CS670/470 Term Project Phase 1: Advanced Machine Learning in Time Series Analysis Using Auto-Regressive Integrated Moving Averages(ARIMA)

Yong Zhuang, Wei Ding, UMass-Boston 04/09/2019

1 Educational Goal

Design and implement advanced machine learning algorithms for time series analysis; Practice how to apply the Auto-Regressive Integrated Moving Averages(ARIMA) algorithm for time series analysis.

2 Details

Project Goal: Using state-of-the-art machine learning libraries to apply the ARIMA algorithm for streamflow forecasting.
Project Data: Twenty-one years from 1996 to 2016, daily streamflow of the Ganges river in India.
Due Date: 2:00 pm, ~~April 18~~ **April 23**, 2019
Programming language: Python.
Package and framework: scikit-learn, statsmodels, pandas, numpy, colab.

3 Tasks

3.1 Task 1

Read the online lecture notes at https://colab.research.google.com/drive/1IMsoRnkzt7-GMtzYBCKgQ-px7Si89t1Y to learn the workflow of time series analysis and the principle of the ARIMA model.

3.2 Task 2

Load the streamflow data set from http://kdl.cs.umb.edu/CS670/data/Ganges_1996_2016.csv. Split it into 2 parts, (1996 – 2011) for training and (2012 – 2016) for testing.

3.3 Task 3

Convert training set into stationary time series. Create and tune your ARIMA model on the training set in order to find the best values of parameters (p,d,q). You can use the Auto-correlation analysis and Partial Auto-correlation analysis to find the best value of p and q, or you can use Grid Search to find the best values of parameters. Any work of data preprocessing work which can improve the model's performance is encouraged.

3.4 Task 4

Train your model on the training set with the best values of parameters you got. Evaluate your ARIMA model on the test set, and calculate the root mean square error (RMSE). Draw a curve to show the original data and your prediction.

4 Submission Requirements

1. One team only need to submit one solution (colab notebook).
2. Share your team colab notebook with teaching.yong@gmail.com.
3. Submit your team colab notebook at online blackboard through the team lead's UMassOnline account.