

BIT Recruitent AI assignment

Introduction

One of our clients gave Bit the assignment to create a model which predicts the stock market of certain share. Since you haven't got much experience doing this, you decide to simplify the problem and first see if you can predict another similar signal. This signal is similar in pattern yet less noisy. On the internet, you have found a dataset on which you try to fit your predictive model, with the goal to predict the future signal.

Assignment

In this assignment, you will implement a program that aims to predict a periodical signal. A training sample of the data is given. The program should allow to do predictions for the future course of this data; meaning that datapoints beyond the range of the traingsdata should be predicted. Three different traing samples are provided:

- One with a 1-hour interval of datapoints
- One with a 5-hour interval of datapoints
- One with a 30-minutes interval of datapoints

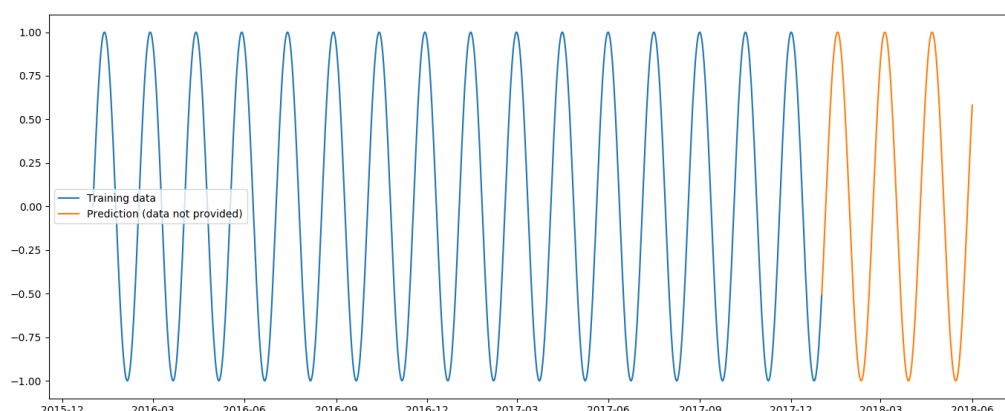
You can decide yourself which dataset you will use. The first colomn of the data contains the timestamp of the measurement, the second colomn contains the signal value at that time. All three datasets consist of the same timeframe and only differ in ΔT . i.e. if one would plot all signals one would get very similar graphs.

We advise you to start with the following steps:

1. Load the CSV-file into your python file and plot the signal, this to get a feeling for the actual behavior of the signal over time.
2. Choose a model which you want to implement. You're free to choose which model, yet we propose to use something different from the ARIMA models. We have seen many attempts using this method and it often didn't succeed. We advise you to do some research before starting with the implementation of the first method!
3. Consider how to define both your inputs and your targets. Again you're free to choose whatever you like, yet it would be nice to have a clear explanation ready.
4. Consider how you perform your prediction. Will you consider a new prediciton each timestep or will you predict time batches?

Requirements for the program:

- Code should be written in Python 3.
- Code should be executable from a *main.py* file.
- The model should make predictions for the next 6 months given the training sample.
- The code should include plotting functionalities for both the training data and the predicted signal. Caution: the figure below is not a correct representation of your data sample.
- ARIMA is a famous method for timeseries predictions. We prefer to see the use of another method over the use of ARIMA!



Deliverables:

- Code in a zip-file.
- The same zip-file should include both the plot of the trainingssample aswell as the plot of the 6-month prediction in a similar format as seen below.

Points of interest

- Readability of the code.
- Consider normalisation of the data
- Think about your code structure and try to work with classes.
- Dependent on the model you choose, training can consume a lot of time. Consider this from the start!

After the weekend we will see eachother again and you will give us a demonstration of what you have done! Good luck!