Predicting Gold Prices Using Machine Learning

Kasoji Prashanthi chary

Abstract

This study aims to predict future gold prices using machine learning techniques. By leveraging historical gold prices and external financial indicators, including oil prices and the US dollar index, regression models were employed to generate price forecasts. Linear regression served as the baseline model, while decision trees, enhanced with hyperparameter tuning, were applied for more advanced predictions. The results, evaluated using R-squared and Mean Absolute Error (MAE), demonstrated that model tuning significantly improved prediction accuracy. These insights into gold price trends can aid investors and market analysts in navigating the volatile gold market.

Introduction

Background: Gold is a critical asset in the global economy, often seen as a hedge against market volatility

Objective: The purpose of this project is to predict gold prices using machine learning, aiding investors in making informed decisions.

Significance: Machine learning models offer potential for more accurate forecasts by capturing non-linear relationships in financial data.

Methodology – Dataset

- **Data Sources:** Historical gold prices, oil prices, and the US dollar index.
- **Features:**
- Gold prices (target variable)
- Oil prices, US dollar index, and other economic indicators (predictor variables)
- **Preprocessing:**
- Data cleaning (handling missing values)
- Feature scaling (normalization)
- Data split into training and test sets

Methodology – Machine Learning Models

- **Baseline Model:**
- Linear Regression: Simple and interpretable but limited for non-linear relationships.
- **Tuned Model:**
- Decision Tree Regressor: Flexibility in handling non-linear data.
- Hyperparameter tuning was performed using grid search for optimizing tree depth and splitting criteria.

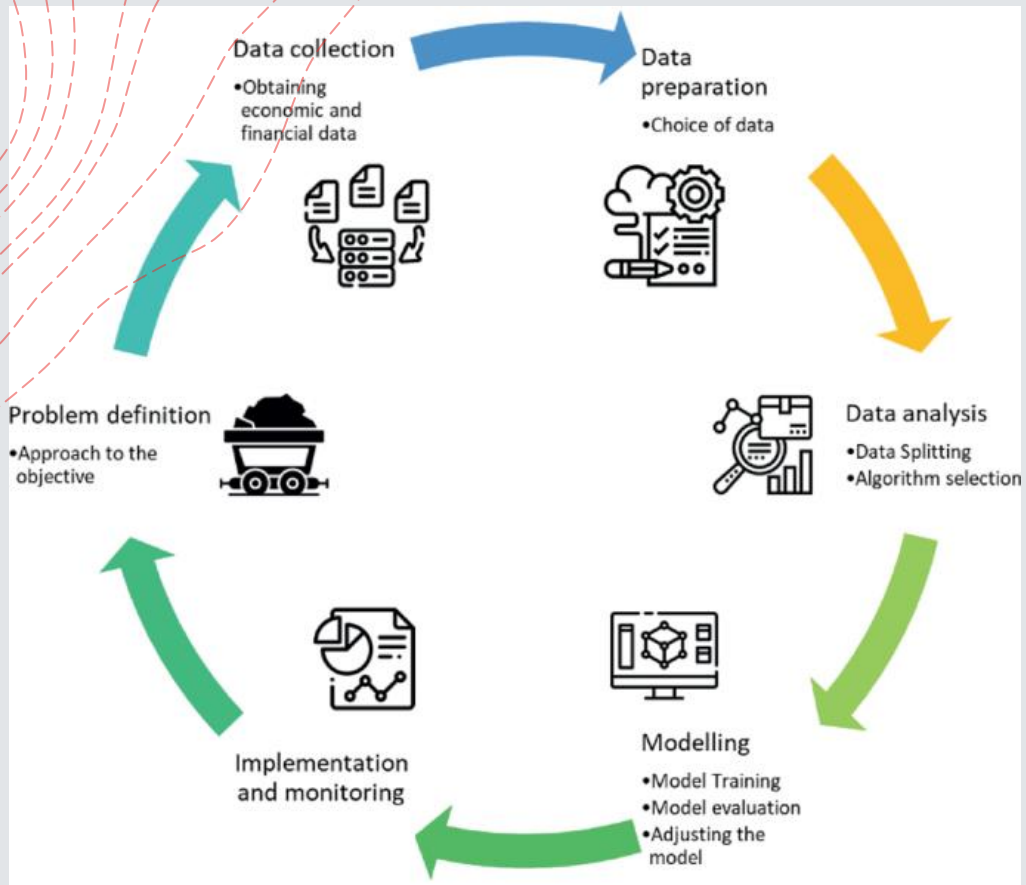
Methodology – Model Optimization

Cross-Validation:

5-fold cross-validation was used to ensure model generalizability.

Hyperparameter Tuning:

Grid search tested different configurations for the decision tree regressor to find optimal model settings.



Results

- **Linear Regression:**
  - R-squared: 0.65
  - MAE: 30.15
- **Decision Tree (Tuned):**
  - R-squared: 0.82
  - MAE: 22.85
- **Visualization:**
  - Line chart comparing actual vs. predicted gold prices.
  - Feature importance plot for the decision tree model.

Conclusion

The tuned decision tree model offered a robust solution for predicting gold prices, outperforming the baseline linear regression.

Applications:

This model can be applied in financial institutions for gold price forecasting.

Future Research:

Incorporating more financial indicators like inflation rates or geopolitical events could further improve prediction accuracy.

THANK YOU