

★ Get unlimited access to the best of Medium for less than \$1/week. [Become a member](#)



AWS Sage Maker



Rajuhegde

Follow

6 min read · Apr 3, 2024



Build, Train and Deploy a ML model on Amazon Sage Maker

The article is for those Machine learning practitioners who know the model building and even they have deployed some projects on other platforms but want to learn how to deploy on major cloud platforms like AWS.

Introduction to Amazon Web Services (AWS)

What is AWS? — AWS is a cloud computing service that provides on-demand computing resources for storage, networking, Machine learning, etc on a pay-as-you-go pricing model. AWS is a premier cloud computing platform around the globe, and most organization uses AWS for global networking and data storage.

These services or building blocks are designed to work with each other, and result in applications which are sophisticated and highly scalable.

Amazon's Web Services have a series of optimized services specifically tailored for Artificial Intelligence and Machine Learning Algorithms. These fit into three major tiers, as follows:

Application Services — These are domain-based services which allow us to very quickly generate predictions with pre-trained models using simple API calls.

Platform Services — Unlike Application services, platform services allow us to build our customized Machine Learning and AI solutions through optimized and scalable options. The SageMaker service that we will discuss in this article, falls in this tier.

Frameworks and Hardware — The tiers mentioned above run on top of the frameworks and hardware tier. This layer provides a wide range of optimized deep learning tools like TensorFlow, Keras, Pytorch and Apache MXNet. Options of compute options (GPU, CPU) are also available.

Some Benefits of Using AWS SageMaker

- Highly Scalable
- Fast Training
- Maintains Uptime — Process keeps on running without any stoppage.
- High Data Security

