Day 14

AI-Powered Analytics

Databricks Genie & Mosaic AI

Databricks 14-Day AI Challenge

Yash Kavaiya

Agenda

- **AI-Powered Analytics**
  - What it is and why it matters
  - Traditional vs AI workflows
- **Databricks Genie**
  - Natural Language to SQL
  - Genie Data Spaces
- **Mosaic AI Platform**
  - Foundation Models
  - Model Serving & Vector Search
- **Generative AI Integration**
  - AI Functions in SQL
  - RAG Applications
- **AI-Assisted Analysis**
  - Sentiment Analysis
  - Automated Insights
- **Best Practices**
  - Data quality and governance
  - Cost optimization

What is AI-Powered Analytics?

Definition

The convergence of AI and traditional data analytics — enabling natural language interactions, intelligent recommendations, and automated analytical tasks.

Business Value:

- Question to insight: hours → seconds
- Empowers non-technical users
- Data democratization
- Faster decision-making

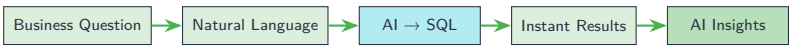
Technical Benefits:

- ▷ Automatic query optimization
- ▷ Intelligent data discovery
- ▷ Pattern recognition at scale
- ▷ Predictive insights generation

Traditional vs AI-Powered Analytics

Traditional: Hours to Days

AI-Powered: Seconds to Minutes

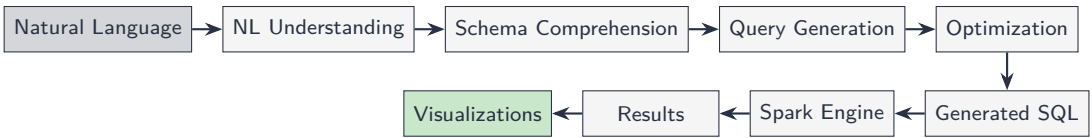


Key Insight: AI-powered analytics compresses multi-step workflows into streamlined, accessible processes!

Databricks Genie: Natural Language to SQL

What is Genie?

An AI-powered conversational interface that transforms natural language questions into SQL queries — like having a 24/7 data analyst!



Genie Data Spaces

What is a Data Space?

A curated environment defining the scope of data Genie can access — a focused "sandbox" for specific business domains.

Aspect	Without Data Spaces	With Data Spaces
Scope	Access to all tables	Focused on relevant data
Accuracy	May query wrong tables	Precise table selection
Security	Broad access concerns	Controlled exposure
Performance	Searches entire catalog	Faster comprehension
Context	Generic understanding	Domain-specific knowledge

Genie Query Examples

"Total revenue by category"

- "total revenue" → SUM aggregation
- "by category" → GROUP BY

"Highest conversion rate?"

- Conversion = Purchases/Views
- ORDER BY DESC + LIMIT

$$\text{Conv Rate} = \frac{\text{Purchases}}{\text{Views}} \times 100$$

"Daily purchases trend?"

- Time series analysis
- COUNT by DATE

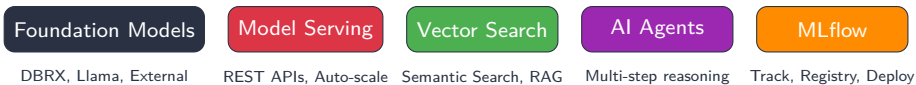
"Viewed but never purchased?"

- LEFT JOIN with NULL check
- Remarketing opportunities!

Mosaic AI Platform

What is Mosaic AI?

Databricks' comprehensive AI/ML platform for training, deploying, and managing AI models at enterprise scale.



Training
Fine-tuning & Pre-training

Serving
Real-time, Batch, Streaming

Tools
Feature Store, Evaluation

Foundation Model APIs

Available Models:

- **DBRX** — Databricks' own LLM
  - Optimized for enterprise
- **Llama Models** — Meta's open-source
  - Various sizes available
- **External Models** — AI Gateway
  - OpenAI, Anthropic, etc.

Key Benefits:

- ✓ No infrastructure management
- ✓ Pay-per-use pricing
- ✓ Unified API interface
- ✓ Built-in governance
- ✓ Enterprise security

Model Serving Types

Serving Type	Use Case	Latency	Scaling
Real-time	Interactive apps	< 100ms	Auto-scale
Batch	Large-scale scoring	Minutes	Parallel jobs
Streaming	Continuous predictions	Seconds	Event-driven

Tip: Use `scale_to_zero_enabled=True` to minimize costs when endpoints are idle!

Vector Search for Semantic Retrieval

What is Vector Search?

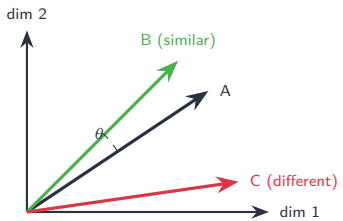
Enables semantic search by storing and querying high-dimensional vector embeddings — essential for RAG applications!

