From Patterns to Profits: Walmart Sales Forecasting

Predictive Analytics Team Project

Chen-Ju Wang
Shang-Chien Wang
Ya-Chin Hsu
Yi Hsiang Yen

Business Context

Walmart drives over $1.27B in U.S. daily sales and serves 37M customers each day, highlighting the massive potential for impact through smarter inventory planning.

By predicting next month's item-level sales, Walmart can better anticipate demand shifts, align inventory with promotions and seasonal trends, and reduce costly stockouts or overstocks.

Ultimately protecting revenue and enhancing the customer experience.

37M
daily Customer

1.27B
Sales Per day

data source : Walmart 2025 Annual Report

Forecasting item-level sales enables smarter decisions before demand hits

Our Approach - Predictive Modeling

Our dataset combines structured tabular features (calendar events, SNAP flags, prices) with sequential sales data at the store-item level.

Below are the model that can cover both structured feature learning and time-series forecasting needs.

Tree-Based Model

LightGBM

Neural Network

LSTM

Nerual Network

GRU

Data Structure

Time Span : past 365d before d_1914
Catogory : Hobby, Foods, Households

Seasonality & Time-Based

- Day of week
- month
- year
- weekend/weekday
- Calendar events
- Event Type
- SNAP

Sales & Pricing Signals

- 7 day moving averages price/sales volume
  - mean, max, min, std
- 28 day moving averages price / sales volume
  - mean, max, min, std
- daily price / sales volume
  - mean, max, min, std, lag 1 day
- Price value difference
- Price percentage difference

Geographic/Product Identifiers

- State_id
- Store_id
- Cat_id
- Dept_id

Model 1 : LightGBM

Rationale

LightGBM is a decision tree–based model that learns by repeatedly asking "yes or no" questions to split data into smaller, more accurate groups.

Each new tree corrects the errors of the previous one, improving accuracy.

Benefits

Fast & scalable ideal for large tabular datasets like sales and promotions.

Can quickly identifies key sales drivers (e.g., holidays, discounts) and delivers interpretable results to support decision-making.

LightGBM Forecasting Workflow

Data Segmention

- Split dataset into 3 product categories: HOBBIES, HOUSEHOLD, and FOODS.
- Each category is trained with a separate LightGBM model.

Feature Engineering

- Encode Categorical Feature
  - store_id
  - state_id
  - cat_id
  - dept_id
- Compute & Encode Time Feature
  - month
  - year
  - w_day
  - is_weekend
  - is_event
  - event_type
  - snap_CA/TX/WI
- Compute Price and Sales Feature
  - min/max/mean/std of past 28d
  - 28d moving avg. of price/sales
  - 1d lag of price/sales
  - value/pct difference of price

LightGBM Model

1. Start Training
2. Generate Training & Test Sets
3. Validate Model Performance
4. Calculate Forecast Error
5. Optimize Model Parameters
6. Final Training & Prediction

Model 2 : LSTM

Rationale

LSTM is a type of neural network designed to learn from sequences of data, like time series.

It remembers patterns over long periods by deciding what to keep, update, or forget at each time step—making it ideal for predicting what comes next.

Benefits

Great for capturing time-based patterns in sales data, like seasonality and trends.

Learns from past behavior to predict future demand—even when influenced by complex, long-term dependencies such as holidays or promotions.

LSTM Forecasting Workflow

Data Segmention → **Data Preprocessing** → **Sequence Modeling Setup** → **LSTM Model**

Data Segmention

- Split dataset into 3 product categories: HOBBIES, HOUSEHOLD, and FOODS & 10 stores: CA_1, CA_2, CA_3, CA_4, TX_1, TX_2, TX_3, WI_1, WI_2, WI_3
- Each category & store trained with a separate LSTM model for better pattern learning

