Big Data Analytics Capstone Project

Customer 360° Analytics Pipeline for Omnichannel Retail

"Turning raw customer data into intelligent insights — building a unified view of your retail customers through a modern data pipeline."

Presented by
Koyalkar Sri Harsha
24MBMB15

Project Overview

The **Customer 360° Analytics Pipeline for Omnichannel Retail** aims to revolutionize customer engagement by integrating data from all retail channels. This project focuses on gathering comprehensive insights into customer behavior, preferences, and interactions, allowing businesses to create personalized experiences that enhance customer satisfaction and drive sales across platforms.
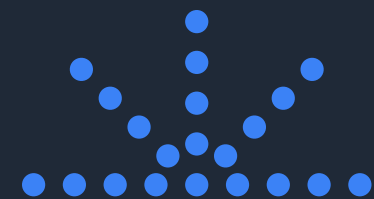
Challenges Addressed:

❌ Fragmented Data: Customer information scattered across multiple systems

❌ Poor Personalization: Inability to create targeted marketing campaigns

❌ Reactive Churn Management: Losing customers before identifying risk

❌ Limited Insights: No unified view of customer lifetime value

❌ Manual Processes: Time-consuming data preparation and analysis
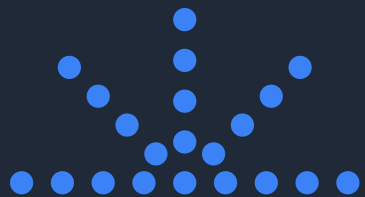
Solution:

✅ **Unified Customer 360° View: Single source of truth for all customer data**

✅ **Predictive Analytics: ML models for churn, CLV, and segmentation**

✅ **Real-time Insights: Interactive dashboards for data-driven decisions**

✅ **Scalable Architecture: Medallion architecture with Delta Lake**

✅ **Automated Pipeline: End-to-end orchestration with minimal manual intervention**

Personas
Diverse personas enhance targeted marketing and segmentation strategies.

Patterns
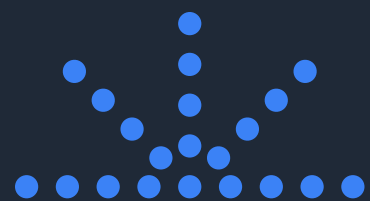**Analyzing behavior** patterns reveals customer preferences and trends.
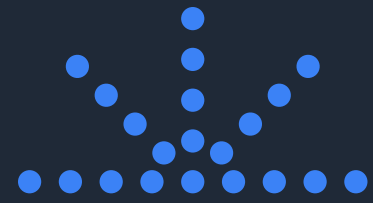

Interactions
Understanding interactions leads to improved customer experiences.

Trends
Tracking trends helps anticipate future customer needs effectively.

🤖 ML Models

1. Churn Prediction

2. Customer Segmentation
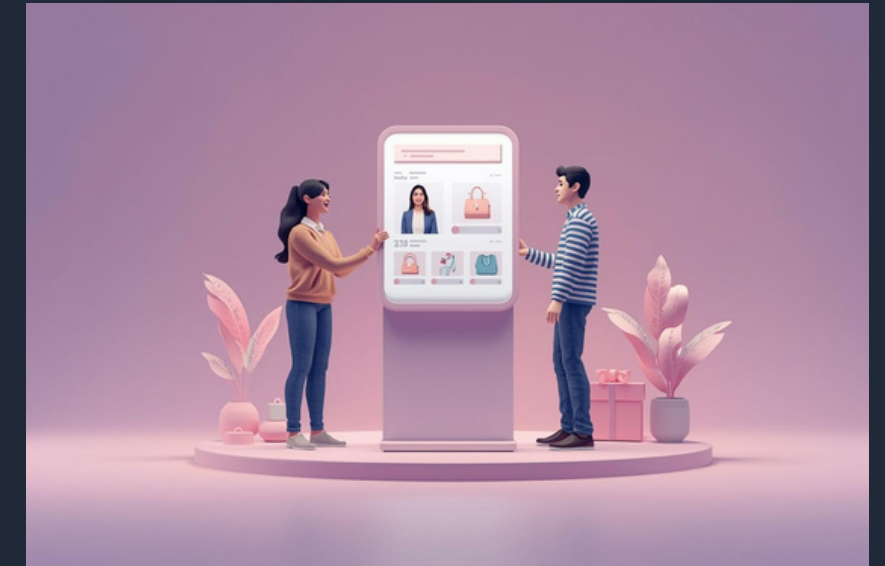
3. CLV Prediction

4. Sales Forecasting

5. Anomaly Detection

Cross-Channel Consistency
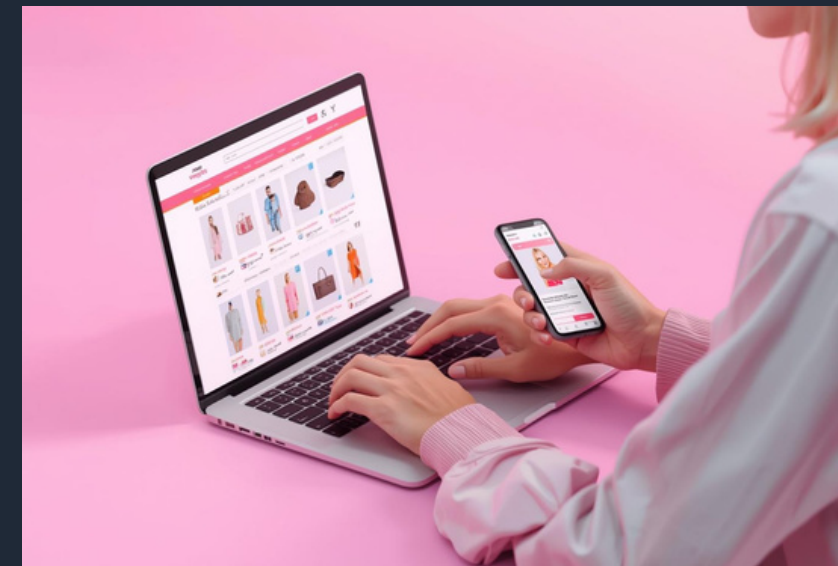Ensures brand message remains unified across all platforms.

