**Lacak eksperimen Azure Databricks dalam Azure Machine Learning**

Azure Machine Learning adalah platform cloud yang skalabel untuk pelatihan, penerapan, dan pengelolaan solusi pembelajaran mesin.

**Tujuan pembelajaran**

Setelah menyelesaikan modul ini, Anda akan dapat:

* Jelaskan Azure Machine Learning
* Jalankan eksperimen Azure Databricks di Azure Machine Learning
* Mencatat metrik dalam Azure Machine Learning dengan MLflow
* Jalankan pipeline Azure Machine Learning pada komputasi Azure Databricks

1. **Pengantar**

Jika Anda memilih untuk melatih model menggunakan Azure Databricks dan melacak pekerjaan Anda menggunakan MLflow, Anda dapat menambahkan integrasi dengan Azure Machine Learning untuk menyimpan metrik dan artefak pelatihan model dan menyimpan gambaran umum yang jelas tentang pekerjaan Anda. Menggunakan Azure Machine Learning sebagai backend untuk eksperimen MLflow Anda yang berjalan di komputasi Azure Databricks memberi Anda manfaat memiliki ruang kerja yang terpusat dan skalabel di mana Anda dapat mengakses semua aset Anda untuk menjalankan eksperimen atau meninjaunya. Dalam modul ini, Anda akan belajar tentang integrasi antara semua produk ini dan bagaimana Anda dapat mengelola pekerjaan Anda dari ruang kerja Azure Machine Learning.

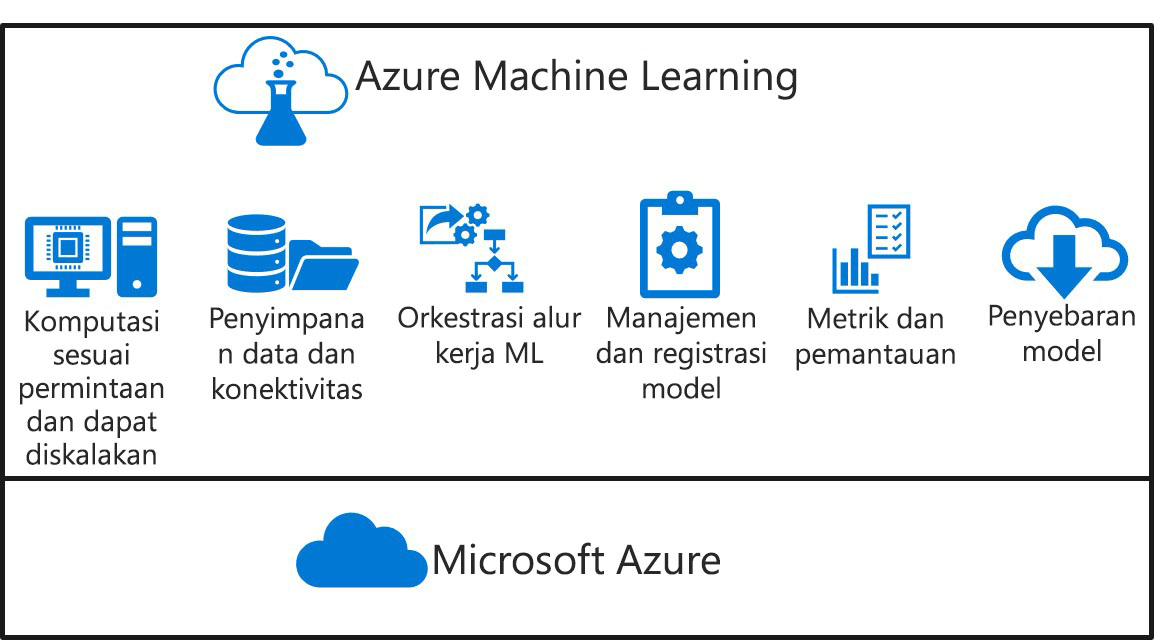
1. **Tujuan pembelajaran**

Setelah menyelesaikan modul ini, Anda akan dapat:

* Jelaskan Azure Machine Learning.
* Jalankan eksperimen.
* Mencatat metrik dengan MLflow.
* Jalankan Pipeline Step pada Databricks Compute.

