

```

!pip install --upgrade scikit-learn scikeras[tensorflow]
!pip install scikeras[tensorflow]

import numpy as np
import pandas as pd
from sklearn.preprocessing import MinMaxScaler
from sklearn.model_selection import train_test_split
from keras.models import Sequential
from keras.layers import LSTM, Dense, Dropout
import matplotlib.pyplot as plt

from google.colab import files

# Upload the dataset manually
uploaded = files.upload()

# Assuming the dataset is named 'stock_prices.csv'
df = pd.read_csv("HistoricalData_1736175120859.csv")
print(df.head())

data=df

# Assume 'Close' column contains the stock prices
target = df['Close/Last'].values

# Preprocess the data
data['Date'] = pd.to_datetime(data['Date'])
for col in ['Close/Last', 'Open', 'High', 'Low']:
    # Ensure the column is a string before applying .str.replace
    data[col] = data[col].astype(str).str.replace('$', '', regex=False).astype(float)

# Sort by date
data.sort_values(by='Date', inplace=True)

# Feature selection
features = ['Open', 'High', 'Low', 'Volume']
target = 'Close/Last'

# Normalize the data
scaler = MinMaxScaler()
# Ensure all columns are numeric and handle missing values
data[features + [target]] = data[features + [target]].apply(pd.to_numeric, errors='coerce')
data = data.dropna() # Drop rows with NaN values
scaled_data = scaler.fit_transform(data[features + [target]])

# Prepare the sequences for LSTM
def create_sequences(data, seq_length):
    X, y = [], []
    for i in range(len(data) - seq_length):
        X.append(data[i:i + seq_length, :-1])
        y.append(data[i + seq_length, -1])
    return np.array(X), np.array(y)

sequence_length = 10 # Hyperparameter
data_sequences, target_sequences = create_sequences(scaled_data, sequence_length)

# Split the data into training and testing sets
X_train, X_test, y_train, y_test = train_test_split(data_sequences, target_sequences, test_size=0.2, random_state=42)

# Define the LSTM model
def build_lstm(optimizer='adam', dropout_rate=0.2, units=50):
    model = Sequential()
    model.add(LSTM(units=units, return_sequences=True, input_shape=(X_train.shape[1], X_train.shape[2])))
    model.add(Dropout(dropout_rate))
    model.add(LSTM(units=units))

```

```

model.add(Dense(units=units))
model.add(Dropout(dropout_rate))
model.add(Dense(1))
model.compile(optimizer=optimizer, loss='mean_squared_error')
return model

#unit means neurons

# Hyperparameter tuning manually
param_grid = {
    'optimizer': ['adam', 'rmsprop'],
    'dropout_rate': [0.2, 0.3],
    'units': [50, 100],
    'batch_size': [16, 32],
    'epochs': [50, 100]
}

best_loss = float('inf')
best_params = None
best_model = None

for optimizer in param_grid['optimizer']:
    for dropout_rate in param_grid['dropout_rate']:
        for units in param_grid['units']:
            for batch_size in param_grid['batch_size']:
                for epochs in param_grid['epochs']:
                    print(f"Training with optimizer={optimizer}, dropout_rate={dropout_rate}, units={units}, batch_s:
model = build_lstm(optimizer=optimizer, dropout_rate=dropout_rate, units=units)
model.fit(X_train, y_train, epochs=epochs, batch_size=batch_size, verbose=0)
loss = model.evaluate(X_test, y_test, verbose=0)
print(f"Loss: {loss}")

                    if loss < best_loss:
                        best_loss = loss
                        best_params = {
                            'optimizer': optimizer,
                            'dropout_rate': dropout_rate,
                            'units': units,
                            'batch_size': batch_size,
                            'epochs': epochs
                        }
                        best_model = model

print("Best parameters found: ", best_params)

# Predict
predicted = best_model.predict(X_test)

# Inverse transform to get actual values
# Create a dummy array to match the scaler's input shape for inverse transformation
dummy_features = np.zeros((predicted.shape[0], len(features))) # Matching the feature columns
dummy_data = np.hstack((dummy_features, predicted.reshape(-1, 1)))

# Inverse transform to get actual predicted values
predicted_actual = scaler.inverse_transform(dummy_data)[:, -1]

# Similarly, process the test target values
dummy_test_features = np.zeros((y_test.shape[0], len(features)))
dummy_test_data = np.hstack((dummy_test_features, y_test.reshape(-1, 1)))
y_test_actual = scaler.inverse_transform(dummy_test_data)[:, -1]

```

3/5

Requirement already satisfied: tensorboard-data-server<0.8.0,>=0.7.0 in /usr/local/lib/python3.10/dist-packages (from tensorboard<2.18,>=2.17->tensorflow>=2.1)

Requirement already satisfied: Werkzeug>=1.0.1 in /usr/local/lib/python3.10/dist-packages (from tensorboard<2.18,>=2.17->tensorflow>=2.1)

Requirement already satisfied: Markdown-It-Py>=2.2.0 in /usr/local/lib/python3.10/dist-packages (from rich->keras>=3.2.0->scikeras[tensorflow])

Requirement already satisfied: pygments<3.0.0,>=2.13.0 in /usr/local/lib/python3.10/dist-packages (from rich->keras>=3.2.0->scikeras[tensorflow])

Requirement already satisfied: mdurl<=0.1 in /usr/local/lib/python3.10/dist-packages (from Markdown-It-Py>=2.2.0->rich->keras>=3.2.0->scikeras[tensorflow])

Requirement already satisfied: MarkupSafe>=2.1.1 in /usr/local/lib/python3.10/dist-packages (from Werkzeug>=1.0.1->tensorboard<2.18,>=2.17->tensorflow>=2.1)

