

A PROJECT REPORT ON

“Gold Price Prediction System

Using Machine Learning”



Session: 2025-2026

DEPARTMENT OF INFORMATION TECHNOLOGY
HALDIA INSTITUTE OF TECHNOLOGY

Submitted To:- Debasish Sahoo

Submitted By :-

1. Swagatam jana – 10300223179
2. Spandan Maity -- 10300223168
3. Soumadwip Manna -- 10300223163

Gold Price Prediction System using Machine Learning

1. Introduction

2. Gold is one of the most valuable and widely traded commodities in the world. Its price fluctuates daily due to various economic factors such as inflation, stock market trends, currency exchange rates, crude oil prices, and global economic uncertainty. Accurate gold price prediction is important for investors, traders, financial institutions, and policy makers.
3. The Gold Price Prediction System is a Machine Learning-based solution designed to analyze historical financial data and predict future gold prices. The system uses structured financial indicators and historical gold price trends to forecast upcoming prices with high accuracy.

The system is designed to:

- Predict future gold prices
- Analyze historical trends
- Identify influencing economic factors
- Provide data-driven investment insights
- Support financial decision-making

This project integrates Machine Learning models with a Flask backend API and a React frontend to create a complete end-to-end intelligent gold price prediction system.

4. Objectives of the Project

5. The primary objectives of this project are:
6.
 - I. To analyze historical gold price data using Machine Learning techniques.
 - II. To identify important financial indicators affecting gold prices.
 - III. To build a regression-based prediction model.
 - IV. To compare multiple ML algorithms for best accuracy.
 - V. To evaluate model performance using standard metrics.
 - VI. To provide accurate future price predictions.
 - VII. To demonstrate real-world financial forecasting using ML.

7. Dataset Description

The dataset used in this project contains historical gold price records along with related economic indicators.

Features Included:

Financial Indicators:

- SPX (S&P 500 Index)
- USO (Crude Oil Prices)
- SLV (Silver Price)
- EUR/USD (Currency Exchange Rate)

Target Variable:

- GLD (Gold Price)

The dataset contains several years of historical data collected from financial markets.

Data Characteristics:

- Time-series structured data
- Numeric financial indicators
- No categorical variables
- Daily price records

The dataset was cleaned and preprocessed before training the models.

4. System Architecture

The system follows a simple Machine Learning workflow architecture:

Data Layer

- Historical gold price dataset
- Financial indicators dataset

Processing Layer

- Data Cleaning
- Handling Missing Values
- Feature Selection
- Train-Test Split

Machine Learning Layer

- Model Training using scikit-learn
- Model Evaluation
- Performance Comparison

Output Layer

- Predicted Gold Price
- Model Accuracy Score
- Visualization Graphs

5. Machine Learning Models Used

5.1 Linear Regression

Type: Regression

Purpose: Establish a linear relationship between gold price and financial indicators.

Reason: Simple and interpretable baseline model.

5.2 Random Forest Regressor

Type: Regression

Reason:

- Handles non-linear relationships
- Reduces overfitting
- High prediction accuracy

This model performed best in most cases.

5.3 Model Evaluation Metrics

The models were evaluated using:

- Mean Absolute Error (MAE)
- Mean Squared Error (MSE)
- R² Score

R² Score measures how well the model explains the variance in gold prices.

6. Feature Engineering

Feature engineering was performed to improve prediction accuracy.

Correlation Analysis

A correlation heatmap was created to identify relationships between:

- Gold Price
- Crude Oil Price
- Stock Market Index
- Currency Exchange Rate
- Silver Price

Observations:

- Gold price has strong correlation with Silver price.
- Moderate correlation with USD exchange rate.
- Inverse relationship with stock market trends in some periods.

No categorical encoding was required since all features were numeric.

7. Results and Output

The system provides:

- Predicted Gold Price
- Actual vs Predicted comparison graph
- Model Accuracy (R^2 Score)
- Error metrics (MAE, MSE)

Random Forest Regressor achieved higher accuracy compared to Linear Regression.

The predicted prices closely followed actual gold price trends, demonstrating strong model performance.

8. Applications and Use Cases

- Investment planning
- Stock market analysis
- Financial risk assessment
- Commodity trading strategies
- Portfolio management
- Economic forecasting

This system can assist investors in making data-driven decisions.

9. Limitations

- Market volatility can affect prediction accuracy
- Model depends on historical data patterns
- Sudden global events (war, pandemic, crisis) may cause unpredictable fluctuations
- Limited financial indicators used

10. Future Enhancements

- Integration with real-time stock market API
 - Deep Learning models (k-fold cross validation)
 - Web-based dashboard for investors
 - Deployment on cloud platform
 - Inclusion of more macroeconomic indicators
-

11. Conclusion

The Gold Price Prediction System demonstrates the practical application of Machine Learning in financial forecasting. By analyzing historical financial data and training regression models, the system can accurately predict future gold prices.

Random Forest Regressor provided strong performance due to its ability to handle complex, non-linear relationships between financial indicators and gold prices.