
aws
INNOVATE
AI/ML EDITION

aws INNOVATE
AI/ML EDITION

Create, train, and deploy machine learning (ML) models using familiar SQL commands

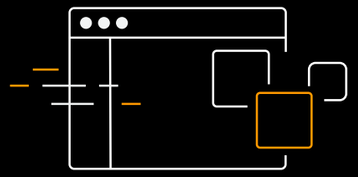
Suman Debnath
Principal Developer Advocate
AISPL

Challenges of data analytics at scale

Variety of sources and data types

Multiple analytics needs

Data volume and velocity



Slow performance

Difficult to manage systems

Complex to scale

Increasing and unpredictable cost

Inflexible tools

Security and compliance

Amazon Redshift

THE MOST WIDELY USED CLOUD DATA WAREHOUSE, WITH TENS OF THOUSANDS OF CUSTOMERS

ANALYZE ALL YOUR DATA

Take a **lake house approach** by analyzing all your data across your data warehouse, your Amazon S3 data lake, and operational databases with consistent security and governance policies

PERFORMANCE AT ANY SCALE

Get up to **3x better price performance** than other cloud data warehouses with a **self-tuning** system, and boost queries up to **10x with AQUA**

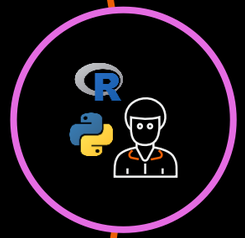
LOWER YOUR COSTS

Start small and pay only for what you use with **predictable** monthly costs; Amazon Redshift is at least **50% less expensive** than other cloud data warehouses

Personas that use AWS Analytics services

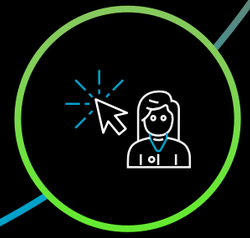
Data engineers/Database developers

Data analysts



Data scientists

BI professionals

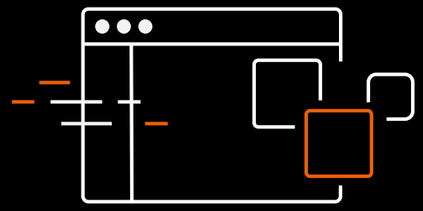


Administrators

Why Machine Learning ?

Transform customer experience

Improve business operations



Better and faster decision-making

Innovate product or service

ML in a data warehouse

Customer churn detection

Predict if a sales lead will close

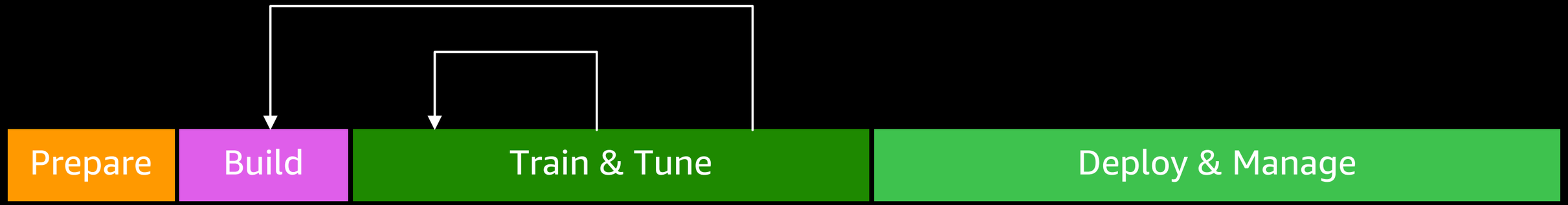
Price/revenue prediction

Product recommendation

Fraud detection

Customer lifetime value prediction

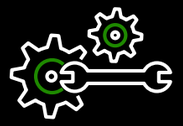
ML workflow can be complex and iterative



