CSE-403 PROJECT PROPOSAL

On

FORECASTING FINANCIAL SERIES USING CLUSTERING METHODS AND SUPPORT VECTOR REGRESSION

SUBMITTING TO: **Prof. KAPIL AHUJA**

By:

RAHUL KUMAR	**190001049**
HARSH VARDHAN AGRAWAL	**190005016**
VIVEK BHUSHAN	**190005043**

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

INDIAN INSTITUTE OF TECHNOLOGY INDORE

Proposal: Two-Stage Model for Forecasting Financial Time Series

Introduction:

Financial time series forecasting has been an important topic in the finance industry for many years. Accurate forecasting of financial time series is critical for making informed decisions related to investments, trading, and risk management. In this proposal, we aim to replicate and extend the findings of the paper "Forecasting financial series using clustering methods and support vector regression" by Lucas F.S. Viela and Rafael C. Leme (2018).

Objective:

The main objective of this proposed project is to replicate the findings of the paper by Lucas and Rafael (2018) and to understand the effect of clustering on patterns before using Support Vector Regressions (SVR).

Methodology:

In this project, we plan to use the same financial time series data used in the paper, which is composed of the values of an equity fund of a Brazilian bank. We will implement the two-stage model proposed in the paper, which uses clustering methods to segment the time series into its various contexts in the first stage, and then employs support vector regressions (SVRs) to forecast future values of the series in the second stage.
To replicate the results of the paper, we will use the same clustering and SVR algorithms as used in the paper, namely the K-means clustering algorithm and the SVM regression algorithm. We will evaluate the performance of the proposed two-stage model by comparing it to the hierarchical model (HM) presented in the literature. Additionally, we will also compare the proposed model to other popular models such as the support vector machine (SVM) and the multilayer perceptron (MLP) models.

To extend the research, we plan to conduct further analysis of the clustering and forecasting methods used in the paper. Specifically, we aim to investigate the impact of using different clustering algorithms on the forecasting accuracy of the proposed model. We will also explore the effect of varying the number of clusters used in the model on its performance. Furthermore, we will investigate the relationship between the clusters formed by the model and the volatility of the financial time series.

Expected Outcomes:

We expect that our work will be consistent with the findings of the paper by Mendes and Lemonge (2018). Additionally, we anticipate that our analysis of the clustering and forecasting methods will provide valuable insights into the mechanisms underlying the model's performance.

Conclusion:

The replication and extension of the two-stage model for forecasting financial time series proposed by Lucas and Rafael (2018) will provide valuable insights into the use of clustering methods in financial time series forecasting. The proposed project will contribute to the development of more accurate forecasting models, which will aid in making informed decisions related to investments, trading, and risk management.