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

**Classification**
- Player churn prediction
- Classify warranty claims
- Defective part identification
- Identify game hackers
- Predict when failure will occur
- IoT digital twin failure prediction
- Predict air pollution
- Return on advertising spend prediction
- Utilization demand forecasting

**Anomaly Detection**
- Detect anomalies in supplies
- Predict assembly line jam

**Recommender System**
- Identify similar users
- Recommend movies to viewers
- Suggest substitute products
- Recommend new products
- Loan default prediction
- Demand forecasting

**Regression**
- Predict flight delay
- Loan amount prediction
- Rain fall amount prediction

**Timeseries Forecasting**
- Identify similar users
- Recommend movies to viewers
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
HeatWave AutoML now supports text columns in the database
Text columns are converted into numerical vectors and treated as regular features
Text can of type – TINYTEXT, TEXT, MEDIUMTEXT or LONGTEXT

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

mysql>
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

1. Logs: Raw MySQL logs
2. Log Parser: Mask known repetitive and uninformative patterns
3. Log Sequence: Group logs to provide context to the ML model
4. Feature Embedding: Convert textual data into numerical vectors
5. Anomaly Detection Model: Mark anomalous log sequences
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 1: Employee Assistant

Vector store + Natural language interface

Employee Assistant: Improve employee productivity

User’s PDF Contracts in Object store

Vector store ingest

Retrieval Augmented Generation

**Use Case 1: Employee Assistant**

**Vector store + Natural language interface**

**Employee Assistant: Improve employee productivity**

- **User’s PDF Contracts in Object store**
- **Vector store ingest**
- **Retrieval Augmented Generation**
Use case 2: Personalization

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

- Anomaly Detection
  - Detect anomalous sequences of logs
  - Unsupervised Anomaly Detection

- HeatWave
  - Generative AI
  - Continuously ingest unstructured text logs

- Augmented prompt

- LLM
  - Summarize incidents from sequences of logs

- Produce incident reports in natural language for an operator

- Generate Incident summaries
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