Data Science Internship Assessment Report

1. Introduction

This project is part of the Web3 Trading Team internship assessment. The objective is to analyze the relationship between trader behavior and market sentiment (Fear vs Greed). The datasets provided include a Bitcoin Market Sentiment dataset and a Historical Trader Data dataset from Hyperliquid. The goal is to uncover hidden patterns and signals that align or diverge from market sentiment, which can ultimately inform smarter trading strategies.

2. Data Overview

Two key datasets were used:
1. Bitcoin Market Sentiment Dataset: Contains daily classifications of 'Fear' or 'Greed'.
2. Hyperliquid Trader Data: Contains trading information such as account, symbol, execution price, size, side, leverage, closedPnL, and time.

The datasets were cleaned, standardized, and merged on date fields to align trades with daily sentiment.

3. Methodology

The workflow followed these steps:
- Data Cleaning: Standardized column names, parsed timestamps, and converted numeric fields.
- Data Integration: Aligned trades with corresponding market sentiment based on daily timestamps.
- Feature Engineering: Derived metrics such as signed trade size, notional value, and PnL percentage.
- Exploratory Data Analysis (EDA): Examined distributions of profitability, leverage, and trade volume across sentiments.
- Statistical Analysis: Conducted Mann-Whitney U tests and Cohen’s d effect size to quantify differences between Fear and Greed.
- Visualization: Generated comparative plots for PnL, leverage, and rolling averages to support insights.

4. Machine Learning Model

A simple predictive model was developed to classify sentiment impact based on trade features. For demonstration purposes, logistic regression was applied using features such as leverage, notional value, and trade size. The model aimed to distinguish between Fear and Greed sentiment days, showing how trader behaviors correlate with sentiment.

Key steps:
- Feature Selection: ClosedPnL, leverage, size, pnl_pct.
- Train-Test Split: 80-20 ratio.
- Logistic Regression Model: Provided baseline predictive capability.
- Evaluation Metrics: Accuracy and confusion matrix were used. The model achieved moderate accuracy, indicating that sentiment alone cannot fully explain profitability but still influences trader behavior trends.

5. Key Insights

- Traders tend to use higher leverage during Greed phases, often leading to higher volatility in profits and losses.
- Profitability was not consistently higher during Greed; some over-leveraged trades caused increased losses.
- Trade volumes were larger during Greed phases, suggesting more aggressive market participation.
- Fear phases showed relatively lower leverage usage and smaller trade sizes, indicating cautious behavior.
- Rolling average PnL trends demonstrated that high-leverage Greed days often led to weaker next-day performance.

6. Recommendations

- Implement leverage caps or stricter risk management during Greed phases.
- Monitor high-leverage Greed days as potential signals for upcoming volatility.
- Develop trading strategies that hedge or scale exposure during extreme sentiment days.
- Combine sentiment data with technical indicators for stronger predictive models.

7. Conclusion

This analysis highlights the significant interplay between trader behavior and market sentiment. While Greed phases encourage higher risk-taking, they do not always lead to better profitability. Fear phases promote caution but can help preserve capital. These findings can support the creation of robust trading strategies that account for psychological and behavioral market drivers.