Final Project Group Report

GROUP NO. 02

Group Members:

- John Butoto
- Godfrey Enosh
- Deogratias Nyamhanga
- Oprah Malisa
- Adili Said

INTRODUCTION

For companies and business entities, the difference between the assets and the liabilities is known as the net assets or the net worth of the company. The term NAV is applied to the fund valuation and pricing, which is arrived at by dividing the difference between assets and liabilities by the number of shares held by the investors.

Net Asset Value (NAV) is the net value of an investment fund's assets less its liabilities, divided by the number of shares outstanding. Most commonly used in the context of a mutual fund or an exchange-trade fund.

NAV is commonly used as per share value calculated for a mutual fund or exchange-trade fund.

Problem Statement:

Significant numbers of investors have encountered challenges in making well informed decisions regarding Net Asset Value (NAV) prices resulting in financial losses.

Main Objective:

The objective of this project is to develop a machine learning model that can accurately forecast Net Asset Value (NAV) prices for various investment categories over a future time horizon of six months. This forecasting model aims to provide investors and financial analysts with reliable predictions of NAV prices, thereby assisting them in making informed investment decisions across a diverse range of investment categories. The model will leverage historical NAV price data, along with relevant market indicators and economic factors, to generate forecasts that enhance the ability to navigate the dynamic landscape of financial markets

Significance of the Research:

The research conducted on the development of a machine learning model for forecasting Net Asset Value (NAV) prices for the next six months holds substantial significance in the domain of finance and investment. The following key points underscore the importance of this study:

- Empowering Informed Investment Decisions:

  At the core of this research lies the objective to empower investment professionals, fund managers, and individual investors with a tool that can facilitate more informed and data-driven investment decisions. Accurate NAV forecasts serve as a compass, guiding stakeholders in the complex landscape of financial markets. The ability to anticipate future NAV values with precision can have a profound impact on investment strategies, leading to potentially higher returns and minimized losses.

- Mitigating Financial Risk:

  One of the fundamental benefits of accurate NAV predictions is the ability to mitigate financial risk. In a world where market dynamics can change rapidly, the foresight provided by this research enables investors and financial institutions to develop risk management strategies that are both proactive and effective. This can prove invaluable in safeguarding investments during market downturns and economic uncertainties.

- Competitive Edge in the Financial Industry:

  In the fiercely competitive financial industry, staying ahead of the curve is essential. By providing financial institutions with a sophisticated NAV forecasting model, this research equips them with a competitive edge. The ability to make more precise investment decisions and offer innovative financial products can attract and retain clients, positioning organizations as leaders in their respective markets.

- Streamlining Portfolio Diversification:

  Effective portfolio diversification is a hallmark of prudent investment practices. The model developed in this research can optimize portfolio diversification strategies by offering insights into how different asset classes and sectors are likely to perform in the near future. This can result in portfolios that are more resilient to market fluctuations.

- Boosting Investor Confidence:

  Accurate NAV forecasts are instrumental in boosting investor confidence. When investors have access to reliable predictions of their investment performance, they are more likely to make informed decisions and remain committed to their investment strategies. This, in turn, can foster trust in financial products and institutions.

- Time Efficiency and Automation:

  The automation of NAV forecasting through machine learning not only enhances accuracy but also saves valuable time for financial analysts. This time-saving aspect can enable professionals to focus on higher-level tasks, such as strategic planning, client relationships, and exploring innovative investment opportunities.

- Boosting Investor Confidence:

  Accurate NAV forecasts are instrumental in boosting investor confidence. When investors have access to reliable predictions of their investment performance, they are more likely to make informed decisions and remain committed to their investment strategies. This, in turn, can foster trust in financial products and institutions.

