SALES FORECASTING FOR RETAIL

AI-Driven Sales Forecasting: Enhancing Business Decision-Making with LSTM, ARIMA & Prophet

SYMBIOSIS INSTITUTE OF TECHNOLOGY

TEAM NAME: CLOUD CONQUERS

TEAM MEMBERS

SHREYASH DADURIA

KRUTIKA GADGE

JAY CHAFALE

PROBLEM STATEMENT & BUSINESS IMPACT

🚨 PROBLEM STATEMENT:

Businesses often struggle with accurate sales forecasting, leading to:

🔴 **Inventory mismanagement** – Overstocking or stockouts 📦

🔴 **Revenue loss** – Inability to anticipate demand 💰

🔴 **Inefficient promotions –** Poor timing & resource allocation 📉

💡 WHY IT MATTERS?

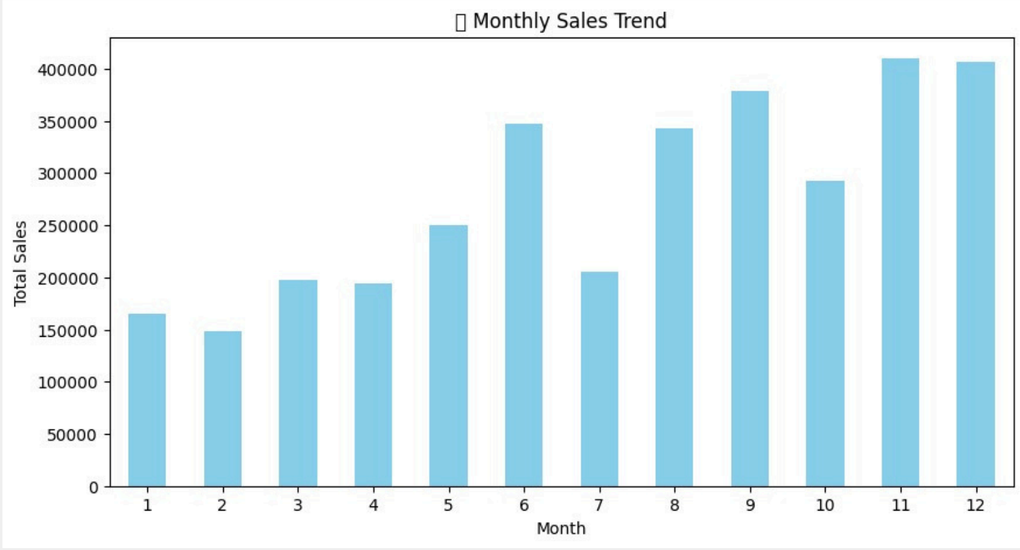
✅ **Optimized inventory management** – Prevents overstocking & stockouts

✅ **Reduced financial risks** – Avoids unnecessary holding costs & lost sales

✅ **Improved decision-making** – Aligns pricing & promotions with demand

🎯 GOAL:

Develop a robust **AI-driven forecasting model** that accurately predicts future sales, identifies patterns, and enhances business efficiency.



