## Artificial Intelligence and Machine Learning

Project Report

Semester-IV (Batch-2022)

Email spam detection

A red and white sign

Description automatically generated with low confidence

**Supervised By: Submitted By:**

Rajeev Bharadwaj Suhavi jindal

2210990867

**G-13**

**Department of Computer Science and Engineering**

Chitkara University Institute of Engineering & Technology,

## Chitkara University, Punjab

**Table of content:**

|  |  |  |
| --- | --- | --- |
| **s.no** | **content** | **Page no.** |
| **1** | **Introduction** | **1-2** |
| **2** | **Problem Definition and Requirements** | **2-4** |
| **3** | **Proposed Design/Methodology** | **5-6** |
| **4** | **code** | **6-22** |
| **5** | **Result** | **23** |
| **6** | **References** | **24** |

**Introduction**

Email has become an indispensable communication tool in today's digital age, facilitating efficient and instantaneous communication across the globe. However, with the convenience of email comes the persistent challenge of spam. Spam emails, often unsolicited and irrelevant to the recipient, inundate inboxes, causing annoyance, wasting time, and posing security risks.

The proliferation of spam emails has prompted the development of sophisticated spam detection systems aimed at filtering out unwanted messages before they reach the recipient's inbox. Email spam detection refers to the process of automatically identifying and classifying incoming emails as either spam or legitimate (ham) based on various attributes and characteristics.

**Objective**

The objective of email spam detection is twofold: to enhance user experience by minimizing the intrusion of unwanted emails and to mitigate potential risks associated with malicious spam content, such as phishing attempts, malware distribution, and fraudulent schemes.

Traditional approaches to email spam detection typically rely on rule-based filtering, where predefined rules or patterns are used to flag emails as spam based on criteria such as keywords, sender reputation, and message content. While effective to some extent, rule-based systems often struggle to adapt to evolving spam tactics and may generate false positives or false negatives.

In recent years, artificial intelligence (AI) and machine learning (ML) techniques have emerged as powerful tools for email spam detection, offering more adaptive and robust solutions. ML algorithms can analyze large volumes of email data, learn from patterns and trends, and automatically detect spam with high accuracy. Common ML approaches for email spam detection include Naive Bayes, Support Vector Machines (SVM), Decision Trees, and Neural Networks.

**Significance**

The significance of email spam detection extends beyond individual users to encompass organizations and businesses, where email security is critical for safeguarding sensitive information, maintaining reputation, and ensuring regulatory compliance. By deploying effective spam detection systems, organizations can mitigate the risks of data breaches, financial fraud, and reputational damage associated with spam emails.

In this context, the aim of this project is to develop an advanced email spam detection system leveraging AI and ML techniques to accurately identify and filter out spam emails while minimizing false positives and ensuring timely delivery of legitimate messages. By harnessing the power of AI and ML, we aim to enhance email security, streamline communication processes, and improve overall user productivity and satisfaction.

**Problem Definition and Requirements**

**Problem Statement:**

The problem addressed in this project is the detection and filtering of spam emails from legitimate ones in an email system. Spam emails pose significant challenges to users and organizations, including annoyance, potential security threats, and reduced productivity. The primary objective is to develop an efficient and accurate spam detection system using artificial intelligence and machine learning techniques to automatically classify incoming emails as spam or legitimate.

**Requirements:**

**1. Software Requirements:**

- Programming Language: Python

- Machine Learning Libraries: scikit-learn, TensorFlow, Keras

- Natural Language Processing (NLP) Libraries: NLTK, SpaCy

- Data Visualization: Matplotlib, Seaborn

- Development Environment: Jupyter Notebook, PyCharm

**2.** **Hardware Requirements:**

- Processor: Intel Core i5 or equivalent

- RAM: Minimum 8GB

- Storage: Sufficient disk space for datasets and model files

**3. Data Sets:**

- Training Data: A labeled dataset consisting of both spam and legitimate emails for model training. The dataset should be representative and diverse, covering various types of spam and legitimate messages.

- Testing Data: Separate labeled dataset for evaluating the performance of the trained model. It should contain a mix of spam and legitimate emails that were not included in the training set.

**4. Feature Extraction:**

- Extract relevant features from email messages, such as sender address, subject line, message body, attachments, and metadata.

- Utilize techniques such as Bag-of-Words (BoW), Term Frequency-Inverse Document Frequency (TF-IDF), and word embeddings for feature representation.

**5. Machine Learning Algorithms:**

- Experiment with various machine learning algorithms suitable for text classification, including:

- Naive Bayes

- Support Vector Machines (SVM)

- Random Forest

- Gradient Boosting

- Neural Networks (e.g., LSTM, CNN)

- Evaluate the performance of each algorithm based on metrics such as accuracy, precision, recall, and F1 score.

**6. Model Evaluation:**

- Split the dataset into training and testing sets for model training and evaluation, respectively.

- Utilize cross-validation techniques to assess model performance and generalization ability.

- Fine-tune hyperparameters to optimize model performance and prevent overfitting.

**Proposed Design/Methodology**

**Design Overview:**

The proposed design for the email spam detection project involves several key components, including data preprocessing, feature extraction, model training, evaluation, and deployment. The project will leverage artificial intelligence and machine learning techniques to develop an effective spam detection system.

**1. Data Preprocessing:**

- Tokenization: Split email messages into individual words or tokens.

- Normalization: Convert tokens to lowercase and remove punctuation to standardize text.

- Stopword Removal: Eliminate common stopwords (e.g., "the", "and") that do not contribute to classification.

- \*\*Stemming/Lemmatization:\*\* Reduce words to their root forms to improve feature representation.

**2. Feature Extraction:**

- Bag-of-Words (BoW): Represent email messages as vectors of word frequencies.

- TF-IDF (Term Frequency-Inverse Document Frequency): Assign weights to words based on their frequency in the document and inverse frequency across all documents.

- Word Embeddings: Utilize pre-trained word embeddings (e.g., Word2Vec, GloVe) to capture semantic relationships between words.

