

13. AWS AI 서비스

2강. 수요 예측 : Amazon Forecast

학습목표

- Amazon Forecast 개요, 주요 기능, 알고리즘과 도메인에 대해 설명할 수 있다.
- Amazon Forecast를 사용하여 전기 사용량 예측을 할 수 있다.

학습내용

- Amazon Forecast 개요
- Amazon Forecast 주요 기능
- Amazon Forecast 알고리즘과 도메인
- 실습. Forecast를 사용하여 전기 사용량 예측하기

■ 세상을 잇(IT)다!

- 예측 기술은 왜 어려울까?
 - ✓ 정기적인 외부 요인 고려
 - 계절에 따른 판매량
 - ✓ 비정기적인 외부 요인 고려
 - 비정기적으로 계획해서 이벤트가 발생하는 경우
 - ✓ 메타데이터에 대한 고려
 - 예측하려는 메타데이터에 대한 관리 유무 예측 정확도가 달라짐
 - ✓ 과거 데이터의 보관 유무
 - 과거데이터가 없는 경우 예측이 불가능

1. Amazon Forecast 개요

- Amazon Forecast
 - ✓ 시계열 데이터를 기반으로 매우 정확한 예측을 위해 머신러닝을 사용하는 완전관리형 서비스
 - ✓ 머신러닝의 지식이 없더라도 기록 데이터와 예측에 영향을 줄 수 있는 추가 데이터만 제공하면 사용 가능

- Amazon Forecast 내부 구조



- Amazon Forecast 장점



✓ 정확한 예측

- 머신러닝을 통한 50% 더 정확한 예측
- 머신러닝을 사용해 시계열 데이터와 기타변수(제품 기능, 스토어 위치 등)가 서로 영향을 주는 방식을 자동으로 검색
- 사용하는 데이터에 따라 고유한 모델을 구축하므로 비즈니스별로 사용자 지정 방식의 예측이 가능

✓ 시간 단축

- AWS Forecast를 사용하면 몇 달씩 엔지니어링 과정이 걸리던 예측 정확도 수준을 몇 시간 만에 달성할 수 있음
- 예측 모델을 구축/교육/튜닝/배포하는 데 필요한복잡한 머신러닝을 자동으로 처리하므로 정확한 예측을 빠르게 생성 가능

✓ 모든 시계열 예측

- 소매, 물류, 금융, 광고 성과 등 거의 모든 산업 및 사용 사례에 맞는 예측을 만들어 낼 수 있음
- 머신러닝을 통해 모든 기록 시계열 데이터를 사용할 수 있으며, 내장된 알고리즘으로 구성된대규모 라이브러리를 사용하여 특정 예측 유형에 가장 적합한 데이터를 자동으로 결정

✓ 높은 보안성

- Amazon Forecast에서 처리하는 모든 콘텐츠는 Amazon Key Management 서비스를 통해 고객 키를 사용하여 암호화되며, 저장 시에는 서비스를 사용 중인 AWS 리전에서 암호화
- 관리자는 IAM 권한 정책을 통해 액세스를 제어할 수 있으므로 민감한 정보는 안전하게 기밀로 유지됨

2. Amazon Forecast 주요 기능



• 자동화된 머신러닝

- ✓ Amazon Forecast에는 머신러닝을 관리하는 AutoML 기능이 포함
- ✓ S3에 데이터를 제공하면 Amazon Forecast에서 자동으로 데이터를 로드 및 검사하고, 올바른 알고리즘을 선택하고, 모델을 교육하고, 정확도 매트릭스를 제공하고, 예측을 생성할 수 있음

- 예측 시각화
 - ✓ Amazon Forecast 및 관련 정확도 지표는 콘솔에서 이해하기 쉬운 그래프와 표로 시각화됨
 - ✓ 시각화를 통해 각 예측의 세부 사항을 빠르게 이해하고 조정이 필요한지 결정할 수 있음
- 확률론적 예측 생성
 - ✓ 포인트 예측을 생성하는 다른 예측 솔루션과 달리 Amazon Forecast는 10%, 50%, 90%의 서로 다른 세 가지 분위에서 확률론적 예측을 생성
 - ✓ 또한, '평균' 예측을 포함하여 1%와 99% 사이의 분위기를 선택할 수 있음

3. Amazon Forecast 알고리즘과 도메인

- Amazon Forecast 도메인
 - ✓ Forecast 데이터 세트를 생성할 때는 도메인과 데이터 세트 유형을 선택해야 함
 - ✓ Forecast는 소매 수요나 웹 트래픽을 예측하는 등 다양한 사용 사례에 따라 필요한 도메인을 제공하며 사용자 지정 도메인을 생성할 수 있음
- 미리 정의된 데이터 세트 도메인

