QF201, Fall 2022
Problem set #7

Due Saturday, October 15 at 23:59:59 MSK at Gradescope

If you have a question redarding this task, we highly encourage you to post your question on our Piazza forum, at https://piazza.com/talantiuspeh.ru/fall2022/qf201

All students must prepare an .ipynb file with the solution. Please provide necessary comments for each subtask via markdown-interface of ipython notebook. Please provide reasonable visualization for each subtask (plots, scatters).

The .ipynb file needs to be sent via Telegram to @evgeny_zavalnyuk. PDF version of the notebook (interface «Download as PDF») must be submitted to gradescope for tracking and evaluation.

Dataset
https://gitlab.com/evgeny.zavalnyuk/sirius_ml_labs/-/tree/main/ps4/mdata
This dataset has to be downloaded locally [git clone …]

Description of the task
This is an extension of ps #4,#5, #6
Please use
https://gitlab.com/evgeny.zavalnyuk/sirius_ml_labs/-/blob/main/ps7/ps7.ipynb
as a template of your solution

Before starting to solve the set, we higly recommend to look over the case study
https://towardsdatascience.com/pytorch-lstms-for-time-series-data-cd16190929d7

(2 points)

Run LSTM model on the market microstructure features, with two additional liner layer on top of it, and hyperbolic tangent activation function.

Use 70% of the dataset for train purpose, 30% for validation.

Parameters of the model are defined in the ps7.ipynb.

Please enjoy the given interface to track the progress of the learning.