Sentiment Analysis on IMDB Reviews

Introduction

Sentiment analysis is a technique for analyzing a piece of text to determine the sentiment contained within it. It accomplishes this by combining machine learning and natural language processing (NLP). In this project, I am trying to predict the number of positive and negative reviews based on sentiments by using LSTM models. The main goal is to estimate the sentiment many movie reviews from the Internet Movie Database (IMDb).

Dataset Used

I'm using IMDB dataset has 50K movie reviews for natural language processing or Text analytics. This is a dataset for binary sentiment classification containing substantially more data than previous benchmark datasets. They provide a set of 25,000 highly polar movie reviews for training and 25,000 for testing. I'm using a number of Python tools and frameworks dedicated to text analytics, natural language processing, and machine learning.

Training & Testing

I've generated a model based on the training dataset. I used the model on the test dataset and got an accuracy of 98%. The additional data is presented in the code file.

Results

In the end, I used different reviews of different movies, both positive and negative and the model predicted the right sentiment every time.

```
In [58]: review = str(input('Movie Review: '))
```

```
Movie Review: I walked out of the cinema after about 40 minutes, when I went to see this film approximately 25 years ago. I attempted to watch it again at Christmas, thinking maybe I was too impulsive or didn't really give it a chance. Turns out I was right the first time. Who gives Tim Burton money to make films? I hope they can sleep at night. Definitely not for me...
```

```
In [62]: if result >= 0.7:
             print('This review is positive')
         else:
             print('This review is negative')
```

```
This review is negative
```

```
In [ ]:
```