

we first load the dataset and split it into training and testing sets. we then use count vectorizer to convert the text data into numerical data, which can be used by the machine learning model.