



Amazon SageMaker ▾

Overview

Features ▾

Pricing

FAQs

IT Ops

Data Science

Developer Resources

Getting Started

Customers

Partners

Amazon SageMaker Autopilot

Automatically create machine learning models with full visibility

[Get Started with Amazon SageMaker Autopilot](#)

FEATURED EVENT

SageMaker Fridays Are Back

Join us each week to learn how to use machine learning to solve major challenges specific to your industry.

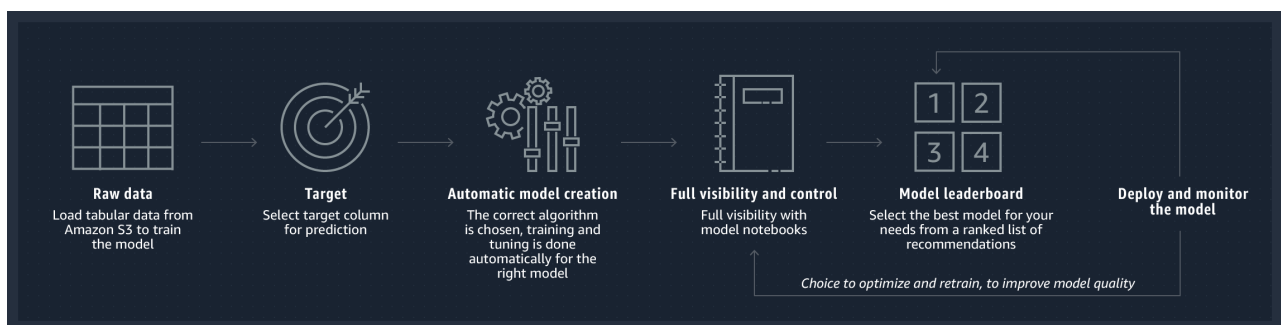
[Learn more »](#)

Amazon SageMaker Autopilot automatically builds, trains, and tunes the best machine learning models based on your data, while allowing you to maintain full control and visibility.

Building machine learning (ML) models requires you to manually prepare features, test multiple algorithms, and optimize hundreds of model parameters in order to find the best model for your data. However, this approach requires deep ML expertise. If you don't have that expertise, you could use an automated approach (AutoML), but AutoML approaches typically provide very little visibility into the impact of your features for model predictions. As a result, you may have less trust in it because you can't recreate it and you can't learn how it makes predictions.

Amazon SageMaker Autopilot eliminates the heavy lifting of building ML models, and helps you automatically build, train, and tune the best ML model based on your data. With SageMaker Autopilot, you simply provide a tabular dataset and select the target column to predict, which can be a number (such as a house price, called regression), or a category (such as spam/not spam, called classification). SageMaker Autopilot will automatically explore different solutions to find the best model. You then can directly deploy the model to production with just one click, or iterate on the recommended solutions with Amazon SageMaker Studio to further improve the model quality.

How it works



Key features

Automatic data pre-processing and feature engineering

You can use Amazon SageMaker Autopilot even when you have missing data. SageMaker Autopilot automatically fills in the missing data, provides statistical insights about columns in your dataset, and automatically extracts information from non-numeric columns, such as date and time information from timestamps.

Amazon SageMaker Autopilot Data Exploration

This report provides insights about the dataset you provided as input to the AutoML job. It was automatically generated by the AutoML training job: **airline-sampled-mc-2**.

As part of the AutoML job, the input dataset was randomly split into two pieces, one for **training** and one for **validation**. The training dataset was randomly sampled, and metrics were computed for each of the columns. This notebook provides these metrics so that you can:

1. Understand how the job analyzed features to select the candidate pipelines.
2. Modify and improve the generated AutoML pipelines using knowledge that you have about the dataset.

We read **5846** rows from the training dataset. The dataset has **25** columns and the column named **DelayType** is used as the target column. This is identified as a **MulticlassClassification** problem. Here are **5** examples of labels: `['SecurityDelay', 'WeatherDelay', 'NASDelay', 'CarrierDelay', 'LateAircraftDelay']`.

