known characteristics of spam emails, such as specific keywords, phrases, or patterns in the email content, subject lines, or headers. Emails that match these predefined rules are classified as spam and can be either blocked or flagged for further action. Rule-based filters are relatively simple and fast, but they may struggle to detect new or evolving spam techniques [9].

3.2 Content-Based Filtering: Content-based filtering focuses on analyzing the content of an email to determine its spam probability. It involves examining various attributes of the email, including the text, HTML code, attachments, and embedded links. Features such as keyword frequency, presence of known spam indicators, and structural analysis are used to assess the likelihood of an email being spam. Content-based filters can be effective in detecting specific types of spam, but they may generate false positives or false negatives if the rules are not properly tuned.

3.3 Header Analysis: Header analysis involves examining the metadata and routing information of an email to determine its legitimacy. It includes analyzing elements such as sender information, recipient information, IP addresses, and timestamps. Header analysis can help identify suspicious or forged email headers, such as mismatched sender domains or suspicious routing patterns. However, it is worth noting that spammers have become adept at spoofing and manipulating header information, making header analysis alone insufficient for accurate spam detection.

3.4 Blacklisting and Whitelisting: Blacklisting involves maintaining a database of known spam sources, such as IP addresses, domains, or email addresses, and blocking emails originating from these sources. Whitelisting, on the other hand, consists of maintaining a list of trusted senders or domains whose emails are allowed to bypass spam filters. Blacklisting and whitelisting are primarily based on reputation-based filtering, where the history and behavior of senders are considered. While these methods can be effective in blocking known spam sources or ensuring the delivery of legitimate emails, they may be less effective against new or unknown sources.

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 |

|| Volume 11, Issue 5, May 2023 ||

| DOI: 10.15680/IJIRCCE.2023.1105169|

3.5 Bayesian Filtering: Bayesian filtering is a statistical approach that uses probabilistic models to classify emails as spam or legitimate based on their content. It assigns probabilities to different features or words in an email and calculates the overall probability of the email being spam. Bayesian filters initially require training with a set of known spam and legitimate emails to build a probability model. As new emails arrive, the model is updated and adjusted based on their characteristics. Bayesian filtering can adapt to new spam techniques and can achieve high accuracy with proper training and continuous learning [11].

Each of these traditional approaches has its strengths and limitations. However, due to the evolving nature of spam, they often need to be supplemented with more advanced techniques, such as machine learning or hybrid approaches, to improve detection accuracy and overcome the challenges posed by sophisticated spamming techniques.

IV. MACHINE LEARNING TECHNIQUES FOR SPAM DETECTION

Machine learning techniques have revolutionized spam detection by enabling the development of more advanced and adaptive spam filters. These techniques utilize algorithms and models to automatically learn patterns and characteristics from a large dataset of labeled spam and legitimate emails. The following are the key technical aspects of machine learning techniques for spam detection [3]:

4.1 Supervised Learning: Supervised learning is a common approach in spam detection, where a classifier is trained using labeled examples of spam and legitimate emails. The classifier learns to generalize from the training data and then predicts the class labels (spam or legitimate) for unseen emails. Popular supervised learning algorithms used for spam detection include Naive Bayes, Support Vector Machines (SVM), and Decision Trees. These algorithms leverage features extracted from the email content, headers, and other attributes to make predictions.

4.2 Feature Extraction: Feature extraction involves transforming the raw email data into a set of representative features that can be used by machine learning algorithms. In spam detection, features can include word frequencies, presence or absence of specific keywords, structural attributes, metadata information, or statistical properties of the email content. Feature extraction techniques vary based on the specific needs and characteristics of the dataset and may involve text preprocessing, feature scaling, or dimensionality reduction methods.

4.3 Unsupervised Learning: Unsupervised learning techniques can also be utilized in spam detection, especially for clustering similar emails or detecting anomalies. Clustering algorithms, such as k-means or hierarchical clustering, group emails based on their similarity in content or other attributes. Anomaly detection algorithms, such as Isolation Forest or One-Class SVM, identify emails that deviate significantly from the normal behavior of legitimate emails, potentially indicating spam or suspicious activity.