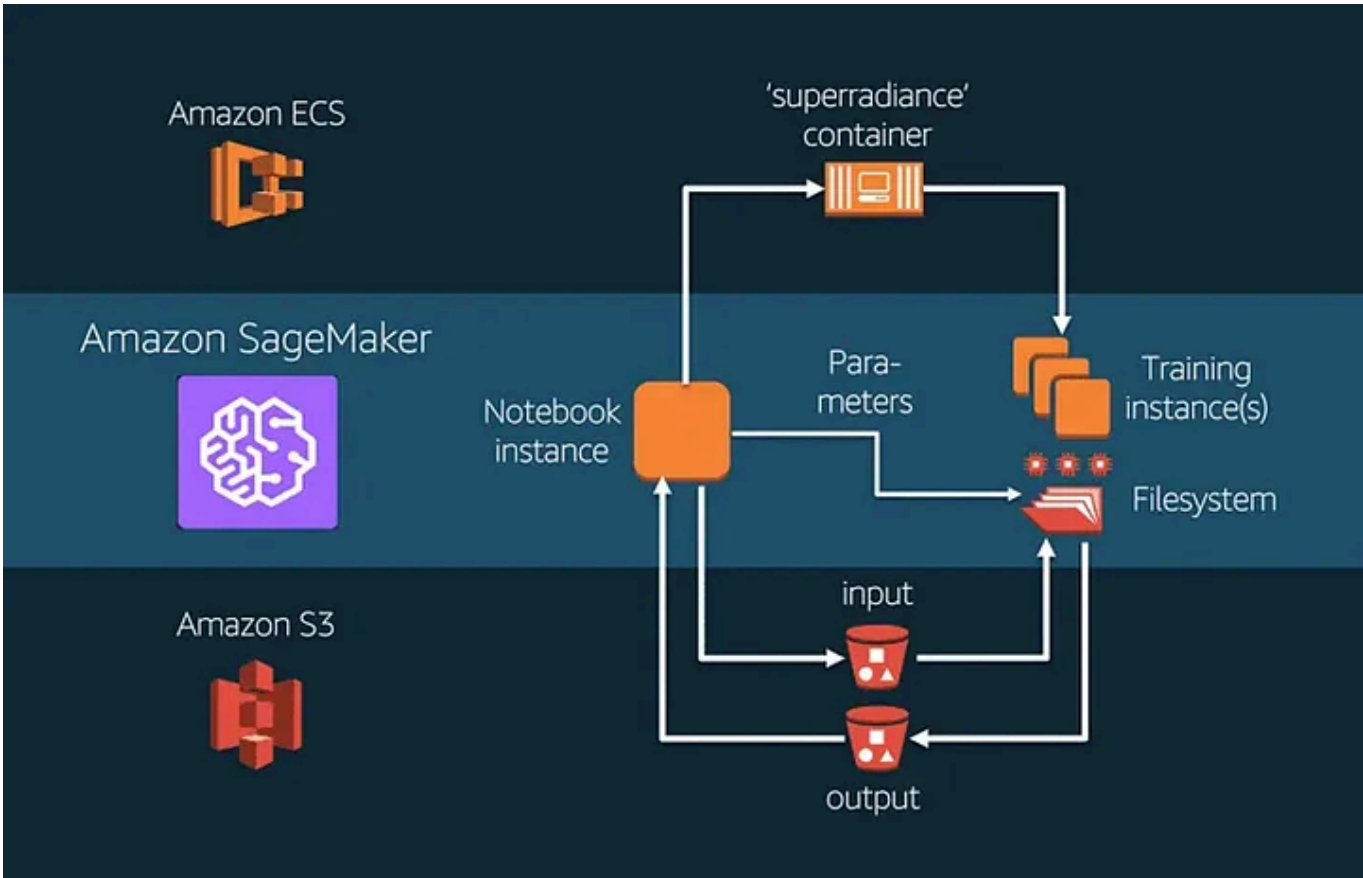


Fig 1: Amazon SageMaker

Data Analytics Pipeline :

Data is a strategic asset of every organization. As data continues to grow, databases are becoming increasingly pivotal to understanding data and converting it to valuable insights. Before data can be analyzed, it needs to be generated, collected, stored and processed. You can think of this as an analytics pipeline that extracts data from source systems, processes the data, and then loads it into data stores where it can be analyzed. Analytics pipelines are designed to handle large volumes of incoming data from heterogeneous sources such as databases, applications, and devices.



Fig 2 : Data Analytics Pipeline — Image — AWS

SageMaker combines these modules and works with three major components:

Build — Involves data extraction from S3, Docker or any other storage option used. Processing and feature engineering follow.

Train — This component combines model tuning and model training

Deploy — This component allows us to deploy the predictions and save them to the preferred storage location.

These components are independent of each other and can be used separately or even in required combinations.

Machine Learning with the SageMaker

Machine Learning is the hottest topic in the current era and the leading cloud provider Amazon web service (AWS) provides lots of tools to explore Machine Learning, creating models with a high accuracy rate. This article makes you familiar with one of those services on AWS i.e Amazon Sagemaker which helps in creating efficient and more accurate rate Machine learning models the other benefit is that you can use other AWS services in your model such as S3 bucket, amazon Lambda for monitoring the performance of your ML model you can use AWS Cloudwatch which is a monitoring tool.

The SageMaker comes with a lot of built-in optimized ML algorithms which are widely used for training purposes. Now to build a model, we need data. We can either collect and prepare training data by ourselves or we can choose from the Amazon S3 buckets which are the storage service (kind of like harddrives in your system) inside the AWS SageMaker. Lets see how we can make use of this service to build an end-to-end ML project.

AWS SageMaker simplifies ML modeling into three steps: preparation, training and deployment.

These components are independent of each other and can be used separately or even in required combinations.

The management of SageMaker components is extremely easy through Amazon SageMaker Console which has a very clean layout, making options for the different components easily accessible and configurable.

Data pipelines for machine learning

Training pipelines and inference pipelines are both needed in order to continually train machine learning models.