Suggested Action Items

- Look for sections like this for recommended actions that you can take.

Contents

- 1. Dataset Sample
- 2. Column Analysis


Dataset Sample

The following table is a random sample of **10** rows from the training dataset. For ease of presentation, we are only showing **20 of the 25** columns of the dataset.

Suggested Action Items

- Verify the input headers correctly align with the columns of the dataset sample. If they are incorrect, update the header names of your input dataset in Amazon Simple Storage Service (Amazon S3).

	Year	Month	DayofMonth	DayOfWeek	DepTime	CRSDepTime	ArrTime	CRSArrTime	UniqueCarrier	FlightNum	...	DepDelay	Origin	Dest	Distance	TaxiIn	TaxiOut	Cancelled	CancellationCode	Diverted	DelayType
0	2008	2	19	2	1219.0	1203	1446.0	1421	FL	1695	...	16.0	EWB	ATL	745	5.0	21.0	0		0	CarrierDelay
1	2008	2	19	2	1545.0	1430	1650.0	1534	YV	2826	...	75.0	ONT	LAS	197	11.0	12.0	0		0	LateAircraftDelay
2	2008	1	11	5	948.0	915	1026.0	948	FL	188	...	33.0	ATL	STL	483	4.0	14.0	0		0	CarrierDelay
3	2008	3	27	4	1943.0	1720	2211.0	1930	MQ	4007	...	143.0	ORD	CMH	296	8.0	29.0	0		0	CarrierDelay
4	2008	2	23	6	1742.0	1729	2012.0	1957	FL	625	...	13.0	PHL	MCO	861	8.0	13.0	0		0	CarrierDelay
5	2008	2	18	1	1201.0	1130	1313.0	1234	DL	462	...	31.0	SAV	ATL	215	9.0	21.0	0		0	CarrierDelay
6	2008	1	2	3	1027.0	945	1132.0	1055	MQ	3403	...	42.0	DFW	HOU	247	3.0	15.0	0		0	CarrierDelay
7	2008	1	13	7	1527.0	1510	1839.0	1820	MQ	4153	...	17.0	ORD	PVD	849	3.0	19.0	0		0	LateAircraftDelay
8	2008	2	23	6	1812.0	1640	1855.0	1721	OH	5146	...	92.0	ABE	LGA	82	5.0	8.0	0		0	NASDelay
9	2008	1	17	4	644.0	640	751.0	734	UA	1235	...	4.0	DSM	DEN	589	7.0	34.0	0		0	NASDelay

 [Click to enlarge](#)

Automatic ML model selection

Amazon SageMaker Autopilot automatically infers the type of predictions that best suit your data, such as binary classification, multi-class classification, or regression. SageMaker Autopilot then explores high-performing algorithms such as gradient boosting decision tree, feedforward deep neural networks, and logistic regression, and trains and optimizes hundreds of models based on these algorithms to find the model that best fits your data.

