Instructor: Alina Vereshchaka

Project 2 -- Time-series Prediction

Deadline: July 21, Tue, 11:59pm

Description

This project is focused on forecasting a time-series data.

Dataset

Requirements to the dataset:

- It has to be different from the dataset used for Project 1 and used as examples in class (exceptions can be discussed)
- Represent the real-world data
- Contain at least 5k entries

Possible resources includes:

- Open Data Buffalo - https://data.buffalony.gov/
- Google Dataset - https://datasetsearch.research.google.com/
- Google Trends - https://trends.google.com/trends/?geo=US
- Yahoo Finance - https://finance.yahoo.com/

Tasks

Part I: Prepare the dataset for training [20 points]

1. Choose the dataset

2. Extract and describe the main statistics about the dataset and provide visual representation of the dataset.
3. Preprocess the dataset for training (e.g. cleaning and filling the missing variables, split between training/testing/validation)

Part II: Classical time series forecasting methods [40 points]

1. Choose the features and targets in the dataset.
2. Apply statistical algorithms (min 3 algorithms) to forecast the values. Possible algorithms include: ARIMA, VAR, SARIMAX, etc.
3. Provide the comparison of the results of different statistical models you have used. This can be in the form of graph representation and your reasoning about the results.

Part III: Deep learning time series forecasting methods [40 points]

1. Apply MLP to predict the value. Show the results on 3 different MLP setups (#layers, activation functions, learning rate, layers structures, etc)
2. Apply RNN or LSTM architecture to predict the value.
3. Plot the graphs (predicted vs true values, accuracy, loss)
4. Discuss and provide the results of predicting the values using different deep learning structures.

Submit the Project

- Submit at **UBLearns > Assignments**
- Complete your project in either Jupyter Notebook or python script.
- In your report include the answers to questions for each part. You can complete the report in a separate pdf file or in Jupyter notebook along with your code.
- Include all the references that have been used to complete the project.

Important Information

This project can be done in a team of up to two people. The standing policy of the Department is that all students involved in an academic integrity violation (e.g. plagiarism in any way, shape, or form) will receive an F grade for the course. Refer to

the [Academic Integrity website](#) for more information.

Late Days Policy

You can use up to 5 late days throughout the course that can be applied to any project. You don't have to inform the instructor, the late submission will be tracked in UBlearns. If you work in teams the late days used will be subtracted from both partners. E.g. you have 4 late days and your partner has 3 days left. If you submit one day after the due date, you will have 3 days and your partner will have 2 days left.

Important Dates

July 15, 11:59pm - Register your team (https://forms.gle/pwJUxu9WZxRnwXuD6)

July 21, 11:59pm - Project 2 is Due