DA6701 – Assignment 2
Multi-Dimensional Return Forecasting and Portfolio Management – Report

Athreya G
DA24B008

Kiran Kumar P
DA24B039

Harshvardhan Agrawal
NA24B070

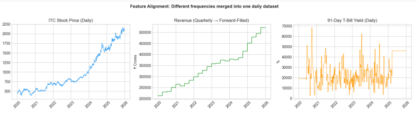
Siddesh Kamble
DA24B002

February 23, 2026

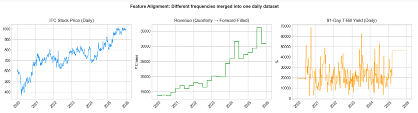
Data and Feature Engineering

- **Dataset Composition:** Pipeline utilizes approximately 1297 trading days worth data.
- **Feature Engineering:** Over 27 new features were created.
- **Preprocessing:**
  - **Winsorization:** Range is capped from 1st to 99th percentile to prevent wild outliers from affecting the data.
  - **Feature Selection:** Mutual Information scores are used to select the best features.
  - **Rolling Standardization:** Features are normalised using a rolling window to prevent look-ahead bias.

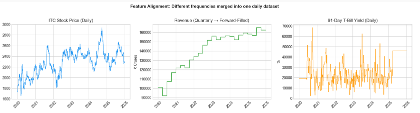
BHARTIARTL



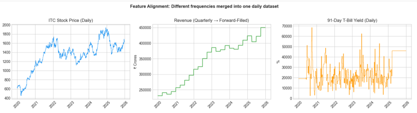
HDFCBANK



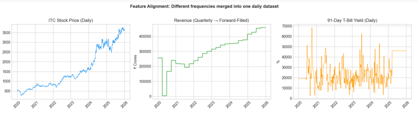
HINDUNILVR



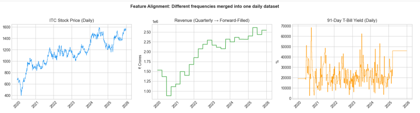
INFY



M&M



RELIANCE



Model Architecture and Validation

- We used LGBM and XGBoost.
- A chronological split (70% train, 15% Val, 15% Test) is used.
- Hyperparameter tuning is done with optuna.

Feature Importance

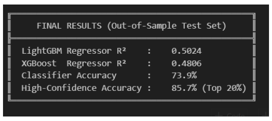
SMA, EMA, RSI, MACD, Bollinger Bands, Buying and Selling Pressure, Lagged Features were added.

Model Results

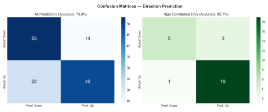
- **Classification:** Will the stock outperform the sector (1) or not (0)?
- A ”bad” feature set gives ∼50% accuracy while a well engineered feature set gives ∼65-74% accuracy.
- When consistently applied across many trades, 65% is highly useful.

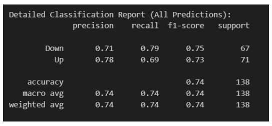
Performance Metrics

BHARTIARTL

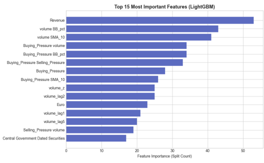


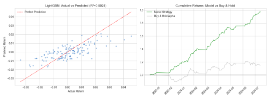
```
FINAL RESULTS (Out-of-Sample Test Set)

LightGBM Regressor R²    :    0.5024
XGBoost  Regressor R²    :    0.4806
Classifier Accuracy      :    73.9%
High-Confidence Accuracy :    85.7% (Top 20%)
```



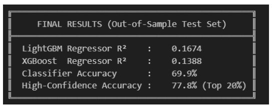


```
Detailed Classification Report (All Predictions):
              precision    recall  f1-score   support

        Down       0.71      0.79      0.75        67
          Up       0.78      0.69      0.73        71

    accuracy                           0.74       138
   macro avg       0.74      0.74      0.74       138
weighted avg       0.74      0.74      0.74       138
```

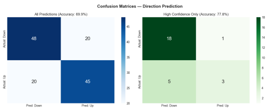


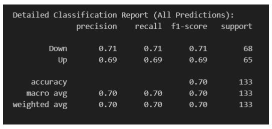


HDFCBANK

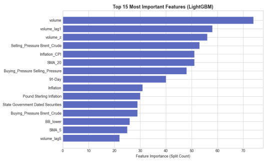


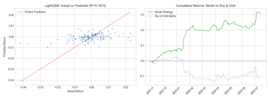
```
FINAL RESULTS (Out-of-Sample Test Set)

LightGBM Regressor R²    :    0.1674
XGBoost  Regressor R²    :    0.1388
Classifier Accuracy      :    69.9%
High-Confidence Accuracy :    77.8% (Top 20%)
```