<https://aws.amazon.com/sagemaker/autopilot/>

3/12

TRIAL COMPONENTS

1 row selected0/20 filters

Q Search column name to start

Add chartDeploy model

Trial name	ObjectiveMetric	validation:accuracy	validation:merror	validation:macro_recall	validation:objective_f1...	validation:mult
tuning-job-1-afb5da873b2e4c16a4-249-c971f83e-aws-trial	0.70051	0.70051	0.29949			
tuning-job-1-afb5da873b2e4c16a4-248-8b75985a-aws-trial	0.66297	0.66297	0.33703			
tuning-job-1-afb5da873b2e4c16a4-251-1f5e410c-aws-trial	0.20051194727420807			0.15630197525024414	2.1411008834838867	0.2005119472
tuning-job-1-afb5da873b2e4c16a4-247-dfee99f-aws-trial	0.69625	0.69625	0.30375			
tuning-job-1-afb5da873b2e4c16a4-245-8a9e4ff4-aws-trial	0.69454	0.69454	0.30546			
tuning-job-1-afb5da873b2e4c16a4-250-5393ab3b-aws-trial	0.64505	0.64505	0.35495			
tuning-job-1-afb5da873b2e4c16a4-246-d9a06691-aws-trial	0.67236	0.67236	0.32765			
tuning-job-1-afb5da873b2e4c16a4-243-1944b28b-aws-trial	0.69198	0.69198	0.30802			
tuning-job-1-afb5da873b2e4c16a4-244-99ee4a6a-aws-trial	0.69198	0.69198	0.30802			
tuning-job-1-afb5da873b2e4c16a4-242-20856684-aws-trial	0.67065	0.67065	0.32935			
tuning-job-1-afb5da873b2e4c16a4-235-6beceaae-aws-trial	0.70734	0.70734	0.29266			
tuning-job-1-afb5da873b2e4c16a4-241-684c9a2b-aws-trial	0.69795	0.69795	0.30205			
tuning-job-1-afb5da873b2e4c16a4-240-a1323164-aws-trial	0.6220099925994873	0.6220099925994873	0.3779900074005127			
tuning-job-1-afb5da873b2e4c16a4-239-24678556-aws-trial	0.69966	0.69966	0.30034			
tuning-job-1-afb5da873b2e4c16a4-238-9dd93711-aws-trial	0.67918	0.67918	0.32082			
tuning-job-1-afb5da873b2e4c16a4-237-4ee36a8a-aws-trial	0.6851500272750854	0.6851500272750854	0.31485000252723694			
tuning-job-1-afb5da873b2e4c16a4-236-f118ec26-aws-trial	0.69625	0.69625	0.30375			
tuning-job-1-afb5da873b2e4c16a4-234-b0000dcd-aws-trial	0.7133100032806396	0.7133100032806396	0.28668999671936035			
tuning-job-1-afb5da873b2e4c16a4-233-451c99bd-aws-trial	0.7090399861335754	0.7090399861335754	0.29096001386642456			
tuning-job-1-afb5da873b2e4c16a4-232-16f3d8e7-aws-trial	0.70563	0.70563	0.29437			
tuning-job-1-afb5da873b2e4c16a4-230-8513450a-aws-trial	0.7167199850082397	0.7167199850082397	0.28327998518943787			
tuning-job-1-afb5da873b2e4c16a4-231-512d9de5-aws-trial	0.69113	0.69113	0.30887			
tuning-job-1-afb5da873b2e4c16a4-229-38a1bd73-aws-trial	0.69113	0.69113	0.30887			
tuning-job-1-afb5da873b2e4c16a4-227-b630bbe3-aws-trial	0.7099000215530396	0.7099000215530396	0.2901000824928284			
tuning-job-1-afb5da873b2e4c16a4-226-ae7b9917-aws-trial	0.69198	0.69198	0.30802			
tuning-job-1-afb5da873b2e4c16a4-228-6532284c-aws-trial	0.6006799936294556	0.6006799936294556	0.39932000637054443			
tuning-job-1-afb5da873b2e4c16a4-224-d76ecc6a-aws-trial	0.69539	0.69539	0.30461			
tuning-job-1-afb5da873b2e4c16a4-225-70971c62-aws-trial	0.7107499837875366	0.7107499837875366	0.289249986410141			

Click to enlarge

Model leaderboard

Amazon SageMaker Autopilot allows you to review all the ML models that are automatically generated for your data. You can view the list of models, ranked by metrics such as accuracy, precision, recall, and area under the curve (AUC), review model details such as the impact of features on predictions, and deploy the model that is best suited to your use case.