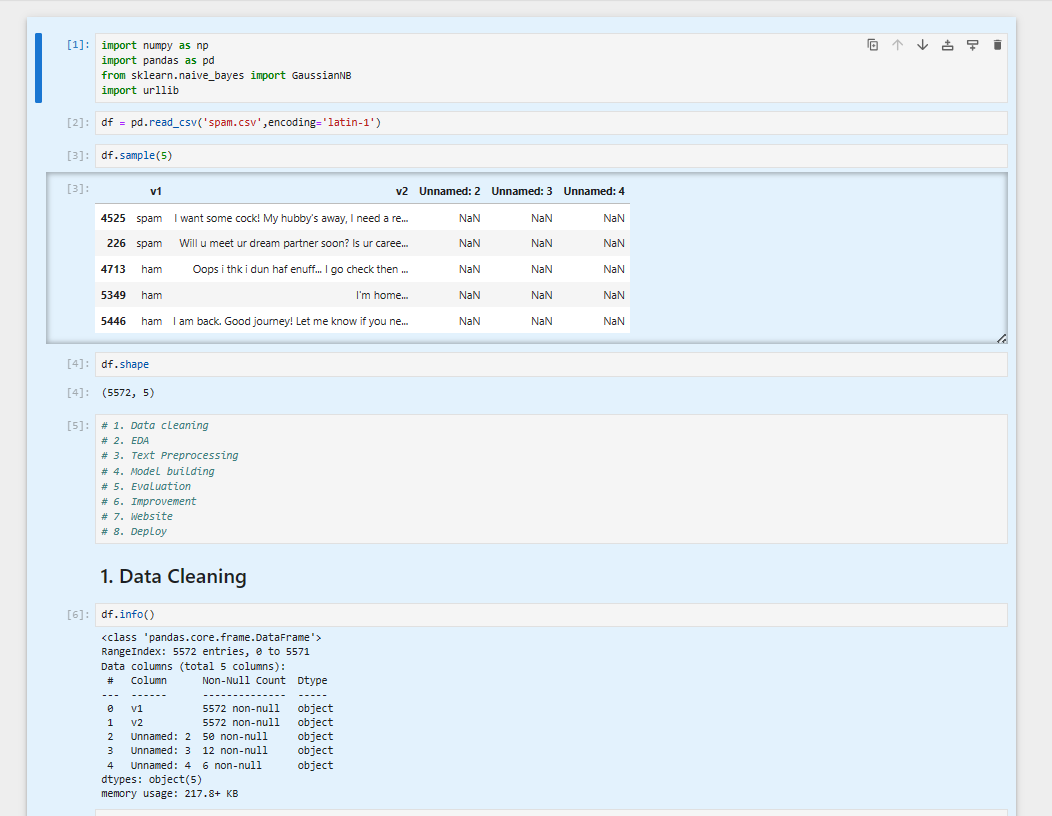
**3. Model Selection and Training:**

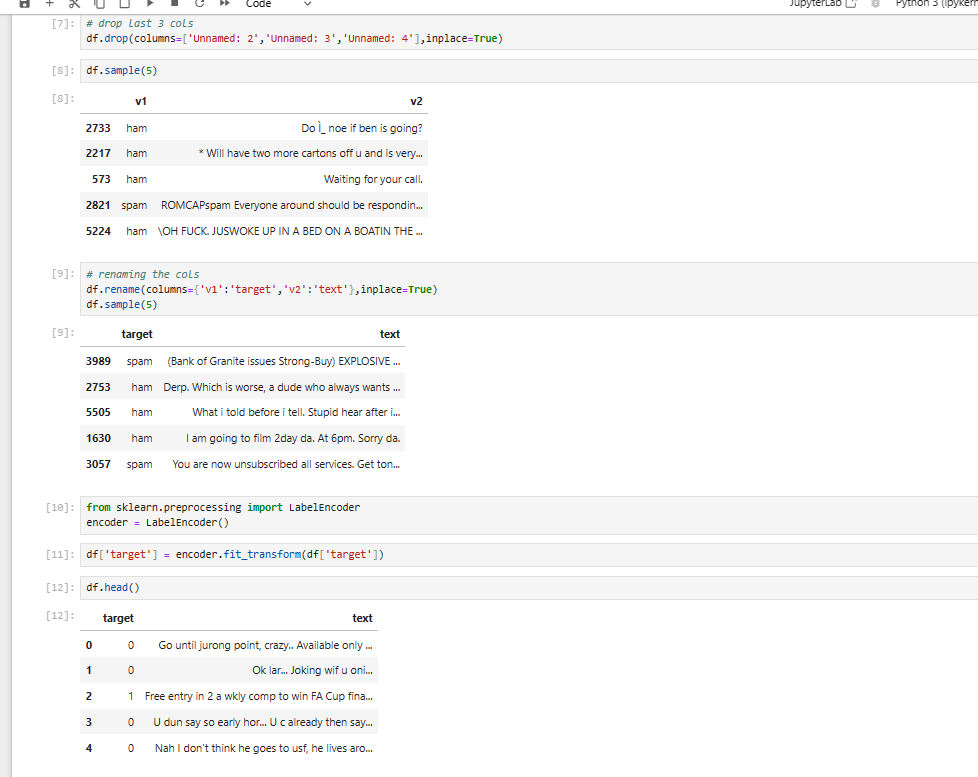
- Algorithm Selection: Experiment with multiple machine learning algorithms suitable for text classification, including Naive Bayes, Support Vector Machines (SVM), Random Forest, and Neural Networks.

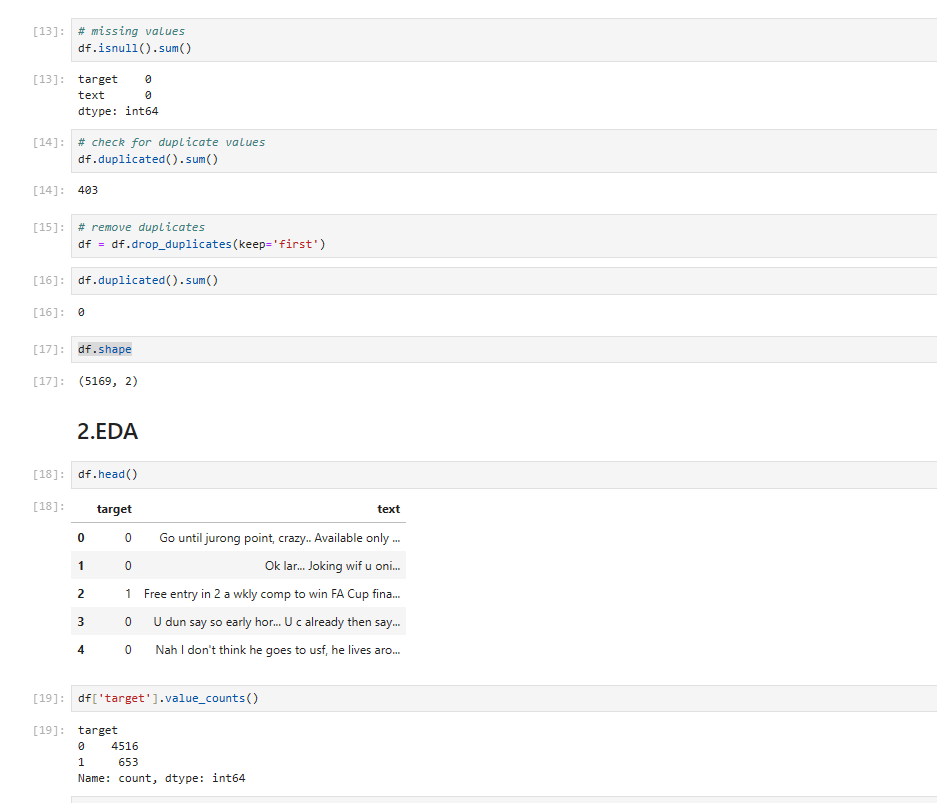
**4. Model Evaluation:**

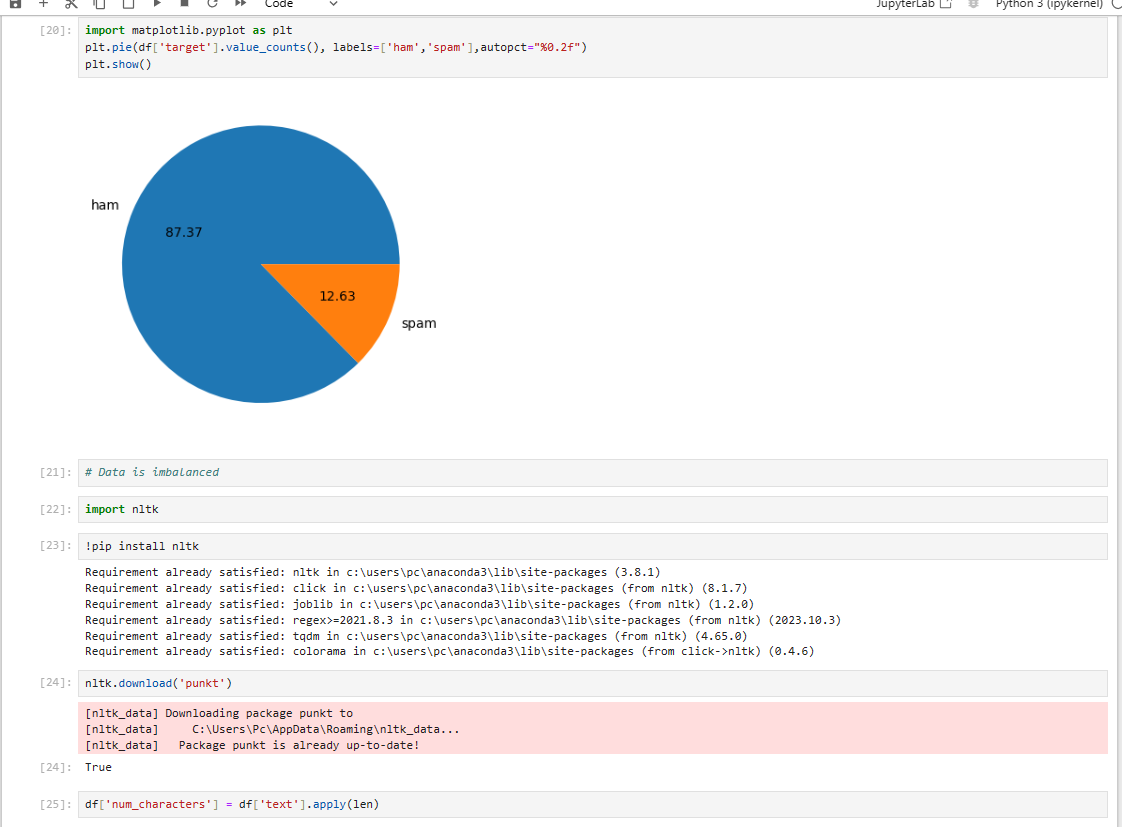
- Performance Metrics: Evaluate the trained models using standard performance metrics such as accuracy, precision, recall, and F1 score.

