COMP 6940 Assignment 5

Date Due: Friday 27th April, 11:59pm

Part 1: Literature Review 30 marks

In no more than **500 words** write a literature review of the following paper: <u>Load Forecasting using Deep Neural Networks</u>

Part 2: Forecasting 40 marks

Context

To better follow energy consumption, the government wants energy suppliers to install smart meters in every home in England, Wales and Scotland. There are more than 26 million homes for the energy suppliers to get to, with the goal of every home having a smart meter by 2020.

This roll out of meter is lead by the European Union who asked all member governments to look at smart meters as part of measures to upgrade our energy supply and tackle climate change. After an initial study, the British government decided to adopt smart meters as part of their plan to update our ageing energy system.

In this dataset, you will find a <u>refactorised version of the data</u> from the London data store, that contains the energy consumption readings for a sample of 5,567 London Households that took part in the UK Power Networks led Low Carbon London project between November 2011 and February 2014. The data from the smart meters seems associated only to the electrical consumption.

There is infomations on the ACORN classification details that you can find in this <u>report</u> or the website of CACI.

Weather data for London area was added, the data was collected from the <u>darksky api</u>.

Dataset

There are two datasets that will be used for this assignment:

1. <u>Daily Weather data for the period November 2011 until April 2014.</u>
2. <u>Energy Consumption data for an area for the period November 2011 until April 2014.</u>

Tasks

The **"temperatureHigh" and "time"** data from the weather dataset should be used for analysis. The **"energy_sum" and "day"** data from the energy dataset should be used for analysis.

Ensure that there is only one row for a given date.

1. Identify any trends in the datasets and discuss if trends in weather are related to trends in energy consumption.

2. Identify Seasonal patterns in the datasets and discuss if seasonality in weather are related to seasonality in energy consumption.
3. Forecast using simple exponential smoothing the expected **temperature** for 2014, use the RMS metric to indicate the accuracy of your forecast.
   1. Choose the alpha that results in the lowest RMS
   2. Plot a graph showing the actual value and the forecasted values
4. Forecast using simple exponential smoothing the expected **energy consumption** for 2014, use the RMS metric to indicate the accuracy of your forecast.
   1. Choose the alpha that results in the lowest RMS
   2. Plot a graph showing the actual value and the forecasted values
5. Combine the initial datasets using an appropriate column as the index. Indicate why you chose that index.
6. Plot separate graphs showing the **"temperatureHigh"** and **"energy_sum"** data from the combined dataset.
7. Discuss what observations can be made from the plot with the combined data vs the plots from the individual datasets.

Part 3: Prediction 30 marks

This section relies on the combined dataset from part 2 question 5.

Tasks

1. Create a new column in the dataset, derived from the **"energy_sum"** values, which indicates if energy consumption was high or low. An appropriate threshold should be used to determine whether or not energy consumption was high or low. Justify the threshold value you chose.
2. Split the dataset into a train and test set. The train set should be all the data before 2014 and the test set should be all data from 2014.
3. Choose the appropriate columns to be used for training. Give a brief explanation why those columns were chosen.
4. Use an appropriate classifier to predict the energy consumption for 2014.
5. Show the performance of the classifier using an appropriate metric.

Submission Details:

Students should submit either the following:

• An exported Jupyter notebook of the assignment with notes

or

• A python script file with the assignment and appropriate comments.

Email submissions to:

nchamansingh@gmail.com

With the following subject format: *< StudentID > COMP6940 Assignment 5*

Late Submissions:

A daily penalty for late submissions of 15% per day or partial day