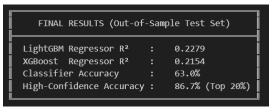


```
Detailed Classification Report (All Predictions):
              precision    recall  f1-score   support

        Down       0.71      0.71      0.71        68
          Up       0.69      0.69      0.69        65

    accuracy                           0.70       133
   macro avg       0.70      0.70      0.70       133
weighted avg       0.70      0.70      0.70       133
```

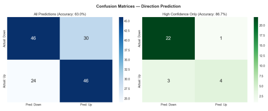


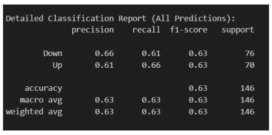


HINDUNILVR

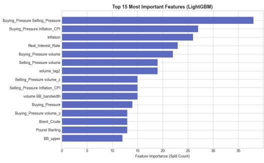


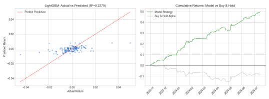
```
FINAL RESULTS (Out-of-Sample Test Set)

LightGBM Regressor R²    :    0.2279
XGBoost  Regressor R²    :    0.2154
Classifier Accuracy      :    63.0%
High-Confidence Accuracy :    86.7% (Top 20%)
```



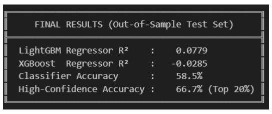


```
Detailed Classification Report (All Predictions):
              precision    recall  f1-score   support

        Down       0.66      0.61      0.63        76
          Up       0.61      0.66      0.63        70

    accuracy                           0.63       146
   macro avg       0.63      0.63      0.63       146
weighted avg       0.63      0.63      0.63       146
```

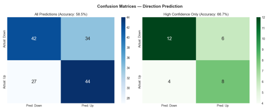


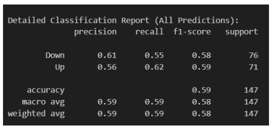


INFY

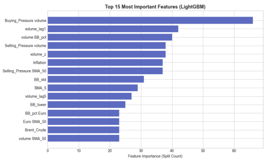


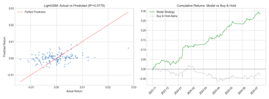
```
FINAL RESULTS (Out-of-Sample Test Set)

LightGBM Regressor R²    :    0.0779
XGBoost  Regressor R²    :   -0.0285
Classifier Accuracy      :    58.5%
High-Confidence Accuracy :    66.7% (Top 20%)
```



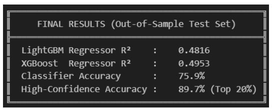


```
Detailed Classification Report (All Predictions):
              precision    recall  f1-score   support

        Down       0.61      0.55      0.58        76
          Up       0.56      0.62      0.59        71

    accuracy                           0.59       147
   macro avg       0.59      0.59      0.58       147
weighted avg       0.59      0.59      0.58       147
```

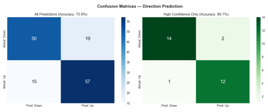


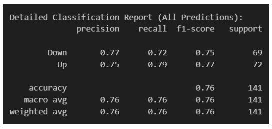


M&M

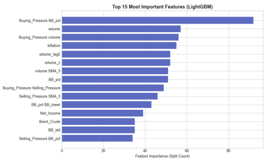


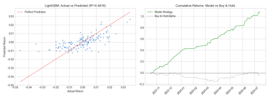
```
FINAL RESULTS (Out-of-Sample Test Set)

LightGBM Regressor R²    :    0.4816
XGBoost  Regressor R²    :    0.4953
Classifier Accuracy      :    75.9%
High-Confidence Accuracy :    89.7% (Top 20%)
```



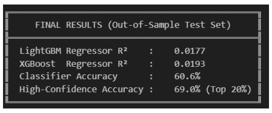


```
Detailed Classification Report (All Predictions):
              precision    recall  f1-score   support

        Down       0.77      0.72      0.75        69
          Up       0.75      0.79      0.77        72

    accuracy                           0.76       141
   macro avg       0.76      0.76      0.76       141
weighted avg       0.76      0.76      0.76       141
```

