

**Task Assignment: Email Text Classification**

**Scenario:**

Our organization receives a large volume of emails every day, which need to be sorted and classified into different categories based on their content. The current process for sorting these emails is time-consuming and error-prone, and we want to implement an automated system to streamline the process.

**Goals:**

The goal of this project is to develop a text classification system that can accurately classify emails into different categories based on their content. This will help to ensure that the organization is responding appropriately to each email and improve the efficiency and accuracy of the email sorting process.

**Data:**

The sample dataset for this project is available in the attached file.

**Insights:**

1. Data Cleaning: The first step in this project is to clean and preprocess the email data. This will involve removing any unnecessary characters, such as punctuation and special symbols, converting all words to lowercase, and removing any stop words or irrelevant words.

2. Text Classification: There are several algorithms that can be used for text classification, such as Naive Bayes, Support Vector Machines (SVM), or Convolutional Neural Networks (CNNs). The selected algorithm will be used to classify each email into the appropriate category based on its content.

3. Model Training and Testing: The dataset will be split into training and testing sets, and the selected algorithm will be trained on the training set. The performance of the algorithm will be evaluated on the testing set using appropriate metrics, such as accuracy, precision, and recall.

4. Performance Evaluation: The text classification algorithm will be evaluated for its effectiveness in accurately classifying emails into the appropriate categories. Any necessary adjustments will be made to the preprocessing or classification steps to improve the accuracy and relevance of the results.

**Hot Fix:**

The text classification algorithm will be evaluated and adjusted as needed to ensure that it is accurately classifying emails into the appropriate categories. Additionally, the performance of the algorithm will be continually monitored and evaluated to ensure that it is providing accurate and timely classification results. Finally, the results of the text classification will be used to improve the organization's response to emails and ensure that customers are receiving appropriate and timely responses.