Universiti Teknologi MARA

Forecasting Malaysian Gold Price Using Random Forest

Muhammad Nur Firmanrulah bin Samsudin

Thesis submitted in fulfilment of the requirements for Bachelor of Computer Science (Hons.) College of Computing, Informatics, and Mathematics

January 2024

ACKNOWLEDGEMENT

In the name of Allah SWT, the Most Merciful and Forgiving. All praise and thanks to Allah SWT and His great mercy for allowing me to finish this thesis. Thank Allah SWT for all the possibilities, challenges, and strength that have gone into finishing this endeavour.

Firstly, I would like to convey my deepest appreciation to my project supervisor, Dr Khairul Anwar for helping me finishing up this project by giving me guidance, beneficial information, assistance, and tips regarding this project. Thank you, Dr, for advising and supporting me throughout this project and for devoting me your precious time in helping me.

In addition to that, I would also like to express my special thanks to Madam Ummu Fatihah Binti Mohd Bahrin for your exceptional consideration, expertise, and support during Project Formulation (CSP600) and Project (CSP650) subject. The ongoing attention to our study has had a huge influence on my overall experience of learning.

Finally, not to forget, I would also like to express gratitude to my mother and my family for their countless support and encouragement towards me during the completion of this project. Moreover, all my friends and classmate also had play part in this project. Thank you for always giving advice, and acknowledgement when it was required.

ABSTRACT

Gold is a yellow valuable metal that is used to make coins, jewellery, attractive artefacts, and many other things. Gold is the most popular and outperforms other metals when used as an investment instruments. The gold prices are influenced by supply and demand. Estimating its future pricing remains a difficult undertaking due to the complex and volatile structure of financial markets. Previously, manual prediction being done to forecast gold prices. Developing this model can save their time in predicting gold prices. Random forest appears to be the best model for predicting gold prices. Dataset is gathered from multiple sources and being merge into one file. Dataset being split into training and testing for ratio 90/10. This ratio being chosen after some experiment being held. The training set will use to generate subset for each decision tree. After that, random forest will be created to add tree into forest until number of trees reached. Next, random forest will make prediction from decision tree and will aggregate predictions from all trees. Lastly, the prediction result will be print in user interface. The coefficient of determination, mean absolute error, mean square error and root mean square error being used to evaluate the random forest model. The accuracy for the best model being tune is 99.71% and producing the smaller error. In conclusion, the model, with an accuracy of 99.71% and minimized errors, offers a time-saving and reliable tool for forecasting gold prices, showcasing its efficiency in financial market analysis and investment decision-making.

TABLE OF CONTENTS

CHAPTER THREE: METHODOLOGY