Choose Files HistoricalData...5120859.csv

• **HistoricalData_1736175120859.csv**(text/csv) - 1056 bytes, last modified: 1/6/2025 - 100% done

Saving HistoricalData_1736175120859.csv to HistoricalData_1736175120859.csv

	Date Close/Last	Volume	Open	High	Low
0	01/03/2025	\$243.36	40244110	\$243.36	\$244.18
1	01/02/2025	\$243.85	55740730	\$248.93	\$249.10
2	12/31/2024	\$250.42	39480720	\$252.44	\$253.28
3	12/30/2024	\$252.20	35557540	\$252.23	\$253.50
4	12/27/2024	\$255.59	42355320	\$257.83	\$258.70

Training with optimizer=adam, dropout_rate=0.2, units=50, batch_size=16, epochs=50

/usr/local/lib/python3.10/dist-packages/keras/src/layers/rnn/rnn.py:204: UserWarning: Do not pass an `input_shape`/`input_dim` argument to `super().__init__()` (`**kwargs`)

Loss: 0.20043589174747467

Training with optimizer=adam, dropout_rate=0.2, units=50, batch_size=16, epochs=100

Loss: 0.06393007934093475

Training with optimizer=adam, dropout_rate=0.2, units=50, batch_size=32, epochs=50

Loss: 0.1561945676803589

Training with optimizer=adam, dropout_rate=0.2, units=50, batch_size=32, epochs=100

Loss: 0.011438247747719288

Training with optimizer=adam, dropout_rate=0.2, units=100, batch_size=16, epochs=50

WARNING:tensorflow:5 out of the last 5 calls to <function TensorFlowTrainer.make_test_function.<locals>.one_step_on_iterator at 0x7b2f2b

Loss: 0.0756458193063736

Training with optimizer=adam, dropout_rate=0.2, units=100, batch_size=16, epochs=100

WARNING:tensorflow:6 out of the last 6 calls to <function TensorFlowTrainer.make_test_function.<locals>.one_step_on_iterator at 0x7b2f2a

Loss: 0.021798714995384216

Training with optimizer=adam, dropout_rate=0.2, units=100, batch_size=32, epochs=50

Loss: 0.04470527544617653

Training with optimizer=adam, dropout_rate=0.2, units=100, batch_size=32, epochs=100

Loss: 0.01936744712293148

Training with optimizer=adam, dropout_rate=0.3, units=50, batch_size=16, epochs=50

Loss: 0.2128603458404541

Training with optimizer=adam, dropout_rate=0.3, units=50, batch_size=16, epochs=100

Loss: 0.0332423560321331

Training with optimizer=adam, dropout_rate=0.3, units=50, batch_size=32, epochs=50

Loss: 0.18422016501426697

Training with optimizer=adam, dropout_rate=0.3, units=50, batch_size=32, epochs=100

Loss: 0.013124839402735233

Training with optimizer=adam, dropout_rate=0.3, units=100, batch_size=16, epochs=50

Loss: 0.12137708067893982

Training with optimizer=adam, dropout_rate=0.3, units=100, batch_size=16, epochs=100

Loss: 0.007250993046909571

Training with optimizer=adam, dropout_rate=0.3, units=100, batch_size=32, epochs=50

Loss: 0.09030996263027191

Training with optimizer=adam, dropout_rate=0.3, units=100, batch_size=32, epochs=100

Loss: 0.022385787218809128

Training with optimizer=rmsprop, dropout_rate=0.2, units=50, batch_size=16, epochs=50

Loss: 0.061324313282966614

Training with optimizer=rmsprop, dropout_rate=0.2, units=50, batch_size=16, epochs=100

Loss: 0.03329877182841301

Training with optimizer=rmsprop, dropout_rate=0.2, units=50, batch_size=32, epochs=50

Loss: 0.09121855348348618

Training with optimizer=rmsprop, dropout_rate=0.2, units=50, batch_size=32, epochs=100

Loss: 0.034052763134241104

Training with optimizer=rmsprop, dropout_rate=0.2, units=100, batch_size=16, epochs=50

Loss: 0.03172944858670235

Training with optimizer=rmsprop, dropout_rate=0.2, units=100, batch_size=16, epochs=100

Loss: 0.06026472896337509

Training with optimizer=rmsprop, dropout_rate=0.2, units=100, batch_size=32, epochs=50

Loss: 0.01985284872353077

Training with optimizer=rmsprop, dropout_rate=0.2, units=100, batch_size=32, epochs=100

Loss: 0.003987314645200968

Training with optimizer=rmsprop, dropout_rate=0.3, units=50, batch_size=16, epochs=50

Loss: 0.1313382238149643

Training with optimizer=rmsprop, dropout_rate=0.3, units=50, batch_size=16, epochs=100

Loss: 0.0412243977189064

Training with optimizer=rmsprop, dropout_rate=0.3, units=50, batch_size=32, epochs=50

Loss: 0.1383463740348816

Training with optimizer=rmsprop, dropout_rate=0.3, units=50, batch_size=32, epochs=100

Loss: 0.1530868113040924

Training with optimizer=rmsprop, dropout_rate=0.3, units=100, batch_size=16, epochs=50

Loss: 0.02452634461224079

Training with optimizer=rmsprop, dropout_rate=0.3, units=100, batch_size=16, epochs=100

Loss: 0.0005738771869800985

Training with optimizer=rmsprop, dropout_rate=0.3, units=100, batch_size=32, epochs=50

Loss: 0.21933765709400177

Training with optimizer=rmsprop, dropout_rate=0.3, units=100, batch_size=32, epochs=100

Loss: 0.0764743834733963

Best parameters found: {'optimizer': 'rmsprop', 'dropout_rate': 0.3, 'units': 100, 'batch_size': 16, 'epochs': 100}

1/1 — 0s 251ms/step

