International Journal of Progressive Sciences and Technologies (IJPSAT)
ISSN: 2509-0119.

Vol. 42 No. 2 January 2024, pp. 191-200

Assessing Manual Dataset Creation For Xauusd Market Prediction : A Comparative Study Logistic Regression And Decision Tree Model

Andri Purwoko[1], Bambang Suharjo[2], Richardus Eko Indrajit[3], H.A. Danang Rimbawa[4]

[1]Cyber Defense Engineering
Republic of Indonesia Defense University
Jakarta, Indonesia
andripurwo82@gmail.com

[2]Faculty of Science and Defense Technology
Republic of Indonesia Defense University
Jakarta, Indonesia
bambang_suharjo@tnial.mil.id

[3]Pradita University
Banten, Indonesia
indrajit@alumni.harvard.edu

[4]Faculty of Science and Defense Technology
Republic of Indonesia Defense University
Jakarta, Indonesia
hadr71@idu.ac.id

Abstract— This study aims to develop a simplified dataset for more effective market prediction, focusing on the Forex trading of XAUUSD (Gold/USD). The dataset was gathered from the TradingView platform, covering the period from March 4, 2023, to December 21, 2023. The data collection method involved intensive observation of daily and weekly charts, utilizing Daily and Weekly Moving Average (MA) indicators and the concept of breakout. The analysis focused on measuring the distance between the Daily MA at the beginning and end of the period (start and stop), and utilizing this data for entry strategy in the following three time periods. The trading strategy adopted involves the simultaneous use of Buy and Sell orders, with a Stop Loss (SL) to Take Profit (TP) ratio of 1:2. TP was adjusted to accommodate aggressive price movements, while SL remained constant. The collected data was meticulously recorded and stored in Excel format for further analysis.

With the prepared dataset, this research applies two AI models, Logistic Regression and Decision Tree, to predict the best trading decision – Buy or Sell. The study aims not only to create a useful dataset for market prediction but also to compare the effectiveness of two different AI methods in the context of Forex trading of XAUUSD. The results are expected to provide insights into which model is more accurate and efficient in analyzing and predicting market trends, with practical implications for traders and market analysts.

Keywords— Market Prediction, XAUUSD (Gold/USD), Simplified Dataset, Logistic Regression, Decision Tree Models, Data Analysis.

I. INTRODUCTION

According to Bisnis.com, the stock market and foreign exchange, or FOREX market, are among the most popular financial markets in the world of investing. Investments in these two markets can yield high returns but also come with high risks. The primary activity in both markets is trading [1]. Weerathunga and Silva [2] suggest that the FOREX market is one of the largest in the world. Foreign exchange involves simultaneous buying and selling of foreign currencies. These transactions are typically executed in pairs [3]. The FOREX market is the largest financial market, with daily transaction values exceeding 2 trillion US dollars. Unlike traditional financial markets with physical locations, the FOREX market operates electronically through networks of banks, companies, and individual traders. Additionally, it functions 24 hours a day during business days [3]. Foreign exchange trading not only involves foreign currencies but also encompasses commodities such as Gold, Silver, and Oil. Among these, gold is one of the most valuable commodities globally [4]. As stated by Astonacci.com, gold ranks as one of the strongest investments after the USD [5]. Investors have begun to trade gold against foreign currencies, including the Euro, Swiss Franc, Australian Dollar, and notably against the USD. The high risk associated with FOREX trading has motivated many researchers to develop robust models for predictive analysis.

In the dynamic spheres of FOREX and commodities trading, particularly in the XAUUSD (Gold/USD) market, the creation of predictive datasets has become a cornerstone of financial analytics. This process has evolved from the early stages of electronic trading, where manual data collection was the norm. Traders and analysts would meticulously record market dynamics, focusing on key indicators such as price movements and trading volumes. This traditional approach laid the groundwork for understanding complex market behaviors, paving the way for the development of effective trading strategies.

Backtesting is a critical method in this context, serving as a bridge between historical market analysis and the formulation of reliable trading strategies. It involves a retrospective examination of market data, with a focus on indicators like Daily and Weekly Moving Averages (MAs) and breakout patterns. The primary goal of backtesting is to determine the probability of various trading outcomes and to calculate the win-rate of different strategies. This method allows traders to simulate trading strategies against historical data, offering a risk-free environment to gauge their potential effectiveness. A successful backtest provides a solid foundation for a trading strategy, indicating its feasibility and the associated risks, win-rates, and probabilities. This step is crucial for any trader seeking to develop a robust approach to the market.

Once a trading strategy has passed the backtesting phase with promising probabilities and win-rates, the integration of Artificial Intelligence (AI) becomes a logical next step. AI models, such as Logistic Regression and Decision Tree models, are employed to further refine these strategies. By applying AI, traders aim to enhance the predictive power of their strategies, potentially increasing their effectiveness. The use of AI in this context is not just about enhancing returns; it also serves as a test to validate whether the back tested strategies align with the complex, real-world market dynamics. The expectation is that AI can provide a more nuanced and sophisticated analysis, leading to better-informed trading decisions. This study, therefore, not only explores the effectiveness of traditional backtesting methods but also investigates the impact of AI integration on the overall performance and reliability of trading strategies in the XAUUSD market.

