COMMODITY PRICE FORECASTING USING MULTI-MARKET FINANCIAL TIME SERIES DATA

GROUP MEMBERS

SAMWEL ONGECHI

MARILYN AKINYI

ERICK MAUTI

LYDIAH CHUMBA

ROSE MIRITI

ISAAC ONYANGO

RODGERS OTIENO

PRESENTED BY: GROUP 5

TM : GEORGE KAMUNDIA

BUSINESS UNDERSTANDING

- Volatility in commodity markets impacts global trade and investments
- Forecasting helps producers, traders, and policymakers make informed decisions.

BUSINESS PROBLEM

Predicting commodity returns is particularly challenging due to:

- Market volatility
- Time-lagged effects
- Dependencies across different financial instruments

How can we leverage time series data to predict commodity price movements?

STAKEHOLDERS

- Commodity traders – optimize trading strategies
- Investors –To manage portfolio risk
- Policy makers and regulators – anticipate inflationary trends
- Financial Analysts – provide data-driven forecasts

OBJECTIVES

- To preprocess and integrate multi-market financial time series data for commodities.
- To explore and analyze patterns, correlations, and volatility in commodity returns.
- To develop predictive models for selected commodity prices and spreads.
- To evaluate model accuracy and stability using appropriate performance metrics.
- To generate insights that can support trading strategies and risk management.

DATA UNDERSTANDING

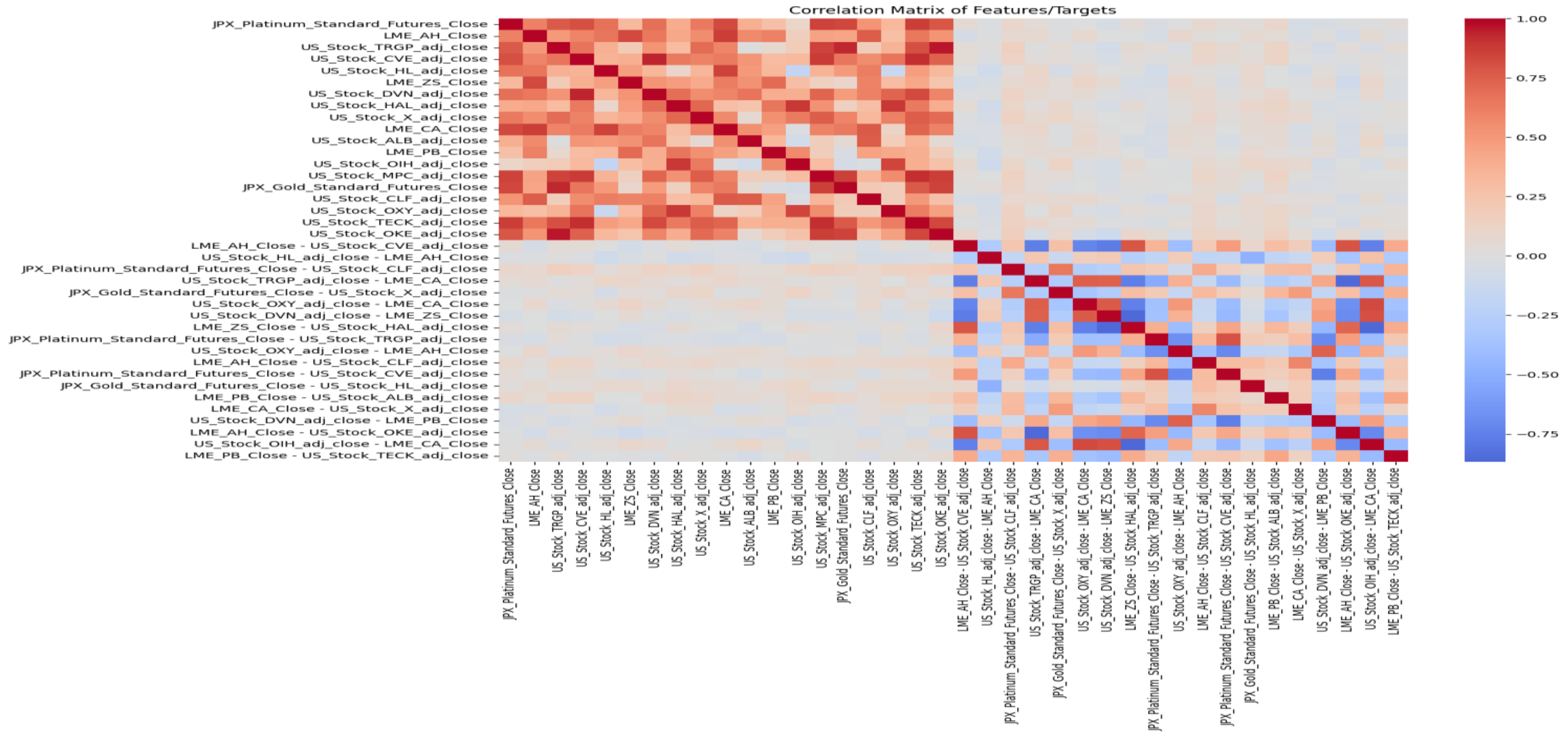
- **Sources:** Kaggle (London Metal Exchange (LME), Japan Exchange Group (JPX), U.S. stock markets.
- **Data used:** Train.csv, Target pairs and Train_labels.csv
- **Time horizon**: Daily
- **Challenges**: missing values, seasonality, volatility, external shocks

