

## 9. 세이지메이커 스튜디오 I

### 1강. 세이지메이커 스튜디오란?

#### 학습목표

- 세이지메이커 스튜디오의 개요를 설명할 수 있다.
- 세이지메이커 스튜디오의 주요 기능을 설명할 수 있다.
- 세이지메이커 스튜디오의 실험정보 모니터링 기능에 대해 설명할 수 있다.

#### 학습내용

- 세이지메이커 스튜디오 개요
- 세이지메이커 스튜디오 주요 기능
- 실험정보 모니터링 기능

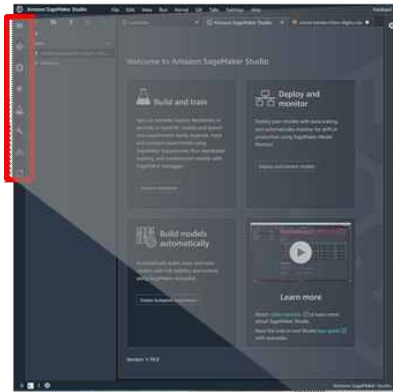
#### ■ 세상을 잇(IT)다!

- 세이지메이커 스튜디오
  - ✓ Amazon SageMaker Studio 출시
  - ✓ Amazon SageMaker Studio는 ML 개발에 필요한 모든 도구를 통합
  - ✓ 개발자는 단일 통합 비주얼 인터페이스 내에서 코드를 작성하고 실험을 추적하며 데이터를 시각화하고 디버깅 및 모니터링을 수행할 수 있으므로 개발자 생산성이 크게 향상됨
  - ✓ ML 워크플로의 모든 단계가 환경 내부에서 추적되므로 개발자는 단계 사이를 빠르게 앞뒤로 이동할 수 있으며 단계를 복제, 조정 및 재생함
  - ✓ 개발자는 신속하게 변경 사항을 적용하고 결과를 관찰하며 더 빠르게 반복 작업을 수행할 수 있으므로 고품질 ML 솔루션의 출시 시간이 단축됨

#### 1. 세이지메이커 스튜디오 개요





- Amazon SageMaker Studio
  - ✓ 기계학습을 위한 통합 개발 도구
  - ✓ ML 코드 작성, 실험을 추적하여 데이터 시각화, 디버깅, 모니터링을 수행


- Amazon SageMaker Studio UI
  - ✓ Amazon SageMaker Studio 시작화면



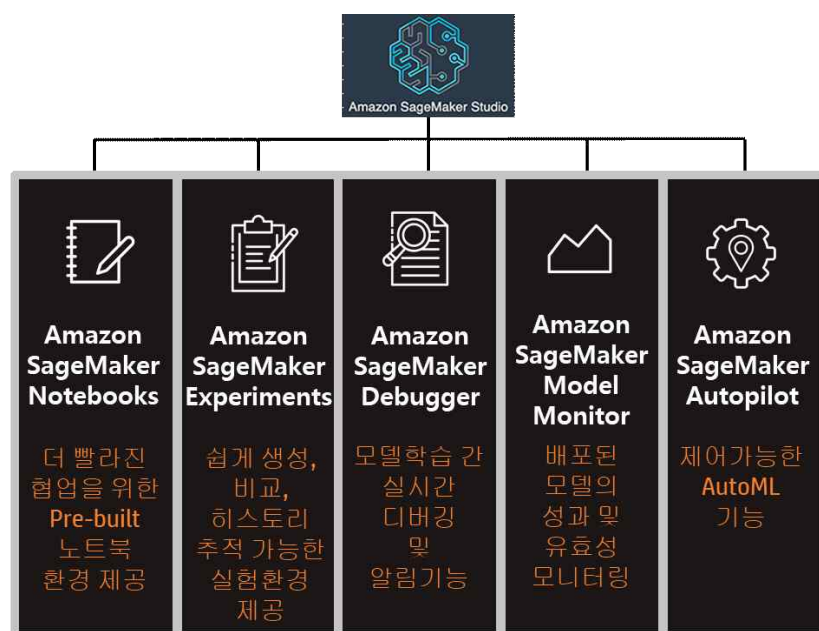
- 파일브라우저
- Git
- 터미널, 커널 이미지 실행
- 명령
- 세이지메이커 실험 목록
- 노트북도구
- 세이지메이커 엔드포인트목록
- 탭열기

