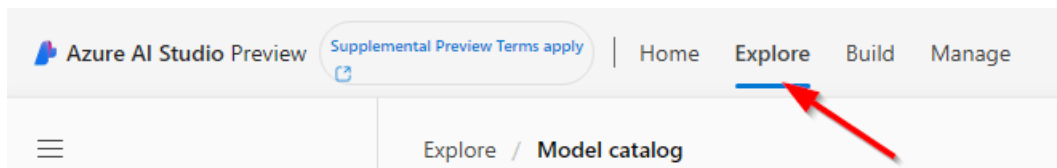


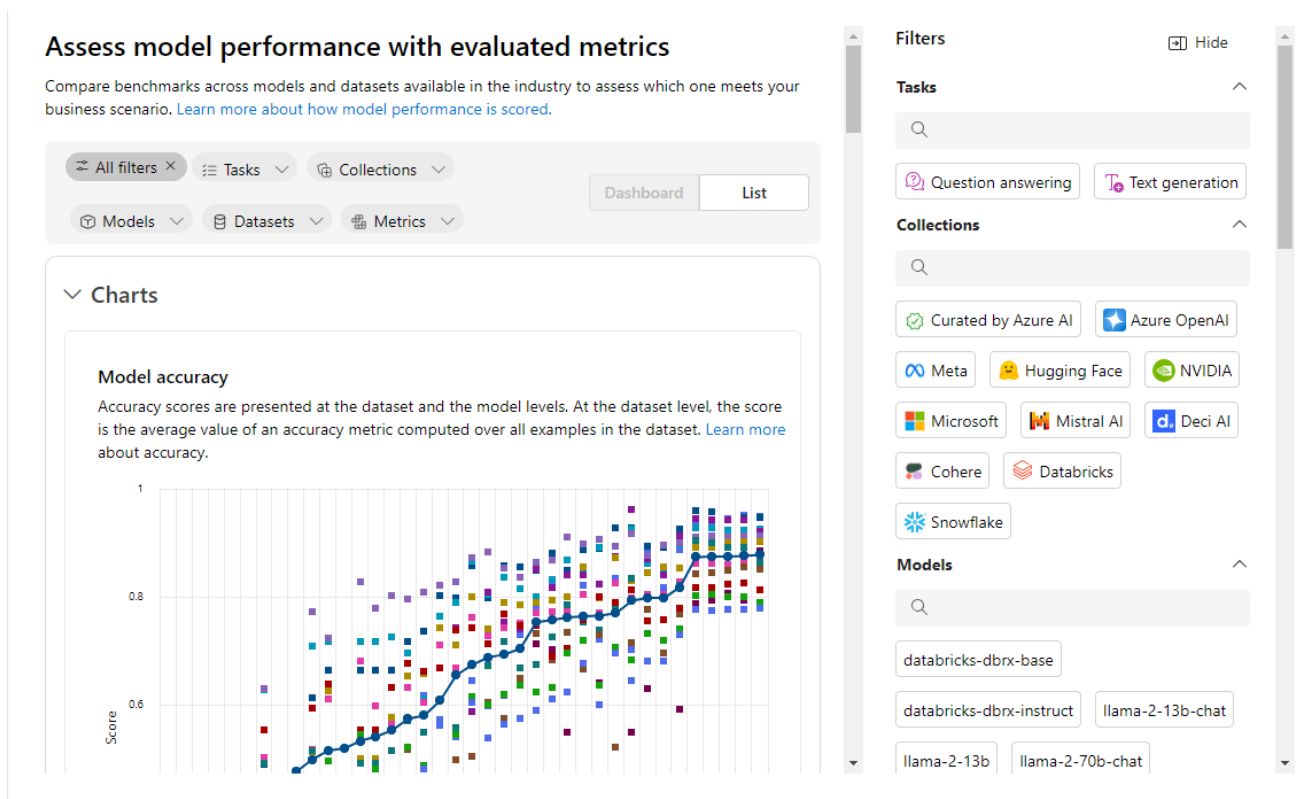
Exercise 4- AI Studio Explore

Model benchmarks

1. Click on the Explore tab in the navigation



2. Click on the Model benchmarks link in the left hand navigation



3. Scroll down in the middle section to look at the **Model accuracy** section. You will notice that the **gpt-4-turbo-2024-04-09** is the most accurate model but all of the gpt-4 models are fairly close.

