

# From Days to Minutes - Automate Machine Learning in your Enterprise with HeatWave AutoML

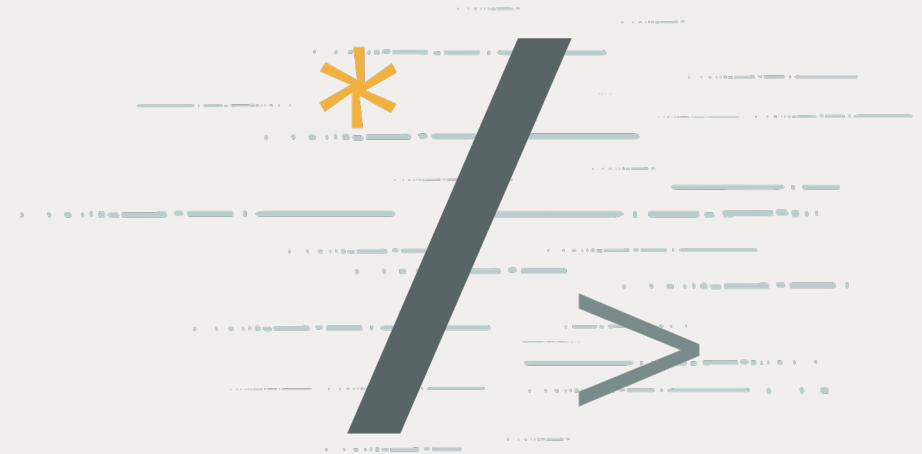
---

Sandeep Agrawal  
Developer, MySQL HeatWave

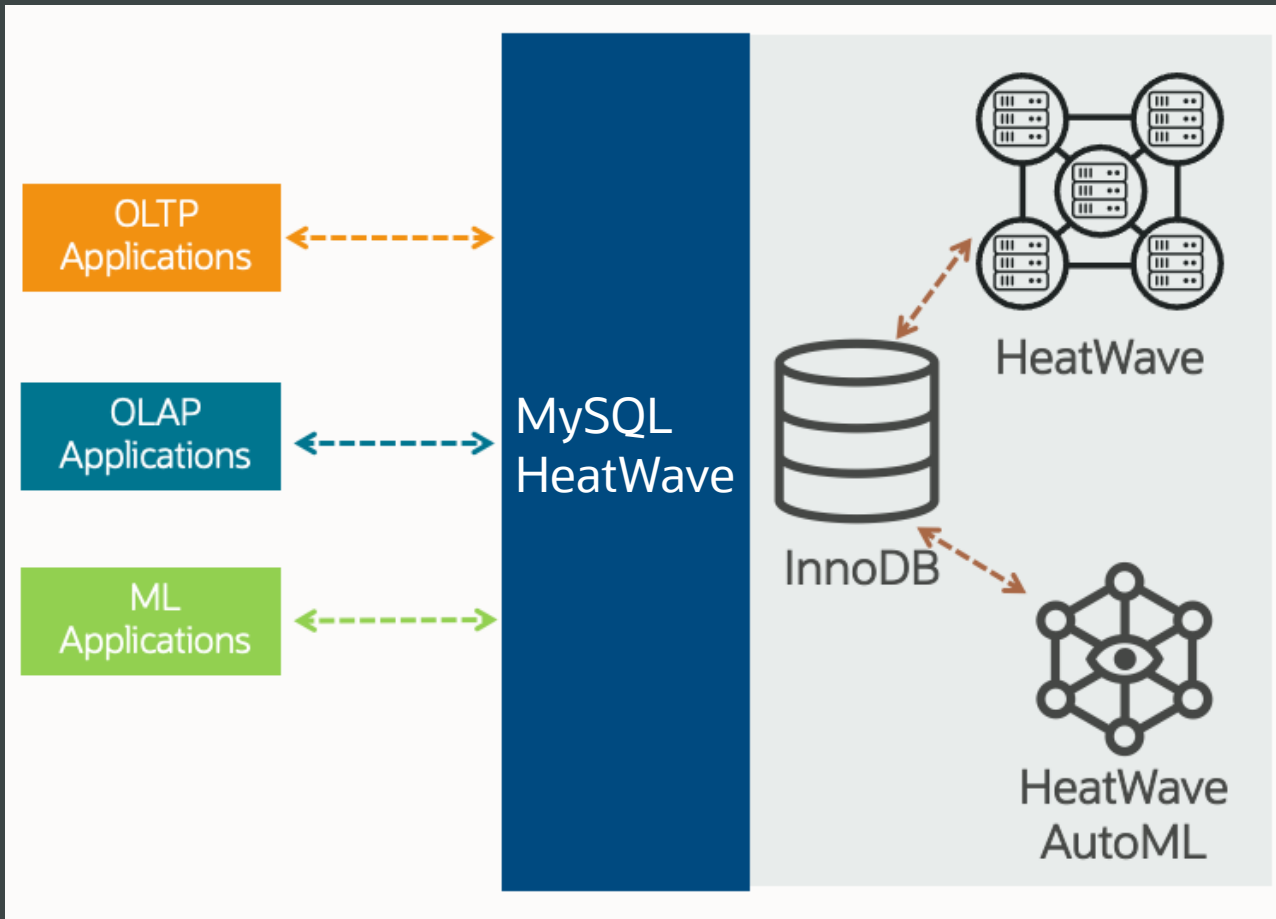
# Safe harbor statement

---

The following is intended to outline our general product direction. It is intended for information purposes only, and may not be incorporated into any contract. It is not a commitment to deliver any material, code, or functionality, and should not be relied upon in making purchasing decisions. The development, release, timing, and pricing of any features or functionality described for Oracle's products may change and remains at the sole discretion of Oracle Corporation.