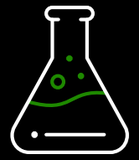
Collect and prepare training data

Choose or build a ML algorithm

Set up and manage environments for training



Train, debug, and tune models



Manage training runs

Deploy model in production

Monitor models

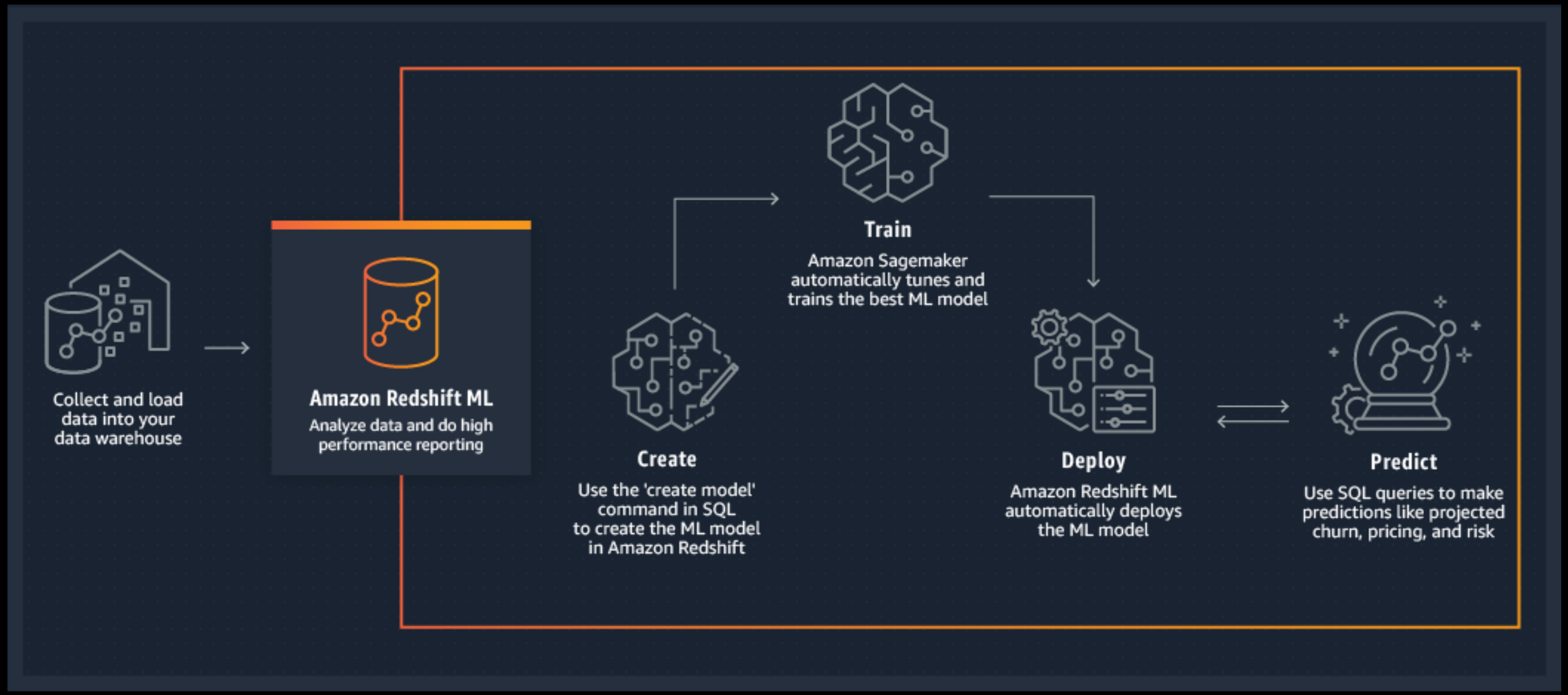
Validate predictions

Scale and manage the production environment

aws INNOVATE
AI/ML EDITION

Amazon Redshift ML (Preview)

