Prepare Your Data for **Machine Learning Analytics**

Unify data science and engineering for successful machine learning initiatives with Databricks and AWS

Databricks' Unified Analytics Platform on Amazon Web Services (AWS) allows data teams to collaborate across the machine learning lifecycle. Create analytic workflows that go from ETL and interactive exploration to machine learning production leveraging the scalability and security of AWS.

Founded by the original creators of Apache Spark™, Databricks provides:

- Highly reliable and performant data pipelines
- Productive data science capability at scale

Streamline data pipelines

- Simplify machine learning workflows with scalable and repeatable data pipelines that combine batch and streaming data
- Use data lakes for machine learning data. Delta Lake optimizes your data lake to make it reliable and high performing.

Collaborate on data pipelines and machine learning jobs

- Bring data science and engineering teams together
- Use notebooks to share and explore code in the programming language of your choice with support for SQL, R, Python, Java, and Scala
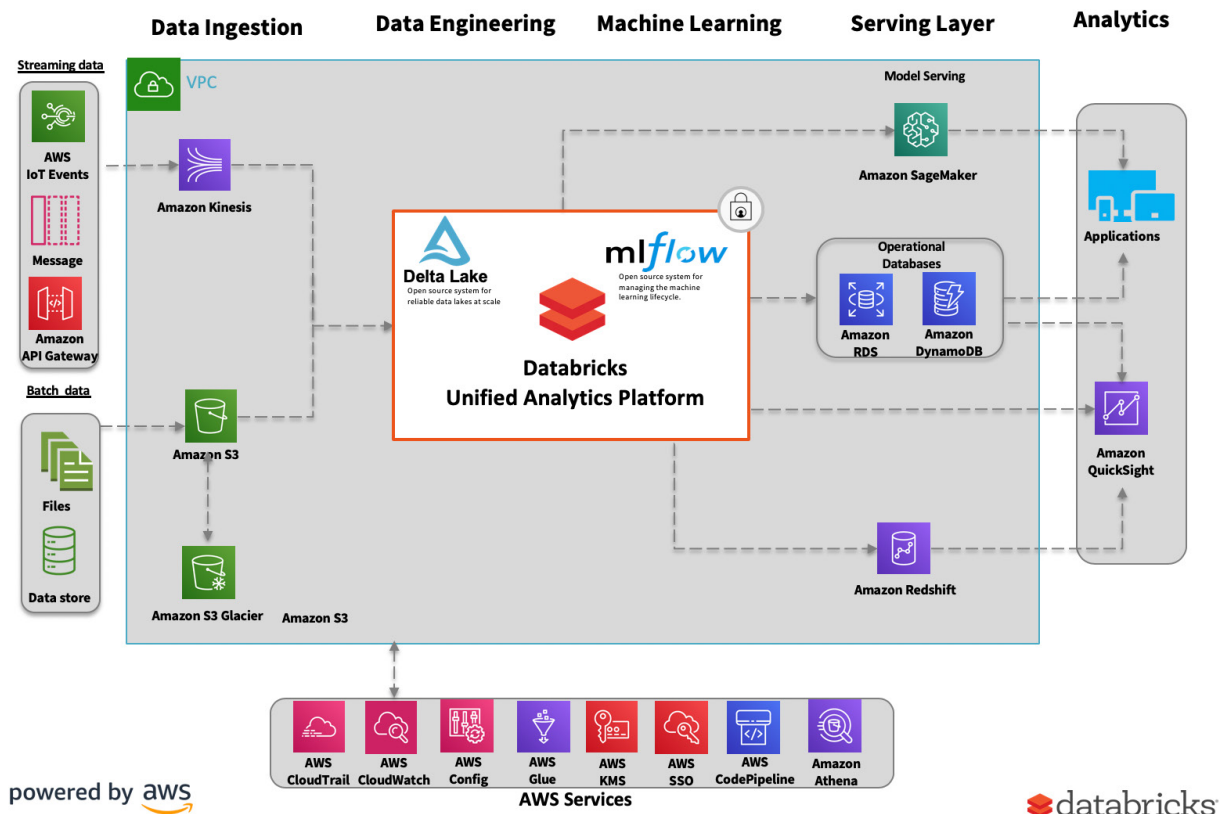- Make exploratory analysis seamless with in-notebook visualizations

Databricks differentiators

- Get results 10-100 times faster than open source Apache Spark
- Collaborate across the data and machine learning lifecycles using notebooks and integrations including data connectors and machine learning frameworks (TensorFlow, Keras, XGBoost, and scikit-learn)
- Leverage support for multiple industries including e-commerce, healthcare and life sciences, financial services, oil and gas, and entertainment

Databricks and AWS

- Enable data scientists, data engineers, and business analysts to iterate in real-time using one shared notebook; team members can see each other's comments and questions and give and get answers
- Manage the complete machine learning lifecycle from experiments to production with MLflow

- Distribute models with Amazon SageMaker, a fully-managed service that helps you build, train, and deploy machine learning models quickly
- Remove operational complexity with a serverless infrastructure that scales up and down to meet the needs of your teams

The Databricks Unified Analytics Platform

Databricks' Unified Analytics Platform runs on AWS and integrates with numerous AWS services.

Approachable AI and machine learning for your business

Eliminate the complexities of physical infrastructure, streamline inefficient machine learning workflows and processes, and create an environment that fosters collaboration across data teams to accelerate data-driven innovation.

TRY DATABRICKS ON AWS >

About Databricks

Databricks' Unified Analytics Platform, from the original creators of Apache Spark, unifies data science and engineering across the Machine Learning lifecycle from data preparation, to experimentation and deployment of machine learning applications.