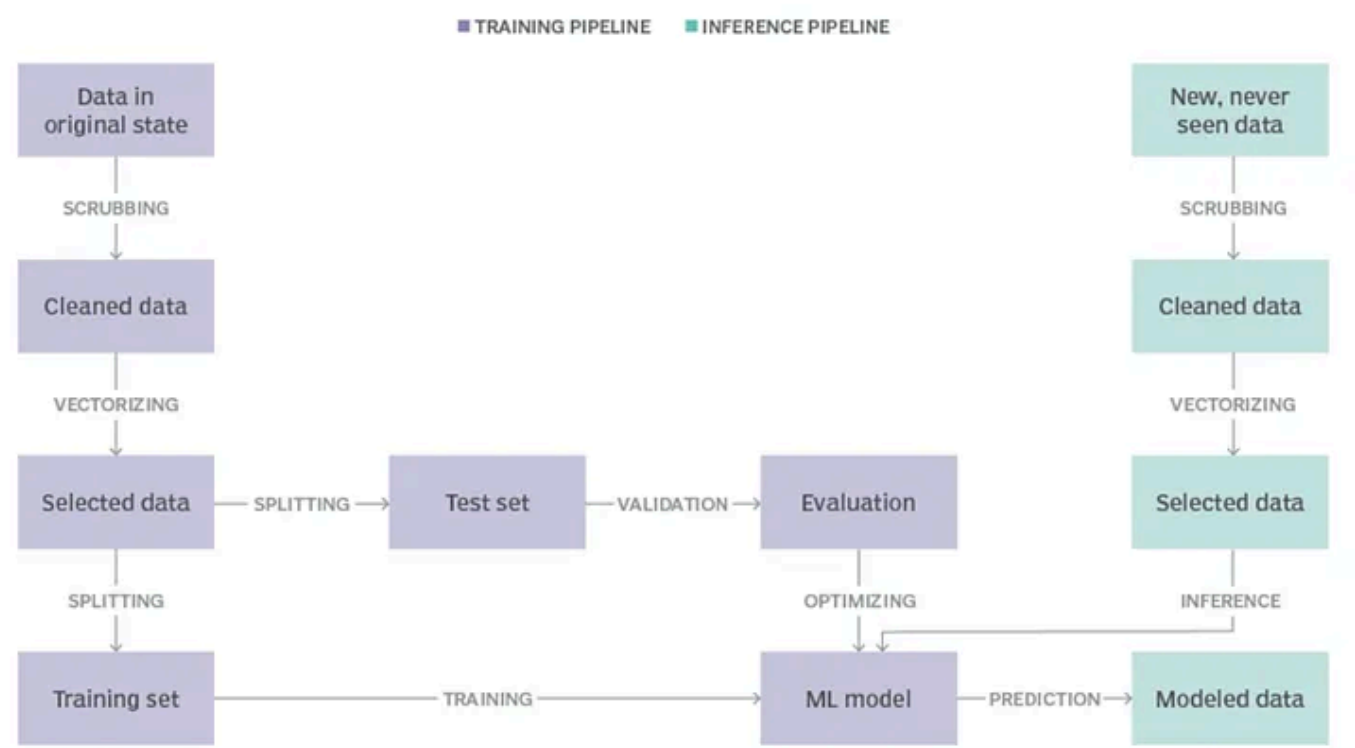


Fig 3 : Tech Target

BUILD

The build phase initiates the first interaction of Data with the Pipeline. The easiest way to do this is to generate SageMaker’s Notebook Instances. This not only enables the integration of the required code but also facilitates clear documentation and visualization.

Amazon SageMaker makes it easy to build ML models and get them ready for training by providing everything you need to quickly connect to your training data and select and optimize the best algorithm and framework for your application. Amazon SageMaker includes hosted Jupyter notebooks that make it easy to explore and visualize your training data stored on Amazon S3. You can connect directly to data in S3, or use AWS Glue to move data from Amazon RDS, Amazon DynamoDB, and Amazon Redshift into S3 for analysis in your notebook.

TRAIN

We can begin training our model with a single click in the Amazon SageMaker console. Amazon SageMaker manages all the underlying infrastructure for you and can easily scale to train models at the petabyte scale. To make the training process even faster and easier, AmazonSageMaker can automatically tune your model to achieve the highest possible accuracy.

Setting up the training module in SageMaker is extremely easy and feasible. The primary attractions of the training component in SageMaker are as follows:

Minimal Setup requirements — on creating a simple training job, SageMaker takes care of the hardware requirements and backend jobs like fetching storage and instances.

Dataset Management — SageMaker takes care of streaming data and also helps manage distributed computing facilities which can help increase the speed of training.

Containerization — All models in SageMaker, whether it is an in-built model like XGBoost or K-Means Cluster, or a custom model integrated by the user, are stored in Docker containers. SageMaker efficiently manages and integrates the containers without any external aid from users.

DEPLOY

There are several deployment options in SageMaker. With SageMaker's UI, it is a one-step deployment process, providing high reliability with respect to quality, scalability and high throughput facilities.

Several models can be deployed using the same end-point (the point of deployment) so that the model can go through A/B testing which is supported by SageMaker.

Once model is trained and tuned, Amazon SageMaker makes it easy to deploy in production so you can start running and generating predictions on new data (a process called inference). Amazon SageMaker deploys your model on an auto-scaling cluster of Amazon EC2 instances that are spread across multiple availability zones to deliver both high performance and high availability. Amazon SageMaker also includes built-in A/B testing capabilities to help you test your model and experiment with different versions to achieve the best results.

Amazon SageMaker takes away the heavy lifting of machine learning, so you can build, train, and deploy machine learning models quickly and easily.