# MySQL HeatWave AutoML



- Fully automated, highly scalable, cost-efficient & superior performance
- In-database machine learning
- Database developer and business analyst friendly
- Model agnostic explanations for model and predictions
- No additional cost for HeatWave customers



# Classification and regression

More than ten most popular algorithms for classification and regression

Classification use-case example:

- Classify loan application into loan approve or deny
- Explain why loan was approved or denied based on applicant data
- Determine what can the applicant do to change the outcome

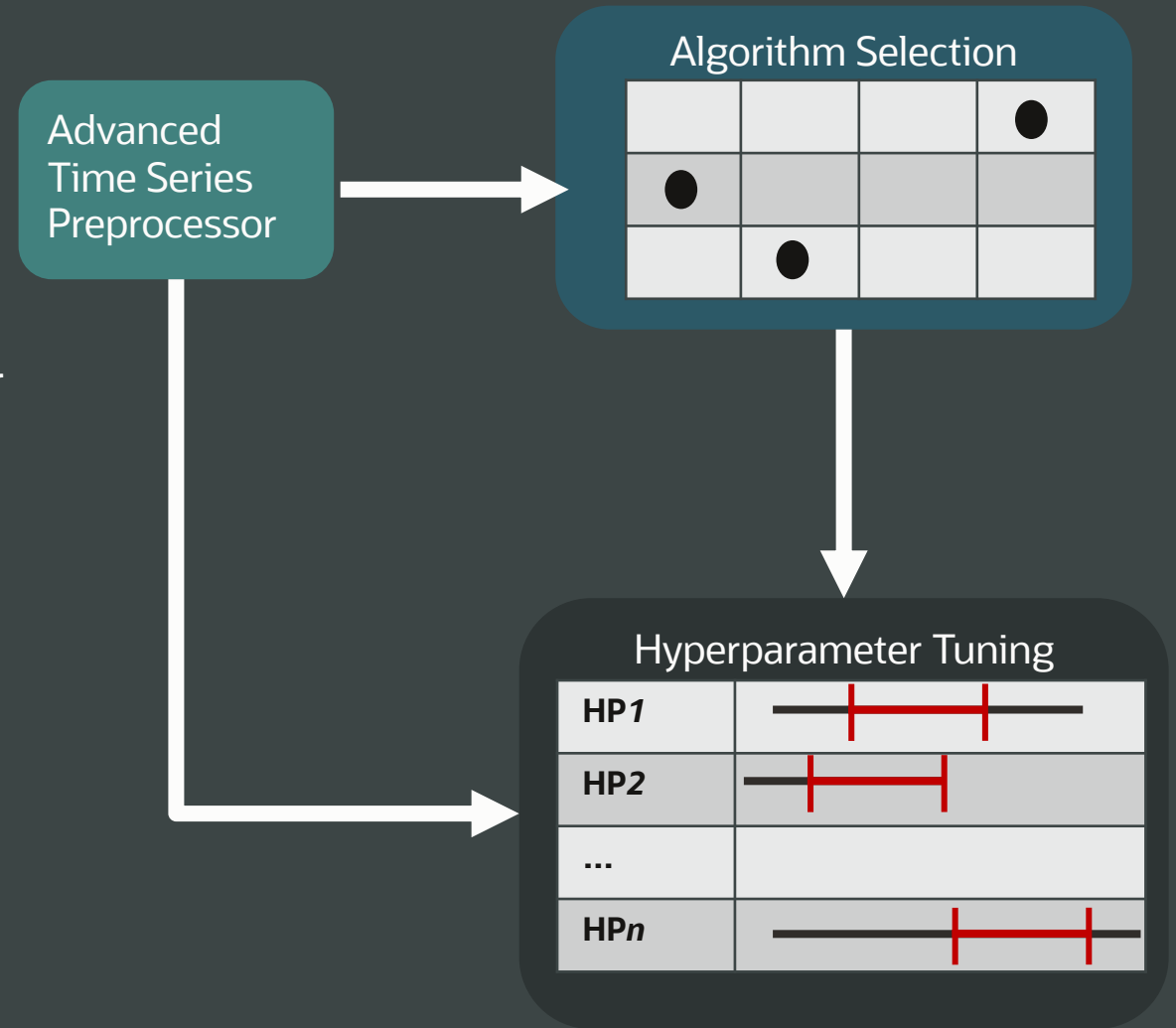
Regression use-case example:

- Provide APR for a loan applicant based on their data
- Explain what contributed to this APR



# Forecasting

- Predicts values for time series observations based on multiple variables
- Utilities: forecast electricity demand in winter considering other energy sources
- Commodities: forecast seasonal demand for corn considering demand other grains
- Novel automation techniques developed in HeatWave AutoML



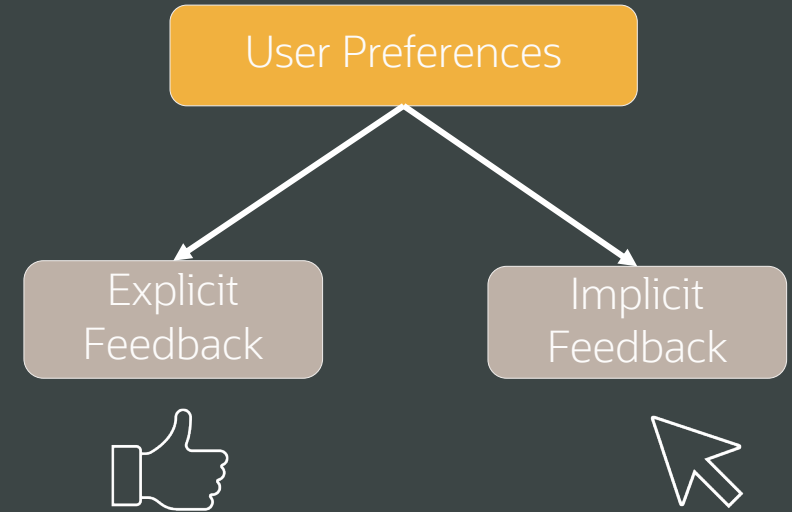
# Recommender system

PREDICTS RATINGS, TOP K ITEMS, SIMILAR USERS AND ITEMS

Recommend new products to customers based on purchase history

Recommender system provides multiple benefits in e-commerce:

- Personal and dynamic recommendations based on prior interactions
- Maximize conversion rates and increase in orders
- Increase customer satisfaction and improve retention
- Simplify product search and generate more revenue



