

Capstone Project Proposal: Text Classification for Spam Detection

Problem Statement:

The proliferation of spam messages in communication channels poses a significant challenge, leading to user inconvenience, privacy concerns, and potential security threats. The aim of this project is to develop an efficient text classification model for spam detection, leveraging various techniques, including Bag of Words (BoW), TF-IDF, Word2Vec, and advanced BERT embeddings.

Context:

Spam messages are pervasive across email, messaging apps, and other communication platforms. Accurate detection is crucial to filter out unwanted content and enhance the overall user experience. The project targets creating a robust and versatile model that can generalize well to different types of spam messages.

Criteria for Success:

- Achieve high accuracy in distinguishing between ham and spam messages.
- Develop interpretable models for analysis and insights.
- Implement three different models for performance analysis.
- Compare and contrast the effectiveness of Naive Bayes, Random Forest, and Neural Network models.

Scope of Solution Space:

The project will focus on the development and comparison of three text classification models, each utilizing different approaches. The scope includes preprocessing and cleaning of the Kaggle dataset, implementing Naive Bayes with BoW and TF-IDF, Random Forest with Word2Vec, and a Single Layer Neural Network with BERT embeddings.

Constraints:

- Availability and quality of the Kaggle dataset.
- Computational resources for training and evaluating models.
- Ethical considerations regarding data privacy and handling.

Stakeholders:

- End-users who will benefit from reduced exposure to spam.
- Communication platform providers seeking improved spam filters.
- Researchers and data scientists exploring diverse text classification techniques.

Data Sources:

The project will utilize a Kaggle dataset containing labeled examples of spam and ham messages. The dataset will be preprocessed to extract relevant features and ensure compatibility with the chosen models.

Approach:

1. **Data Collection, Preprocessing and EDA:** Obtain and clean the Kaggle dataset, preparing it for training and perform EDA by creating basic visualizations.
2. **Model Development:**
 - Implement Naive Bayes using BoW and TF-IDF.
 - Train Random Forest classifier using Word2Vec.
 - Build a Single Layer Neural Network with BERT embeddings.
3. **Performance Analysis:**
 - Evaluate each model's performance using metrics such as accuracy, precision, recall, and F1 score.
 - Compare and contrast the strengths and weaknesses of each model.

Deliverables:

- Jupyter notebooks containing code for each model.
- A comprehensive report detailing the project methodology, results, and analysis.
- Visualizations illustrating the performance metrics and comparative analysis.

By undertaking this project, I aim to contribute valuable insights into the effectiveness of various text classification techniques for spam detection, providing actionable recommendations for real-world applications.