CAPSTONE PROJECT

SPAM CLASSIFICATION

Presented By: B.Sasidharan,
JAYA ENGINEERING COLLEGE,
3rd year-EIE.

edunet foundation

OUTLINE

- Problem Statement
- Proposed System/Solution
- System Development Approach
- Algorithm & Deployment
- Result
- Conclusion
- Future Scope
- References

PROBLEM STATEMENT

Develop a machine learning model to classify SMS messages as spam or non-spam . The goal is to accurately identify spam messages to prevent users from receiving unwanted or potentially harmful content. The model should be trained on a dataset of labeled SMS messages and evaluated based on its ability to correctly classify messages as spam or non-spam.

PROPOSED SOLUTION

1.Choose a suitable dataset with labeled SMS messages and understand its structure and label distribution.

2.Load the dataset and preprocess the text data, including removing unnecessary characters, converting text to lowercase, and tokenizing messages.

3.Utilize TF-IDF vectorization to convert text data into numerical features and split the dataset into training and testing sets.

4.Select a machine learning algorithm for text classification and train the model using TF-IDF vectorized features.

5.Evaluate model performance using metrics like accuracy, precision, recall, and F1-score, and fine-tune the model as needed

SYSTEM APPROACH

Building the proposed solution would involve a combinatiow of data processing, feature engineering, and machine learning. Here are the key system and library requirements:

System Requirements:

1. **Hardware**:

   A computer with sufficient processing power, preferably with multiple cores or a GPU for faster training of machine leaming models.

   Adequate RAM to handle the size of the dataset and computational requirements.

2. Software:

   An operating system compatible with the required machine learning libraries (e.g., Windows, Linux, macOS)

ALGORITHM & DEPLOYMENT

Algorithm Selection

Data Exploration:

Explore the spam classification dataset's structure, features, and target variable(s)

Identify potential patterns, correlations, and outliers

Problem Formulation:

Define the problem: Predict spam message coming time, ideal timing per day, and likelihood of special requests hased on historical data

Algorithm Selection:

Regreson tasks (e.g., predicting daily rates):
Consider lineat regression, decision trees, or ensemble methods (XGBoost. LightGBM
Classification tasks (eg, predicting special requests):
Consider logistic regression, decision trees, or random forests.

RESULT

1.Accuracy:

1. The accuracy score indicates the percentage of correctly classified instances (spam or ham) out of the total number of instances in the test set.

2.Classification Report:

1. The classification report provides detailed metrics such as precision, recall, F1-score, and support for each class (spam and ham).
2. Precision represents the ratio of true positive predictions to the total predicted positives.
3. Recall represents the ratio of true positive predictions to the total actual positives.
4. F1-score is the harmonic mean of precision and recall, providing a balanced measure of the model's performance.

3.Confusion Matrix:

1. The confusion matrix visualizes the model's performance by showing the counts of true positive, false positive, true negative, and false negative predictions.
2. It helps identify the model's strengths and weaknesses in classifying spam and ham messages.

4.Cross-Validation Scores:

1. Cross-validation scores provide an estimate of the model's generalization performance across different training and validation splits.
2. The mean cross-validation score indicates the average performance of the model across multiple validation folds.

CONCLUSION

- The implemented spam classification model demonstrates strong performance in accurately distinguishing between spam and non-spam (ham) messages.
- With high accuracy and balanced precision, recall, and F1-score metrics, the model effectively identifies spam while minimizing false positives and false negatives.
- The confusion matrix provides a clear visualization of the model's performance, showcasing its ability to make correct predictions while highlighting areas for potential improvement.
- Cross-validation further confirms the model's stability and generalization across different data subsets.
- Overall, the model presents a reliable solution for filtering spam messages, with opportunities for further optimization to enhance its effectiveness in real-world applications.

FUTURE SCOPE

1. **Real-Time Monitoring:** Implement real-time monitoring and feedback mechanisms to continuously evaluate the model's performance, adapt to changing spam patterns, and mitigate concept drift over time.

2. **Multi-Class Classification:** Extend the model to handle multi-class classification tasks, where messages can be classified into multiple categories (e.g., spam, promotional, personal) to provide more nuanced filtering capabilities.

3. **Real-Time Monitoring:** Implement real-time monitoring and feedback mechanisms to continuously evaluate the model's performance, adapt to changing spam patterns, and mitigate concept drift over time.

4. **User Feedback Integration:** Incorporate user feedback mechanisms to collect labeled data and refine the model's predictions based on user interactions and preferences.

REFERENCES

- https://www.kaggle.com/datasets
- https://pandas.pydata.org/pandas-docs/stable/user guide/index.html
- https://seaborn.pydata.org/

edunet foundation

THANK YOU