Walmart Sales Forecasting App

This Streamlit application demonstrates the analysis and forecasting of Walmart sales data using machine learning (Random Forest) and time series (ARIMA) techniques.

Features

- **Data Upload & Processing**: Upload and process the required CSV files for analysis
- **Data Exploration**: Interactive visualizations of store types, holiday impact, time trends, and external factors
- **Model Training**: Train Random Forest and Time Series models with customizable parameters
- **Predictions & Evaluation**: Generate predictions and evaluate model performance using WMAE

Setup Instructions

1. Installation

First, clone the repository (if applicable) or create a new directory and add the provided files:

```
mkdir walmart_app
cd walmart_app
```

Create the following files in this directory:

- `app.py` : Main Streamlit application
- `model.py` : Model training and prediction logic
- `utils.py` : Helper functions for data processing
- `requirements.txt` : Application dependencies

2. Create a Virtual Environment (Recommended)

```
# On Windows
python -m venv venv
venv\Scripts\activate

# On macOS/Linux
python3 -m venv venv
source venv/bin/activate
```

3. Install Dependencies

```
pip install -r requirements.txt
```

4. Run the Application

```
streamlit run app.py
```

The application will launch in your default web browser at http://localhost:8501

5. Using the Application

1. **Data Upload**:
    - Upload the three required CSV files: `stores.csv`, `train.csv`, and `features.csv`
    - Click "Process Data" to merge and clean the datasets
2. **Data Exploration**:
    - Navigate to the "Data Exploration" page to view visualizations
    - Explore different analysis aspects using the tabs
3. **Model Training**:
    - Set model parameters for Random Forest and Time Series models
    - Click "Train Models" to train the selected models
4. **Predictions & Evaluation**:
    - Generate predictions and evaluate model performance
    - Compare results from different models

- Download predictions as CSV

Data Files

The application requires three CSV files:

1. **stores.csv**: Contains information about the Walmart stores
   - Columns: Store, Type, Size
2. **train.csv**: Contains historical sales data
   - Columns: Store, Dept, Date, Weekly_Sales, IsHoliday
3. **features.csv**: Contains additional features
   - Columns: Store, Date, Temperature, Fuel_Price, MarkDown1-5, CPI, Unemployment, IsHoliday

Sample Workflow

1. Upload the three required CSV files on the "Data Upload" page
2. Process the data by clicking the "Process Data" button
3. Explore the data through visualizations on the "Data Exploration" page
4. Set model parameters and train models on the "Model Training" page
5. Generate predictions and evaluate models on the "Predictions & Evaluation" page
6. Download predictions for further analysis

Running Your Walmart Sales Forecasting App

Here's a step-by-step guide to set up and run the application:

Step 1: Create Project Structure

First, create a directory for your application and set up the file structure:

```
walmart_app/
├── app.py                  # Main Streamlit app
├── model.py                # Model training & prediction logic
├── utils.py                # Helper functions
├── requirements.txt        # Dependencies
└── sample_data/            # Optional folder for sample data
```

Step 2: Copy Code Files

Copy the code from each artifact into the respective files:

- Copy the `app.py` content into your app.py file
- Copy the `model.py` content into your model.py file
- Copy the `utils.py` (including the continuation) into your utils.py file
- Copy the `requirements.txt` content into your requirements.txt file

Step 3: Set Up Virtual Environment

Create and activate a virtual environment to isolate your dependencies:

```
# On Windows
python -m venv venv
venv\Scripts\activate

# On macOS/Linux
python3 -m venv venv
source venv/bin/activate
```

Step 4: Install Dependencies

Install all required packages:

```
pip install -r requirements.txt
```

Step 5: Run the Application

Start the Streamlit application:

```
streamlit run app.py
```

The application will automatically open in your default web browser at

Local URL: http://localhost:8501

Network URL: http://139.13.79.22:8501

Step 6: Use the Application

1. Upload your three CSV files (`stores.csv` , `train.csv` , and `features.csv`) on the Data Upload page
2. Process the data by clicking the "Process Data" button
3. Navigate through the tabs to explore visualizations and train models
4. Generate and evaluate predictions
5. Download results for further analysis

Key Features of Your Application

1. **Interactive UI**: Streamlit provides a clean interface with tabs, file uploaders, and controls
2. **Comprehensive Analysis**: Includes all the data exploration visualizations from your original code
3. **Multiple Models**: Implements both Random Forest and Time Series (ARIMA) models
4. **Custom Parameters**: Allows adjusting model parameters without code changes
5. **Performance Evaluation**: Calculates and displays WMAE (Weighted Mean Absolute Error)
6. **Data Download**: Provides downloadable prediction results

This application effectively showcases your Walmart sales analysis work in an interactive format that can be shared with others. The code structure follows best practices with separate modules for different functionality, making it easy to maintain and extend.