- Adaptation to Market Volatility:

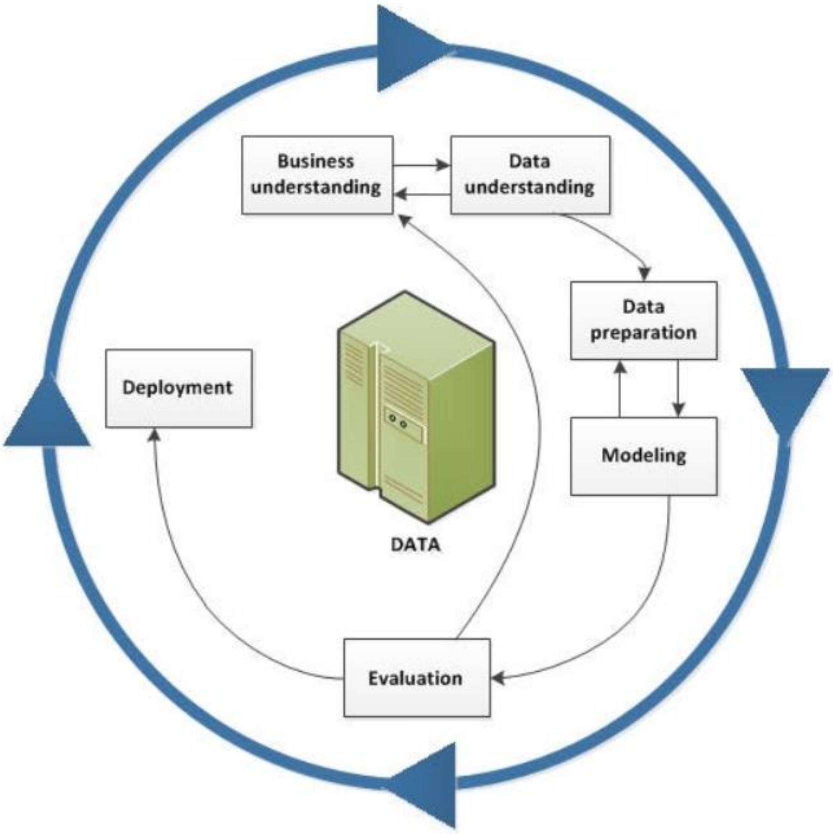
  The unpredictable nature of financial markets demands agility. Your research equips investors with a tool to adapt swiftly to market volatility. By providing timely and accurate NAV predictions, stakeholders can make agile investment decisions, adjusting their strategies to align with changing market dynamics.

- Advancing Financial Research:

  This research contributes to the ever-evolving field of financial machine learning and data-driven investment strategies. The methodologies and insights developed in this study can inspire further research and innovation in finance, driving the industry forward.

METHODOLOGY

-Methodology we used in our project is Cross-Industry Standard Process for Data Mining (CRISP-DM) Methodology



❖ Business Understanding:

In the realm of UTT Asset Management and Investment Services (UTT AMIS), the dynamic nature of asset valuations within each fund or scheme introduces a notable risk factor for potential investors. This risk is inherent in the possibility of incurring financial losses stemming from the inherent uncertainty and the absence of reliable, foresighted predictions regarding the future value trajectories of these funds and schemes. Consequently, investors face the challenge of making well-informed decisions amid the complexity of asset valuation fluctuations, which may ultimately impact the profitability of their investments.

❖ Data Understanding:

Data collection involves acquiring historical NAV data from UTT AMIS MANAGEMENT AND INVESTOR SERVICES. A thorough exploration of the dataset is conducted to uncover underlying patterns and identify potential anomalies. We plot different graphs to present our dataset such to discover if there are any potential anomalies.

❖ Data preparation:

Data preparation begins with rigorous data cleaning procedures to address missing values and outliers, ensuring the dataset's integrity. Our dataset didn’t contain any missing values or duplicated data, so we move to next stage of feature engineering. Feature engineering techniques are applied to create meaningful variables which can aid in improving prediction accuracy.

-

❖ Modeling:

The selection of appropriate machine learning algorithms is a critical step. We consider time series forecasting methods, including xgboost and Facebook Prophet algorithm, to build our predictive models. The dataset is split into training and testing sets to facilitate model evaluation. Extensive model training and fine-tuning are carried out using the training data, with a focus on optimizing predictive accuracy.