II. LITERATURE REVIEW

Research on FOREX, not only uses a machine learning approach but also uses a model from Reinforcement learning (RL) as done by [6] using RL to predict FOREX, which uses a values-based approach and a policy-based approach. [7] uses the Deep learning method to improve prediction results on the stock market with several technical analyzes, namely Stochastic %K, Stochastic %D, Momentum, Rate of Change, William's %R, A/D Oscillator, and Disparity 5. [8] used the Long-Short Term Memory (LSTM) model with the Forex Loss Function (FLF-LSTM) indicator in making predictions on FOREX. Based on the results of the research, the approach of the FLF-LSTM model was able to reduce the error by 13% when compared to using the ARIMA indicator. Research on FOREX with commodities was also carried out by [9] using the Deep Reinforcement Learning approach, which succeeded in proving that Deep Reinforcement Learning can increase the ability to trade automatically. Other indicators were also introduced by [10] in predicting the movement of stock market prices using deep learning with engineering features. The method introduced by [10] is a Multi-Filter Neural Network (MFNN) with financial time series features and price movement prediction. This pilot study aims to see a model from machine learning that has a fairly high level of accuracy in making FOREX predictions. [11] comparing the moving average convergence/divergence (MACD) indicator with the faster take profit signal feature with the simple MACD indicator and the results obtained between both MACD does not have a significant difference. [12] introduces LR2GBDT, a new method that combines logistic regression (LR) with gradient boosted decision trees (GBDT) to forecast stock index changes. This approach aims to improve prediction accuracy in stock market trading. [13] focuses on predicting financial distress in companies using decision tree methods and logistic regression. It examines data from the Taiwan Stock Exchange and finds that decision trees are more accurate for short-term predictions, while logistic regression performs better in the long-term. The study highlights the effectiveness of combining AI methods with traditional statistical approaches in financial distress prediction. [14] explores the use of data mining techniques for predicting stock prices, focusing on both technical and fundamental information. The authors develop a framework for classifying industrial stock performances and design a trading strategy. Their methodology aims to outperform the Australian market. Simulation results show that their selected stock portfolios exceed the Australian All-Ordinaries Index, demonstrating the effectiveness of their analytical approach in stock selection and trading strategy.

III. Research Method

This study employs a comprehensive backtesting methodology, utilizing historical data from the XAUUSD (Gold/USD) market, obtained from the TradingView platform. The data encompasses a specific period from March 4, 2023, to December 21, 2023. Central to this methodology are the Daily and Weekly Moving Average (MA) indicators and breakout patterns, which are instrumental in the formation of the dataset. The backtesting approach involves a detailed analysis of three preceding time periods to inform each trading decision. This historical examination is integral in determining the Stop Loss (SL) and Take Profit (TP) thresholds and in identifying the starting price for each trading instance. A unique aspect of this strategy is the simultaneous execution of buy and sell orders, adhering to a risk-reward ratio of 2:1 for TP and SL. Notably, the TP is flexible to adjust to market fluctuations, whereas the SL is fixed, ensuring consistent risk management.

The progression of each trade is meticulously monitored and documented, extending over the subsequent three time periods. This documentation includes critical information such as entry and exit points, attainment of TP or SL, and the overall duration of each trade. Such comprehensive data collection is vital for capturing the nuances of market behavior under the applied trading strategy. All collected data for this study have been meticulously gathered and systematically logged in an Excel spreadsheet, resulting in the creation of a comprehensive dataset that encompasses the entire duration of the defined research period. This dataset not only acts as a rich historical archive, detailing market reactions to the various strategies implemented, but it also forms the foundational platform for subsequent, more in-depth analysis. This further analysis leverages cutting-edge Artificial Intelligence (AI) techniques, offering a sophisticated approach to interpreting and understanding the data. The underlying structure and approach of this study are visually represented and can be comprehensively understood by examining the framework depicted in the figure provided below:

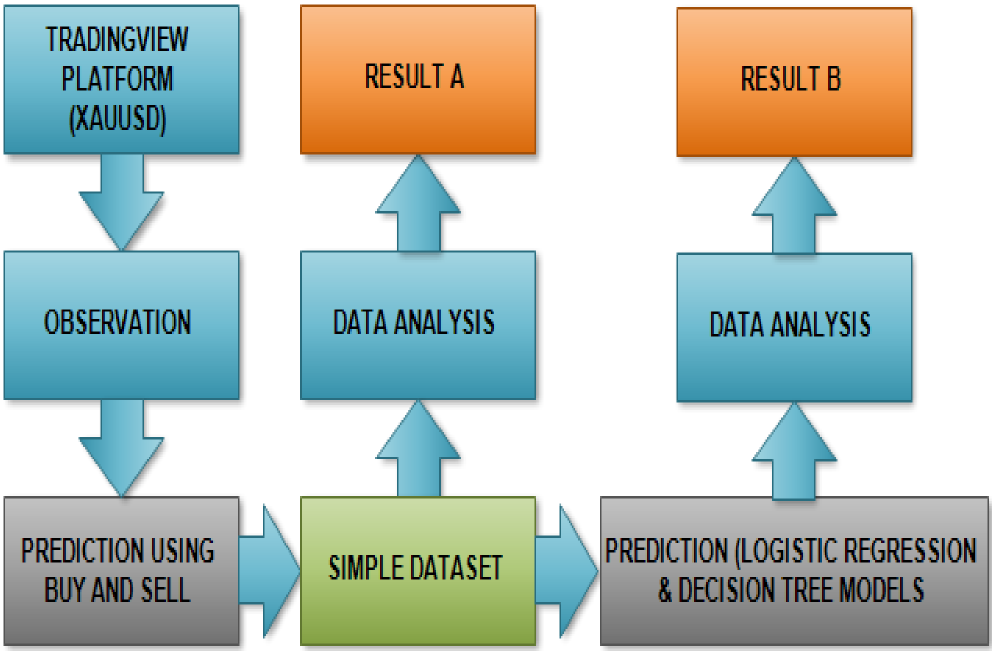


Fig 1: Framework XAUUSD prediction

The framework provided outlines a structured approach to analyzing and predicting stock market performance, specifically applied to the XAUUSD pair on the TradingView platform. It begins with a period of observation, where market behaviors and trends are closely monitored. From this observational data, two distinct predictive paths are followed. The first path employs a straightforward technique, utilizing buy and sell signals to inform predictions. These signals are compiled into a simple dataset, which is then analyzed to yield Result A. This result likely represents the effectiveness of a basic trading strategy based solely on these direct signals.