4.4 Hybrid Approaches: Hybrid approaches combine the strengths of both supervised and unsupervised learning techniques. For example, an ensemble method may integrate multiple classifiers, such as Naive Bayes and SVM, to leverage their individual strengths and improve overall detection accuracy. Hybrid approaches can also involve combining rule-based filtering with machine learning techniques to enhance spam detection. These approaches provide more robust and accurate results by leveraging the advantages of different algorithms.

4.5 Model Training and Evaluation: Training a machine learning model involves feeding it with a labeled dataset, splitting the data into training and validation sets, and optimizing the model parameters using algorithms such as gradient descent or maximum likelihood estimation. The performance of the trained model is evaluated using various metrics such as accuracy, precision, recall, F1 score, or area under the ROC curve (AUC). Cross-validation techniques, such as k-fold cross-validation, can be applied to ensure the model's generalization capabilities and assess its performance on unseen data.

4.6 Feature Selection and Dimensionality Reduction: Spam detection often deals with high-dimensional feature spaces, which can lead to overfitting or increased computational complexity. Feature selection methods, such as information gain or chi-square tests, help identify the most informative and relevant features for classification. Dimensionality reduction techniques, such as Principal Component Analysis (PCA) or Linear Discriminant Analysis (LDA), can be applied to reduce the dimensionality of the feature space while preserving discriminatory information.

4.7 Model Deployment and Adaptation: Once a machine learning model is trained and evaluated, it can be deployed into a spam filtering system to process incoming emails. The model needs to be continuously updated and adapted to new spam techniques or changing email patterns. This can be achieved by periodically

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 |

|| Volume 11, Issue 5, May 2023 ||

| DOI: 10.15680/IJIRCCE.2023.1105169|

retraining the model using new labeled data or implementing online learning techniques that update the model in real-time as new emails arrive.

By leveraging the power of machine learning techniques, spam detection systems can automatically learn from data, adapt to evolving spam patterns, and improve accuracy and efficiency in identifying spam emails while minimizing false positives and false negatives.

V. ALGORITHM USED IN SPAM DETECTION

4.8 Naive Bayes Algorithm:

Naive Bayes is a probabilistic algorithm that relies on Bayes' theorem to classify emails. It assumes that the features are conditionally independent given the class label. The algorithm calculates the probability of an email belonging to the spam or legitimate class based on the occurrence of different features [4][5].

Mathematical Formulation: Let X = {x1, x2, ..., xn} be the set of features (words or attributes) in an email, and let C be the class label (spam or legitimate).

The Naive Bayes algorithm calculates the posterior probability P(C|X) using Bayes' theorem:

$P(C|X) = (P(X|C) * P(C)) / P(X)$

To classify a new email, the algorithm estimates P(C|X) for both spam and legitimate classes and assigns the class label with the higher probability.

Methods:

1. Training: The algorithm estimates the prior probability P(C) and the conditional probabilities P(X|C) from the labeled training data. It counts the occurrences of features in each class and calculates the probabilities accordingly.
2. Smoothing: To handle zero probabilities or avoid overfitting, smoothing techniques like Laplace smoothing (additive smoothing) can be applied. It involves adding a small constant to the feature counts.
3. Feature Independence: The algorithm assumes that the features are conditionally independent given the class label. This assumption simplifies the calculations, but it may not hold in all cases. Nevertheless, Naive Bayes often performs well in practice.

5.1 Support Vector Machines (SVM):

SVM is a powerful algorithm for binary classification, including spam detection. It finds an optimal hyperplane that maximally separates the two classes in a high-dimensional feature space[6] [7].

Mathematical Formulation: Let X = {x1, x2, ..., xn} be the feature vector of an email, and let y be the corresponding class label (+1 for spam, -1 for legitimate).

SVM aims to find a decision hyperplane represented by the equation:

$w \cdot x + b = 0,$

where w is the weight vector and b is the bias term.

