

# AWS Sagemaker

For Big Data Analytics

# Agenda

What is AWS?

ML Pipeline

What is Sagemaker?

What Sagemaker offers?

Sagemaker Demo

# What is AWS?

Amazon Web Services (AWS) is a secure cloud services platform, offering compute power, database storage, content delivery and other functionalities to help businesses scale and grow.



# ML Pipeline

Data  
Preprocessing

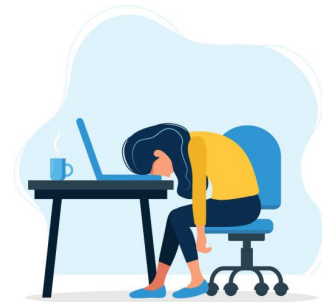
Choose and  
Train Model

Deploy Model

Monitor and  
Scale

# Pain Points in Current Pipeline

- The entire process is highly iterative
- Demands heavy and scalable infrastructure
- Scaling has 3 Dimensions (Data, Training cycles, Inference)
- Each Phase require a person with completely different skillset
- Requires collaboration between engineers and data scientists



# Sagemaker to the Rescue!

Amazon SageMaker is a fully managed machine learning service.

With Amazon SageMaker, data scientists and developers can quickly build and train machine learning models, and then deploy them into a production-ready hosted environment.



**Amazon  
SageMaker**

# What SageMaker does?

- Groundtruth labelling
- Choose and Optimize ML Models
- Setup and Manage training environment
- Train & Tune Models
- Deploy ML model to Production
- Scale and Manage production environment

Demo



# References

Code: <https://github.com/nikhilkoditala/aws-sagemaker>

AWS Sagemaker: <https://aws.amazon.com/sagemaker/>

Questions?