Data Preprocessing

- Categorical data :
  - one-hot encoding
- Numeric data :
  - Normalization

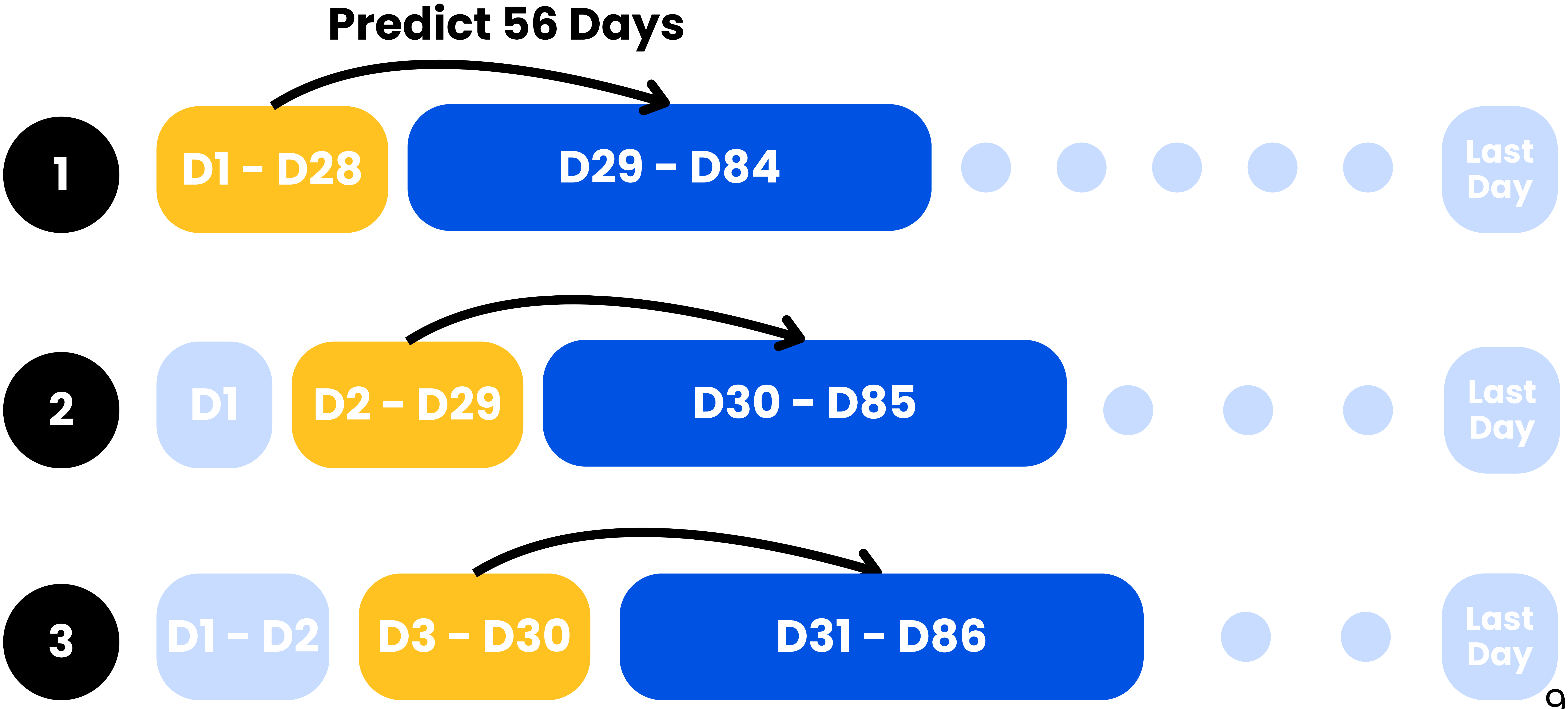
Sequence Modeling Setup

- Input sequence: 28 days → Output forecast: 56 days ahead
- Generate Sliding Window Samples

LSTM Model

- Set Training Parameters
- Start Training
- Final Training & Prediction

LSTM Sliding Window Prediction Principle



Model 3 : GRU

Rationale

GRU is a neural network model designed to capture patterns over time in sequential data like daily sales.

It uses gates to decide what information to keep or forget, making it efficient for learning short-term trends.