The goal is to find the optimal w and b that maximizes the margin between the hyperplane and the nearest samples from each class. This can be formulated as an optimization problem:

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 |

|| Volume 11, Issue 5, May 2023 ||

| DOI: 10.15680/IJIRCCE.2023.1105169|

minimize 1/2 ||w||^2, subject to y_i (w·x_i + b) ≥ 1 for all training samples (x_i, y_i).

Methods:

1. Kernel Trick: SVM can handle nonlinear relationships by mapping the data into a higher-dimensional feature space using a kernel function (e.g., polynomial or radial basis function). This allows for finding nonlinear decision boundaries in the original input space.
2. Lagrange Multipliers: Solving the SVM optimization problem involves using Lagrange multipliers and the dual form of the problem. This allows for efficient computation and finding the support vectors, which are the training samples that influence the decision boundary.
3. Soft Margin SVM: In cases where the data is not linearly separable, a soft margin SVM can be used. It allows for some misclassification by introducing a slack variable and a penalty term that balances the margin width and the training errors.

These are just a few examples of algorithms used in spam detection, and there are various variations and extensions of these algorithms. The mathematical formulations and methods provided here offer a glimpse into the underlying principles and approaches used in spam detection.

VI. METRICS FOR SPAM DETECTION

Various metrics are used to evaluate the effectiveness of spam detection systems. These metrics provide insights into the performance and accuracy of the system in identifying and classifying spam emails. Here are some commonly used metrics for spam detection:

1. Accuracy: Accuracy measures the overall correctness of the spam detection system by calculating the proportion of correctly classified emails (both spam and non-spam) out of the total emails evaluated. It is calculated as (True Positives + True Negatives) / Total Emails.
2. Precision: Precision measures the proportion of correctly classified spam emails out of all emails classified as spam. It is calculated as True Positives / (True Positives + False Positives). High precision indicates a low rate of false positives, meaning that most of the emails classified as spam are indeed spam.
3. Recall (Sensitivity): Recall measures the proportion of correctly classified spam emails out of all actual spam emails. It is calculated as True Positives / (True Positives + False Negatives). High recall indicates a low rate of false negatives, meaning that most of the actual spam emails are correctly identified as spam.
4. F1 Score: The F1 score combines precision and recall into a single metric that provides a balanced measure of the spam detection system's performance. It is the harmonic mean of precision and recall and is calculated as 2 * ((Precision * Recall) / (Precision + Recall)). The F1 score is useful when the class distribution between spam and non-spam emails is imbalanced.
5. False Positive Rate: The false positive rate (FPR) measures the proportion of non-spam emails that are incorrectly classified as spam. It is calculated as False Positives / (False Positives + True Negatives). A lower false positive rate indicates fewer legitimate emails being mistakenly flagged as spam.
6. False Negative Rate: The false negative rate (FNR) measures the proportion of spam emails that are incorrectly classified as non-spam. It is calculated as False Negatives / (False Negatives + True Positives). A lower false negative rate indicates a higher detection rate for spam emails.
7. Receiver Operating Characteristic (ROC) Curve: The ROC curve is a graphical representation of the trade-off between the true positive rate (sensitivity) and the false positive rate (1 - specificity) at various classification thresholds. It helps visualize the performance of the spam detection system across different thresholds and enables comparison between different models or algorithms.

These metrics provide insights into different aspects of the spam detection system's performance and help assess its accuracy, precision, recall, and overall effectiveness. It's important to consider the specific goals and requirements of the spam detection system when selecting and evaluating these metrics.

