MAJOR PROJECT

- **Project Name:**

Machine Learning May Major Project

- **Project Description:**

SENTIMENT ANALYSIS - Sentiment analysis studies the subjective information in an expression, that is, the opinions, appraisals, emotions, or attitudes towards a topic, person or entity. Expressions can be classified as positive, negative, or neutral. For example: “I really like the new design of your website!” → Positive.

- You are provided with a dataset of the ‘restaurant review’ taken from kaggle. Build a machine learning model by using Support vector classifier(SVM) and count vectorizer using two
methods(direct method, pipeline method) to predict the label of the review either as positive or negative.

- Also use MultinomialNB (Naïve bayes) and count vectorizer using two methods(direct method, pipeline method) to classifiy if the sentiment is positive or negative and compare the accuracies of all 4 models of the models

- https://www.kaggle.com/d4rklucif3r/restaurant-reviews

- [[NOTE :There are two features - ‘review’ - the sentence and ‘sentiment’ - the label for the review. 1 means positive review and 0 means negative review.]]

You are supposed to

1.Create a dataframe

2.process the data and do visualizations (represent using matplotlib / seaborn the number of positive reviews and negative reviews)(ex-use a bar graph)

3.create svc model and count vectorizer separately (method 1)

4.(method 2) create a pipeline with Vectorization model and ML algorithm to predict the final sentiment.

5.create NB model and count vectorizer separately (method 1)
6.(method 2) create a pipeline with Vectorization model and ML algorithm to predict the final sentiment for the multinomialNB and CountVectorizer
7.Use joblib to create and save it as a model (USE THE MODEL WITH THE HIGHEST ACCURACY)(joblib is similar to pickle)
8.Using the new model created using joblib, predict the output of a new review
9.Create a streamlit webapp for sentiment analysis using the joblib model (pipeline model)

Create a few lines about sentiment analysis and twitter sentiments (own words/ sentences) in the colab notebook itself