SageMaker Workflow

1. Create Notebook Instance
2. Understanding Libraries Used
3. Creating S3 Buckets
4. Loading Data into S3
5. Building and Training the Model

6. Deployment

7. Prediction

8. Clean Up

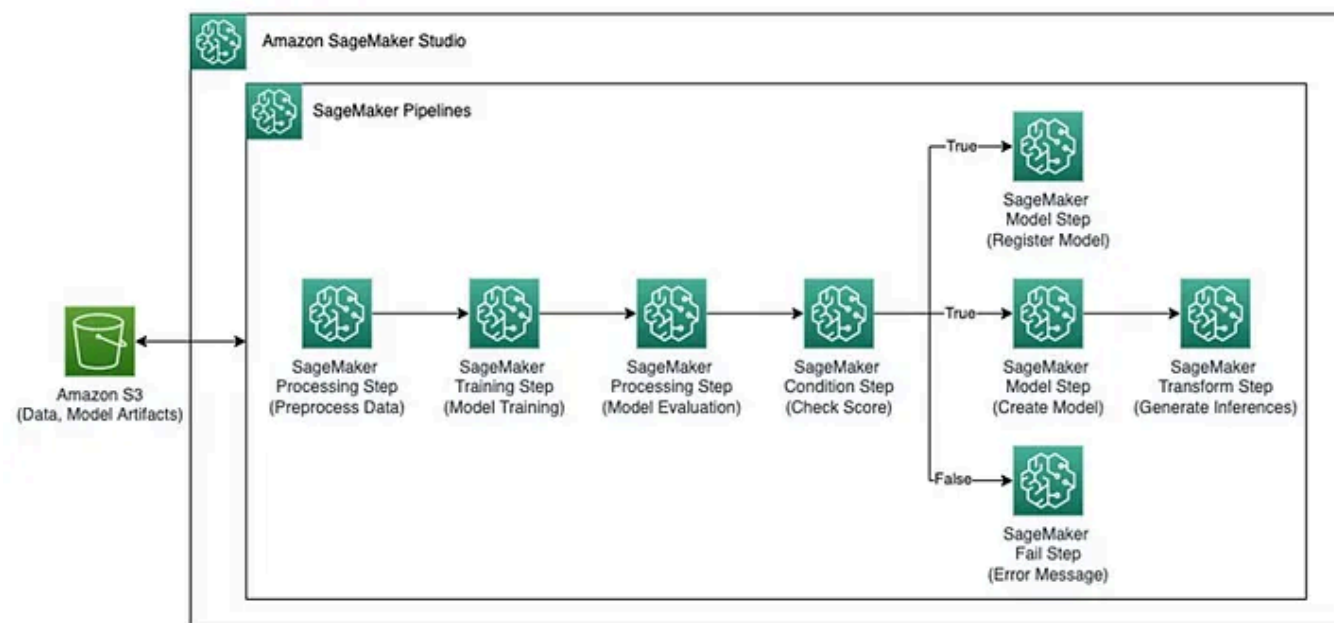


Fig 4 : AWS Workflow

The AWS platform provides a range of tools and resources for building and optimizing machine learning models, as well as for managing the end-to-end machine learning workflow. By using AWS SageMaker, the company was able to improve its customer service operations and ensure that its customers were receiving timely and effective .

Reference :

<https://medium.com/analytics-vidhya/data-analytics-lifecycle-overview-aws-2bb73af90bad>

<https://towardsdatascience.com/aws-sagemaker-db5451e02a79>


<https://pub.towardsai.net/deploying-machine-learning-models-as-a-service-using-aws-ec2-97a6434194ca>

<https://medium.com/analytics-vidhya/aws-sagemaker-service-1639d778aa2f>

<https://www.techtarget.com/searchaws/definition/Amazon-SageMaker>

<https://www.simplilearn.com/tutorials/aws-tutorial/aws-sagemaker>

<https://www.geeksforgeeks.org/what-is-sagemaker-in-aws/>




Written by Rajuhegde

3 followers · 3 following

Follow

Data Scientist | Software Testing Automation | Business Analyst| AI Subject Expert| Data Story Telling| Project Consulting | ML & DL Expert | Mentor | Blogger

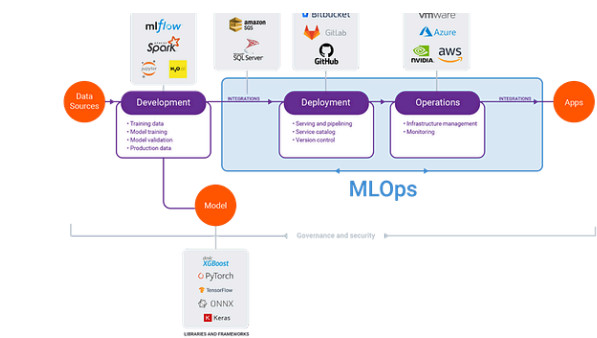
No responses yet



lury Miguel

What are your thoughts?

