

AWS의 ML 파이프라인 이해하기

최영준

AI/ML Expert SA

AWS



AWS 클라우드의 다양하고 폭넓은 서비스



AWS 활용 = 레고 블록 조립

비즈니스 요구 사항에 맞는 수백여개 이상의 서비스 조립을 통해 유연한 활동 가능

필요에 따른 다양한 구현 가능



purpose-built data services

Business Intelligence and Machine Learning



QuickSight
Visualizations



Data Exchange
Data exchange



SageMaker
ML



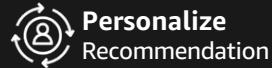
Comprehend
NLP



Transcribe
Speech-to-text



Textract
Extract text



Personalize
Recommendation



Forecast
Forecasts



Translate
Translation



CodeGuru
Code reviews



Kendra
Enterprise search

+Many more

AI 모델링/ MLOps / 시각화 Analytics



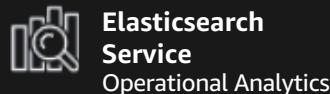
Redshift
Data warehousing



EMR
Hadoop + Spark



Athena
Interactive analytics

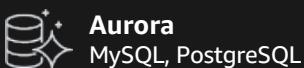


Elasticsearch Service
Operational Analytics

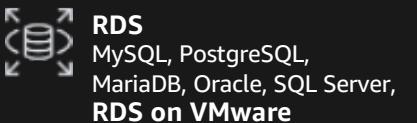


Kinesis Data Analytics
Real time

DB화/데이터분석



Aurora
MySQL, PostgreSQL



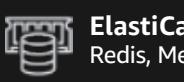
RDS
MySQL, PostgreSQL,
MariaDB, Oracle, SQL Server,
RDS on VMware



DynamoDB
Key value, Document



DocumentDB
Document



ElastiCache
Redis, Memcached



Keyspaces
(For Apache Cassandra)
Wide column



Neptune
Graph



QLDB
Ledger Database



Timestream
Time Series



Managed Blockchain



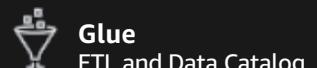
Blockchain
Templates

Databases

Data Lake



Lake Formation
Data Lakes



Glue
ETL and Data Catalog

데이터 저장



S3/Glacier

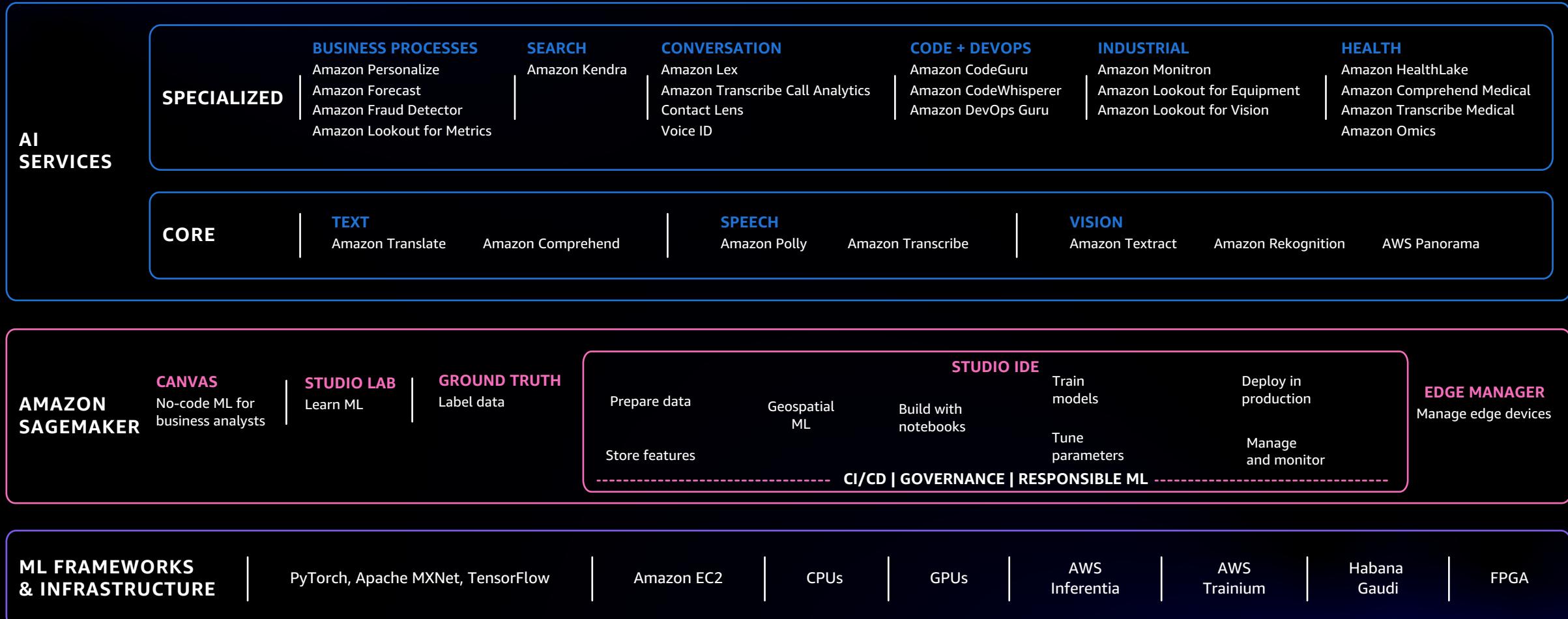
데이터 습득

Database Migration Service | Snowball | Snowmobile | Kinesis Data Firehose | Kinesis Data Streams | Managed Streaming for Kafka



AI/ML Stack

BROADEST AND MOST COMPLETE SET OF MACHINE LEARNING CAPABILITIES



MLOps 사례

- <https://aws.amazon.com/ko/blogs/tech/aws-mlops-use-case/>
- <https://aws.amazon.com/ko/blogs/tech/nongshim-amazon-sagemaker-material-forecast-mlops/>

AWS 개발 도구

1. AWS CLI 사용

AWS Command Line Interface

AWS Command Line Interface(AWS CLI)는 AWS 서비스를 관리하는 통합 도구입니다. 하나의 도구만 다운로드하여 구성하면 여러 개의 AWS 서비스를 명령줄에서 제어하고 스크립트를 통해 자동화할 수 있습니다.

AWS CLI v2는 개선된 설치 프로그램, AWS IAM Identity Center(AWS SSO의 후속 서비스)와 같은 새로운 구성 옵션, 다양한 상호작용 기능을 비롯한 여러 가지 [새로운 기능](#)을 제공합니다.



Windows

[64비트 Windows 설치 프로그램](#)을 다운로드해서 실행합니다.

MacOS

[MacOS PKG 설치 프로그램](#)을 다운로드해서 실행합니다.

Linux

[Linux 설치 프로그램](#)을 다운로드하고 압축을 해제한 다음 설치합니다.

Amazon Linux

AWS CLI는 [Amazon Linux AMI](#)에 미리 설치되어 제공됩니다.

출시 정보

최신 버전에 대한 자세한 내용은 [출시 정보](#)를 확인하십시오.

2. AWS Console 사용

EC2 > Instances > Launch an instance

Launch an instance [Info](#)

Amazon EC2 allows you to create virtual machines, or instances, that run on the AWS Cloud. Quickly get started by following the simple steps below.

Name and tags [Info](#)

Name

e.g. My Web Server

Add additional tags

▼ Application and OS Images (Amazon Machine Image) [Info](#)

An AMI is a template that contains the software configuration (operating system, application server, and applications) required to launch your instance. Search or Browse for AMIs if you don't see what you are looking for below.

Search our full catalog including 1000s of application and OS images

Recents | My AMIs | **Quick Start**



S



Browse more AMIs

Including AMIs from AWS, Marketplace and the Community

▼ Summary

Number of instances [Info](#)

1

Software Image (AMI)

Amazon Linux 2 Kernel 5.10 AMI... [read more](#)
ami-0aa7d40eeeae50c9a9

Virtual server type (instance type)

t2.micro

Firewall (security group)

New security group

Storage (volumes)

1 volume(s) - 8 GB

ⓘ Free tier: In your first year includes 750 hours of t2.micro (or t3.micro in the Regions in which t2.micro is unavailable) instance usage on free tier AMIs per month, 30 GiB of EBS storage, 2 million I/Os, 1 GB of snapshots, and 100 GB of bandwidth to the internet.

Cancel

Launch instance

<https://aws.amazon.com/ko/cli/?pg=developertools>

AWS 개발 도구

3. AWS SDK

<https://aws.amazon.com/ko/tools>

AWS에서의 구축을 위한 도구

AWS에서 애플리케이션 개발 및 관리를 위한 도구

프로그래밍 언어 기준으로 찾아보기

원하는 프로그래밍 언어에서 손쉽게 AWS 기반 애플리케이션 개발

[C++](#) [Go](#) [Java](#) [JavaScript](#) [Kotlin](#) [.NET](#) [Node.js](#) [PHP](#) [Python](#) [Ruby](#) [Rust](#) [Swift](#)

C++를 사용하여 구축 시작

⚙️ 애플리케이션 구축

C++ 특정 API 및 개발 환경에 통합된 친숙한 도구를 사용하여 애플리케이션 개발

[C++용 AWS SDK »](#)

☁️ IDE를 사용하여 AWS에 구축

AWS에서 코드 작성, 디버그 및 배포를 위해 널리 사용되는 IDE(통합 개발 환경) 사용

[AWS Cloud9 »](#)

▣ 시작하기

AWS에서 C++를 시작하는데 도움이 되는 설명서 및 샘플 코드에 액세스

[C++용 개발자 안내서 »](#)

👥 커뮤니티에 연결

AWS에서 C++ 기반 애플리케이션을 구축하는데 도움이 되도록 대화에 참여하거나 질문, 지침 및 리소스를 찾아볼 수 있습니다.

[개발자 블로그 »](#)

[GitHub의 AWS »](#)

[스택 오버플로 »](#)

AWS Python SDK – boto3 (<http://boto3.readthedocs.io>)



이미지 : 위키피디아 Amazon river dolphin

Boto3 Docs 1.26.60 documentation

Docs / Boto3 documentation

Boto3 documentation

You use the AWS SDK for Python (Boto3) to create, configure, and manage AWS services, such as Amazon Elastic Compute Cloud (Amazon EC2) and Amazon Simple Storage Service (Amazon S3). The SDK provides an object-oriented API as well as low-level access to AWS services.

