KNIT Finance

Portfolio Performance Analysis

Authored by: Yuvaraj C U | yuvarajcu@gmail.com

Problem Statement:

The objective of this analysis is to evaluate the performance of multiple Binance accounts over a 90-day period using historical trade data and rank them based on multiple performance metrics, including Return on Investment (ROI), Profit and Loss (PnL), Sharpe Ratio, Maximum Drawdown (MDD), Win Rate, and Win Positions. These metrics provide insight into profitability, risk management, and consistency of the trading strategies associated with each portfolio. The final ranking of each portfolio is derived through a weighted composite score based on profitability, risk-adjusted performance, and trade success metrics.

Objective:

1. Calculate relevant financial metrics for each account.
2. Rank accounts based on these metrics.
3. Provide insights and a list of the top 20 performing accounts.

Dataset Overview:

The dataset contains the following:

- **Port_IDs**: Unique identifiers for each account.
- **Trade_History**: Details of trades, including timestamp, asset, trade side (BUY/SELL), price, quantity and few other relevant details

2. Methodology

The analysis was structured into three key steps: Data Handling and Cleaning, Feature Engineering, and Ranking Algorithm Development.

Step 1: Data Exploration and Cleaning

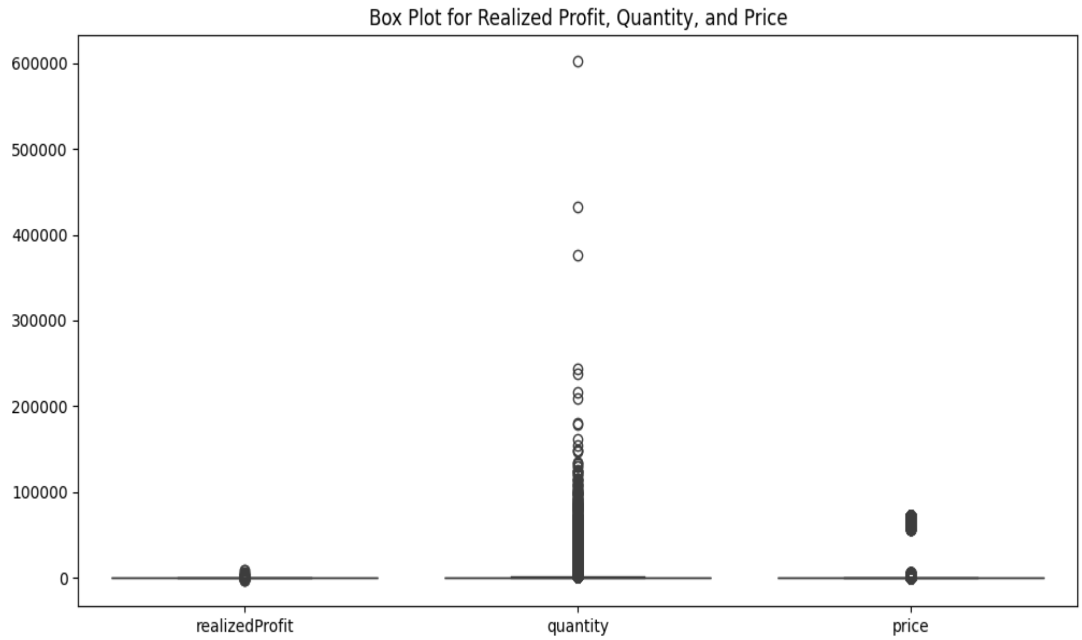
To begin, the initial dataset was loaded, cleaned, and transformed:

1. **Data Loading and Cleaning**:
    a. The dataset was imported, and initial inspections included checks for missing values and duplicated entries.
    b. Missing values were handled by removing rows with missing data, ensuring a cleaner dataset for accurate metric calculations.
2. **Flattening the Trade History**:
    a. Each portfolio contains a 'Trade_History' column that stores trade records in a nested format. We utilized the '*ast.literal_eval()*' function to safely parse this column, expanding each trade into separate records for detailed per-trade analysis.
    b. This created a new DataFrame, '*df_parsed*', where each row represents an individual trade, tagged with its corresponding 'Port_ID' for portfolio identification.
3. **Data Validation**:
    a. Further validation included checking for column types, data completeness, and missing values in the expanded dataset *(df_parsed)*, laying the foundation for reliable feature engineering.

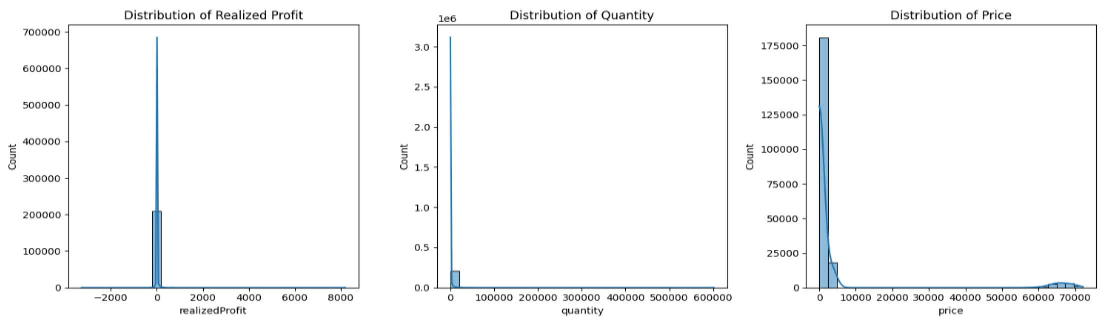
Step 2: Exploratory Data Analysis (EDA)

A detailed EDA was conducted to understand the dataset distribution and identify any potential anomalies:

1. **Data Types and Summary Statistics**:
    a. Basic statistics (mean, standard deviation, min, max) for columns like realizedProfit, quantity, and price were inspected.
    b. Visualization methods, such as box plots and histograms, revealed the distribution patterns and outliers for numerical columns.
2. **Visualizing Key Variables**:
    a. **Box Plots**: Used to visualize potential outliers in realizedProfit, quantity, and price.



b. **Histograms**: Generated to display the distribution of key metrics (e.g., realizedProfit, quantity, price), helping in understanding the spread and skewness of data values.



c. **Datetime Conversion**: The time field was converted from Unix milliseconds to a datetime format, aiding in time-series calculations like Sharpe Ratio and Maximum Drawdown.

