



aws INNOVATE

AI/ML EDITION

24 February 2022

Amazon SageMaker를 이용한 시계열 예측 모델 개발 및 MLOps 구성

강민재

프로토타이핑 엔지니어

AWS



Agenda

Why Forecast?

Amazon SageMaker를 이용한 시계열 예측 모델 개발

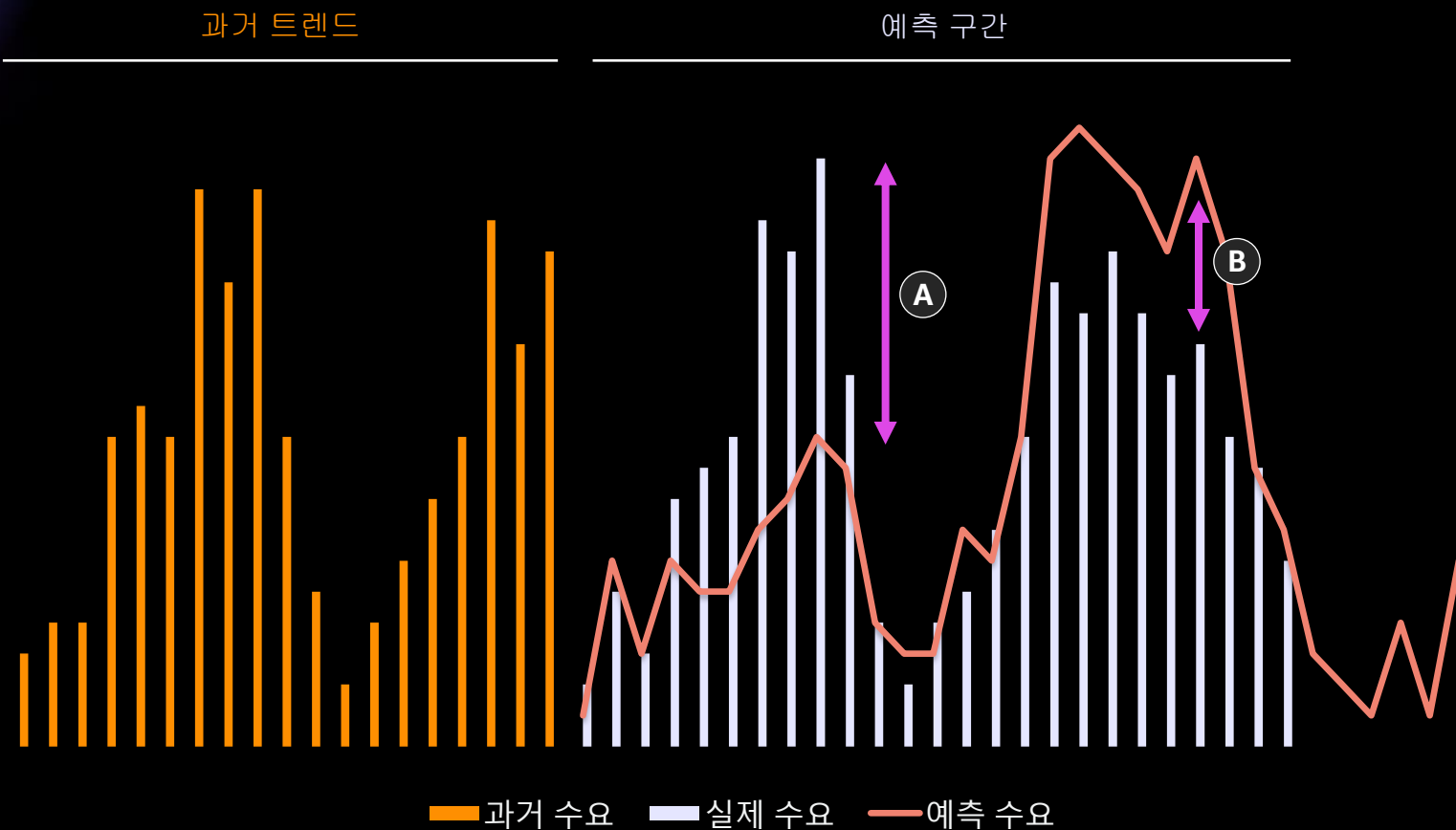
시계열 예측 모델을 위한 MLOps Pipeline

SageMaker Pipelines를 이용한 시계열 예측 모델의 MLOps 구성

Why Forecast?

시계열 예측이란?

