SmartBridge Internz Artificial Intelligence Project

Time Series Analysis of Gold price using Prophet

Madhumitha Kuduva, 20BCE7646, VIT-AP

Mavillapalli Puneeth , 20BEC0079, VIT-VELLORE

Mithisha Brilent Tavares, 20BCI7052, VIT-AP

Yashashvi Kala, 20BIT0180, VIT-VELLORE

3-7-2023

Content:

1 INTRODUCTION

1.1 Overview:

The project is focused on implementing a time series analysis for gold price prediction using the Prophet library in Python. Creating a forecasting model that can correctly project future gold prices based on previous data is the aim.

In order to guarantee that the data is in the right shape for analysis, the project entails gathering historical gold price data and preparing it. After that, the data is investigated using statistical calculations and visualizations to spot patterns and trends.

To train the prediction model, the Prophet library is used. Prophet is renowned for its ease of use and adaptability when working with time series data.

The model may be used to forecast future gold prices once it has been trained and validated. A specific date can be entered by users for prediction. Based on the learned patterns, the model forecasts gold prices and generates predictions.

Users can interact with the model using a web interface that is also part of the project. Users may enter their preferred prediction criteria into the interface and look at the projected gold prices.

The project's overall goal is to give customers a trustworthy and easy-to-use tool for gold price prediction by utilizing time series analytic methods and the Prophet library's forecasting capabilities.

1.2 Purpose:

The goal of this project is to develop a Python-based time series analysis tool called Prophet for gold price prediction. The project has several uses and can be useful in a variety of circumstances:

- **Making Informed Investment Decisions:** Investors, traders, and financial analysts may use gold price forecasts to make wise investment choices. The ability to predict future gold prices might be useful in spotting market opportunities and dangers.

- **Risk management:** Organizations engaged in the gold trade or sectors impacted by gold prices might make use of the forecasts for risk management. Companies are able to modify their strategy, control inventories, and protect themselves from price swings by obtaining an estimate of future gold prices.

- **Financial Planning:** This project might be useful to people or families interested in gold investments or financial planning connected to gold. People can plan their savings, retirement funds, or other financial objectives involving gold assets by having knowledge of future gold prices.

- **Research and Analysis:** This project may be used by academics, economists, and researchers who are interested in examining the trends and patterns of gold price changes. It is possible to support research projects, statistical analysis, and modeling pertaining to the gold market by having access to reliable gold price projections.

- **For educational purposes**, this project may be used to teach Python programming, time series analytic methods, and how to use the Prophet library in forecasting software. The project code and approach may be explored and studied by students or anyone else interested in learning about and using time series forecasting.

2 LITERATURE SURVEY

2.1 Existing Problem:

Depending on the unique objectives and data properties, there are a variety of existing systems and methods for predicting the price of gold using time series analysis. Among the often employed techniques are:

- **Moving Average (MA):** This method predicts future gold prices by averaging a defined window of previous gold prices. It offers a straightforward and understandable strategy but could miss subtle data trends.

- Time series forecasting frequently makes use of **ARIMA (Autoregressive Integrated Moving Average)** models. They apply the ideas of moving average, differencing, and autoregression to the data in order to capture trends, seasonality, and random fluctuations. ARIMA models need parameter tweaking and data stationarity.

- Based on weighted averages of previous data, **exponential smoothing (ES)** is used. They give more weight to current observations while progressively diminishing the weight of older data points.

- Regression models, random forests, support vector machines (SVM), and gradient boosting are a few examples of **machine learning techniques** that may be used to forecast gold prices. To capture complicated correlations in the data, these methodologies take into account

a variety of input features, including past pricing, economic indicators, sentiment analysis, or other pertinent elements.

- Recurrent neural networks (RNN) and long short-term memory (LSTM) networks are two examples of **deep learning models** that have demonstrated promising results in time series forecasting. These models can recognize complicated patterns in the data as well as sequential linkages.

2.2 Proposed Solution:

Using the Prophet Library to forecast gold prices is the suggested option for this project. Facebook created Prophet, a potent time series forecasting tool that offers a specific method for modeling and forecasting time series data.

The method involves the following steps:

1. Data gathering: Compile historical gold price information from the website goldpricez.com, ideally in the form of a time series with dates and associated gold values.

2. Data Preprocessing: Perform the required preprocessing operations on the data, including converting dates to the correct date/time format, resolving missing values or outliers, and making sure the data is in a format that will be useful for analysis.

