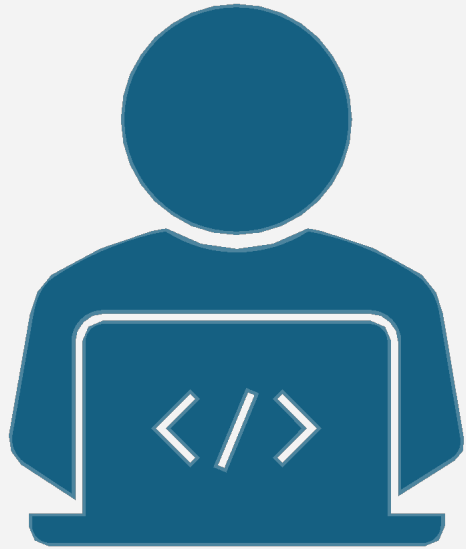


Team Members:

1) Sushanth Reddy Papareddy - 00868162

2) Nijaz Salihovic - 00681449



SMS Spam Detection Project

- This project focuses on creating an efficient system to identify spam messages using machine learning techniques in Python, with a strong emphasis on TensorFlow-based deep learning models. The goal is to leverage advanced AI methodologies to analyze text patterns and accurately distinguish spam from legitimate messages.

PROJECT OBJECTIVES



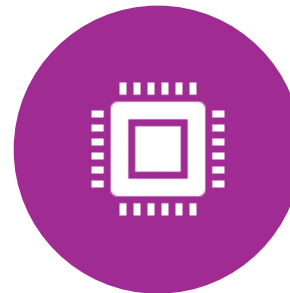
Explore a wide range of text classification techniques, including traditional machine learning methods (e.g., Naïve Bayes, Support Vector Machines, and Decision Trees) and advanced deep learning architectures (e.g., LSTMs, CNNs, and Transformer-based models) using TensorFlow.



Develop and train multiple models using labeled SMS datasets to compare their efficiency in distinguishing spam from legitimate messages.

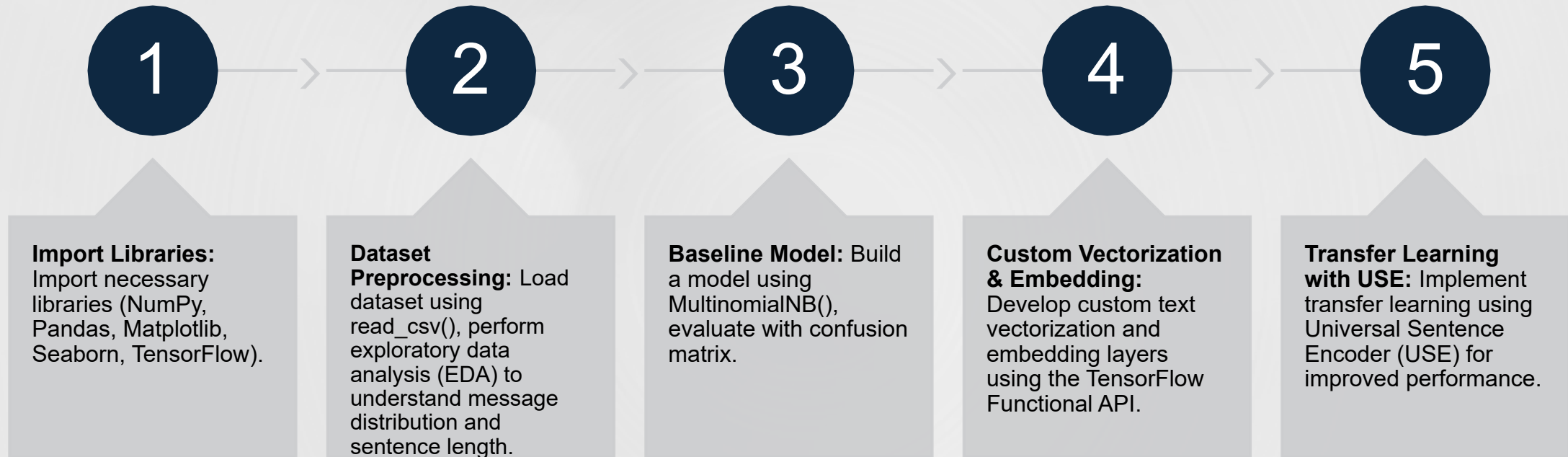


Evaluate and benchmark model performance using key metrics like accuracy, precision, recall, and F1-score to determine the most effective spam detection approach.



Deploy the best-performing model as a real-time spam detection system, integrating it into an application or API for practical use.

Project Approach



Tools : NumPy, Pandas, Matplotlib, TensorFlow, Seaborn, Universal Sentence Encoder (USE).

Techniques: Data Preprocessing(reading csv file), Machine Learning Model, Custom Text Vectorization, Transfer Learning.

Expected Deliverables:



Preprocessed Dataset: A well-processed dataset ensuring balanced feature representation to improve model learning and performance.



Baseline Model Evaluation Report: A comprehensive report detailing the confusion matrix analysis and assessing the effectiveness of the initial model.



Trained Deep Learning Models: Advanced models incorporating customized text vectorization and embedding layers, trained for optimal spam detection.



Comprehensive Model Evaluation Metrics: Performance analysis of each model using key metrics such as **F1-score, recall, accuracy, and precision**.



Relevance: These deliverables are crucial in assessing the effectiveness of various models for SMS spam detection, directly contributing to the project's objectives.

Evaluation Methodology

- Evaluate the model's performance using essential metrics, including accuracy, precision, recall, and F1-score.
- Given the dataset's imbalance, emphasize minimizing false positives and false negatives to improve classification reliability.
- Use the F1-score as the primary metric for model selection, as it offers a balanced assessment of precision and recall.
- This evaluation method ensures the identification of the most efficient model for SMS spam detection by thoroughly analyzing both precision and recall.

