



# AI Governance AI asset report

---

Car rental review satisfaction  
analysis

Dec 17, 2023

# AI use case

|          |                             |
|----------|-----------------------------|
| Name     | Car rental review           |
| Status   | In_operation                |
| Approach | Generation with IBM Granite |
| Version  | 0.0.3                       |

## Foundation Model

|                  |                         |
|------------------|-------------------------|
| Foundation Model | granite-13b-instruct-v2 |
| Publisher        | IBM                     |

## Prompt Template

|                      |   |
|----------------------|---|
| Prompt Template name | Car rental review satisfaction analysis |
| Prompt Template ID   | ba4ce1ca-f7af-4544-8879-5c2a8e6a88e4    |
| Task type            | <div>Classification</div>               |
| Created by           | Shuvanker Ghosh (sghosh@us.ibm.com)     |
| Date                 | Dec 14 2023, 19:44 PM GMT               |

## Prompt

Determine if the customer was satisfied with the experience based on the comment. Return simple yes or no.

|           |   |
|-----------|---|
| Review:   | I thought the representative handled the initial situation badly. The company was out of cars, with none coming in that day. Then the representative tried to find us a car at another franchise. There they were successful. |
| Satisfied | no  |

|           |  |
|-----------|--|
| Review:   | The car rental agent was very helpful. He worked with us to get a bigger car for very little additional fees |
| Satisfied | yes  |

Input to test:  
{Review}

Variables

Review

|               |                 |
|---------------|-----------------|
| Default Value | Agent was nasty |
|---------------|-----------------|

Prompt parameters

Decoding

|                    |        |
|--------------------|--------|
| Decoding method    | Greedy |
| Temperature        | 0.7    |
| Top P              | 1      |
| Top K              | 50     |
| Repetition Penalty | 1      |
| Random seed        | None   |

Stopping criteria

|                |      |
|----------------|------|
| Stop sequences | None |
|----------------|------|

|            |    |
|------------|----|
| Min tokens | 0  |
| Max tokens | 20 |

# Evaluation

## Development

Project name: **Model Lifecycle Management with Watsonx Demo**

|            |   |
|------------|---|
| Project ID | 26329aec-3a9c-4399-ad0a-8316c6729832    |
| Created by | Shuvanker Ghosh (sghosh@us.ibm.com)     |
| Date       | Dec 04 2023, 12:59 PM GMT               |
| Asset name | Car rental review satisfaction analysis |

## Validate

Space name: **Gen AI Model Deployments for Validation**

|                       |   |
|-----------------------|---|
| Space ID              | 8a4a457f-aa96-4dbc-a1dd-15bae050c340    |
| Created by            | Shuvanker Ghosh (sghosh@us.ibm.com)     |
| Date                  | Dec 05 2023, 17:43 PM GMT               |
| Asset name            | Car rental review satisfaction analysis |
| Number of deployments | 1                                       |

## Evaluation results

# Deployment #1: Car rental review satisfaction analysis validation deployment

|               |                                      |
|---------------|--------------------------------------|
| Deployment ID | aca1553f-90bd-4a6e-b7c4-e2a0602d2cd0 |
| Created       | Dec 13 2023, 14:00 PM GMT            |
| Modified      | Dec 17 2023, 14:55 PM GMT            |
| Type          | Online                               |

## Scoring endpoints

<https://us-south.ml.cloud.ibm.com/ml/v1-beta/deployments/aca1553f-90bd-4a6e-b7c4-e2a0602d2cd0/generation/text> ,  
[https://us-south.ml.cloud.ibm.com/ml/v1-beta/deployments/aca1553f-90bd-4a6e-b7c4-e2a0602d2cd0/generation/text\\_stream](https://us-south.ml.cloud.ibm.com/ml/v1-beta/deployments/aca1553f-90bd-4a6e-b7c4-e2a0602d2cd0/generation/text_stream)

