First 5 rows of the dataset:

Date Open High Low Close Adj Close Volume

0 8/19/2004 2.490664 2.591785 2.390042 2.499133 2.499133 897427216

1 8/20/2004 2.515820 2.716817 2.503118 2.697639 2.697639 458857488

2 8/23/2004 2.758411 2.826406 2.716070 2.724787 2.724787 366857939

3 8/24/2004 2.770615 2.779581 2.579581 2.611960 2.611960 306396159

4 8/25/2004 2.614201 2.689918 2.587302 2.640104 2.640104 184645512

Last 5 rows of the dataset:

Date Open High Low Close Adj Close Volume

5122 2024-12-24 00:00:00+00:00 194.839996 196.110001 193.779999 196.110001 196.110001 10403300

5123 2024-12-26 00:00:00+00:00 195.149994 196.750000 194.380005 195.600006 195.600006 12046600

5124 2024-12-27 00:00:00+00:00 194.949997 195.320007 190.649994 192.759995 192.759995 18891400

5125 2024-12-30 00:00:00+00:00 189.800003 192.550003 189.119995 191.240005 191.240005 14264700

5126 2024-12-31 00:00:00+00:00 191.080002 191.960007 188.509995 189.300003 189.300003 17466900

Dataset Information:

<class 'pandas.core.frame.DataFrame'>

RangeIndex: 5127 entries, 0 to 5126

Data columns (total 7 columns):

# Column Non-Null Count Dtype

--- ------ -------------- -----

0 Date 5127 non-null object

1 Open 5127 non-null float64

2 High 5127 non-null float64

3 Low 5127 non-null float64

4 Close 5127 non-null float64

5 Adj Close 5127 non-null float64

6 Volume 5127 non-null int64

dtypes: float64(5), int64(1), object(1)

memory usage: 280.5+ KB

Summary Statistics:

Open High Low Close Adj Close Volume

count 5127.000000 5127.000000 5127.000000 5127.000000 5127.000000 5.127000e+03

mean 47.795052 48.301175 47.314533 47.816238 47.793686 1.141486e+08

std 46.328424 46.843789 45.865890 46.359642 46.306432 1.486301e+08

min 2.470490 2.534002 2.390042 2.490913 2.490913 1.584340e+05

25% 13.130409 13.248591 12.985950 13.122065 13.122065 2.734500e+07

50% 27.654345 27.818624 27.345249 27.632925 27.632925 5.330600e+07

75% 62.660749 63.459000 62.215750 63.094000 63.094000 1.394226e+08

max 197.250000 201.419998 194.979996 196.660004 196.660004 1.650833e+09

The Columns:

Index(['Date', 'Open', 'High', 'Low', 'Close', 'Adj Close', 'Volume'], dtype='object')

Data types:

Date object

Open float64

High float64

Low float64

Close float64

Adj Close float64

Volume int64

dtype: object

Missing Values Count:

Date 0

Open 0

High 0

Low 0

Close 0

Adj Close 0

Volume 0

dtype: int64

Missing Values Count After Handling:

Date 0

Open 0

High 0

Low 0

Close 0

Adj Close 0

Volume 0

dtype: int64

Dataset Shape After Cleaning (Rows, Columns): (5127, 7)

Date column converted and new time-based features added.

Moving Averages, Daily Returns, and Volatility Metrics added.

Data split into training and testing sets.

Model Mean Absolute Error Mean Squared Error R² Score Training Time (s)

0 Linear Regression 0.197588 0.137925 0.999905 0.001005

1 Ridge Regression 0.197632 0.137915 0.999905 0.001000

2 Random Forest 0.259306 0.261978 0.999820 0.849062

3 Gradient Boosting 0.329720 0.320786 0.999780 0.397674

4 Support Vector Regression 17.655519 839.079749 0.424353 0.389313

Iteration: 5, Func. Count: 33, Neg. LLF: 31744.744117809943

Iteration: 10, Func. Count: 58, Neg. LLF: 31744.735314322275

Iteration: 15, Func. Count: 83, Neg. LLF: 31743.598163968636

Iteration: 20, Func. Count: 108, Neg. LLF: 31742.29648671578

Iteration: 25, Func. Count: 137, Neg. LLF: 31742.296396557093

Optimization terminated successfully (Exit mode 0)

Current function value: 31742.2963956122

Iterations: 25

Function evaluations: 138

Gradient evaluations: 25

Generated Plots:

- stl\_decomposition.png

- acf\_plot.png

- pacf\_plot.png

- prophet\_components.png

- garch\_volatility.png

Epoch 1/50

240/240 ━━━━━━━━━━━━━━━━━━━━ 4s 6ms/step - loss: 0.0053 - val\_loss: 0.0071

Epoch 2/50

240/240 ━━━━━━━━━━━━━━━━━━━━ 1s 4ms/step - loss: 6.6886e-05 - val\_loss: 0.0078

Epoch 3/50

240/240 ━━━━━━━━━━━━━━━━━━━━ 1s 4ms/step - loss: 6.6202e-05 - val\_loss: 0.0088

Epoch 4/50

240/240 ━━━━━━━━━━━━━━━━━━━━ 1s 4ms/step - loss: 6.2930e-05 - val\_loss: 0.0029

Epoch 5/50

240/240 ━━━━━━━━━━━━━━━━━━━━ 1s 5ms/step - loss: 6.2752e-05 - val\_loss: 0.0023

Epoch 6/50

240/240 ━━━━━━━━━━━━━━━━━━━━ 1s 5ms/step - loss: 6.8752e-05 - val\_loss: 0.0036

Epoch 7/50

240/240 ━━━━━━━━━━━━━━━━━━━━ 1s 5ms/step - loss: 5.4961e-05 - val\_loss: 0.0037

Epoch 8/50

240/240 ━━━━━━━━━━━━━━━━━━━━ 1s 5ms/step - loss: 4.3563e-05 - val\_loss: 0.0036

Epoch 9/50

240/240 ━━━━━━━━━━━━━━━━━━━━ 1s 5ms/step - loss: 5.1766e-05 - val\_loss: 0.0024

Epoch 10/50

240/240 ━━━━━━━━━━━━━━━━━━━━ 1s 5ms/step - loss: 4.8364e-05 - val\_loss: 0.0011

Epoch 11/50

240/240 ━━━━━━━━━━━━━━━━━━━━ 1s 5ms/step - loss: 5.5529e-05 - val\_loss: 0.0014

Epoch 12/50

240/240 ━━━━━━━━━━━━━━━━━━━━ 1s 5ms/step - loss: 4.9432e-05 - val\_loss: 0.0019

Epoch 13/50

240/240 ━━━━━━━━━━━━━━━━━━━━ 1s 5ms/step - loss: 4.2758e-05 - val\_loss: 0.0012

Epoch 14/50

240/240 ━━━━━━━━━━━━━━━━━━━━ 1s 5ms/step - loss: 4.8008e-05 - val\_loss: 0.0020

Epoch 15/50

240/240 ━━━━━━━━━━━━━━━━━━━━ 1s 4ms/step - loss: 4.5917e-05 - val\_loss: 0.0013

Epoch 16/50

240/240 ━━━━━━━━━━━━━━━━━━━━ 1s 5ms/step - loss: 3.2023e-05 - val\_loss: 6.4886e-04

Epoch 17/50

240/240 ━━━━━━━━━━━━━━━━━━━━ 1s 5ms/step - loss: 3.8262e-05 - val\_loss: 0.0017

Epoch 18/50

240/240 ━━━━━━━━━━━━━━━━━━━━ 1s 5ms/step - loss: 4.0564e-05 - val\_loss: 0.0015

Epoch 19/50

240/240 ━━━━━━━━━━━━━━━━━━━━ 1s 5ms/step - loss: 3.8206e-05 - val\_loss: 4.8812e-04

Epoch 20/50

240/240 ━━━━━━━━━━━━━━━━━━━━ 1s 5ms/step - loss: 2.7689e-05 - val\_loss: 0.0014

Epoch 21/50

240/240 ━━━━━━━━━━━━━━━━━━━━ 1s 5ms/step - loss: 2.9519e-05 - val\_loss: 8.6350e-04

Epoch 22/50

240/240 ━━━━━━━━━━━━━━━━━━━━ 1s 5ms/step - loss: 2.8326e-05 - val\_loss: 0.0011

Epoch 23/50

240/240 ━━━━━━━━━━━━━━━━━━━━ 1s 4ms/step - loss: 2.6703e-05 - val\_loss: 0.0011

Epoch 24/50

240/240 ━━━━━━━━━━━━━━━━━━━━ 1s 5ms/step - loss: 2.5547e-05 - val\_loss: 5.0211e-04

Epoch 25/50

240/240 ━━━━━━━━━━━━━━━━━━━━ 1s 5ms/step - loss: 2.5049e-05 - val\_loss: 7.4344e-04

Epoch 26/50

240/240 ━━━━━━━━━━━━━━━━━━━━ 1s 5ms/step - loss: 1.9188e-05 - val\_loss: 0.0013

Epoch 27/50

240/240 ━━━━━━━━━━━━━━━━━━━━ 1s 5ms/step - loss: 2.4858e-05 - val\_loss: 3.4555e-04

Epoch 28/50

240/240 ━━━━━━━━━━━━━━━━━━━━ 1s 5ms/step - loss: 3.4669e-05 - val\_loss: 0.0013

Epoch 29/50

240/240 ━━━━━━━━━━━━━━━━━━━━ 1s 5ms/step - loss: 1.8877e-05 - val\_loss: 8.1678e-04