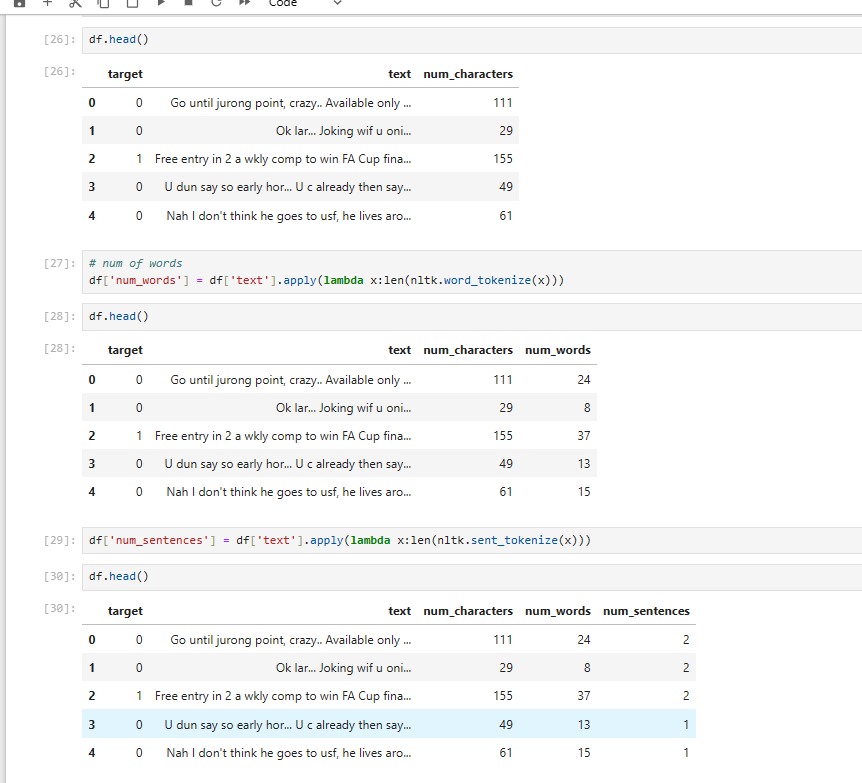
**Code:**

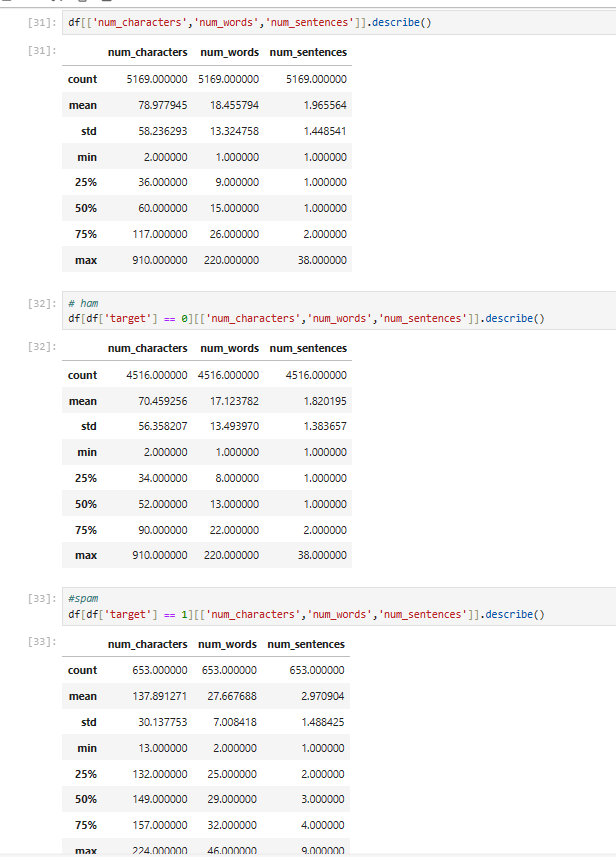
****

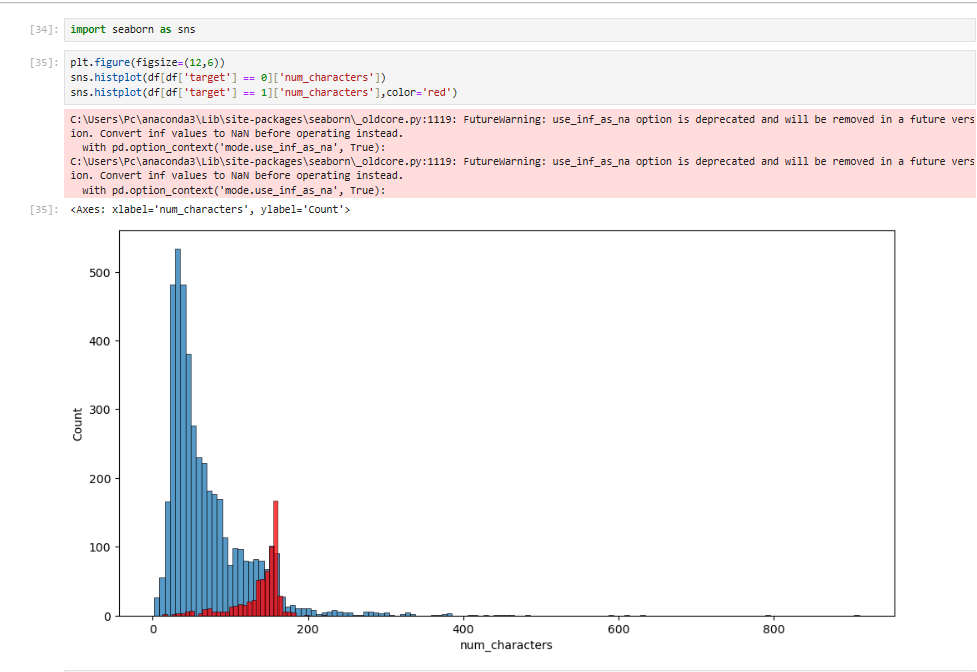
****

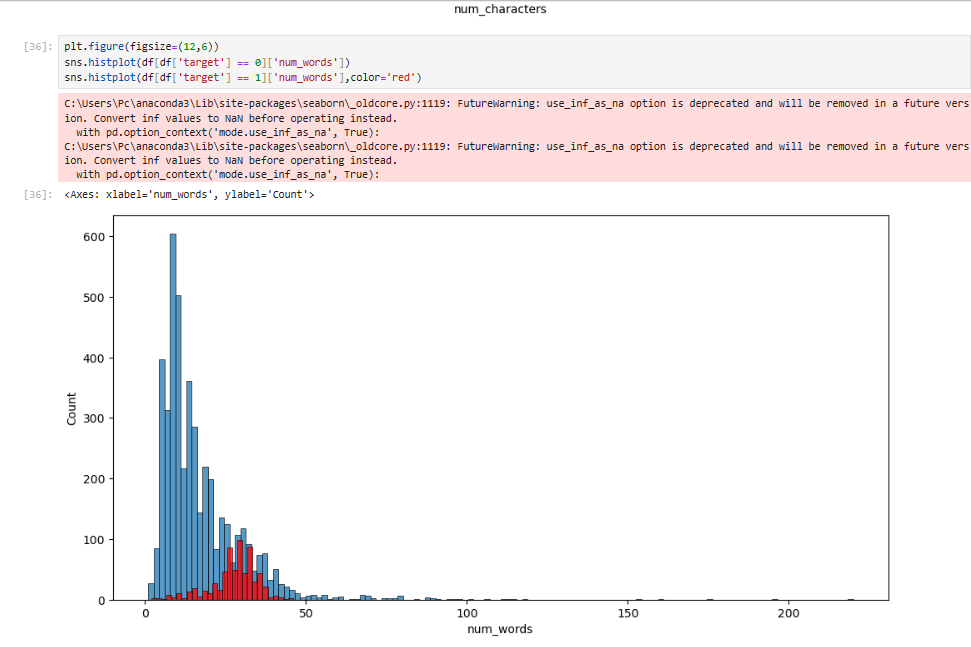
****

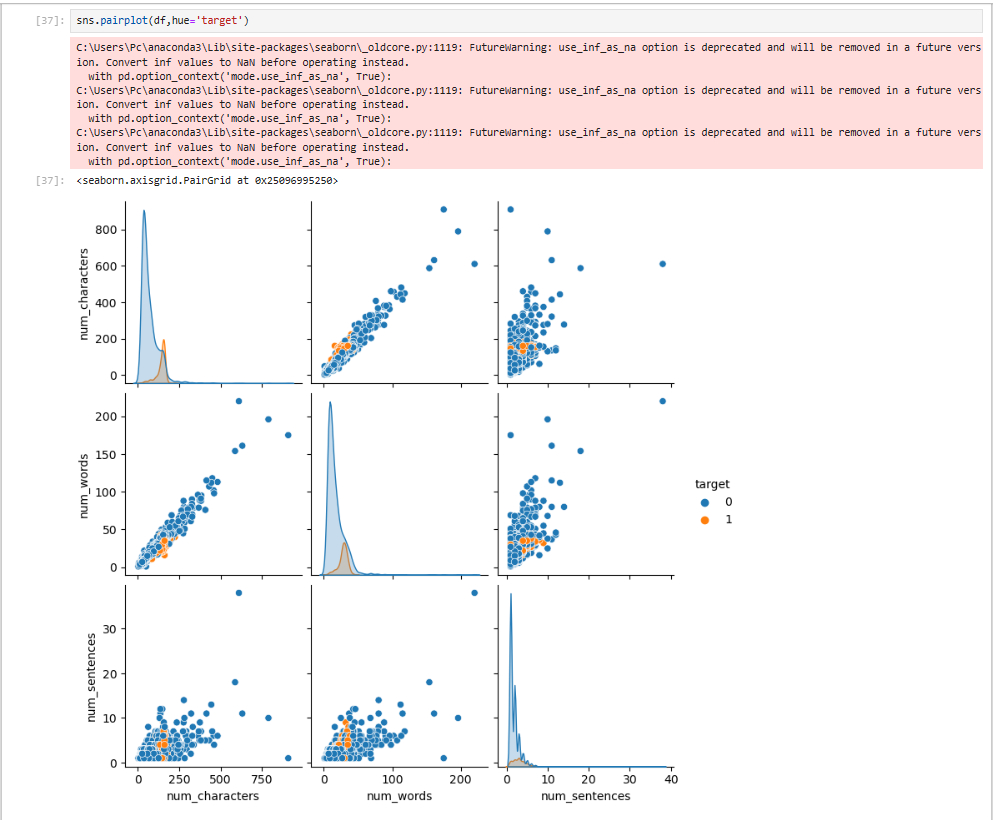
****

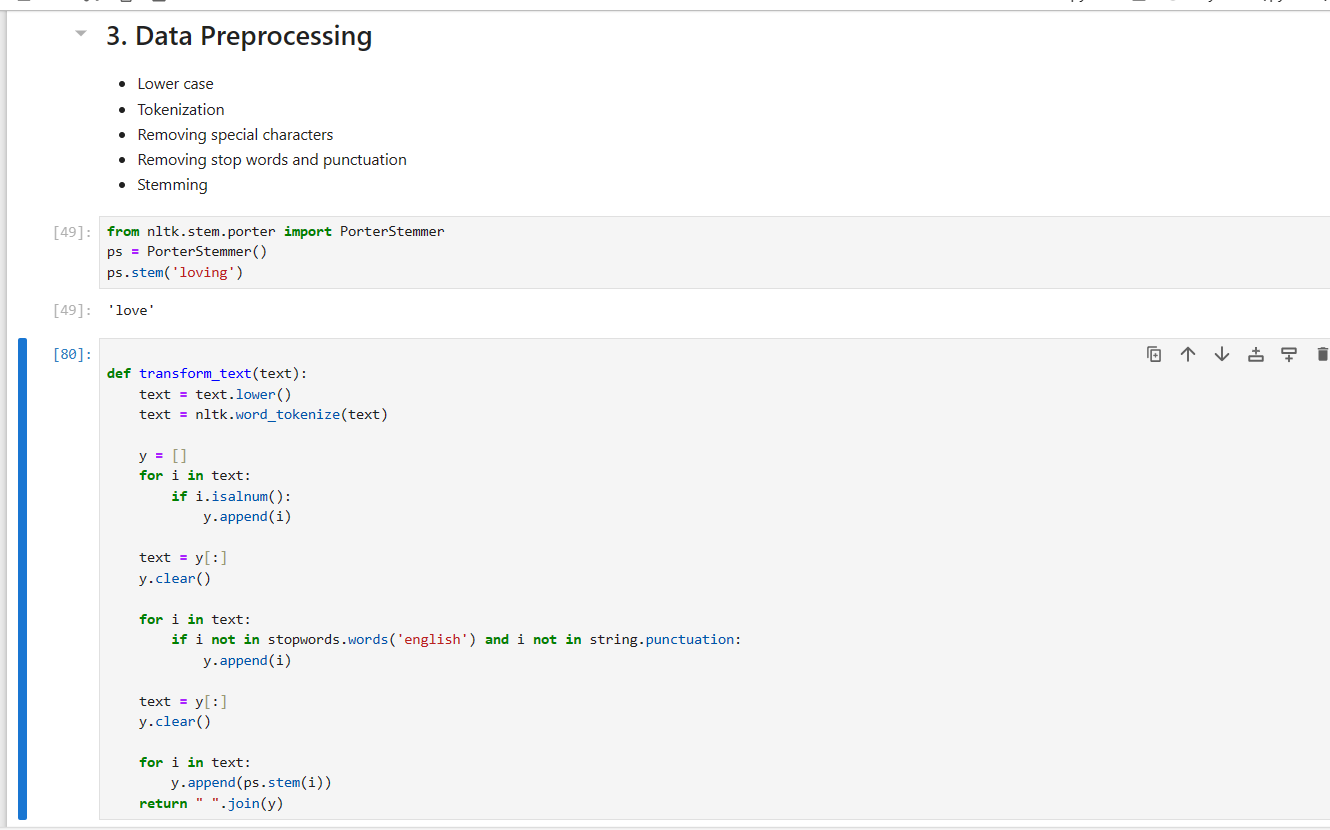