- ✓ 왼쪽 사이드 바 아이콘

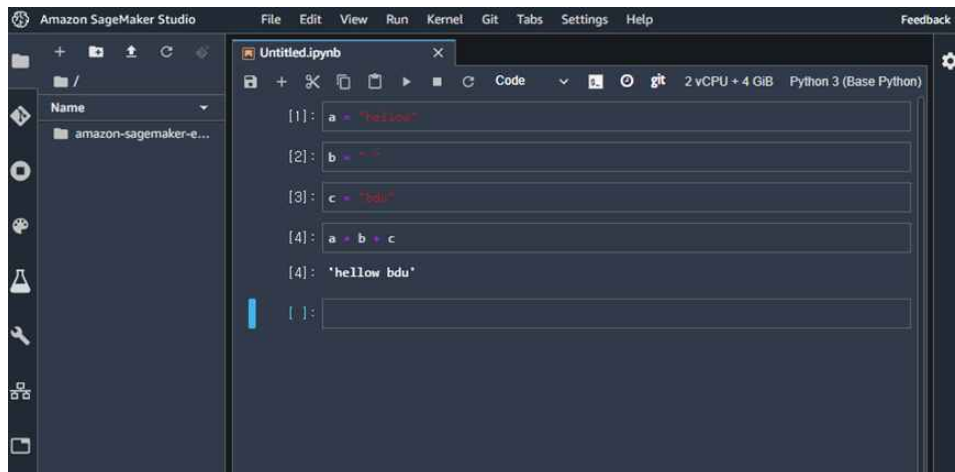
 <p>파일 브라우저</p>	<ul style="list-style-type: none"> <li>- 파일을 Studio에 추가 하거나 새 탭에서 파일을 열 수 있음</li> <li>- 파일 브라우저의 상단에 있는 메뉴에서 더하기(+) 기호를 선택하여 Studio 시작 관리자를 열 수 있음</li> </ul>
 <p>Git</p>	<ul style="list-style-type: none"> <li>- Git 리포지토리에 연결하여 전체 Git 도구 및 연산에 액세스할 수 있음</li> </ul>
 <p>터미널, 커널 이미지 실행</p>	<ul style="list-style-type: none"> <li>- 노트북, 터미널, 커널 및 이미지를 비롯한 개별리소스를 종료하거나 이러한 범주 중 하나에 있는 모든 리소스를 동시에 종료할 수 있음</li> </ul>
 <p>명령</p>	<ul style="list-style-type: none"> <li>- 대부분의 메뉴 명령을 사용할 수 있음 (콘솔, 확장관리자, 파일작업, HELP 등)</li> </ul>

 <p>세이지메이커 실험 목록</p>	<ul style="list-style-type: none"> <li>- 실험은 관련 목적을 가진 여러 시험으로 구성</li> <li>- 시험은 데이터 사전 처리 작업 및 훈련 작업과 같은 하나 이상의 시험 구성 요소로 구성</li> <li>- 실험 브라우저를 사용하여 이러한 엔터티(Entity)에 대한 자세한 정보를 볼 수 있음</li> </ul>
 <p>노트북 도구</p>	<ul style="list-style-type: none"> <li>- 고급 도구 섹션을 통해 노트북의 메타데이터에 액세스할 수 있음</li> <li>- 노트북이 열려있을 때만 표시</li> </ul>
 <p>세이지메이커 엔드포인트 목록</p>	<ul style="list-style-type: none"> <li>- 배포된 모델의 엔드포인트 목록을 제공</li> <li>- Studio 랜딩 페이지에서 배포 및 모니터링을 선택할 때만 선택 가능</li> </ul>
 <p>탭열기</p>	<ul style="list-style-type: none"> <li>- 열린 탭의 목록을 제공하며, 이 목록은 열려 있는 탭이 여러 개인 경우에 유용함</li> </ul>