## Evaluation Activity

|  |   |
|--|---|
| OpenScale instance ID                                      | cc5e126e-17ee-493d-bff5-697cb87fd277  |
| OpenScale service provider                                 | SPACE_8a4a457f-aa96-4dbc-a1dd-15bae050c340  |
| Last evaluation  | Dec 15 2023, 04:35 AM GMT   |
| Evaluation data  | car_rental_review_payload_dataset.csv   |
| Evaluation data type                                       | Offline   |
| <div>&lt;h4&gt; Feedback data/ Test data &lt;/h4&gt;</div> |   |
| Feedback data location                                     | db2   |
| Database   | bludb   |
| Schema   | aiopenscale00   |
| Table  | Feedback_48a55460-f552-48a0-a49d-8dee66bd87fb   |
| Host   | 24013820-da9f-4e69-a7d6-17efc5d30a33.c8l9ggds0kmvoig3l8kg.databases.<br>appdomain.cloud |
| Port   | 31,613  |

## Payload load/ Transaction storage

|                      |               |
|----------------------|---------------|
| Payload log location | db2           |
| Database             | bludb         |
| Schema               | aiopenscale00 |

|       |   |
|-------|---|
| Table | Payload_48a55460-f552-48a0-a49d-8dee66bd87fb  |
| Host  | 24013820-da9f-4e69-a7d6-17efc5d30a33.c8l9ggsd0kmvoig3l8kg.databases.appdomain.cloud |
| Port  | 31,613  |

### Quality evaluation

|                   |                           |
|-------------------|---------------------------|
| Breach status     | RED                       |
| Records evaluated | 486                       |
| Last evaluation   | Dec 15 2023, 04:35 AM GMT |

| Quality Metric                      | Value  |
|-------------------------------------|--------|
| Accuracy                            | 0      |
| Weighted True Positive Rate (wTPR)  | 0      |
| Weighted False Positive Rate (wFPR) | 0      |
| Weighted recall                     | 0      |
| Weighted precision                  | 0      |
| Weighted F1-Measure                 | 0      |
| Label skew                          | -0.257 |
| Matthews correlation coefficient    | 0      |

### Model Health evaluation

|                 |                           |
|-----------------|---------------------------|
| Breach status   | RED                       |
| Last evaluation | Dec 15 2023, 04:36 AM GMT |

| Model Health Metric       | Value |
|---------------------------|-------|
| Total records             | 486   |
| Maximum record throughput | 4.63  |
| Median API throughput     | 3.65  |

| Model Health Metric        | Value  |
|----------------------------|--------|
| Median input token count   | 114    |
| Average output token count | 3      |
| Minimum API throughput     | 0.503  |
| Median record latency      | 274    |
| Minimum output token count | 3      |
| Users                      | 1      |
| Maximum API latency        | 1,988  |
| Maximum API throughput     | 4.63   |
| Total scoring requests     | 486    |
| Median records             | 1      |
| Median record throughput   | 3.65   |
| Minimum API latency        | 216    |
| Average input token count  | 114    |
| Average record throughput  | 2.949  |
| Maximum input token count  | 114    |
| Median output token count  | 3      |
| Total output token count   | 1,458  |
| Maximum records            | 1      |
| Minimum record latency     | 216    |
| Average record latency     | 339.06 |
| Median API latency         | 274    |
| Minimum records            | 1      |
| Minimum input token count  | 114    |
| Average API throughput     | 2.949  |
| Maximum record latency     | 1,988  |
| Average API latency        | 339.06 |
| Total input token count    | 55,404 |
| Minimum record throughput  | 0.503  |
| Maximum output token count | 3      |

| Model Health Metric | Value |
|---------------------|-------|
| Average records     | 1     |

## Operate

### Space name: Gen AI Model Deployments for Production

|                       |   |
|-----------------------|---|
| Space ID              | 566d4549-253a-489c-8173-d3a9c92bb027    |
| Created by            | Shuvanker Ghosh (sghosh@us.ibm.com)     |
| Date                  | Dec 14 2023, 19:42 PM GMT               |
| Asset name            | Car rental review satisfaction analysis |
| Number of deployments | 1                                       |

## Evaluation results

### Deployment #1: Car rental review satisfaction analysis production deployment

|               |                                      |
|---------------|--------------------------------------|
| Deployment ID | 22905dc9-3d72-427a-91d2-8e3ddbae8b3b |
| Created       | Dec 14 2023, 19:48 PM GMT            |
| Modified      | Dec 17 2023, 14:55 PM GMT            |
| Type          | Online                               |