3. Model Training: To build a forecasting model, use the Prophet library. Prophet uses an additive regression model to analyze the data and identify trends, seasonality, and holiday impacts. These elements are automatically found by the library and added to the model.

4. Fit the Prophet model to the historical data on gold prices. Based on the supplied data, the model will calculate the trend, seasonality, and other pertinent factors.

5. Future Period Prediction: Utilize the learned model to produce predictions for upcoming intervals. Indicate how many future timeframes or exact dates you want to use to make your gold price prediction.

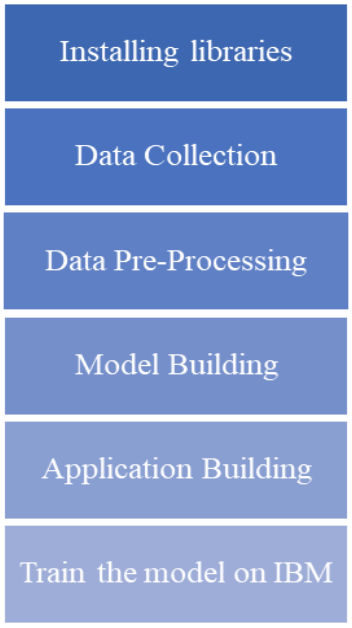
6. Visualization: To see the trend and expected values, visualize the historical data and projected gold prices.

7. Evaluation: Utilize the evaluation measure Mean Absolute Error (MAE) to judge the performance and accuracy of the gold price forecasts. This procedure aids in evaluating the model's performance.

Prophet's suggested approach has a number of benefits, including the capacity to manage intricate time series patterns, adaptability in adding seasonality and holiday impacts, and automated trend change detection. It makes time series forecasting easier to use and more accessible to users of all skill levels by providing a user-friendly interface.

3 THEORITICAL ANALYSIS

3.1 Block diagram:



3.2 Hardware and Software design:

Hardware Requirements:

· Computer or Server: A device that can execute the necessary software and carry out data processing operations.

· Storage that is adequate for the project's supplementary datasets and the historical gold price information.

· Processing power: A CPU with enough horsepower to effectively perform model training, data preparation, and prediction activities.

Software Requirements:

- Python: A programming language used for implementing the project.
- Jupyter Notebook or any Python IDE: To write and execute the project code.
- Python libraries:
    1. pandas: For data manipulation and preprocessing.
    2. prophet: The time series forecasting library used for gold price prediction.
    3. matplotlib or plotly: For data visualization.
    4. Flask: To develop and deploy the web application.
- Web development technologies: HTML, CSS, and JavaScript for designing the user interface and creating interactive elements in the web application.

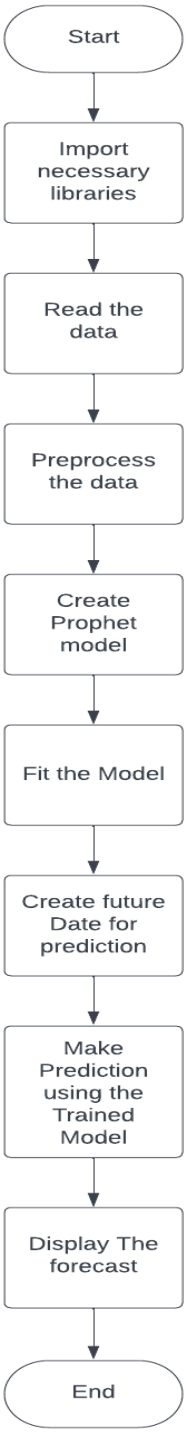
4 EXPERIMENTAL INVESTIGATIONS

During the development of the code for Time Series Analysis and Gold Price Prediction using Prophet, several key steps were undertaken to analyze the data and ensure accurate predictions. These steps include:

1. Data Loading and Preprocessing: The initial step involved loading the gold price dataset and preprocessing it to ensure it was in the appropriate format for analysis. The 'Date' column was converted to the datetime format, and the column names were adjusted to match the expected format of Prophet.
2. Exploratory Data Analysis (EDA): Before proceeding with the time series analysis, it is essential to conduct exploratory data analysis. This includes examining the distribution of gold prices, identifying any missing values or outliers, and assessing the general patterns and trends within the data. EDA helps in understanding the characteristics of the dataset and identifying any preprocessing steps required to improve the analysis.
3. Model Training and Evaluation: The next step involved creating an instance of the Prophet model and fitting it to the preprocessed dataset. The model learns from the historical data to capture trends, seasonality, and other patterns. To evaluate the model's performance, various evaluation metrics specific to time series forecasting, such as Mean Absolute Error (MAE) or Root Mean Squared Error (RMSE), can be calculated by comparing the predicted values to the actual values. Here we used MAE.