Assess model performance with evaluated metrics

Compare benchmarks across models and datasets available in the industry to assess which one meets your business scenario. [Learn more about how model performance is scored.](#)

Filters: All filters x Tasks Collections

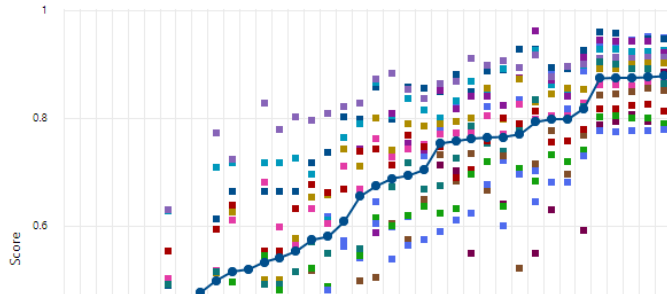
Models Datasets Metrics

Dashboard List

Charts

Model accuracy

Accuracy scores are presented at the dataset and the model levels. At the dataset level, the score is the average value of an accuracy metric computed over all examples in the dataset. [Learn more about accuracy.](#)



Filters

Hide

Tasks

Q

Question answering

Text generation

Collections

Q

Curated by Azure AI

Azure OpenAI

Meta

Hugging Face

NVIDIA

Microsoft

Mistral AI

Deci AI

Cohere

Databricks

Snowflake

Models

Q

databricks-dbrx-base

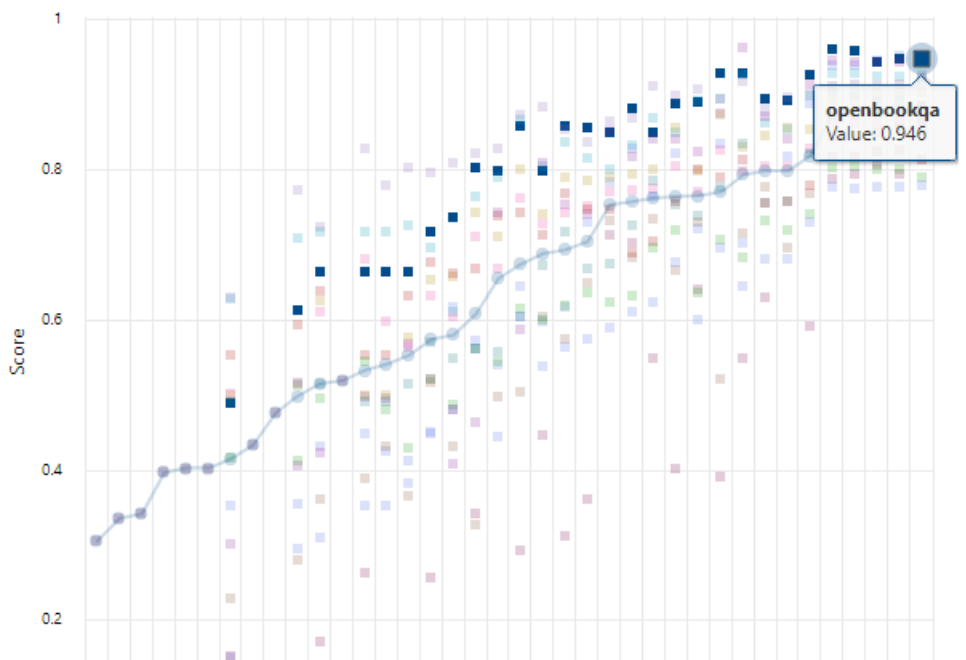
databricks-dbrx-instruct

llama-2-13b-chat

llama-2-13b

llama-2-70b-chat

4. Using your mouse mouse-over the top blue squares till you see the **openbookqa** dataset highlight about accuracy.



5. In the far right pain you can filter on the following:

- Tasks
- Collections
- Models
- Inference tasks

- Fine-tuning tasks
- Licenses