Master's Degree in Computer Science, University of Udine

Course of Deep Learning, Academic Year 2021/2022

Practice Assignment

Course held by: Prof. Giuseppe Serra, Dr. Beatrice Portelli,
Dr. Giovanni D'Agostino

Dipartimento di Scienze Matematiche, Informatiche e Fisiche

Udine, 26/11/2021

Instructions

You can do this assignment **either alone or in group of 2**. Please submit no more than one submission per group.

Be sure to write your code in **Python**; you can make full use of the libraries seen in class (pytorch, pandas, sklearn, numpy, matplotlib...); please justify the usage of eventual additional libraries you used in your code.

Your source code should be appropriately (but not excessively) commented; your code can be presented as either a **.py** (python source code) or a **.ipynb** (Jupyter Notebook) file. If you have more than one source code file please name them in such a way that it is easy to understand what's inside each file and which file contains the code to run your work.

Alongside your code please submit a **short written description** (preferably in **.pdf**) of what you have done, including used methods, results, discussion, conclusions and eventual references; in case of multiple source code files please summarize the content of each .py/.ipynb file. The report *should not exceed 4 pages*.

Put both the source code in a **.zip** file named after your student id: "studentid.zip" for an individual work and "idstudent1_idstudent2.zip" for a work in group of two, e.g. 012345_678901.zip". The .zip file must be uploaded to the appropriate folder of the E-Learning page of this course no later than the specified deadline.

DEADLINE: Thursday December 2nd, 2021 (you have time up until midnight).

Assignment

You are given a real-life weather dataset, which you can find in the file *"weather_dataset.csv"*. The data consists in hourly measurements of weather features in a city; the five measured variables are:

- Wind speed in $0.1m/s$
- Wind direction in degrees (360 North, 90 East, 0 No wind)
- Temperature in $0.1°C$
- Dew Point in $0.1°C$
- Air Pressure in $0.1hPa$

The dataset has the shape ($Timesteps \times Features$). Use the last 168 samples (one week of recording) as test case, and the remaining ones for training your

model. Scale your data before training and scale them back to be able to compare your predictions with real measurements. Think about what kind of neural networks model could be useful for time series data. Although RNNs are an option, it is possible to use other models discussed in the course lectures such as CNNs. Try your preferred model and compare its performance in terms of *mean absolute error* with real test records. Be sure to:

(a) Implement a (deep) neural networks model to predict the *Temperature* of the city.

(b) The model should run in a recursive fashion. Meaning the predictions of temperature should be sent back and used for up-comping ahead predictions. For other variables you may use the real measurements. Run your model to predict the temperate of the city for one week (i.e. your model should predict the temperature for 168 samples in a recursive fashion). Note that you are not allowed to use the real 168 temperature samples in your model.

(c) Use visualisation techniques to plot your results in addition to reporting the obtained numbers in a Table for comparison purpose.

(d) (*Optional)* Can you come up with a strategy to determine or visualise which features among the 5 measured features might be more relevant for the outputs of the model? In addition, experiments how many previous time steps you need to feed in to your network to give you the best possible answer you can obtain by your model?

Good luck!