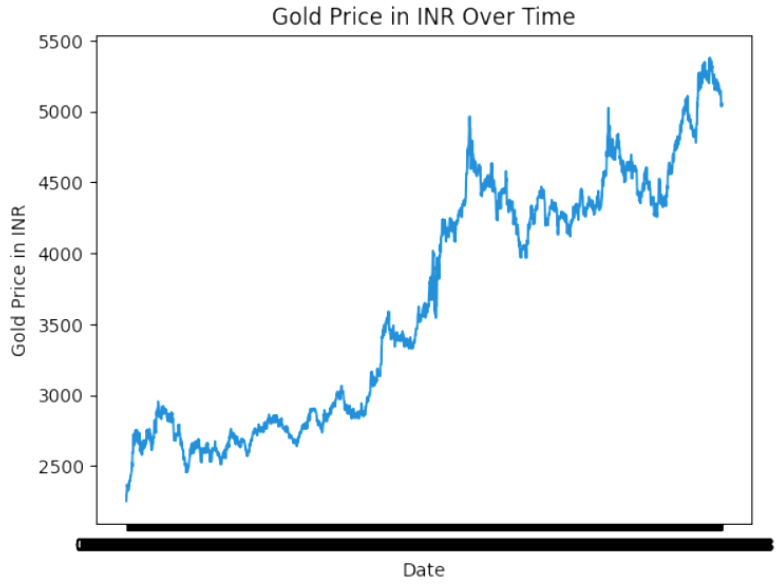
4. Future Date Generation and Predictions: To make future predictions, a set of future dates was generated using the **make_future_dataframe()** function. These dates extended beyond the existing dataset, allowing the model to forecast gold prices for a specified number of periods into the future. The model's **predict()** function was then used to generate predictions for these future dates.
5. Visualization of Results: Visualizing the forecasted values is crucial for understanding the predicted trends and patterns. The **plot()** function provided by Prophet was used to plot the historical gold prices, the predicted values, and the uncertainty intervals. This visualization helps in assessing the model's ability to capture the underlying trends and provides a visual representation of the future predicted prices.

5 FLOWCHART

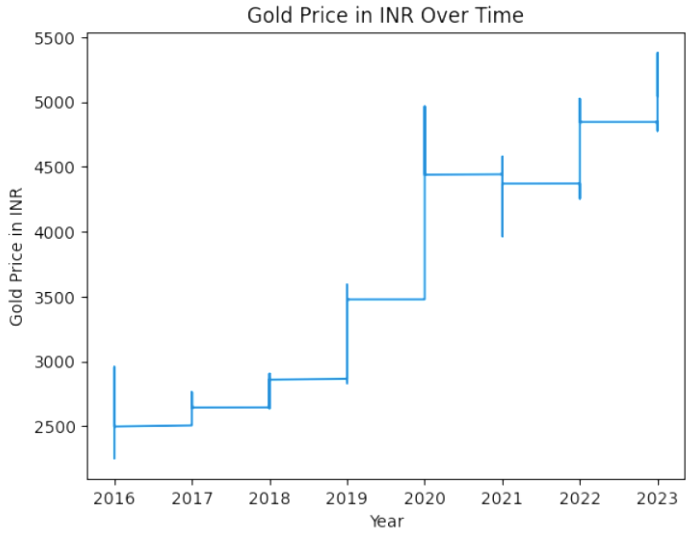

Start
Import necessary libraries
Read the data
Preprocess the data
Create Prophet model
Fit the Model
Create future Date for prediction
Make Prediction using the Trained Model
Display The forecast
End