In parallel, the second path takes a more refined analytical approach. It harnesses advanced statistical models, namely logistic regression and decision tree models, to predict market movements. This sophisticated method integrates a broader range of data, potentially encompassing both technical and fundamental analysis, to create a more comprehensive prediction model. The analysis of this complex dataset leads to Result B, which is indicative of the performance of a more nuanced trading strategy. Ultimately, the framework compares the outcomes of both methodologies—Result A from the simple buy/sell strategy and Result B from the advanced model-based strategy—to determine which provides a more accurate forecast of stock performance, offering insights into the potential benefits of integrating complex analytical techniques in stock market predictions.

DATE	START	STOP	TP	SL	PREDICTION	RESULT	PERSENTAGE	GROW
04/03/2023	1827.066	1852.521	1853.430	1813.429	1	1	1,39%	35,65%
	1827.066	1840.399	1813.429	1853.430	-1	1	-0,73%	35,65%
07/03/2023	1852.521	1829.187	1898.279	1829.187	1	-1	-1,26%	35,65%
	1852.521	1814.338	1806.459	1875.551	-1	-1	2,06%	35,65%
09/03/2023	1814.338	1910.704	1842.218	1799.490	1	1	5,31%	35,65%
	1814.338	1828.581	1785.853	1828.581	-1	1	-0,79%	35,65%
17/03/2023	1920.418	1941.478	2013.906	1873.160	1	1	1,10%	35,65%
	1920.418	1967.162	1826.416	1967.162	-1	1	-2,43%	35,65%
22/03/2023	1941.478	1977.194	2050.383	1886.829	1	1	1,84%	35,65%
	1941.478	1995.865	1832.312	1995.865	-1	1	-2,80%	35,65%
25/03/2023	1978.237	1963.449	2007.812	1963.449	1	-1	-0,75%	35,65%
	1978.237	1948.661	1948.661	1993.024	-1	-1	1,50%	35,65%
30/03/2023	1964.966	1984.303	1998.333	1947.903	1	1	0,98%	35,65%
	1964.966	1981.649	1931.599	1981.649	-1	1	-0,85%	35,65%
04/04/2023	1984.347	2006.846	2001.971	1975.534	1	1	1,13%	35,65%
	1984.347	1993.159	1966.722	1993.159	-1	1	-0,44%	35,65%
10/04/2023	2006.479	2015.038	2085.547	1966.944	1	1	0,43%	35,65%
	2006.479	2015.038	1927.410	2046.013	-1	1	-0,43%	35,65%
13/04/2023	2014.852	2033.053	2033.053	2006.126	1	1	0,90%	35,65%
	2014.852	2024.077	1996.900	2024.077	-1	1	-0,46%	35,65%
18/04/2023	1994.887	2005.775	2005.775	1989.710	1	1	0,55%	35,65%
	1994.887	2000.241	1984.356	2000.241	-1	1	-0,27%	35,65%
21/04/2023	2005.061	1991.317	2032.013	1991.317	1	-1	-0,69%	35,65%
	2005.061	1977.930	1977.930	2018.626	-1	-1	1,35%	35,65%
26/04/2023	1997.921	1989.532	2014.699	1989.532	1	-1	-0,42%	35,65%
	1997.921	1989.353	1981.143	2006.310	-1	-1	0,43%	35,65%
29/04/2023	1989.353	2056.063	2016.452	1975.833	1	1	3,35%	35,65%
	1989.353	2002.913	1962.294	2002.913	-1	1	-0,68%	35,65%
04/05/2023	2054.901	2031.703	2101.299	2031.703	1	-1	-1,13%	35,65%
	2054.901	2008.186	2008.186	2077.782	-1	-1	2,27%	35,65%
09/05/2023	2021.888	2015.512	2051.012	2007.068	1	-1	-0,32%	35,65%
	2021.888	2036.536	1992.593	2036.536	-1	1	-0,72%	35,65%
12/05/2023	2015.685	1997.246	2052.218	1997.246	1	-1	-0,91%	35,65%
	2015.685	1990.008	1979.151	2033.951	-1	-1	1,27%	35,65%
17/05/2023	2009.653	1990.008	2048.944	1990.008	1	-1	-0,98%	35,65%
	2009.653	1970.362	1970.362	2029.299	-1	-1	1,96%	35,65%
22/05/2023	1981.219	1962.091	2019.476	1962.091	1	-1	-0,97%	35,65%
	1981.219	1957.610	1942.962	2000.348	-1	-1	1,19%	35,65%
25/05/2023	1957.783	1940.205	1992.937	1940.205	1	-1	-0,90%	35,65%
	1957.783	1942.273	1922.628	1975.360	-1	-1	0,79%	35,65%
30/05/2023	1942.273	1982.425	1982.425	1922.455	1	1	2,07%	35,65%
	1942.273	1962.263	1902.293	1962.263	-1	1	-1,03%	35,65%
02/06/2023	1977.600	1960.546	2012.620	1960.546	1	-1	-0,86%	35,65%
	1977.600	1942.709	1942.709	1995.005	-1	-1	1,76%	35,65%
07/06/2023	1963.528	1956.794	1976.996	1956.794	1	-1	-0,34%	35,65%
	1963.528	1950.060	1950.060	1970.262	-1	-1	0,69%	35,65%
12/06/2023	1959.637	1951.556	1975.799	1951.556	1	-1	-0,41%	35,65%
	1959.637	1943.475	1943.475	1967.718	-1	-1	0,82%	35,65%
15/06/2023	1942.428	1934.347	1959.039	1934.347	1	-1	-0,42%	35,65%
	1942.428	1951.406	1925.966	1951.406	-1	1	-0,46%	35,65%
20/06/2023	1952.155	1947.665	1961.134	1947.665	1	-1	-0,23%	35,65%
	1952.155	1913.994	1943.176	1956.644	-1	-1	1,95%	35,65%
23/06/2023	1914.294	1914.755	1940.931	1900.975	1	1	0,02%	35,65%
	1914.294	1927.574	1887.598	1927.574	-1	1	-0,69%	35,65%
28/06/2023	1914.755	1919.535	1962.553	1890.857	1	1	0,25%	35,65%
	1914.755	1919.535	1871.738	1943.434	-1	1	-0,25%	35,65%
03/07/2023	1919.535	1915.624	1946.693	1906.282	1	-1	-0,20%	35,65%
	1919.535	1915.624	1892.595	1933.005	-1	-1	0,20%	35,65%
06/07/2023	1915.190	1925.401	1942.565	1901.502	1	1	0,53%	35,65%
	1915.190	1929.094	1901.502	1929.094	-1	1	-0,73%	35,65%
11/07/2023	1925.184	1960.163	1942.782	1916.711	1	1	1,82%	35,65%
	1925.184	1933.874	1907.803	1933.874	-1	1	-0,45%	35,65%
14/07/2023	1960.163	1978.413	1988.841	1945.824	1	1	0,93%	35,65%