# Anomaly detection

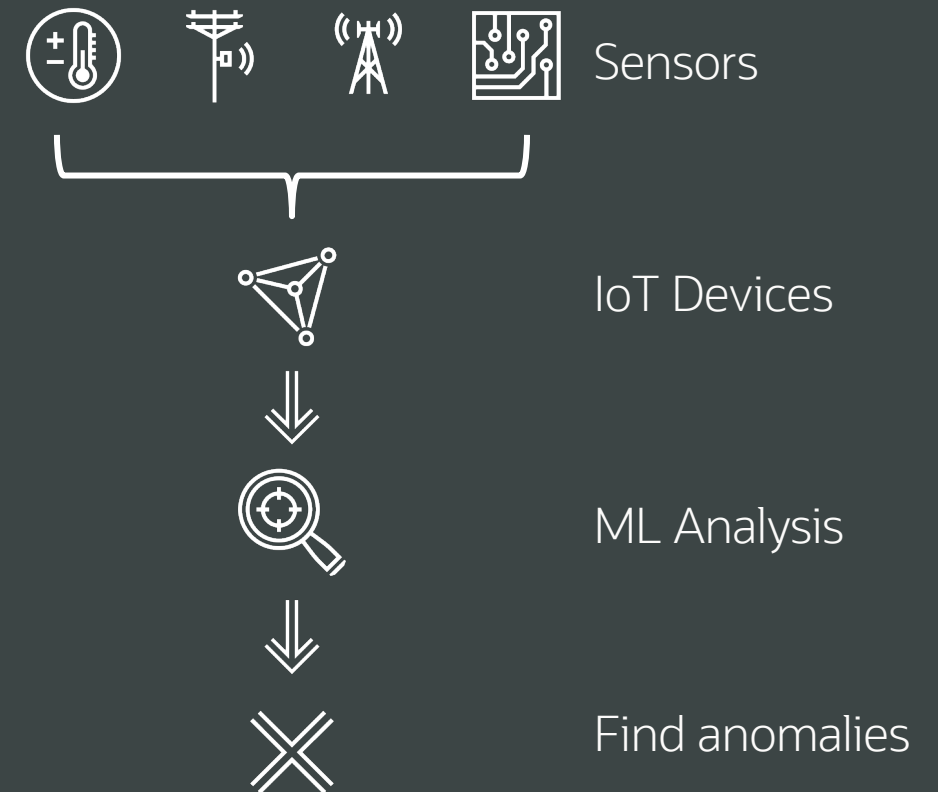
Training a model on unlabeled data is extremely difficult

HeatWave AutoML facilitates training on unlabeled data

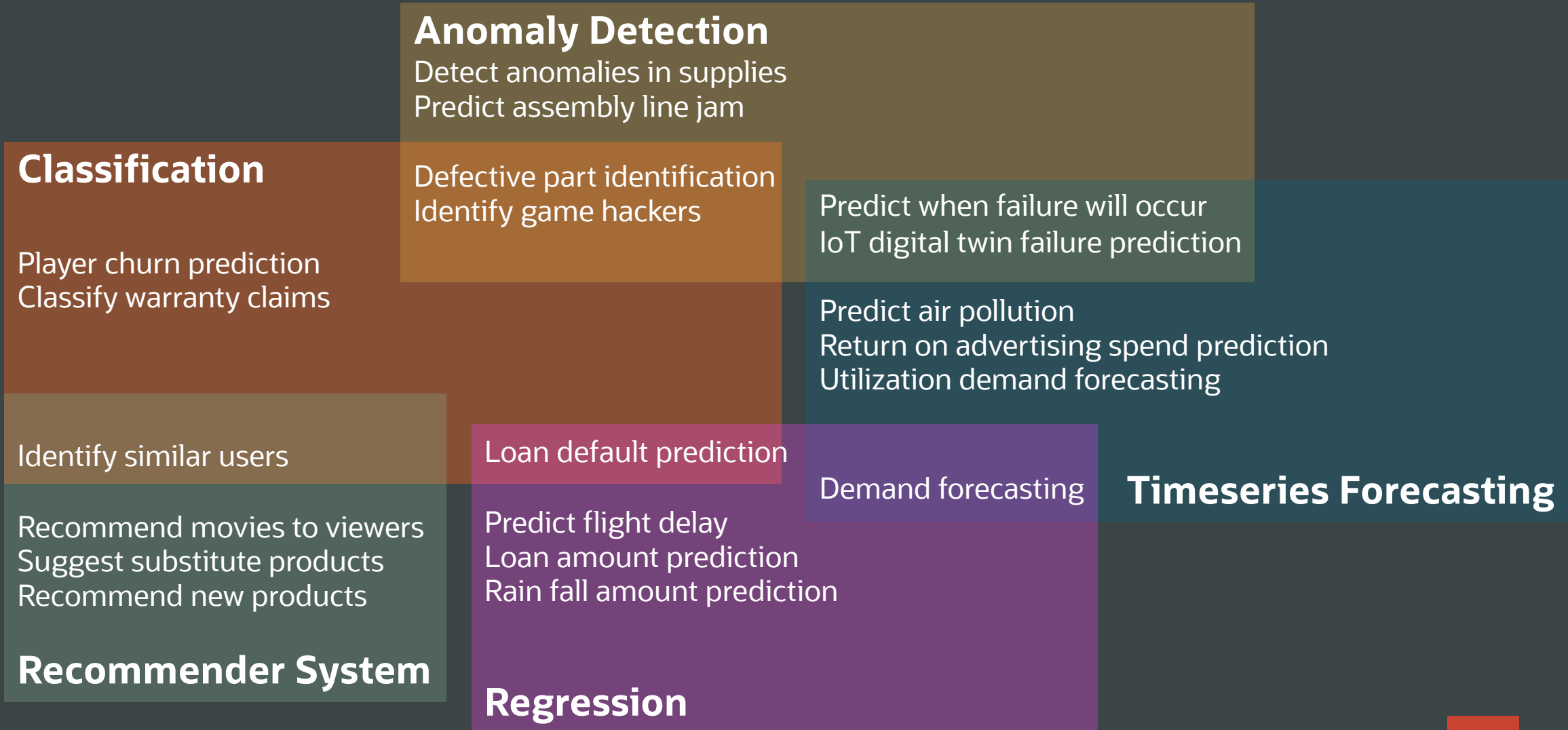
Proprietary algorithm detects multiple types of anomalies

Applications span variety of fields:

- Fraud detection in banking and finance
- Failure detection and prevention in manufacturing
- Root cause analysis based on log analysis



# HeatWave AutoML use-cases





# HeatWave AutoML functionality exposed via SQL

**Any SQL client can leverage HeatWave AutoML**

ML\_TRAIN – build an optimized model for a given table

ML\_EXPLAIN – train additional model or prediction explainers

ML\_SCORE – evaluate a trained model

ML\_PREDICT\_ROW / TABLE – inference on a row / table

ML\_EXPLAIN\_ROW / TABLE – explanations of predictions on a row / table

# Text column support

HeatWave AutoML now supports text columns in the database

Text columns are converted into numerical vectors and treated as regular features