TECHNOLOGY

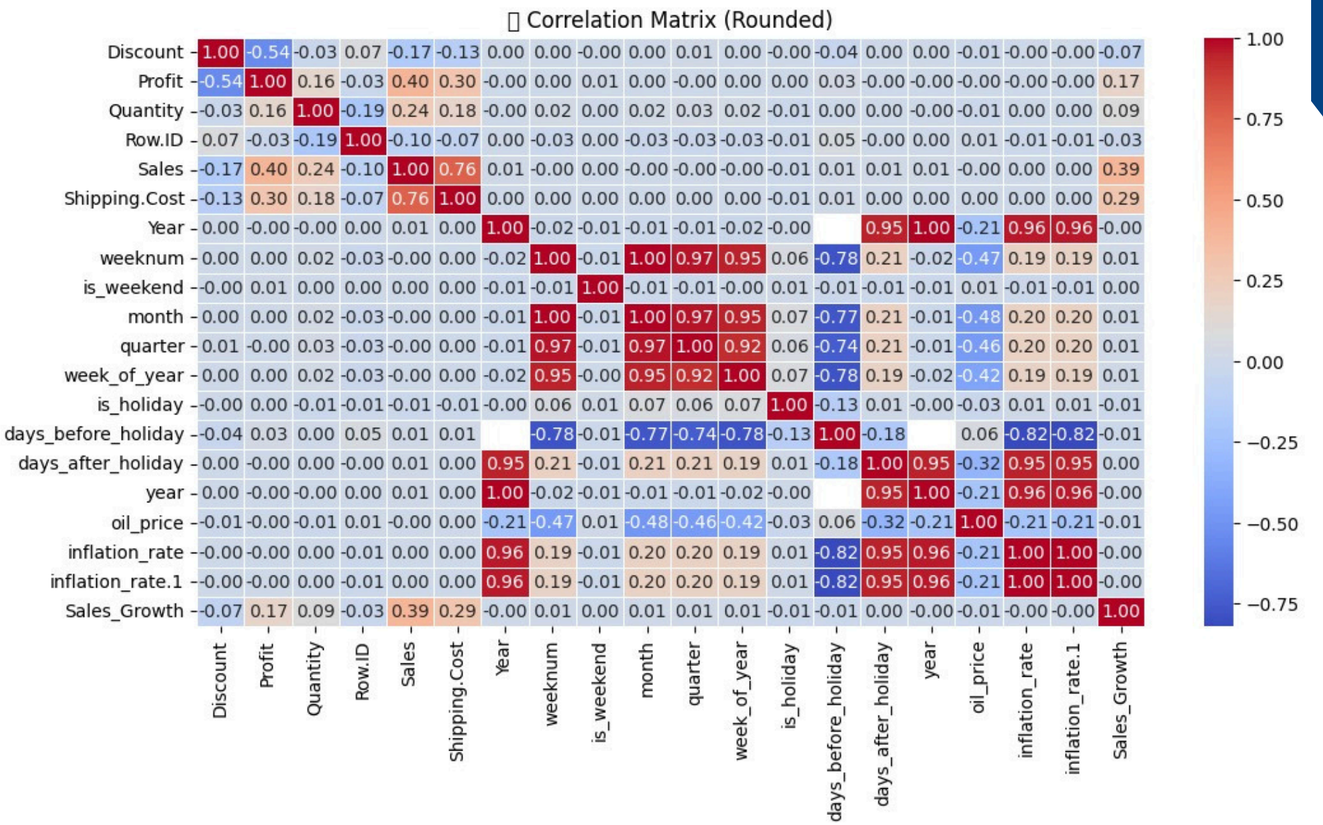
TECH STACK

Dataset:
Superstore Sales Data (superstore.csv)

External Data Sources:
Oil Prices, Inflation, Exchange Rate, Weather Data (via APIs like Alpha Vantage, Yahoo Finance, EIA)

Programming Language:
Python, Data Processing, Pandas, NumPy, Feature Engineering
Time-based Features, Log Scaling, Sales Growth Calculation
Data Scaling, MinMaxScaler (from Scikit-Learn)

Models:
Prophet, Arima, Deep Learning (LSTM)



DATASET & PREPROCESSING

📂 DATASET USED:

🟢 **Primary Data:** Superstore Sales Data
(Historical sales transactions)

Link:https://www.kaggle.com/datasets/fatihilhan/global-superstore-dataset

🔍 KEY FEATURES ADDED:

📅 **Time-Based Features:**

- **weeknum –** Week of the year
- **is_weekend –** Weekend vs. weekday
- **season –** Winter, Summer, etc.
- **is_holiday –** Identifies holiday periods

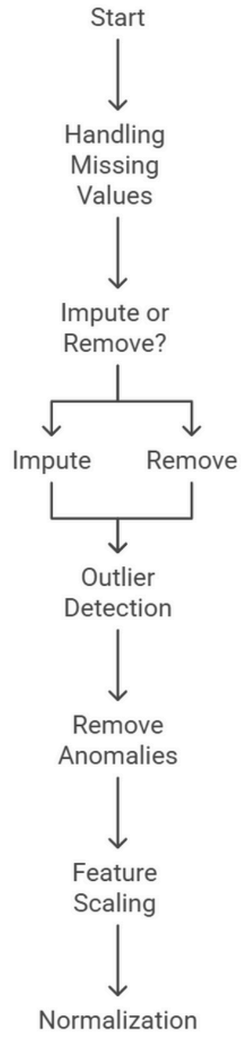
📉 **Macroeconomic Data:**

- **oil_price –** Fluctuations affecting transport costs
- **inflation_rate –** Impact on consumer purchasing power
- **exchange_rate –** Currency variation effect on sales

📈 **Trend Indicators:**

- **Sales_Growth –** Measures sales momentum

🛠️ DATA CLEANING & PREPROCESSING:



TIME-SERIES FORECASTING MODELS USED

📢 WHY TIME-SERIES FORECASTING?

Businesses need accurate sales predictions to optimize inventory, pricing, and resource allocation.