EXPLORATORY DATA ANALYSIS(EDA)

- Understand the distributions of target variables and features.
- Identify trends, patterns, and outliers in the data.
- Explore relationships between features and targets, including correlations and potential predictive signals.
- Visualize time series behaviors since this is a temporal dataset.

This step helps inform **feature engineering, model selection, and further preprocessing decisions**.

EDA CONT 'D



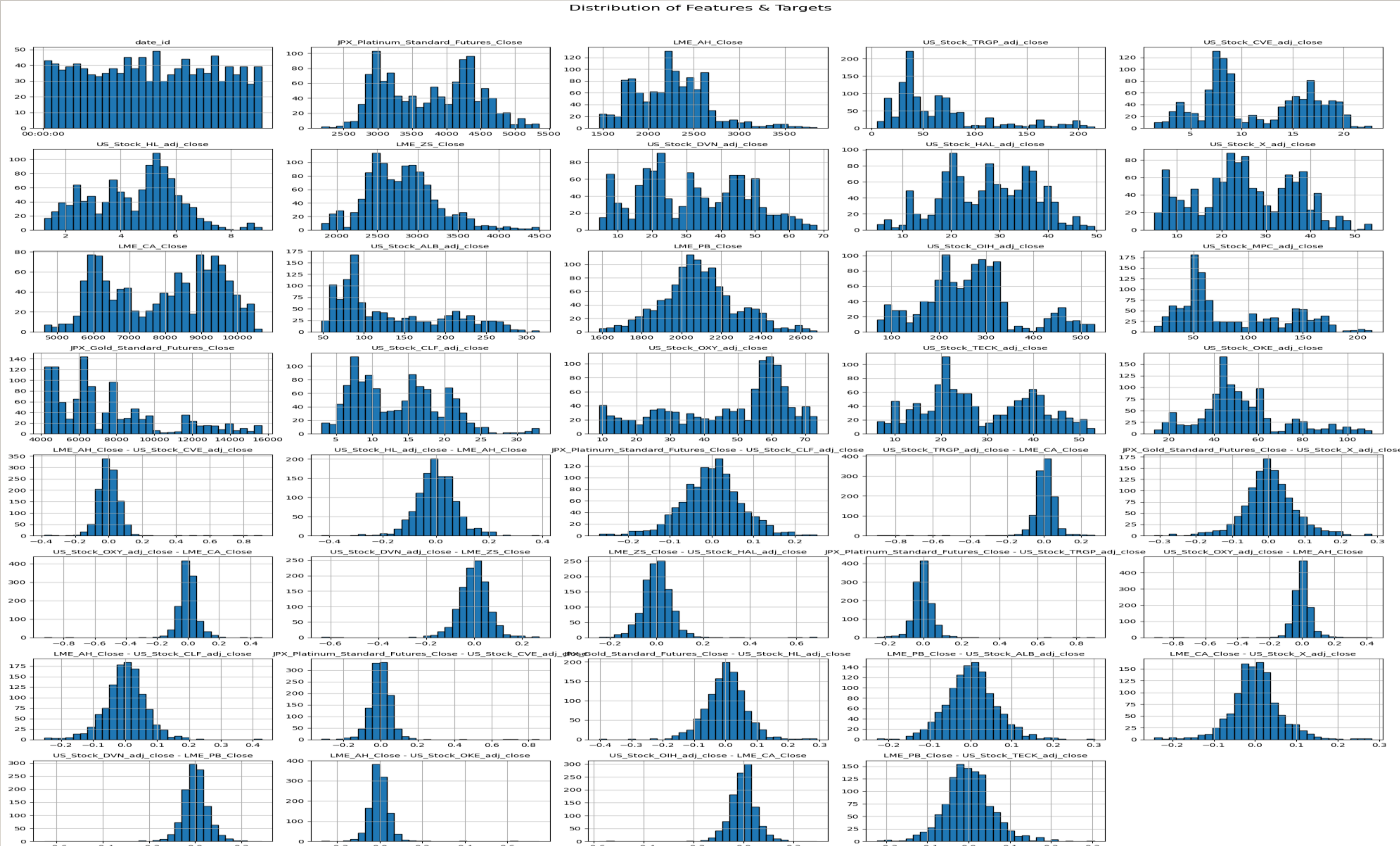
Correlation Matrix of Features/Targets
JPX_Platinum_Standard_Futures_Close
LME_AH_Close
US_Stock_TRGP_adj_close
US_Stock_CVE_adj_close
US_Stock_HL_adj_close
LME_ZS_Close
US_Stock_DVN_adj_close
US_Stock_HAL_adj_close
US_Stock_X_adj_close
LME_CA_Close
US_Stock_ALB_adj_close
LME_PB_Close
US_Stock_OIH_adj_close
US_Stock_MPC_adj_close
JPX_Gold_Standard_Futures_Close
US_Stock_CLF_adj_close
US_Stock_OXY_adj_close
US_Stock_TECK_adj_close
US_Stock_OKE_adj_close
LME_AH_Close - US_Stock_CVE_adj_close
US_Stock_HL_adj_close - LME_AH_Close