Text can be of type – TINYTEXT, TEXT, MEDIUMTEXT or LONGTEXT

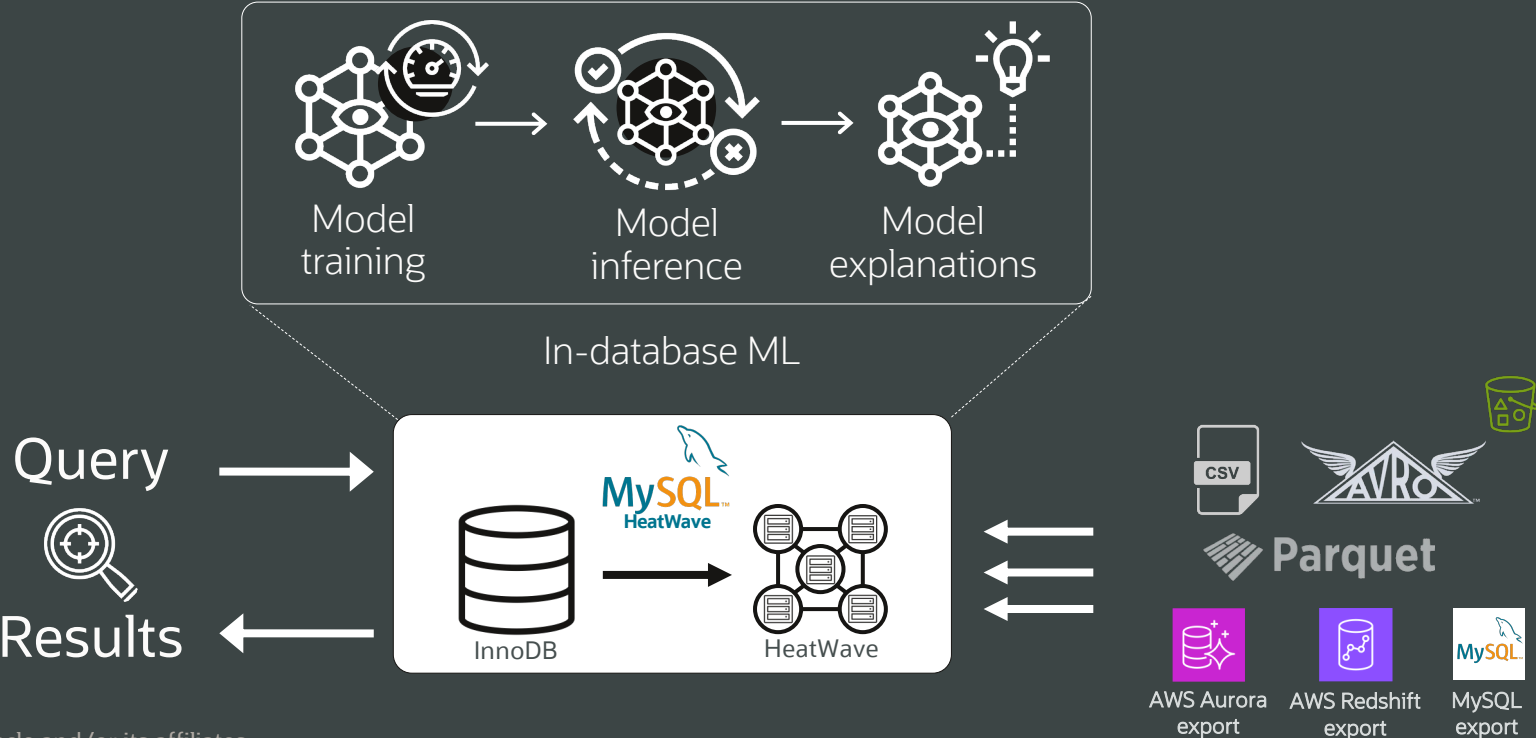
```
mysql>
CREATE TABLE `20newsgroups_train` (`text` LONGTEXT DEFAULT NULL, `category` VARCHAR(255)
DEFAULT NULL);
CREATE TABLE `20newsgroups_test` LIKE `20newsgroups_train`;

mysql-js>
util.importTable("20newsgroups_train.csv",{table: "20newsgroups_train", dialect:
"csv-unix", skipRows:1})
util.importTable("20newsgroups_test.csv",{table: "20newsgroups_test", dialect: "csv-unix",
skipRows:1})

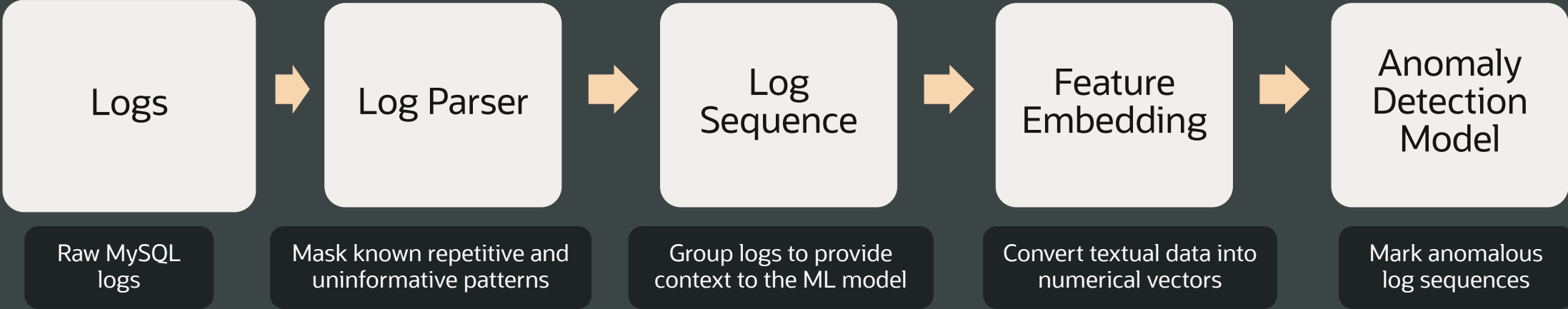
mysql> CALL sys.ML_TRAIN('mlcorpus.`20newsgroups_train`', 'category', NULL, @model);
mysql> CALL sys.ML_PREDICT_TABLE('mlcorpus.`20newsgroups_test`', @model, 'mlcorpus.
`20newsgroups_predictions`', NULL);
```

# Lakehouse support

- Build ML models on data in object store or the database
- Use the same APIs as used for data in the database
- No additional cost – reuse the same HeatWave Cluster