과거 트렌드로부터 미래의 트렌드를 예측하는 기술



시계열 예측 문제 해결의 목표:

- Ⓐ 과소 예측으로 인한 비즈니스 기회 손실
 - Ⓑ 과다 예측으로 인한 비즈니스 자원 낭비
- 의 최소화

시계열 예측 모델 적용 사례

INVENTORY PLANNING



상품 수요 예측

적정 수준의 재고 확보를 통한 효율적인 재고 관리

WORKFORCE PLANNING



인력 수요 예측

효율적인 인력 관리 및 고용 계획 수립을 통한
고객 만족도 향상

CAPACITY PLANNING



설비 등의 인프라 수요 예측

자본 활용률 증대

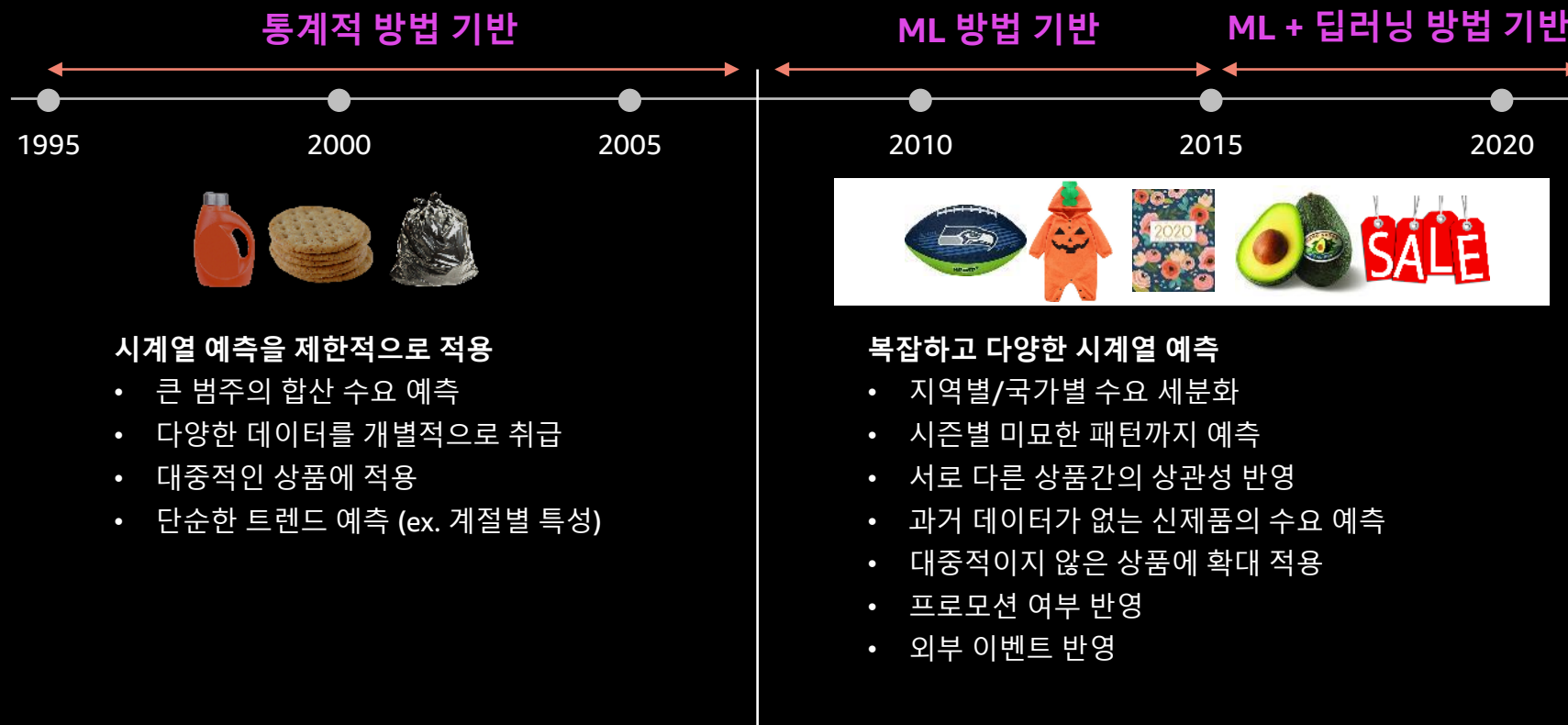
FINANCIAL PLANNING



매출, 영업 이익 등의 재무 지표 예측

효율적인 현금 흐름 관리

지난 20+년간의 AMAZON.COM 사례로 살펴보는 시계열 예측 기술 발전의 역사



Amazon SageMaker를 이용한 시계열 예측 모델 개발

Amazon SageMaker 소개

Key Benefits

ML 모델을 빠르게 구축, 훈련 및 배포할 수 있도록 하는 완전 관리형 서비스



- ML 모델 개발 가속화
- 엔드투엔드(end-to-end) ML 워크플로의 모든 단계를 자동화하고 확장할 수 있는 20개 이상의 도구 제공



- 데이터 과학자들의 생산성 향상
- ML 만을 위해 설계된 IDE인 SageMaker Studio 제공



- ML 모델 개발, 운영에 필요한 비용 절감
- ML 워크플로에 최적화된 기능과 도구들을 통합하여 비용 최적화된 ML 어플리케이션 및 시스템 구축