Step 3: Feature Engineering

In this step, we derived key metrics for each portfolio. Each metric provides insight into different aspects of portfolio performance.

1. **Return on Investment (ROI)**:
    a. **Definition**: ROI is the return generated relative to the investment, reflecting portfolio profitability.
    b. **Calculation**: ROI for each trade was calculated as:

Result(ROI) = $\frac{Realized\ Profit}{Quantity \cdot price}$

c. **Averaging ROI**: By averaging ROI for each portfolio, we assess profitability over the portfolio's entire trade history, providing a standardized measure of return on a per-trade basis.

2. **Profit and Loss (PnL)**:
   a. **Definition**: The sum of realized profits across all trades in a portfolio.
   b. **Significance**: While ROI standardizes return relative to investment, PnL provides a raw, cumulative measure of profit, giving a straightforward sense of a portfolio's net outcome.

3. **Sharpe Ratio**:
   a. **Definition**: The Sharpe Ratio is a risk-adjusted return metric, indicating how much excess return a portfolio generates per unit of risk (volatility).
   b. **Calculation**: Sharpe Ratio is calculated as:
   Sharpe Ratio = $\frac{MeanDailyReturn}{StandardDeviationofDailyReturn}\sqrt{2}52$
   c. **Interpretation**: Higher Sharpe Ratios signify better risk-adjusted performance, making this an essential measure for evaluating portfolio stability.

4. **Maximum Drawdown (MDD)**:
   a. **Definition**: MDD measures the largest peak-to-trough drop in the cumulative portfolio value, reflecting maximum potential loss.
   b. **Significance**: MDD provides insight into the worst-case scenario in terms of risk, helping assess portfolio resilience.
   c. **Implementation**: We used cumulative profits for each portfolio to compute MDD, identifying peak-to-trough drops within the trade history.

5. **Win Rate, Win Positions and Total Positions**:
   a. **Definition**: Win Rate is the percentage of profitable trades, while the total positions and the winning positions are the number of trades and winning trades respectively for each account.
   b. **Significance**: These metrics are valuable for understanding trading consistency and frequency of profitable trades.

The next step was to Create a data frame for the Metrics scores with respective Portfolio ID, which we shall use it to rank those accounts based on their performances.

Step 4: Ranking Algorithm

To determine the final ranking of each portfolio, we created a composite score that combines metrics from three key areas:

1. **Metric Categories**:
    a. **Profitability**: Includes ROI and PnL.
    b. **Risk-Adjusted Performance**: Comprises the Sharpe Ratio and normalized MDD.
    c. **Trade Success**: Consists of Win Rate and Win Positions.
2. **Feature Scaling and Normalization**:
    a. Metrics were normalized using the MinMaxScaler for comparability across different ranges.
    b. **Inversion of MDD**: Since MDD represents risk (where lower values are preferable), it was inverted before normalization.
3. **Feature Weights via Random Forest**:
    a. We employ a RandomForestRegressor model to derive feature importance within each category. This allowed for assigning weights based on the feature's contribution to the respective target metric (ROI for profitability, Sharpe Ratio for Risk-Adjusted performance and Win Rate for Trade Success).
    b. These weights were then used to calculate weighted scores within each category.
4. **Calculating Category Scores**:
    a. **Profitability Score**: Sum of weighted ROI and PnL values.
    b. **Risk-Adjusted Score**: Sum of weighted Sharpe Ratio and MDD values.
    c. **Trade Success Score**: Sum of weighted Win Rate and Win Positions.
5. **Composite Score Calculation**:
    a. The final score for each portfolio was computed as a weighted average of the three category scores:
    Overall Composite Score = $(0.4\times\text{Profitability Score}) + (0.3\times\text{Risk-Adjusted Score}) + (0.3\times\text{Trade Success Score})$
    b. We assign weights of 0.4, 0.3, and 0.3 to emphasize profitability while balancing risk and consistency.
6. **Final Ranking**:
    a. Portfolios were ranked based on their composite score, with higher scores indicating better performance across all metrics.

b. Csv files were created containing the metrics scores and the rankings of the top 20 calculated

Conclusion

The analysis presented an in-depth performance evaluation for each portfolio, covering profitability, risk, and trade success. By integrating various metrics, we achieved a balanced ranking that captures both the profitability and stability of each portfolio. Key insights include:

- **High ROI Portfolios**: Portfolios with a consistently high ROI but moderate Sharpe Ratios may be profitable but riskier.
- **High Sharpe Ratio Portfolios**: Portfolios with high Sharpe Ratios and low MDD are risk-averse, yielding stable, albeit possibly lower, returns.
- **Balanced Portfolios**: Portfolios ranking well across profitability, risk-adjusted performance, and trade success provide an ideal blend of return and stability, offering robust investment choices.