# MySQL log anomaly detection pipeline



# Generative AI

---



# Generative AI in HeatWave enables new use cases



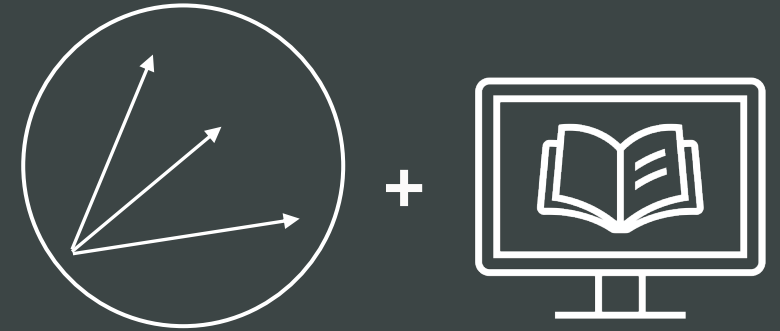
## Content generation & summarization

- Generate insights from enterprise documents
- Generate blogs from pdf instruction manuals
- Summarize logs



## Retrieval Augmented Generation (RAG)

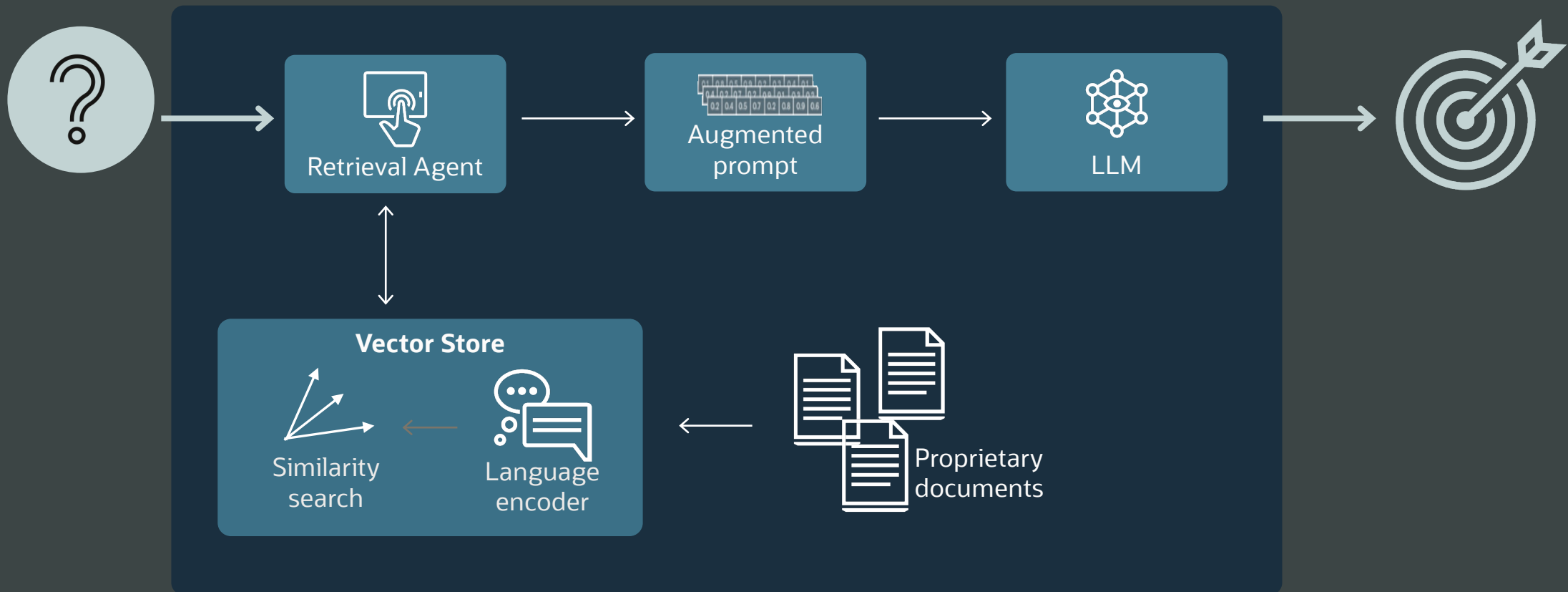
- Search on public and private enterprise data
- Search on unstructured data in vector store



## Natural language interaction

- Natural language interaction with unstructured data
- Content retrieval and response in natural language