Note
Documentation and developers tend to refer to the AWS SDK for Python as "Boto3," and this documentation often does so as well.

Quickstart

- [Quickstart](#)
 - [Installation](#)
 - [Configuration](#)
 - [Using Boto3](#)
- [A sample tutorial](#)
 - [SQS](#)
 - [Creating a queue](#)
 - [Using an existing queue](#)
 - [Sending messages](#)
 - [Processing messages](#)
- [Code examples](#)
 - [Amazon CloudWatch examples](#)
 - [Amazon DynamoDB](#)
 - [Amazon EC2 examples](#)
 - [AWS Identity and Access Management examples](#)

SageMaker SDK (<https://sagemaker.readthedocs.io>)

The screenshot shows a web browser displaying the Amazon SageMaker Python SDK documentation. The URL in the address bar is <https://sagemaker.readthedocs.io/en/stable/>. The page has a blue header with the "sagemaker" logo and "stable" branch indicator. A search bar is present in the header. The main content area features a large title "Amazon SageMaker Python SDK" and a brief introduction: "Amazon SageMaker Python SDK is an open source library for training and deploying machine-learned models on Amazon SageMaker." Below this, there's a section titled "Overview" with a bulleted list of topics: "Using the SageMaker Python SDK", "Train a Model with the SageMaker Python SDK", "Using Models Trained Outside of Amazon SageMaker", "SageMaker Automatic Model Tuning", "SageMaker Batch Transform", and "Local Mode". On the right side of the page, there's a "Edit on GitHub" button.

Using the SageMaker Python SDK

Use Version 2.x of the SageMaker Python SDK

APIs

Frameworks

First-Party Algorithms

Workflows

Amazon SageMaker Debugger

Amazon SageMaker Feature Store

Amazon SageMaker Model Monitor

Amazon SageMaker Processing

» Amazon SageMaker Python SDK

Edit on GitHub

Amazon SageMaker Python SDK

Amazon SageMaker Python SDK is an open source library for training and deploying machine-learned models on Amazon SageMaker.

With the SDK, you can train and deploy models using popular deep learning frameworks, algorithms provided by Amazon, or your own algorithms built into SageMaker-compatible Docker images.

Here you'll find an overview and API documentation for SageMaker Python SDK. The project homepage is in Github: <https://github.com/aws/sagemaker-python-sdk>, where you can find the SDK source and installation instructions for the library.

Overview

- Using the SageMaker Python SDK
 - Train a Model with the SageMaker Python SDK
 - Using Models Trained Outside of Amazon SageMaker
 - SageMaker Automatic Model Tuning
 - SageMaker Batch Transform
 - Local Mode

boto3 & SageMaker SDK 데모

The screenshot shows a JupyterLab environment with the title bar "JupyterLab". The browser address bar displays the URL <https://leon.notebook.us-east-1.sagemaker.aws/lab>. The JupyterLab menu bar includes File, Edit, View, Run, Kernel, Git, Tabs, Settings, and Help. A sidebar on the left contains icons for file operations like Open, Save, and Delete, along with AWS-related icons for Lambda, S3, and SageMaker.

The main workspace contains two sections of code:

- 1. boto3 sdk 사용하기**

```
[ *]: import boto3
```

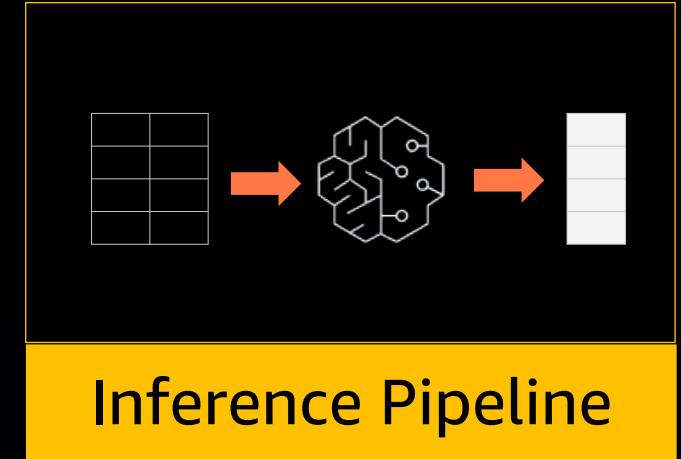
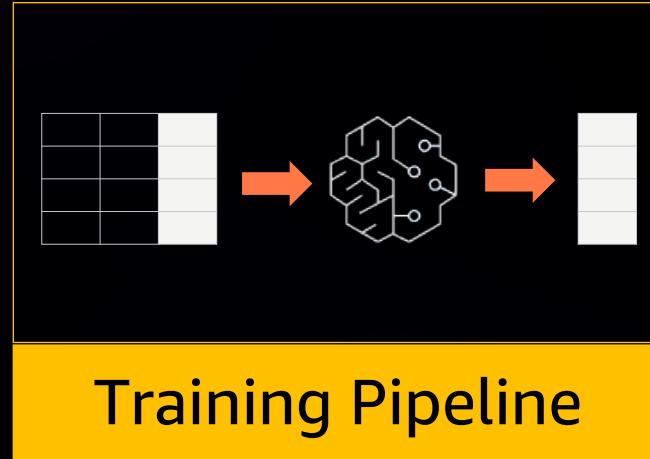
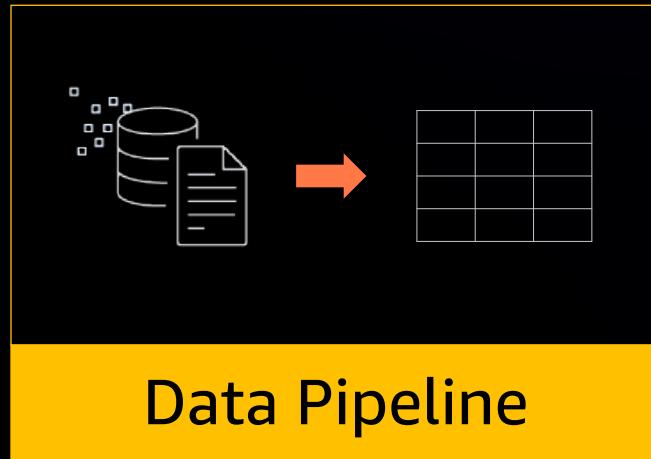
```
[ ]:
```
- 2. SageMaker sdk 사용하기**

```
[ ]:
```

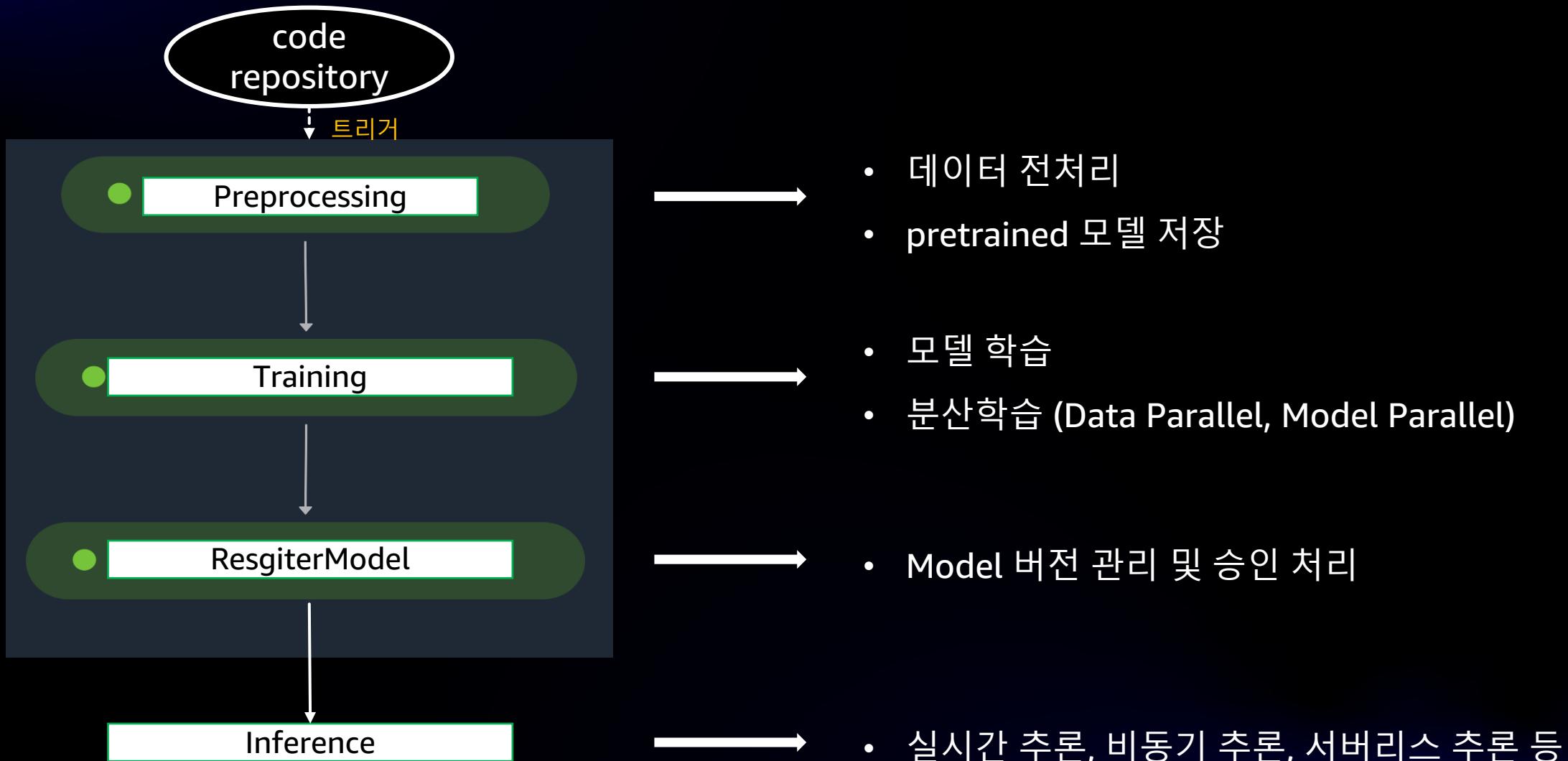
The bottom status bar shows the AWS logo, the current kernel as "conda_python3 | Idle", and the mode as "Mode: Command". The footer includes the copyright notice "© 2022 Amazon Web Services, Inc. or its affiliates. All rights reserved" and the page number "11".