EXPERIMENT: AIRLINE-SAMPLED-MC-2

Open candidate generation notebook ⓘ

Open data exploration notebook ⓘ

Problem typeMulticlassClassification

Trials

Job profile

TRIALS0 rows selected

Deploy model

Trial name	Status	Start time	Objective: Accuracy
tuning-job-1-afb5da873b2e4c16a4-181-26bee647	Completed	1 month ago	0.722699998092651
★ Best: tuning-job-1-afb5da873b2e4c16a4-163-7...	Completed	1 month ago	0.722699998092651
tuning-job-1-afb5da873b2e4c16a4-155-7a3d564e	Completed	1 month ago	0.7218400239944458
tuning-job-1-afb5da873b2e4c16a4-151-a74a1321	Completed	1 month ago	0.7218400239944458
tuning-job-1-afb5da873b2e4c16a4-087-2d8bcb7e	Completed	1 month ago	0.720990002155304
tuning-job-1-afb5da873b2e4c16a4-143-90c8282f	Completed	1 month ago	0.720990002155304
tuning-job-1-afb5da873b2e4c16a4-110-99577f5e	Completed	1 month ago	0.720990002155304
tuning-job-1-afb5da873b2e4c16a4-109-c776de67	Completed	1 month ago	0.720990002155304
tuning-job-1-afb5da873b2e4c16a4-166-04265c01	Completed	1 month ago	0.7201399803161621
tuning-job-1-afb5da873b2e4c16a4-249-c971f83e	Completed	1 month ago	0.7192800045013428
tuning-job-1-afb5da873b2e4c16a4-203-3e177457	Completed	1 month ago	0.7184299826622009
tuning-job-1-afb5da873b2e4c16a4-245-8a9e4ff4	Completed	1 month ago	0.7184299826622009
tuning-job-1-afb5da873b2e4c16a4-211-3bda6dc4	Completed	1 month ago	0.7175800204277039
tuning-job-1-afb5da873b2e4c16a4-132-a34915b9	Completed	1 month ago	0.7175800204277039
tuning-job-1-afb5da873b2e4c16a4-178-2fbbec23	Completed	1 month ago	0.7175800204277039
tuning-job-1-afb5da873b2e4c16a4-113-ad2f5aa6	Completed	1 month ago	0.7175800204277039
tuning-job-1-afb5da873b2e4c16a4-130-07225d3c	Completed	1 month ago	0.7175800204277039
tuning-job-1-afb5da873b2e4c16a4-235-6beceaae	Completed	1 month ago	0.7175800204277039
tuning-job-1-afb5da873b2e4c16a4-230-8513450a	Completed	1 month ago	0.7167199850082397
tuning-job-1-afb5da873b2e4c16a4-212-55b029a4	Completed	1 month ago	0.7167199850082397
tuning-job-1-afb5da873b2e4c16a4-194-f56fcc18	Completed	1 month ago	0.7167199850082397
tuning-job-1-afb5da873b2e4c16a4-192-11f7443b	Completed	1 month ago	0.7167199850082397
tuning-job-1-afb5da873b2e4c16a4-165-86bf295d	Completed	1 month ago	0.7167199850082397
tuning-job-1-afb5da873b2e4c16a4-244-99ee4a6a	Completed	1 month ago	0.7167199850082397
tuning-job-1-afb5da873b2e4c16a4-221-09826f70	Completed	1 month ago	0.7167199850082397
tuning-job-1-afb5da873b2e4c16a4-213-a39ef205	Completed	1 month ago	0.7167199850082397

 Click to enlarge

Automatic notebook creation

You can automatically generate a Amazon SageMaker Studio Notebook for any model Amazon SageMaker Autopilot creates and dive into the details of how it was created, refine it as desired, and recreate it from the notebook at any point in the future.

Amazon SageMaker Autopilot Candidate Definition Notebook

This notebook was automatically generated by the AutoML job **airline-sampled-mc-2**. This notebook allows you to customize the candidate definitions and execute the SageMaker Autopilot workflow.

The dataset has **25** columns and the column named **DelayType** is used as the target column. This is being treated as a **MulticlassClassification** problem. The dataset also has **5** classes. This notebook will build a **MulticlassClassification** model that maximizes the **"ACCURACY"** quality metric of the trained models. The **"ACCURACY"** metric provides the percentage of times the model predicted the correct class.

As part of the AutoML job, the input dataset has been randomly split into two pieces, one for **training** and one for **validation**. This notebook helps you inspect and modify the data transformation approaches proposed by Amazon SageMaker Autopilot. You can interactively train the data transformation models and use them to transform the data. Finally, you can execute a multiple algorithm hyperparameter optimization (multi-algo HPO) job that helps you find the best model for your dataset by jointly optimizing the data transformations and machine learning algorithms.

