UG Project Registration Form  
 ----------------------------  
To the student: Complete this form, then email it to your g.roberts@cs.ucl.ac.uk and your internal supervisor. Please use the subject line "UG Project Registration Form".

To the supervisor:  Please add your comments (if any) to the bottom of the form, then email a copy to g.roberts@cs.ucl.ac.uk.  Please use the subject line "UG Registration Confirmation".

**Student's Name:** Keshav Aggarwal

**Degree Programme:** MEng Computer Science

**Supervisor(s) Name:** Fabio Caccioli, Simone Righi

**Project Working Title:** Analysing and Improving Statistical and Machine Learning Forecasting for Finance

**Brief Outline of Aims:** To successfully apply state of the art Statistic and Machine Learning techniques in the context of forecasting Financial data, enhance accuracy using Sliding Window approach and perform comparative analysis.

I intend to analyse the use of certain statistical and machine learning algorithms to forecast financial time-series data majorly pertaining to the Bonds and Commodities markets. Using ‘*Statistical and Machine Learning forecasting methods: Concerns and ways forward*’ (Makridakis, Spiliotis and Assimakopoulos, 2018) as the foundation, my work will specifically revolve around forecasting financial time-series data with emphasis on measuring the accuracy of the prediction and with an eye on the complexity of the methods used. This is will involve reproducing some of the paper’s results and carrying out comparative performance evaluation using suitable accuracy measures. This will be followed by extending the work to make use of the Sliding Window Framework approach to enhance the accuracy of ML algorithms and re-establish the performance levels to make suitable comparisons. I will then study the output of these parameters. Time permits, I will carry out automated data pre-processing in order to speed up computations and allow better data modelling.

**Intended goals:**

* Evaluation and collection of suitable datasets pertaining to the Bonds and Commodities markets for different time-periods and frequencies (potentially from Bloomberg and Reuters terminal)
* Brief literature review of statistical and machine learning algorithms and their application in forecasting data
* Implement a suitable testing environment to reproduce the work done by the aforementioned paper
* Implement Machine Learning algorithms (Bayesian Neural Network, K-Nearest Neighbour Regression, Support Vector Regression) and Classic Statistical methods (Naïve 2 (RW Benchmark), Damped, ETS) for time-series forecasting
* Run each method above through accuracy measures such as MaxAPE, sMAPE, MASE and RMSE and carry out a comparative analysis.
* Extend on the methods proposed by the peers using Sliding Window Framework (and automated Pre-processing) to enhance the accuracy and improve the complexity of the methodologies
* Re-evaluate the methods to re-establish the performance levels for each algorithm

SUPERVISOR'S COMMENTS:

Second Marker Suggestions (by Supervisor):

**Bibliography:**

1. Makridakis, S., Spiliotis, E. and Assimakopoulos, V. (2018). *Statistical and Machine Learning forecasting methods: Concerns and ways forward*. [online] Available at: https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0194889.