❖ Evaluation:

Model performance is rigorously evaluated using the testing dataset. We employ established evaluation metrics such as Mean Absolute Error (MAE) and Root Mean Square Error (RMSE) to assess the accuracy of our forecasts. Adjustments to the models are made as necessary to improve performance, with a commitment to delivering accurate NAV predictions.

❖ Deployment:

Upon achieving satisfactory model performance, the selected model is deployed in a suitable production environment for real-time NAV predictions. Ongoing monitoring and maintenance ensure the model's continued relevance and accuracy in the face of evolving market dynamics.

❖ Documentation:

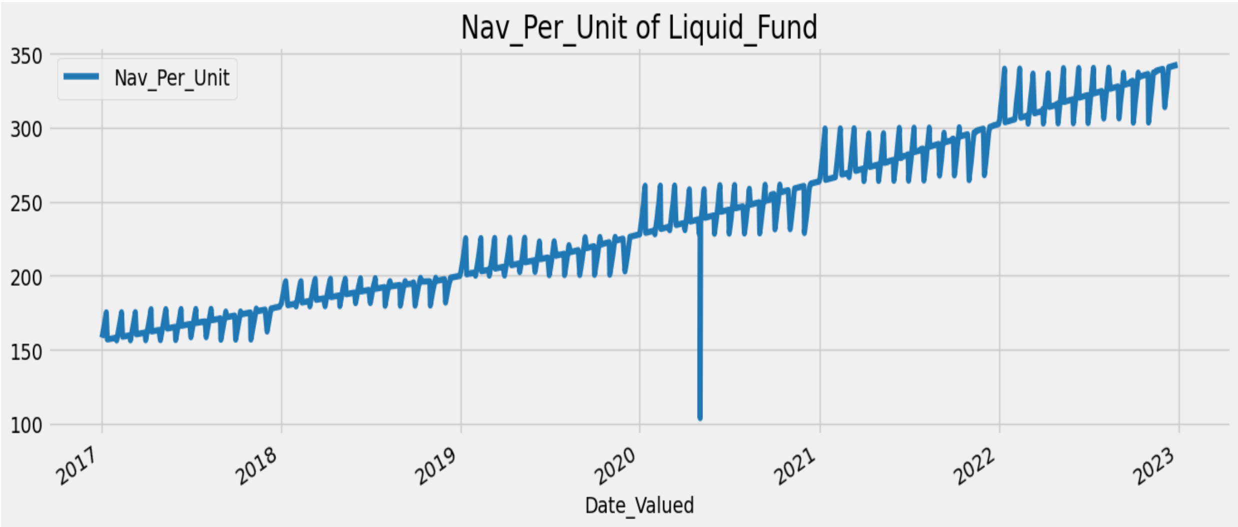
The research process is meticulously documented, encompassing data sources, preprocessing methodologies, and model specifications. Detailed findings, insights, and recommendations are communicated effectively to stakeholders.

Exploratory Data Analysis (EDA)

The Exploratory Data Analysis (EDA) phase of this research constitutes a pivotal stage in the research process, providing critical insights into the nature and characteristics of the data. EDA serves as the foundational step that precedes the development of predictive models and statistical inferences, enhancing the overall understanding of the dataset.