ML Pipelines

- ML projects can be divided into 3 main pipelines

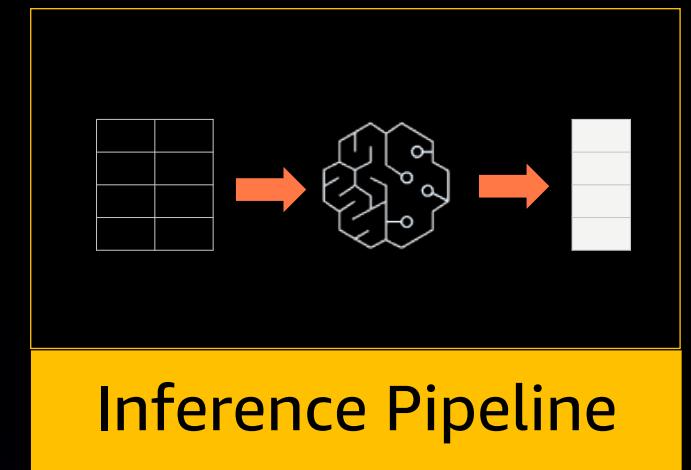
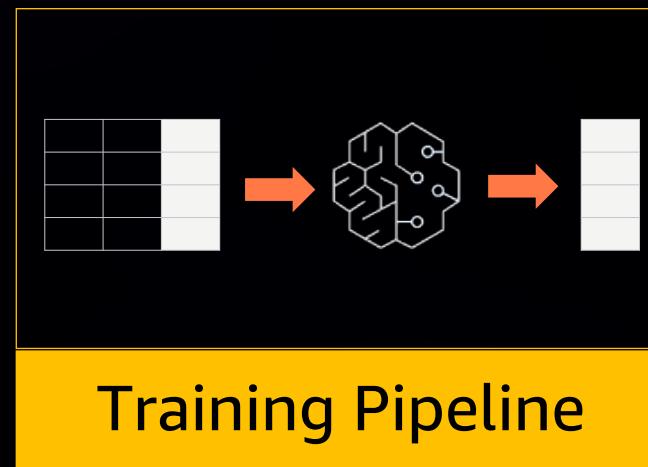
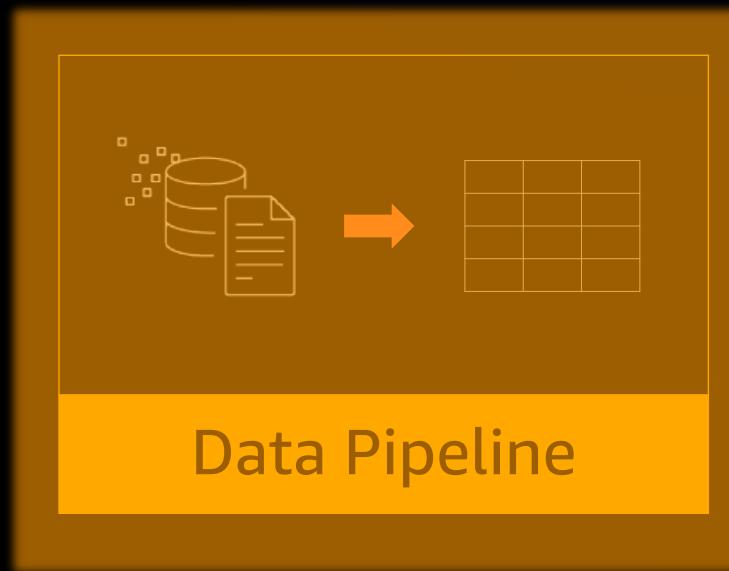


ML Pipelines



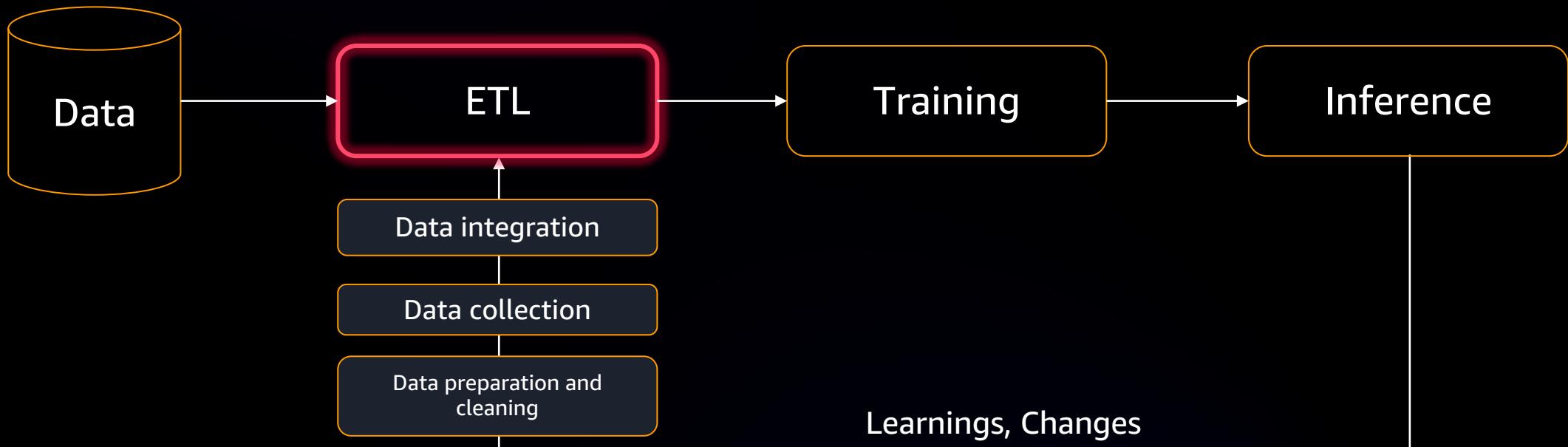
ML Pipelines

- ML projects can be divided into 3 main pipelines



Data Pipeline

서로 다른 소스에서 원시 데이터를 수집하여 저장 및 분석을 위한 데이터를 최종 위치로 이동하는 파이프라인



Data Pipeline



Amazon SageMaker Data Wrangler

	2 Categories	11 Categories
123		
1		Allen, Miss. Elisabet
1		Allison, Master. Hud
0		Allison, Miss. Helen
0		Allison, Mr. Hudson
0		Allison, Mrs. Hudson
1		Anderson, Mr. Harry
1		Andrews, Miss. Korr
0		Andrews, Mr. Thomas
1		Appleton, Mrs. Edward
0		Artagaveytia, Mr. Ra

ADD TRANSFORM

CUSTOM

Custom formula
Define a new column using a Spark SQL expression to query data in the current dataframe.

Custom transform
Use Pyspark, Pandas, or Pyspark (SQL) to define custom transformations.

STANDARD

Balance data
Balance the data for binary classification problems using random oversampling, random undersampling or SMOTE.

Dimensionality Reduction
For the top K principal components, trains a model to project vectors to a lower dimensional space.

Encode categorical
Convert categorical variables to numeric or vector representations.

Featurize date/time
Encode date/time values to numeric and vector representations.

ENCODE CATEGORICAL

Convert categorical variables to numeric or vector representations. [Learn more.](#)

Transform [One-hot encode](#)

Input columns [Select...](#)

CUSTOM TRANSFORM

Use Pyspark, Pandas, or Pyspark (SQL) to define custom transformations. [Learn more.](#)

Name

Optional

Python (PySpark) [Python \(PySpark\)](#)

Search example snippets

Your custom transform

```
1 # Table is available as variable `df`  
2
```

Clear **Preview** **Add**

Amazon SageMaker Data Wrangler

The screenshot displays the Amazon SageMaker Data Wrangler interface. On the left, a histogram titled "Histogram: Late flight day of week and morning" shows the count of records versus DayOfWeek(binned) for flights where LateFlight is false. The x-axis ranges from 1.0 to 7.0, and the y-axis ranges from 0 to 80. The distribution shows peaks around Day 3 and Day 6. On the right, the "Data flow" editor shows a workflow for joining two datasets. The flow starts with two "Source" nodes: "S3: income.csv" and "S3: airlines-2008.csv". Both pass through "Data types" steps (represented by boxes with "#.#") and then "Transform" steps (represented by boxes with '</>'). The "income.csv" path continues through a "Custom SQL" step (with '</>') and then a "Join" step (with two overlapping circles) to produce the final output "joined_df". The "airlines-2008.csv" path continues through a "Steps (5)" step (represented by a box with a plus sign) and then a "Join" step to also contribute to "joined_df". A "Create job" button is located in the top right corner of the editor.

Amazon SageMaker Data Wrangler

QUICK MODEL

Quick model provides an estimate of the expected predicted quality of a model that you train. The data is split into training and validation folds where Data Wrangler uses 80% of the sample to train the XGBoost model with the default hyper-parameters. It applies early stopping on the validation fold.

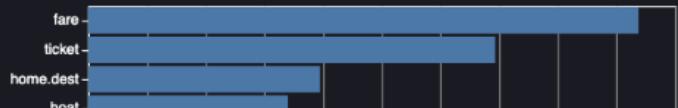
Metric	Validation scores	Train scores
R2	0.955	0.998
MSE	0.0306	0.00139
RMSE	0.175	0.0372
MAE	0.0777	0.0202
Max error	0.835	0.287
Median absolute error	0.0142	0.00847

FEATURE SUMMARY

See a summary of the features ordered by the prediction power. Prediction power is measured by fitting a model for each feature separately on the training fold after applying minimal feature pre-processing.

- The scores are normalized to the range [0,1].
- Higher prediction power scores, toward 1, indicate columns that are more useful for prediction.
- Lower scores, toward 0 point to columns that contain little useful information for prediction.
- A score of 1 implies perfect predictive abilities, which often indicates an error called target leakage.