## 2. 세이지메이커 스튜디오 주요 기능



- Amazon SageMaker Notebook
  - ✓ AWS SSO로 간편한 로그인 가능
  - ✓ 안전한 접근과 권한제어
  - ✓ 자유로운 컴퓨팅 인스턴스 변경
  - ✓ 원클릭으로 URL 공유
  - ✓ 노트북상에서 제공되는 손쉬운 작업재현과 리뷰



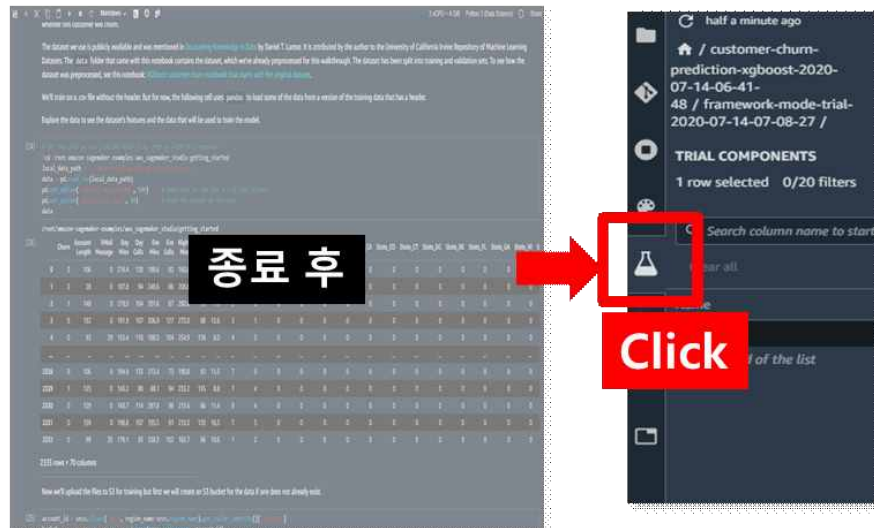
### 생각해보기

- 세이지메이커 노트북 인스턴스 vs 세이지메이커 (스튜디오) 노트북
  - ✓ Studio 노트북을 시작하는 것이 인스턴스 기반 노트북을 시작하는 것보다 빠름
  - ✓ SageMaker Studio의 통합기능 중 하나인 노트북 공유 가능
  - ✓ SageMaker Studio 노트북은 최신 Amazon SageMaker Python SDK가 사전 설치되어 제공
  - ✓ SageMaker Studio 노트북은 Studio 내에서 액세스할 수 있음
    - Studio를 나가지 않고 모델을 빌드, 훈련, 디버그, 추적 및 모니터링 가능
  - ✓ AWS SSO를 사용하는 경우 고유한 URL을 통해 SSO 자격 증명을 사용하여 SageMaker Studio에 직접 액세스
  - ✓ 노트북을 실행하기 위해 AWS Management 콘솔과 상호 작용할 필요가 없음
  - ✓ Studio 노트북에는 사전 정의된 SageMaker 이미지 설정 세트가 탑재되어 있어 보다 빠르게 시작할 수 있음
  - ✓ Studio 노트북은 로컬 모드를 지원하지 않으나 노트북 인스턴스를 사용하여 데이터 세트의 샘플을 로컬에서 훈련시킨 다음 Studio 노트북에서 동일한 코드를 사용하여 전체 데이터 세트를 훈련시킬 수 있음