RETAIL도메인	소매 수요 예측
INVENTORY_PLANNING 도메인	공급망 및 재고 계획 수립
EC2_CAPACITY 도메인	Amazon EC2용량 예측
WORK_FORCE 도메인	인력 계획
WEB_TRAFFIC 도메인	향후 웹 트래픽 예측
METRICS 도메인	수익 및 현금 흐름과 같은 지표 예측
CUSTOM 도메인	기타 모든 유형의 시계열 예측

- Amazon Forecast 알고리즘
 - ✓ Amazon Forecast는 다양한 알고리즘을 사용하여 모델을 학습한 다음, 이 모델을 사용하여 입력 데이터에 대한 예측을 생성
 - ✓ Amazon Forecast에서는 특정 알고리즘을 선택하거나, AutoML을 선택하여 데이터 세트에 적합한 알고리즘을 적용할 수 있음
- 미리 정의된 알고리즘

Autoregressive Integrated Moving Average(ARIMA)	일반적으로 사용되는 시계열 예측용 로컬 통계 알고리즘
CNN-QR	CNN을 사용하여 스칼라(1차원) 시계열 예측을 위한 머신러닝 알고리즘
DeepAR+	RNN을 사용하여 스칼라(1차원) 시계열을 예상하는 지도 학습 알고리즘
Exponential Smoothing(ETS)	시계열 예상에 일반적으로 사용되는 로컬 통계 알고리즘
Non-Parametric Time Series(NPTS)	과거 관측치로부터 샘플링하여 주어진 시계열의 미래 가치 분포를 예측
Prophet	로컬 베이지안 구조의 시계열 모델

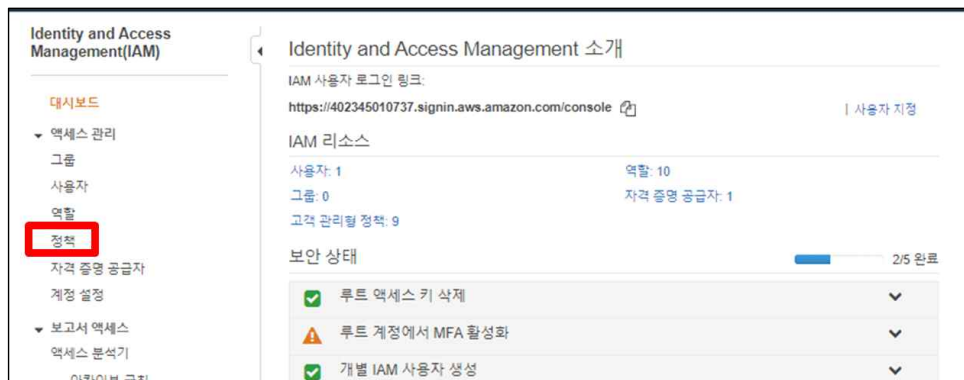
4. 실습. Amazon Forecast를 사용하여 전기 사용량 예측하기

- Forecast 사전 준비
 - ✓ 아래 주소에서 데이터 세트 다운로드 후 압축을 풀고 로컬 컴퓨터에 저장
 - ✓ https://docs.aws.amazon.com/ko_kr/forecast/latest/dg/samples/electricityusagedata.zip
 - 이 연습에서는 개별 가정용 전력 소비량 데이터 세트를 사용합니다. (Dua, D. 및 Karra Taniskidou, E. (2017). UCI Machine Learning 리포지토리[<http://archive.ics.uci.edu/ml>]. Irvine, CA: University of California, School of Information and Computer Science.) AWS에서는 시간별 사용량 데이터를 집계합니다.

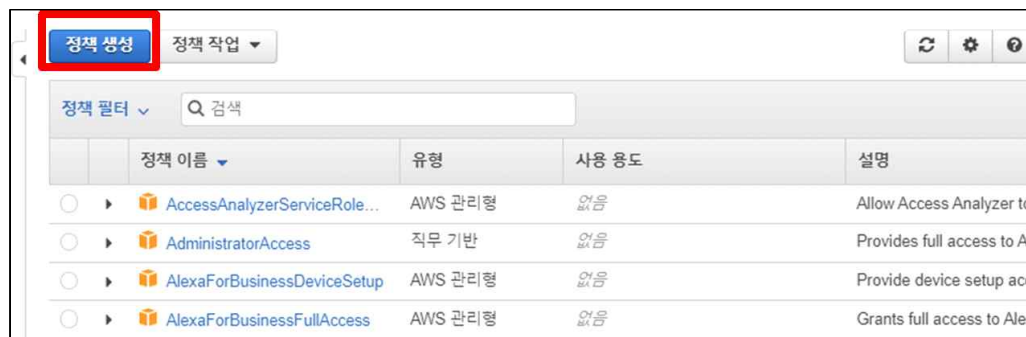
- ✓ S3에 새로운 버킷(기존 버킷도 가능)을 생성 후 압축이 해제된 csv 파일을 업로드