Feature	Prediction power	Type	Valid	Missing	Outliers	Skew
fare	0.938	numeric	99.9%	0	1	4.36
ticket	0.693	numeric	73.1%	0	0	0
home.dest	0.395	text	100%	0	0	0
boat	0.34	categorical	100%	0%	0	1
name	0.166	text	100%	0%	0	0
cabin	0.14	text	100%	0%	0	0
embarked	0.093	categorical	100%	0%	0	0
age	0.0819	numeric	79.9%	20.1%	0	0
survived	0.0781	binary	100%	0%	0	0
sibsp	0.0238	numeric	100%	0%	0	0
parch	0.0105	numeric	100%	0%	0	0

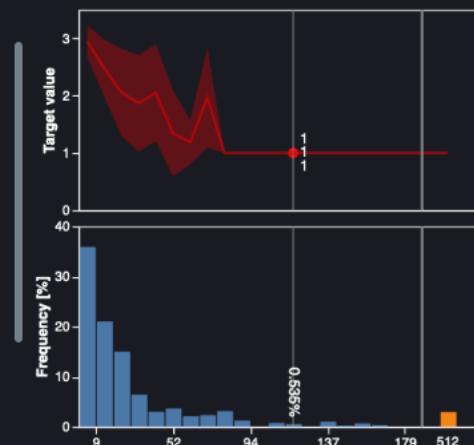


FEATURE DETAILS

fare

numeric

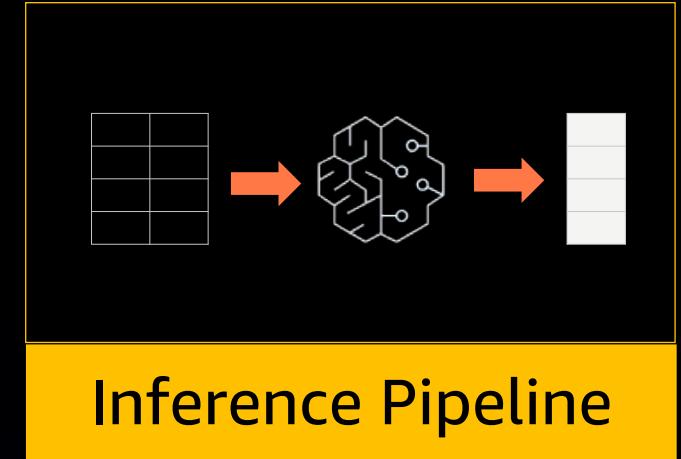
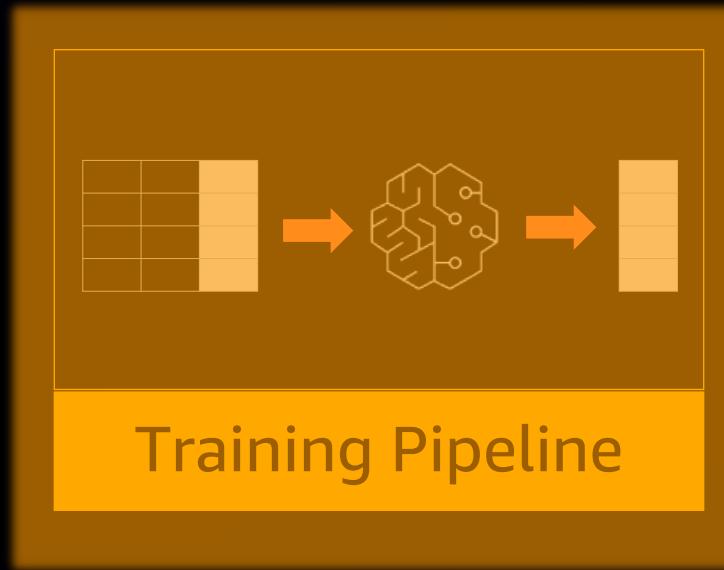
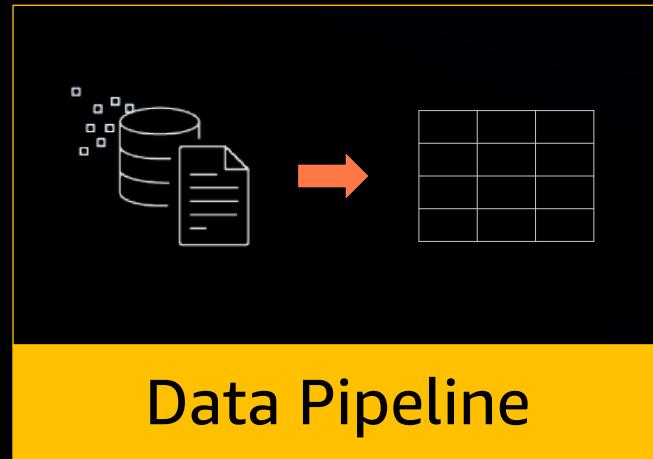
key	value
Feature name	fare
Type	numeric
Prediction power	0.938
Valid	99.9%
Missing	0.0764%
Outliers	2.91%
Min	0
Max	512
Mean	33.3
Median	14.5
Skew	4.36



Histogram of fare with the corresponding target distribution. The lower - the histogram of fare frequency, the upper - the corresponding target average with a confidence interval. Orange bars contain outliers and the value below them is the outlier count.

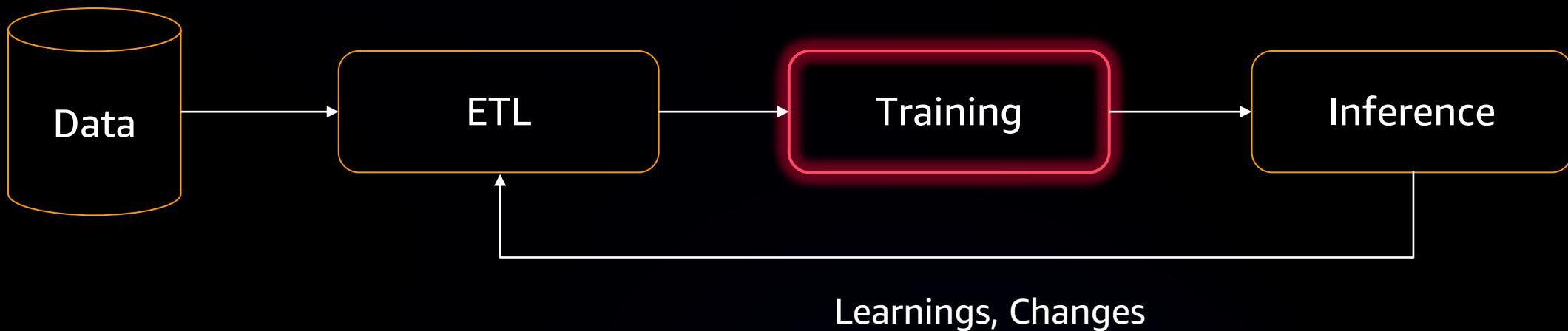
ML Pipelines

- ML projects can be divided into 3 main pipelines

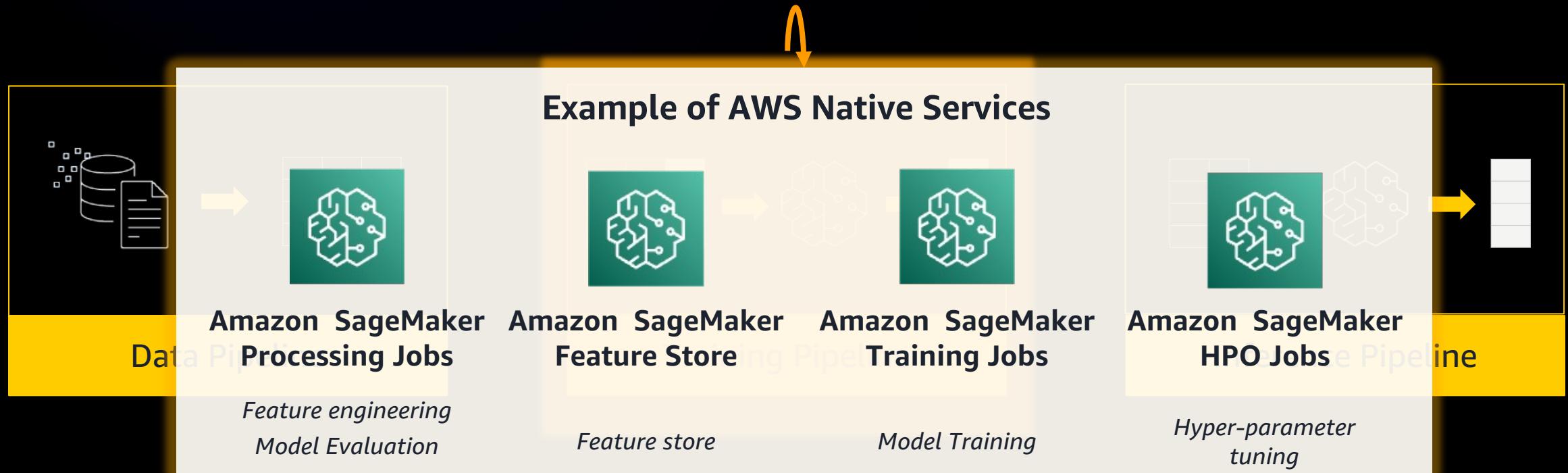


Training Pipeline

피처 엔지니어링, ML 모델 훈련, 조정(HPO) 및 성능 평가를 담당하는 ML 파이프라인

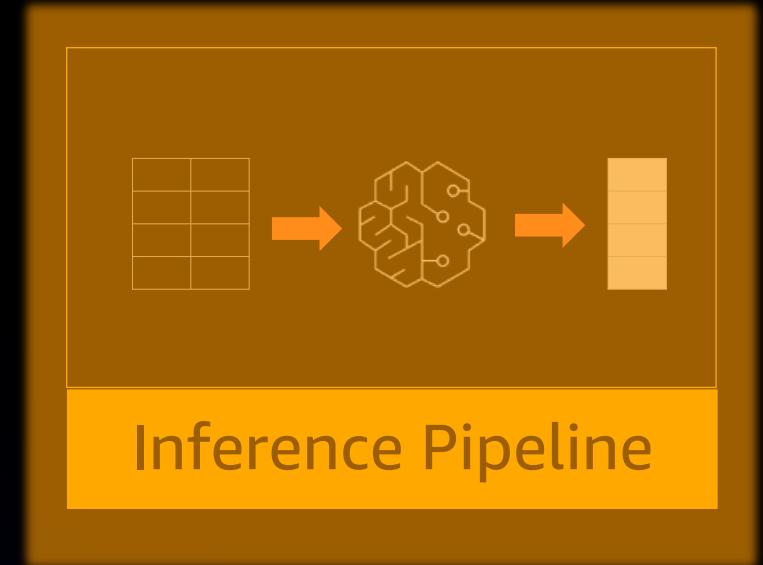
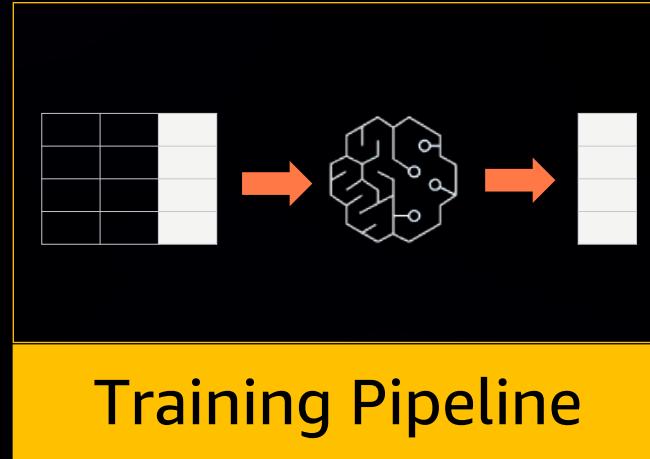
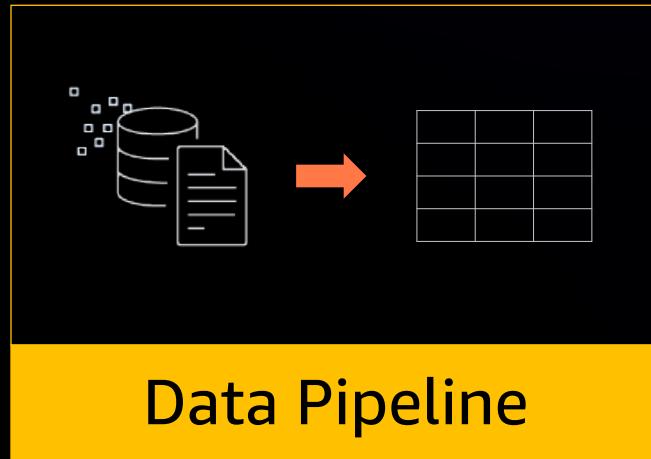