- Amazon SageMaker Experiments
  - ✓ 실험의 모델 파라미터, 성능지표 등을 기록하고 관리
  - ✓ 반복적인 실험과 시도(Trial)들을 체계적으로 관리
  - ✓ 원하는 만큼 반복적으로 실험하고, 실험 간 비교 평가
  - ✓ API 호출로 제공되는 실험 트래킹과 분석
- Amazon SageMaker Debugger
  - ✓ TensorFlow, Apache MXnet, Pytorch, XGBoost 등의 프레임워크 지원
  - ✓ 학습이 진행되는 동안 텐서들의 모니터링 변화 수집 및 집계 가능
  - ✓ 룰로 정의한 학습의 실패상황들을 감지하여 실시간 알림
  - ✓ 커스텀 룰을 생성하여 사용 가능
- Amazon SageMaker Monitor
  - ✓ 엔드포인트를 모니터링 하며 주기적으로 데이터를 S3에 수집
  - ✓ 모델에 사용되는 입력 특징들의 통계적인 분석 진행
  - ✓ 입력 데이터를 모니터링 하여 모델링 대상의 통계적 특성 변화 감지
  - ✓ 변화감지 시 Amazon CloudWatch와 Amazon SageMaker Studio를 통해 알림
- Amazon SageMaker Autopilot
  - ✓ 테이블 형태 데이터 지원
  - ✓ Regression과 Classification 모델 지원
  - ✓ Feature 생성, 알고리즘 선택, 하이퍼파라미터 튜닝 자동화
  - ✓ 모든 실험과정은 세이지메이커 스튜디오 내에서 확인가능
  - ✓ 노트북형태로 소스코드 지원

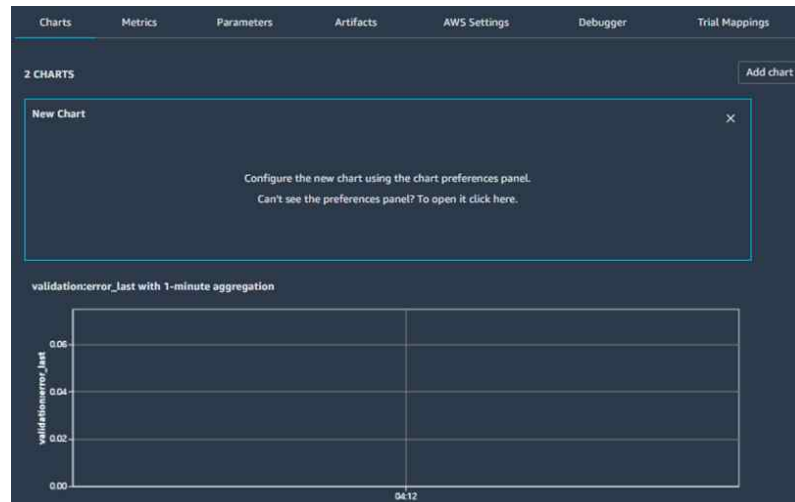
### 3. 실험정보 모니터링 기능

- ✓ 모델 훈련 종료 후 실험 아이콘을 클릭하여 실험 정보 확인 가능



- ✓ Describe Trial Component 탭에는 7개의 헤더가 있음

- 차트(Charts) : 고유한 차트 작성



- 매트릭스(Metrics) : 실험 실행 중 Tracker에서 기록되는 지표

Charts	Metrics	Parameters	Artifacts	AWS Settings	Debugger	Trial Mappings
Name	Minimum	Maximum	Standard Deviation	Final value		
validation:error	0.061562	0.121622	0.012694315843445752	0.069069		
train:error	0.030004	0.093871	0.013608132137118562	0.030433		

