Politecnico di Milano

Artificial Neural Networks and Deep Learning

Second Homework

Assignment: Time Series Forecasting

Anas Bahtaoui, Théo Saulus
Centrale International Team

January 21, 2022

Description of the homework

In this second homework, we aim to forecast time series composed of 7 variables which names do not provide information regarding the nature of the data studied.

Our approach was to work first with direct forecasting of the whole sequence to predict. Then, we used an autoregressive method, to incrementally build the prediction based on previous predictions and data. We finally tried to implement more complex models based on attention mechanism and seq2seq principles, but to no avail.

Data preparation

The dataset provided showed time series labelled with uninformative names, which prevents any attempt to draw intuitive conclusion regarding the possibility to make use of any correlation in the prediction process. This is however of no influence on the direct and autoregressive methods that we are to use.

First, we divided our time series in training and validation sets, by only using a proportionally small validation set: 1000 values out of the 67000+ available. This choice is based on the idea that the unknown test set used in Codalab is the continuation of the available data, which means that the very last data of the model is especially important to make further predictions. Thus, training the model on as much data as possible seems sensible.

Then, we normalised data based on training data (and obviously not based on the whole dataset, to prevent any use of test data when computing the validation accuracy). This normalisation allows to use as input values in the networks quantities that are easier to compute on, and less prone to explode or vanish.

First models

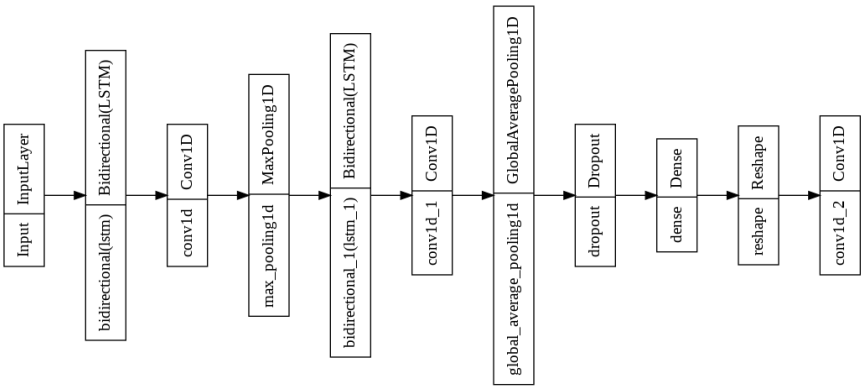
We first explored some very basic models available in keras like GRU and LSTM just by curiosity however they didn't give us any intersting results.

Our first attempt was via direct forcasting.

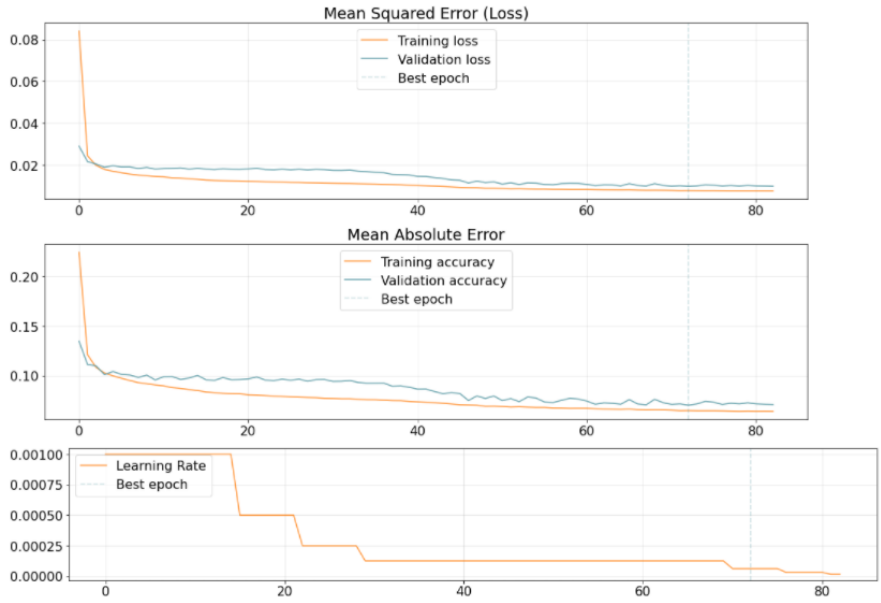
Direct Forecasting

We chose a window size 300 and stride 15 was what gave us the best results having experimented with different other values.