Training Pipeline



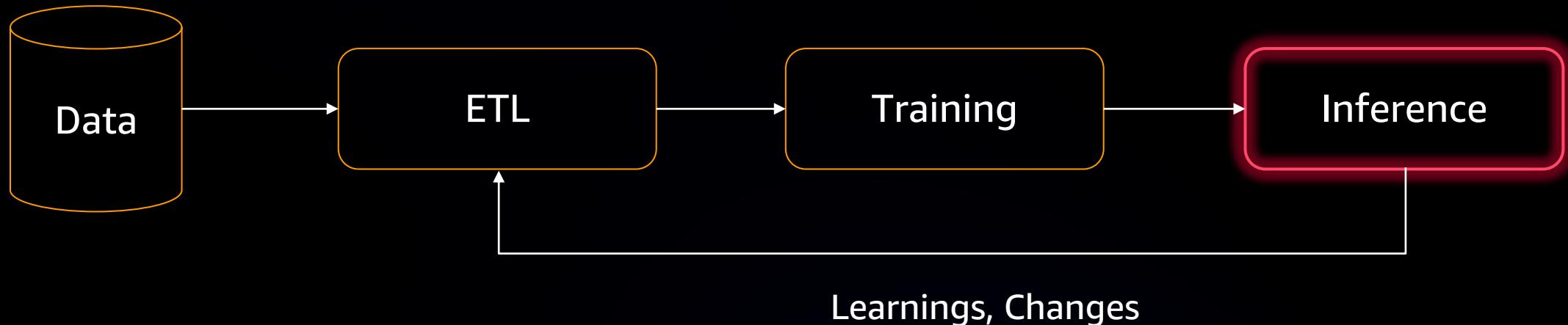
ML Pipelines

- ML projects can be divided into 3 main pipelines



Inference Pipeline

데이터에 대한 추론 요청을 처리하는 단계들의 선형적 시퀀스로 구성된 ML 파이프라인
컨테이너에 패키징된 사전 훈련된 알고리즘을 정의하고 배포하는 데 사용 가능

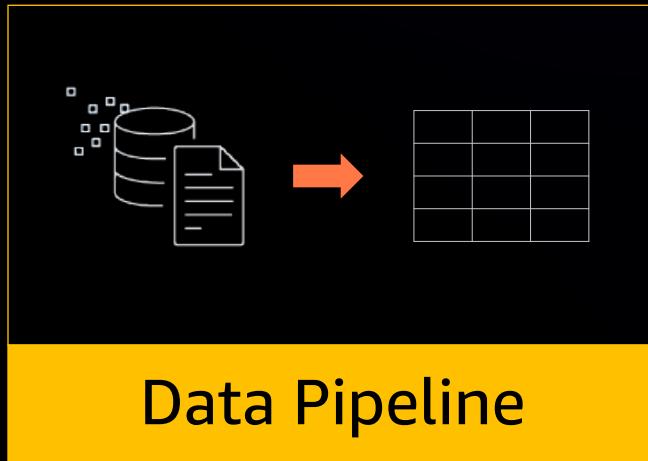


Inference Pipeline

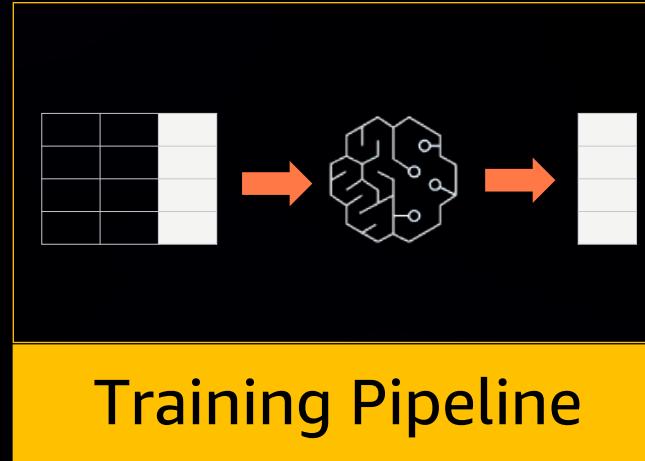


ML Pipelines

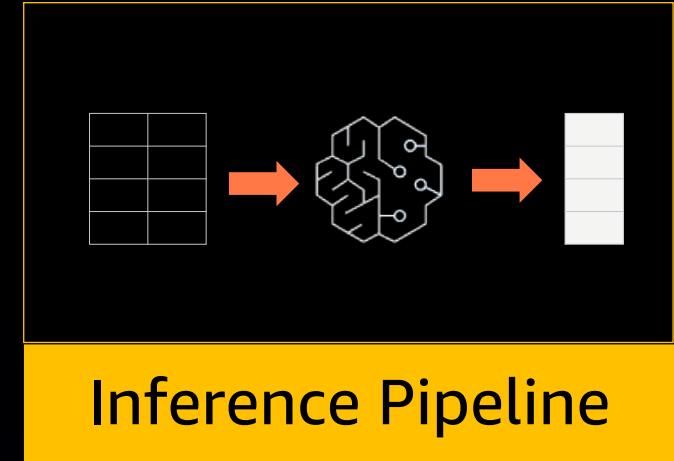
- ML projects can be divided into 3 main pipelines