- ✓ IAM 콘솔에 로그인 → 왼쪽 대시보드에서 정책 클릭



- ✓ 정책생성 클릭



- ✓ json 탭을 선택하고 bucket-name을 s3 버킷의 이름으로 바꾼 다음 정책 검토 선택

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Action": [
        "s3:Get*",
        "s3:List*",
        "s3:PutObject"
      ],
      "Resource": [
        "arn:aws:s3:::bucket-name",
        "arn:aws:s3:::bucket-name/*"
      ]
    }
  ]
}
```



- ✓ 정책 검토의 이름에 정책이름을 입력하고 정책에 대한 설명을 입력한 후 오른쪽 하단의 정책 생성 클릭

정책 생성

1 2

정책 검토

이름*
영숫자 및 '-', '_', '@', '.' 문자를 사용합니다. 최대 128자입니다.

설명
최대 1000자입니다. 영숫자 및 '-', '_', '@', '.' 문자를 사용합니다.

요약
이 정책은 권한을 제공하지 않는 일부 작업, 리소스 또는 조건을 정의합니다. 액세스를 허용하려면 정책에 해당 리소스 또는 조건을 갖는 작업이 있어야 합니다. 자세한 내용은 다음을 선택하십시오 (나머지 표시, 자세히 알아보기)

필터:

서비스 ▼ 액세스 레벨 리소스

허용 (1 / 237 서비스) 나머지 236 표시

서비스	액세스 레벨	리소스
S3	제한: 목록, 읽기, 쓰기	여러 값을 설정할 수 있습니다

취소 이전 **정책 생성**

- ✓ 탐색창에서 역할을 선택 역할 만들기 선택

Identity and Access Management(IAM)

대시보드

- ▼ 액세스 관리
 - 그룹
 - 사용자
 - 역할**
 - 정책
 - 자격 증명 공급자
 - 계정 설정
- ▼ 보고서 액세스
- 액세스 분석기

• IAM 역할 설명서
 • 자습서: AWS 계정 간 액세스 권한 위임
 • 역할에 대한 일반적인 시나리오

역할 만들기 역할 삭제

검색

역할 이름 ▼	신뢰할 수 있는 개체	마지막 활동 ▼
<input type="checkbox"/> AmazonSageMak...	AWS 서비스: sagemaker	5 일
<input type="checkbox"/> AmazonSageMak...	AWS 서비스: sagemaker	없음
<input type="checkbox"/> AmazonSageMak...	AWS 서비스: sagemaker	5 일

- ✓ 신뢰할 수 있는 유형의 개체 선택에 aws 서비스 선택 → 사용사례선택에 Forecast 선택 → 검토 세션까지 계속 다음 클릭

역할 만들기 1 2 3

신뢰할 수 있는 유형의 개체 선택

AWS 서비스
EC2, Lambda 및 기타

다른 AWS 계정
귀하 또는 타사 소유

웹 ID
Cognito 또는 OpenID 공급자

SAML 2.0 연동
귀사 디렉터리

AWS 서비스가 사용자를 대신하여 작업을 수행하도록 허용합니다. 자세히 알아보기

사용 사례 선택

일반 사용 사례

EC2
Allows EC2 instances to call AWS services on your behalf.

Lambda
Allows Lambda functions to call AWS services on your behalf.

또는 서비스를 선택하여 해당 서비스의 사용 사례 확인

API Gateway	CodeGuru	ElastiCache	Kinesis	RoboMaker
AWS Backup	CodeStar Notifications	Elastic Beanstalk	Lake Formation	S3
AWS Chatbot	Comprehend	Elastic Container Service	Lambda	SMS
AWS Support	Config	Elastic Transcoder	Lex	SNS
Amplify	Connect	ElasticLoadBalancing	License Manager	SWF
AppStream 2.0	DMS	Forecast	Machine Learning	SageMaker
AppSync	Data Lifecycle Manager	GameLift	Macie	Security Hub
Application Auto Scaling	Data Pipeline	Global Accelerator	Managed Blockchain	Service Catalog

취소 다음: 권한

aws 서비스 리소스 그룹 jyunseo 글로벌 지원

역할 만들기 1 2 3

검토

생성하기 전에 아래에 필요한 정보를 입력하고 이 역할을 검토하십시오.

역할 이름*

영숫자 및 '+', '@', '_' 문자를 사용합니다. 최대 64자입니다.

역할 설명