Improved Engagement
Interactive strategies boost customer interest and foster loyalty.

Brand Loyalty
Strengthening customer relationships through reliable brand experiences.

Enhanced Insights
Analyzing customer interactions drives better decision-making and strategies.

Personalized Offers
Customers receive tailored promotions and incentives based on behavior.
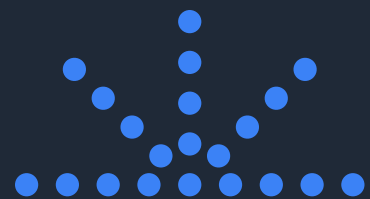
🛠️ Technical Achievements

Architecture: Scalable Medallion architecture (Bronze-Silver-Gold)

Technology Stack: PySpark, SQL, Delta Lake, MLlib, Databricks

Data Processing: 70,000+ records across 21 tables (Bronze: 7, Silver: 8, Gold: 6)
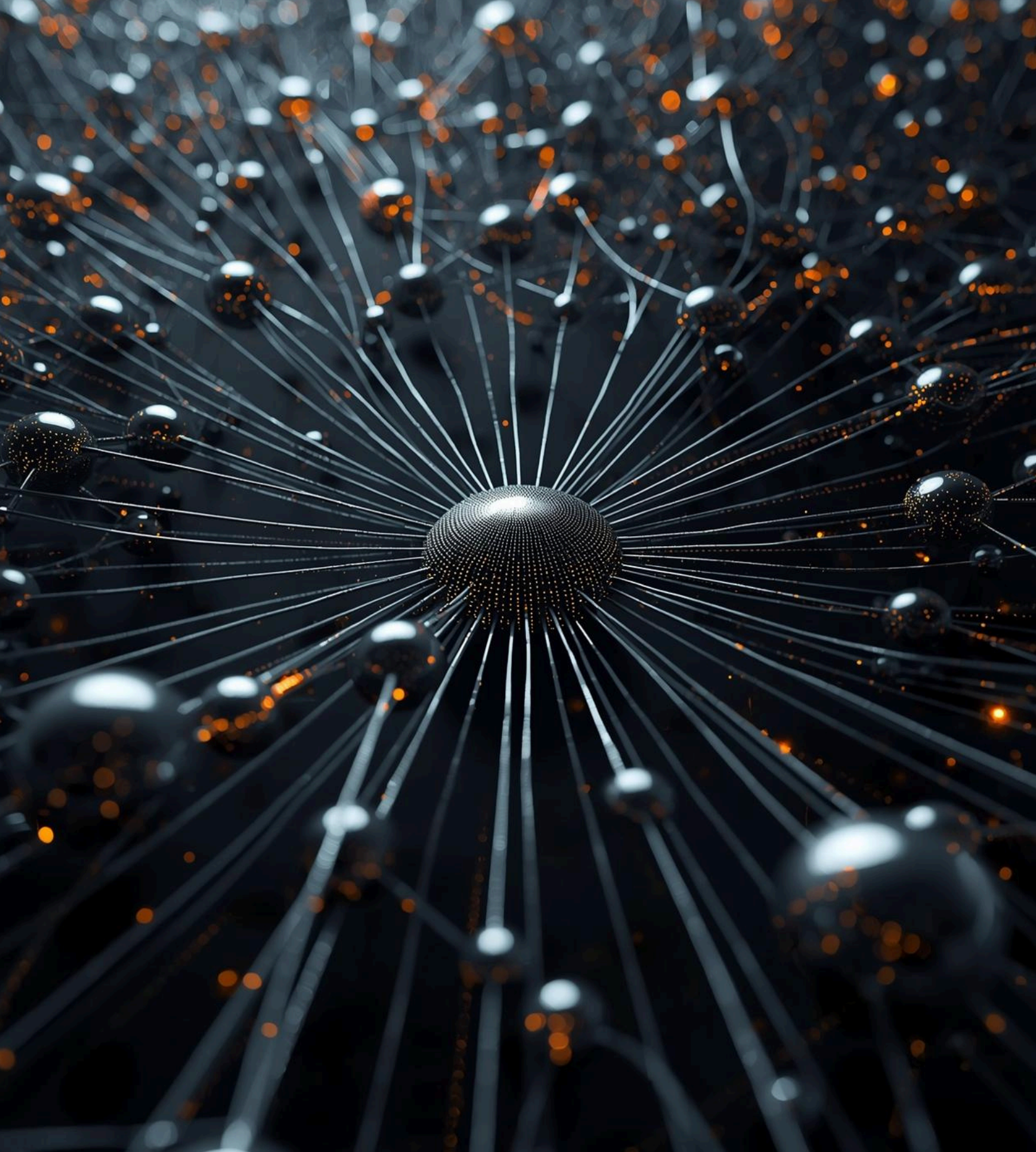
ML Pipeline: 5 models with MLflow tracking & explainable AI

Performance: Sub-minute query response times with Delta optimizations

Project Conclusion

Successfully delivered a production-ready, scalable, and automated Customer 360° Analytics Pipeline that transforms fragmented data into actionable insights. The solution enables predictive analytics, personalized experiences, and data-driven decision making - positioning the organization to reduce churn by 15-20%, increase CLV by 25%, and improve forecast accuracy by 30%.

Thank You