JPX_Platinum_Standard_Futures_Close - US_Stock_CLF_adj_close
US_Stock_TRGP_adj_close - LME_CA_Close
JPX_Gold_Standard_Futures_Close - US_Stock_X_adj_close
US_Stock_OXY_adj_close - LME_CA_Close
US_Stock_DVN_adj_close - LME_ZS_Close
LME_ZS_Close - US_Stock_HAL_adj_close
JPX_Platinum_Standard_Futures_Close - US_Stock_TRGP_adj_close
US_Stock_OXY_adj_close - LME_AH_Close
LME_AH_Close - US_Stock_CLF_adj_close
JPX_Platinum_Standard_Futures_Close - US_Stock_CVE_adj_close
JPX_Gold_Standard_Futures_Close - US_Stock_HL_adj_close
LME_PB_Close - US_Stock_ALB_adj_close
LME_CA_Close - US_Stock_X_adj_close
US_Stock_DVN_adj_close - LME_PB_Close
LME_AH_Close - US_Stock_OKE_adj_close
US_Stock_OIH_adj_close - LME_CA_Close
LME_PB_Close - US_Stock_TECK_adj_close
1.00
0.75
0.50
0.25
0.00
−0.25
−0.50
−0.75

EDA CONT 'D

- Applied 0.9 correlation threshold to flag highly correlated pairs.
- A total of 5 pairs exceeded this threshold, meaning they may introduce multicollinearity risks if included together in models.