최대 1000자입니다. 영숫자 및 '+', '@', '_' 문자를 사용합니다.

신뢰할 수 있는 개체 AWS 서비스: forecast.amazonaws.com

정책 AmazonS3FullAccess

권한 경계 권한 경계가 설정되지 않았습니다

태그가 추가되지 않았습니다.

취소 이전 역할 만들기

- ✓ Role name에 대한 설명을 업데이트한 다음 Create role을 선택
 - (역할 이름)에 역할의 이름을 입력(예: Forecast Role). Role description (역할 설명)
- ✓ Summary(요약)에서 Role ARN(역할 ARN) 값을 복사해 저장합니다. 데이터 세트를 Amazon Forecast로 가져오려면 이 값이 필요합니다.

The first screenshot shows the 'Create role' process in the AWS IAM console. The role name 'ForecastRole' is entered, and the role description is 'Allows Amazon Forecast to access data in S3.' The role is created successfully.

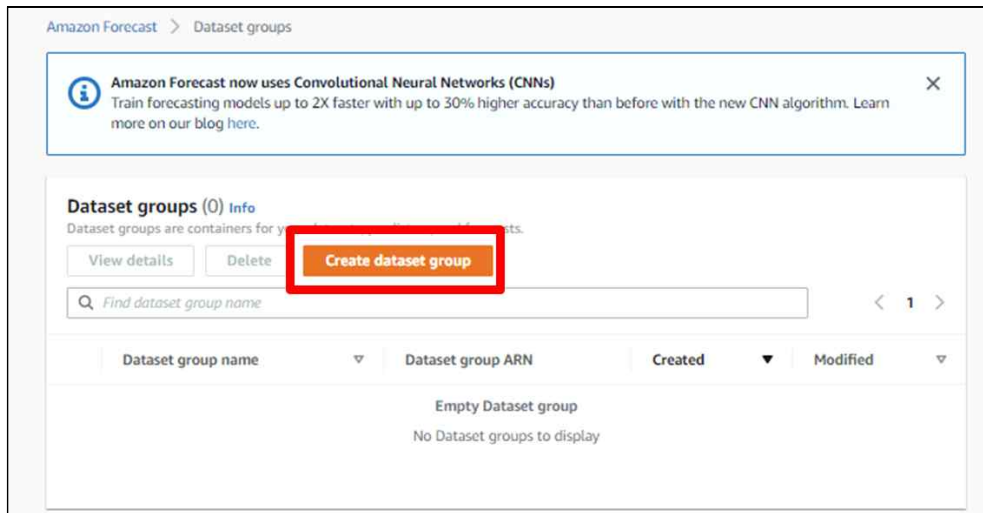
The second screenshot shows the 'Summary' page for the 'ForecastRole'. The role ARN is highlighted: `arn:aws:iam::402345010737:role/ForecastRole`. The role description is 'Allows Amazon Forecast to access data in S3.' The role was created on 2020-08-15 21:17 UTC+0900.

Below the summary, the 'Permissions policies' section shows that the role has the 'AmazonS3FullAccess' policy attached. The 'Permissions boundary' is not set.

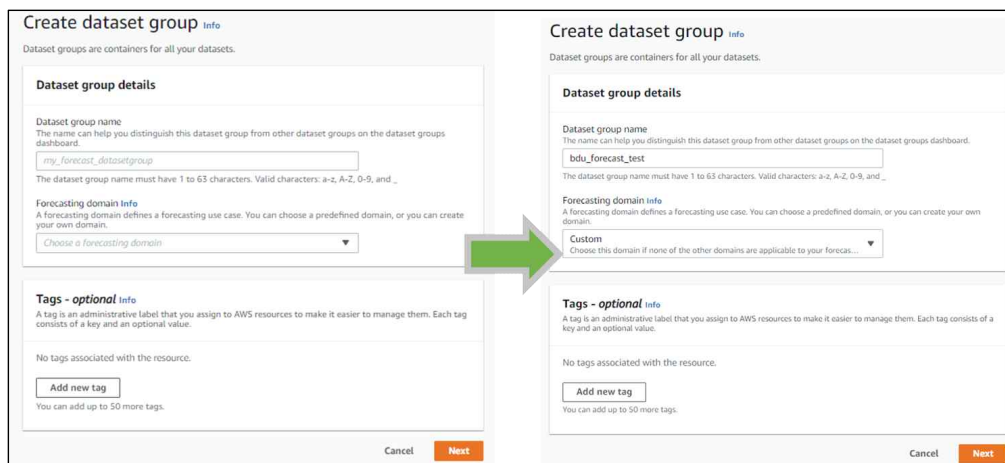
- 훈련 데이터 가져오기
 - ✓ Forecast 콘솔 접속 후 데이터 세트 생성 클릭

The screenshot shows the Amazon Forecast console. The 'Dataset groups' section is highlighted. The 'Create new dataset group' button is visible, which is used to create a new dataset group for training and forecasting.

- ✓ Create dataset group 클릭



- ✓ Dataset group name = 데이터 세트 그룹의 이름을 입력
- ✓ Forecasting domain(예상 도메인) = 드롭다운 메뉴에서 사용자 지정을 선택 입력 및 선택 후 [next] 클릭



- ✓ 데이터 세트 이름 및 데이터 빈도 선택 후 next 클릭
- ✓ 데이터 스키마 순서 바꾸기

	A	B	C	D	
1	2014-01-01 01:00	2.538071	client_0		
2	2014-01-01 01:00	23.64865	client_1		
3	2014-01-01 01:00	0	client_2		
4	2014-01-01 01:00	144.8171	client_3		

Dataset details

Dataset name
The name can help you distinguish this dataset from other datasets on your Datasets dashboard.

The dataset name must have 1 to 63 characters. Valid characters: a-z, A-Z, 0-9, and _

Frequency of your data
This is the frequency at which entries are registered into your data file.

Your data entries have a time interval of

Data schema [Info](#)
To help Amazon Forecast understand the fields in your data, you must define the schema. Specify the attributes in the same order as they appear in your CSV file. The pre-populated attributes below are required for the chosen domain.