****

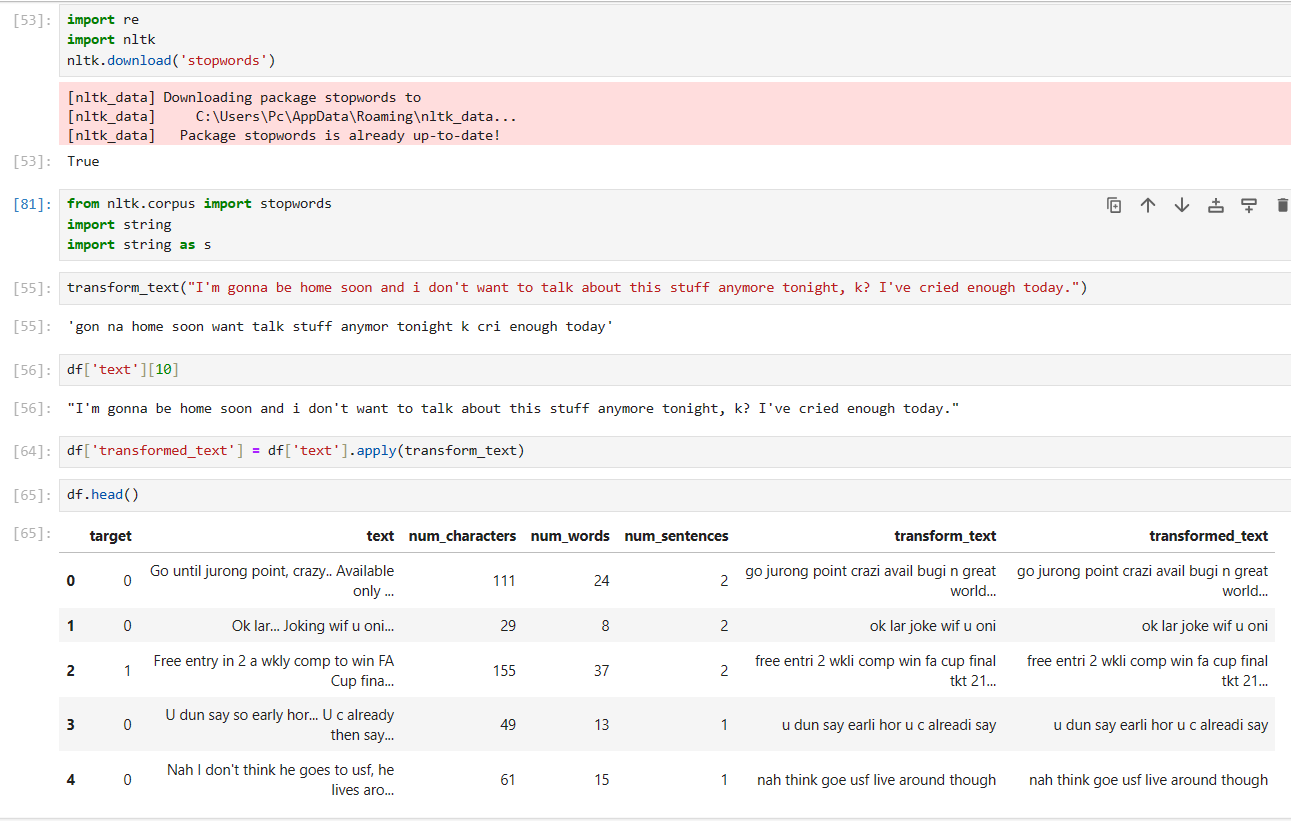
****

****

****

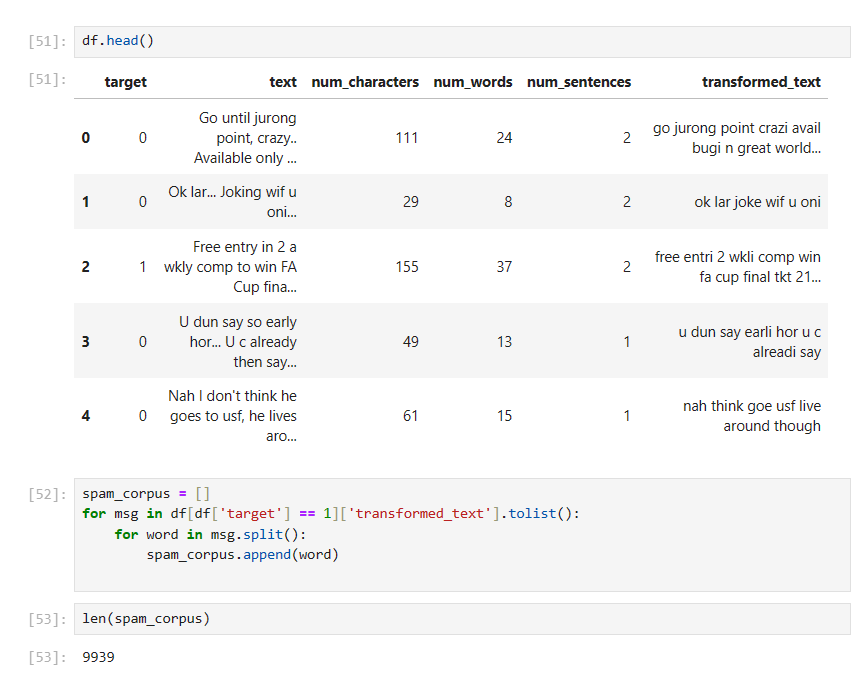
****

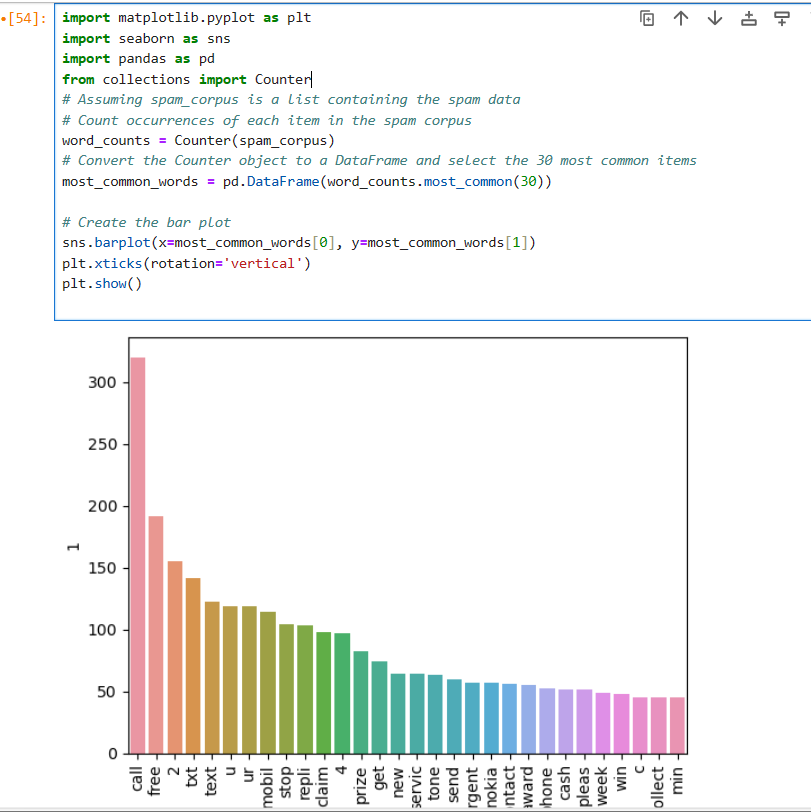
****

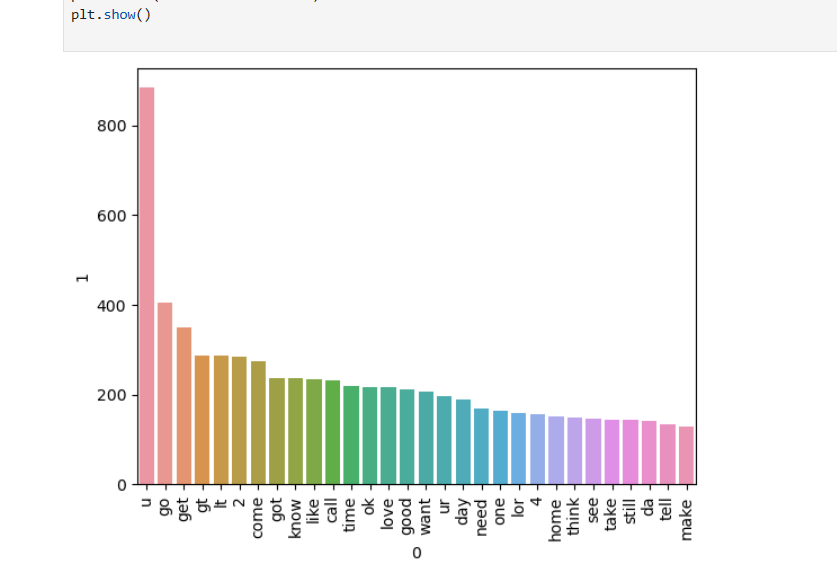
****

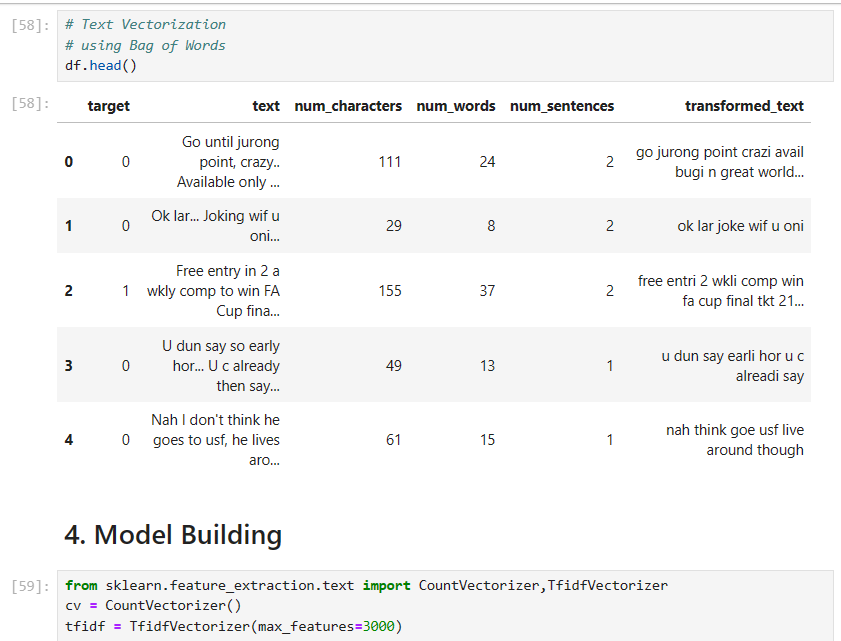