Data Pipeline



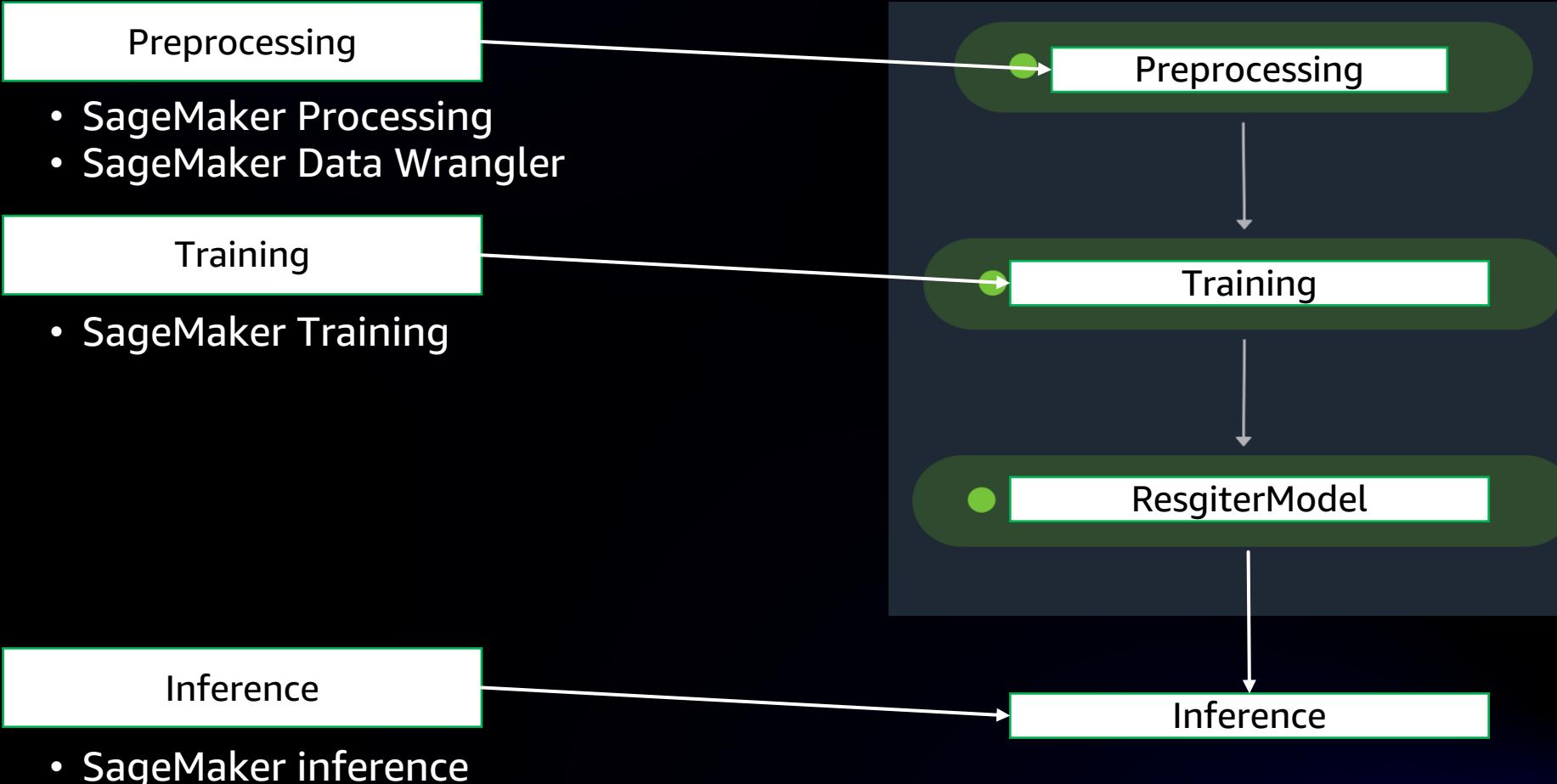
Training Pipeline



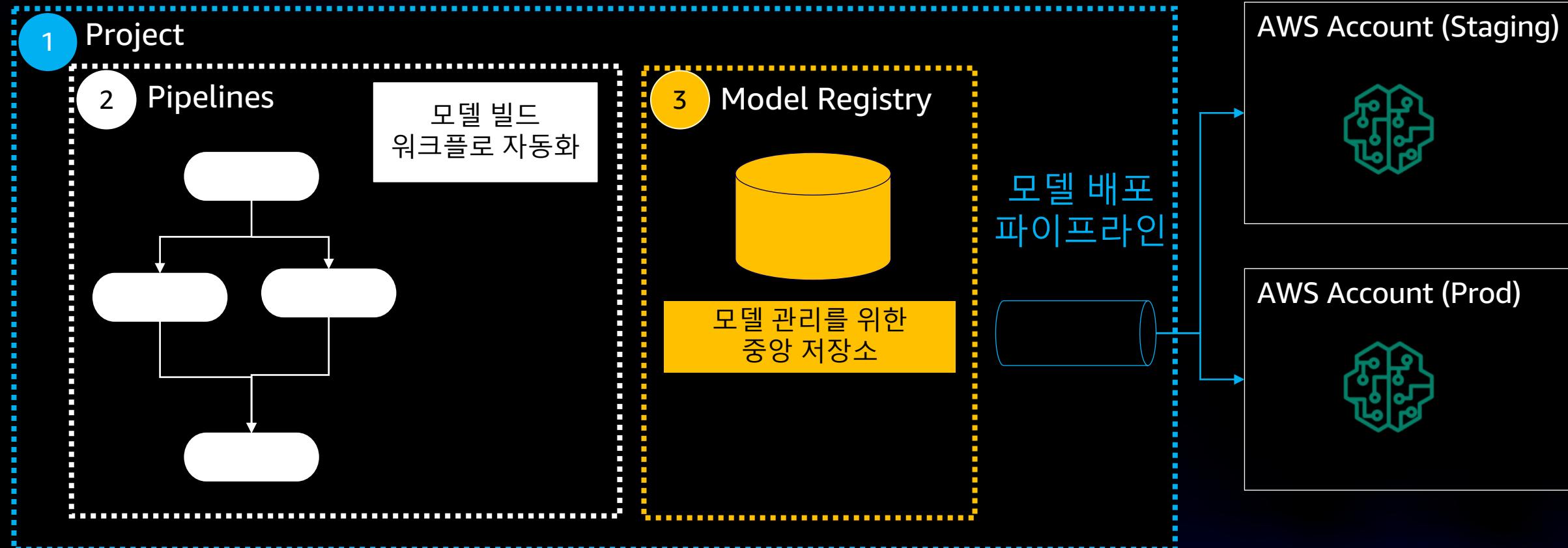
Inference Pipeline

MLOps

ML pipeline 구성요소



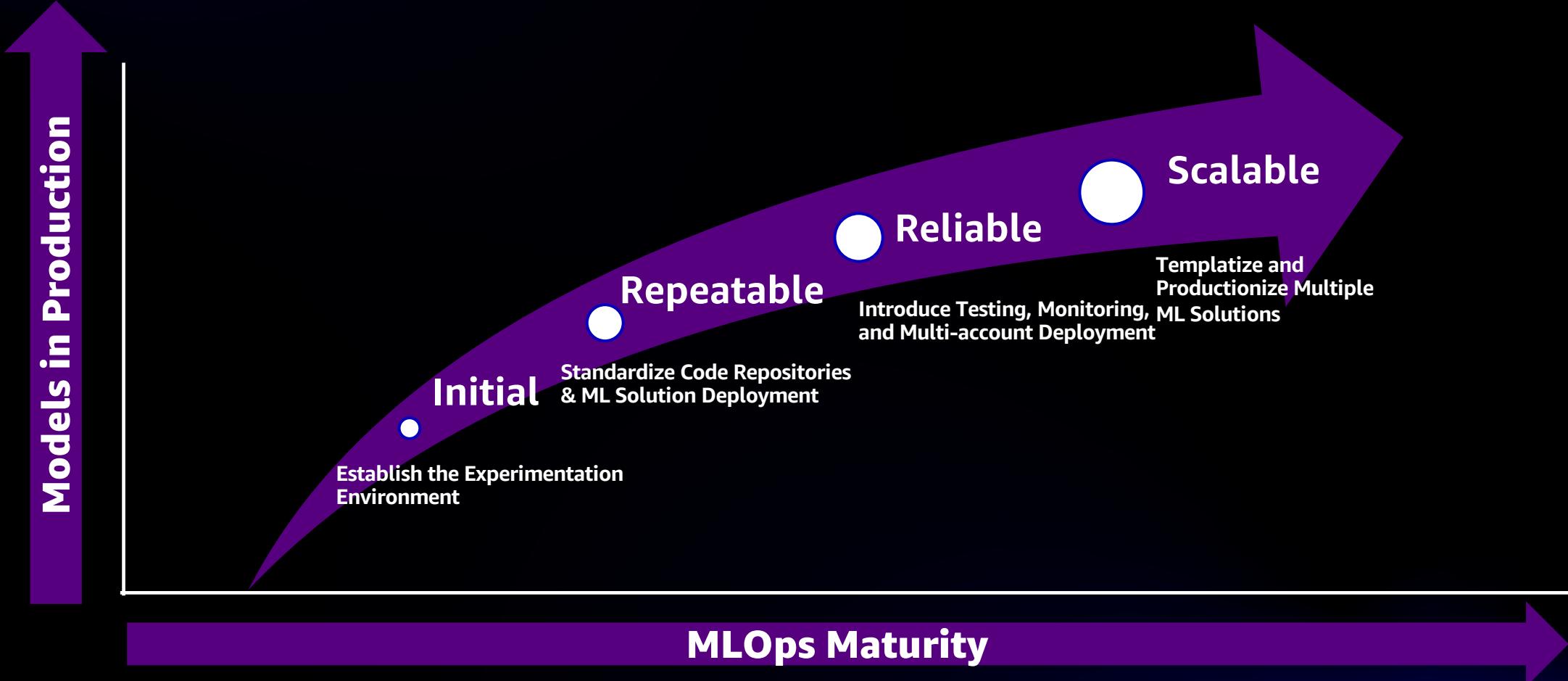
Amazon SageMaker Pipelines



Reference: <https://aws.amazon.com/blogs/machine-learning/multi-account-model-deployment-with-amazon-sagemaker-pipelines/>

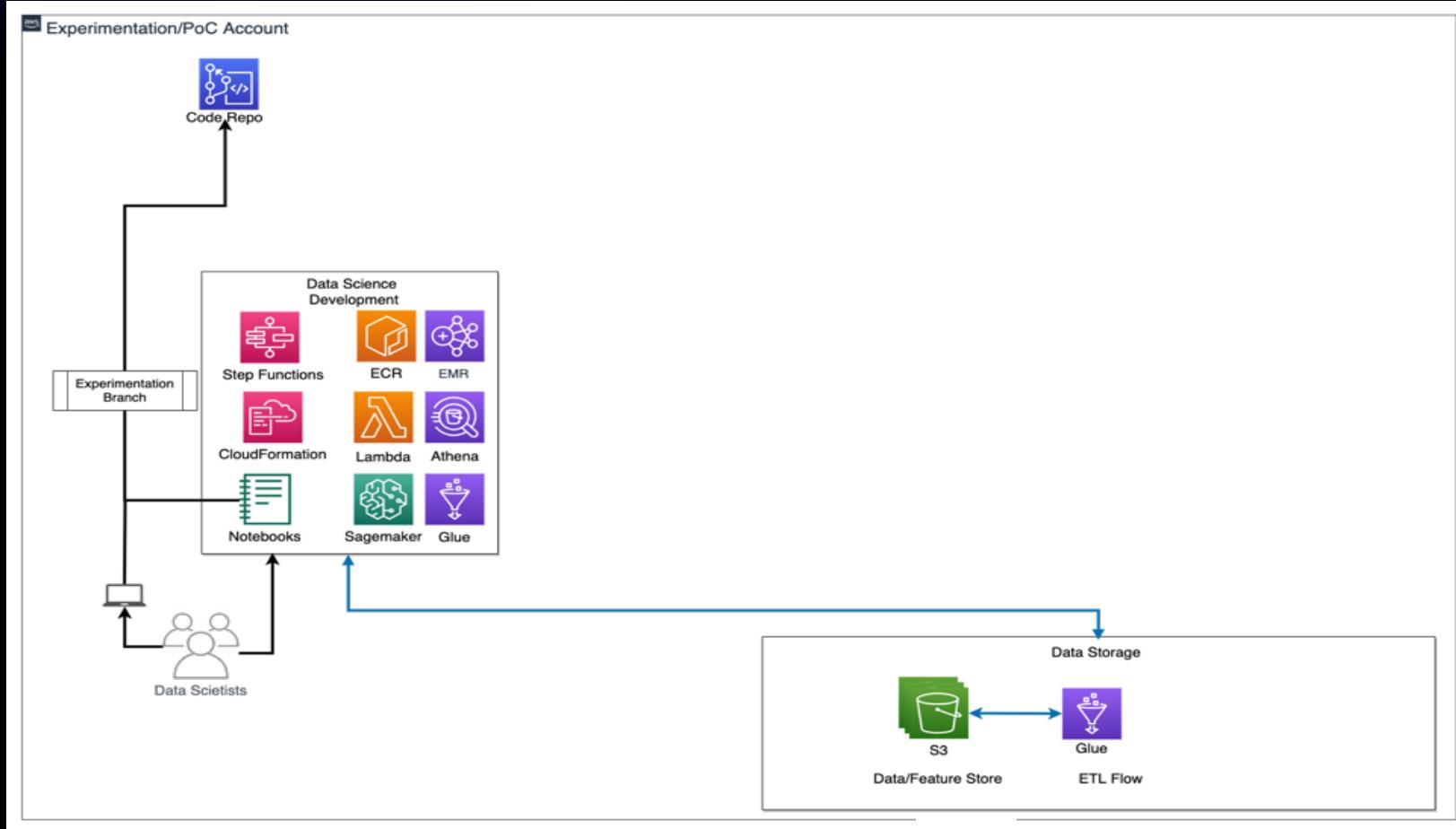
How to do MLOps?

