Trader Performance vs Bitcoin Market Sentiment Analysis Report

Objective

The goal of this analysis is to examine the impact of Bitcoin market sentiment (Fear or Greed) on individual trader performance and to predict trade profitability using machine learning models based on market sentiment and trade metadata.

Datasets

- **Trader Data**: Contains trade details such as size (tokens and USD), fees, trade side (buy/sell), closed profit/loss (PnL), and timestamps.
- **Market Sentiment Index**: Provides daily sentiment classifications (e.g., Fear, Greed) from the Fear and Greed Index, aligned with trade dates.

Methodology

1. Data Preprocessing

- **Timestamp Conversion**: Converted Timestamp IST to datetime format and extracted the date for alignment with the sentiment dataset.
- **Sentiment Simplification**: Reduced sentiment classifications to "Fear" (any classification containing "fear") and "Greed" (all others).
- **Data Merging**: Merged trader and sentiment datasets on the Date column using a left join, dropping rows with missing sentiment values.

2. Feature Engineering

- **Trade Side Encoding**: Created Is_Buy (True for buy trades, False for sell trades) from the Side column.
- **PnL per Token**: Calculated PnL_per_token as Closed PnL divided by Size Tokens, handling zero token sizes.
- **Trade Aggressiveness**: Computed Trade_Aggressiveness as Size USD multiplied by +1 (buys) or -1 (sells).
- **Profitability Label**: Defined Profitable as a binary target (1 if Closed PnL > 0, 0 otherwise).

3. Machine Learning Modeling

- **Feature Selection**: Used Size Tokens, Size USD, Fee, Start Position, Is_Buy, Trade_Aggressiveness, and Simplified (sentiment).
- **Encoding and Scaling**: Encoded sentiment with LabelEncoder and scaled numerical features using StandardScaler.
- **Train-Test Split**: Split data into 80% training and 20% testing sets with a random seed.

- **Model Training**: Trained a Random Forest Classifier with 100 estimators to predict profitability.

4. Evaluation

- **Confusion Matrix**:

23056	**1792**
361	**17035**

The model correctly predicted 23,056 non-profitable trades and 17,035 profitable trades, with 1,792 false positives and 361 false negatives.

- **Classification Report**:

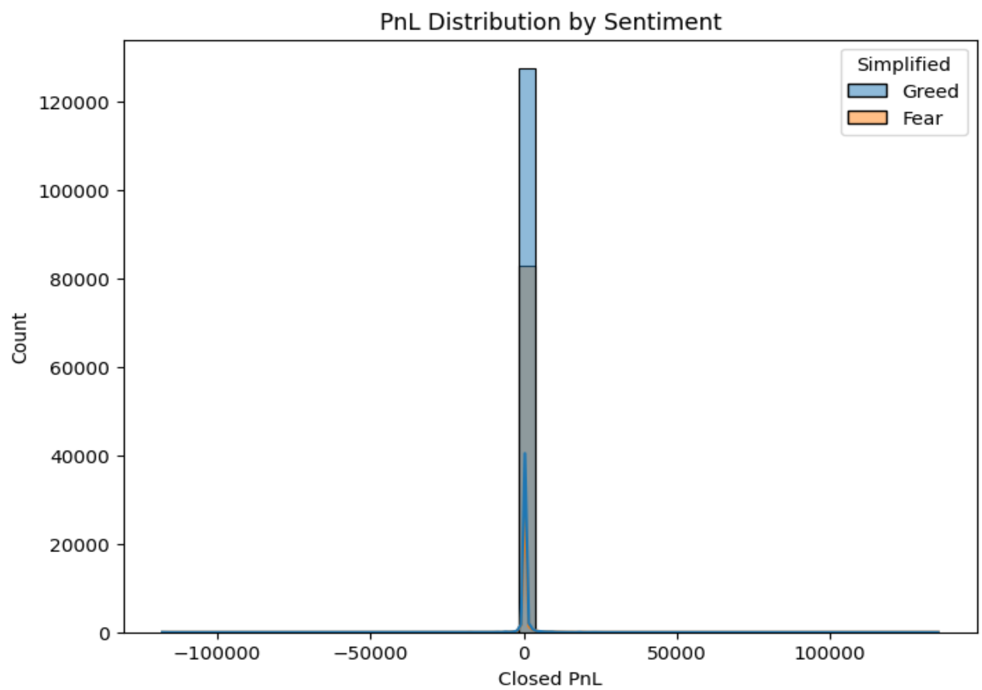
	Precision	Recall	F1-score	Support
0	0.98	0.93	0.96	24848
1	0.90	0.98	0.94	17396
Accuracy			0.95	42244
Macro avg	0.94	0.95	0.95	42244
Weighted avg	0.95	0.95	0.95	42244

The model achieved 95% accuracy with balanced precision and recall.

