Quinnel Soft Company

Title: Time Series Prediction using Deep Learning

Name : Pavuluru Charan Kumar
Department: B.E ECE
Mail Id: mailto:charan.ece.saveetha@gmail.com
Course Work: Github

Date : 29/05/2024

Course Outline

1. Title Page
    a. Project Title
    b. Your Name
    c. Date
2. Table of Contents
3. Introduction
    a. Background
    b. Objective
4. Dataset Description
    a. Data Source
    b. Data Features
5. Data Analysis
    a. Visualizations
6. Data Preprocessing
    a. Handling Missing Values
    b. Feature Scaling
    c. Sequence Creation
7. Model Selection
    a. Choice of Model
8. Model Training
    a. Training Process
9. Model Evaluation
    a. Evaluation Metrics
    b. Results
10. Conclusion
11. References

Table of Contents

Summary :

This report presents the development and evaluation of a deep learning model for predicting retail sales based on historical time series data. The model is designed to forecast future sales with high accuracy using LSTM (Long Short-Term Memory) neural networks implemented in TensorFlow.

Introduction :

Objective: Develop a predictive model using a deep learning framework (TensorFlow or PyTorch) to forecast future sales from historical time series data. You are to select a deep learning algorithm of your choice to accurately predict future time steps.

Dataset Description :

Data Source:

The dataset consists of monthly sales data from a major retail chain for the years 2015 to 2020.

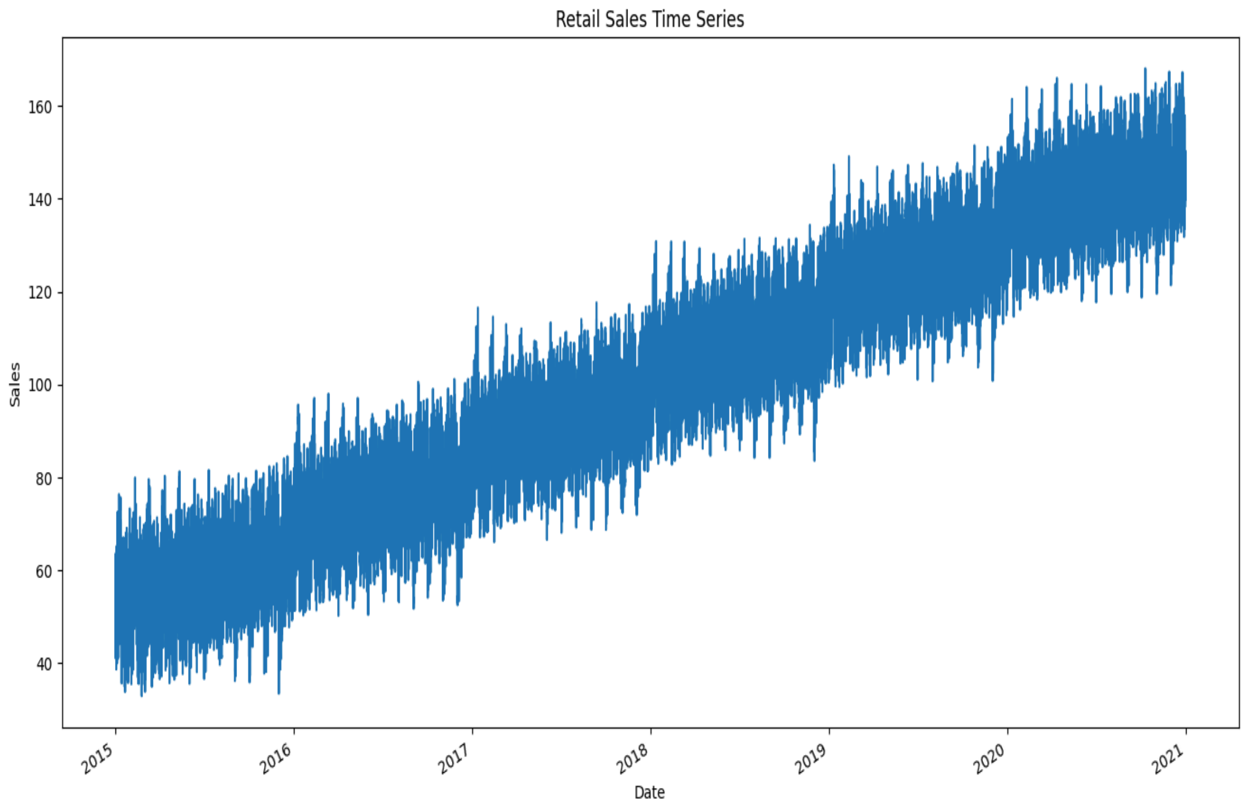
Data Features:

Month: Month of the sales data record.