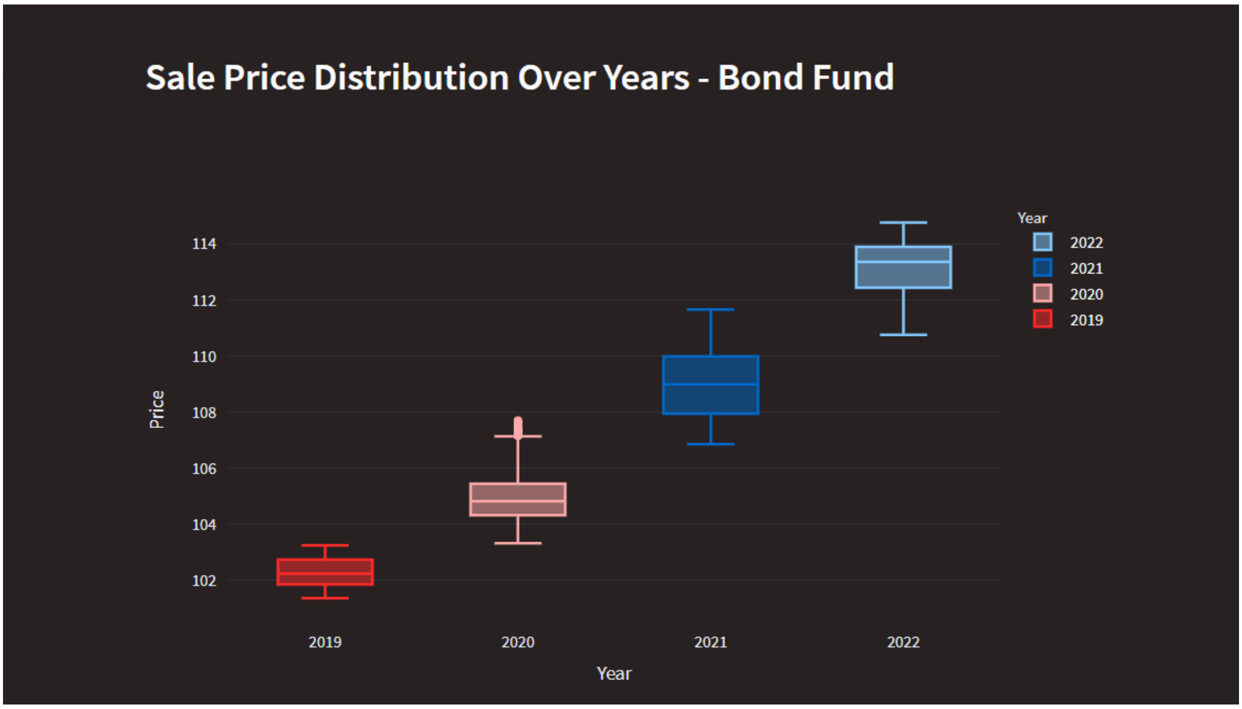
1. Seasonality and Trends:

The project's objective to forecast NAV prices over a future time horizon of six months. During your exploratory data analysis (EDA), if you observe recurring patterns in NAV prices over shorter time intervals (e.g., monthly or quarterly) or long-term trends that impact NAV, this would be an indication of seasonality and trends. Prophet excels at capturing these patterns, making it a suitable choice.



2. Missing Data and Outliers:

In data cleaning no missing values or duplicated data were found. However, during EDA, Outliers were. Outliers can be particularly important in financial data, where extreme market events can lead to unusual NAV price movements. Prophet is relatively robust to outliers, which is an advantage in such cases.



3. Change Points:

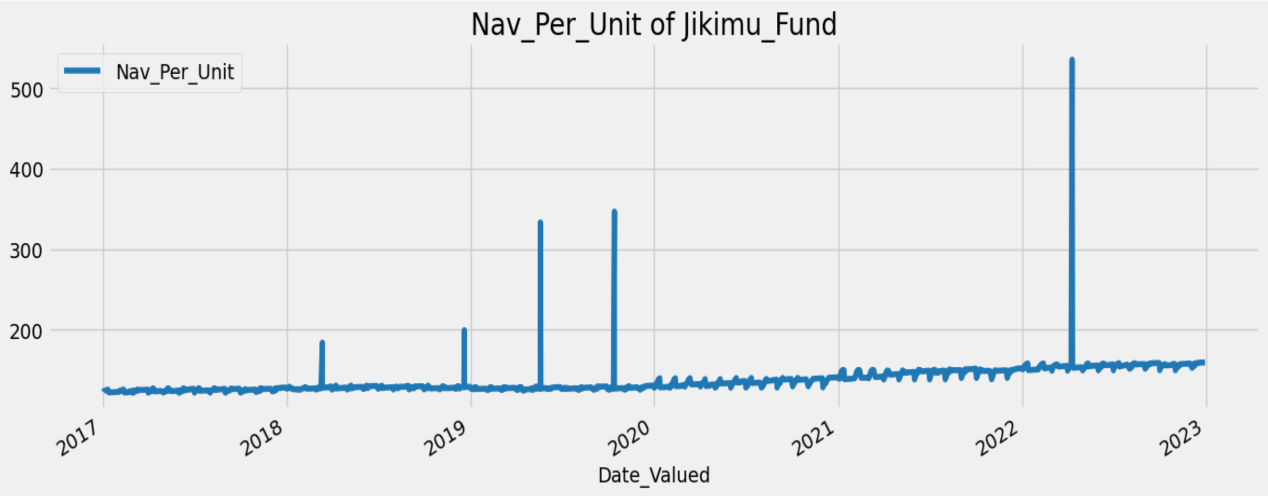
In financial markets, there can be abrupt changes in asset prices due to various factors, such as economic announcements or geopolitical events. EDA uncovers instances of significant changes or shifts in NAV values, Prophet's feature for modeling change points can help capture these shifts.