# Vector store provides context to LLM for more relevant results

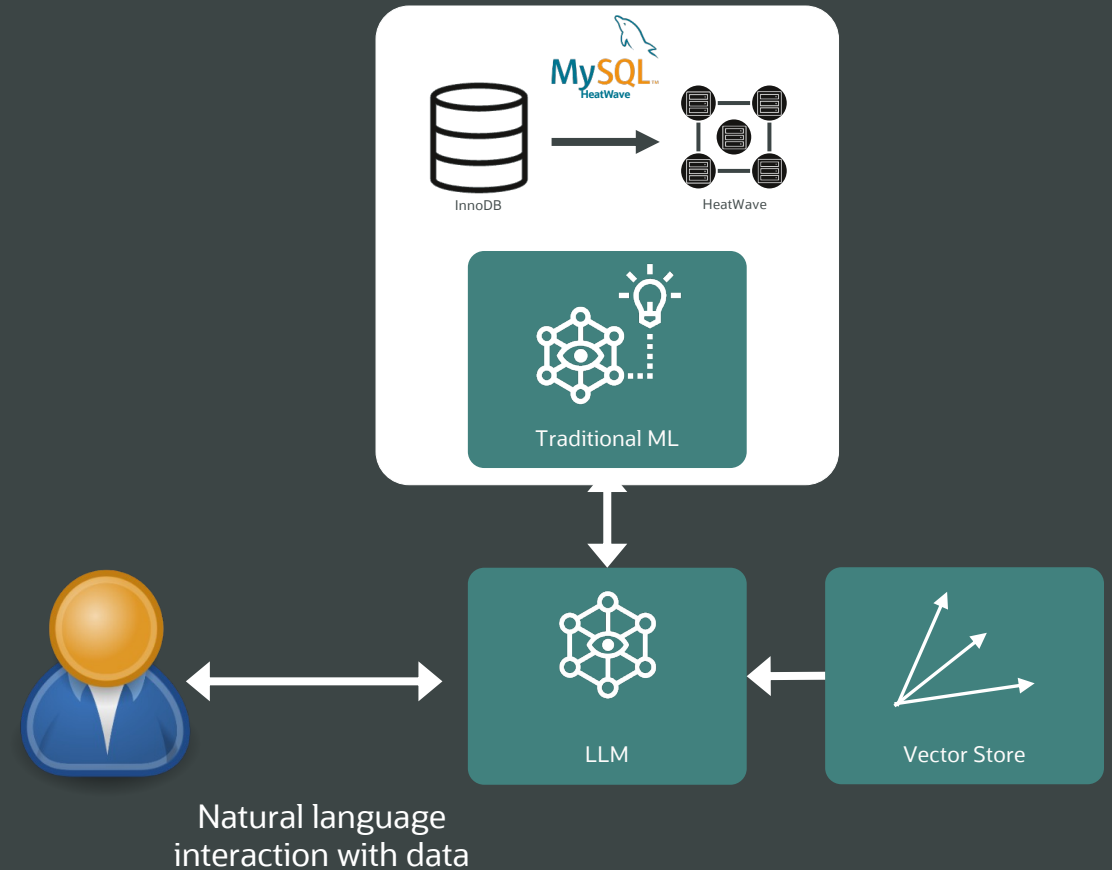


# Synergy of Generative AI and AutoML

## A differentiator in HeatWave

Multiple advantages of combining HeatWave AutoML with Generative AI:

- More accurate LLM results by filtering irrelevant data
- Faster LLM inference due to smaller search space







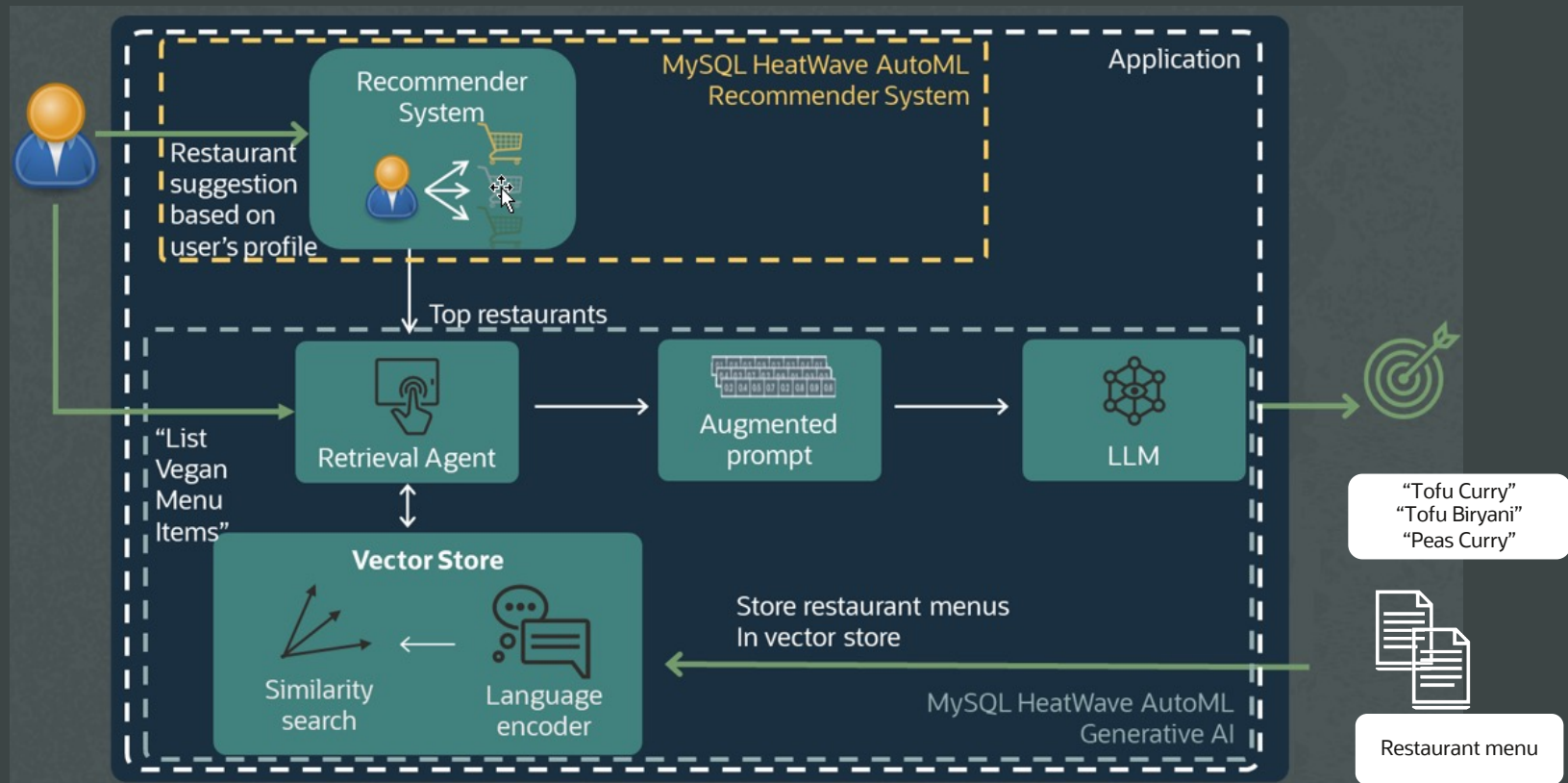
# Use case 2 : Personalization

## Recommender System + RAG



Recommend, Retrieve, and Generate descriptions of dishes based on user preference

