Movie Review Sentiment Analysis

Tomal Majumder Toushik
Student ID:1405094
Mainul Hossain Yaad
Student ID:1405068
Akib Jawad Nafis
Student ID:1405112

February 5, 2019

1 Introduction

This document discusses three issues.

- How to download the project and set up the environment to run the project.
- How to predict the sentiment of movie reviews.
- How to reproduce the results.

2 Setting Up the Project and Environment

Download the full project folder from https://drive.google.com/drive/folders/1x88pL8n0WeYNFLChd-YzShiIFSGPcjNP?usp=sharing
::::::::::::::::::How to install in Linux::::::::::::::::::::::::::
We have used Numpy,Pandas,Sklearn,tensorflow,Keras libraries and python version 3.6. To create runnable environment you have to download all the libraries and python 3.6. To do this, simply follow.

- Download python from https://www.python.org/downloads/release/python-372/ and install using the given readme file inside it.
- after installation you can run python3 –version to check if you have the installation correct.
- Assuming you've installed it correctly. Just run these command in terminal
    1. pip3 install numpy

2. pip3 install pandas
3. pip3 install keras
4. pip3 install tensorflow
5. pip3 install sklearn

- After these you're ready to run the program.

3 Running Sentiment Analysis and Reproducing Results

You can check the dataset in "data.csv" and "dataTest.csv".
To run the program simply run ***./script.sh***

Check the "output.txt" for results.

4 Our Results

4.1 LSTM

No of LSTM Layer	Epoch	Test data Accuracy
1	3	81.768
1	4	79.892
2	3	86.828
2	4	86.996

Table 1: LSTM output by changing no of layers and epoch

4.2 SVM

C	Accuracy
0.01	89.024
0.05	88.992
0.25	89.024
0.5	89.024
1.0	89.024

Table 2: SVM output by changing value of C

Best C is 0.01 and test accuracy with it is 89.896

C	Accuracy
0.01	88.992
0.05	89.856
0.25	89.968
0.5	89.856
1.0	89.824

Table 3: Regression output by changing value of C

4.3 Regression

Best C is 0.25 and test accuracy with it is 89.816