DISTRIBUTION OF FEATURES AND TARGETS

EDA CONT 'D

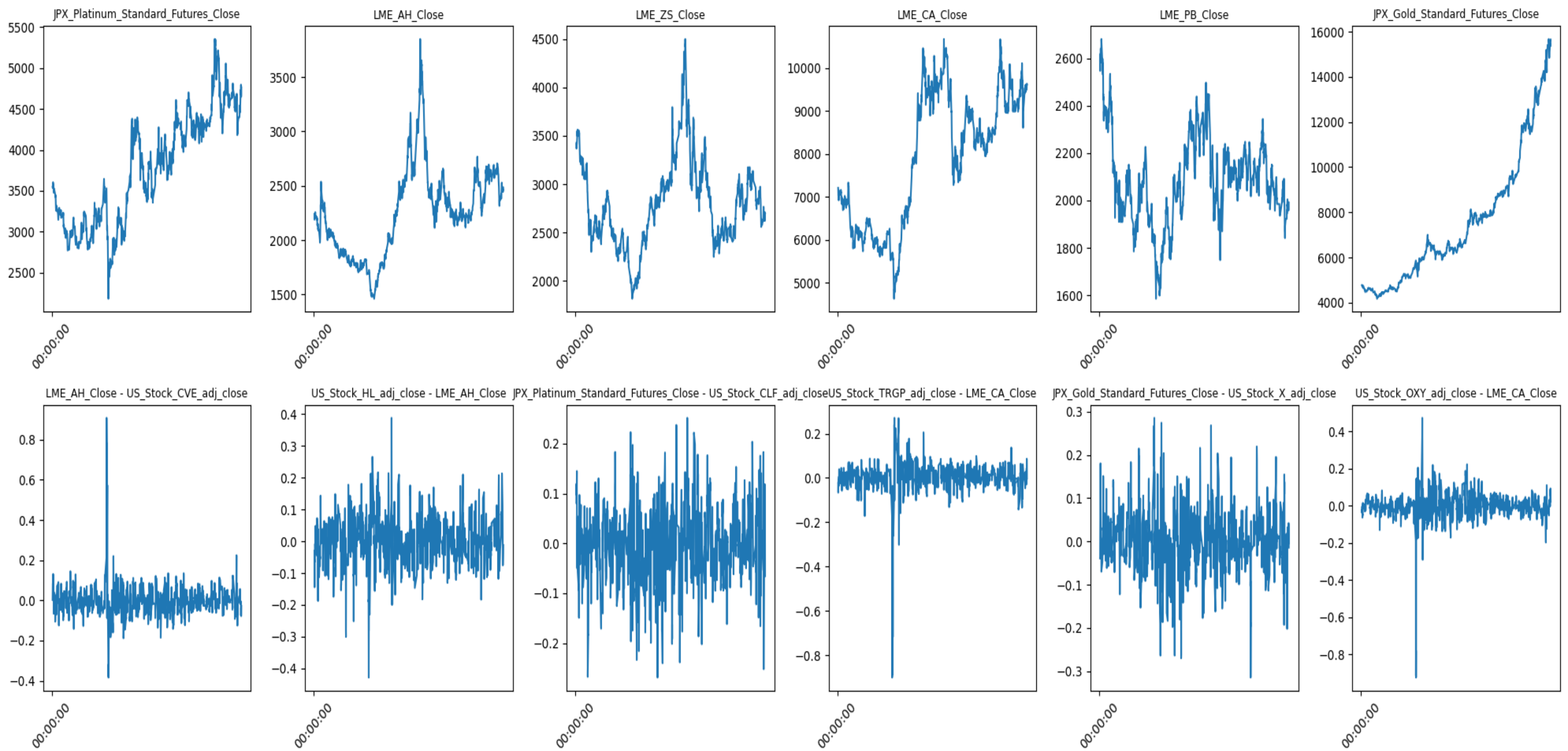


EDA CONT ‘D

- The histograms above illustrate the distributions of our merged dataset’s features and targets.
- dataset combines skewed raw price distributions with nearly-normal return distributions.
- To improve model performance, we should consider applying transformations to raw prices and using returns directly as predictors.

TARGET TRENDS OVER TIME

EDA CONT ‘D



TARGET TRENDS OVER TIME

- **Overall trend** : To form the baseline for prediction.
- **Volatility:** Indicates market risk and uncertainty which impacts model choice.
- **Seasonality/patterns :** Provides insights for timing decisions, e.g when to buy/sell commodities.
- **Outliers:** Could be caused by market shocks, policy changes, or data errors.

