Time Series Forecasting Project

Data Science International Summer School
Predeal
August 28, 2018

Razvan C. Bunescu
School of Electrical Engineering and Computer Science
http://ace.cs.ohio.edu/~razvan
bunescu@ohio.edu
Time Series Forecasting

**Task:** Train and evaluate an RNN-based model for time series prediction.

**Models:** LSTMs or GRUs implemented in PyTorch, or in other high-level package, e.g. Keras:
- Compare against other baselines, e.g. ARIMA, GP, SVR, …

**Data:** Start with a simple artificial time series:
1. Sine function.
2. Sine + linear.
3. Sine with amplitude growing linearly in time.
   • Add noise …
Time Series Forecasting

**Data**: Continue with a real-life dataset. Identify a dataset of interest or use an existing dataset:

1. **Univariate time series**:
   - Airline passengers (from James’ Amazon Drive).
   - More than 500 time series, including finance and economic: [https://datamarket.com/data/list/?q=provider:tsdl](https://datamarket.com/data/list/?q=provider:tsdl)

2. **Multivariate time series**:
   - Predict ship’s position [https://www.myshiptracking.com/](https://www.myshiptracking.com/)
Time Series Forecasting

- **Implementation:**
  1. Starting from a PyTorch tutorial on RNNs/LSTMs.
  2. Starting from the `LSTMPredict.ipynb` Keras code from James’ folder (rewrite in PyTorch?).
  3. Start from my PyTorch code for sine prediction: [http://ace.cs.ohio.edu/~razvan/courses/mlds18](http://ace.cs.ohio.edu/~razvan/courses/mlds18)

- **Tuning:**
  - capacity, regularization, pre-training, …

- **Data pre-processing:**
  - Scaling?, Predict absolute value vs difference?, Extrapolation of missing data?