

Project 1

Sequential Models in NLP

Sentiment Classification

Objective

The objective of this project is to build a text classification model that analyses the customer's sentiments based on their reviews in the IMDB database. The model uses a complex deep learning model to build an embedding layer followed by a classification algorithm to analyze the sentiment of the customers.

Dataset

The Dataset of 50,000 movie reviews from IMDB, labelled by sentiment (positive/negative). Reviews have been preprocessed, and each review is encoded as a sequence of word indexes (integers). For convenience, the words are indexed by their frequency in the dataset, meaning the word that has index 1 is the most frequent word. Use the first 20 words from each review to speed up training, using a max vocab size of 10,000. As a convention, "0" does not stand for a specific word, but instead is used to encode any unknown word.

Steps

Steps & marks are mentioned in the question notebook.

Support

You can raise a support query if you need any help.

Happy Learning!