- Cloud IT, DevOps 에 최적화된 ML 환경 구축
- 타 AWS 서비스들과의 결합을 통해 보안, 관제 등 운영 및 유지보수에 필수적인 인프라 구축

SageMaker를 이용한 시계열 예측 모델 개발

SageMaker Built-in Algorithms

Amazon SageMaker에 기본 내장된 알고리즘

- Amazon SageMaker에 완전 최적화되어 손쉬운 사용 가능
- 다양한 ML 태스크를 위한 10가지 이상의 알고리즘 제공
- DeepAR, XGBoost를 활용해 시계열 예측 모델 구현
- Low-level 커스터마이징은 어려움

SageMaker를 이용한 시계열 예측 모델 개발



시계열 예측 모델 개발을 위한 Apache
MxNet 기반의 ML 툴킷

- DeepAR, Prophet, LSTM 등 기존 유명 알고리즘부터 N-BEATS, TFT 등의 SOTA 알고리즘까지 다양한 시계열 특화 알고리즘을 빌트인으로 제공
- Amazon SageMaker와의 높은 호환성으로 효율적인 시계열 예측 모델 개발 및 운영 가능
- 오픈 소스 라이브러리로 모델 아키텍처를 Low-level 커스터마이징 가능

시계열 예측 모델을 위한 **MLOps Pipeline**

Amazon SageMaker Pipelines

AWS의 CI/CD 서비스와 SageMaker 워크플로를 통합하여 완전 자동화된 MLOps 구축



SageMaker Python SDK를 이용해 쉽고 빠르게 ML 워크플로 구축



실시간으로 파이프라인의 실행 상태, 학습 지표 등 모니터링



모델 레지스트리를 통한 효율적인 모델 관리



실행 이력, 데이터, 모델 학습 이력 등 추적 기능

ML 워크플로 구축

- SageMaker Python SDK를 이용해 모델 개발과 유사한 방식으로 MLOps 파이프라인 구축
- 데이터 전처리, 학습, 모니터링, 승인, 배포 등 CI/CD 및 MLOps의 전 과정을 구현 가능

```
# Example processing step for feature engineering

step_process = ProcessingStep(
    name="PreprocessAbaloneData",
    processor=sklearn_processor,
    outputs=[
        ProcessingOutput(output_name="train", source="/opt/ml/processing/train"),
        ProcessingOutput(output_name="validation", source="/opt/ml/processing/validation"),
        ProcessingOutput(output_name="test", source="/opt/ml/processing/test"),
    ],
    code=os.path.join(BASE_DIR, "preprocess.py"),
    job_arguments=["--input-data", input_data],
)

# Example training step

step_train = TrainingStep(
    name="TrainAbaloneModel",
    estimator=xgb_train,
    inputs={
        "train": TrainingInput(
            s3_data=step_process.properties.ProcessingOutputConfig.Outputs[
                "train"
            ].S3Output.S3Uri,
            content_type="text/csv",
        ),
        "validation": TrainingInput(
            s3_data=step_process.properties.ProcessingOutputConfig.Outputs[
                "validation"
            ].S3Output.S3Uri,
            content_type="text/csv",
        ),
    },
)
```

모니터링

- 코드로 정의한 파이프라인 구조를 다이어그램 형태로 확인
- 파이프라인 실행 상태를 SageMaker Studio에서 실시간으로 모니터링
- 각 실행 단계별 출력, 상세 정보, 실행 로그, 메트릭을 실시간으로 확인
- 각 실행 이력과 상세 정보 저장

The screenshot displays the SageMaker Studio interface during a pipeline execution. At the top, a status bar shows a green dot indicating success, the date and time '2/22/2021, 10:57 AM', and the elapsed time '12m46s'. Below this, the pipeline diagram shows five steps: 'PreprocessAbaloneData', 'TrainAbaloneModel', 'EvaluateAbaloneModel', 'CheckMSEAbaloneEvaluation', and 'RegisterAbaloneModel'. The 'TrainAbaloneModel' step is highlighted with a green border. To the right, a panel for 'TrainAbaloneModel' provides detailed metrics and information. The 'Metrics' table shows 'train:rmse' as 1.68371 and 'validation:rmse' as 2.22083. The 'Files' section lists 'model.tar.gz' with its S3 bucket URI. Below this, a section titled 'Customer Churn Prediction - Boosted Trees' provides a pipeline description. At the bottom, the 'Executions' table lists 12 individual execution steps with their status, start time, elapsed time, name, library, and parameters.