MLOps Maturity Curve



Initial Phase

Establish Experimentation Environment

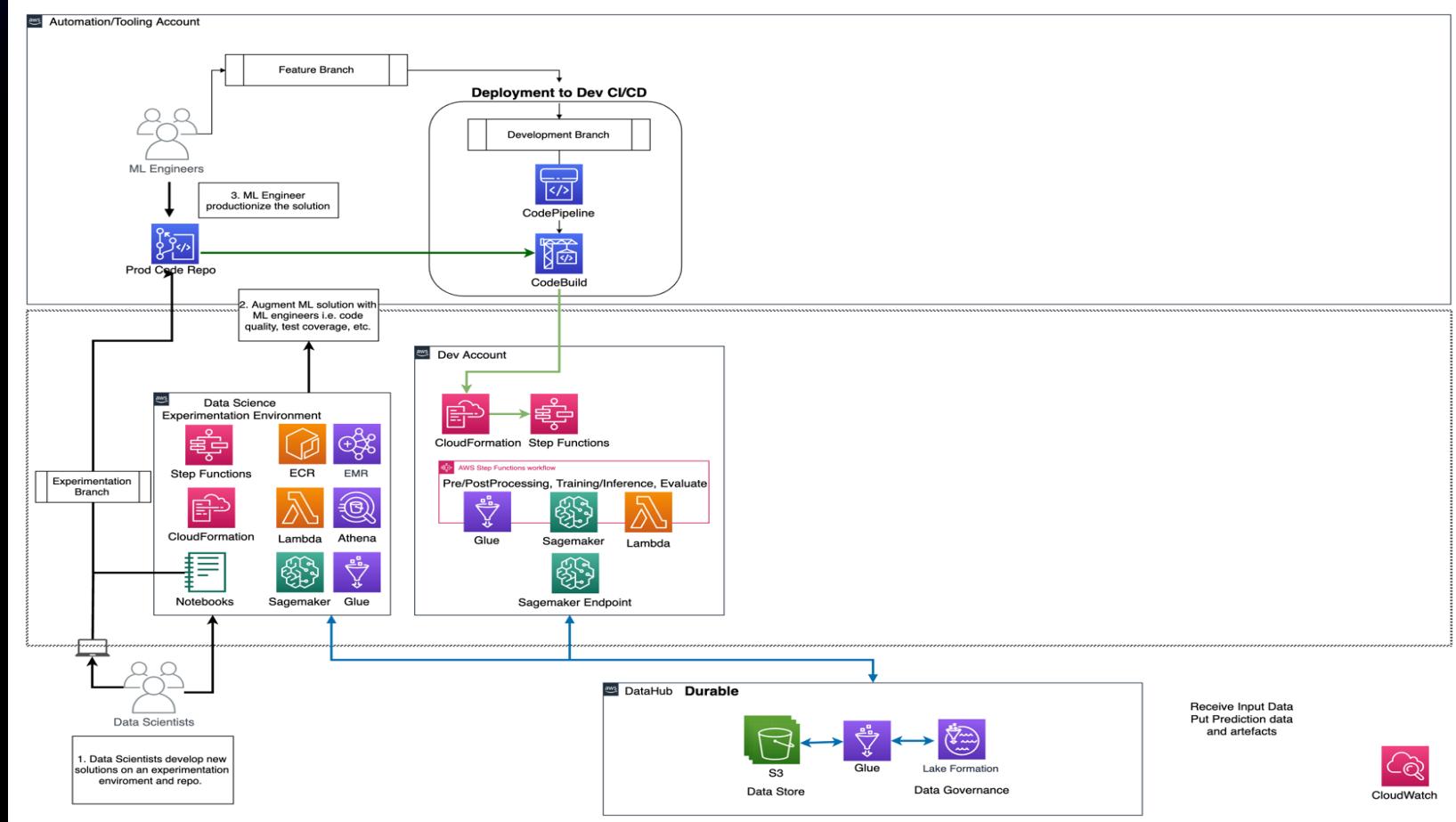


The Gray Area:

- Orchestrators
- Deployment/Pipelines
- Defined Roles & Responsibilities: DS, MLE, DevOps
- Environments: Dev, Pre-prod, Prod
- Model & Metadata: model registry, model

Repeatable Phase

First steps on MLOps



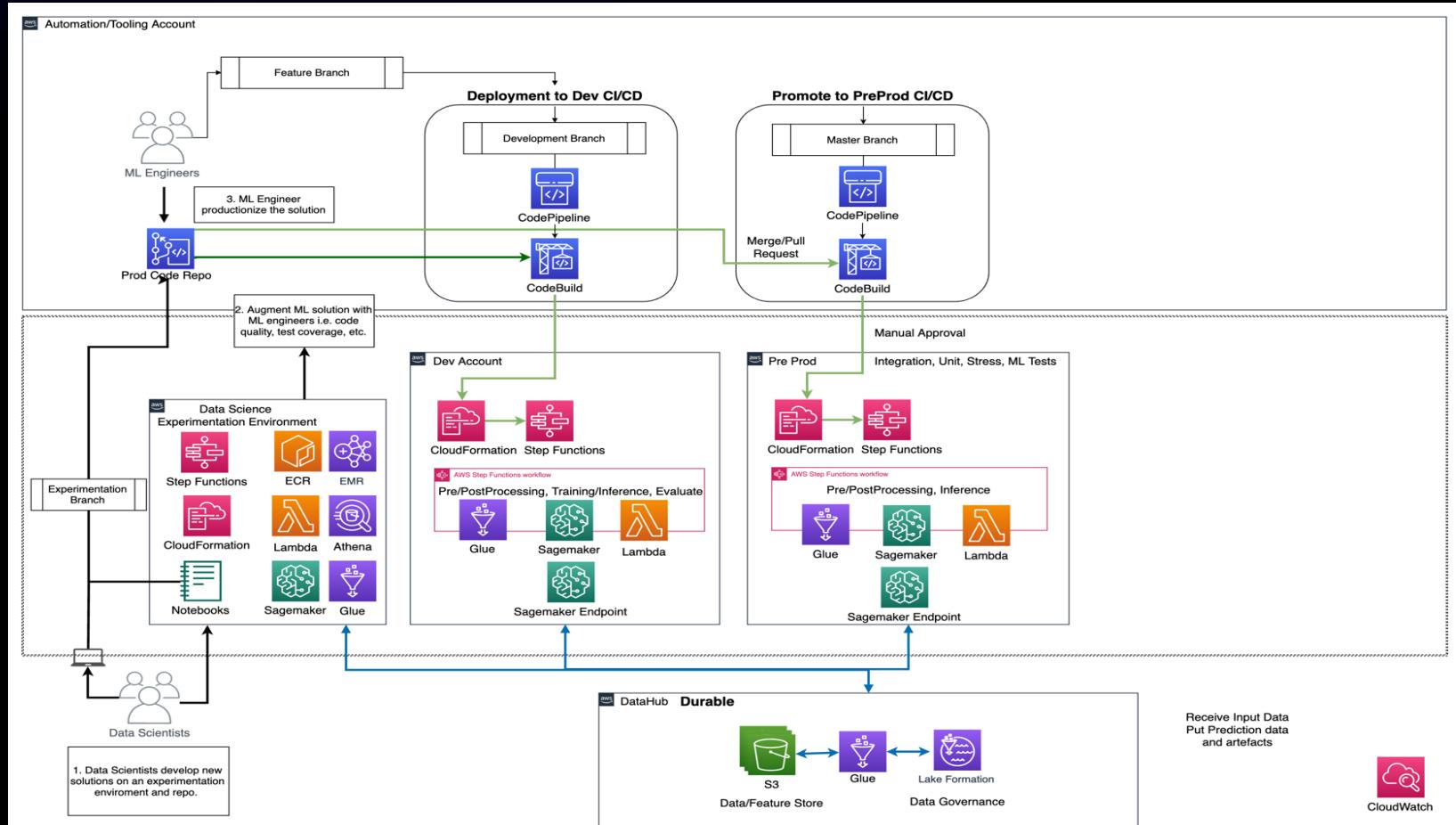
The Gray Area:

Orchestrators

- Deployment/Pipelines: **Dev**, Pre-prod, Prod
- Defined Roles & Responsibilities: **DS**, **MLE**, **DevOps**, Biz Owners, SMEs, Security
- Environments: **Dev**, Pre-prod, Prod
- Model & Metadata: model registry, model

Reliable Phase 1/2

Understand the need of MLOps



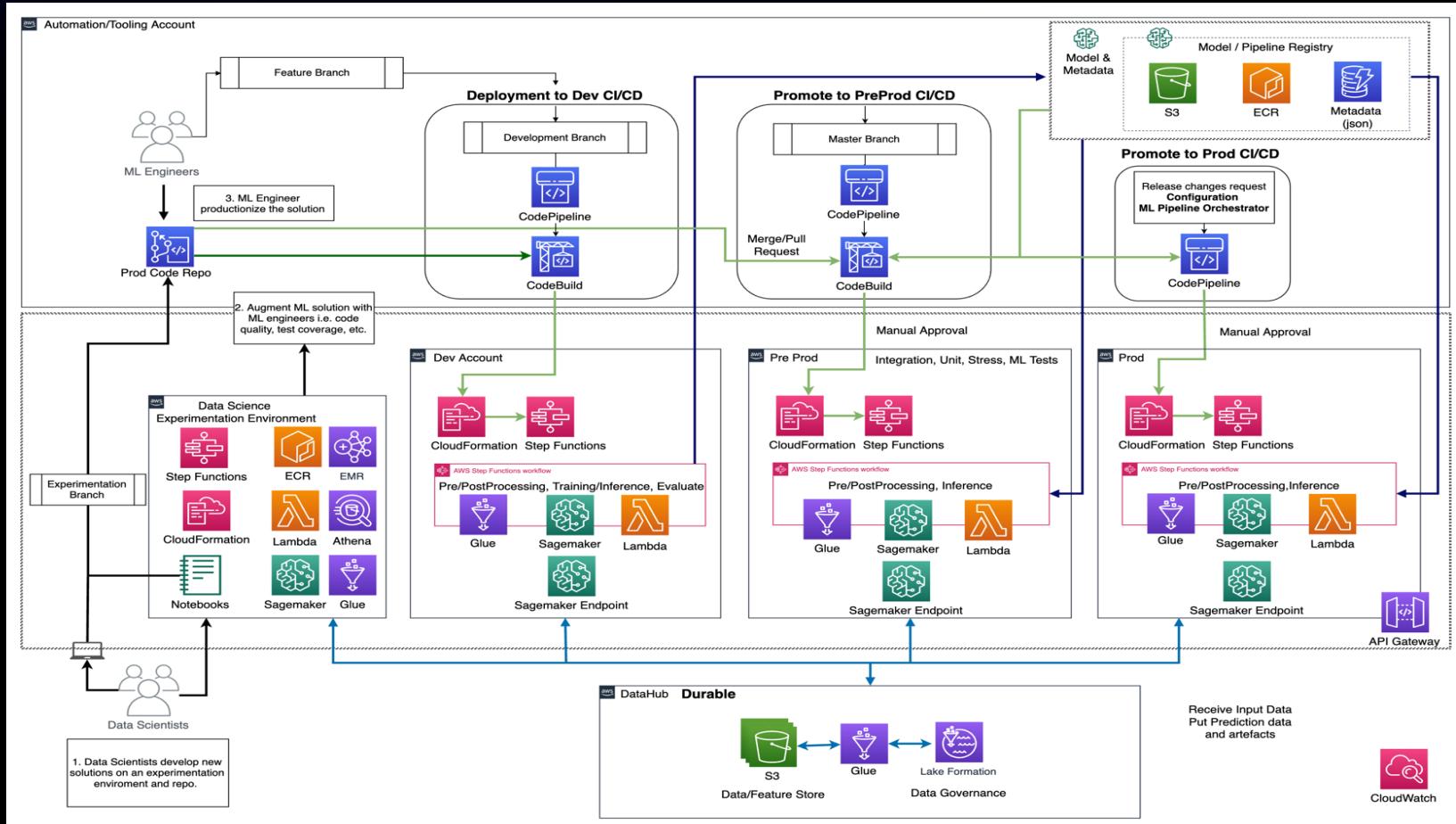
The Gray Area:

