**Petrol Price Prediction: Time Series Analysis with LSTM**

**Overview**

This document outlines the steps involved in predicting petrol prices in Maharashtra using a time series forecasting model. The model is built using Long Short-Term Memory (LSTM) neural networks to capture temporal patterns in historical petrol prices.

**Data Loading and Initial Exploration**

**Load the Dataset:**

* + The loaded petrol price dataset is in CSV file format and collected from Kaggle. Here is the link.

<https://www.kaggle.com/datasets/abhijitdahatonde/top-15-indian-states-petrol-prices-2017-2022>

**Initial Data Exploration:**

* + Display the first few rows of the dataset to understand its structure.
  + Select relevant columns ('date' and 'Maharashtra') for analysis.

**Date Formatting:**

* + Convert the 'date' column to a datetime object using the specified format ('%Y\_%b').
  + Format the 'date' column to 'YYYY-MM'.
  + Sort the dataset based on the 'date' column.

**Indexing and Visualization:**

* + Set the 'date' column as the index for time series analysis.
  + Visualize the entire dataset to observe the trend in petrol prices over time.

**Time Series Analysis**

**Average Price Calculation:**

* + Compute the average petrol price for each year (2017-2022) for initial insights.

**Dataset Visualization:**

* + Plot the complete time series dataset to identify any observable patterns.
  + Plot the average petrol prices over the specified years for a holistic view.

**Data Preprocessing**

**Train-Test Split:**

* + Split the dataset into training and testing sets. (First 55 data points for training, the rest for testing)

**MinMax Scaling:**

* + Utilize MinMaxScaler to scale the training and testing datasets for compatibility with the LSTM model.

**LSTM Model Building**

**Time Series Generator:**

* + Create a TimeSeriesGenerator for training the LSTM model with a sequence length of 12.

**LSTM Model Architecture:**

* + Design an LSTM model with two layers of 500 units each and a Dense output layer.

**Model Compilation:**

* + Compile the model using the Adam optimizer and Mean Squared Error (MSE) as the loss function.

**Model Training:**

* + Train the LSTM model on the training data for 120 epochs.

**Loss Visualization:**

* + Plot the loss curve to assess model convergence during training.

**Prediction Preparation:**

* + Prepare the last batch of training data for predicting future petrol prices.

**Prediction Loop:**

* + Iterate through the test set, making predictions using the LSTM model and updating the current batch.

**Inverse Scaling:**

* + Inverse scale the predicted petrol prices to obtain real-world values.

**Visualization of Prediction:**

* + Plot the actual vs predicted petrol prices for the test set.

**Model Evaluation**

**Root Mean Squared Error (RMSE):**

* + Calculate the RMSE to quantify the prediction accuracy.

**Conclusion:**

* + Summarize the findings, model performance, and potential areas for improvement.