- 파라미터(Parameters) : 하이퍼파라미터 값 및 인스턴스 정보

Charts	Metrics	Parameters	Artifacts	AWS Settings	Debugger	Trial Mappings
		Name	Value			
		SageMaker.ImageUri	141502667606.dkr.ecr.eu-west-1.amazonaws.com/sagemaker-xgboost:0.90...			
		SageMaker.InstanceCount	1			
		SageMaker.InstanceType	ml.m4.xlarge			
		SageMaker.VolumeSizeInGB	30			
		eta	0.2			
		gamma	4			
		max_depth	5			
		min_child_weight	10			
		num_round	600			
		objective	"binary:logistic"			
		sagemaker_container_log_level	20			
		sagemaker_enable_cloudwatch_metrics	false			
		sagemaker_job_name	"demo-xgboost-customer-churn-2020-07-14-07-08-27-610"			
		sagemaker_program	"xgboost_customer_churn.py"			
		sagemaker_region	"eu-west-1"			
		sagemaker_submit_directory	"s3://sagemaker-studio-eu-west-1-402345010737/demo-xgboost-custome..."			
		silent	0			
		subsample	0.8			

- 아티팩트(Artifacts) : 입력 데이터 세트 및 출력 모델용 S3 위치

Charts	Metrics	Parameters	Artifacts	AWS Settings	Debugger	Trial Mappings
<b>Input artifacts (2)</b>						
train						
s3://sagemaker-studio-eu-west-1-402345010737/xgboost-churn/train						
validation						
s3://sagemaker-studio-eu-west-1-402345010737/xgboost-churn/validation/						
<b>Output artifacts (2)</b>						
SageMaker.DebugHookOutput						
s3://sagemaker-studio-eu-west-1-402345010737/xgboost-churn/output						
SageMaker.ModelArtifact						
s3://sagemaker-studio-eu-west-1-402345010737/xgboost-churn/output/demo-xgboost-customer-churn-2020-07-14-07-08-27-610/output/model.tar.gz						

- AWS 설정(AWS Settings) : 작업 이름, ARN, 상태, 생성 시간, 인스턴스 등 다양한 정보 모니터링

Charts	Metrics	Parameters	Artifacts	AWS Settings	Debugger	Trial Mappings
JOB SETTINGS						
Job name		ARN			Status	
demo-xgboost-customer-churn-2020-07-14-07-08-27-610		arn:aws:sagemaker:eu-west-1:402345010737:training-job/demo-xgboost-customer-churn-2020-07-14-07-08-27-610			Completed	
Creation time		Last modified time			Tuning job source/parent	
5 hours ago		5 hours ago			-	
IAM role ARN						
arn:aws:iam::402345010737:role/service-role/AmazonSageMaker-ExecutionRole-20200713T162151						
USAGE						
Training time (seconds)		Billable time (seconds)			Managed spot training savings	
84		84			0%	
RESOURCES						
Name		Training image			Input mode	
-		141502667606.dkr.ecr.eu-west-1.amazonaws.com/sagemaker-xgboost:0.90-2-cpu-py3			File	
Enable SageMaker metrics time series						
false						

- 디버거 (Debugger) : 디버거 규칙 및 발견된 문제 목록

Charts	Metrics	Parameters	Artifacts	AWS Settings	Debugger	Trial Mappings
						⚙
Status	Last modified		Rule name		Job ARN	
No Issues Found	4 hours ago		LossNotDecreasing		arn:aws:sagemaker:eu-west-1:4023...	
No Issues Found	4 hours ago		Overtraining		arn:aws:sagemaker:eu-west-1:4023...	
No Issues Found	4 hours ago		Overfit		arn:aws:sagemaker:eu-west-1:4023...	

- 시도 매핑(Trial Mappings) : 변환 작업 중 '시도 매핑'에서 이 작업 중에 사용된 '시도(Trial)'에 대한 세부 정보를 얻음

