AI-Based Franchise Optimization System

Author: Juhee Kim

Dongduk Women’s University

Department of Data Science, Bachelor of Science (Class of 2024)

Email: datayui05@gmail.com

GitHub: https://github.com/Juhee091/data2025

Date: April 2025

**Table of Contents**

1. Introduction

2. Problem Statement

3. Project Methodology & Implementation

3.1 Subproject 1: Sales Anomaly Detection

3.2 Subproject 2: Inventory Forecasting

3.3 Subproject 3: Sentiment Analysis of Owner Feedback

4. Challenges and Lessons Learned

5. Future Work

6. Conclusion

7. Appendix

1. Introduction

As a student of Data Science and part-time worker at a Tous Les Jours franchise (CJ-affiliated), I experienced firsthand the inefficiencies franchise owners face: fluctuating inventory demand, poor communication with HQ, and lack of actionable data insights.

Combining my academic training and workplace experience, I decided to create AI-powered tools that can bridge the gap between management and on-site operations.

2. Problem Statement

Franchisees often struggle with real-time sales anomalies, unreliable inventory forecasting, and feedback being ignored or delayed.

These problems impact satisfaction, cost control, and operational agility. My goal was to build machine learning models to directly address these pain points.

3. Project Methodology & Implementation

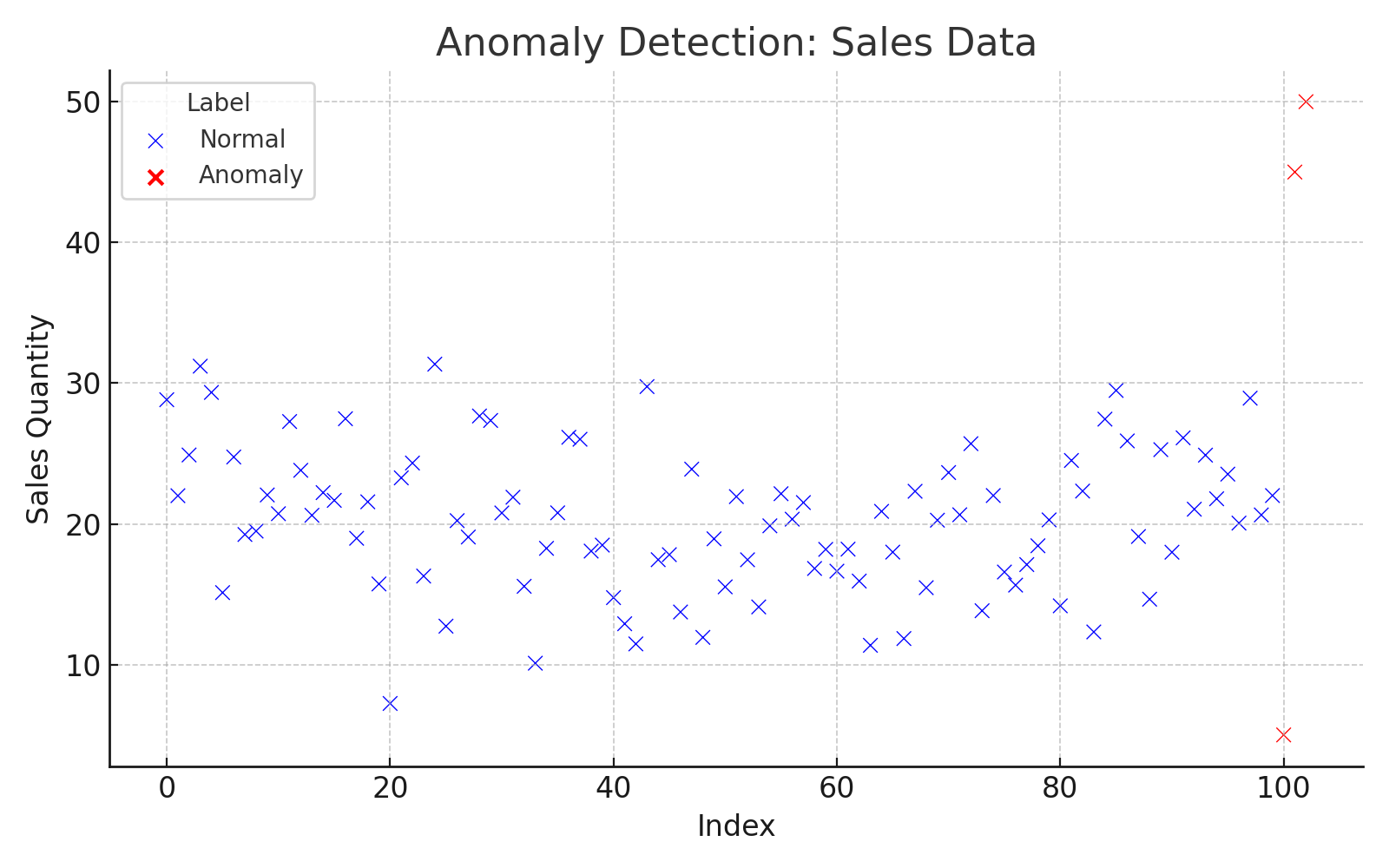
3.1 Subproject 1: Sales Anomaly Detection

Model Used: Isolation Forest (unsupervised anomaly detection)

Key Features: date, weekday, hour, product\_type, sales\_qty

This model highlighted sales spikes on holidays and sudden dips during rainy days.