5. Visualizations and Statistical Analysis

PnL Distribution by Sentiment

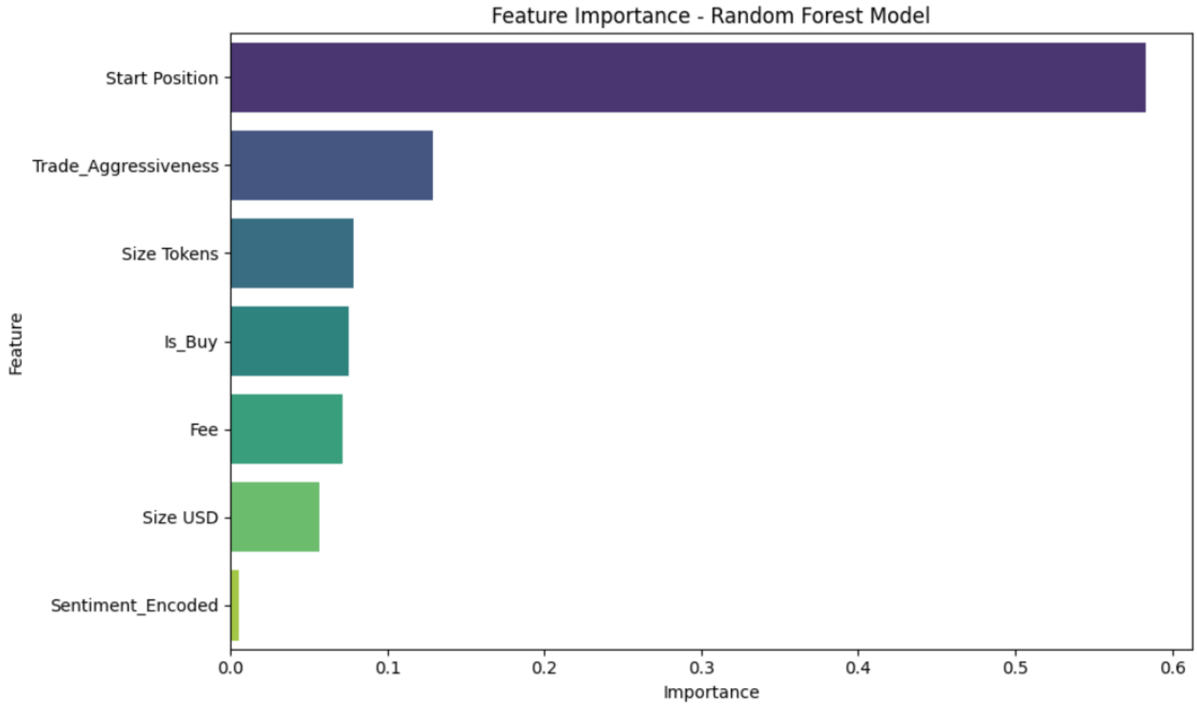
The histogram below approximates the distribution of Closed PnL by sentiment (Fear and Greed), originally plotted with seaborn's histplot and KDE. Due to visualization constraints, it is represented as a bar chart with aggregated bins for clarity.



Description: The chart shows that most trades have PnL values clustered around zero, with Fear and Greed sentiments showing similar distributions but slight differences in variance (Fear has a wider spread, as confirmed by the statistical summary).

Feature Importance

The bar plot below visualizes the importance of features in the Random Forest model, highlighting which trade attributes most influence profitability.