1. **Jelaskan Azure Machine Learning**

Azure Machine Learning adalah platform untuk mengoperasikan beban kerja pembelajaran mesin di cloud.



Dibangun pada platform cloud Microsoft Azure, Azure Machine Learning memungkinkan Anda untuk mengelola:

* Komputasi sesuai permintaan yang skalabel untuk beban kerja pembelajaran mesin.
* Penyimpanan data dan konektivitas untuk menyerap data dari berbagai sumber.
* Orkestrasi alur kerja pembelajaran mesin untuk mengotomatiskan pelatihan model, penyebaran, dan proses manajemen.
* Registrasi dan pengelolaan model, sehingga Anda dapat melacak beberapa versi model dan data tempat mereka dilatih.
* Metrik dan pemantauan untuk eksperimen pelatihan, himpunan data, dan layanan yang dipublikasikan.
* Penyebaran model untuk inferensi real time dan batch.

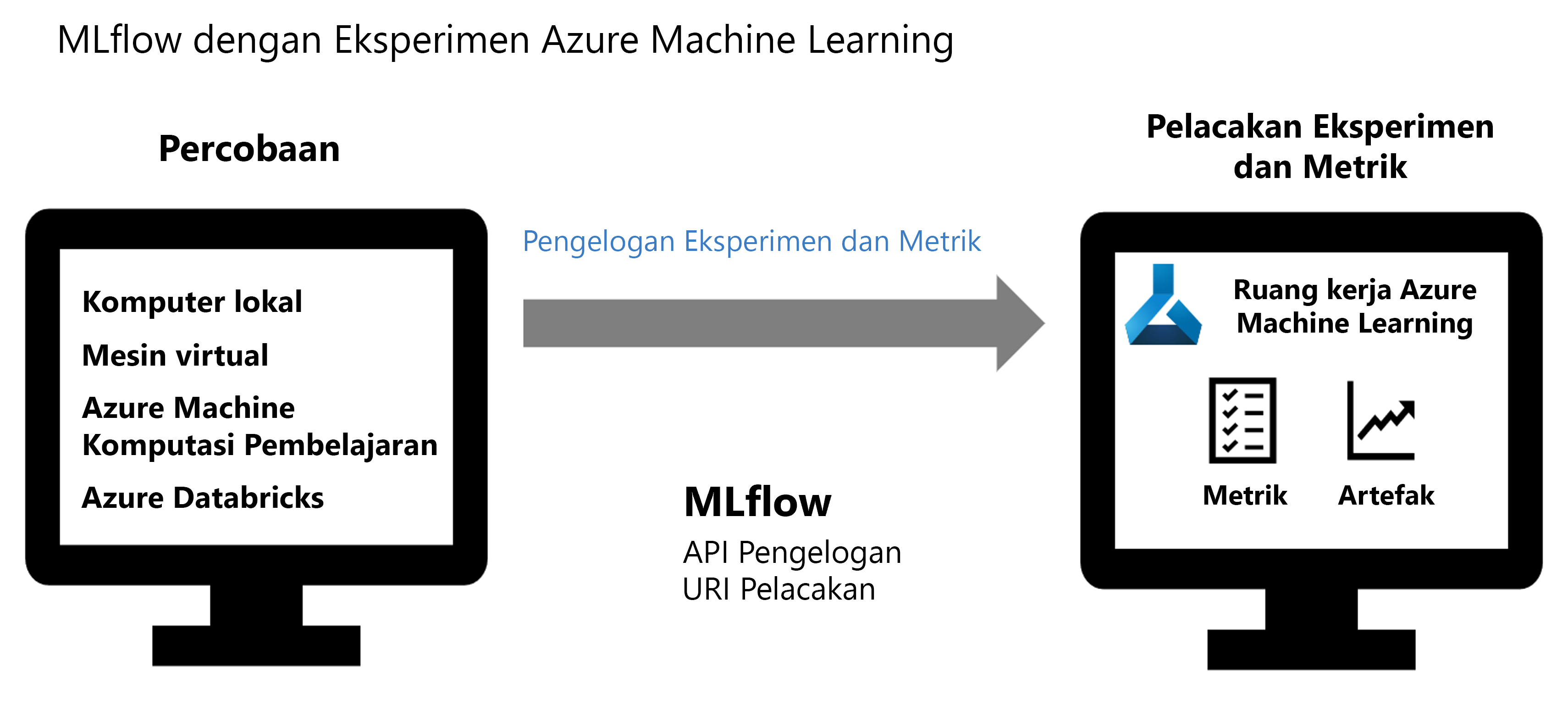
1. **Jalankan eksperimen Azure Databricks di Azure Machine Learning**

**MLflow** ( <https://www.mlflow.org/> ) adalah pustaka sumber terbuka untuk mengelola siklus hidup eksperimen pembelajaran mesin.

