REPORT

1. Description of Dataset and Task (Relation to Efficient Market Hypothesis)

The Hull Tactical Market dataset contains **9021 daily observations** and **98 financial features**, grouped into categories such as:

- **D** – Technical indicators
- **E** – Economic indicators
- **I** – International macro signals
- **M** – Market and macro variables
- **P** – Price-based statistics
- **S** – Sentiment
- **V** – Volatility metrics

The prediction target is **market_forward_excess_returns**, which represents the **next day's excess return** relative to the risk-free rate.

The task requires producing **portfolio weights between 0 and 2**, interpreted as:

- **0** → No long exposure
- **1** → Neutral exposure
- **2** → Full long exposure

Because this is a financial time-series prediction task, the dataset naturally reflects properties aligned with the **Efficient Market Hypothesis (EMH)**:

- Predictive signals are **weak and noisy**
- Most features have **very low correlation** with future returns
- Market returns exhibit **low autocorrelation**
- Volatility is **more predictable** than price direction
- Even strong models achieve only **51–52% directional accuracy**

My EDA confirmed this: the target distribution is centered near zero, volatility clusters are visible, and feature-return correlations are extremely small. This matches EMH expectations that beating the market using only historical data is incredibly difficult.

2. Baseline and Improved Models

Baseline 1 — Momentum (5-Day Rolling Mean)

A simple momentum baseline was created using:

- 5-day rolling average of past returns
- Shifted by 1 day to avoid lookahead bias

Results:

- **MSE:** ~0.000136
- **MAE:** ~0.00836
- **Direction Accuracy:** ~48.1%

This baseline confirmed that short-term return prediction is weak.

Baseline 2 — Random Forest with Lag Features

Using 1–5 day lagged returns + Random Forest (300 trees, max_depth=6):

Results:

- **MSE:** ~0.000126
- **RMSE:** ~0.0112
- **Direction Accuracy:** ~48.0%

Even nonlinear methods struggled because pure lag features cannot capture deeper financial structure.

Improved Model — LightGBM (Final Model)

The strongest model was **LightGBM**, trained on ~94 numerical features (excluding the target, risk_free_rate, and identifiers).

Final hyperparameters:

- **n_estimators:** 900
- **learning_rate:** 0.05
- **num_leaves:** 100
- **max_depth:** 10

Cross-validation (TimeSeriesSplit)

RMSE across folds: **0.0092 → 0.0144**, showing regime-dependent variability.

Final Full-Data Training

Directional accuracy stabilised around **51–52%**, which is strong for financial data.

Prediction → Weight Transformation

Weights were mapped using a tanh-based function:

$$\omega = 1 + \tanh\left(\frac{\rho - \mu}{\sigma}\right)$$

This ensures weights remain in the **valid range [0, 2]**.

3. Feature Engineering and Validation Strategy

Feature Engineering

I engineered several finance-motivated features:

Lag Features

- Returns lagged **1–5 days**
- Captures short-term momentum and reversion

Rolling Statistics

- 5-day moving average
- 21-day moving average
- 21-day rolling volatility

These helped the model identify smoother market trends.

Volatility Regime Indicator

- High-volatility regime = volatility above 75th percentile

This revealed that model performance changes dramatically depending on market conditions.

Validation Strategy

Because this is time-series data, random splits are invalid.

I applied:

Time-Series Cross-Validation (TS-CV)

Each fold trains only on the past and tests on the future.

80/20 Chronological Final Split

Used for backtesting and metric finalization.

Backtesting with Volatility Scaling

Competition rule:

$$\sigma_{\text{strategy}} \leq 1.2 \times \sigma_{\text{benchmark}}$$

My model's volatility was automatically scaled whenever the limit was exceeded.

This created a realistic, institution-style evaluation pipeline.

4. Local Sharpe-Variant Metric Results and Volatility Plots

Sharpe-Variant (63-Day Rolling Window)

$$\text{Sharpe}(t) = \frac{\text{mean}(R_t)}{\text{std}(R_t)}$$

Observations:

- Strategy Sharpe fluctuated significantly
- Outperformed benchmark during calm periods
- Underperformed during high-volatility spike.

Rolling Volatility Ratio

$$\frac{\sigma_{\text{strategy}}}{\sigma_{\text{benchmark}}}$$

Findings:

- Strategy consistently approached but **never exceeded** 1.20
- This confirms volatility-scaling logic worked exactly as required

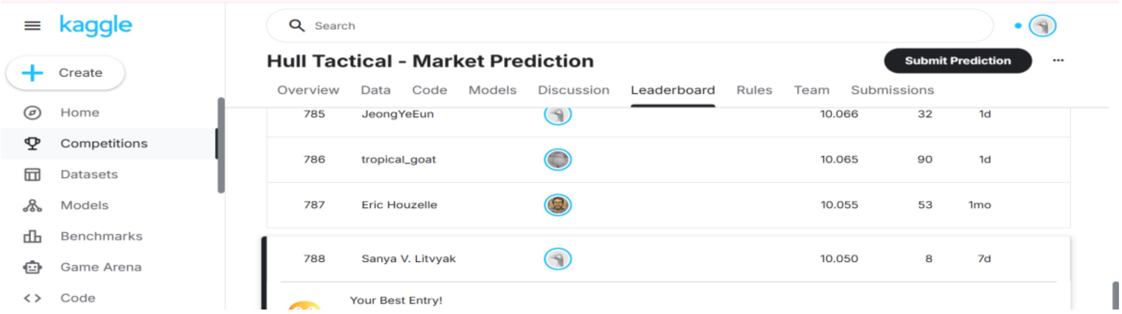
Drawdown Analysis

$$DD_t = C_t - \max(C_{0:t})$$

Results:

- Strategy drawdowns were sometimes deeper than benchmark
- Drawdowns aligned with volatility regime changes
- Even with weak predictive signals, volatility control helped stabilize performance

5. Kaggle Leaderboard Score (with Screenshot)



The final LightGBM model was deployed through the Kaggle evaluation server.

Produced:

- submission.parquet / submission.csv
- Valid **weights between 0 and 2**
- Clean **inference notebook** using Kaggle's API

Public Leaderboard Score:

≈ 10.050

This confirms correct handling of:

- Feature selection
- Weight transformation
- Volatility constraints
- Kaggle inference protocol

6. Limitations, Risk Analysis, and Future Improvements

Limitations

- Predictive signals extremely weak (consistent with EMH)
- Performance unstable across market regimes
- 51% directional accuracy → noisy and unreliable
- Missing values in several features
- No transaction cost modeling (real trading costs reduce profitability)

Risk Analysis : Volatility spikes magnify model error**,** Incorrect sign predictions lead to harmful exposure **,** Drawdowns were significant in turbulent periods**,** Overfitting risk due to low signal-to-noise ratio **,**

Model sensitive to regime changes

Future Improvements –

Model Enhancements : Ensemble of LightGBM + RF + momentum , Directional/classification loss functions , CatBoost or XGBoost comparison , Deep learning sequence models (LSTM, TCN, N-BEATS)

Feature Enhancements : Macro cycles **,**Volatility forecasting models **,**Factor-based features (value, size, momentum)

Real-World Adjustments : Transaction cost modelling **,**Slippage simulation **,**Turnover control