**Learning From this project**

**1. Data Exploration through Visualization:**

- I learned how to visually explore dataset characteristics, such as class distribution, using histograms and boxplots to gain insights into the underlying patterns.

**2. Text Preprocessing Techniques:**

- I gained hands-on experience in preprocessing text data, including tokenization, converting text to lowercase, and removing stopwords, which is crucial for building effective natural language processing models.

**3. TF-IDF Vectorization:**

- I understood the process of converting text data into numerical features using the TF-IDF vectorization technique, a fundamental step in training machine learning models for text classification.

**4. Model Training and Evaluation:**

- I learned to train a Multinomial Naive Bayes classifier on the TF-IDF vectors and evaluate its performance using metrics such as accuracy and a detailed classification report.

**5. NLTK Compatibility:**

- I explored the integration of a Scikit-learn classifier with NLTK, enabling the use of the classifier within the NLTK framework.

**6. Challenges in Data Visualization:**

- I encountered and addressed challenges related to visualizing class distribution histograms, specifically dealing with the categorical nature of the 'label' column.

**7. Comprehensive Understanding of the Text Classification Process:**

- The project provided a holistic understanding of the end-to-end process of building a text classification model, from data exploration and preprocessing to model training and evaluation.

**8. Practical Application of Data Science Techniques:**

- I applied various data science techniques, including visualization, feature engineering, and machine learning, to a real-world problem of spam detection in SMS messages.