Status	Start time	Elapsed time	Name	library_name	num_trees	depth
Failed	08/22/20 13:55 PST	12m 55s	Execution 12	XGBoost	99	5
Completed	08/22/20 13:54 PST	5m 39s	Execution 11	XGBoost	98	3
Stopped	08/22/20 13:53 PST	4m 39s	Execution 10	LightGBM	80	2
Completed	08/22/20 13:52 PST	1m 39s	Execution 9	XGBoost	63	5
Running...	08/22/20 13:51 PST	11m 39s	Execution 8	LightGBM	100	5
Running...	08/22/20 13:50 PST	1m 39s	Execution 7	XGBoost	70	3
Running...	08/22/20 13:45 PST	1m 39s	Execution 6	XGBoost	80	1
Stopped	Open execution details	39s	Execution 5	LightGBM	78	5
Completed	Stop execution	39s	Execution 4	XGBoost	80	3
Running...	08/22/20 13:42 PST	1m 39s	Execution 3	XGBoost	49	4
Running...	08/22/20 13:41 PST	1m 39s	Execution 2	XGBoost	80	5
Running...	08/22/20 13:40 PST	1m 39s	Execution 1	XGBoost	80	2

모델 레지스트리

- 학습 결과 생성된 모델들의 저장소
- 모델의 파일 및 메타 데이터와 관련 메트릭 등을 조회
- 팀원과 모델을 공유하고 필요시 SageMaker Python SDK 또는 SageMaker Studio를 통해 배포

The screenshot displays the Amazon SageMaker Studio interface. On the left, the 'Components and registries' sidebar shows a search bar and a list of model groups. The main panel shows the 'Recommendations Model - Latin America' group, which contains 6 versions. A table lists the versions with their status, step, description, and modification date. Version 6 is currently in 'Pending' status with the step 'Staging'. A context menu is open for version 6, showing options to 'Open model version' and 'Update model version status...'. The bottom status bar indicates 'Python 3 (Data Science) | Idle'.

Name	Status	Step	Description	Status updated by	Modified on	Actions
version 6	Pending	Staging	New model with SKLe...	Pending	10/10/20	⋮
version 5	Approved	Production	Model updated on 8/1...	Jen Cabro		Open model version ↗
version 4	Approved	Archived	Model updated on 7/15...	Jen Cabro		Update model version status...
version 3	Approved	Archived	Model updated on 6/15...	Jen Cabro		
version 2	Approved	Archived	Model built on 5/15/20...	Jen Cabro	10/10/20	
version 1	Approved	Archived	Model built on 4/15/20...	Jen Cabro	10/10/20	

AWS Step Functions

FULLY MANAGED STATE MACHINE ON AWS

AWS Step Functions

AWS 서비스들을 연결해 워크플로를 구성할 수 있는 서버리스 서비스

Use Cases

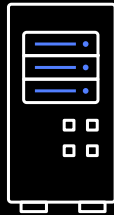
Data
Processing



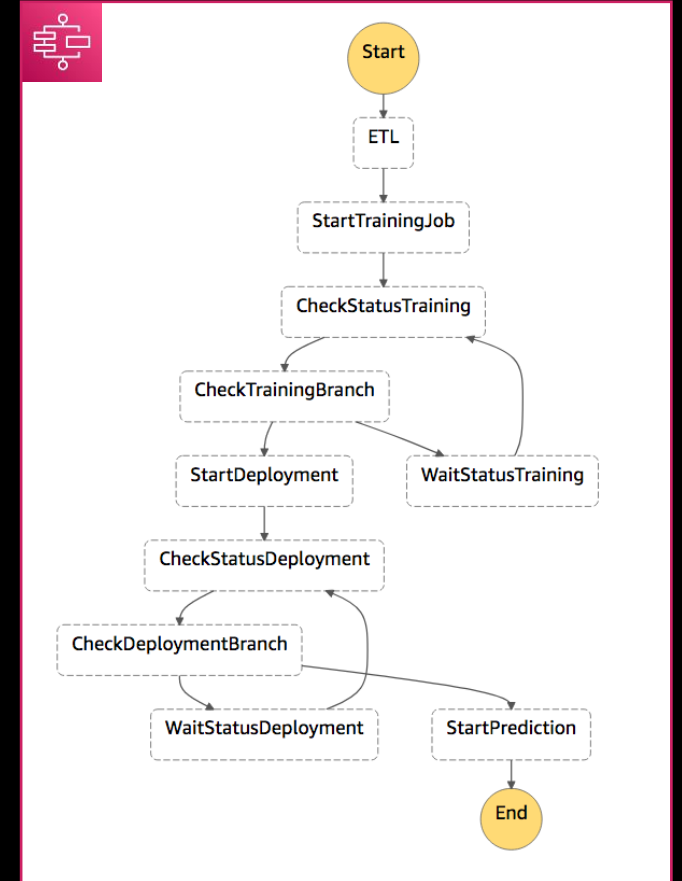
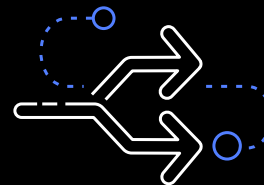
IT & Security
Automation