6 RESULT

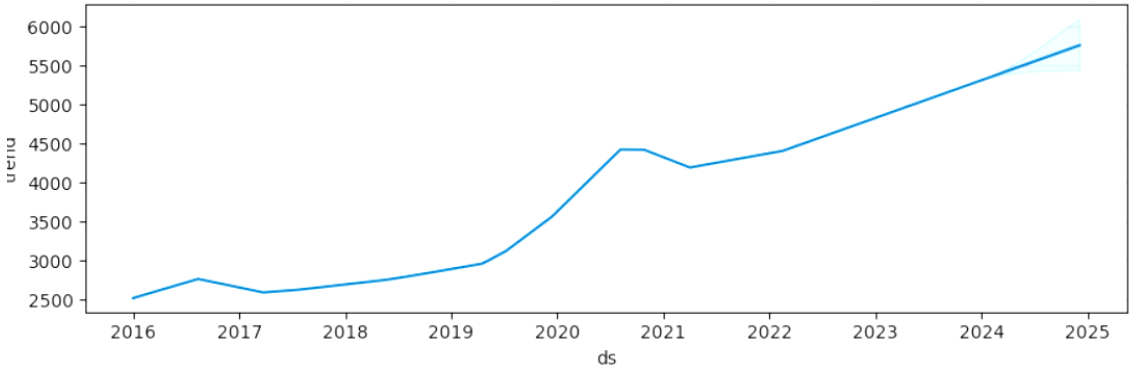
Variation of the gold price over time:

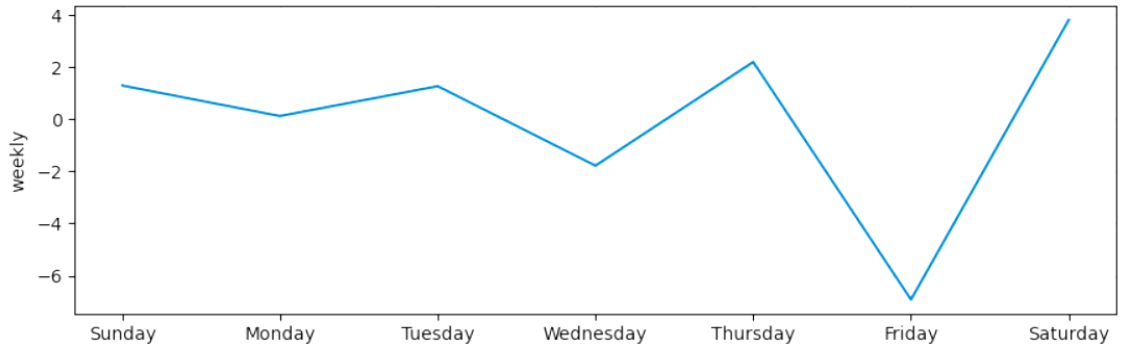


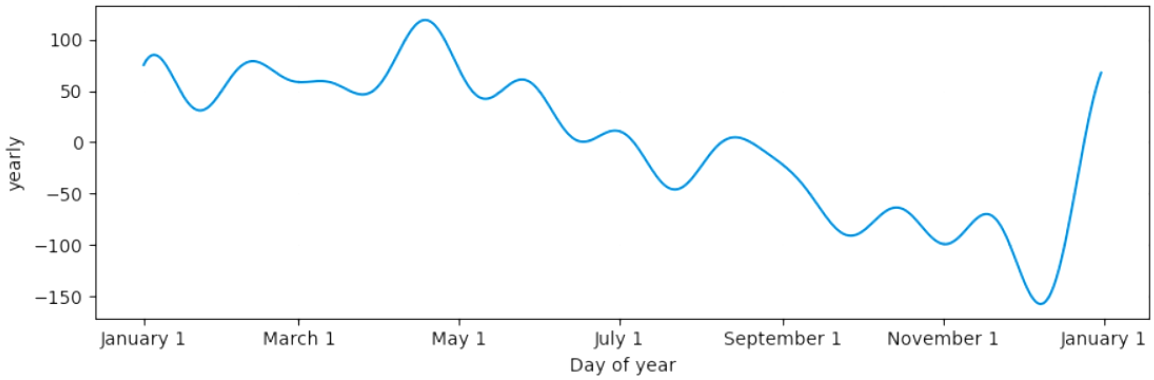
Change in the gold prices w.r.t. each individual year:

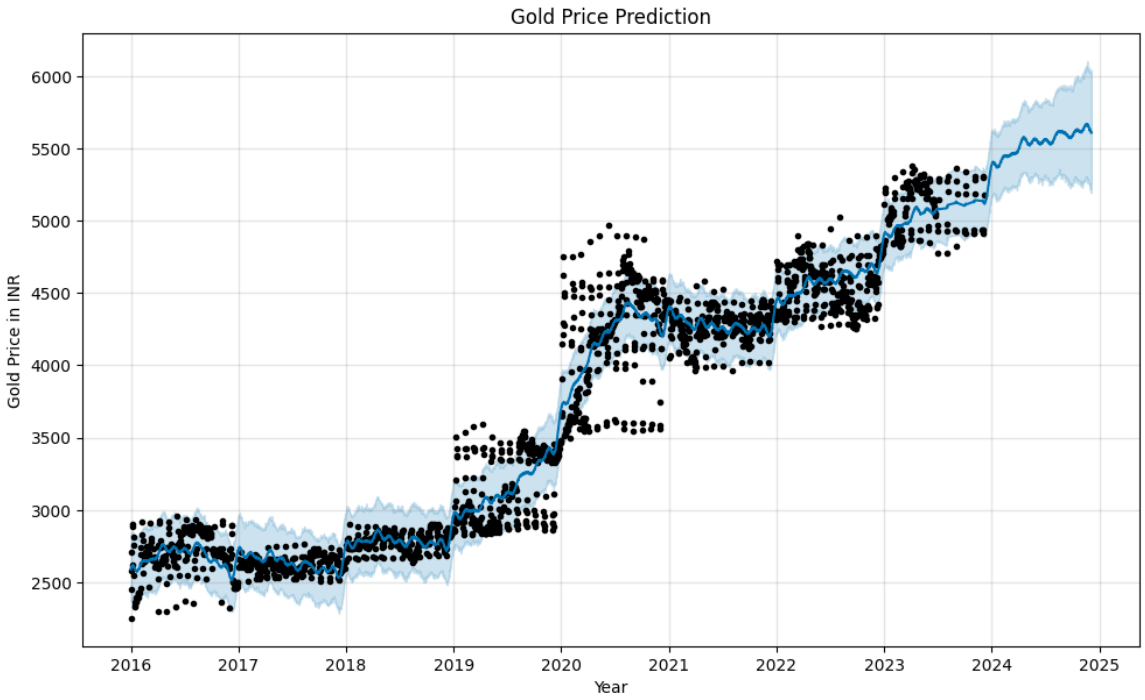


The forecast of gold prices:


6000
5500
5000
4500
4000
3500
3000
2500
2016
2017
2018
2019
2020
2021
2022
2023
2024
2025
ds


4
2
0
−2
−4
−6
weekly
Sunday
Monday
Tuesday
Wednesday
Thursday
Friday
Saturday


100
50
0
−50
−100
−150
yearly
January 1
March 1
May 1
July 1
September 1
November 1
January 1
Day of year



7 ADVANTAGES & DISADVANTAGES

Advantages of the model:

1. Simple Implementation: Prophet is designed to provide a straightforward and easy-to-use interface for time series analysis. It abstracts away many of the complexities involved in modeling and forecasting, making it accessible to users with varying levels of expertise.

2. Automatic Seasonality Detection: Prophet automatically detects and incorporates seasonality patterns present in the data. It can handle multiple types of seasonality, such as daily, weekly, monthly, and yearly patterns, without requiring explicit manual specification.

3. Flexibility in Handling Missing Data: Prophet is equipped to handle missing data and outliers, which are common issues in real-world time series datasets. It utilizes advanced imputation techniques to fill in missing values, ensuring a more robust and accurate analysis.

4. Incorporation of Additional Covariates: Prophet allows the inclusion of external variables or covariates that may influence the gold price, such as economic indicators, geopolitical events, or market factors. This flexibility enables more comprehensive and contextual predictions.

5. Uncertainty Estimation: Prophet provides a mechanism for estimating and quantifying the uncertainty associated with its predictions. It generates prediction intervals that capture the range of possible future values, allowing for a more comprehensive understanding of the forecasted gold prices.

Disadvantages of the model:

1. Limited Customization: While Prophet offers a user-friendly interface, it may lack the flexibility and customization options required for certain advanced modeling techniques. Users with specific requirements or complex data patterns might find the predefined modeling capabilities of Prophet restrictive.

2. Dependence on Historical Patterns: Time series analysis heavily relies on historical patterns to make predictions. However, if the gold market experiences significant changes or unexpected events that deviate from historical patterns, the accuracy of the predictions may be compromised

3. Sensitivity to Outliers: While Prophet can handle outliers to a certain extent, extreme or influential outliers can still affect the model's performance. Careful preprocessing and outlier detection techniques are necessary to ensure accurate predictions.

4. Limited Consideration of External Factors: Although Prophet allows for the inclusion of external covariates, its focus primarily lies on capturing and forecasting inherent time series patterns. Complex relationships with external factors may not be fully captured or may require additional advanced modeling techniques.