Fig 2: Natural Simple Dataset XAUUSD

start	stop	TP	SL	predictions	result
1827.066	1852.521	1853.430	1813.429	1	1
1827.066	1840.399	1813.429	1853.430	0	1
1852.521	1829.187	1898.279	1829.187	1	0
1852.521	1814.338	1806.459	1875.551	0	0
1814.338	1910.704	1842.218	1799.490	1	1
1814.338	1828.581	1785.853	1828.581	0	1
1920.418	1941.478	2013.906	1873.160	1	1
1920.418	1967.162	1826.416	1967.162	0	1
1941.478	1977.194	2050.383	1886.829	1	1
1941.478	1995.865	1832.312	1995.865	0	1
1978.237	1963.449	2007.812	1963.449	1	0
1978.237	1948.661	1948.661	1993.024	0	0
1964.966	1984.303	1998.333	1947.903	1	1
1964.966	1981.649	1931.599	1981.649	0	1
1984.347	2006.846	2001.971	1975.534	1	1
1984.347	1993.159	1966.722	1993.159	0	1
2006.479	2015.038	2085.547	1966.944	1	1
2006.479	2015.038	1927.410	2046.013	0	1
2014.852	2033.053	2033.053	2006.126	1	1
2014.852	2024.077	1996.900	2024.077	0	1
1994.887	2005.775	2005.775	1989.710	1	1
1994.887	2000.241	1984.356	2000.241	0	1
2005.061	1991.317	2032.013	1991.317	1	0
2005.061	1977.930	1977.930	2018.626	0	0
1997.921	1989.532	2014.699	1989.532	1	0
1997.921	1989.353	1981.143	2006.310	0	0
1989.353	2056.063	2016.452	1975.833	1	1
1989.353	2002.913	1962.294	2002.913	0	1
2054.901	2031.703	2101.299	2031.703	1	0
2054.901	2008.186	2008.186	2077.782	0	0
2021.888	2015.512	2051.012	2007.068	1	0
2021.888	2036.536	1992.593	2036.536	0	1
2015.685	1997.246	2052.218	1997.246	1	0
2015.685	1990.008	1979.151	2033.951	0	0
2009.653	1990.008	2048.944	1990.008	1	0
2009.653	1970.362	1970.362	2029.299	0	0
1981.219	1962.091	2019.476	1962.091	1	0
1981.219	1957.610	1942.962	2000.348	0	0
1957.783	1940.205	1992.937	1940.205	1	0
1957.783	1942.273	1922.628	1975.360	0	0
1942.273	1982.425	1982.425	1922.455	1	1
1942.273	1962.263	1902.293	1962.263	0	1
1977.600	1960.546	2012.620	1960.546	1	0
1977.600	1942.709	1942.709	1995.005	0	0
1963.528	1956.794	1976.996	1956.794	1	0
1963.528	1950.060	1950.060	1970.262	0	0
1959.637	1951.556	1975.799	1951.556	1	0
1959.637	1943.475	1943.475	1967.718	0	0
1942.428	1934.347	1959.039	1934.347	1	0

