ML Model Deployment on AWS for Customer Churn Prediction
Project Overview

Business Overview
Churn rate, also known as turnover or customer churn, is the rate at which a customer suspends a transaction with a company. Most commonly, it is expressed as the percentage of service subscribers who canceled their subscription within a certain time period. In this project, We aim to deploy a model which predicts whether the customer is going to churn in the near future or not. Amazon Web Services(AWS) is used as a cloud provider. We would advise you to have a basic understanding of Customer Churn

Prediction Analysis using Ensemble Techniques before jumping into this project.

Aim
To deploy a model on AWS which predicts whether the customer is going to churn in the near future or not.

Tech Stack
➔ Language: Python
➔ Libraries: Flask, gunicorn
➔ Services: Flask, Docker, AWS, Gunicorn, Terraform

Prerequisites
It is advisable to have a basic knowledge of the following services to get an understanding of the project.

- Flask
- Terraform
- AWS s3
- Aws ECR
- AWS ECS
- AWS EC2 Load balancer
- AWS Code commit
- AWS Code Build
- AWS Code Deploy
- AWS Code Pipeline

Project Takeaways

1. Understanding various services provided by AWS
2. What is Terraform and how to use terraform to create AWS services?
3. How to create an AWS s3 Bucket?
4. How to create a commit repository and commit code in it?
5. How to create an ECR repository?
6. Converting ML application to Flask application
7. How to deploy the application using the Gunicorn web server?
8. Building Docker image
9. How to create an ECS Cluster?
10. Testing the docker container in ECS
11. How to create a load balancer in the EC2 service?
12. How to create and build a project in Code Build?
13. Understanding ECS Cluster Task Definition
14. How to create Code Pipeline?
15. How to store the Terraform state in the AWS s3 backend bucket?