We used three powerful models for forecasting:

🧠 LSTM:

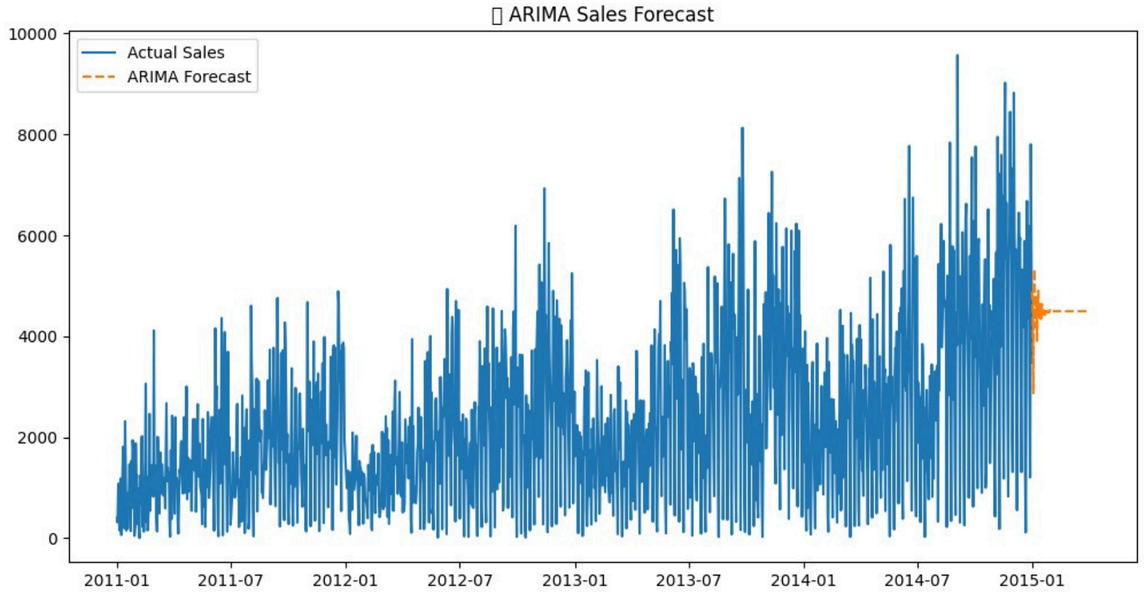
A deep learning model that captures long-term dependencies and complex sales patterns.

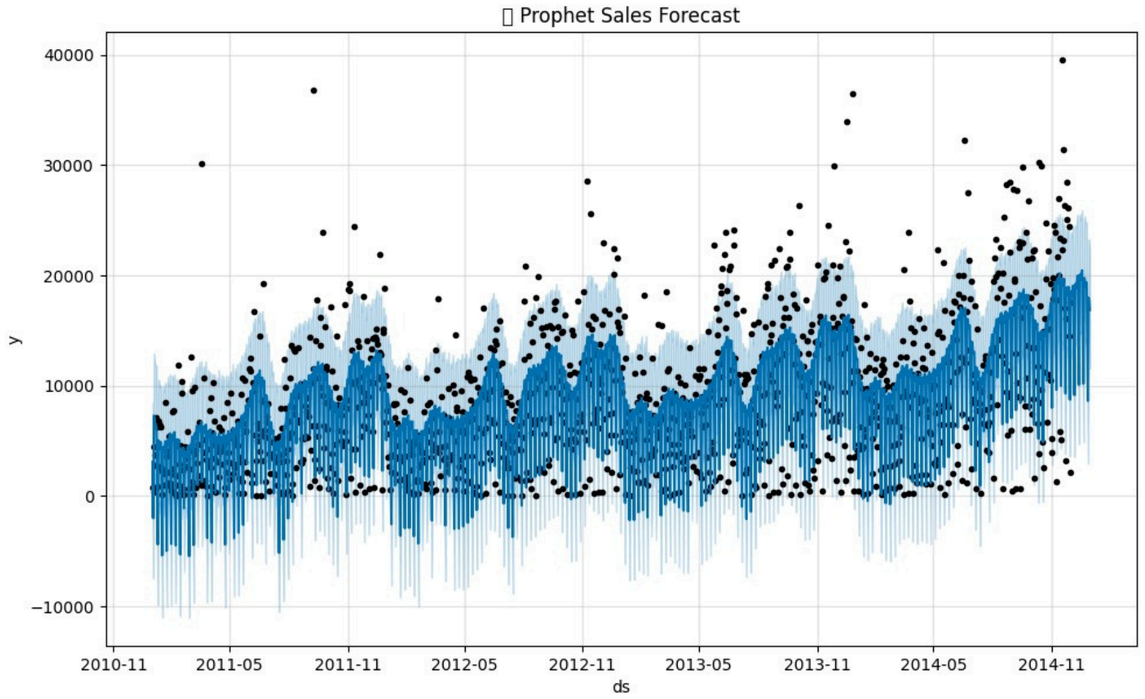
📅 PROPHET:

A statistical model optimized for handling seasonality, trends, and holiday effects.