Fig 3: Modification Simple Dataset XAUUSD for LR & DT

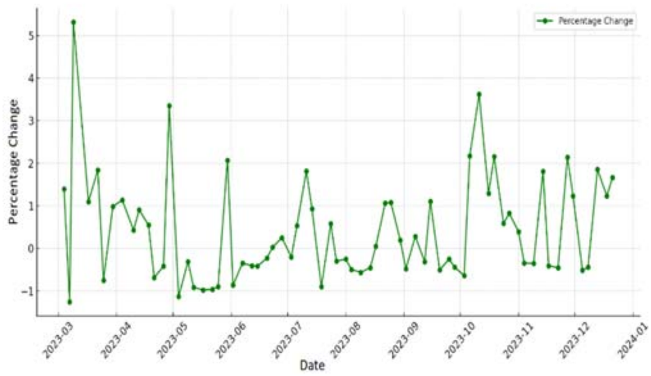


Fig 3: Price Movement Chart

Fig 4: Growth Percentage Chart

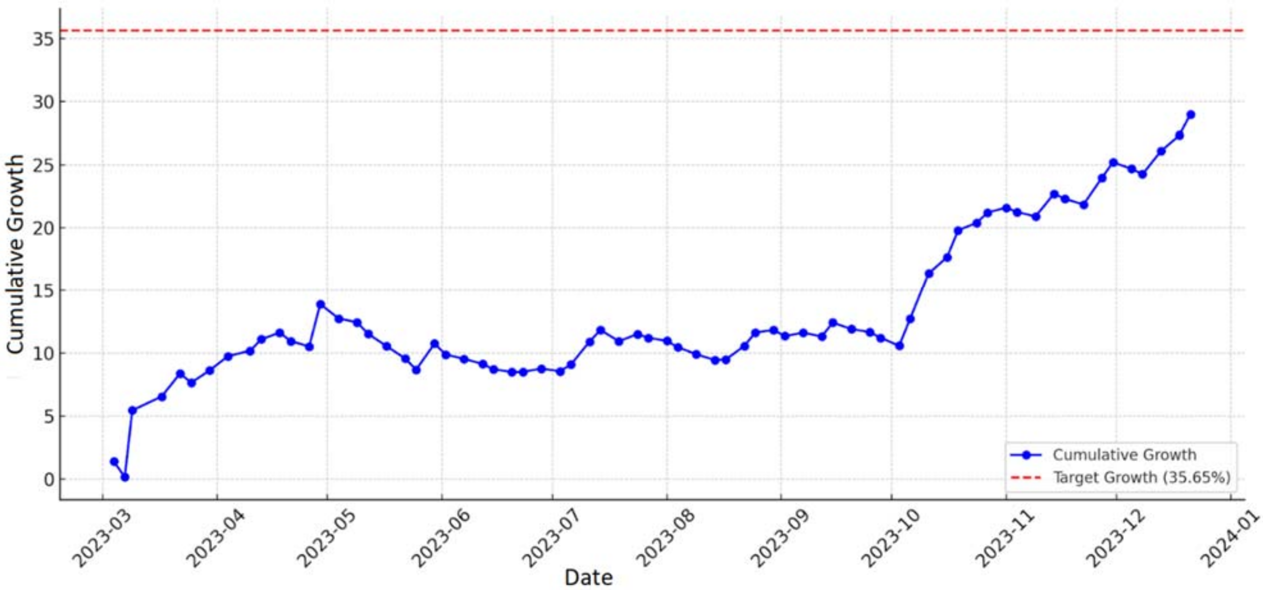
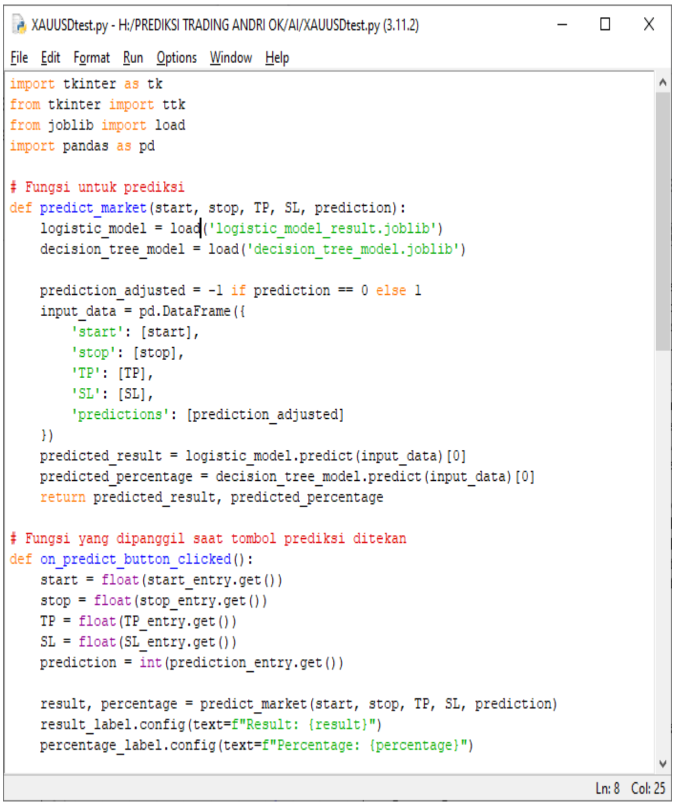