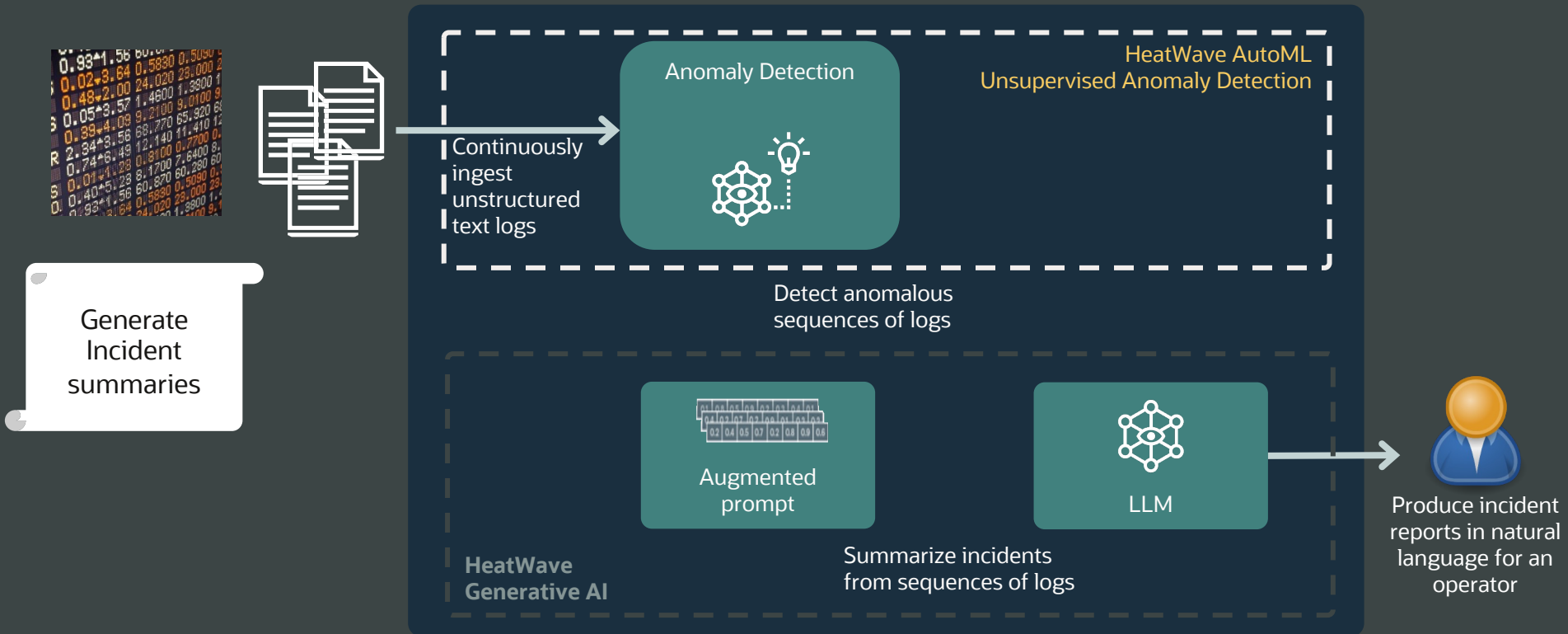
### Personalized Menu: HW AutoML + Retrieval Augmented Generation



# Use Case 3: Report Generation

## Anomaly Detection + Content generation

### Report Generation: HW AutoML + Summarization of anomalous logs



# Generator Example

## Incident Report

### Report Generation: Summarization of anomalous logs

```
{"analyticsClusterId":"","component":"ml_driver","level":"INFO","ts":1688777883620,"msg":"[8145:140643603412800] [fit_transform:140] [ 1.939] fit_transform: Remove non-categorical columns with > 20% missing values. Dropping 0 columns"}
{"analyticsClusterId":"","component":"ml_driver","level":"INFO","ts":1688777883805,"msg":"[8145:140643603412800] [fit_transform:147] [ 1.939] fit_transform: Remove constant columns. Dropping 0 columns"}
{"analyticsClusterId":"","component":"ml_driver","level":"INFO","ts":1688777964868,"msg":"[8145:140643603412800] [fit_transform:227] [ 173.397] fit_transform: After feature engineering and transformations. Updated shape : (157776, 72976)"}
{"analyticsClusterId":"","component":"ml_driver","level":"INFO","ts":1688777964868,"msg":"[8145:140643603412800] [fit_transform:231] [ 173.397] fit_transform: col_types_: ['text', 'text']"}
{"analyticsClusterId":"","component":"ml_driver","level":"INFO","ts":1688777964879,"msg":"[8145:140643603412800] [fit_transform:247] [ 173.397] fit_transform: class counts : 0 157494\n1 282\nName: label, dtype: int64"}
{"analyticsClusterId":"","component":"ml_driver","level":"INFO","ts":1688777975619,"msg":"[8145:140641670100736] [run:272] [ 249.292] (249.29190063476562, 249.29190063476562) (current,max) GB for process 8145 is NOT within the threshold 248.94811630249023 GB memory usage"}
{"analyticsClusterId":"","component":"ml_driver","level":"INFO","ts":1688777975619,"msg":"[8145:140641670100736] [run:279] [ 249.292] Sending SIGTERM to the main thread"}
{"analyticsClusterId":"","component":"ml_driver","level":"INFO","ts":1688777975725,"msg":"[8145:140641670100736] [run:272] [ 250.038] (250.03799438476562, 250.03799438476562) (current,max) GB for process 8145 is NOT within the threshold 248.19862747192383 GB memory usage"}
{"analyticsClusterId":"","component":"ml_driver","level":"INFO","ts":1688777975725,"msg":"[8145:140641670100736] [run:279] [ 250.038] Sending SIGTERM to the main thread"}
{"analyticsClusterId":"","component":"ml_driver","level":"INFO","ts":1688777975830,"msg":"[8145:140641670100736] [run:272] [ 250.782] (250.78213500976562, 250.78213500976562) (current,max) GB for process 8145 is NOT within the threshold 247.45304489135742 GB memory usage"}
{"analyticsClusterId":"","component":"ml_driver","level":"INFO","ts":1688777975830,"msg":"[8145:140641670100736] [run:279] [ 250.782] Sending SIGTERM to the main thread"}
{"analyticsClusterId":"","component":"ml_driver","level":"INFO","ts":1688777975936,"msg":"[8145:140641670100736] [run:272] [ 251.528] (251.52822875976562, 251.52822875976562) (current,max) GB for process 8145 is NOT within the threshold 246.70550918579102 GB memory usage"}
{"analyticsClusterId":"","component":"ml_driver","level":"INFO","ts":1688777975936,"msg":"[8145:140641670100736] [run:279] [ 251.528] Sending SIGTERM to the main thread"}
{"analyticsClusterId":"","component":"ml_driver","level":"INFO","ts":1688777976042,"msg":"[8145:140641670100736] [run:272] [ 252.274] (252.27432250976562, 252.27432250976562) (current,max) GB for process 8145 is NOT within the threshold 245.95797729492188 GB memory usage"}
{"analyticsClusterId":"","component":"ml_driver","level":"INFO","ts":1688777976042,"msg":"[8145:140641670100736] [run:279] [ 252.274] Sending SIGTERM to the main thread"}
{"analyticsClusterId":"","component":"ml_driver","level":"INFO","ts":1688777976148,"msg":"[8145:140641670100736] [run:272] [ 253.02] (253.02041625976562, 253.02041625976562) (current,max) GB for process 8145 is NOT within the threshold 245.21044158935547 GB memory usage"}
{"analyticsClusterId":"","component":"ml_driver","level":"INFO","ts":1688777976148,"msg":"[8145:140641670100736] [run:279] [ 253.02] Sending SIGTERM to the main thread"}
{"analyticsClusterId":"","component":"ml_driver","level":"INFO","ts":1688777976253,"msg":"[8145:140641670100736] [run:272] [ 253.767] (253.76651000976562, 253.76651000976562) (current,max) GB for process 8145 is NOT within the threshold 244.46290588378906 GB memory usage"}
{"analyticsClusterId":"","component":"ml_driver","level":"INFO","ts":1688777976253,"msg":"[8145:140641670100736] [run:279] [ 253.767] Sending SIGTERM to the main thread"}
{"analyticsClusterId":"","component":"ml_driver","level":"INFO","ts":1688777976359,"msg":"[8145:140641670100736] [run:272] [ 254.513] (254.51260375976562, 254.51260375976562) (current,max) GB for process 8145 is NOT within the threshold 243.71537017822266 GB memory usage"}
{"analyticsClusterId":"","component":"ml_driver","level":"INFO","ts":1688777976359,"msg":"[8145:140641670100736] [run:279] [ 254.513] Sending SIGTERM to the main thread"}
{"analyticsClusterId":"","component":"ml_driver","level":"INFO","ts":1688777976465,"msg":"[8145:140641670100736] [run:272] [ 255.261] (255.26065063476562, 255.26065063476562) (current,max) GB for process 8145 is NOT within the threshold 242.96588134765625 GB memory usage"}
{"analyticsClusterId":"","component":"ml_driver","level":"INFO","ts":1688777976465,"msg":"[8145:140641670100736] [run:279] [ 255.261] Sending SIGTERM to the main thread"}
```