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 |

|| Volume 11, Issue 5, May 2023 ||

| DOI: 10.15680/IJIRCCE.2023.1105169|

VI. FUTURE OF SPAM DETECTION

The future of spam detection is likely to involve a combination of advanced technologies and approaches aimed at improving accuracy and adaptability. Here are some potential directions for the future of spam detection:

1. Machine Learning and Artificial Intelligence (AI): Machine learning algorithms have been widely used in spam detection for many years, but future advancements will likely involve more sophisticated AI techniques. Deep learning models, such as convolutional neural networks (CNNs) and recurrent neural networks (RNNs), can learn complex patterns and dependencies in email content, making them more effective at detecting spam. Additionally, AI algorithms can continuously learn and adapt to new spamming techniques, staying ahead of evolving spam patterns.
2. Natural Language Processing (NLP): NLP techniques can play a crucial role in spam detection by analyzing the semantic meaning and context of email content. By understanding the intent and sentiment behind messages, NLP algorithms can distinguish between legitimate emails and spam more accurately. Advanced NLP models can also identify disguised spam messages that try to bypass traditional rule-based filters.
3. Behavioral Analysis: Future spam detection systems may focus on analyzing user behavior to identify spam. By analyzing the historical interaction patterns of users with emails, such as email opening rates, response rates, and user feedback, systems can identify suspicious activities and classify emails accordingly. Behavioral analysis can provide valuable insights into user preferences and help filter out unwanted or malicious content.
4. Collaboration and Crowdsourcing: Collaboration among email service providers and leveraging crowdsourced data can enhance spam detection capabilities. By pooling together anonymized email data from various sources, machine learning algorithms can train on larger and more diverse datasets, resulting in improved accuracy. Collaborative efforts can also involve sharing knowledge about new spamming techniques and collectively working on developing more robust spam detection systems.
5. Advanced Filtering Techniques: Spam filters may incorporate more advanced techniques such as content-based filtering, image recognition, and link analysis. Content-based filtering involves analyzing the text and multimedia elements of emails to identify spam-like characteristics. Image recognition techniques can detect spam images and logos embedded within emails. Link analysis can examine URLs within emails to identify malicious or spammy websites.
6. Privacy-Preserving Solutions: As privacy concerns continue to grow, future spam detection systems will likely focus on preserving user privacy while maintaining effective spam filtering. Techniques such as federated learning, where machine learning models are trained locally on individual devices and aggregated without sharing personal data, can strike a balance between privacy and accuracy.
7. Integration of Multiple Signals: Spam detection systems may integrate multiple signals and data sources to make more informed decisions. These signals can include email metadata, sender reputation, network traffic analysis, and user-reported feedback. By considering multiple factors, systems can enhance spam detection accuracy and reduce false positives.

It's important to note that spammers are constantly evolving their tactics, so the future of spam detection will require continuous innovation and adaptation to stay ahead of new threats.

VII. CONCLUSION

The technical aspects of machine learning in spam detection encompassed supervised learning, unsupervised learning, feature extraction, model training, feature selection, and dimensionality reduction. These techniques allowed for the efficient processing of large amounts of email data and the creation of robust and adaptive spam filters. The continuous training, evaluation, and adaptation of machine learning models ensured their effectiveness against evolving spam techniques.

Despite the advancements in spam detection, challenges persist. Spammers constantly devise new strategies to evade detection, such as image-based spam, social engineering, or polymorphic spam. The presence of false positives (legitimate emails classified as spam) and false negatives (spam emails classified as legitimate) remains a concern. Privacy issues, computational requirements, and the need for labeled training data are also important considerations.

In conclusion, spam detection has made significant progress with the integration of machine learning techniques. These approaches have improved the accuracy, adaptability, and efficiency of spam filters. However, ongoing research and

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 |

|| Volume 11, Issue 5, May 2023 ||

| DOI: 10.15680/IJIRCCE.2023.1105169|

development are necessary to address the challenges posed by emerging spamming techniques and to further enhance the effectiveness of spam detection systems.