Fig 5: Cumulative Growth Chart

In this study, we have crafted a Python-based predictive program that incorporates two distinct models: Logistic Regression and Decision Tree, tailored for analyzing a dataset from the XAUUSD market. This dataset is meticulously curated through a detailed backtesting process using TradingView data, focusing on crucial indicators like Daily and Weekly Moving Averages and breakout patterns. The program is designed to intake specific market parameters – 'start', 'stop', Stop Loss (SL), Take Profit (TP), and a binary 'prediction' indicator – ensuring alignment with the dataset's structure for effective processing. The program features the Logistic Regression model, implemented using Python's scikit-learn library for its effectiveness in binary classification tasks, and the Decision Tree model, known for its interpretability and capability in handling non-linear relationships. The output generated includes a binary 'Result' from the Logistic Regression model, indicating the predicted market direction, and a 'Percentage' from the Decision Tree model, representing the prediction's confidence level. A critical addition to this program is a user-friendly form for entering new data for real-time predictive analysis. This functionality allows users to feed fresh market data into the models, enabling on-the-spot predictions. Moreover, the program plays a pivotal role in testing the validity and accuracy of the data. Each prediction made by the program is meticulously recorded and stored. This stored data facilitates a crucial comparative analysis to evaluate the growth in prediction accuracy. The study aims to compare the percentage growth in prediction accuracy between the initial manual dataset formation and the predictions made using the AI-powered Logistic Regression and Decision Tree models. This comparison is vital to assess the added value and efficacy of AI models in enhancing the predictive analysis of the XAUUSD market. Through this methodological approach, the program not only serves as a tool for current market analysis but also establishes a framework for assessing the evolution and improvement in market prediction techniques. By comparing the manually formed dataset predictions with those generated by advanced AI models, the study seeks to highlight the progression and potential of AI integration in financial market analysis. This comprehensive methodology underscores the importance of AI in elevating the accuracy and reliability of predictive models in Forex trading. For a visual representation, see the figure below:



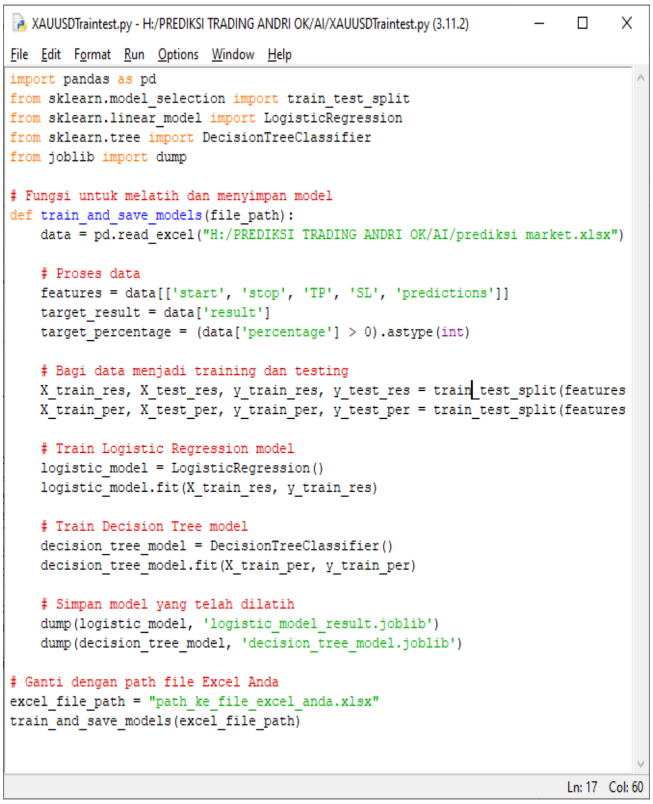
XAUUSDtest.py - H:/PREDIKSI TRADING ANDRI OK/AI/XAUUSDtest.py (3.11.2)

File Edit Format Run Options Window Help

```
import tkinter as tk
from tkinter import ttk
from joblib import load
import pandas as pd

# Fungsi untuk prediksi
def predict_market(start, stop, TP, SL, prediction):
    logistic_model = load('logistic_model_result.joblib')
    decision_tree_model = load('decision_tree_model.joblib')

    prediction_adjusted = -1 if prediction == 0 else 1
    input_data = pd.DataFrame({
        'start': [start],
        'stop': [stop],
        'TP': [TP],
        'SL': [SL],
        'predictions': [prediction_adjusted]
    })
    predicted_result = logistic_model.predict(input_data)[0]
    predicted_percentage = decision_tree_model.predict(input_data)[0]
    return predicted_result, predicted_percentage

# Fungsi yang dipanggil saat tombol prediksi ditekan
def on_predict_button_clicked():
    start = float(start_entry.get())
    stop = float(stop_entry.get())
    TP = float(TP_entry.get())
    SL = float(SL_entry.get())
    prediction = int(prediction_entry.get())

    result, percentage = predict_market(start, stop, TP, SL, prediction)
    result_label.config(text=f"Result: {result}")
    percentage_label.config(text=f"Percentage: {percentage}")
```

Ln: 8 Col: 25

Fig 3: building models



XAUUSDTraintest.py - H:/PREDIKSI TRADING ANDRI OK/AI/XAUUSDTraintest.py (3.11.2)