****

****

****

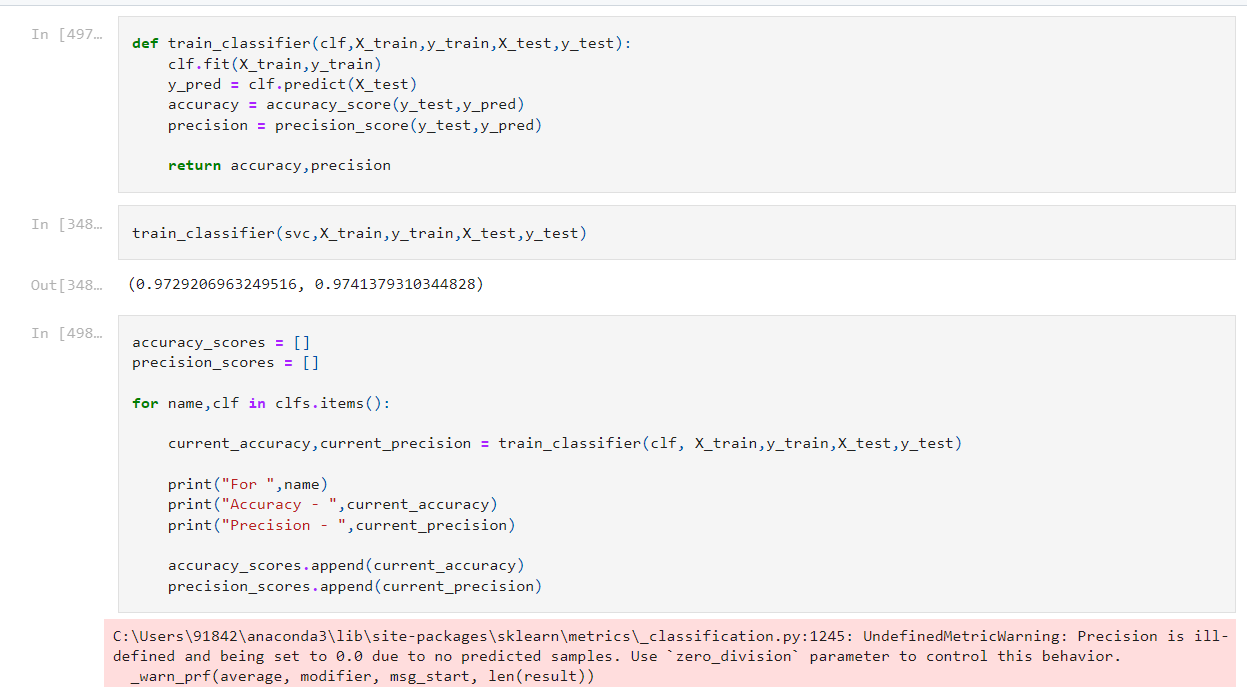
****

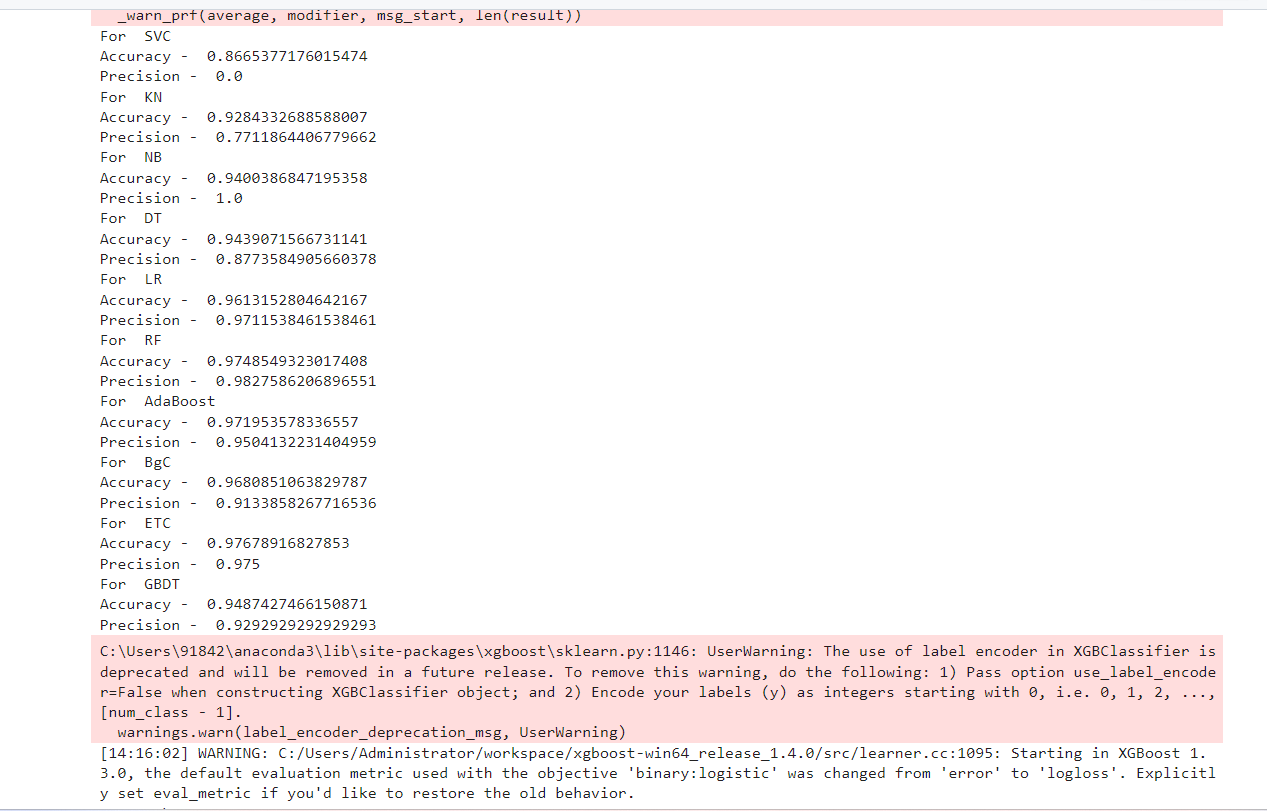
****

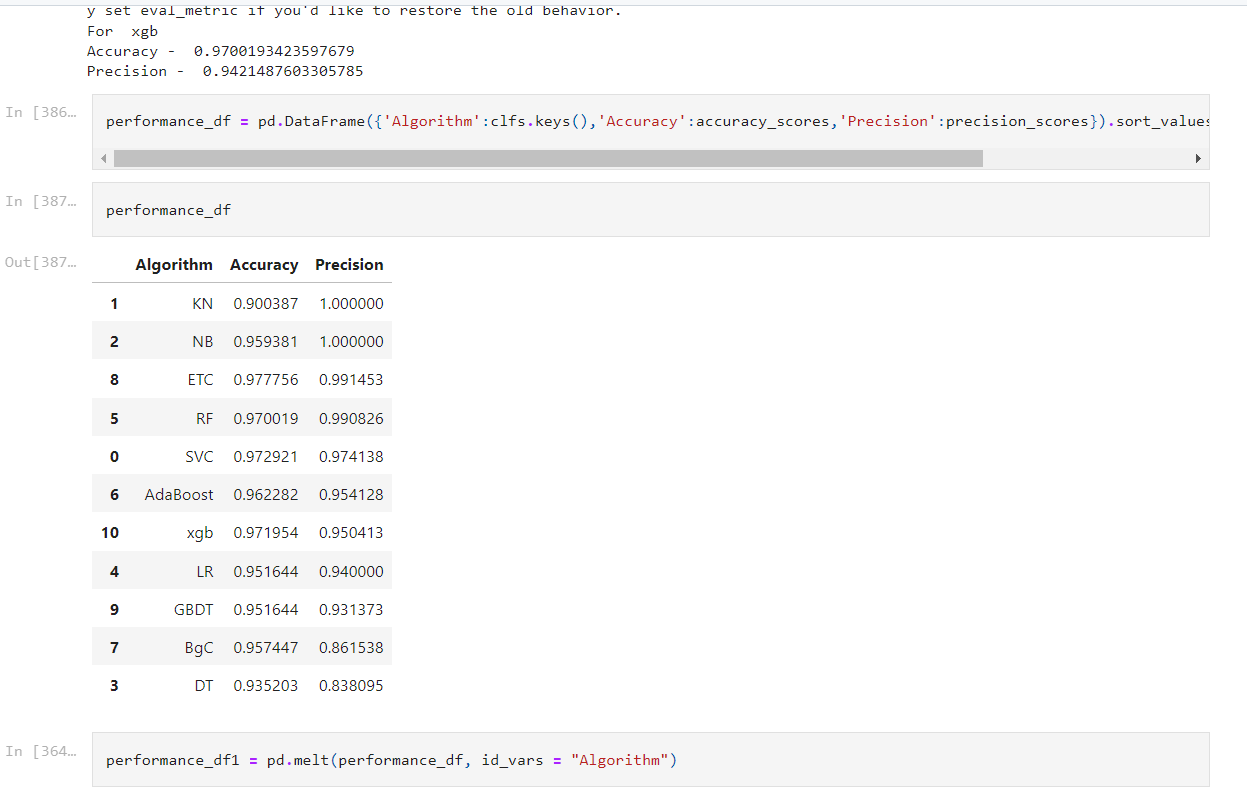
****

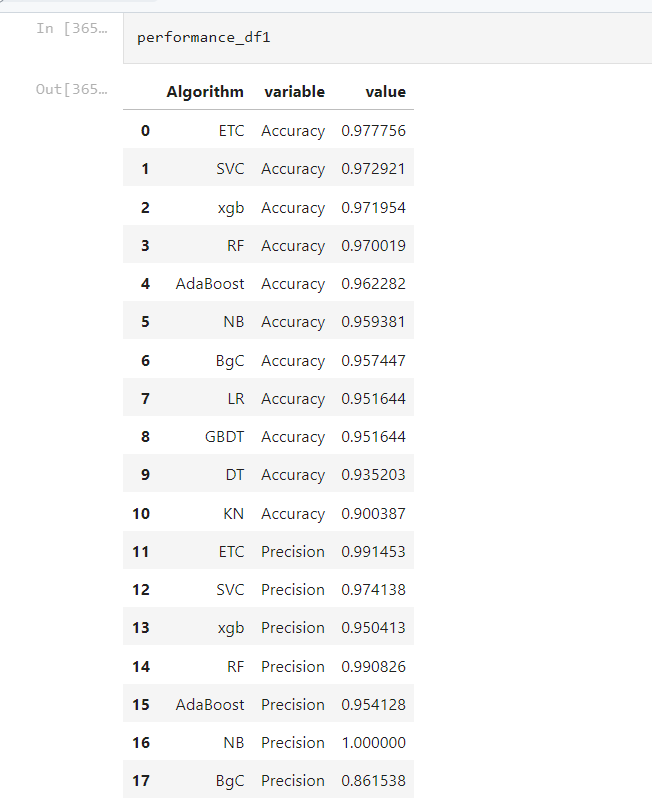
****

****

****

****

****

****

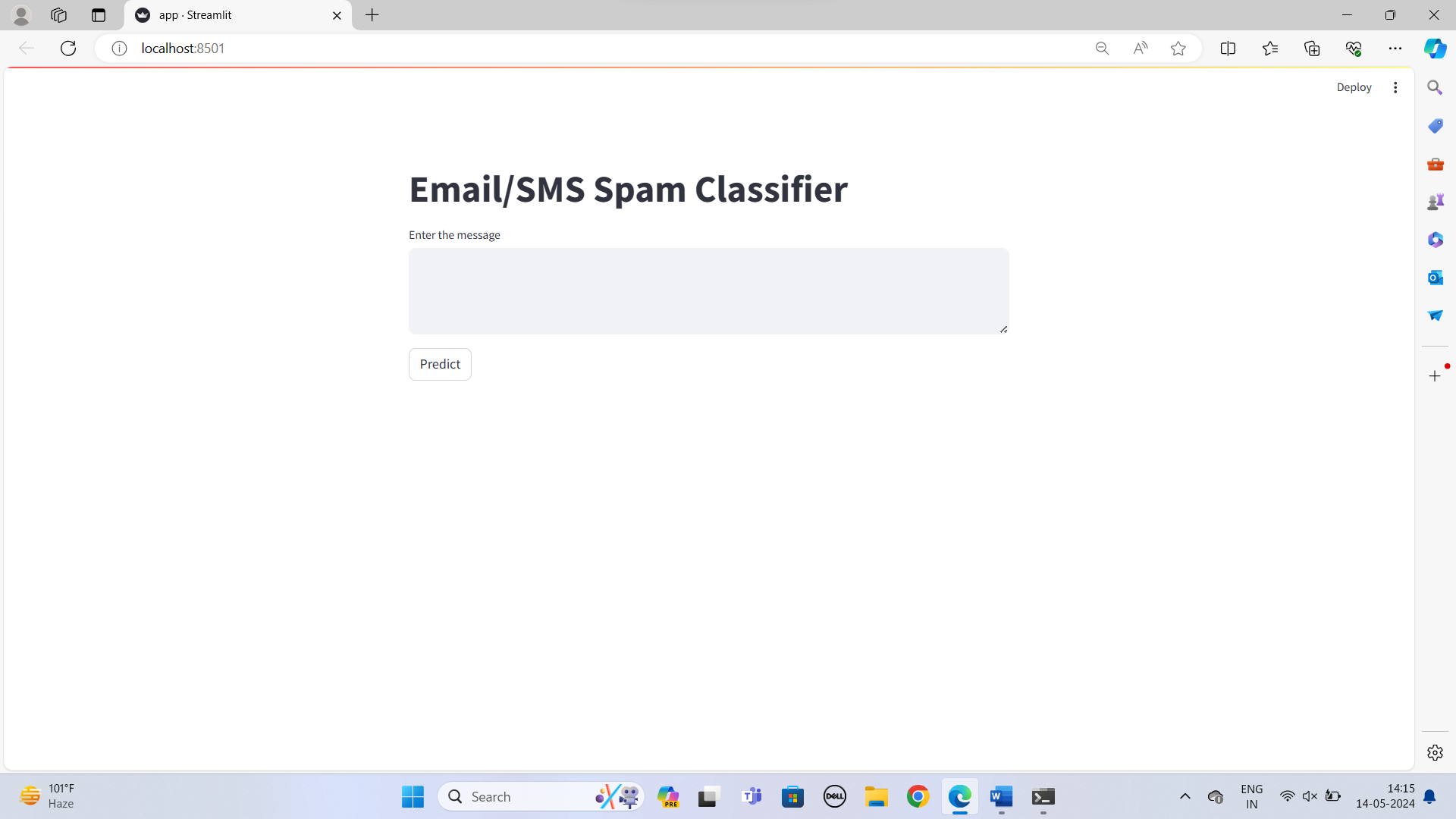
