Ways to Integrate SageMaker with Databricks

**1. Databricks → SageMaker Training**

* Use Databricks to **clean & transform raw data**.
* Save final training datasets into **S3** (Parquet, Delta, CSV).
* Call **SageMaker training jobs** directly from a Databricks notebook using boto3 or sagemaker SDK.
* SageMaker pulls training data from S3.

✅ Use case: Big data preprocessing with Spark → ML training on managed SageMaker infra.

**2. Databricks → SageMaker Inference**

* Train your models in Databricks (PyTorch, TensorFlow, MLlib, etc.).
* Export model artifacts to **S3**.
* Deploy them in SageMaker as:
  + **Real-time endpoints** (REST API for predictions).
  + **Batch Transform** (large-scale offline inference).
* Databricks applications or jobs then **invoke the SageMaker endpoint**.

✅ Use case: Databricks ML development, but scalable inference via SageMaker.

**3. Shared MLOps with MLflow + SageMaker**

* Databricks comes with **MLflow** built-in.
* You can log and register models in MLflow.
* Export registered models to **S3** and deploy in SageMaker.
* SageMaker Model Registry can also be synced with MLflow.

✅ Use case: Cross-platform model lifecycle management.