**Pelacakan MLFlow** ( <https://mlflow.org/docs/latest/quickstart.html#using-the-tracking-api> ) adalah komponen MLflow yang mencatat dan melacak metrik dan artefak model pelatihan Anda, apa pun lingkungan eksperimen Anda.

Pendekatan yang disarankan untuk menjalankan Eksperimen Azure Machine Learning (AML) di kluster Azure Databricks adalah dengan menggunakan Pelacakan MLflow dan menghubungkan Azure Machine Learning sebagai backend untuk eksperimen MLflow.

Diagram berikut mengilustrasikan bahwa dengan MLflow Tracking, Anda melacak metrik eksekusi eksperimen dan menyimpan artefak model di ruang kerja Azure Machine Learning Anda.



1. **Lacak Eksperimen AML di Azure Databricks**

Saat menjalankan eksperimen AML di Azure Databricks, ada tiga langkah utama:

1. Konfigurasikan URI pelacakan MLflow untuk menggunakan AML.
2. Konfigurasikan eksperimen MLflow.
3. Jalankan eksperimen Anda.
4. **Konfigurasikan URI pelacakan MLflow untuk menggunakan AML**

Untuk mengonfigurasi Pelacakan MLflow dan menghubungkan Azure Machine Learning sebagai backend untuk eksperimen MLFlow, Anda harus mengikuti langkah-langkah ini seperti yang ditunjukkan dalam cuplikan kode:

* Dapatkan objek ruang kerja AML Anda.
* Dari objek ruang kerja AML Anda, dapatkan alamat URI pelacakan unik.
* Siapkan URI pelacakan MLflow agar mengarah ke ruang kerja AML.



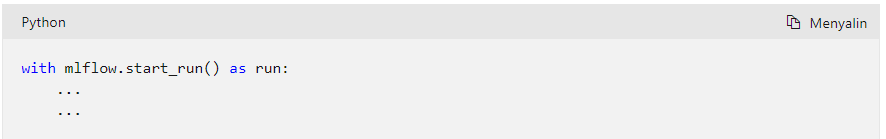
1. **Konfigurasikan eksperimen MLflow**

Berikan nama untuk eksperimen MLflow seperti yang ditunjukkan di bagian ini. Perhatikan bahwa nama eksperimen yang sama akan muncul di Azure Machine Learning.



1. **Jalankan eksperimen Anda**

Setelah eksperimen disiapkan, Anda dapat memulai latihan dengan **start\_run()** seperti yang ditunjukkan di bagian ini:

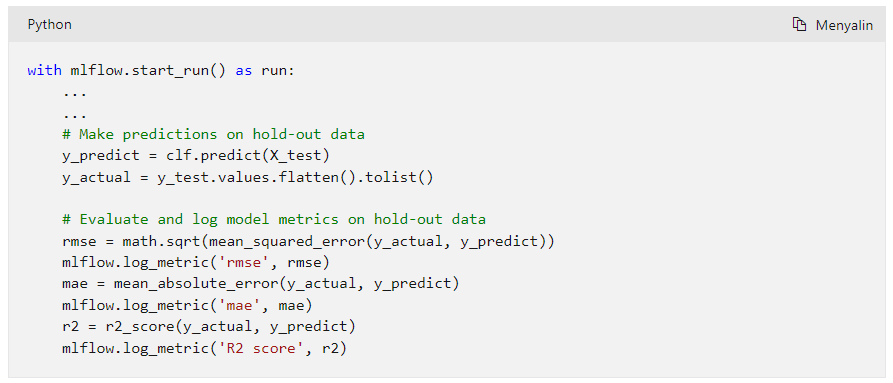


Pelatihan model dan kode logging Anda disediakan dalam blok **with**.

