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:

```bash
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)

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

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

3. Install Dependencies

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

4. Run the Application

```bash
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