 **Available Knobs** Look for sections like this for recommended settings that you can change.

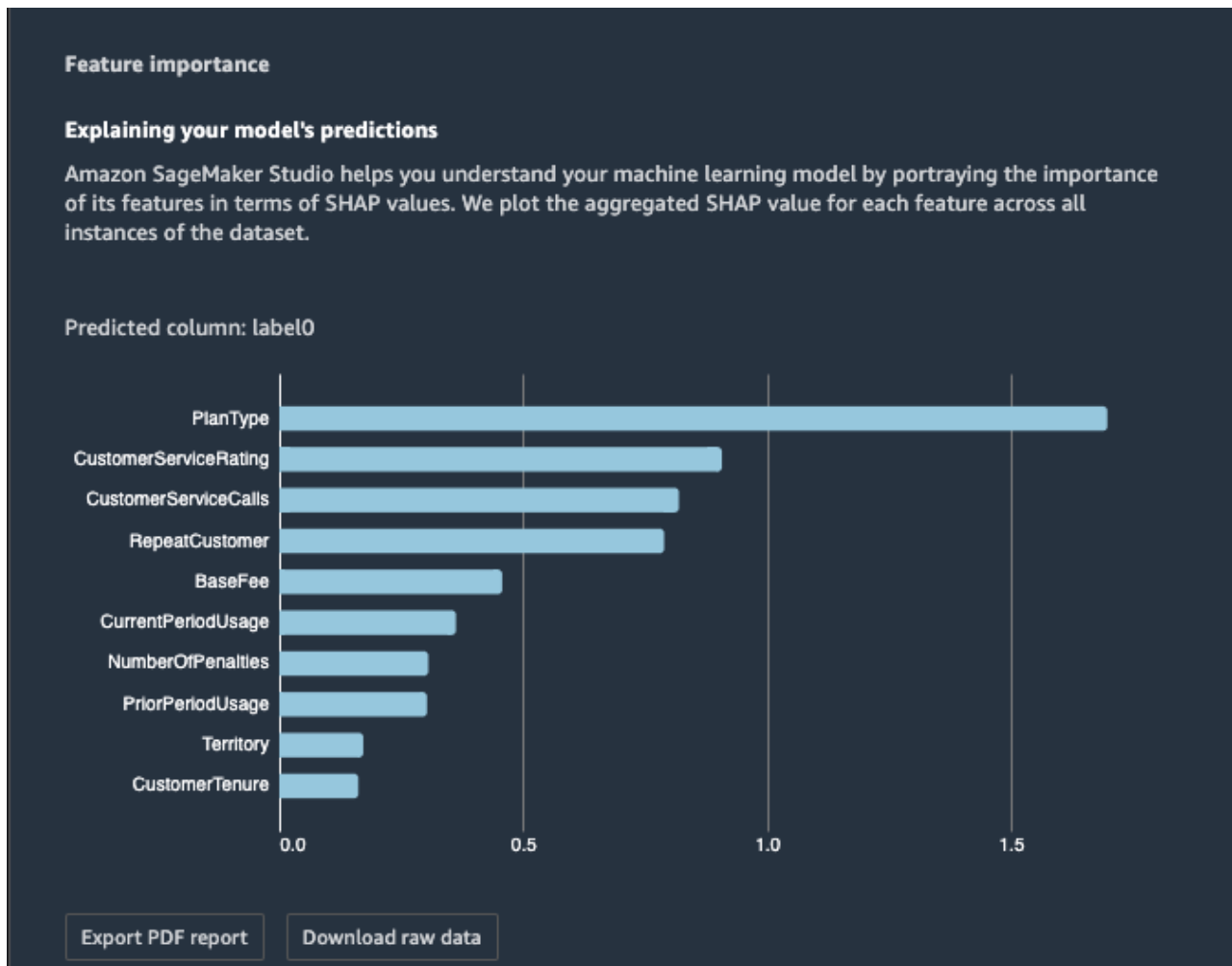
Contents

1. Sagemaker Setup
 - A. Downloading Generated Candidates
 - B. SageMaker Autopilot Job and Amazon Simple Storage Service (Amazon S3) Configuration
2. Candidate Pipelines
 - A. Generated Candidates
 - B. Selected Candidates
3. Executing the Candidate Pipelines
 - A. Run Data Transformation Steps
 - B. Multi Algorithm Hyperparameter Tuning
4. Model Selection and Deployment
 - A. Tuning Job Result Overview
 - B. Model Deployment