Epoch 30/50

240/240 ━━━━━━━━━━━━━━━━━━━━ 1s 5ms/step - loss: 1.8138e-05 - val\_loss: 8.7481e-04

Epoch 31/50

240/240 ━━━━━━━━━━━━━━━━━━━━ 1s 5ms/step - loss: 1.7419e-05 - val\_loss: 6.0498e-04

Epoch 32/50

240/240 ━━━━━━━━━━━━━━━━━━━━ 1s 5ms/step - loss: 1.6016e-05 - val\_loss: 6.3097e-04

Epoch 33/50

240/240 ━━━━━━━━━━━━━━━━━━━━ 1s 5ms/step - loss: 1.7024e-05 - val\_loss: 5.9787e-04

Epoch 34/50

240/240 ━━━━━━━━━━━━━━━━━━━━ 1s 5ms/step - loss: 1.8575e-05 - val\_loss: 6.5494e-04

Epoch 35/50

240/240 ━━━━━━━━━━━━━━━━━━━━ 1s 5ms/step - loss: 1.9553e-05 - val\_loss: 6.6566e-04

Epoch 36/50

240/240 ━━━━━━━━━━━━━━━━━━━━ 1s 5ms/step - loss: 1.9953e-05 - val\_loss: 5.6824e-04

Epoch 37/50

240/240 ━━━━━━━━━━━━━━━━━━━━ 1s 5ms/step - loss: 1.5140e-05 - val\_loss: 7.1310e-04

Epoch 38/50

240/240 ━━━━━━━━━━━━━━━━━━━━ 1s 5ms/step - loss: 1.7569e-05 - val\_loss: 3.5670e-04

Epoch 39/50

240/240 ━━━━━━━━━━━━━━━━━━━━ 1s 5ms/step - loss: 1.8924e-05 - val\_loss: 6.7853e-04

Epoch 40/50

240/240 ━━━━━━━━━━━━━━━━━━━━ 1s 5ms/step - loss: 2.0824e-05 - val\_loss: 4.3125e-04

Epoch 41/50

240/240 ━━━━━━━━━━━━━━━━━━━━ 1s 5ms/step - loss: 1.5176e-05 - val\_loss: 0.0018

Epoch 42/50

240/240 ━━━━━━━━━━━━━━━━━━━━ 1s 5ms/step - loss: 1.7451e-05 - val\_loss: 6.1244e-04

Epoch 43/50

240/240 ━━━━━━━━━━━━━━━━━━━━ 1s 5ms/step - loss: 1.3566e-05 - val\_loss: 4.8458e-04

Epoch 44/50

240/240 ━━━━━━━━━━━━━━━━━━━━ 1s 5ms/step - loss: 1.6551e-05 - val\_loss: 9.7485e-04

Epoch 45/50

240/240 ━━━━━━━━━━━━━━━━━━━━ 1s 5ms/step - loss: 1.7413e-05 - val\_loss: 4.6558e-04

Epoch 46/50

240/240 ━━━━━━━━━━━━━━━━━━━━ 1s 5ms/step - loss: 1.8673e-05 - val\_loss: 3.7124e-04

Epoch 47/50

240/240 ━━━━━━━━━━━━━━━━━━━━ 1s 5ms/step - loss: 1.8689e-05 - val\_loss: 8.8473e-04

Epoch 48/50

240/240 ━━━━━━━━━━━━━━━━━━━━ 1s 5ms/step - loss: 1.6848e-05 - val\_loss: 0.0012

Epoch 49/50

240/240 ━━━━━━━━━━━━━━━━━━━━ 1s 5ms/step - loss: 2.0852e-05 - val\_loss: 3.9647e-04

Epoch 50/50

240/240 ━━━━━━━━━━━━━━━━━━━━ 1s 5ms/step - loss: 1.5662e-05 - val\_loss: 9.1770e-04

30/30 ━━━━━━━━━━━━━━━━━━━━ 0s 8ms/step

LSTM Model Performance Metrics

Model Mean Absolute Error Mean Squared Error R² Score Training Time (s)

0 LSTM Deep Learning 0.024082 0.000918 0.968313 58.556184

Decision Tree Model Performance:

Mean Absolute Error (MAE): 1.00

Mean Squared Error (MSE): 1.60

R² Score: 1.00

ADF Statistic: 0.9659276819321151

p-value: 0.9938931143442359

Visualization Completed.

SHAP feature importance image saved as 'feature\_importance.png'

Model, features, and scaler saved.

