

Project 2: Walmart Store Sales Forecasting

Contents

Overview	1
Datasets	1
Objective	2

Overview

Given historical sales data from 45 Walmart stores spread across different regions, your task is to predict the future weekly sales for every department in each store.

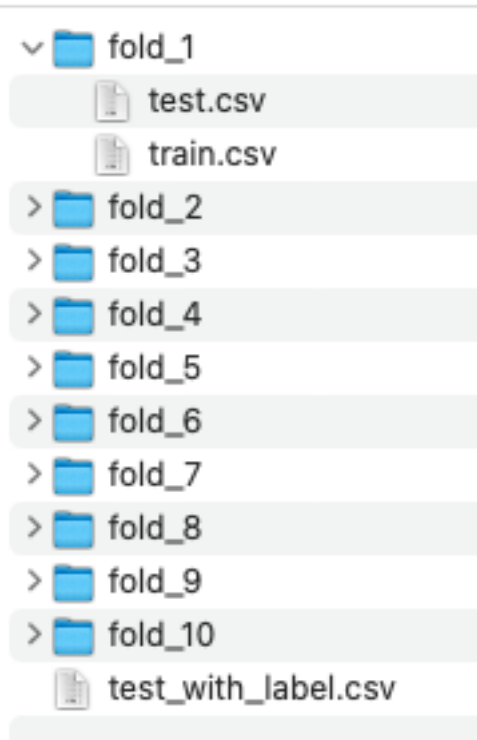
The dataset is from <https://www.kaggle.com/c/walmart-recruiting-store-sales-forecasting>.

Please Note::

1. We will **only** be using the **training data** in this project.
2. Kaggle's competition includes an additional CSV file with features such as temperature, fuel price, CPI, etc. This feature set is **not** utilized in our project.

Datasets

Download the dataset from the link: [proj2.zip]. Extract the zip to find 10 folders and a file named `test_with_label.csv`. In each folder, there are two files: `train.csv` and `test.csv`.



- **train.csv** contains 5 columns (“Store”, “Dept”, “Date”, “Weekly_Sales”, “IsHoliday”) and ranges
 - from 2010-02 (February 2010) to 2011-02 (February 2011) in fold_1,
 - from 2010-02 to 2011-04 in fold_2,
 - from 2010-02 to 2011-06 in fold_3,
 -
 - from 2010-02 to 2012-08 in fold_10.
- **test.csv** contains 4 columns (“Store”, “Dept”, “Date”, “IsHoliday”) and ranges
 - from 2011-03 to 2011-04 in fold_1,
 - from 2011-05 to 2011-06 in fold_2,
 - from 2011-07 to 2011-08 in fold_3,
 -
 - from 2012-09 to 2012-10 in fold_10.
- **test_with_label.csv** is formatted similarly to **train.csv** and ranges from 2011-03 to 2012-10.

Objective

Predict the weekly sales for the subsequent two months for every combination of Store, Dept, and Date in **test.csv** using the historical data from **train.csv**.