Machine
Learning



Microservice
Orchestration



AWS Step Functions

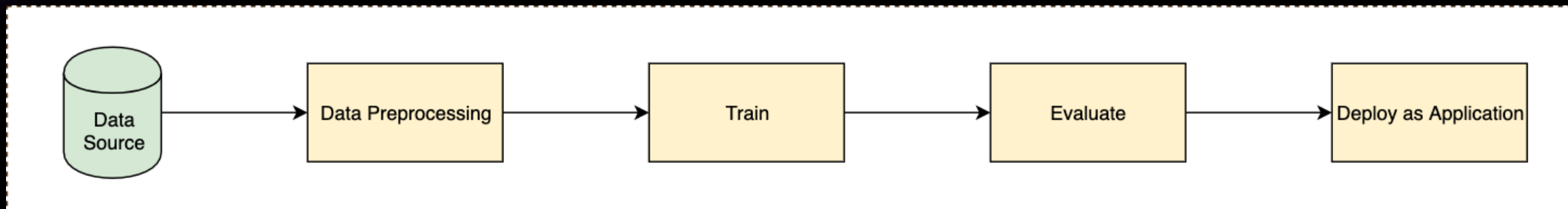
FULLY MANAGED STATE MACHINE ON AWS

- SageMaker 워크플로와 다른 AWS 서비스간의 유연한 결합
- Step Functions Data Science Python SDK를 이용해 구성
- 실시간 실행 상태 추적, 실행 이력, 로그 관리 등 강력한 모니터링 기능
- MSA 어플리케이션을 작성하고, 중앙 관제하기 위한 빌트인 기능 제공
- Pay Per Use 과금
- AWS CDK, Terraform, Serverless Framework 등의 IaC 프레임워크를 통해 전체 파이프라인 구성

SageMaker Pipelines를 이용한 시계열 예측 모델의 MLOps 구성

일반적인 **MLOps** 파이프라인 구조

MLOps Pipeline Workflow



- 데이터 인입 시 전처리, 재학습, 평가, 배포의 과정이 한번에 발생
- 실시간 추론을 위해 모델에 필요 리소스를 할당한 상태로 대기

시계열 예측 모델의 **MLOps** 구성

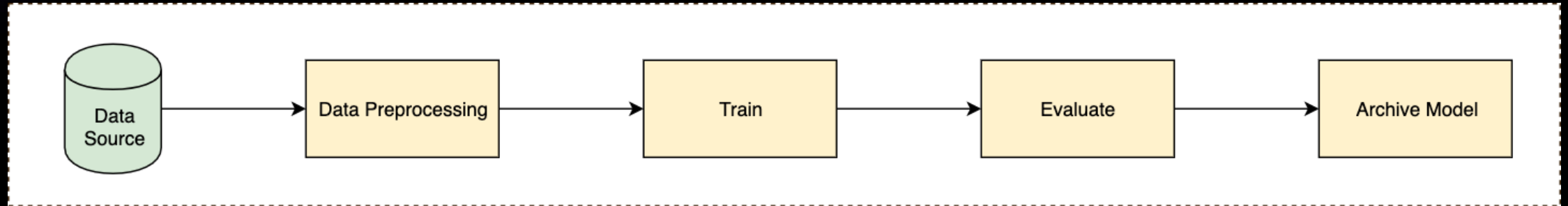
REQUIREMENTS

- 시계열 데이터의 Frequency에 따라 재학습, 추론의 주기 결정
- 실시간 추론을 위해 모델이 리소스를 가지고 대기할 필요가 없음
- 한번 추론 시 전체 시계열의 예측치를 모두 업데이트: Batch Inference
- 재학습은 Long Term, 추론은 Short Term: 파이프라인 분리