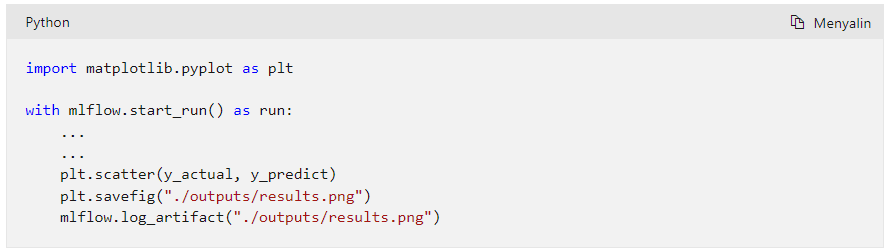
**4. Mencatat metrik dalam Azure Machine Learning dengan MLflow**

Di unit sebelumnya, kami membahas cara menyiapkan Azure Machine Learning sebagai backend untuk eksperimen MLflow. Kami juga melihat cara memulai pelatihan model Anda di Azure Databricks sebagai eksperimen MLflow. Di bagian ini, kita akan melihat cara membuat log metrik dan artefak model ke MLflow logging API. Metrik dan artefak yang dicatat ini kemudian ditangkap di ruang kerja Azure Machine Learning yang menyediakan lokasi terpusat, aman, dan skalabel untuk menyimpan metrik dan artefak pelatihan.

Dalam eksperimen MLflow, setelah Anda melatih dan mengevaluasi model, Anda dapat menggunakan MLflow logging API, mlflow.log\_metric(), untuk mulai mencatat metrik model seperti yang ditunjukkan di bagian ini:

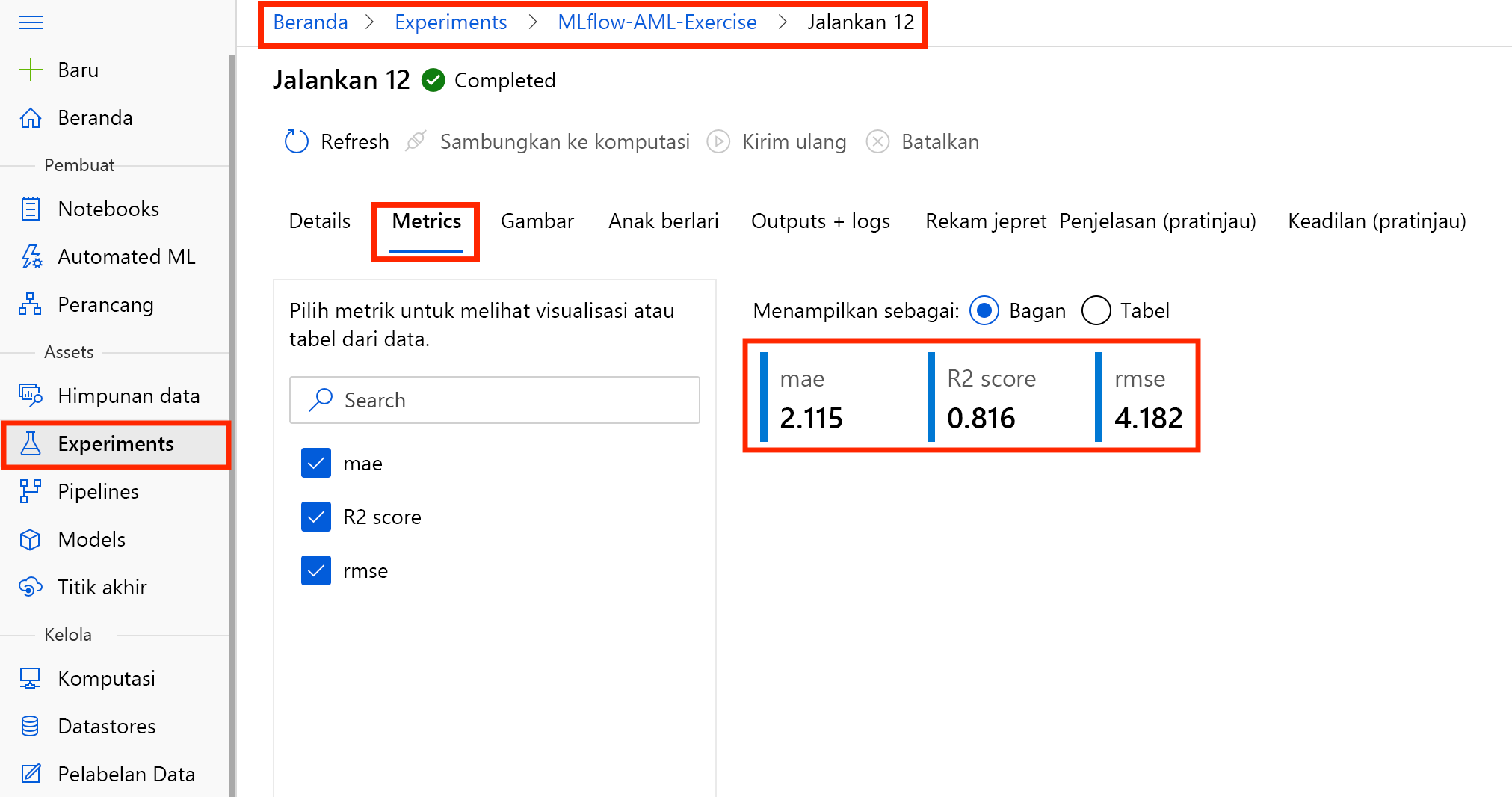


Selanjutnya, Anda dapat menggunakan **log\_artifact()** API MLflow untuk menyimpan artefak model seperti kurva **Predicted vs True** seperti yang ditunjukkan:

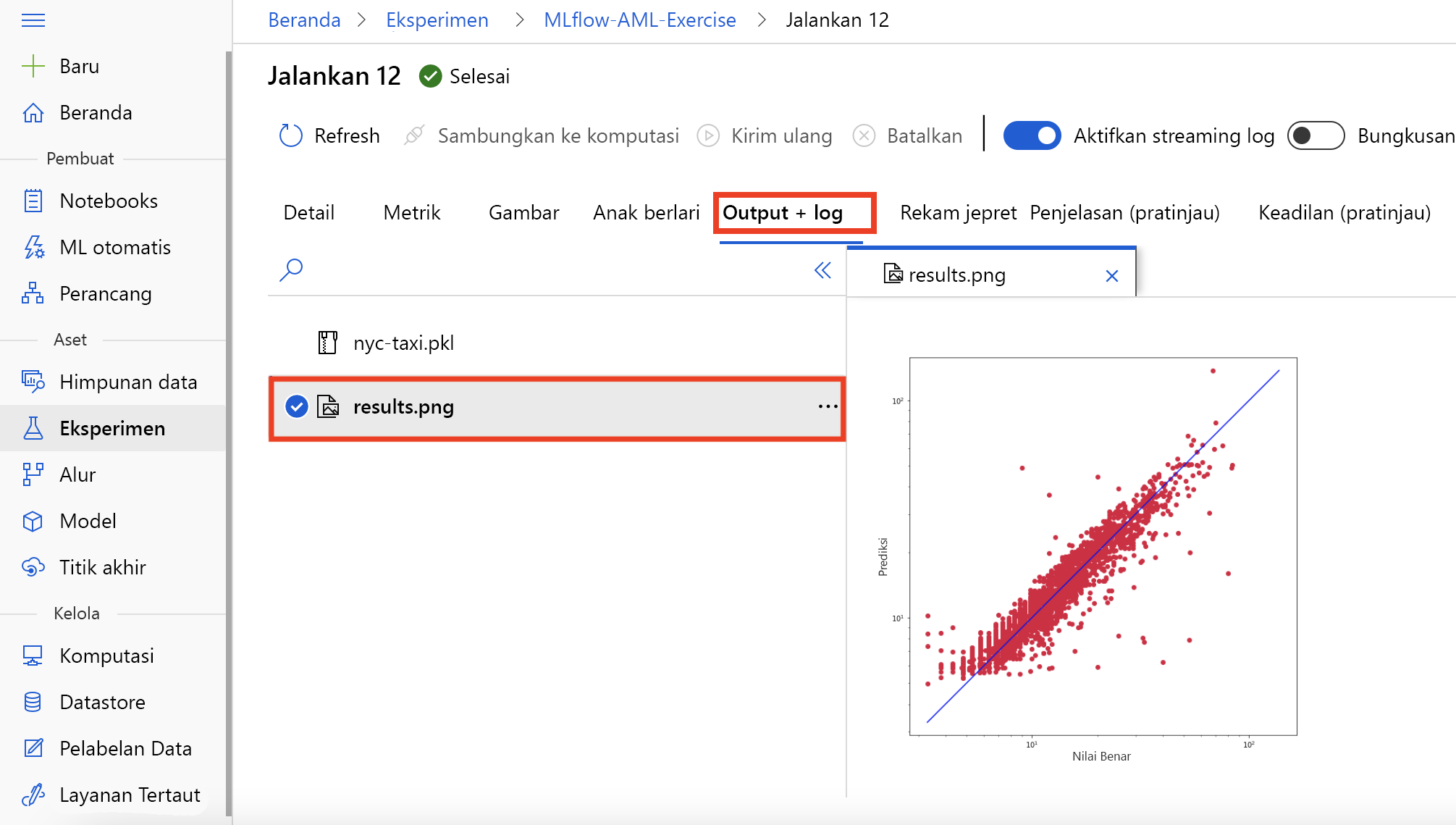


1. **Meninjau metrik dan artefak eksperimen di Azure ML Studio**

Karena Azure Machine Learning disiapkan sebagai backend untuk eksperimen MLflow, Anda dapat meninjau semua metrik dan artefak pelatihan dari dalam Studio Azure Machine Learning. Dari dalam studio, navigasikan ke tab **Experiments**, dan buka eksperimen yang sesuai dengan eksperimen MLflow. Di tab **Metrics** saat menjalankan, Anda akan mengamati metrik model yang dicatat melalui API pelacakan MLflow.



Selanjutnya, saat Anda membuka tab **Outputs + logs**, Anda akan mengamati artefak model yang dicatat melalui API pelacakan MLflow.



Singkatnya, menggunakan integrasi MLflow dengan Azure Machine Learning, Anda dapat menjalankan eksperimen di Azure Databricks dan memanfaatkan kemampuan ruang kerja Azure Machine Learning dari solusi terpusat, aman, dan skalabel untuk menyimpan metrik dan artefak pelatihan model.

1. **Jalankan pipeline Azure Machine Learning pada komputasi Azure Databricks**

Azure Machine Learning mendukung berbagai jenis komputasi untuk eksperimen dan pelatihan. Secara khusus, Anda dapat menjalankan **pipeline Azure Machine Learning** pada komputasi Databricks.

1. **Apa itu pipeline Azure Machine Learning?**

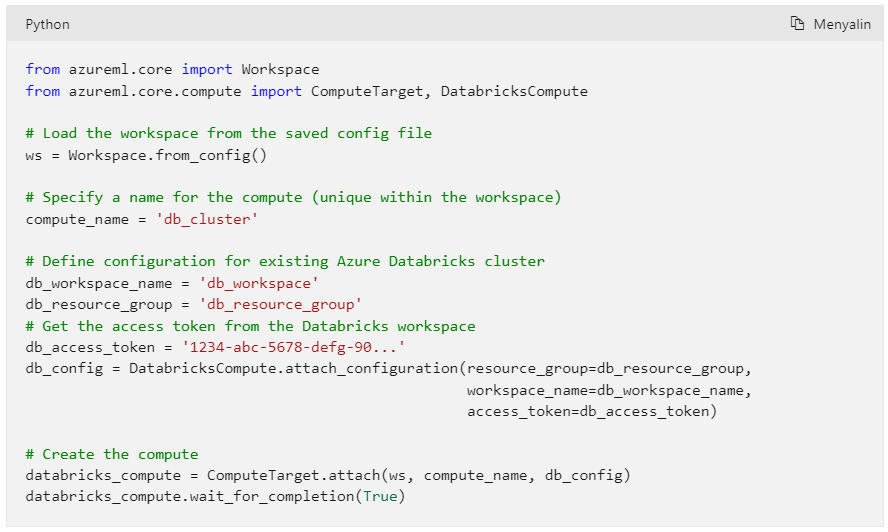
Di Azure Machine Learning, alur adalah alur kerja tugas pembelajaran mesin ketika setiap tugas diimplementasikan sebagai langkah. Langkah-langkah dapat diatur secara berurutan atau paralel, sehingga memungkinkan Anda untuk membangun logika alur yang mutahir untuk mengatur operasi pembelajaran mesin. Setiap langkah dapat dijalankan pada target komputasi tertentu, sehingga memungkinkan untuk menggabungkan berbagai jenis pemrosesan yang diperlukan untuk mencapai tujuan keseluruhan.

