Report: Analysis of Binance Trade Data

1.Introduction

This report details the analysis of historical trade data from various Binance accounts over a 90-day period. The goal was to calculate key financial metrics for each account, rank them based on performance, and identify the top 20 accounts. The analysis provides insights into profitability, risk management, and trade consistency across accounts.

2.Methodology

2.1 Data Preparation

- Data Loading: The dataset was loaded into a pandas DataFrame. It contained two columns: Port_IDs (unique account identifiers) and Trade_History (a list of trades with details such as timestamp, symbol, side, price, fee, quantity, and realized profit).
- Data Cleaning:
    - Missing values in the Trade_History column were dropped.
    - The Trade_History column, initially stored as a string, was converted into a list of dictionaries using ast.literal_eval.
- Data Transformation:
    - The Trade_History column was exploded to create individual rows for each trade.
    - The nested dictionaries in Trade_History were normalized into separate columns using pd.json_normalize.

2.2 Feature Engineering

The following metrics were calculated for each account (Port_IDs):

1. ROI (Return on Investment):
    - Formula: ROI = (Net Profit / Total Investment) * 100
    - Net Profit: Sum of realizedProfit minus sum of fee.
    - Total Investment: Sum of quantity for all BUY trades.
2. PnL (Profit and Loss):
    - Formula: PnL = Sum of realizedProfit - Sum of fee.
3. Sharpe Ratio:
    - Formula: Sharpe Ratio = Mean of realizedProfit / Standard Deviation of realizedProfit.
    - Handled cases where standard deviation was zero by setting the Sharpe Ratio to 0.

4. MDD (Maximum Drawdown):
    - Formula: MDD = (Peak of Cumulative PnL - Trough of Cumulative PnL) / Peak of Cumulative PnL.
    - Handled cases where the peak was zero by setting MDD to 0.
5. Win Rate:
    - Formula: Win Rate = (Number of Winning Trades / Total Number of Trades) * 100.
6. Win Positions:
    - Number of trades with positive realizedProfit.
7. Total Positions:
    - Total number of trades.

2.3 Ranking Algorithm

- A composite score was calculated for each account using weighted metrics:
    - ROI: 30 percent
    - PnL: 25 percent
    - Sharpe Ratio: 20 percent
    - MDD: 15 percent (lower MDD is better, so 1 - MDD was used)
    - Win Rate: 10 percent
- Accounts were ranked based on the composite score in descending order.

3. Findings

3.1 Top 20 Accounts
The top 20 accounts, ranked by their composite score, are as follows:

Rank | Port_IDs | ROI | PnL | Sharpe Ratio | MDD | Win Rate | Win Positions | Total Positions | Composite Score
1 | 4020204877254599680 | 0.733 | 83347.15 | 0.0607 | 1.1066 | 32.94 | 1993 | 6050 | 20840.30
2 | 3999240873283311617 | 1.141 | 44446.92 | 0.2275 | 0.9970 | 52.32 | 2366 | 4522 | 11117.35
3 | 4021669203289716224 | 0.859 | 27963.77 | 0.0683 | 0.9985 | 31.39 | 528 | 1682 | 6994.35
...

3.2 Key Observations

1. High PnL Accounts:
    - The top-ranked account achieved the highest PnL due to a large number of trades. However, its ROI and Sharpe Ratio were relatively low, indicating lower profitability per trade and poor risk-adjusted returns.

2. Balanced Performance:
    - The second-ranked account demonstrated a balance between ROI, PnL, and Sharpe Ratio. Its Win Rate was significantly higher, indicating more consistent profitability.
3. High ROI Accounts:
    - Some accounts achieved high ROI but had fewer trades, suggesting concentrated, high-performing trades.
4. Risk Management:
    - Accounts with lower MDD values managed drawdowns effectively, while others experienced significant losses.
5. Trade Consistency:
    - Accounts with higher Win Rates demonstrated consistent profitability, while those with lower Win Rates struggled to maintain profitability.
6. Assumptions and Limitations
7. Risk-Free Rate:
    - The Sharpe Ratio calculation assumed a risk-free rate of 0.
8. Trade Classification:
    - Trades were classified as BUY or SELL based on the side column. More complex trade types were not considered.
9. Data Quality:
    - Missing or inconsistent data in the Trade_History column was dropped.
10. Time Period:
    - The analysis was limited to a 90-day period.
11. Conclusion
    The analysis successfully identified the top 20 accounts based on their financial performance. Key metrics such as ROI, PnL, Sharpe Ratio, MDD, and Win Rate were used to evaluate and rank the accounts.

Github link: https://github.com/RadhaVaishnavi/Primetrade.ai