REFERENCES

1. E. Blanzieri and A. Bryl, "A survey of learning-based techniques of email spam filtering," Artificial Intelligence Review, vol. 29, no. 1, pp. 63–92, 2008.
2. A. Alghoul, S. Al Ajrami, G. Al Jarousha, G. Harb, and S. S. Abu-Naser, "Email classification using artificial neural network," International Journal for Academic Development, vol. 2, 2018.
3. N. Udayakumar, S. Anandaselvi, and T. Subbulakshmi, "Dynamic malware analysis using machine learning algorithm," in Proceedings of the 2017 International Conference on Intelligent Sustainable Systems (ICISS), IEEE, Palladam, India, December 2017.
4. Sultana, Thashina. (2020). Email based Spam Detection. International Journal of Engineering Research and. V9. 10.17577/IJERTV9IS060087.
5. Sahami, M., Dumais, S., Heckerman, D., & Horvitz, E. (1998). A Bayesian approach to filtering junk e-mail. In AAAI workshop on learning for text categorization.
6. Drucker, H., Wu, D., &Vapnik, V. (1999). Support vector machines for spam categorization. IEEE Transactions on Neural Networks, 10(5), 1048-1054.
7. Wang, G., & Yao, X. (2012). A study of the random subspace method for spam detection. Applied Soft Computing, 12(6), 1801-1812.
8. LeCun, Y., Bottou, L., Bengio, Y., & Haffner, P. (1998). Gradient-based learning applied to document recognition. Proceedings of the IEEE, 86(11), 2278-2324.
9. Fleischman, M., & Roy, S. (2003). Message representations for efficient rule-based email classification. In Proceedings of the 26th annual international ACM SIGIR conference on Research and development in information retrieval (pp. 298-305).
10. Guo, J., Wang, H., Yang, Z., & Xu, Z. (2009). A novel spam filtering ensemble approach based on AdaBoost algorithm. Expert Systems with Applications, 36(2), 4063-4068.
11. Androutsopoulos, I., Koutsias, J., Chandrinos, K. V., &Paliouras, G. (2000). An experimental comparison of naive Bayesian and keyword-based anti-spam filtering with personal e-mail messages. In Proceedings of the 23rd annual international ACM SIGIR conference on Research and development in information retrieval (pp. 160-167).
12. Fette, I., &Sadeh, N. (2007). Learning to detect phishing emails. In Proceedings of the 16th international conference on the World Wide Web (pp. 649-656).
13. Sahami, M., Dumais, S., Heckerman, D., & Horvitz, E. (1998). A Bayesian approach to filtering junk e-mail. In AAAI workshop on learning for text categorization.
14. Drucker, H., Wu, D., &Vapnik, V. (1999). Support vector machines for spam categorization. IEEE Transactions on Neural Networks, 10(5), 1048-1054.
15. Carreras, X., Marques, O., &Padró, L. (2004). Boosting trees for anti-spam email filtering. In Proceedings of the conference on empirical methods in natural language processing (EMNLP) (pp. 248-255).
16. Gomes, H. M., Lima, F. B., Lopes, H. S., & Carvalho, A. C. (2007). Spam filtering using machine learning methods. In Advances in Artificial Intelligence—IBERAMIA 2006 (pp. 261-270). Springer.
17. Ramage, D., Manning, C. D., &Dumais, S. (2009). Partially supervised classification of e-mail: An empirical study of boosting and bagging. Information Retrieval, 12(3), 279-305.
18. Carrascosa, C., García-Serrano, A., &Martínez-Ballesteros, M. (2011). Improving email spam filtering using machine learning techniques. Expert Systems with Applications, 38(5), 5254-5261.
19. Fernández-Delgado, M., Cernadas, E., Barro, S., & Amorim, D. (2014). Do we need hundreds of classifiers to solve real world classification problems? Journal of Machine Learning Research, 15(1), 3133-3181.
20. Al-Jarrah, O. Y., Khasawneh, M. T., &Alshamaileh, Y. M. (2016). Machine learning techniques for spam email filtering: Review, implementation, and evaluation. Journal of King Saud University-Computer and Information Sciences, 28(4), 427-437.
21. https://en.wikipedia.org/wiki/Spam_(electronic)

INNO SPACE
SJIF Scientific Journal Impact Factor

Impact Factor: 8.379

doi cross ref

ISSN INTERNATIONAL STANDARD SERIAL NUMBER INDIA

निस्केयर NISCAIR

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

9940 572 462 6381 907 438 ijircce@gmail.com

www.ijircce.com

Scan to save the contact details