Benefits

GRU is lightweight and faster to train than LSTM, making it a strong choice for time-series forecasting.

It captures short-term trends and seasonality with fewer parameters, enabling faster training and solid forecasting performance.

GRU Forecasting Workflow

Data Segmention

- Split dataset into 3 product categories: HOBBIES, HOUSEHOLD, and FOODS
- Each category trained with a separate GRU model for better pattern learning

Feature Engineering

- Categorical data :
  - one-hot encoding
- Numeric data :
  - Normalization
  - Log transform if variances are large

Sequence Modeling Setup

- Input sequence: 28 days → Output forecast: 28 days ahead
- Used a custom generator to batch sequential inputs and target outputs

GRU Model

Start Training
↓
Generate Training & Test Sets
↓
Validate Model Performance
↓
Calculate Forecast Error
↓
Optimize Model Parameters
↓
Final Training & Prediction

Insights & Recommendation

1

LSTM as Forecasting Model

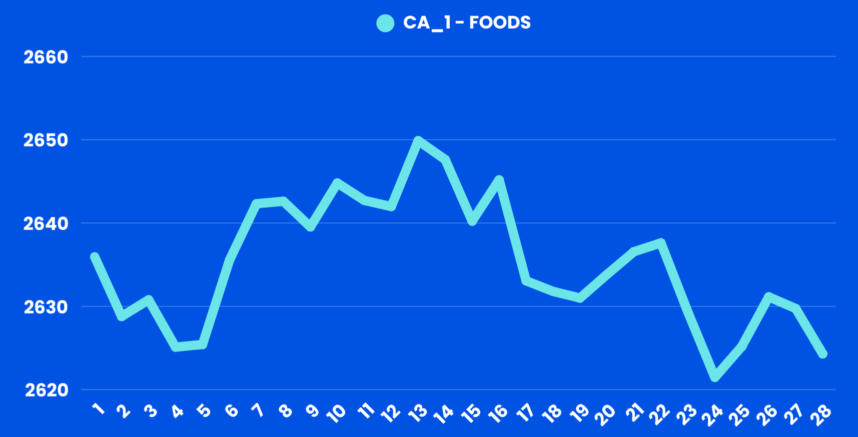
- Better captures the pattern of long-term sequential data
- Learns from past behavior and accommodate complex features such as promotion, holiday etc.
- Relatively better accuracy with RMSSE of 1.50406

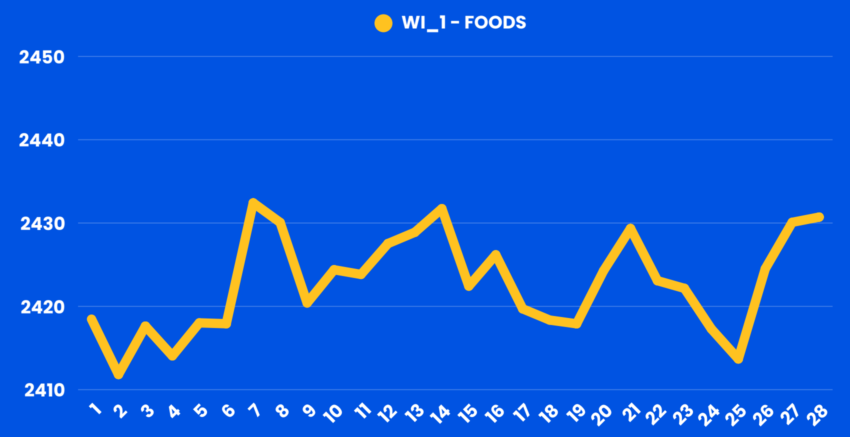
2

Strategize Demand Planning

Walmart can allocate inventory based on forecast results (down to item-level sales forecasting). For example:

- High demand in week 2 for store CA_1
- High demand on weekends for store WI_1





3

Automation as Next Step

Developing an automated pipeline to import historical data and generate sales forecasts can enable Walmart to dynamically and efficiently adjust its inventory plans

Thank You