```

1 {
2   "Attributes": [
3     {
4       "AttributeName": "timestamp",
5       "AttributeType": "timestamp"
6     },
7     {
8       "AttributeName": "target_value",
9       "AttributeType": "float"
10    },
11    {
12      "AttributeName": "item_id",
13      "AttributeType": "string"
14    }
15  ]
16 }

```

- ✓ Import target time series data(대상 시계열 데이터 가져오기) 페이지에서 Dataset import job details(데이터 세트 가져오기 작업 세부 정보)에 다음 정보를 입력 후 Start import(가져오기시작) 선택

Import target time series data [Info](#)

Dataset import details

Dataset import name
The name can help you distinguish this dataset import from other imports on your dataset detail page.

The dataset import name must have 1 to 63 characters. Valid characters: a-z, A-Z, 0-9, and _

Timestamp format [Info](#)
This is the format of the timestamp in your dataset. The format that you enter here must match the format in your data file.

IAM Role [Info](#)
Dataset groups require permissions from IAM to read your dataset files in S3. Choose or create a role using this control.

Custom IAM role ARN

Data location [Info](#)
The location is the path to the file in your S3 bucket that contains your data.

Your files must be in CSV format.

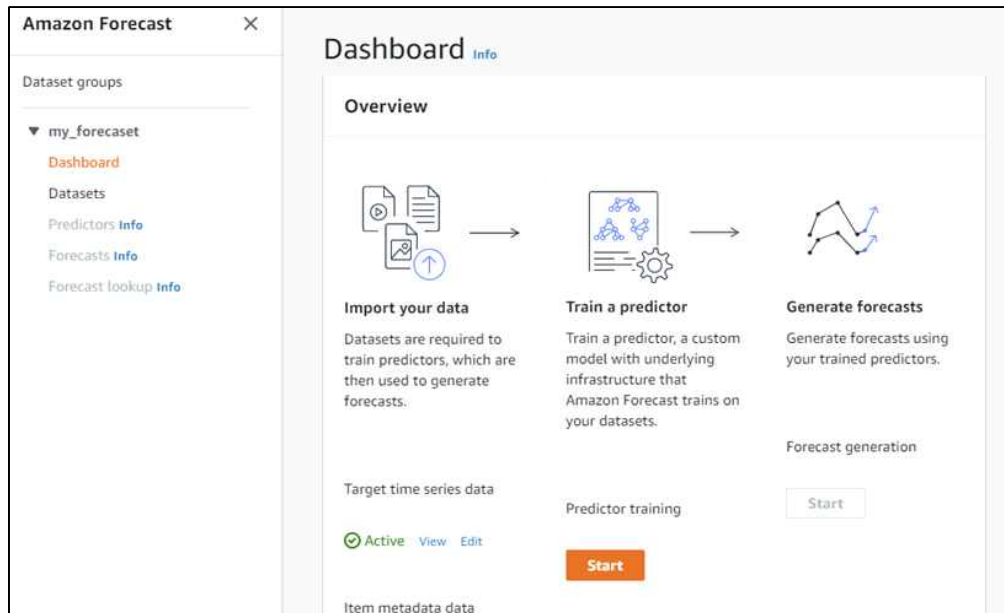
Tags - optional [Info](#)

A tag is an administrative label that you assign to AWS resources to make it easier to manage them. Each tag consists of a key and an optional value.

No tags associated with the resource.

You can add up to 50 more tags.

- 예측기 학습
 - ✓ 대상 시계열 데이터 세트를 가져오기가 완료되면 데이터 세트 그룹의 대시보드에 다음과 비슷한 화면이 나타남
 - ✓ Train a predictor(예측기 교육)에서 시작을 선택



- ✓ 예측기 교육 페이지에서 세부정보 입력

Train predictor Info

Amazon Forecast uses an algorithm to train a predictor on the data in your dataset group.

Predictor details

Predictor name
The name can help you distinguish this predictor from your other predictors.