## Scoring endpoints

<https://us-south.ml.cloud.ibm.com/ml/v1-beta/deployments/22905dc9-3d72-427a-91d2-8e3ddbae8b3b/generation/text> ,  
[https://us-south.ml.cloud.ibm.com/ml/v1-beta/deployments/22905dc9-3d72-427a-91d2-8e3ddbae8b3b/generation/text\\_stream](https://us-south.ml.cloud.ibm.com/ml/v1-beta/deployments/22905dc9-3d72-427a-91d2-8e3ddbae8b3b/generation/text_stream)

## Evaluation Activity

|                            |  |
|----------------------------|--|
| OpenScale instance ID      | cc5e126e-17ee-493d-bff5-697cb87fd277       |
| OpenScale service provider | SPACE_566d4549-253a-489c-8173-d3a9c92bb027 |



|                        |   |
|------------------------|---|
| Feedback data location | db2   |
| Database               | bludb   |
| Schema                 | aiopenscale00   |
| Table                  | Feedback_eaf6cb14-80d9-41c5-a919-0b7573f459af   |
| Host                   | 24013820-da9f-4e69-a7d6-17efc5d30a33.c8l9ggds0kmvoig3l8kg.databases.<br>appdomain.cloud |
| Port                   | 31,613  |

## Payload load/ Transaction storage

|                      |   |
|----------------------|---|
| Payload log location | db2   |
| Database             | bludb   |
| Schema               | aiopenscale00   |
| Table                | Payload_eaf6cb14-80d9-41c5-a919-0b7573f459af  |
| Host                 | 24013820-da9f-4e69-a7d6-17efc5d30a33.c8l9ggds0kmvoig3l8kg.databases.<br>appdomain.cloud |
| Port                 | 31,613  |

## Quality evaluation

|                   |                           |
|-------------------|---------------------------|
| Breach status     | GREEN                     |
| Records evaluated | 486                       |
| Last evaluation   | Dec 15 2023, 22:30 PM GMT |

| Quality Metric                      | Value |
|-------------------------------------|-------|
| Accuracy                            | 0.994 |
| Weighted True Positive Rate (wTPR)  | 0.994 |
| Weighted False Positive Rate (wFPR) | 0.007 |
| Weighted recall                     | 0.994 |
| Weighted precision                  | 0.994 |

| Quality Metric                   | Value  |
|----------------------------------|--------|
| Weighted F1-Measure              | 0.994  |
| Label skew                       | -0.266 |
| Matthews correlation coefficient | 0.987  |

Drift v2 evaluation

|                   |                           |
|-------------------|---------------------------|
| Records evaluated | 486                       |
| Last evaluation   | Dec 15 2023, 22:32 PM GMT |

| Drift v2 Metric       | Value |
|-----------------------|-------|
| Output metadata drift | 0.149 |
| Input metadata drift  | 0.151 |
| Prediction drift      | 0     |

Model Health evaluation

|                 |                           |
|-----------------|---------------------------|
| Last evaluation | Dec 17 2023, 04:56 AM GMT |
|-----------------|---------------------------|

| Model Health Metric        | Value     |
|----------------------------|-----------|
| Total records              | 3         |
| Maximum record throughput  | 3,964.939 |
| Median API throughput      | 642.557   |
| Median input token count   | 117       |
| Average output token count | 3         |
| Minimum API throughput     | 371.72    |
| Median record latency      | 1.556     |
| Minimum output token count | 3         |
| Users                      | 1         |
| Maximum API latency        | 2.69      |
| Maximum API throughput     | 3,964.939 |

| Model Health Metric        | Value   |
|----------------------------|---------|
| Total scoring requests     | 3       |
| Median records             | 1       |
| Median record throughput   | 642.557 |
| Minimum API latency        | 0.252   |
| Average input token count  | 117     |
| Average record throughput  | 666.861 |
| Maximum input token count  | 117     |
| Median output token count  | 3       |
| Total output token count   | 9       |
| Maximum records            | 1       |
| Minimum record latency     | 0.252   |
| Average record latency     | 1.5     |
| Median API latency         | 1.556   |
| Minimum records            | 1       |
| Minimum input token count  | 117     |
| Average API throughput     | 666.861 |
| Maximum record latency     | 2.69    |
| Average API latency        | 1.5     |
| Total input token count    | 351     |
| Minimum record throughput  | 371.72  |
| Maximum output token count | 3       |
| Average records            | 1       |