 Click to enlarge

Feature importance

Amazon SageMaker Autopilot provides an explainability report, generated by Amazon SageMaker Clarify, that makes it easier for you to understand and explain how models created with SageMaker Autopilot make predictions. You can also identify how each attribute in your training data contributes to the predicted result as a percentage. The higher the percentage, the more strongly that feature impacts your model's predictions.



[Click to enlarge](#)

Easy integration with your applications

You can use the Amazon SageMaker Autopilot application programming interface (API) to easily create models and make inferences right from your applications, such as your data analytics and data warehousing tools.

Use Cases

Price predictions

Price prediction models are used heavily in financial services, real estate, and energy and utilities to predict the price of stocks, real estate, and natural resources. Amazon SageMaker Autopilot can predict future prices to help you make sound investment decisions based on your historical data such as demand, seasonal trends, and price of other commodities.

Churn prediction

Customer churn is the loss of customers or clients, and every company looks for ways to eliminate it. Models automatically generated by Amazon SageMaker Autopilot help you understand churn patterns. Churn prediction models work by first learning patterns in your existing data and identifying patterns in new datasets so you can get a prediction about customers mostly likely to churn.

Risk assessment

Risk assessment requires identifying and analyzing potential events that may negatively impact individuals, assets, and your company. Models automatically generated by Amazon SageMaker Autopilot predict risks as new events unfold. Risk assessment models are trained using your existing datasets so you can get optimized predictions for your business.

Customers



"Sisense's new ML service powered by Amazon SageMaker Autopilot was exactly what we needed to keep ahead of the curve in customer service during this COVID-19 pandemic. Skullcandy was able to gain deep insights into our customers' needs, improve our issue resolution, and increase customer satisfaction scores."

Mark Hopkins, Chief Information Officer, Skullcandy Inc.



"Previously, we would simply pick two restaurants that looked similar, but now we have a true understanding of the relationships between our menu items, customers, and locations. Amazon SageMaker Autopilot, which powers Domo's new ML capability, has been a force multiplier for our marketing and purchasing teams to try new ideas and improve our customers' experience."

Sean Thompson, IT Director, Freddy's



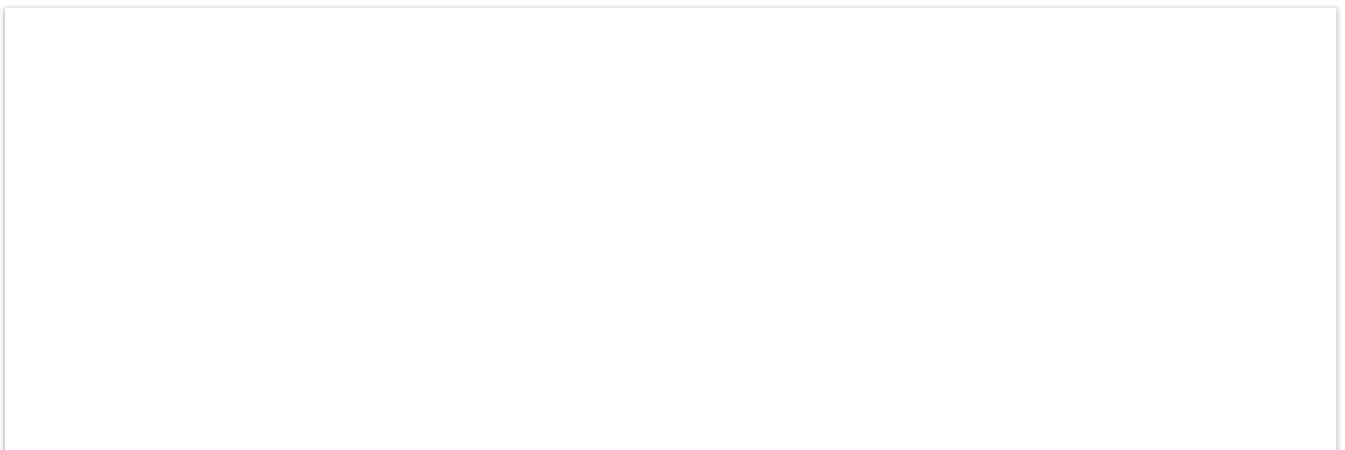
"The primary goal in demographic mapping is optimizing across both accuracy and scale. While this is generally difficult, we were able to use Amazon SageMaker Autopilot with our comprehensive training data and sophisticated features to produce better models that improved our prediction accuracy by 137%."