****

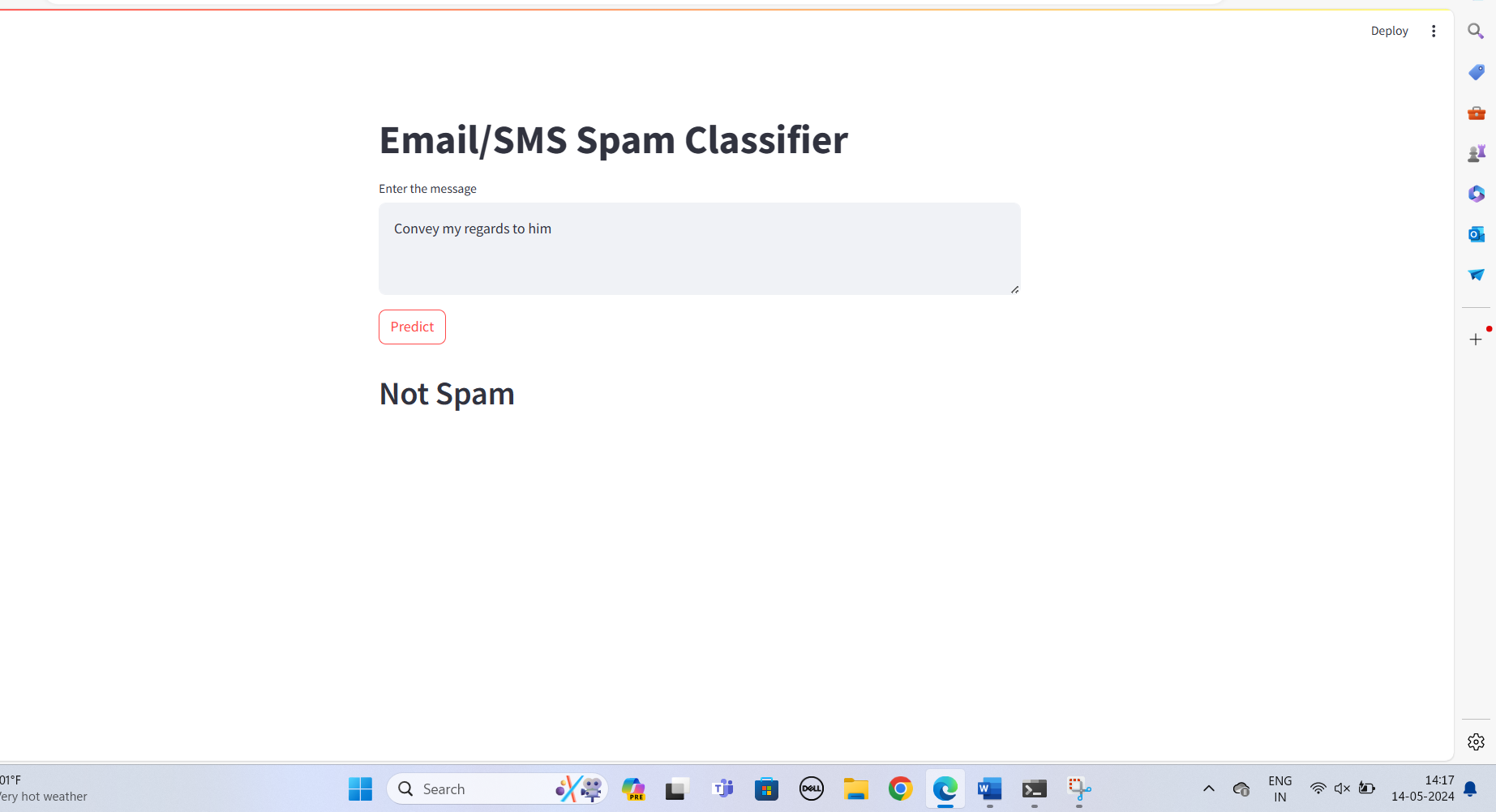
****

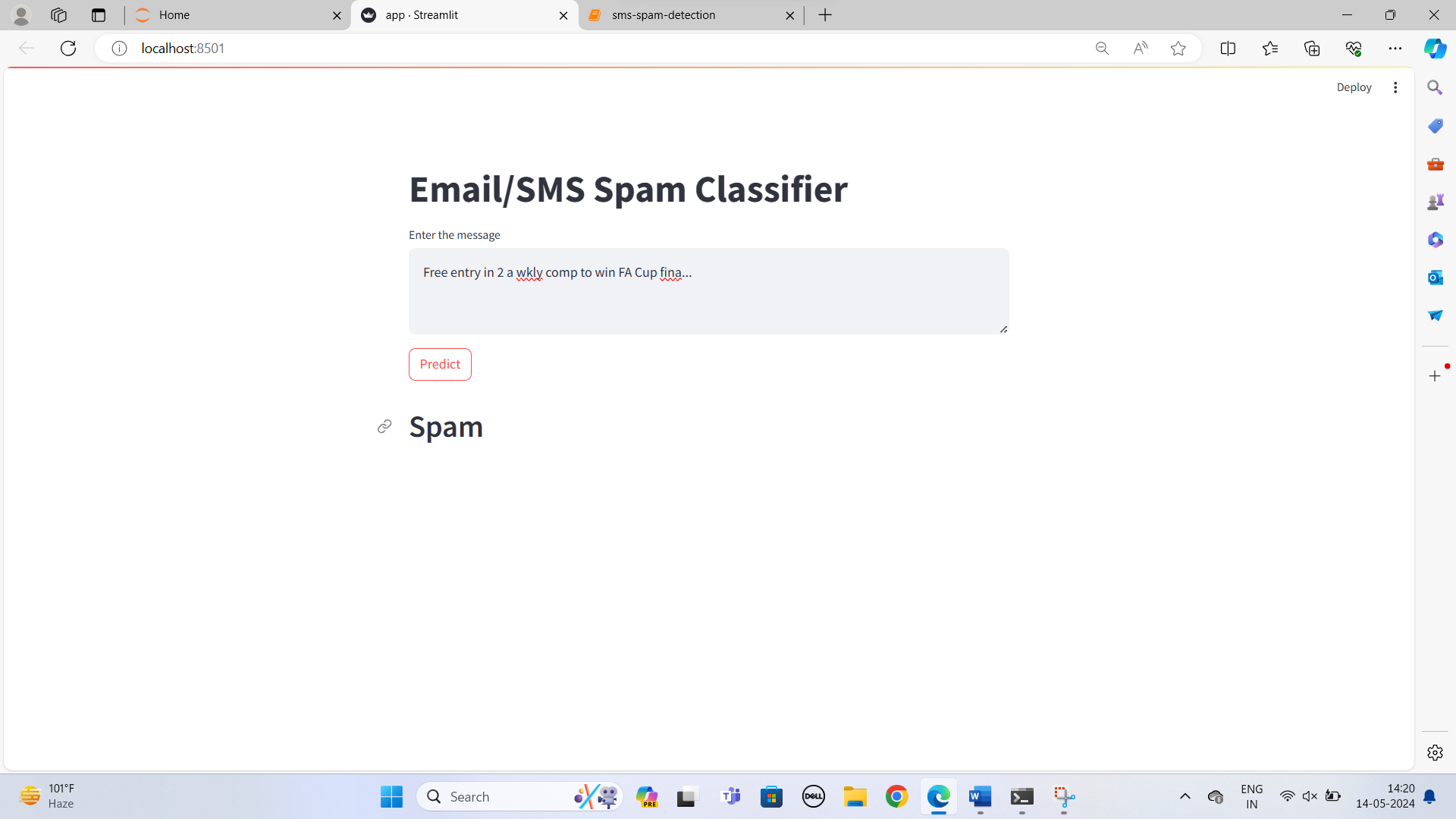
****

****

**Result:**

****

****

****

**References:**

**Aiml notes:** [**https://www.geeksforgeeks.org/aiml-introduction/#:~:text=AIML%20(Artificial%20Intelligence%20Markup%20Language,XML%20(eXtensible%20Markup%20Language**](https://www.geeksforgeeks.org/aiml-introduction/#:~:text=AIML%20(Artificial%20Intelligence%20Markup%20Language,XML%20(eXtensible%20Markup%20Language)**).**

**Youtube:**[**https://www.youtube.com/watch?v=YncZ0WwxyzU&pp=ygUUZW1haWwgc3BhbSBkZXRlY3Rpb24%3D**](https://www.youtube.com/watch?v=YncZ0WwxyzU&pp=ygUUZW1haWwgc3BhbSBkZXRlY3Rpb24%3D)

**Dataset:** [**https://www.kaggle.com/datasets/uciml/sms-spam-collection-dataset**](https://www.kaggle.com/datasets/uciml/sms-spam-collection-dataset)