File Edit Format Run Options Window Help

```
import pandas as pd
from sklearn.model_selection import train_test_split
from sklearn.linear_model import LogisticRegression
from sklearn.tree import DecisionTreeClassifier
from joblib import dump

# Fungsi untuk melatih dan menyimpan model
def train_and_save_models(file_path):
    data = pd.read_excel("H:/PREDIKSI TRADING ANDRI OK/AI/prediksi market.xlsx")

    # Proses data
    features = data[['start', 'stop', 'TP', 'SL', 'predictions']]
    target_result = data['result']
    target_percentage = (data['percentage'] > 0).astype(int)

    # Bagi data menjadi training dan testing
    X_train_res, X_test_res, y_train_res, y_test_res = train_test_split(features
    X_train_per, X_test_per, y_train_per, y_test_per = train_test_split(features

    # Train Logistic Regression model
    logistic_model = LogisticRegression()
    logistic_model.fit(X_train_res, y_train_res)

    # Train Decision Tree model
    decision_tree_model = DecisionTreeClassifier()
    decision_tree_model.fit(X_train_per, y_train_per)

    # Simpan model yang telah dilatih
    dump(logistic_model, 'logistic_model_result.joblib')
    dump(decision_tree_model, 'decision_tree_model.joblib')

# Ganti dengan path file Excel Anda
excel_file_path = "path_ke_file_excel_anda.xlsx"
train_and_save_models(excel_file_path)
```

Ln: 17 Col: 60

Fig 3: Prediction form

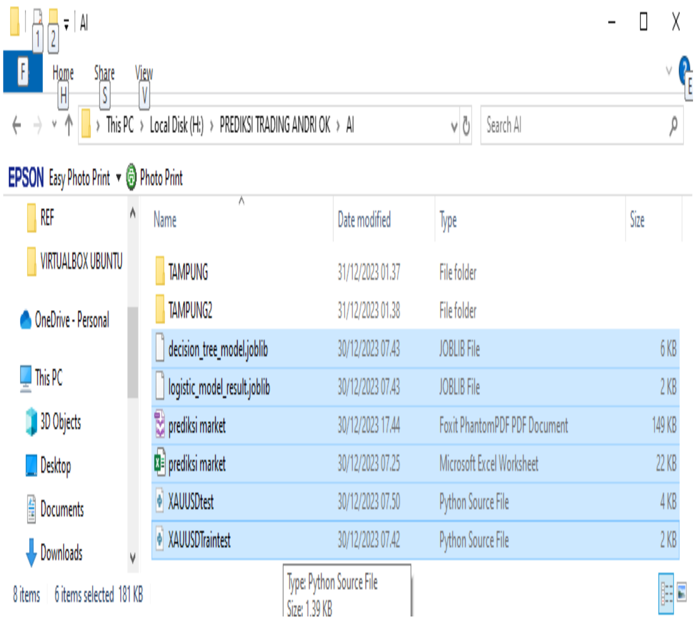


Fig 3: Models storage location

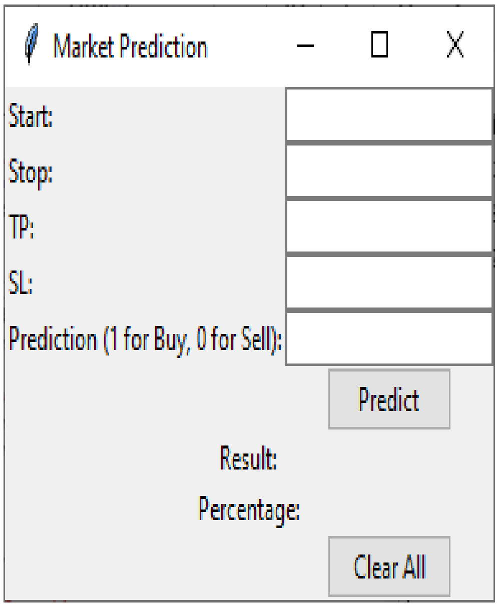


Fig 3: Prediction form format

IV. Result and Discussion

by presenting a thorough and detailed examination of the results obtained from our investigative modeling efforts. this research has led to the exploration and comparison of two different predictive modeling paradigms. On the one hand, we have what we call the "buy and sell of Model" predictive approach, a method that represents an early attempt at Predictive Analytics in this domain. On the other hand, a more advanced approach combines Logistic Regression methodologies and robust decision tree models. not only to present an enumeration of the findings from these two approaches, but also to offer an in-depth analysis by juxtaposing the performance of each. Through a lens that focuses on a variety of key performance indicators, including accuracy, precision, recall, F1 score, and the total number of correct and incorrect predictions, we seek to uncover the strengths and limitations of each model. Our discussion is designed to go beyond surface-level metrics, seeking to provide an understanding of what these results mean in the broader context of predictive modeling. We believe that this analysis will not only validate the effectiveness of the chosen methodology but also shed light on the way forward for future research in this exciting and growing field.

TABLE I. TESTING RESULTS

Metric	**Prediction of Model**	
	Buy & Sell	***Logistic Regression & Decision Tree***
Accuracy	45.00%	92.65%
Precision	51.43%	92.31%
Recall	45.57%	94.74%
F1-Score	48.32%	93.51%

Total Predictions	140	68
Correct Predictions	63	63
Incorrect Predictions	77	5

The comparative analysis of the two models, namely the "Observation of Model" and the "Logistic Regression & Decision Tree" model, reveals compelling insights into their respective performances across various metrics. The most notable difference lies in accuracy, where the combined model demonstrates exceptional performance at 92.65%, more than double the accuracy of the Observation of Model, which stands at 45.00%. This stark contrast underscores the enhanced predictive prowess of the Logistic Regression + Decision Tree model, indicating a substantial improvement in its ability to correctly interpret and forecast outcomes. Moving beyond accuracy, precision, and recall metrics further illuminate the superiority of the combined model. It achieves impressively high scores in both precision and recall, showcasing its effectiveness in identifying positive outcomes accurately while minimizing false positives. These metrics emphasize the refined capability of the Logistic Regression & Decision Tree model to sift through data accurately, ensuring that its predictions are both relevant and reliable.

