In [1]: In [2]:	stock price perdiction by using LSTM method import pandas as pd import pandas as pd	
Out[2]:	df = pd.read_csv('MSFT.csv') df Date Open High Low Close Adj Close Volume	
	<th></th>	
Out[3]:	df = df[['Date', 'Close']] df Date Close 0 1986-03-13 0.097222 1 1986-03-14 0.100694 2 1986-03-17 0.102431 3 1986-03-18 0.099826 4 1986-03-19 0.098090	
In [4]:	9405 2023-07-10 331.829987 9406 2023-07-11 332.470001 9407 2023-07-12 337.200012 9408 2023-07-13 342.660004 9409 2023-07-14 345.239990 9410 rows × 2 columns df['Date']	
Out[4]: In [5]:	0	
Out[5]: In [6]:	<pre>def str_to_datetime(s): split = s.split('-') year, month, day = int(split[0]), int(split[2]) return datetime.datetime(year=year, month=month, day=day) datetime_object = str_to_datetime('1986-03-19') datetime_object datetime.datetime(1986, 3, 19, 0, 0)</pre> df	
Out[6]:	Date Close 0 1986-03-13 0.097222 1 1986-03-14 0.100694 2 1986-03-17 0.102431 3 1986-03-18 0.099826 4 1986-03-19 0.098090 9405 2023-07-10 331.829987 9406 2023-07-11 332.470001	
In [7]:	9407 2023-07-12 337.200012 9408 2023-07-13 342.660004 9409 2023-07-14 345.239990 9410 rows × 2 columns df['Date'] = df['Date'].apply(str_to_datetime) df['Date'] C:\Users\durga prasad\AppData\Local\Temp\ipykernel_97680\2565755782.py:1: SettingWithCopyWarning: A value is trying to be set on a copy of a slice from a DataFrame. Try using loc_frow indexer coll indexer value instead	
Out[7]:	Try using .loc[row_indexer, col_indexer] = value instead See the caveats in the documentation: https://pandas.pydata.org/pandas-docs/stable/user_guide/indexing.html#returning-a-view-versus-a-copy df['Date'] = df['Date'].apply(str_to_datetime) 1	
In [8]: Out[8]:	Name: Date, Length: 9410, dtype: datetime64[ns] df.index = df.pop('Date') df Close Date 1986-03-13 0.097222 1986-03-14 0.100694 1986-03-17 0.102431 1986-03-18 0.099826	
	1986-03-19 0.098090 2023-07-10 331.829987 2023-07-11 332.470001 2023-07-12 337.200012 2023-07-13 342.660004 2023-07-14 345.239990 9410 rows × 1 columns	
In [9]: Out[9]:	<pre>import matplotlib.pyplot as plt plt.plot(df.index, df['Close']) [<matplotlib.lines.line2d 0x2029bd19400="" at="">] 350 - 300 - 250 -</matplotlib.lines.line2d></pre>	
	200 - 150 - 100 - 50 -	
In [10]:	<pre>import numpy as np def df_to_windowed_df(dataframe, first_date_str, last_date_str, n=3): first_date = str_to_datetime(first_date_str) last_date = str_to_datetime(last_date_str) target_date = first_date dates = [] X, Y = [], []</pre>	
	<pre>last_time = False while True: df_subset = dataframe.loc[:target_date].tail(n+1) if len(df_subset) != n+1: print(f'Error: Window of size {n} is too large for date {target_date}') return values = df_subset['Close'].to_numpy() x, y = values[:-1], values[-1] dates.append(target_date) X.append(x) Y.append(y)</pre>	
	<pre>next_week = dataframe.loc[target_date:target_date+datetime.timedelta(days=7)] next_datetime_str = str(next_week.head(2).tail(1).index.values[0]) next_date_str = next_datetime_str.split('T')[0] year_month_day = next_date_str.split('-') year, month, day = year_month_day next_date = datetime.datetime(day=int(day), month=int(month), year=int(year)) if last_time: break target_date = next_date if target_date == last_date: last_time = True</pre>	
	<pre>ret_df = pd.DataFrame({}) ret_df['Target Date'] = dates X = np.array(X) for i in range(0, n): X[:, i] ret_df[f'Target-{n-i}'] = X[:, i] ret_df[['Target'] = Y return ret_df # Start day second time around: '2021-03-25'</pre>	
Out[10]:	windowed_df = df_to_windowed_df(df,	
	4 2021-03-31 236.479996 235.240005 231.850006 235.770004	
In [11]:	<pre>def windowed_df_to_date_X_y(windowed_dataframe): df_as_np = windowed_dataframe.to_numpy() dates = df_as_np[:, 0] middle_matrix = df_as_np[:, 1:-1] X = middle_matrix.reshape((len(dates), middle_matrix.shape[1], 1)) Y = df_as_np[:, -1] return dates, X.astype(np.float32), Y.astype(np.float32) dates, X, y = windowed_df_to_date_X_y(windowed_df) dates.shape, X.shape, y.shape</pre>	