4. Uncertainty Estimation:

Given the importance of risk management in financial markets, it's crucial to provide uncertainty estimates alongside your NAV forecasts. EDA emphasizes the need for understanding the range of possible future NAV values, Prophet's ability to provide uncertainty intervals can be valuable for investors and analysts.

5. Heteroscedasticity:

Financial data often exhibits changing levels of volatility (heteroscedasticity) over time. EDA shows that the variance in NAV data fluctuates, Prophet's modeling of heteroscedasticity can be advantageous for more accurate predictions.



6. Auto-Regression:

While Prophet primarily focuses on capturing seasonality and trends, it can also handle basic auto-regressive patterns. EDA indicates that past NAV values influence future NAV values, Prophet can accommodate this to some extent.

7. Interactions with External Factors:

Incorporate relevant market indicators and economic factors into your forecasting model. Prophet allows for the inclusion of additional regressors, making it a suitable choice when you need to account for external influences on NAV, which aligns with your project's objectives.

8. Ease of Use:

Team lacks extensive expertise in time series modeling, Prophet's user-friendly nature and its reduced need for extensive parameter tuning can be advantageous. This aligns with project's goal of providing a practical tool for investors and analysts.

MODELLING

Modeling refers to the process of creating mathematical representations of real-world phenomena to make informed predictions or decisions. In the context of Net Asset Value, modeling entails constructing intricate mathematical frameworks that analyze historical data, market conditions, and a range of influential variables to estimate the current or future NAV of an investment portfolio, mutual fund, or financial instrument.

These models are indispensable tools for investors, asset managers, and financial institutions seeking to gauge the performance and value of their assets, allocate resources efficiently, and make well-informed investment decisions. In this intricate and dynamic field, the modeling process involves a fusion of data science, statistical analysis, economic insights, and domain expertise, all geared towards providing a deeper understanding of financial markets and enabling more effective asset management strategies.

1. Data Sources and Preprocessing:

Data used to build the model was from UTT asset management website. https://www.uttamis.co.tz/

Data preprocessing initiatives include cleaning, handling duplicates and missing values and feature engineering creating relevant financial ratios and indicators.

2. Time Series Considerations:

The dataset is a time series data (common in market and finance, thus it involves seasonality, trends and autocorrelation as mentioned in exploratory data analysis.

Seasonality, trends, and autocorrelation are key concepts in time series analysis and forecasting.

Seasonality refers to the repetitive and predictable patterns or fluctuations in a time series data set at regular intervals. These intervals can be daily, weekly, monthly, or any other consistent time frame.

Trends in time series data represent the long-term, non-seasonal changes in the data over time. Trends can be upward (indicating growth or expansion), downward (indicating decline or contraction), or flat (indicating stability).

Autocorrelation, also known as serial correlation, is a statistical concept that measures the degree to which values in a time series are correlated with previous values. In other words, it quantifies the relationship between data points at different time lags.