Charts	Metrics	Parameters	Artifacts	AWS Settings	Debugger	Trial Mappings
Name					ARN	
uning-job-1-e6a38d67c4094b088e-250-5e43800d-aws-trial					arn:aws:sagemaker:us-east-2:489534398131:experiment-trial/tuning-job-	
uning-job-1-e6a38d67c4094b088e-249-d71e1581-aws-trial					arn:aws:sagemaker:us-east-2:489534398131:experiment-trial/tuning-job-	
uning-job-1-e6a38d67c4094b088e-248-c1f7f108-aws-trial					arn:aws:sagemaker:us-east-2:489534398131:experiment-trial/tuning-job-	
uning-job-1-e6a38d67c4094b088e-246-3bfccac4-aws-trial					arn:aws:sagemaker:us-east-2:489534398131:experiment-trial/tuning-job-	
uning-job-1-e6a38d67c4094b088e-247-5b530b5c-aws-trial					arn:aws:sagemaker:us-east-2:489534398131:experiment-trial/tuning-job-	
uning-job-1-e6a38d67c4094b088e-245-46c99cb4-aws-trial					arn:aws:sagemaker:us-east-2:489534398131:experiment-trial/tuning-job-	
uning-job-1-e6a38d67c4094b088e-238-da32c295-aws-trial					arn:aws:sagemaker:us-east-2:489534398131:experiment-trial/tuning-job-	
uning-job-1-e6a38d67c4094b088e-241-b15fff3a-aws-trial					arn:aws:sagemaker:us-east-2:489534398131:experiment-trial/tuning-job-	
uning-job-1-e6a38d67c4094b088e-243-89b175a1-aws-trial					arn:aws:sagemaker:us-east-2:489534398131:experiment-trial/tuning-job-	
uning-job-1-e6a38d67c4094b088e-242-95b4e03d-aws-trial					arn:aws:sagemaker:us-east-2:489534398131:experiment-trial/tuning-job-	
uning-job-1-e6a38d67c4094b088e-239-8a924bb7-aws-trial					arn:aws:sagemaker:us-east-2:489534398131:experiment-trial/tuning-job-	
uning-job-1-e6a38d67c4094b088e-240-5b20ce36-aws-trial					arn:aws:sagemaker:us-east-2:489534398131:experiment-trial/tuning-job-	
uning-job-1-e6a38d67c4094b088e-237-7ee7d56b-aws-trial					arn:aws:sagemaker:us-east-2:489534398131:experiment-trial/tuning-job-	
uning-job-1-e6a38d67c4094b088e-236-d382e382-aws-trial					arn:aws:sagemaker:us-east-2:489534398131:experiment-trial/tuning-job-	



**평가하기****1. 세이지메이커 스튜디오의 주요 기능으로 볼 수 없는 것은?**

- ① 세이지메이커 Notebooks
- ② 세이지메이커 Ground Truth
- ③ 세이지메이커 Debugger
- ④ 세이지메이커 Autopilot

- 정답 : ② 번

해설 : 세이지메이커 Ground Truth는 학습 데이터 세트를 신속하게 구축 및 관리하기 위해 제공되는 기능으로 세이지메이커 스튜디오 내부에서 실행되지 않습니다.

**2. 세이지메이커 스튜디오 실험정보 목록에서 제공하는 정보가 아닌 것은?**

- ① 차트
- ② 파라미터
- ③ 아티팩트
- ④ 하이퍼파라미터 튜닝

- 정답 : ④ 번

해설 : 세이지메이커 스튜디오 실험정보 목록에서는 차트, 매트릭스, 파라미터, 아티팩트, AWS 설정, 디버거, 시도매핑의 정보를 확인할 수 있습니다.

**학습정리****1. 세이지메이커 스튜디오 개요**

- 기계학습을 위한 통합 개발 도구
- ML 코드 작성, 실험을 추적하여 데이터 시각화, 디버깅, 모니터링을 수행할 수 있음
- 왼쪽 사이드바 아이콘
  - 파일 브라우저 : 파일 추가 및 열기
  - Git : Git 리포지토리에 연결
  - 터미널, 커널 이미지 실행 : 리소스 종료
  - 명령 : 대부분의 메뉴 명령을 사용할 수 있음
  - 세이지메이커 실험 목록 : 실험 정보 확인
  - 노트북 도구 : 노트북의 메타데이터에 액세스
  - 세이지메이커 엔드포인트 목록 : 엔드포인트 목록 확인
  - 탭 열기 : 열린 탭의 목록 제공