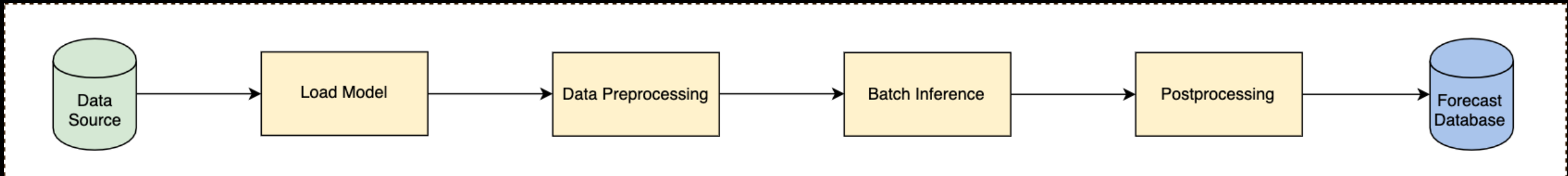
시계열 예측 모델의 MLOps 구성

HIGH LEVEL ARCHITECTURE

MLOps Pipeline Workflow (Model Build)



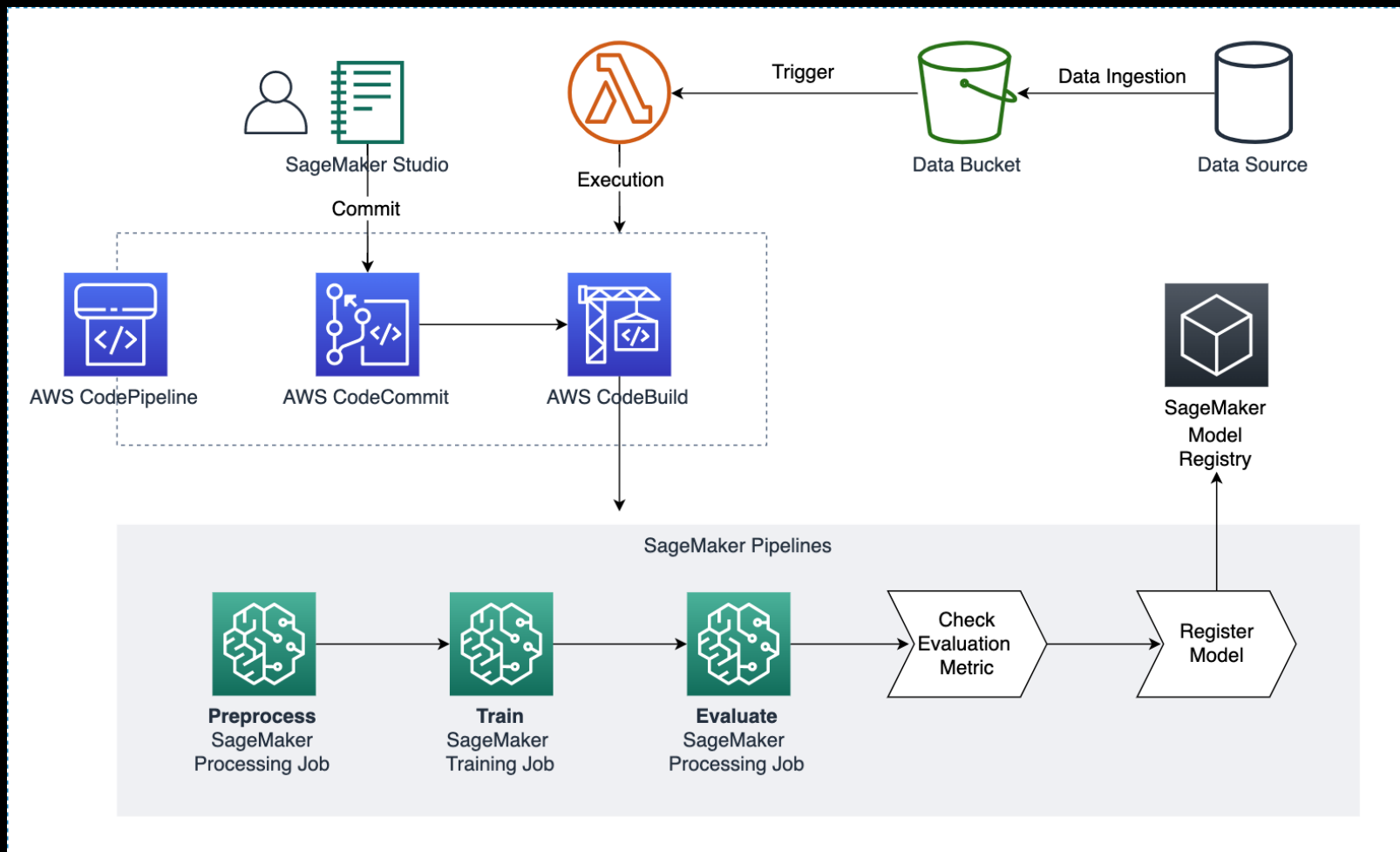
MLOps Pipeline Workflow (Deploy & Infer)