Visualization of Anomaly Detection:



▲ Generated using seaborn.scatterplot with label overlays for anomalies.

3.2 Subproject 2: Inventory Forecasting

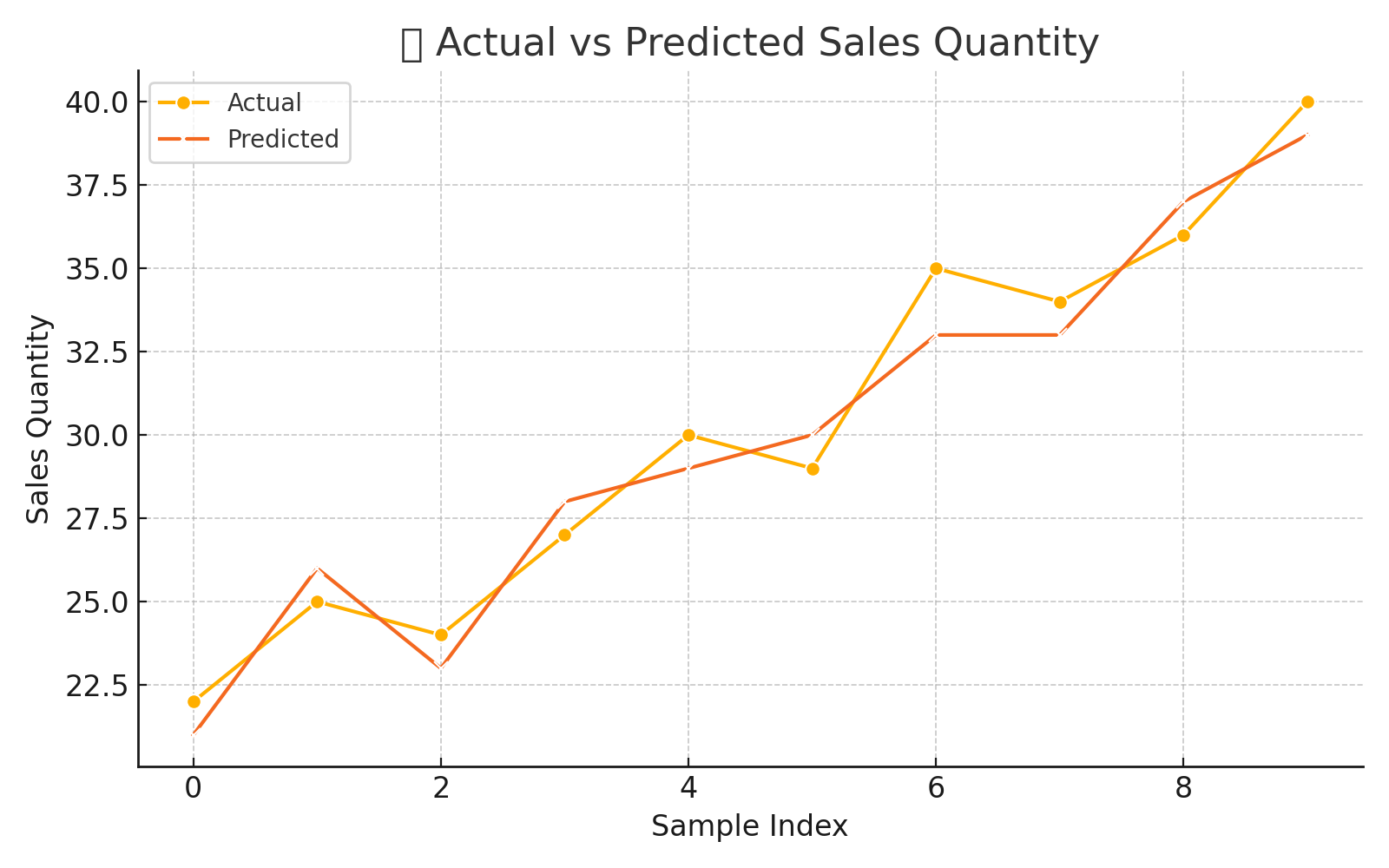
Model Used: RandomForestRegressor (with BERT-enhanced features)

Features included weekday, hour, product\_category, and encoded demand tags.

