Market Sentiment vs Trader Performance

Project Title:
Analyzing the Relationship Between Bitcoin Market Sentiment and Trader Performance on Hyperliquid

Prepared by: Yati Rathod
Tools Used: Python, Pandas, Matplotlib, Seaborn, Scikit-learn
Environment: Google Colab

Project Summary

To explore the relationship between market sentiment (Fear vs. Greed) and trader performance using features derived from the Hyperliquid trader dataset, and evaluate how well trader statistics can predict market sentiment using machine learning.

Dataset Description

➤ **Market Sentiment Dataset (fear_greed_index.csv)**

- Columns: date, classification (e.g., Fear, Greed, Extreme Greed, etc.)
- Cleaned to consolidate into 2 categories: "Fear" and "Greed"

➤ **Trader Dataset (historical_data.csv)**

- Columns include Timestamp IST, PnL, Position, Side, etc.
- Feature engineered to include:
    - trade_date from Timestamp IST
    - Merge with sentiment data on matching dates

Data Preprocessing

- Converted date and Timestamp IST to datetime
- Encoded the "Sentiment" label: Fear → 0, Greed → 1.
- Selected key trader performance features as model inputs.
- Missing values were handled using .fillna(0).

- Merged both datasets on date → merged_df

Modeling and Performance

Model Used: RandomForestClassifier

Target Variable: Sentiment_Label (0: Fear, 1: Greed)

Train/Test Split: 80% training, 20% testing.

Confusion Matrix:

	Predicted: Fear (0)	**Predicted: Greed (1)**
Actual: Fear (0)	45	55
Actual: Greed (1)	39	78

Classification Report:

Metric	**Fear (0)**	**Greed (1)**
Precision	0.54	0.59
Recall	0.45	0.67
F1-score	0.49	0.62

- **Overall Accuracy**: 0.57
- **Macro Avg F1**: 0.56

Insights

- A random forest model can pick up patterns in trader behavior that correlate with public sentiment, but with moderate predictive power.
- The relatively low accuracy (57%) suggests that trader performance alone doesn't fully determine market sentiment—external factors likely play a significant role.
- Trader performance metrics like Win Rate, Average PnL, and Total PnL are moderately predictive of market sentiment.

- **Greed days** are associated with **slightly higher win rates and average profitability**, according to the model and descriptive stats.
- The **Random Forest Classifier** had better recall for **Greed**, suggesting trader behavior is more consistent in optimistic markets.

Conclusion

While the model shows that there is some relation between aggregated trader behavior and daily sentiment trends, the moderate accuracy highlights that additional features (like news volume, social media mentions, or price volatility) could enhance predictive power.

Random Forest worked as a good baseline due to its robustness and non-linearity handling.