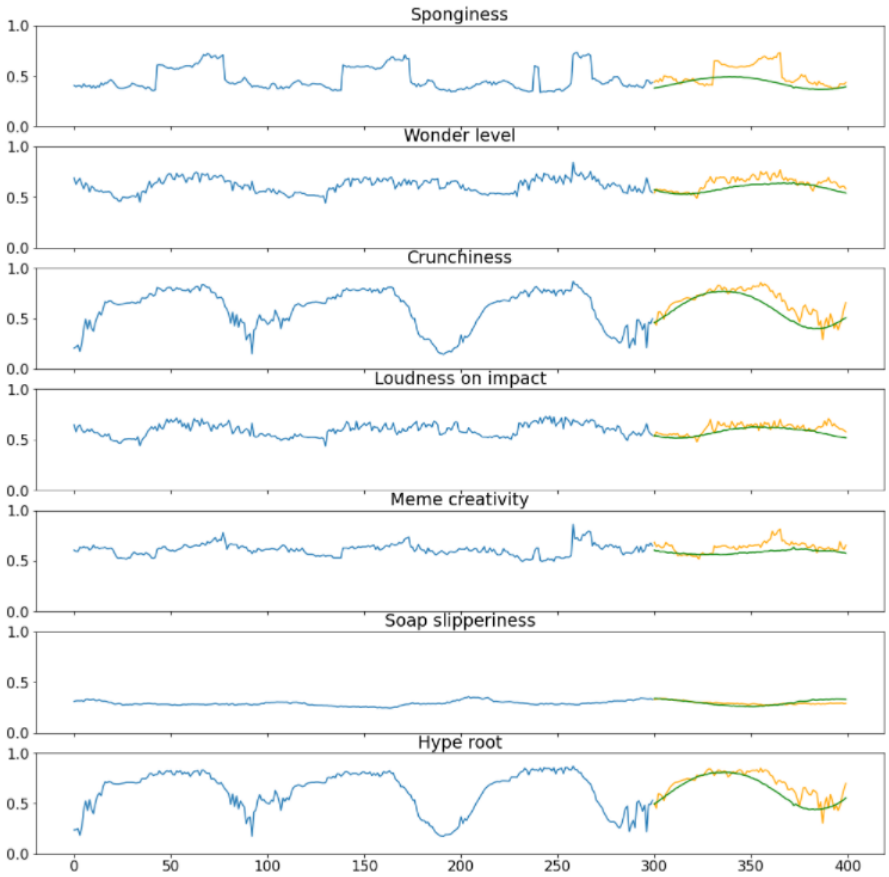
We set a telescope of 100. And then build our training and validation sequences. We built a neural network using bi-directional LSTM layers followed by a convolutional layer and a max pooling layer.



We trained the model for 63 epochs. And plotted the Mean Square Error and the Mean Absolute Error.



We also plotted our test evaluation (comparing prediction and our local test data) as well as the prediction of the future.



Result : We got an RMSE of 4.6 of the hidden test data.