시계열 예측 모델의 MLOps 구성

DEMO ARCHITECTURE POWERED BY SAGEMAKER PIPELINES

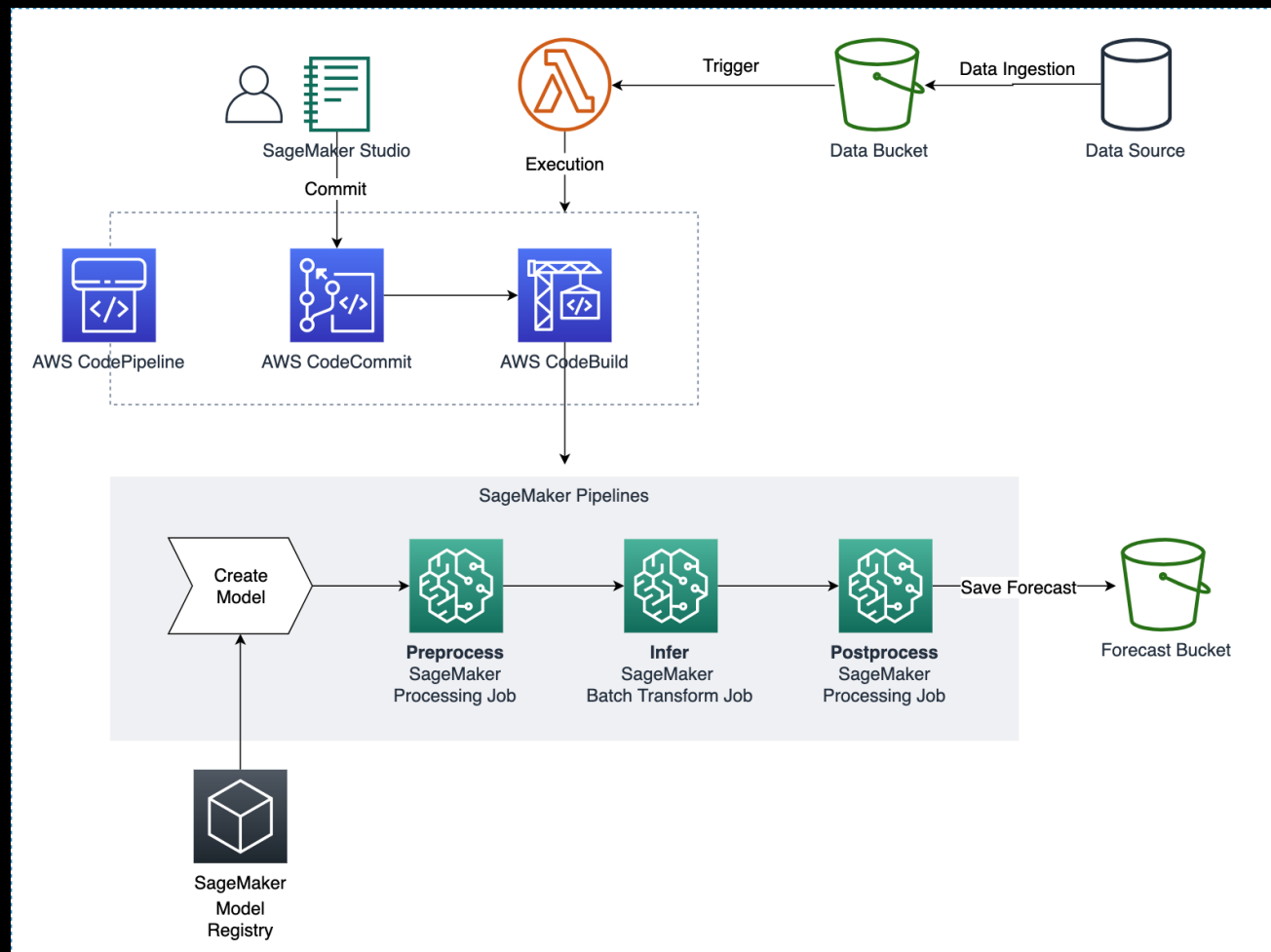
Model Build



시계열 예측 모델의 MLOps 구성

DEMO ARCHITECTURE POWERED BY SAGEMAKER PIPELINES

Model Deploy & Infer



Demo





X

Values are in kW of each 15 min. To convert values in kWh values must be divided by 4.