➡ Orchestrators

- Deployment/Pipelines: **Dev**, **Pre-prod**, **Prod**
- Defined Roles & Responsibilities: **DS**, **MLE**, **DevOps**, **Biz Owners**, **SMEs**, Security, Architects, Operations
- Environments: **Dev**, **Pre-prod**, **Approval**, **Prod**
- Model & Metadata: model registry, model

Reliable Phase 2/2

Best practices on MLOps



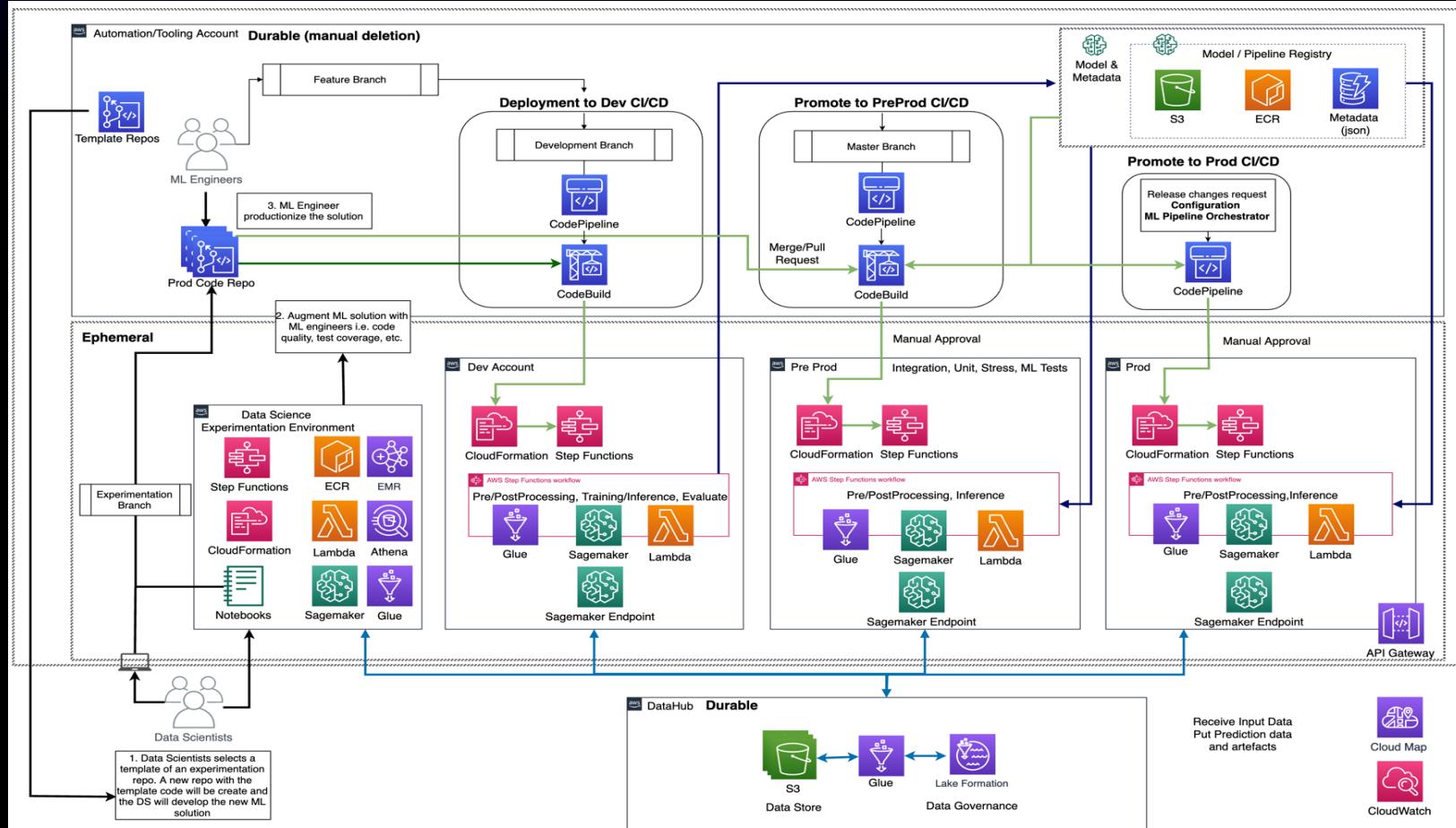
The Gray Area:

Orchestrators

- Deployment/Pipelines: **Dev**, **Pre-prod**, **Prod**
- Defined Roles & Responsibilities: **DS**, **MLE**, **DevOps**, **Biz Owners**, **SMEs**, **Security**, **Architects**, **Operations**
- Environments: **Dev**, **Pre-prod**, **Approval**, **Prod**
- Model & Metadata: **model registry**, **model**

Scalable Phase

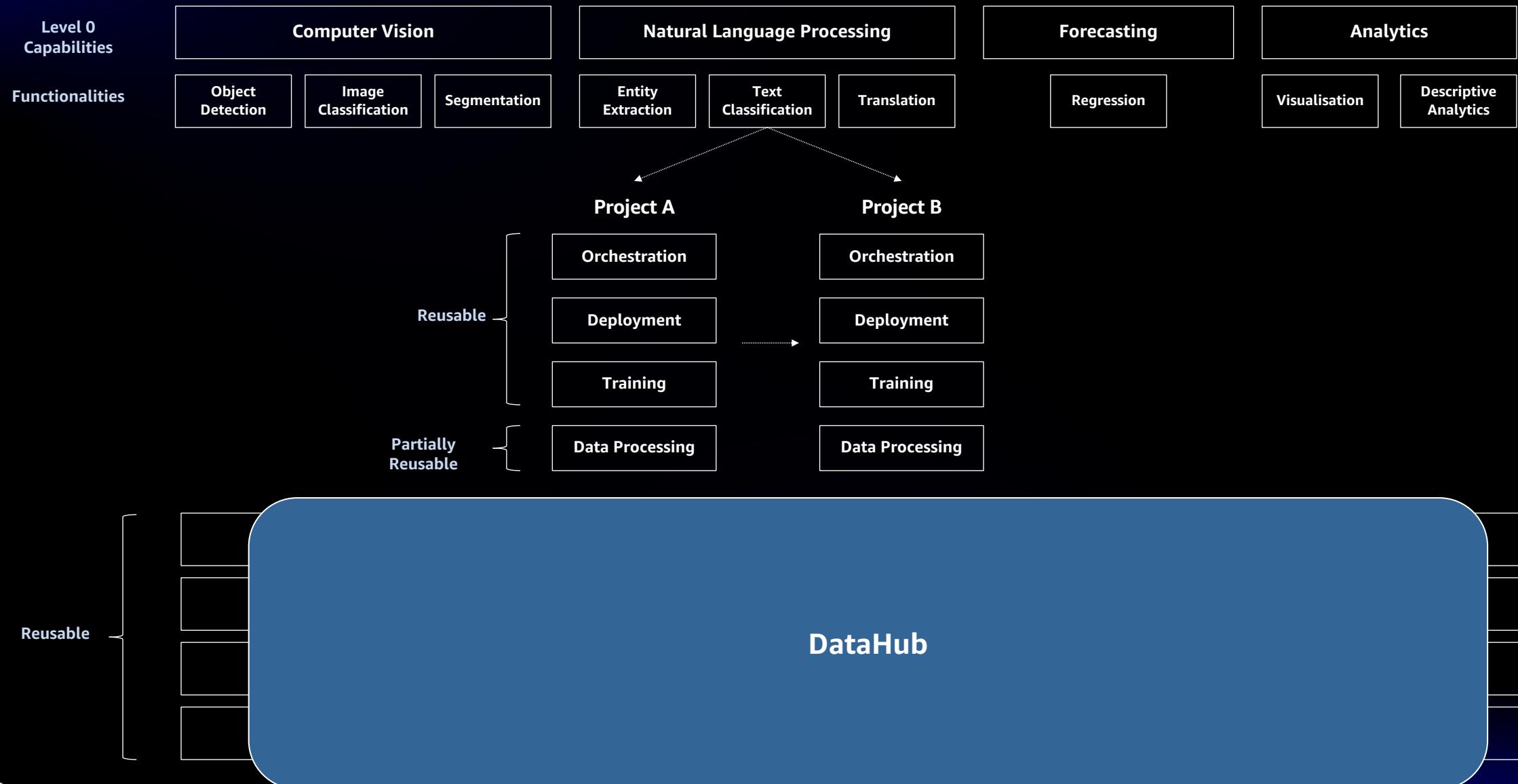
MLOps at scale



Additions:

Scale & Reproducibility by Templates in Service catalog

Reusable ML Template Example



Thank you!

최영준

AI/ML Expert SA

AWS