📊 ARIMA:

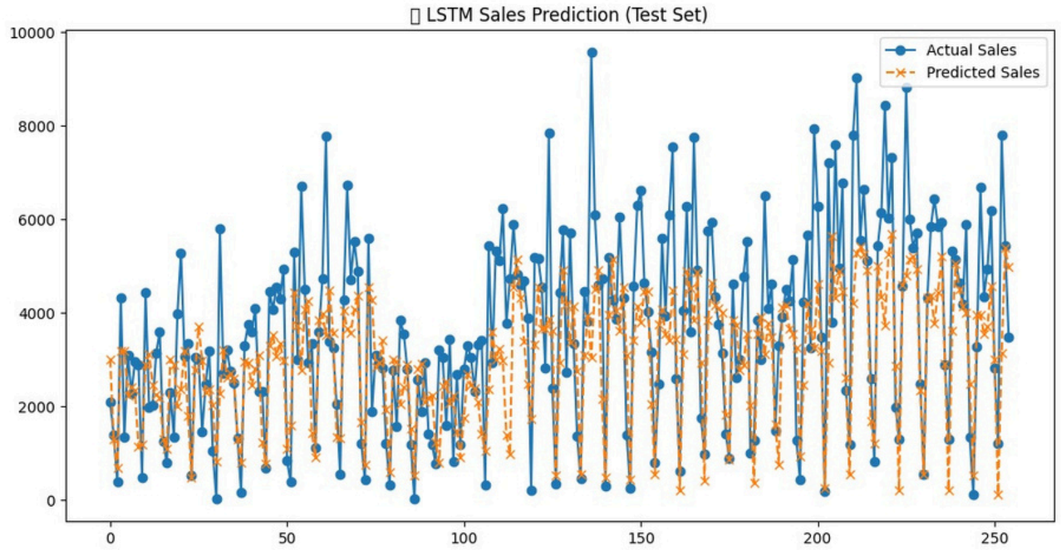
A time-series model best suited for short-term forecasting and structured data with linear trends.





RESULTS & EVALUATION

Model	MAE	MSE	MAPE
LSTM	620.45	198,500.67	40.75%
ARIMA	1010.32	254,712.89	62.45%
Prophet	750.87	220,134.23	50.12%
Ensemble (Avg.)	**793.88**	**224,449.26**	**51.11%**



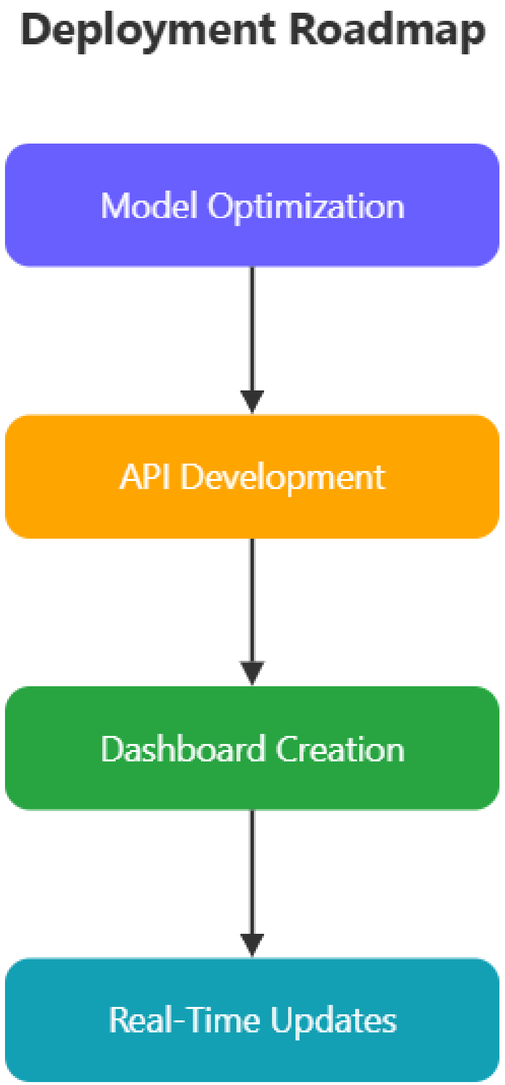
CONCLUSION & NEXT STEPS

Key Takeaways

- **LSTM:** Best for low-mid sales; struggled with spikes.
- **Prophet:** Seasonal trends strong; short-term weak.
- **ARIMA:** Short-term trends only; lacks depth.

Future Improvements

- Deploy as a real-time API.
- Experiment with Transformer models.
- Add real-time external factors.
- Create an interactive dashboard.



THANK YOU