1	"";"MT_001";"MT_002";"MT_003";"MT_004";"MT_005"
2	"2011-01-01 00:15:00";0;0;0;0;0;0;0;0;0;0;0;0;0;0;0;
3	"2011-01-01 00:30:00";0;0;0;0;0;0;0;0;0;0;0;0;0;0;0;
4	"2011-01-01 00:45:00";0;0;0;0;0;0;0;0;0;0;0;0;0;0;0;
5	"2011-01-01 01:00:00";0;0;0;0;0;0;0;0;0;0;0;0;0;0;0;
6	"2011-01-01 01:15:00";0;0;0;0;0;0;0;0;0;0;0;0;0;0;0;
7	"2011-01-01 01:30:00";0;0;0;0;0;0;0;0;0;0;0;0;0;0;0;
8	"2011-01-01 01:45:00";0;0;0;0;0;0;0;0;0;0;0;0;0;0;0;
9	"2011-01-01 02:00:00";0;0;0;0;0;0;0;0;0;0;0;0;0;0;0;
10	"2011-01-01 02:15:00";0;0;0;0;0;0;0;0;0;0;0;0;0;0;0;
11	"2011-01-01 02:30:00";0;0;0;0;0;0;0;0;0;0;0;0;0;0;0;
12	"2011-01-01 02:45:00";0;0;0;0;0;0;0;0;0;0;0;0;0;0;0;
13	"2011-01-01 03:00:00";0;0;0;0;0;0;0;0;0;0;0;0;0;0;0;
14	"2011-01-01 03:15:00";0;0;0;0;0;0;0;0;0;0;0;0;0;0;0;
15	"2011-01-01 03:30:00";0;0;0;0;0;0;0;0;0;0;0;0;0;0;0;
16	"2011-01-01 03:45:00";0;0;0;0;0;0;0;0;0;0;0;0;0;0;0;
17	"2011-01-01 04:00:00";0;0;0;0;0;0;0;0;0;0;0;0;0;0;0;
18	"2011-01-01 04:15:00";0;0;0;0;0;0;0;0;0;0;0;0;0;0;0;
19	"2011-01-01 04:30:00";0;0;0;0;0;0;0;0;0;0;0;0;0;0;0;
20	"2011-01-01 04:45:00";0;0;0;0;0;0;0;0;0;0;0;0;0;0;0;
21	"2011-01-01 05:00:00";0;0;0;0;0;0;0;0;0;0;0;0;0;0;0;
22	"2011-01-01 05:15:00";0;0;0;0;0;0;0;0;0;0;0;0;0;0;0;
23	"2011-01-01 05:30:00";0;0;0;0;0;0;0;0;0;0;0;0;0;0;0;
24	"2011-01-01 05:45:00";0;0;0;0;0;0;0;0;0;0;0;0;0;0;0;
25	"2011-01-01 06:00:00";0;0;0;0;0;0;0;0;0;0;0;0;0;0;0;
26	"2011-01-01 06:15:00";0;0;0;0;0;0;0;0;0;0;0;0;0;0;0;
27	"2011-01-01 06:30:00";0;0;0;0;0;0;0;0;0;0;0;0;0;0;0;
28	"2011-01-01 06:45:00";0;0;0;0;0;0;0;0;0;0;0;0;0;0;0;
29	"2011-01-01 07:00:00";0;0;0;0;0;0;0;0;0;0;0;0;0;0;0;
30	"2011-01-01 07:15:00";0;0;0;0;0;0;0;0;0;0;0;0;0;0;0;

AI & ML 리소스 허브

AWS가 제공하는 AI 및 ML에 관한 다양한 자료들을 통해 더욱 심층적으로 학습해보세요!

- 기계 학습 여정 가이드
- 기계 학습의 7가지 주요 사용 사례
- 데이터, 분석 및 기계 학습을 위한 전략 플레이북
- 올바른 클라우드 서비스 및 인프라를 통한 기계 학습 혁신 가속화 전략 가이드
- 기계 학습에 적합한 컴퓨팅 인프라 선택 가이드
- 컨택트 센터의 서비스 개선 및 비용 절감 방법
- + 외의 다양한 동영상 학습 자료 및 기술 학습 자료!



<https://bit.ly/3yUk0Kx>

리소스 허브 방문하기

AWS Innovate - AI/ML 특집에 참석해주셔서 대단히 감사합니다.

저희가 준비한 강연, 어떻게 보셨나요?
더 나은 세미나를 위하여 **설문을 꼭 작성해 주세요!**



aws-korea-marketing@amazon.com



twitter.com/AWSKorea



facebook.com/amazonwebservices.ko



youtube.com/user/AWSKorea



linkedin.com/company/amazon-web-services



twitch.tv/aws

Thank you!