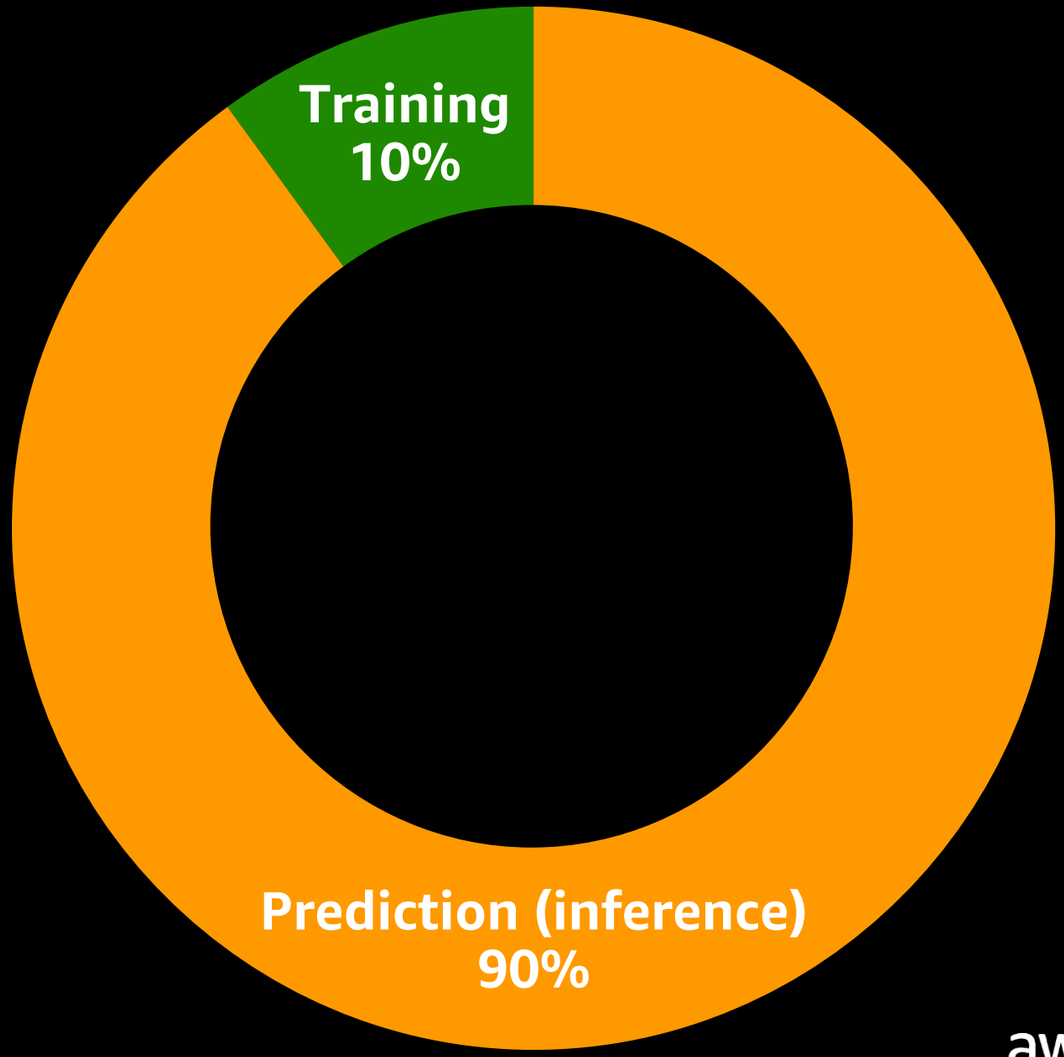
Amazon Redshift ML (preview)



Amazon Redshift ML: Optimized for cost

Typically predictions drive cost in production

You only pay for training while prediction comes at no extra cost when you use Amazon Redshift ML



How Amazon Redshift ML works

TRAIN

```
CREATE MODEL customer_churn
FROM (SELECT c.age, c.zip,
c.monthly_spend, c.monthly_cases,
c.active AS label
FROM customer_info_table c)
TARGET label
FUNCTION predict_customer_churn
```

Amazon Redshift

Runs Amazon SageMaker Autopilot and returns model

Amazon SageMaker

PREDICT

```
SELECT n.id, n.firstName, n.lastName,
predict_customer_churn(n.age,c.zip,..)
AS activity_prediction
FROM new_customers n
WHERE n.marital_status = ‘single’
```

Amazon Redshift

Create, train, and deploy model with a simple SQL command

Auto-selection of model, pre-processing, and training using SageMaker Autopilot

Trained model gets compiled by SageMaker Neo in Amazon Redshift data warehouse so that you can make predictions using SQL

Uses previously built model to predict in-place (inference executed entirely in Amazon Redshift)