How it Works:

- Text → Embeddings (vectors)
- Similar items = close vectors
- Query by similarity, not keywords

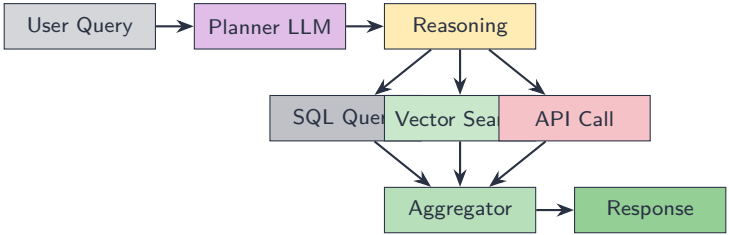
Cosine Similarity:

$$\cos(\theta) = \frac{A \cdot B}{\|A\|\|B\|}$$



Range: -1 to 1 (1 = identical)

AI Agents: Autonomous AI Systems



Capabilities:

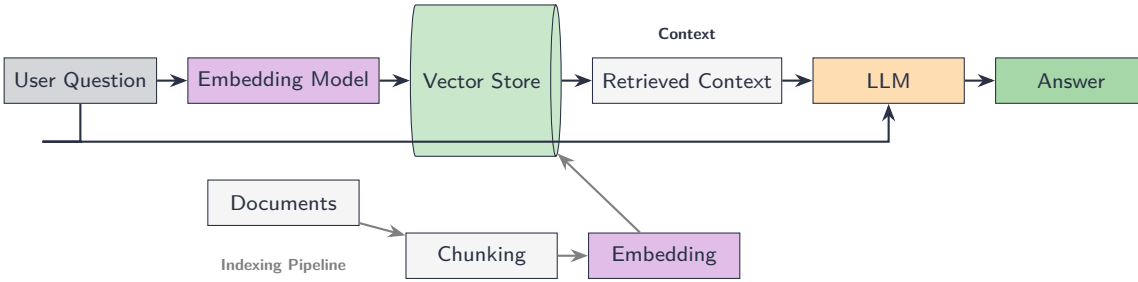
- Multi-step reasoning
- Tool use (SQL, APIs)
- Memory management
- Self-correction

AI Functions in Databricks SQL

Function	Description	Use Case
`ai_analyze_sentiment()`	Detect +/- sentiment	Feedback analysis
`ai_extract_entities()`	Extract named entities	Info extraction
`ai_summarize()`	Generate summaries	Document processing
`ai_classify()`	Categorize text	Ticket routing
`ai_translate()`	Language translation	Globalization
`ai_query()`	Custom LLM queries	Flexible AI tasks

Example: `SELECT ai_analyze_sentiment(review_text) FROM reviews`

RAG: Retrieval-Augmented Generation



How RAG Works:

1. Embed user question
2. Search for relevant docs
3. Provide context to LLM
4. Generate grounded answer

Sentiment Analysis Implementation

How it Works:

- Text → Model → Label + Score
- POSITIVE / NEGATIVE
- Confidence 0-100%

Confidence Score (Softmax):

$$P(\text{sent}|\text{text}) = \frac{e^{z_i}}{\sum_j e^{z_j}}$$

Use Cases:

- ▷ Product review analysis
- ▷ Customer feedback
- ▷ Social media monitoring
- ▷ Support ticket triage

✓ "Amazing product!" → POSITIVE
× "Terrible quality" → NEGATIVE

AI-Assisted Analysis

Definition

Augments human analytical capabilities with AI tools that automatically detect patterns, generate hypotheses, and provide explanations.

Anomaly Detection Identify unusual patterns automatically

Predictive Insights Forecast future trends

Root Cause Analysis Trace back factors causing changes

NL Explanations Convert findings to plain English

Best Practices: Data Quality

Issue	Impact on AI	Solution
Missing values	Incomplete analysis	Imputation or filtering
Inconsistent formats	Parsing errors	Standardization
Duplicate records	Inflated metrics	Deduplication
Outdated data	Irrelevant insights	Regular refresh
Biased samples	Skewed predictions	Stratified sampling

Remember: AI systems are only as good as the data they analyze!

Prompt Engineering for Better Results

Less Effective	More Effective
"Show me data"	"Show monthly revenue by category for Q4 2024"
"Sales info"	"Total sales amount grouped by customer segment"
"Best products"	"Top 10 products with highest profit margin"

LLM Prompt Tips:

1. Be Specific
2. Provide Context
3. Specify Format
4. Use Examples
5. Set Boundaries

Cost Optimization Strategies

Strategy	Description	Savings
Right-sizing	Match compute to workload	20-40%
Caching	Store frequent query results	30-50%
Batch Processing	Combine small requests	15-25%
Model Selection	Use smaller models when appropriate	40-60%

Pro Tip: Start with foundation models, fine-tune only when needed!

Quick Reference: When to Use What

Need	Solution
Business users need self-service analytics	Databricks Genie
Deploy custom ML models as APIs	Mosaic AI Model Serving
Build semantic search or Q&A systems	Vector Search + RAG
Analyze text data (sentiment, classification)	Foundation Model APIs
Track experiments and model versions	MLflow
Create automated insights	AI Functions + Custom Analysis

Key Takeaways

1. Databricks Genie NL to SQL — democratize data access

2. Mosaic AI Comprehensive AI/ML platform

3. Foundation Models No infrastructure management

4. Vector Search Essential for RAG applications

5. AI-Assisted Analysis Amplify human capabilities

6. MLflow Integration Tracking and governance

Congratulations!

14-Day AI Challenge Complete!

Day 14: AI-Powered Analytics

You've mastered:

Databricks Genie ▪ Mosaic AI ▪ Vector Search ▪ RAG ▪ MLflow

Connect with me:

LinkedIn: Yash Kavaiya

Gen AI Guru

Easy AI Labs