 The predictor name must have 1 to 63 characters. Valid characters: a-z, A-Z, 0-9, and _

Forecast horizon Info
This number tells Amazon Forecast how far into the future to predict your data at the specified forecast frequency.

Forecast frequency
This is the frequency at which your forecasts are generated.
 Your forecast frequency is

Algorithm selection Info
An algorithm is used to train your predictor.
☐ Automatic (AutoML)
 Let Amazon Forecast choose the right algorithm for your dataset.
☒ Manual
 Explore the algorithms and choose one.

Algorithm
The algorithm that you want Amazon Forecast to use to train your predictor.

arn:aws:forecast::algorithm/ETS

✓ 예측기 교육 페이지에서 세부정보 입력

Forecast dimensions - optional
Item id is used in training by default. Select additional keys you would like to use to generate a forecast. These keys are fields in your dataset.

Choose a forecast dimension ▼

Country for holidays - optional
The holiday calendar you want to include for model training

Choose a country ▼ Reset

Number of backtest windows - optional Info
This is the number of times that the algorithm splits the input data for use in training and evaluation.

1

Backtest window offset - optional Info
This is the point in the dataset where you want to split the data for model training and evaluation.

36

► **Advanced configurations Info**
Set advanced configurations for your predictor and forecasts.

Tags - optional Info
A tag is an administrative label that you assign to AWS resources to make it easier to manage them. Each tag consists of a key and an optional value.

No tags associated with the resource.

Add new tag


You can add up to 50 more tags.

Cancel Train predictor

- ✓ Predictor training(예측기 교육) 아래에 교육 상태가 표시됨
 - Amazon Forecast에서 예측기 교육을 완료할 때까지 대기

Dashboard Info

Overview




Import your data
Datasets are required to train predictors, which are then used to generate forecasts.

Target time series data

