Knapp, Stephen

Multi-step LSTM Autoencoder/decoder for Multi-variate Time Series

Problem Statement: Compress a multi-variate time series model into a lower dimensional space and tune the model to optimize stable and accurate predictions.

- This experiment uses a multi-step autoencoder/decoder LSTM network which receives a multi-variate time series input to predict future values of temporal slices (in this case temporal slides are weeks).
  - Simply put: the model will predict the next 46 weeks of energy consumption by day of the week based on the last four weeks of data from 8 time series inputs
- Identify the effect of different levels of tuning parameters:
  - # of epochs
  - # of neurons in the LSTM layers
  - # of dense layers in the decoder
  - Dropout
  - Kernel regularization

High Level Overview of Steps:

1. Created Google Colab Notebook: googlecolab.com
2. Loaded dataset from link below to my Google Drive and mounted my drive to my Colab notebook
3. Cleaned/Pre-processed data
4. Converted Data to Daily
5. Created a naïve model as a baseline
6. Ran example provided in reference below and re-ran using 4 weeks of training data
7. Determine optimal number of epochs
8. Experiment with number of neurons in LSTM encoder and decoder layers
9. Experiment with depth of the decoder layer
10. Experiment with using dropout
11. Experiment with kernel regularization
12. Compare model root mean square error boxplots, medians, ranges and stability

Dataset: Data is four years' worth of a single household's power usage sampled every minute. Seven variables originally, with the 8th calculated using the original 7.

URL:
https://archive.ics.uci.edu/ml/datasets/individual+household+electric+power+consumption

Large size: 129MB; Toy size: 6MB

Hardware: Windows 10 Pro, Intel® Core™ i7-6700HQ CPU @ 2.6GHz 2.59GHz, 8GB RAM, 64-bit OS

Software: Python 3.6.7 on Google Colab running a GPU

Reference: Deep Learning for Time Series Forecasting: Predict the Future with MLPs, CNNs and LSTMs in Python by Jason Brownlee; Section 20.8

Acknowledgements of Data Set: UCI machine learning repository, "Household Power Consumption"

Tutorial Links:

Two minute (short): https://youtu.be/oFNoOs1J-30

15 minutes (long): https://youtu.be/FOwgHquJe2g