Walmart Sales Prediction Model

This project demonstrates how to build, train, and deploy a linear regression model to predict weekly sales for Walmart stores based on features such as store number, holiday flag, temperature, fuel price, CPI, and unemployment rate. It also includes functionality for loading new data, making predictions, and saving the results.

Project Structure

- **Training and Model Saving**: The script builds a linear regression model using a dataset containing features related to Walmart stores and sales data. It trains the model, saves it as a pickle file, and demonstrates how to load the model to make predictions.
- **Making Predictions**: The second part of the script demonstrates loading the saved model, making predictions for a new dataset, and saving the predictions to a CSV file.

Files

- walmart_sales_model.pkl: Pickle file containing the trained linear regression model.
- combined_script.pkl: Pickle file containing the entire script for reuse.
- new_walmart_sales_with_predictions.csv: Output file with new sample data and predicted sales.

How to Use

1. Dataset

The dataset should contain the following features:

- Store: The store number.
- Holiday_Flag: A binary flag (1 = holiday week, 0 = non-holiday week).
- Temperature: Average temperature in Fahrenheit.
- Fuel_Price: Average fuel price.
- CPI: Consumer Price Index.
- Unemployment: Unemployment rate.

The target variable is Weekly_Sales, which represents the sales value to predict.

2. Model Training

To train the model:

1. Load the dataset using pandas.
2. Exclude unnecessary features and keep the relevant ones (as defined in the script).
3. Split the data into training and testing sets (80% training, 20% testing).

4. Train a linear regression model on the training data.
5. Save the trained model to a pickle file for later use.

3. Making Predictions with New Data

You can use the saved model to make predictions on new data:

1. Load the new data into a pandas DataFrame.
2. Load the saved model (walmart_sales_model.pkl).
3. Use the predict() method to generate predictions.
4. Save the predictions alongside the new data in a CSV file.

4. Example

Here's an example of how to structure your new data for predictions:

python

Copy code

```
new_data = {
   'Store': [1, 2, 3],
   'Holiday_Flag': [0, 1, 0],
   'Temperature': [85.0, 80.0, 78.0],
   'Fuel_Price': [2.57, 2.67, 2.77],
   'CPI': [211.0, 215.5, 210.7],
   'Unemployment': [6.5, 6.7, 6.9]
}
```

5. Output

The predictions will be saved in a CSV file (new_walmart_sales_with_predictions.csv), with the new data and the predicted Weekly_Sales.

6. Dependencies

- Python 3.x
- pandas
- scikit-learn
- numpy
- pickle