8 APPLICATIONS

A few real life applications of Time Series Analysis for Gold Price Prediction using Prophet are:

1. Financial Investment Decision-Making: Time series analysis using Prophet for gold price prediction can aid investors and financial institutions in making informed decisions regarding gold investments. By accurately forecasting future gold prices, investors can optimize their trading strategies, portfolio allocations, and risk management techniques.

2. Risk Management: Time series analysis can assist in risk assessment and management for various stakeholders involved in the gold market. By understanding and predicting gold price

movements, risk managers can identify potential risks, estimate market volatility, and develop risk mitigation strategies.

3. Trading and Hedging Strategies: Traders and speculators in the commodities market can leverage gold price predictions to devise effective trading and hedging strategies. Time series analysis helps identify patterns, trends, and potential turning points in gold prices, allowing traders to optimize entry and exit points, minimize losses, and maximize profits.

4. Mining Industry Planning: Gold mining companies can utilize time series analysis to forecast gold prices for future mining operations. Accurate predictions can guide decisions related to exploration, production volumes, resource allocation, and overall business planning in the gold mining sector.

5. Jewelry and Retail Pricing: Retailers in the jewelry industry can leverage gold price predictions to adjust their pricing strategies and inventory management. By understanding the expected fluctuations in gold prices, retailers can optimize their pricing decisions, ensure competitive pricing, and manage their inventory levels effectively.

6. Economic Analysis: Gold prices are often indicative of broader economic conditions and market sentiment. Time series analysis of gold prices using Prophet can contribute to economic analysis by providing insights into inflationary pressures, currency fluctuations, geopolitical tensions, and macroeconomic trends.

9 CONCLUSION

In this report, we conducted a time series analysis and gold price prediction using Prophet. The analysis was based on a dataset containing historical gold prices. The goal was to leverage the Prophet library to forecast future gold prices and gain insights into the underlying trends. By plotting the forecasted values, we could observe trends and potential patterns in future price movements. The results of the analysis provide valuable insights into the potential direction of gold prices. However, it is important to note that predictions made by the model are subject to various factors and uncertainties that may impact actual gold prices. External factors such as geopolitical events, economic conditions, and market sentiment can influence gold prices and may not be fully captured by the model. Overall, this analysis demonstrates the potential of Prophet for time series analysis and gold price prediction. By leveraging this powerful library, we can gain valuable insights into gold price trends and make informed decisions in the dynamic and ever-changing market environment.

10 FUTURE SCOPE

Incorporating additional features: One potential future scope for this project is to incorporate additional relevant features that may impact gold prices. For example, you can consider including macroeconomic indicators such as inflation rates, interest rates, GDP growth, and stock market indices. By including these features, the model may be able to capture more complex relationships and provide more accurate predictions. Additionally, you can explore the inclusion of sentiment analysis data from news articles or social media, as market sentiment can play a significant role in gold price movements.

1. Hyperparameter tuning and model optimization: Another future scope is to conduct a thorough hyperparameter tuning and model optimization process. Prophet offers several tunable parameters, such as seasonality settings, trend flexibility, and regularization terms. By performing systematic experimentation and optimization techniques, you can fine-tune these parameters to improve the model's performance and prediction accuracy. Techniques like grid search or Bayesian optimization can be employed to efficiently search for the optimal combination of parameters.

2. Ensemble modeling and model comparison: Ensemble modeling involves combining multiple models to improve predictive accuracy. In the future, you can explore the implementation of ensemble techniques such as stacking or blending. By training and combining multiple Prophet models with different settings or using other time series forecasting algorithms (e.g., ARIMA, LSTM), you can create a more robust and accurate ensemble model. Additionally, it would be beneficial to compare the performance of Prophet with other forecasting methods to assess its effectiveness and determine if alternative models provide better predictions for gold prices.

11 BIBILOGRAPHY

http://goldpricez.com/

https://github.com/facebook/prophet

https://www.geeksforgeeks.org/time-series-data-visualization-in-python/

https://machinelearningmastery.com/time-series-prediction-lstm-recurrent-neural-networks-python-keras/

https://www.kaggle.com/datasets/lakshmi25npathi/gold-price

https://www.kaggle.com/code/itsbitan/gold-price-prediction

12 APPENDIX

Source code:

https://github.com/YashashviKala20BIT0180/VIT-AI/blob/main/gold_price_pred.ipynb