Out[11]: In [12]:	<pre>((252,), (252, 3, 1), (252,)) q_80 = int(len(dates) * .8) q_90 = int(len(dates) * .9) dates_train, X_train, y_train = dates[:q_80], X[:q_80], y[:q_80] dates_val, X_val, y_val = dates[q_80:q_90], X[q_80:q_90], y[q_80:q_90] dates_test, X_test, y_test = dates[q_90:], X[q_90:], y[q_90:] plt.plot(dates_train, y_train) plt.plot(dates_val, y_val)</pre>	
Out[12]:	plt.legend(['Train', 'Validation', 'Test']) <matplotlib.legend.legend 0x2029ef7f100="" at=""> 340 Train Validation Test 320 Test</matplotlib.legend.legend>	
	280 - 240 - 240 - 240 - 240	
In [13]:	<pre>from tensorflow.keras.models import Sequential from tensorflow.keras.optimizers import Adam from tensorflow.keras import layers model = Sequential([layers.Input((3, 1)),</pre>	
	model.fit(X_train, y_train, validation_data=(X_val, y_val), epochs=100) Epoch 1/100 7/7 [===================================	
	Epoch 6/100 7/7 [===================================	
	Epoch 13/100 7/7 [===================================	
	7/7 [===================================	
	7/7 [===================================	
	Epoch 35/100 7/7 [===================================	
	7/7 [===================================	
	7/7 [===================================	
	Epoch 57/100 7/7 [===================================	
	7/7 [===================================	
	Epoch 72/100 7/7 [===================================	
	Epoch 79/100 7/7 [===================================	
	Epoch 86/100 7/7 [===================================	
	7/7 [===================================	
Out[13]: In [14]: Out[14]:	<pre>/// [==================================</pre>	
	320 - 300 - 280 - 260 - 300 -	
In [15]:	240 - 2021-02021-02021-02021-02021-02021-12021-12021-12022-01 val_predictions = model.predict(X_val).flatten() plt.plot(dates_val, val_predictions) plt.plot(dates_val, y_val) plt.legend(['Validation Predictions', 'Validation Observations']) 1/1 [===================================	
Out[15]:	1/1 [========] - 0s 31ms/step <matplotlib.legend.legend 0x202ad7a9e80="" at=""> Validation Predictions Validation Observations 310 - 305 - 305 - 31ms/step</matplotlib.legend.legend>	
In [16]:	300 - 295 - 290 - 290 - 2022-0PQ2-0PQ	
<pre>In [16]: Out[16]:</pre>	test_predictions = model.predict(X_test).flatten() plt.plot(dates_test, test_predictions) plt.plot(dates_test, y_test) plt.legend(['Testing Predictions', 'Testing Observations']) 1/1 [===================================	
	290 - 285 - 280 - Testing Predictions	
In [17]:	plt.plot(dates_train, train_predictions) plt.plot(dates_train, y_train) plt.plot(dates_val, val_predictions) plt.plot(dates_val, val_predictions) plt.plot(dates_test, test_predictions) plt.plot(dates_test, test_predictions) plt.plot(dates_test, y_test) plt.legend(['Training Predictions',	
Out[17]:	'Validation Observations', 'Testing Predictions', 'Testing Observations']) <matplotlib.legend.legend 0x202ad72cdf0="" at=""> 340 - 320 -</matplotlib.legend.legend>	
	Training Predictions Training Observations Validation Predictions Validation Observations Testing Observations Testing Observations Testing Observations Testing Observations	
In [18]:	<pre>from copy import deepcopy recursive_predictions = [] recursive_dates = np.concatenate([dates_val, dates_test]) for target_date in recursive_dates: last_window = deepcopy(X_train[-1]) next_prediction = model.predict(np.array([last_window])).flatten() recursive_predictions.append(next_prediction) last_window[-1] = next_prediction 1/1 [===================================</pre>	
	1/1 [===================================	
	1/1 [===================================	
In [19]:	1/1 [===================================	
Out[19]:		
	320 - 300 - 280 - Training Predictions Training Observations Validation Predictions Validation Predictions Validation Observations Training Durdictions	
In []:	Validation Observations — Testing Predictions — Testing Observations — Recursive Predictions 2021-05 2021-07 2021-09 2021-11 2022-01 2022-03	