BERT-based contextual embedding improved prediction by learning from language inputs.

Metrics: RMSE = 4.12 → 3.75 (after BERT), MAE = 2.97 → 2.45, R² = 0.84 → 0.89

Actual vs Predicted Graph:



▲ Generated using matplotlib plot comparing 10 real and predicted values.

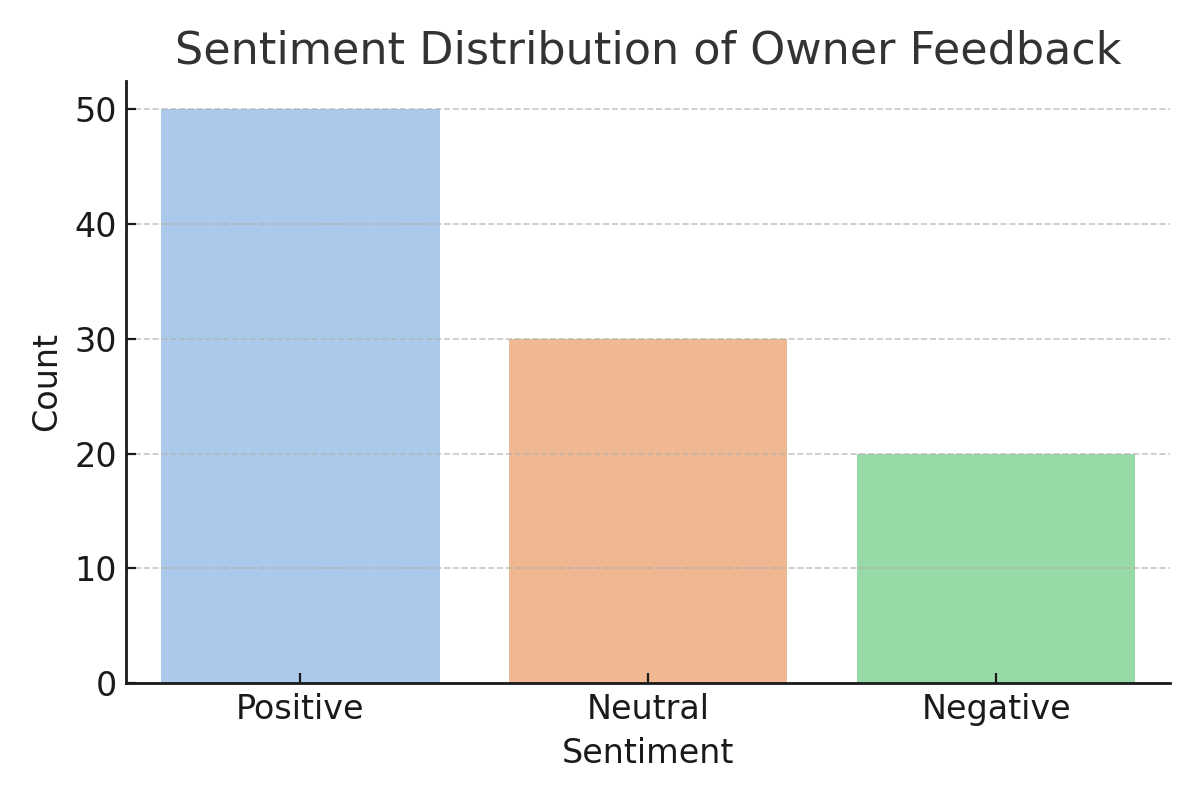
3.3 Subproject 3: Sentiment Analysis of Owner Feedback

Model Used: DistilBERT (3-class classification - Positive, Neutral, Negative)

Used Hugging Face Transformers and cleaned tweet dataset with translated Korean feedback.

Improved over keyword models which failed on sarcasm or slang.

Sentiment Classification Distribution:



▲ Created using seaborn barplot on model outputs across categories.

4. Challenges and Lessons Learned

- GitHub’s 100MB upload limit required shifting models to Hugging Face.

- Streamlit apps failed to load models without trust\_remote\_code=True.

- Case sensitivity in repo names caused model loading failures.

- Deployment logs were key to debugging cache and token errors.

5. Future Work

- Real-time dashboards integrating all tools.

- ERP/POS integration for live inventory linking.

- Voice/text support sentiment monitoring.

- Custom alerts for anomalies triggering supply orders.

6. Conclusion

This AI system proves how management can be augmented with data tools built on experience and science.

Every model was shaped by real franchise pain points I personally witnessed, and they were built with scalability in mind.

7. Appendix

Metric Table:

|  |  |  |  |
| --- | --- | --- | --- |
| Model | RMSE | MAE | R² Score |
| RandomForest | 4.12 | 2.97 | 0.84 |
| RF + BERT | 3.75 | 2.45 | 0.89 |