Forms the basis for data preprocessing and feature engineering

DATA PREPARATION

- Cleaning missing values & outliers
- Checking for duplicates and data types
- Normalization / scaling of variables
- Mapping
- Handling stationarity: differencing, log returns
- Feature engineering: lags, rolling averages, volatility measures

DATA PREPROCESSING AND FEATURE ENGINEERING

- Stationary check
- Spread calculations to create spread columns. These help us see how one asset is performing compared to another.
- Create lag features for our dataset while avoiding data leakage

MODELING

BASELINE MODELS

- Ridge & Lasso Regression
- Linear Regression
- Random Forest LightGBM
- XGboost

ADVANCE MODELS

- Gradient Boosting
- ElasticNet
- SVR
- MLP Regressor

EVALUATION

- Mean Absolute Error (MAE): Average magnitude of prediction errors.
- Root Mean Squared Error (RMSE): Measures error magnitude with greater penalty for larger deviations.
- R-squared (R^2): Proportion of variance explained by the model.
- MAPE (Mean Absolute Percentage Error)
- SMAPE (Symmetric MAPE)
- Directional Accuracy: Percentage of correct predictions in the direction of price movement.

EVALUATION BASELINE MODELS

Model	MAE	RMSE	R2	MAPE	SMAPE	Directional Accuracy	CombinedScore
Ridge	0.0224	0.0332	0.4202	351.06	108.46	0.7228	2.5091
LinearRegressio	0.0224	0.0332	0.4194	351.51	108.46	0.7227	2.5057
RandomForest	0.0232	0.0362	0.3513	346.83	111.77	0.7219	2.2492
LightGBM	0.024	0.0376	0.3046	375.55	113.66	0.7049	2.058
XGBoost	0.0255	0.0397	0.1742	412.73	115.53	0.6927	1.6332
Lasso	0.0297	0.0451	-0.0049	113.09	185.04	0	-0.0117

EVALUATION BASELINE MODELS CONT’D

- Ridge and Linear Regression achieved the highest combined score with directional accuracy of 0.72 and reasonable MAE/RMSE values.
- Tree-based models performed well for most targets.
- Lasso showed weaker performance on several targets, with very low directional accuracy for some folds.

EVALUATION ADVANCED MODELS

Model	MAE	RMSE	R2	MAPE	SMAPE	Directional Accuracy	CombinedScore
GradientBoosting	0.0237	0.0371	0.319	352.72	111.85	0.7102	3.61
SVR	0.031	0.0457	-0.041	420.23	140.69	0.5804	2.19
ElasticNet	0.0297	0.0451	-0.0049	113.09	185.04	0	1.50
MLP	0.125	0.1858	-16.9889	2923.83	155.60	0.5402	-52.50

EVALUATION ADVANCED CONT'D

- Gradient Boosting showed the strongest performance among this set, with an R^2 of 0.32 and directional accuracy of 0.71.
- SVR and ElasticNet underperformed, with negative or near zero R^2 scores. SVR struggled with scalability, while ElasticNet failed to capture complex relationships in spread movements.
- MLP Regressor performed poorly, with very high MAE/RMSE and strongly negative R^2.

TIME SERIES MODELS

- SARIMA (Seasonal AutoRegressive Integrated Moving Average)
- SARIMA + Ridge Hybrid

Model	MAE	RMSE	R2	MAPE	SMAPE	Directional Accuracy	CombinedScore
SARIMA+Ridge	0.0225	0.0333	0.4154	359.1141	108.732	0.7222	2.5036
SARIMA	0.0297	0.0451	-0.0058	114.3862	183.137	0.019	0.0125

TIME SERIES MODEL EVALUATION

- SARIMA alone produced relatively weak performance, with low directional accuracy and near-zero or negative R^2 on several targets.
- The SARIMA + Ridge Hybrid consistently improved accuracy, reducing both MAE and RMSE while significantly boosting directional accuracy to 72 and R^2 of 42 compared to plain SARIMA.

