Programming Assignment – 4 (RNN)

Due: June 23, 2023
Total marks: 100
Assignments turned in after the deadline are subject to a 100% grade penalty.

Overview:

In this assignment, you will create an RNN model to predict the prices of ABC company's stock. You are free to use any techniques. The assignment will be graded based on the performance of your classifiers (using RMSE score), that is how well they perform on unseen test data compared to the performance of your classroom peers.

Data:

A set of training, and test data is available as a compressed ZIP archive on D2L. The grading script will run your model against the Test Dataset and generate the RMSE score. This score will be compared with the accuracies of your peers and ranked. Your grade will depend on your position on the leaderboard (not released) in class.

Program:

You need to create a Python script to generate the predictions for all your test data. You are free to use any RNN Classifier. You are allowed to use any techniques for data preprocessing and cleaning. Your python script should create a **predictions.csv** file after running.

Make sure that the final predictions are in form of a dataframe with the same columns as input data (i.e., date, price).

Submission:

- You are encouraged to submit early and often to iron out any problems, especially issues with the format of the final output. Any sort of formatting issues in the output will lead to deduction in the assignment grade.
- Name the python script as **FirstName_LastName_USCID.py.**
- Failure to format your output correctly may result in very low scores, which will not be changed.
- For full credit, make sure to submit your assignment well before the deadline. The time of submission recorded by the system is the time used for determining late penalties. If

your submission is received late, whatever the reason (including equipment failure and network latencies or outages), it will incur a late penalty.
- If you have any issues with D2L with regards to logging in, submission, code not executing properly, etc., please make a private Piazza post so the instructional team can look into the issue.

Grading:

After the due date, we will train your model again on the training data, and run your model on the **test dataset**, compute the RMSE score and rank it in your class.

- There is a **maximum 7 minutes of time limit** for running your entire python file. Codes that run over 7 minutes will be scored as 0. So, tune your hyper parameters accordingly.
- All the correct working models will get a 100% score.
- In addition, top 20% students will get a **bonus 20% score.**

Notes:

You are free to use any Python Libraries. Use of Pre-trained models is not allowed.

Collaboration and external resources:

- This is an individual assignment. You may not work in teams or collaborate with other students. You must be the sole author of 100% of the code you turn in.
- You may not look for solutions on the web or use code you find online or anywhere else.
- You may not download the data from any source other than the files provided on D2L, and you may not attempt to locate the test data on the web or anywhere else.
- You may use any python packages you like. Make sure not to use any pre-trained models for this assignment. You may use external resources to learn basic functions of Python (such as reading and writing files, handling text strings, and basic math).
- Failure to follow the above rules is considered a violation of academic integrity and is grounds for failure of the assignment, or in serious cases failure of the course.
- We use plagiarism detection software to identify similarities between student assignments, and between student assignments and known solutions on the web. Any attempt to fool plagiarism detection, for example, the modification of code to reduce its similarity to the source, will result in an automatic failing grade for the course.
- Please discuss any issues you have on the Piazza discussion boards. Do not ask questions about the assignment by email; if we receive questions by email where the response

could be helpful for the class, we will ask you to repost the question on the discussion boards.