Examining the F1-Score, a harmonized mean of precision and recall, reinforces the superiority of the combined model, boasting a score of 93.51% compared to 48.32% for the Observation of Model. This superior F1-Score reflects the balanced and nuanced approach of the combined model in handling the trade-off between precision and recall, ensuring a more rounded and reliable predictive performance. Delving into prediction analysis adds an intriguing dimension to the study. Despite making significantly fewer predictions (68 compared to 140 by the Observation of Model), the Logistic Regression & Decision Tree model achieves an equal number of correct predictions (63) while drastically reducing the number of incorrect predictions (only 5 compared to 77 by the Observation of Model). This remarkable efficiency underscores the model's ability to make more accurate predictions with less data, highlighting its potential in reducing the cost and time associated with processing large volumes of data in real-world applications. The findings collectively emphasize the robustness, reliability, and efficiency of the Logistic Regression + Decision Tree model in comparison to the Observation of Model.

V. Conclusion

In the examination of this predictive model, it was found that the Logistic Regression & Decision Tree model consistently outperforms the "Buy & Sell" model across various prediction metrics. The superior accuracy, precision, and recall demonstrated by the combined model provide in-depth insights into its predictive capabilities. This superiority reflects the maturity of the underlying concepts and methodologies, yielding superior results in predictive modeling. From a more practical standpoint, these findings signify a significant contribution to the understanding and application of the Logistic Regression & Decision Tree model. The depth of analysis of prediction metrics indicates that this model not only delivers accurate predictions but also has the ability to navigate the complex trade-offs between precision and recall, resulting in an F1-Score that reflects careful balance.

The model's reliability and predictive efficiency can inspire further research into its internal understanding, results interpretation, and practical applications across various fields. These findings offer a new perspective in concrete data analysis and contribute to sustainable development in predictive modeling. Researchers may explore additional variables, apply the model in diverse contexts, optimize its performance, or integrate it with cutting-edge technologies for enhanced accuracy and efficiency.

References

[1] D. R. Meilanova, “Mengenal Perbedaan antara Pasar Saham dan Forex - Finansial Bisnis.com.”

[2] H. P. S. D. Weerathunga and A. T. P. Silva, “DRNN-ARIMA approach to short-term trend forecasting in forex market,” 18th Int. Conf. Adv. ICT Emerg. Reg. ICTer 2018 - Proc., pp. 287–293, 2019, doi: 10.1109/ICTER.8615580.

[3] M. D. Archer, Getting started in currency trading : winning in today’s Forex market. 2010.

[4] B. Lobel, “What is Gold? Understanding Gold as a Trader’s Commodity,” Nov. 07, 2020. https://www.dailyfx.com/education/commodities/what-is-gold.html (accessed May 27, 2021).

[5] H. Jumaidi, “MENGENAL TRADING GOLD/XAU/USD YUK,” 2020. https://astronacci.com/blog/read/mengenal-trading-gold-xau-usd-yuk (accessed May 27, 2021).

[6] K. S. Zarkias, N. Passalis, A. Tsantekidis, and A. Tefas, “Deep Reinforcement Learning for Financial Trading Using Price Trailing,” ICASSP, IEEE Int. Conf. Acoust. Speech Signal Process.- Proc., vol. 2019-May, pp. 3067–3071, 2019, doi: 10.1109/ICASSP.2019.8683161.

[7] P. Oncharoen and P. Vateekul, “Deep Learning for Stock Market Prediction Using Event Embedding and Technical Indicators,” ICAICTA 2018 - 5th Int. Conf. Adv. Informatics Concepts Theory Appl., pp. 19–24, 2018, doi: 10.1109/ICAICTA.2018.8541310.

[8] S. Ahmed, S. U. Hassan, N. R. Aljohani, and R. Nawaz, “FLF-LSTM: A novel prediction system using Forex Loss Function,” Appl. Soft Comput. J., vol. 97, p. 106780, 2020, doi: 10.1016/j.asoc.2020.106780.

[9] B. A. Usha, T. N. Manjunath, and T. Mudunuri, “Commodity and Forex trade automation using Deep Reinforcement Learning,” 1st Int. Conf. Adv. Technol. Intell. Control. Environ. Comput. Commun. Eng. ICATIECE 2019, pp. 27–31, 2019, doi: 10.1109/ICATIECE45860.2019.9063807.

[10] W. Long, Z. Lu, and L. Cui, “Deep learning-based feature engineering for stock price movement prediction,” Knowledge-Based Syst., vol. 164, pp. 163–173, 2019, doi: 10.1016/j.knosys.2018.10.034.

[11] D. Vezeris, T. Kyrgos, and C. Schinas, “Take Profit and Stop Loss Trading Strategies Comparison in Combination with an MACD Trading System,” J. Risk Financ. Manag., vol. 11, no. 3, p. 56, 2018, doi: 10.3390/jrfm11030056.

[12] Chen, M. Y. (2011). Predicting corporate financial distress based on integration of decision tree classification and logistic regression. *Expert Systems with Applications*, *38*(9), 11261–11272. https://doi.org/10.1016/J.ESWA.2011.02.173

[13] Hargreaves, C., & Hao, Y. (2013). *Prediction of Stock Performance Using Analytical Techniques*. https://doi.org/10.4304/jetwi.5.2.136-142

[14] Zhou, F., Zhang, Q., Sornette, D., & Jiang, L. (2019). Cascading logistic regression onto gradient boosted decision trees for forecasting and trading stock indices. *Applied Soft Computing*, *84*, 105747. https://doi.org/10.1016/J.ASOC.2019.105747