Sales: Total sales value in USD.

Data Analysis:

Visualizations: Retail Sales Time Series



Data Preprocessing:

Handling Missing Values:

# Display the column names
print("Column names in the dataset:")
print(df.columns)
# Display the first few rows of the dataset
print("\nFirst few rows of the dataset:")
print(df.head())

Feature Scaling :

# Strip any leading/trailing spaces from column names
df.columns = df.columns.str.strip()

Sequence Creation:

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

seq_length = 10
X_train, y_train = create_sequences(train['Sales'].values, seq_length)
X_test, y_test = create_sequences(test['Sales'].values, seq_length)
```

Model Selection:

Choice of Model: LSTM was chosen for this task.

Model Training:

```
# Defining the model
model = Sequential([
LSTM(10, return_sequences=True, input_shape=(seq_length, 1)),
Dropout(0.2),
LSTM(10, return_sequences=False),
Dropout(0.2),
Dense(1)
])

# Compile the model
model.compile(optimizer='Adam', loss='mean_squared_error')
```

```
# Train the model
history = model.fit(X_train, y_train, epochs=5, batch_size=32,
validation_split=0.2)

# Make predictions
predictions = model.predict(X_test)
```

Model Evaluation:

Evaluation Metrics:

- Root Mean Square Error (RMSE)
- Mean Absolute Error (MAE)
- Coefficient of Determination (R2)

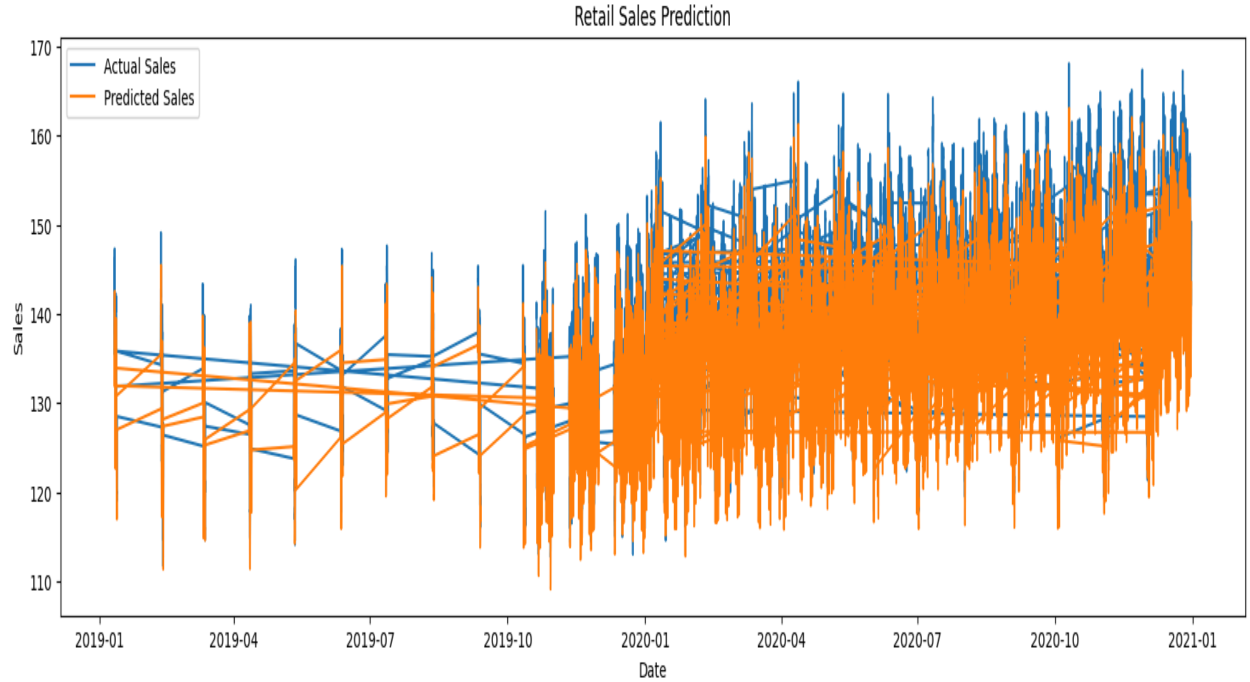
Results:

RMSE: 4.897721874716446

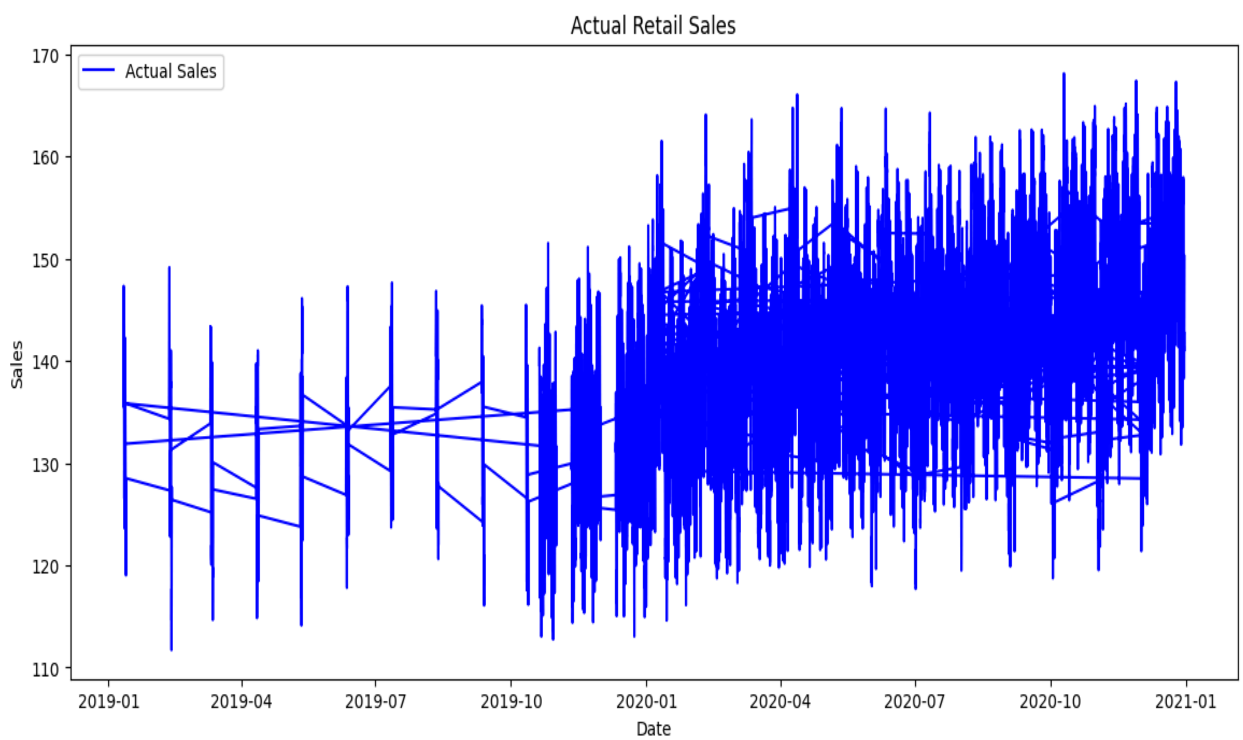
MAE: 4.13038449847004

R^2: 0.7554807950905527

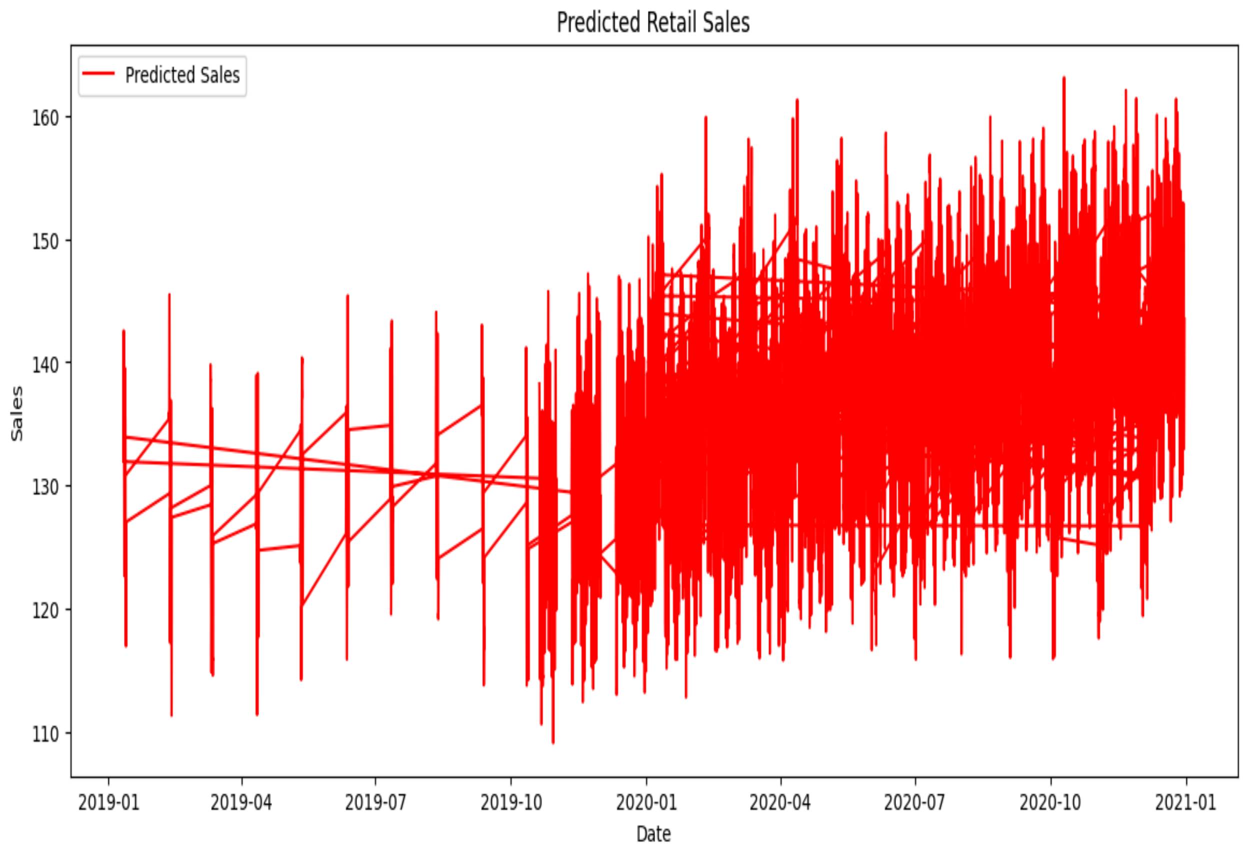
Plots: Actual Sales and Predicted Sales



Actual Sales:



Predicted Sales:



References :

Books and Articles on Time Series Forecasting:

Hyndman, R. J., & Athanasopoulos, G. (2018). Forecasting: principles and practice. OTexts. https://otexts.com/fpp3/

Box, G. E. P., Jenkins, G. M., Reinsel, G. C., & Ljung, G. M. (2015). Time Series Analysis: Forecasting and Control. Wiley.

Chatfield, C. (2003). The Analysis of Time Series: An Introduction. Chapman and Hall/CRC.

Deep Learning and LSTM Models:

Hochreiter, S., & Schmidhuber, J. (1997). Long Short-Term Memory. Neural Computation, 9(8), 1735–1780. https://doi.org/10.1162/neco.1997.9.8.1735

Brownlee, J. (2018). Deep Learning for Time Series Forecasting: Predict the Future with MLPs, CNNs and LSTMs in Python. Machine Learning Mastery.

Conclusion :

I conclude that python scripts for preprocessing, modeling, training, and evaluation. A detailed report outlining your choice of model, methodologies used, and key findings are delivered and outcomes of graphs were plotted and all these materials are well documented and submitted under the github repository.