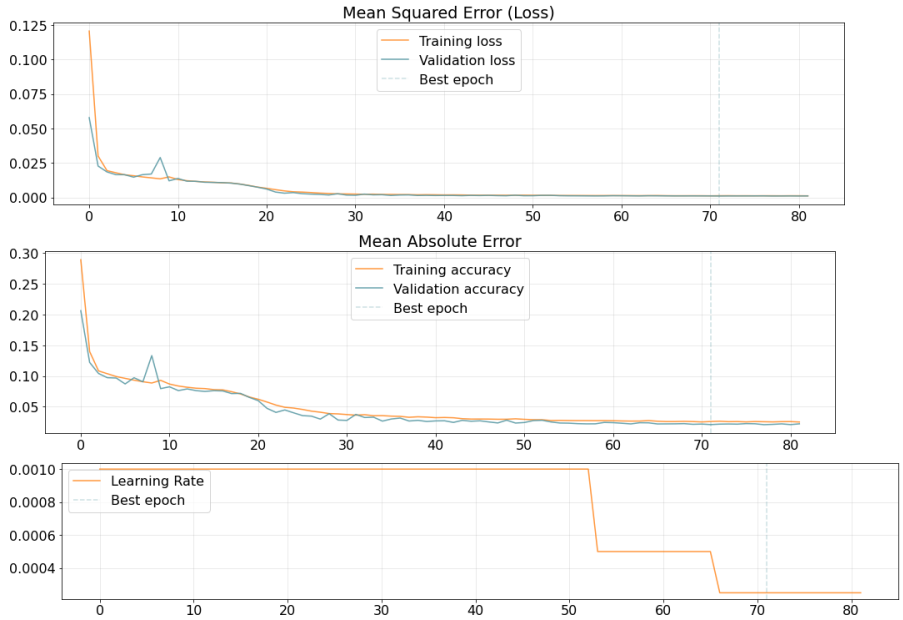
Autoregressive Model

Our second model and also our best model was trying an autoregressive model.

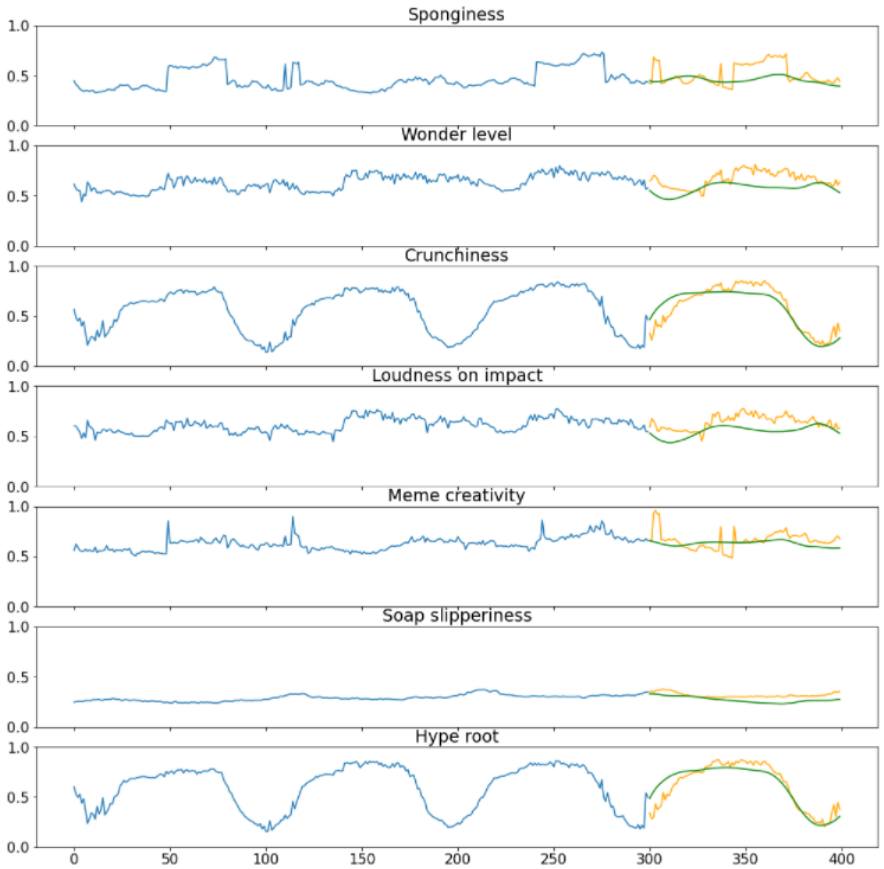
Setting a telescope of 1, we built the following model.

At first we had a patience of 5 but at the end we chose a patience of 10 epochs because we saw that training the model a little longer gave us a better result.

We trained the model for 82 epochs. And similarly we report here the plots of the mean Square Error and the Mean Absolute Error.



The plots of our test evaluation (comparing prediction and our local test data) as well as the prediction of the future.



Result : We were able to get an RMSE of 4.3 on the hidden test data.