XGBoost Model Performance:

Mean Absolute Error : 0.3117

Mean Squared Error : 0.3441

R² Score : 0.9998

1/1 ━━━━━━━━━━━━━━━━━━━━ 0s 25ms/step

1/1 ━━━━━━━━━━━━━━━━━━━━ 0s 21ms/step

1/1 ━━━━━━━━━━━━━━━━━━━━ 0s 21ms/step

1/1 ━━━━━━━━━━━━━━━━━━━━ 0s 22ms/step

1/1 ━━━━━━━━━━━━━━━━━━━━ 0s 18ms/step

1/1 ━━━━━━━━━━━━━━━━━━━━ 0s 18ms/step

1/1 ━━━━━━━━━━━━━━━━━━━━ 0s 10ms/step

1/1 ━━━━━━━━━━━━━━━━━━━━ 0s 10ms/step

1/1 ━━━━━━━━━━━━━━━━━━━━ 0s 18ms/step

1/1 ━━━━━━━━━━━━━━━━━━━━ 0s 19ms/step

1/1 ━━━━━━━━━━━━━━━━━━━━ 0s 14ms/step

1/1 ━━━━━━━━━━━━━━━━━━━━ 0s 19ms/step

1/1 ━━━━━━━━━━━━━━━━━━━━ 0s 18ms/step

1/1 ━━━━━━━━━━━━━━━━━━━━ 0s 21ms/step

1/1 ━━━━━━━━━━━━━━━━━━━━ 0s 18ms/step

1/1 ━━━━━━━━━━━━━━━━━━━━ 0s 18ms/step

1/1 ━━━━━━━━━━━━━━━━━━━━ 0s 17ms/step

1/1 ━━━━━━━━━━━━━━━━━━━━ 0s 17ms/step

1/1 ━━━━━━━━━━━━━━━━━━━━ 0s 19ms/step

1/1 ━━━━━━━━━━━━━━━━━━━━ 0s 15ms/step

1/1 ━━━━━━━━━━━━━━━━━━━━ 0s 19ms/step

1/1 ━━━━━━━━━━━━━━━━━━━━ 0s 18ms/step

1/1 ━━━━━━━━━━━━━━━━━━━━ 0s 21ms/step

1/1 ━━━━━━━━━━━━━━━━━━━━ 0s 18ms/step

1/1 ━━━━━━━━━━━━━━━━━━━━ 0s 15ms/step

1/1 ━━━━━━━━━━━━━━━━━━━━ 0s 10ms/step

1/1 ━━━━━━━━━━━━━━━━━━━━ 0s 13ms/step

1/1 ━━━━━━━━━━━━━━━━━━━━ 0s 16ms/step

1/1 ━━━━━━━━━━━━━━━━━━━━ 0s 22ms/step

1/1 ━━━━━━━━━━━━━━━━━━━━ 0s 18ms/step

[✓] Future prediction image saved as 'future\_prediction.png'

Simulated profitability image saved as 'profitability\_simulation.png'

--- Optimized Classification Results ---

Optimal Threshold: 0.153

Accuracy: 53.88%

precision recall f1-score support

Down 0.00 0.00 0.00 368

Up 0.54 1.00 0.70 430

accuracy 0.54 798

macro avg 0.27 0.50 0.35 798

weighted avg 0.29 0.54 0.38 798

Files generated:

- Feature importance.png

- selected features.pkl

- scaler.pkl

- xgb model.pkl

- future prediction.png

- profitability simulation.png

Date Range in DataFrame: 2004-10-01 00:00:00 to 2023-11-10 00:00:00

=== Last Date Validation ===

Training Period: 2004-10-01 to 2023-11-08

Test Date: 2023-11-10

Linear Regression:

Test Date: 2023-11-10

Actual Close: 134.06

Predicted Close: 133.18

Absolute Error: 0.88

Accuracy: 99.34%

Ridge Regression:

Test Date: 2023-11-10

Actual Close: 134.06

Predicted Close: 133.18

Absolute Error: 0.88

Accuracy: 99.34%

Random Forest:

Test Date: 2023-11-10

Actual Close: 134.06

Predicted Close: 132.77

Absolute Error: 1.29

Accuracy: 99.03%

Gradient Boosting:

Test Date: 2023-11-10

Actual Close: 134.06

Predicted Close: 132.76

Absolute Error: 1.30

Accuracy: 99.03%

Support Vector Regression:

Test Date: 2023-11-10

Actual Close: 134.06

Predicted Close: 57.75

Absolute Error: 76.31

Accuracy: 43.08%

=== Final Report ===

Model Actual Predicted Absolute Error Accuracy

0 Linear Regression 134.06 133.18 0.88 99.34

1 Ridge Regression 134.06 133.18 0.88 99.34

2 Random Forest 134.06 132.77 1.29 99.03

3 Gradient Boosting 134.06 132.76 1.30 99.03

4 Support Vector Regression 134.06 57.75 76.31 43.08