EMAIL SPAM DETECTION WITH ML & NLP

Agenda

- Problem Statement
- Related Work
- Proposed Methodology
- Results
- Conclusions

Problem Statement

Email spam detection is the task of classifying emails as either "spam" or "ham" based on their content. Traditional rule-based spam filters often struggle to keep up with new tactics used by spammers, which makes them ineffective in the long run. By integrating machine learning and NLP techniques, we can build a detection system that is not only accurate but also adaptive, capable of learning and evolving as new spam methods arise.

Related Work

In the contemporary world, emails and messages are essential across all industries, including business and education. These communications fall into two categories: ham (legitimate) and spam (unwanted). Spam messages, or junk email, can harm users by consuming their time and resources and compromising sensitive data. As the volume of spam messages rapidly increases, email and IoT service providers face significant challenges in identifying and filtering spam. Spam filtering, utilizing machine learning and deep learning techniques such as Naive Bayes, decision trees, neural networks, and random forests, is one of the most effective solutions. This study provides a comprehensive survey of these machine learning methods, categorizing them based on their approaches and comparing them on metrics like accuracy, precision, and recall. [1]

Survey of Review Spam Detection Using Machine Learning Techniques" published in the Journal of Big Data focuses on the critical issue of detecting fake online reviews, which can mislead consumers. It surveys various machine learning techniques employed for review spam detection, with an emphasis on supervised learning methods that require labeled data. The scarcity of such labeled data presents a significant challenge. The paper underscores the potential of Big Data analytics for effectively addressing this issue due to the massive volume of online reviews. It also suggests that future research should aim at developing methods capable of handling Big Data while enhancing the accuracy of spam detection techniques. [2]

Related Work

Comparison of Machine Learning Techniques for Spam Detection" published in Multimedia Tools and Applications evaluates the performance of thirteen machine learning classifiers for spam email detection. The classifiers include Adaptive Boosting, Artificial Neural Networks, Decision Trees, K-Nearest Neighbors, and more. The study found that the Random Forest classifier performed best with an accuracy of 99.91% on the Spam Corpus dataset, while the Naïve Bayes classifier had lower accuracy at 87.63%. The research highlights the importance of selecting the right classifier for effective spam detection. [3]

The paper focuses on developing a spam classification model to distinguish between ham and spam emails using various machine learning classifiers. It evaluates classifiers like Bayesian, Naïve Bayes, SVM, J48, and their combinations with Adaboost using metrics such as accuracy (F-measure), False Positive Rate, and training time. The study finds that the SVM classifier, despite having a high training time, achieves the highest accuracy and the lowest false positive rate. The research emphasizes the effectiveness of the SVM classifier, making it the preferred choice for this application. [4]

References

- [1] Teja Nallamothu, Phani and Shais Khan, Mohd (2023) “Machine Learning for SPAM Detection”. Asian Journal of Advances in Research, 6 (1). pp. 167-179. http://eprint.subtopublish.com/id/eprint/3333/
- [2] Michael Crawford, Taghi M. Khoshgoftaar, Joseph D. Prusa, Aaron N. Richter & Hamzah AlNajada “Survey of Review Spam Detection Using Machine Learning Techniques” . https://link.springer.com/article/10.1186/S40537-015-0029-9%20
- [3] Argha Ghosh & A. Senthilrajan “Comparison of Machine Learning Techniques for Spam Detection”. https://link.springer.com/article/10.1007/s11042-023-14689-3
- [4] Shrawan Kumar Trivedi “A study of machine learning classifiers for spam detection”. https://ieeexplore.ieee.org/document/7743279

Proposed Methodology

- **Data Preparation**
  Clean and preprocess the email dataset, removing unnecessary characters and formatting the text for analysis.

- **Feature Selection**
  Extract meaningful features such as keywords, sender reputation, email structure, and word frequency patterns to distinguish spam from legitimate emails.

- **Model Training**
  NAIVE BAYES (NB), LOGISTIC REGRESSION (LR), K-NEAREST NEIGHBOR (KNN), DECISION TREE (DT), RANDOM FOREST (RF)

Proposed Methodology

- **Evaluation**
  Measure the performance of the models using metrics like accuracy, precision, recall, and F1-score.
- **Deployment**
  Implement the selected model in a real-time email filtering system that ensures efficient and accurate spam classification.
- **Model Updates**
  Continuously update the model with new email data to account for evolving spam tactics and maintain its effectiveness.

Results

Random Forest and **Naive Bayes** emerged as the most accurate models, suitable for robust spam detection.

Logistic Regression and **Decision Tree** provide solid alternatives, balancing interpretability and performance.

KNN is effective for smaller datasets but less efficient with larger or more complex data.

Algorithm	Accuracy
Naive Bayes (NB)	98%
Logistic Regression (LR)	96%
K-Nearest Neighbors (KNN)	91%
Decision Tree (DT)	96%
Random Forest (RF)	98%

Conclusion

Each model has its strengths in email spam detection. Naive Bayes is fast and works well with text data. Logistic Regression is simple and reliable for clear classifications. K-Nearest Neighbors uses proximity for decisions, making it effective in small datasets. Decision Trees are easy to interpret and good at finding spam patterns. Random Forest combines multiple trees for high accuracy. The best choice depends on the dataset and whether speed, accuracy, or simplicity is the priority, with ensemble models like Random Forest and Naive Bayes often performing best overall.

Thank you