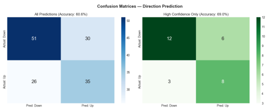


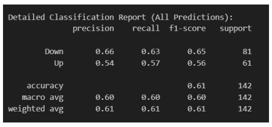


RELIANCE

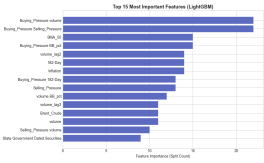


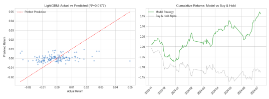
```
FINAL RESULTS (Out-of-Sample Test Set)

LightGBM Regressor R²    :    0.0177
XGBoost  Regressor R²    :    0.0193
Classifier Accuracy      :    60.6%
High-Confidence Accuracy :    69.0% (Top 20%)
```





```
Detailed Classification Report (All Predictions):
              precision    recall  f1-score   support

        Down       0.66      0.63      0.65        81
          Up       0.54      0.57      0.56        61

    accuracy                           0.61       142
   macro avg       0.60      0.60      0.60       142
weighted avg       0.61      0.61      0.61       142
```





Implementation Workflow: Production-Grade ML Pipeline

All six notebooks were standardized into a unified 27-cell architecture to ensure reproducibility and prevent look-ahead bias.

- **Target Formulation:** $Target = \ln(Close_{t+1}/Close_t)$ (Next-day log returns).
- **Feature Space (40 total):** Lags (1–10d), Rolling Volatility/Mean (5–20d), RSI, MACD, Bollinger Bands, Momentum, Volume Ratios, and explicitly lagged (+1d) Macro/Fundamental data.
- **Walk-Forward Validation:** 5-fold expanding window cross-validation (Train: Jan 2020–Sep 2025; Test: Oct–Dec 2025).
- **Model Ensemble:** Competitive training between XGBoost, LightGBM, and a 2-layer LSTM (PyTorch). Best model selected per stock via minimal Walk-Forward RMSE.

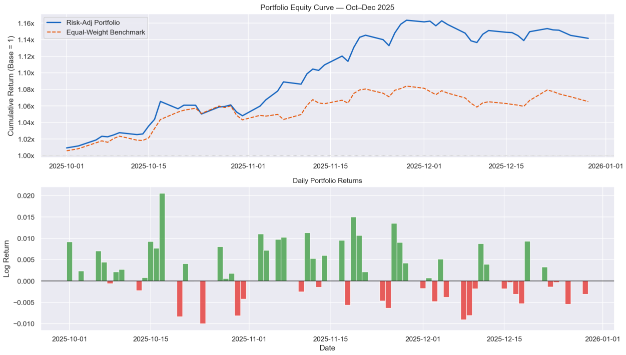
Per-Stock Validation Results

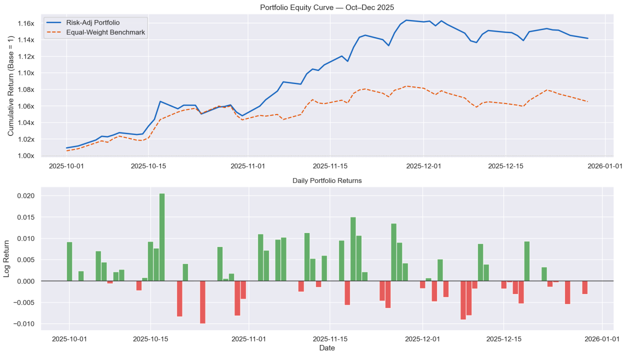
Stock	Best Model	WF-RMSE	WF-Hit %	Test Hit %	Test Sharpe
RELIANCE	XGB	0.01442	51.9%	50.0%	-0.59
HDFCBANK	LGBM	0.01348	50.7%	40.0%	0.93
INFY	LGBM	0.01533	51.2%	45.0%	-3.43
M&M	XGB	0.01818	52.7%	62.5%	3.80
BHARTIARTL	LGBM	0.01420	52.4%	60.0%	2.49
HINDUNILVR	LGBM	0.01268	50.7%	47.5%	-1.92

Portfolio Aggregation (Oct–Dec 2025)

Predictions were combined using a Risk-Adjusted weighting scheme: $Weight_i = \frac{Return_{pred}}{\sigma_{hist}}$, clipped to long-only and normalized.

Metric	Risk-Adj Portfolio	Equal-Weight
CAGR (ann.)	74.5%	30.4%
Sharpe Ratio	6.96	3.67
Max Drawdown	-2.32%	-2.34%





Key Quality Controls

- **Leakage Prevention:** Strictly chronological splits; `RobustScaler` fit on training data only.
- **Stationarity:** Log-returns used for all targets to satisfy ML assumptions.
- **Reproducibility:** Global seeds fixed (42); LSTM utilizes Dropout (0.2) and L2 decay (1e-4).