1. **Menjalankan langkah pipa di Databricks Compute**

Azure Machine Learning mendukung langkah saluran khusus yang disebut DatabricksStep yang dapat digunakan untuk menjalankan buku catatan, skrip, atau JAR yang dikompilasi pada kluster Azure Databricks. Untuk menjalankan langkah pipeline pada kluster Databricks, Anda perlu melakukan langkah-langkah berikut:

1. Lampirkan Azure Databricks Compute ke ruang kerja Azure Machine Learning.
2. Mendefinisikan DatabricksStep dalam sebuah pipeline.
3. Kirim alur.
4. **Melampirkan Azure Databricks Compute**

Contoh kode berikut dapat digunakan untuk melampirkan kluster Azure Databricks yang sudah ada:



1. **Mendefinisikan Databricks Langkah dalam pipeline**

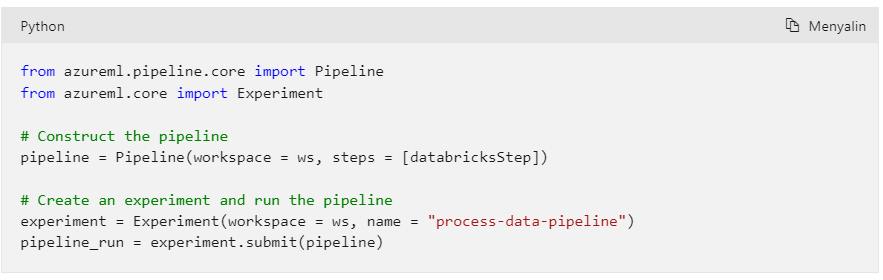
Untuk membuat alur, Anda harus terlebih dahulu menentukan setiap langkah lalu membuat alur yang menyertakan langkah-langkahnya. Konfigurasi spesifik dari setiap langkah tergantung pada jenis langkah. Misalnya, kode berikut mendefinisikan langkah **DatabricksStep** untuk menjalankan skrip python, , pada komputasi Databricks terlampir.



Langkah di atas mendefinisikan konfigurasi untuk membuat kluster pekerjaan Databricks baru untuk menjalankan skrip Python. Kluster dibuat dengan cepat untuk menjalankan skrip dan kluster kemudian dihapus setelah eksekusi langkah selesai.

1. **Kirim alur**

Setelah menentukan langkah, Anda dapat menetapkannya ke pipeline, dan menjalankannya sebagai eksperimen:



**6. Latihan - Gunakan Azure Databricks dengan Azure Machine Learning**

Sekarang, Anda akan menjalankan eksperimen di Azure Machine Learning dari Azure Databricks.

Dalam latihan ini, Anda akan:

* Menjalankan eksperimen Azure ML di Databricks.
* Meninjau metrik eksperimen di Azure ML Studio.

1. **Instruksi**

Ikuti petunjuk berikut untuk menyelesaikan latihan:

* Buka petunjuk latihan di <https://aka.ms/mslearn-dp090.> \ <https://microsoftlearning.github.io/dp-090-databricks-ml/>
* Selesaikan latihan **Menjalankan eksperimen di Azure Machine Learning.**

1. **Ringkasan**

Dalam modul ini, Anda mempelajari cara melacak Azure Databricks di Azure Machine Learning.

Sekarang setelah Anda menyelesaikan modul ini, Anda dapat:

* Jelaskan Azure Machine Learning.
* Jalankan eksperimen.
* Mencatat metrik dengan MLflow.
* Jalankan Pipeline Step pada Databricks Compute.