Anindya Datta, CEO, Mobilewalla

"At RetentionX, we provide one-click business insights to e-commerce companies. To serve our customers, it is important that they can get started quickly and make timely business decisions, however building accurate machine learning models can be costly and take months of trial-and-error. In addition, model accuracy is also highly dependent on the breadth and depth of training data and unique feature set available for each of our customers. With the help of Amazon SageMaker Autopilot, our customers can automatically generate the best ML models based on unique datasets. Thanks to SageMaker Autopilot, we can provide personalized insights to tens of millions of shoppers tapping into the power of AutoML."

Alexander Jost, CEO, RetentionX

Resources for Amazon SageMaker Autopilot



Build high-quality machine learning models easily and quickly with Amazon SageMaker Autopilot (31:39)

Amazon SageMaker Autopilot

Get Started

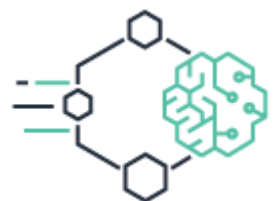
AMAZON SAGEMAKER PARTNERS AVAILABLE

Accelerate building and deploying ML models and solutions



GET EXPERT HELP FOR AMAZON SAGEMAKER

Produce high quality ML models quickly with the help of AWS certified experts



AWS INNOVATE DATA EDITION | AUGUST 19 | FREE ONLINE EVENT

Drive agility, digital transformation, and continuous innovation with your data today.
Register now >>



[Sign In to the Console](#)

Learn About AWS

[What Is AWS?](#)

Resources for AWS

[Getting Started](#)

[Training and Certification](#)

[AWS Solutions Portfolio](#)

Developers on AWS

[Developer Center](#)

[SDKs & Tools](#)

[.NET on AWS](#)

[What Is Cloud Computing?](#)[Architecture Center](#)[Python on AWS](#)[AWS Inclusion, Diversity & Equity](#)[Product and Technical FAQs](#)[Java on AWS](#)[What Is DevOps?](#)[Analyst Reports](#)[PHP on AWS](#)[What Is a Container?](#)[AWS Partner Network](#)[Javascript on AWS](#)[What Is a Data Lake?](#)[AWS Cloud Security](#)[What's New](#)[Blogs](#)[Press Releases](#)

Help

[Contact Us](#)[AWS Careers](#)[File a Support Ticket](#)[Knowledge Center](#)[AWS Support Overview](#)[Legal](#)[Sign In to the Console](#)

Amazon is an Equal Opportunity Employer: *Minority / Women / Disability / Veteran / Gender Identity / Sexual Orientation / Age.*

Language

[عربي |](#)[Bahasa Indonesia |](#)[Deutsch |](#)[English |](#)[Español |](#)[Français |](#)[Italiano |](#)[Português |](#)[Tiếng Việt |](#)[Türkçe |](#)[Русский |](#)[ไทย |](#)[日本語 |](#)[한국어 |](#)[中文 \(简体\) |](#)

[中文 \(繁體\)](#)

[Privacy](#)

|

[Site Terms](#)

|

[Cookie Preferences](#)

|

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.