**EE7207 Assignment: Sentiment Classification for IMDB Movie Reviews**

**Introduction:**

In this assignment, you will be tasked with building sentiment classification models to predict the sentiment of movie reviews. You will work with a dataset stored as csv files containing movie reviews from IMDB.

**Dataset Description:**

**movie\_train.csv**: This file contains movie review text along with their corresponding sentiment labels. Each row represents a review, and the associated sentiment label is provided. You shall use this file for model training.

**movie\_test.csv**: this file contains the test data. You shall use this file for model performance evaluation.

**Sentiment Labels:**

The sentiment labels in the dataset are as follows:

* 0: Negative
* 1: Positive

**Assignment Requirements:**

Your assignment involves the following tasks:

1. **Build a Sentiment Classification Model with Plain Vanilla RNN:** Implement a sentiment classification model using a plain vanilla Recurrent Neural Network (RNN). You will preprocess the text data, tokenize it, and feed it into the RNN for training. Evaluate the model's performance on test data using appropriate evaluation metrics.
2. **Build a Sentiment Classification Model with LSTM:** Develop another sentiment classification model using Long Short-Term Memory (LSTM) networks. Preprocess the data similarly to the first model, train the LSTM model, and evaluate its performance on test data.
3. **Bonus: Build a Sentiment Classification Model with a More Powerful Model:** As an optional task, y­­ou can explore building a sentiment classification model using a more powerful architecture and comparing the performance with the plain vanilla RNN and LSTM models.

**Submission Guidelines:**

You are required to submit the following:

1. A Jupyter notebook containing the code implementation for each sentiment classification model.
2. In the same Jupyter notebook, discuss your approach, including explanations of the techniques used, insights gained from data exploration, details of the selected models, evaluation metrics, performance analysis, and any challenges encountered.

**Submission Deadline:**

11:59pm on April 5, 2024