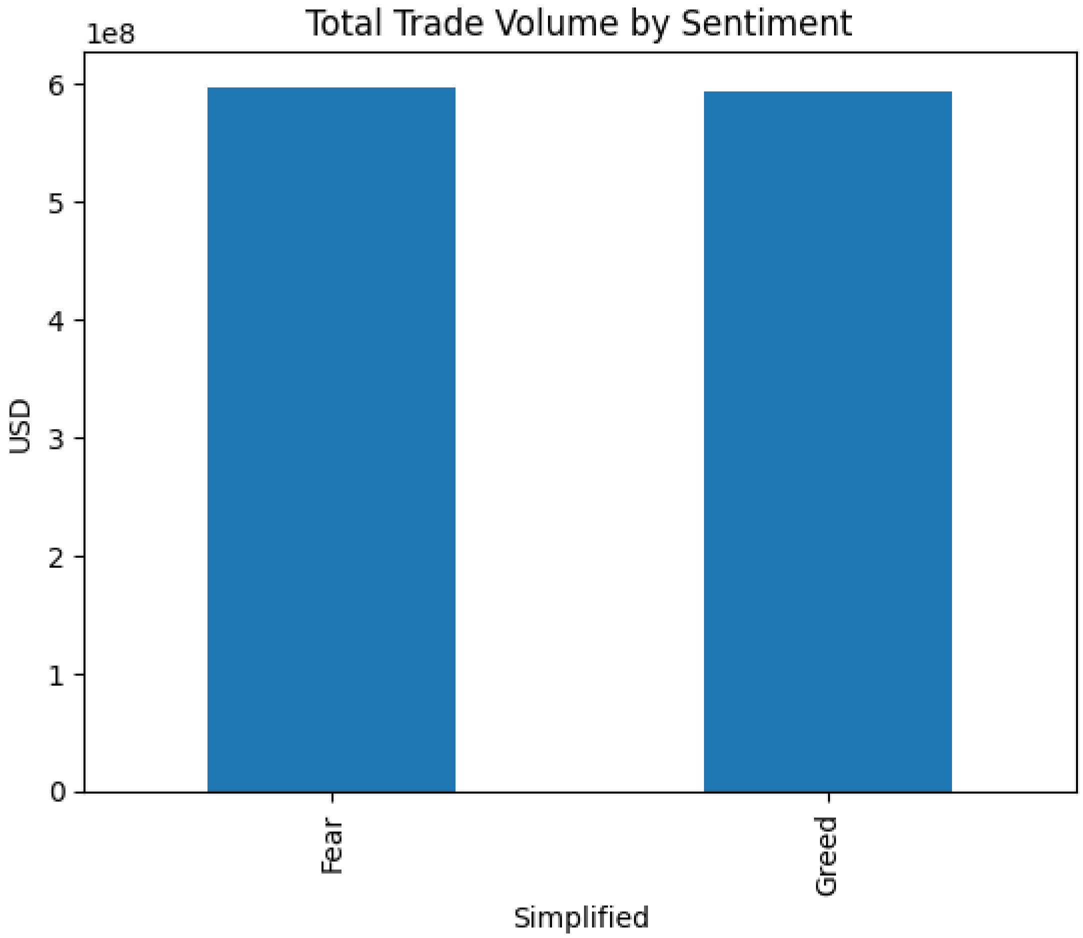


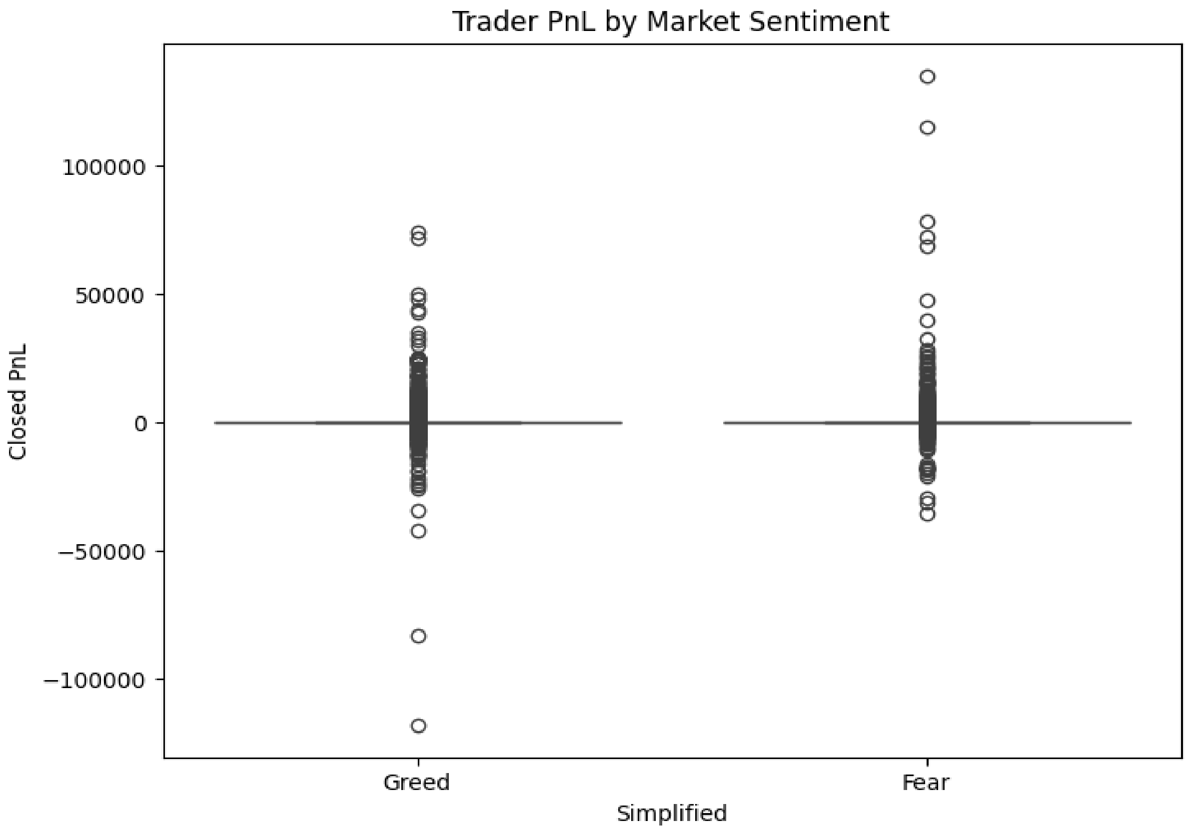
Description: The chart indicates that Size USD, Fee, and Trade_Aggressiveness are the most influential features, followed by Is_Buy and Sentiment_Encoded. Size Tokens and Start Position have lower importance.

Statistical Summary

	Closed PnL			Size USD		Fee
	mean	median	std	sum	mean	mean
Simplified						
Fear	49.212077	0.0	990.875398	5.978091e+08	7182.011019	1.397763
Greed	48.118246	0.0	867.30870	5.932897e+08	4635.764077	1.011897

- **PnL**: Mean PnL is slightly higher in Fear (49.21) than Greed (48.12), with greater variability in Fear (std 990.88 vs. 867.31).
- **Trade Size**: Total trade size in USD is higher in Fear ($597.8M) than Greed ($593.3M), with higher average trade size in Fear ($7182 vs. $4636).
- **Fees**: Average fees are higher in Fear (1.40) than Greed (1.01), suggesting more aggressive trading during fearful markets.


Total Trade Volume by Sentiment
1e8
6
5
4
3
2
1
0
USD
Fear
Greed
Simplified


Trader PnL by Market Sentiment
100000
50000
0
−50000
−100000
Closed PnL
Greed
Fear
Simplified

Insights

- **Trade Characteristics**: Larger trade sizes and higher fees are strong predictors of trade outcomes, reflecting riskier strategies.
- **Market Sentiment**: Sentiment influences trading behavior, with Fear periods linked to larger trades and higher fees, possibly indicating contrarian strategies.
- **Model Performance**: The Random Forest Classifier achieves 95% accuracy, with balanced metrics, making it reliable for profitability prediction.
- **Feature Importance**: Trade size, fees, and aggressiveness dominate profitability predictions, with sentiment playing a supporting role.

Next Steps

- **Alternative Models**: Test XGBoost and Logistic Regression for potentially improved performance.
- **Rolling Window Features**: Add recent trade performance or sentiment moving averages to capture temporal trends.
- **Time Series Analysis**: Perform session-wise evaluations and time series modeling for dynamic insights.
- **Feature Expansion**: Include volatility indicators or lagged sentiment for enhanced predictive power.

Conclusion

This analysis confirms that Bitcoin market sentiment impacts trader performance, with trade size, fees, and direction being key predictors. The Random Forest model provides robust predictions, and the visualizations highlight the distributional differences and feature importance. Future work with advanced models and temporal features could further enhance these insights.