More from Rajuhegde





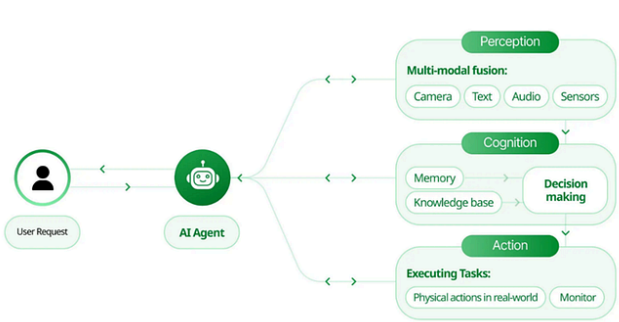
Rajuhegde

MLOps: Continuous Delivery and Automation Pipelines in Machine...

As per the Statista survey, the total amount of data created, captured, copied, and...

Mar 16, 2024





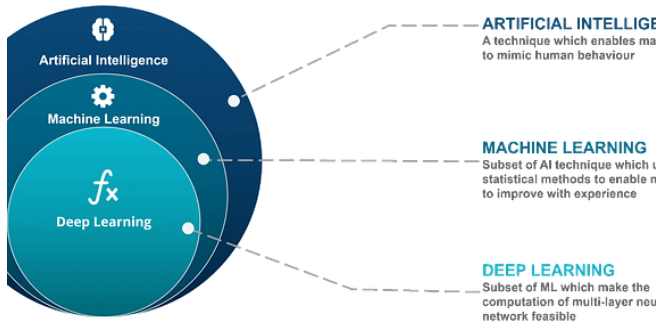
Rajuhegde

Overview on Customizing and Integrating Agentic AI Solutions

Agentic AI, a paradigm shift in artificial intelligence, is designed to mimic the...

Feb 4

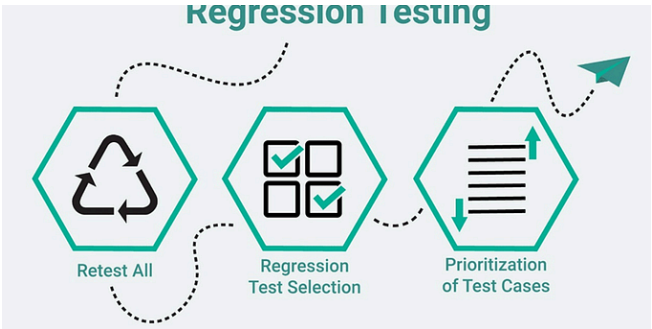


Rajuhegde

Assistances of Artificial Intelligence in future Software Te...

Artificial intelligence (AI) is the main topic of discussion in the present software...

Oct 20, 2021 12



Rajuhegde

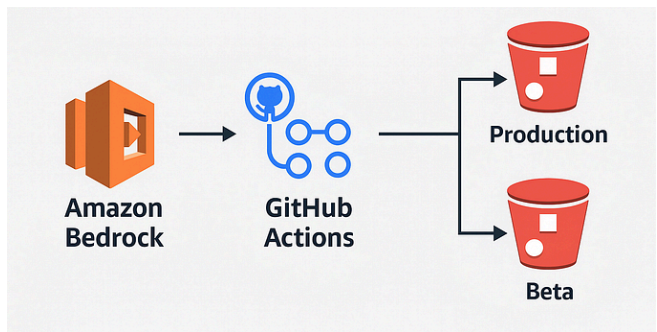
Need of Test Suite Size Reduction in Software Testing

Quality of a Software can be assured by going through the Software Testing process of the...

Mar 16, 2024

See all from Rajuhegde

Recommended from Medium



Eric Francis

Using AWS Bedrock, S3 Buckets and GitHub Actions to help Pixel...

By Eric Francis

Jul 20 4



ThreadSafe Diaries

He Was a Senior Developer, Until We Read His Pull Request

When experience doesn't translate to expertise, and how one code review change...

Aug 3 4.1K 127



Paul issack minoltan

AWS Hands-On | S3 Introduction | S3 Bucket Policy

Notice



TechPreneur

Why Amazon is Quietly Abandoning the Cloud

May 31



For years, Amazon Web Services (AWS) was the undisputed king of cloud computing. It...

★ Jul 11 🖱️ 2.2K 💬 96 📌 ⋮



Pablo Pazos

Why Everyone Is Migrating from MLflow to Weights & Biases (W&B...

“The tooling you use defines the speed at which you innovate.”

★ May 15 🖱️ 13 📌 ⋮



Jordan Gibbs

ChatGPT Is Poisoning Your Brain...

Here’s How to Stop It Before It’s Too Late.

★ Apr 29 🖱️ 24K 💬 1197 📌 ⋮

See more recommendations