Seasonality, trends, and autocorrelation are critical components of time series analysis. Recognizing and understanding these patterns and relationships within time series data is essential for building accurate forecasting models and making informed decisions in various fields, including finance, economics, and environmental science.

3. Feature Selection:

The most important features are often used in making the model. These features influence the predictions. The features used in this model are **date valued** and **NAV per unit**.

4. Model Validation:

Validation includes examinations to evaluate the performance of different data-driven models for forecasting future NAV values, with the aim of refining investment strategies and asset management decisions. Initial attempts involved employing the ARIMA and XGBoost models, both well-established for their predictive capabilities. These models proved inadequate in capturing the nuanced dynamics within NAV data.

It was then that Prophet, a specialized forecasting tool by Facebook, was introduced. To our delight, Prophet demonstrated exceptional success in handling this complex task, showing an innate ability to account for seasonality and abrupt data fluctuations.

5. Model Selection and training:

Algorithm choice for predicting this financial data, considering a number of reasons including the fact that it's a time series problem is **Facebook Prophet.**

Training to get the model

6. Evaluation Metrics:

Different metrics are used to evaluate the performance of the model. Metrics include Mean Absolute Error (MAE), Mean Squared Error (MSE), Root Mean Squared Error (RMSE), and R-squared.

These metrics are relevant to the financial domain and they analyze how the predictive model deviates from the actual values.

After performing evaluation on each Fund using Facebook prophet model performance on each fund occur as follows:

TABLE SHOWING THE METRICS VALUES OF THE MODELS FOR EACH FUND

FUND NAME	R2 SCORE	MAE	MSE	RMSE
Jikimu Fund	0.52823333150	1.434582326634	3.70456244523	1.92472399196
Umoja Fund	0.80459065377	5.06292097054	47.881770551	6.91966549416
Liquid Fund	0.69196440613	2.258588325138	9.787381594546	3.12847911844
Watoto Fund	0.77539891279	3.642584483536	26.4107493160	5.13913896640
Wekeza Fund	0.634836242805	6.16973588770	78.2914801496	8.84824729252
Bond Fund	0.583792269211	1.12684007604	1.56238111334	1.24995244443

7. Hyperparameter Tuning:

Hyperparameter tuning is a critical step in improving the accuracy of Net Asset Value (NAV) prediction models. In this process, various hyperparameters, such as daily weekly yearly seasonality and seasonality mode are fine-tuned to enhance the model's performance.

Techniques like grid search CV and manual techniques external regressor explore different hyperparameter combinations to strike the right balance between model complexity and predictive accuracy.

This optimization process ensures that the model can effectively capture the nuances of NAV fluctuations and generalize well to unseen data, ultimately contributing to more reliable NAV predictions in the field of investment management.

8. Deployment:

Deployment involved creating a web-based application that provides a user-friendly interface for showcasing and predicting the NAV/Unit prices.

The application required online hosting for any user to access via link. The application was hosted on the streamlit cloud.

The end result was a web-based application that is capable of predicting the NAV/scheme for the next six months.

9.Findings:

The project gave a valuable insight into the field of data science and its practical applications, practice and sharpen skills used in real data science jobs.

Prophet model performance: It excels at automatically detecting and modeling seasonality patterns, making it suitable for data with complex periodic trends. Uncertainty. The model provides uncertainty estimates for predictions, helping users understand the range of potential outcomes. It is also suitable for beginners

Fund-Specific Trends: We observed that each fund displayed unique trends and seasonality patterns in their NAV per unit values. Some funds exhibited steady growth, while others displayed some fluctuation

10.Conclusion:

While the Prophet model provided valuable predictions, it's essential to consider potential risks and external factors that may impact fund performance, such as economic conditions and market dynamics.

For funds showing stable growth, long-term investment strategies may be suitable. Conversely, funds with higher volatility may need a more active and risk-aware approach.

The data set had limited features, more features would have helped to analyse and understand better the trends and patterns of the funds and gain more insights for better conclusion and prediction