Creating and training ML model

Specify training data as a table name or SELECT query

TARGET column specifies the column you are trying to predict

FUNCTION specifies the name of the prediction function that will be generated

```
CREATE MODEL demo_ml.customer_churn_auto_model
FROM (SELECT state,
             account_length,
             area_code,
             total_charge/account_length AS average_daily_spend,
             cust_serv_calls/account_length AS average_daily_cases,
             churn
      FROM demo_ml.customer_activity
      WHERE  record_date < '2020-01-01'
     )
TARGET churn
FUNCTION ml_fn_customer_churn_auto
;
```

Creating and training ML model

Optionally specify:

Model type; e.g., XGBOOST

Objective for training; e.g., mean squared error (MSE)

Preprocessors or hyperparameters

```
CREATE MODEL abalone_xgboost_multi_predict_age
FROM ( SELECT length_val,
              shucked_weight,
              viscera_weight,
              shell_weight,
              rings
         FROM abalone_xgb WHERE record_number < 2500 )
TARGET rings FUNCTION ml_fn_abalone_xgboost_multi_predict_age
MODEL_TYPE XGBOOST
OBJECTIVE 'multi:softmax'
PREPROCESSORS 'none'
;
```

Show ML model

Check status of model with SHOW MODEL command

SHOW MODEL ALL shows all your models

Provides status of the models

```
SHOW MODEL demo_ml.customer_churn_auto_model
```

	Key	Value
1	Model Name	customer_churn_auto_model
2	Schema Name	demo_ml
3	Owner	awsuser
4	Creation Time	Thu, 28.01.2021 05:20:33
5	Model State	READY
6	Training Job Status	MaxAutoMLJobRuntimeReached
7	validation:f1	0.662400
8		
9	TRAINING DATA:	
10	Query	SELECT STATE, ACCOUNT_LENGTH, AREA_COD…
11		FROM DEMO_ML.CUSTOMER_ACTIVITY
12		WHERE RECORD_DATE < '2020-01-01'
13	Target Column	CHURN
14		
15	PARAMETERS:	
16	Model Type	auto
17	Problem Type	BinaryClassification
18	Objective	F1
19	Function Name	ml_fn_customer_churn_auto
20	Function Parameters	state account_length area_code average…
21	Function Parameter Types	varchar int4 int4 float8 int4
22	Max Runtime	3600

Using ML model for prediction

The prediction (inference) function is available as a UDF

You can generate prediction from any SQL construct just as you use UDFs today

You can use WLM to prioritize your compute resources for inference function

Prediction function takes all benefits of Amazon Redshift, including the massively parallel processing capability

```
SELECT phone,
       record_date,
       state,
       DEMO_ML.ml_fn_customer_churn_auto(
              state,
              account_length,
              area_code,
              total_charge/account_length ,
              cust_serv_calls/account_length )
          AS active
FROM DEMO_ML.customer_activity
WHERE record_date > '2020-01-01';
```

aws INNOVATE
AI/ML EDITION

Demo

Visit the AI and Machine Learning Resource Hub for more resources

Dive deeper with these resources, get inspired and learn how you can use machine learning to accelerate business outcomes.

- The machine learning journey e-book
- Machine learning enterprise guide
- 7 leading machine learning use cases e-book
- A strategic playbook for data, analytics, and machine learning
- Accelerating ML innovation through security e-book
- … and more!

Visit resource hub »

https://tinyurl.com/aiml-aws

AWS Machine Learning (ML) Training and Certification

Learn like an Amazonian, based on the curriculum we've used to train our own developers and data scientists

AWS is how you build machine learning skills

Courses built on the curriculum leveraged by Amazon's own teams. Learn from the experts at AWS.

Flexibility to learn your way

Learn online with 65+ on-demand digital courses or live with virtual instructor-led training, plus hands-on labs and opportunities for practical application.

Validate your expertise

Demonstrate expertise in building, training, tuning, and deploying machine learning models with an industry-recognized credential.

aws.training/machinelearning

Thank You for Attending AWS Innovate

We hope you found it interesting! A kind reminder to **complete the survey.**
Let us know what you thought of today's event and how we can improve the event experience for you in the future.


aws-apac-marketing@amazon.com


twitter.com/AWSCloud


facebook.com/AmazonWebServices


youtube.com/user/AmazonWebServices


slideshare.net/AmazonWebServices


twitch.tv/aws

aws INNOVATE
AI/ML EDITION

Thank you!