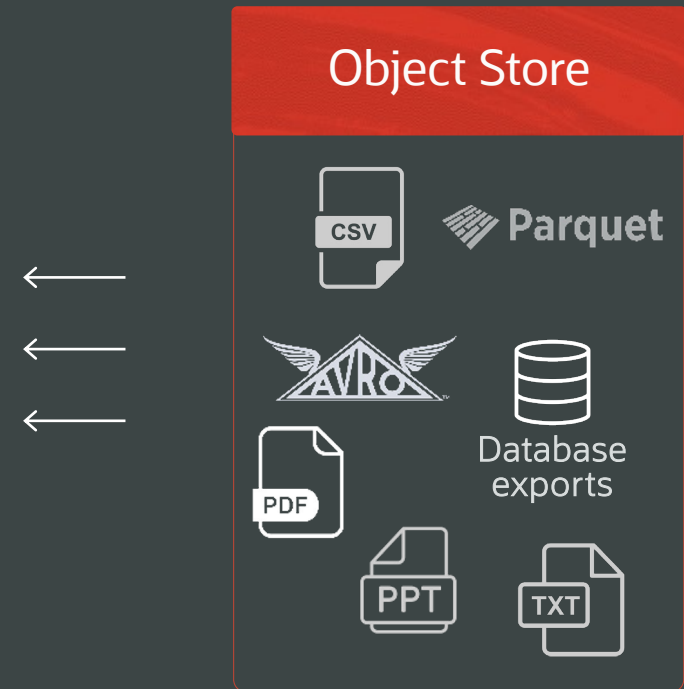
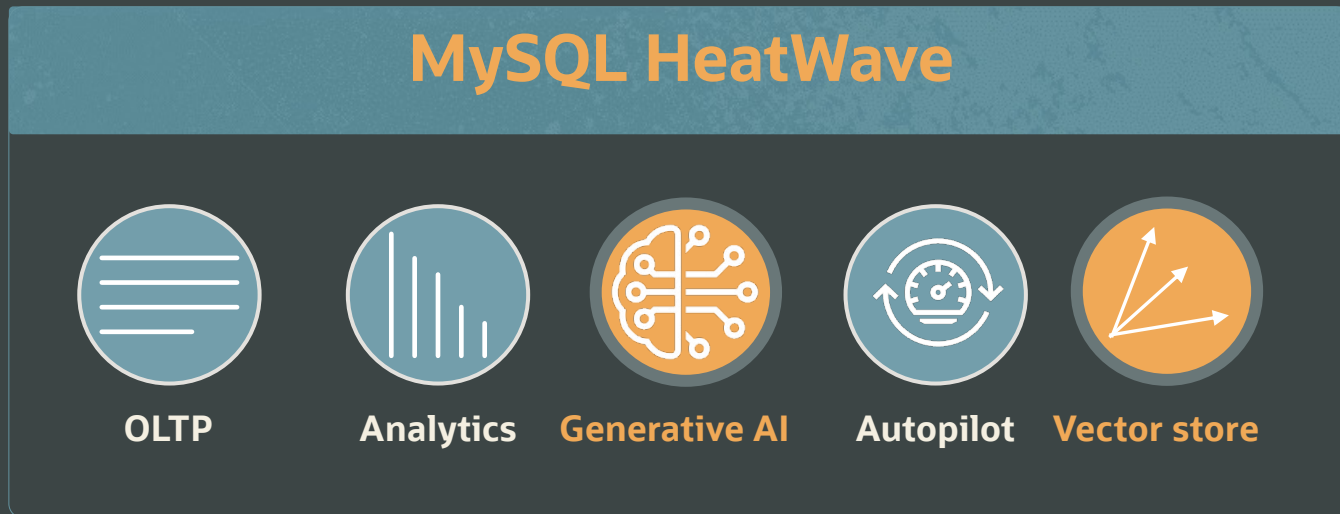
Ask the question (“What is the main problem in the following collection of logs. Provide a two-sentence summary.”)

“The main problem in this collection of logs is that the memory usage of a particular process (with ID 8145) is consistently exceeding the defined memory threshold. This issue is repeatedly highlighted in the logs, and SIGTERM signals are sent to the main thread to terminate the process when it exceeds the memory threshold. Would you like me to help you with anything else regarding this collection of logs?”



# Summary - Generative AI with MySQL HeatWave vector store

- Enables querying of unstructured documents
- Allows semantic search of content
- Users can query unstructured documents in natural language



ORACLE