✓ Active View Edit

Item metadata data


Import



Train a predictor
Train a predictor, a custom model with underlying infrastructure that Amazon Forecast trains on your datasets.

Predictor training

⏸ Create in progress... View

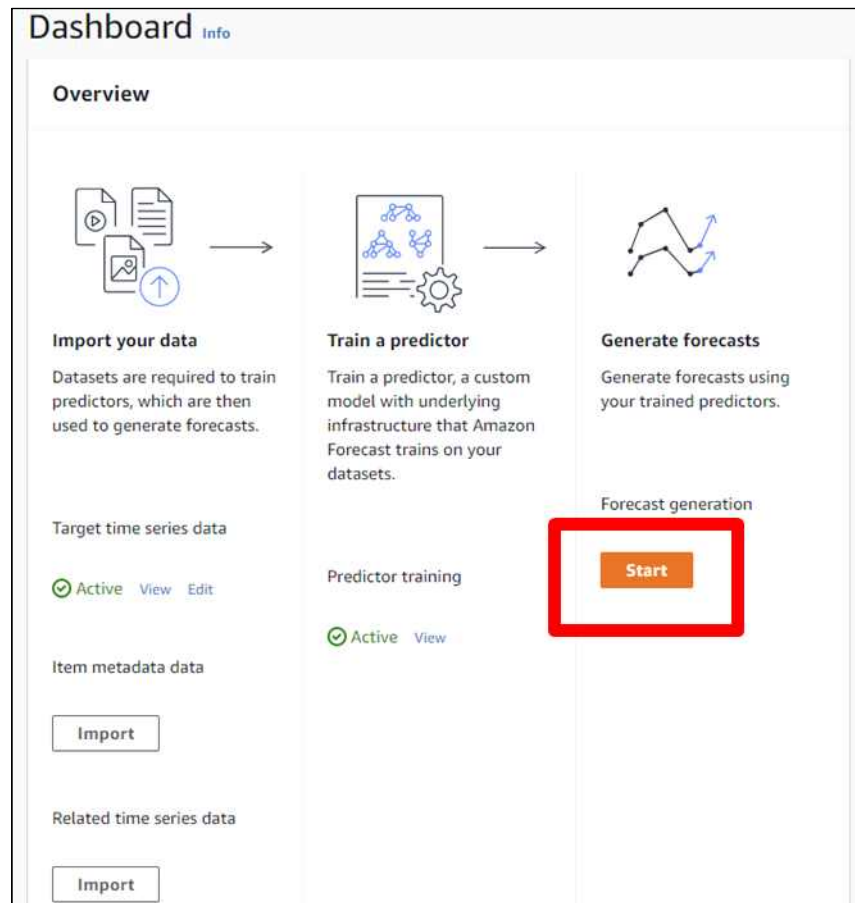


Generate forecasts
Generate forecasts using your trained predictors.

Forecast generation

Start

- 예측 생성 및 시각화
 - ✓ 예측기 교육이 완료되면 데이터 세트 그룹의 대시보드에서 예상생성 클릭



- ✓ Create a forecast(예상 생성) 페이지의 Forecast details(예상 세부 정보)에 다음 정보를 입력
 - 예상 이름 : 예상의 이름을 입력
 - Predictor(예측기) : 드롭다운 메뉴에서 Step 2: Train a Predictor에서 생성한 예측기를 선택
 - 나머지 설정은 선택 항목이므로 기본값을 유지

Create a forecast Info

Use a predictor to create forecasts based on your datasets.

Forecast details

Forecast name
The name can help you distinguish this forecast from your other forecasts.

The forecast name must have 1 to 63 characters. Valid characters: a-z, A-Z, 0-9, and _.

Predictor Info
The predictor that you want to use to create forecasts.

Forecast types - optional Info
Enter up to 5 quantile values between .01 to .99 including 'mean'. By default, Amazon Forecast will generate forecasts for .10, .50 and .90 quantiles.

Separate forecast types with commas.

Tags - optional Info

A tag is an administrative label that you assign to AWS resources to make it easier to manage them. Each tag consists of a key and an optional value.


No tags associated with the resource.

You can add up to 50 more tags.

✓ 데이터 세트 그룹의 대시보드 페이지가 표시

Dashboard Info

Overview




Import your data
Datasets are required to train predictors, which are then used to generate forecasts.

Target time series data

✓ Active [View](#) [Edit](#)


Item metadata data



Train a predictor
Train a predictor, a custom model with underlying infrastructure that Amazon Forecast trains on your datasets.

Predictor training

✓ Active [View](#)



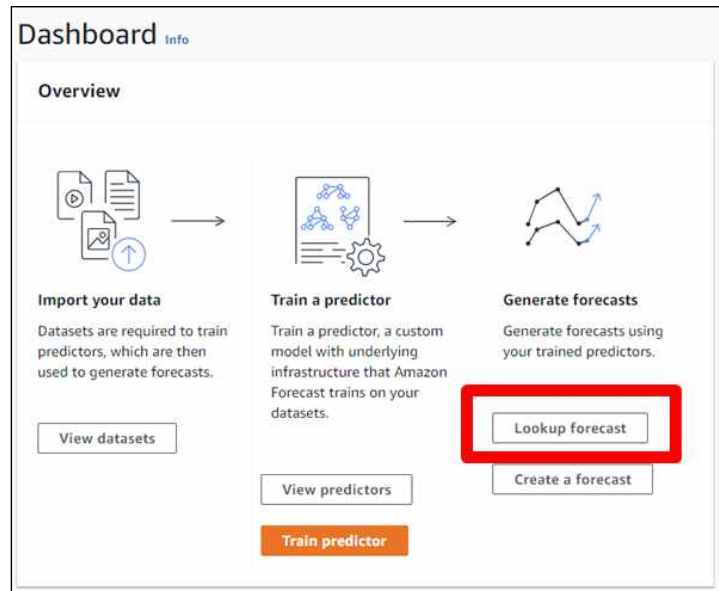