COMBINED MODEL EVALUATION

Model	MAE	RMSE	R2	MAPE	SMAPE	Directional Accurac	CombinedScore
Ridge_Tuned	0.0197	0.0289	0.55	402.395	107.227	0.74	3.6867
Ridge	0.0224	0.0332	0.42	351.062	108.461	0.7228	3.3827
LinearRegressio	0.0224	0.0332	0.419	351.511	108.461	0.7227	3.3807
SARIMA+Ridge	0.0225	0.0333	0.415	359.114	108.732	0.7222	3.3717
RandomForest	0.0232	0.0362	0.351	346.828	111.767	0.7219	3.2334
GradientBoostin	0.0237	0.0371	0.319	352.717	111.854	0.7102	3.1504
LightGBM	0.024	0.0376	0.305	375.549	113.664	0.7049	3.112
XGBoost	0.0255	0.0397	0.174	412.728	115.526	0.6927	2.8353
SVR	0.031	0.0457	-0.04	420.227	140.694	0.5804	2.216
SARIMA	0.0297	0.0451	-0.01	114.386	183.137	0.019	1.5354
ElasticNet	0.0297	0.0451	-0	113.088	185.035	0	1.5113
Lasso	0.0297	0.0451	-0	113.088	185.035	0	1.5113
NaivePersistenc	0.0325	0.0474	-0.15	400.344	152.731	0	1.2184
MLP	0.125	0.1858	-17	2923.83	155.601	0.5402	-30.1365

COMBINED MODEL EVALUATION CONT'D

- Best Model Ridge_Tuned achieved the highest CombinedScore, outperforming both untuned Ridge and advanced ensemble models.
- Strong Baselines: Ridge and LinearRegression ranked closely behind, showing that linear methods remain competitive.
- Hybrid Success: SARIMA+Ridge performed better than standalone SARIMA, proving that combining time-series modeling with machine learning improves results.
- Tree-based Models: RandomForest and GradientBoosting showed strong but slightly weaker performance compared to tuned Ridge.
- Weaker Models: ElasticNet, Lasso, and NaivePersistence underperformed significantly, while MLP collapsed with highly negative R^2.
- SARIMA Alone: Bare SARIMA struggled, but adding Ridge correction significantly boosted its performance.

ENSEMBLE MODELING AND EVALUATION

- Selected the top 3 performing models based on a combination of R^2, MAE, and directional accuracy to do the ensemble modelling.

Rank	Model	R2	Directional Accuracy
1	Ensemble-Weighted-Top3	0.7227	0.8028
2	Ensemble-Mean-Top3	0.6716	0.784
3	Ridge_Tuned	0.5504	0.74
4	Ridge	0.4202	0.7228
5	LinearRegression	0.4194	0.7227
6	SARIMA+Ridge	0.4154	0.7222
7	RandomForest	0.3513	0.7219
8	GradientBoosting	0.319	0.7102

ENSEMBLE CONT’D

- Best single model R^2: 0.5504
- Best ensemble R^2: 0.7227
- R^2 improvement: +0.1723

The ensemble outperforms individual models, providing more accurate predictions.

FEATURE IMPORTANCE

Rank	Feature	Importance
1	US_Stock_OIH_adj_close	0.1853
2	US_Stock_HAL_adj_close	0.1216
3	US_Stock_CVE_adj_close	0.1055
4	LME_CA_Close	0.0887
5	US_Stock_OKE_adj_close	0.0882
6	US_Stock_OXY_adj_close	0.0832
7	LME_ZS_Close	0.0813
8	US_Stock_DVN_adj_close	0.0758
9	LME_PB_Close	0.0579
10	US_Stock_TRGP_adj_close	0.0535
11	LME_AH_Close	0.0491
12	US_Stock_TECK_adj_close	0.0279
13	JPX_Platinum_Standard_Future	0.0276
14	JPX_Gold_Standard_Futures_C	0.0157
15	US_Stock_HL_adj_close	0.0132

RECOMMENDATIONS

- Go with the Ensemble-Weighted-Top3 Model
- Energy Sector Features Matter Most
- Markets Don’t Move in Isolation

LIMITATIONS

Market regimes change

- The model is tuned for today's conditions but markets evolve
- Over-reliance on energy features might reduce accuracy in tech-led or defensive markets.
- External shocks (wars, pandemics, policy changes)

Validation Gaps

- Back tests are based on past data true future-proofing is untested.
- Assumes markets remain liquid execution could fail during crises.

DEPLOYMENT

- We deployed the **Weighted Ensemble Method** https://commodity-price-forecasting.streamlit.app/ as a price forecasting app.

NEXT STEP

- Strengthen Validation Testing
- Smarter Feature Engineering
- Build a Monitoring System