Generate forecasts
Generate forecasts using your trained predictors.

Forecast generation

⌚ Create pending

✓ 단일 항목을 쿼리하려면

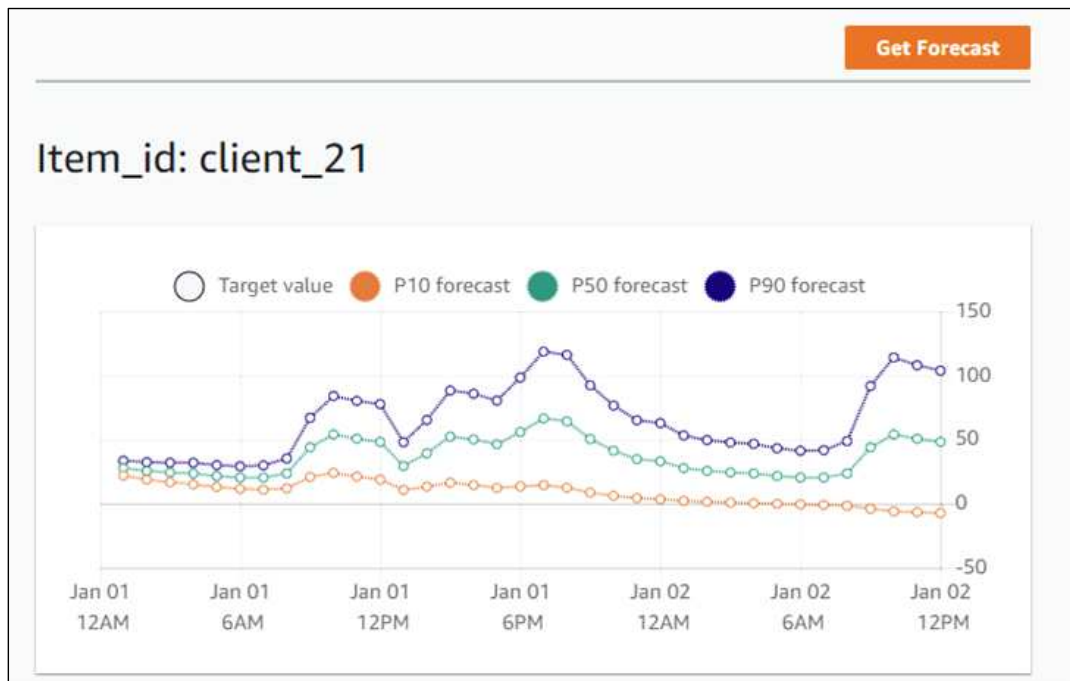
- ① 데이터 세트 그룹의 탐색창에 대시보드가 표시되지 않는다면 대시보드를 선택
- ② 대시보드에서 Generate forecasts(예상 생성) 아래의 Lookup forecast(예상 조회)를 선택합니다. Forecast lookup(예상 조회) 페이지가 표시



- ③ Forecast lookup(예상 조회) 페이지의 Forecast details(예상 세부 정보)에 다음 정보를 입력합니다.

The screenshot shows the 'Forecast lookup' page. The 'Forecast details' section is visible. It includes a 'Forecast' dropdown menu with 'bdu_forecaset' selected. There are 'Start date info' and 'End date info' sections, each with a date and time input field. The 'Forecast key' section has a dropdown menu with 'item_id' selected and a 'Value' input field with 'client_21'. There are buttons for 'Remove forecast key' and 'Add forecast key'. At the bottom right, there is a 'Get Forecast' button.

- ✓ Get Forecast(예상 가져오기)를 선택합니다. 예상이 표시되면 client_21에 의한 전기 사용량 수요 예상을 검토



평가하기

1. Amazon Forecast는 이미지 데이터를 기반으로 매우 정확한 예측을 위해 머신러닝을 사용하는 완전관리형 서비스이다.(O/X)

- 정답 : X

해설 : Amazon Forecast는 시계열 데이터를 기반으로 매우 정확한 예측을 위해 머신러닝을 사용하는 완전관리형 서비스입니다.

2. Forecast에 미리 정의된 데이터 세트도메인 중 소매 수요 예측 시 사용해야 할도메인은?

- ① RETAIL 도메인
- ② EC2 CAPACITY 도메인
- ③ METRICS 도메인
- ④ WORK_FORECE 도메인

- 정답 : ① 번

해설 : Amazon Forecast에서 소매 수요 예측 시사용해야 할 도메인은 RETAIL 도메인입니다.

학습정리

1. Amazon Forecast 개요

- 시계열 데이터를 기반으로 매우 정확한 예측을 위해 머신러닝을 사용하는 완전관리형 서비스

2. Amazon Forecast 주요 기능

- 자동화된 머신러닝
- 예측 시각화
- 확률론적 예측 생성

3. Amazon Forecast 알고리즘과 도메인

- 데이터 세트 도메인
 - RETAIL도메인
 - INVENTORY_PLANNING 도메인
 - EC2 CAPACITY 도메인
 - WORK_FORCE 도메인
 - WEB_TRAFFIC 도메인
 - METRICS 도메인
 - CUSTOM 도메인
- 알고리즘
 - Autoregressive Integrated Moving Average(ARIMA)
 - CNN-QR
 - DeepAR+
 - Exponential Smoothing(ETS)
 - Non-Parametric Time Series(NPTS)
 - Prophet

3. 실습. Forecast를 사용하여 전기 사용량 예측하기

- Forecast 사전 준비
- 교육 데이터 가져오기
- 예측기 교육
- 예측 생성
- 예상 검색하기