

Configuration Manual

MSc Research Project MSc Cloud Computing

Priyal Patil Student ID: 22209573

School of Computing National College of Ireland

Supervisor: Prof. Jitendra Kumar Sharma

National College of Ireland





School of Computing

Student Name:	Priyal Narendra Patil		
Student ID:	22209573		
Programme:	MSc in Cloud Computing	Year:	2024
Module:	MSc Research Project		
Lecturer:	Prof. Jitendra Kumar Sharma		
Submission Due Date:	12/08/2024		
Project Title:	Optimizing Movie Recommendations with ML	Ops in	AWS
Word Count:	1004 Page Count:		

I hereby certify that the information contained in this (my submission) is information pertaining to research I conducted for this project. All information other than my own contribution will be fully referenced and listed in the relevant bibliography section at the rear of the project.

<u>ALL</u> internet material must be referenced in the bibliography section. Students are required to use the Referencing Standard specified in the report template. To use other author's written or electronic work is illegal (plagiarism) and may result in disciplinary action.

Signature: Priyal Patil

Date: 12/08/2024

PLEASE READ THE FOLLOWING INSTRUCTIONS AND CHECKLIST

Attach a completed copy of this sheet to each project (including multiple copies)	
Attach a Moodle submission receipt of the online project submission, to each project (including multiple copies).	
You must ensure that you retain a HARD COPY of the project, both for your own reference and in case a project is lost or mislaid. It is not sufficient to keep a copy on computer.	

Assignments that are submitted to the Programme Coordinator Office must be placed into the assignment box located outside the office.

Office Use Only	
Signature:	
Date:	
Penalty Applied (if applicable):	

Configuration Manual

Priyal Patil Student ID: 22209573

1. Introduction

This manual documents about all the used tools, technologies as well as some steps for building the framework. The manual is further divided into few parts such as environment setup, tools technology used and some screenshot's of code, implementation.

2. Project Requirements

This section has the essential requirements needed to successfully implement the Movie Recommendation System using MLOps tools on AWS. It covers hardware as well as the software specifications, along with the necessary cloud services and development tools:

a. Hardware requirements:

Device: MacBook air m2 with 16GB and 512 gb SSD as a local device. However, the complete project was executed over AWS Public cloud using their hardware in background.

b.<u>Software requirements:</u>

- **Operating System:** macOS Ventura./ Amazon Linux2
- Programming Language: Python 3.11
- **Framework:** Flask Framework
- **IDE:** AWS Cloud9

c. <u>Cloud Services:</u>

- AWS IAM User: Essential as to develop and deploy a serverless function and manage the API gateways as well along with Lambda.
- Sagemaker: SageMaker is a fully managed machine learning (ML) as a service in AWS ("What is Amazon SageMaker? Amazon SageMaker," n.d.). SageMaker simplifies the process of developing, training, and deploying machine learning models on a large scale.
- AWS ECS (Elastic Container Service): ECS is a container management service that is highly scalable and high-performance, it supports Docker containers and enables you to run applications on a managed cluster of servers.
- AWS Lambda: Lambda enables you to execute code without the need to allocate or oversee servers. It is utilized to carry out behind-the-scenes operations in reaction to AWS service events (such as modifications in an S3 bucket).

- AWS CloudWatch: CloudWatch: CloudWatch keeps track of the AWS resources and the applications that run on AWS. It is utilized for gathering and monitoring metrics, logs, setting alarms, and automatically responding to changes in AWS resources, it is essential for managing and optimizing machine learning models.
- AWS Elastic Container Registry (ECR): ECR is a Private Docker container registry that makes it easy for developers to store, manage, and deploy Docker container images.
- AWS Cloud9: Cloud9 is an online integrated development environment (IDE) based on the AWS that allows to write, execute, and troubleshoot code in a web browser
- Simple Storage Solution (S3): S3 is a service for storing objects that provides scalability, data availability, security, and performance. It is utilized in this project for the storing and accessing of various data types, like training datasets, model artifacts, and output results.

d. Development tools

• <u>Github:</u> GitHub is a hosting service for Git repositories that also offers tools for version control and teamwork. Enables multiple developers to collaborate on projects remotely, a critical aspect of team-based software development.

3. Starting ML Pipeline

- 1. Login to AWS Management Console.
- 2. Go to AWS Sagemaker.
- 3. Click on Domains.
- 4. Start a Domain.
- 5. Create an user.
- 6. Launch the studio.

SageMaker Studio		🖛 Brovide feedback 💿
* 2000 * 2000 * 2000 * 2000 * 2000 * 2000 * 2000 * 2000 * 2000	Home Launch workflows, manage your applications and spaces, and view getting started in Durniced Cetting started Overview Start a new ML workflow or jump back into your workflow	naterials
≥ Joins V	Jupyter	
V Motels	JupyterLab Create, manage, and run durable instances of JupyterLab using spaces	Code Editor Bared on Code-OSS, Visual Studio Code Open Source Create, manage, and run durable instances of Code Editor using spaces
e Deployments	View JupyterLab spaces >	View Code Editor spaces >
Collapse Menu	 Prebuilt and automated solutions Deploy built-in algorithms, prebuilt solutions, and example notebooks. Build models and 	d prepare data from a visual interface.

Fig 1: SageMaker Dashboard

- 7. Click on Juypter Space.
- 8. Create a new space.
- 9. Copy my notebooks to the new space.
- 10. Run each cell.
- 11. Create a S3 bucket with names mentioned in my Notebook.

aws	Services	Q Sea	rch				[Al
	🚳 Elastic Bea	anstalk	Amazon SageMaker	🔁 S3	0 RDS	👸 Secrets Mana	ager 📆 Co
=	Amazon S3	Buckets	s > mlops-datasets	-movie >	latestda	tasets/	
	latestd	atase	ets/				
	Objects	Prope	erties				
	Objects Objects are		ental entities stored in A	C mazon S3. Y		opy S3 URI	Copy Ty Coget a
	Q Find o	objects by	prefix				
	N	ame		▲ T	ype		▼
		anime.c	sv	C	SV		
		rating.cs	5V	C	SV		

Fig 2: S3 Bucket

12. Create a Lambda function.

Lambda > Functions > SageMaker-Pipeline-S3-Trigger	
SageMaker-Pipeline-S3-Trigger	
▼ Function overview Info	
Diagram Template	
SageMaker-Pipeline-S3-Trigger	
Layers (0)	
S 3	+ Add destination
+ Add trigger	
Code Test Monitor Configuration Aliases Versions	
Code lest Monitor Connguration Atlases Versions	
Code source Info	
▲ File Edit Find View Go Tools Window Test ▼ Deploy	
Q Go to Anything (Ctrl-P) 🖥 lambda_function × Environment Var × Execution results × 🕀	
* Supstans-Pysites O import botol impor	
<pre>9 pipeline'awa - 'saganakar ekops-train-pipeline' 10 11 12 try'; Start the pipeline execution 11 repose - liont.trut pipeline genution(14 Pipelinetworking) is genution(14 Pipelinetworking) is pipeline's excitation' restands' 15 return (16 'status:Code': 200, 18 'status: code': 200, 19 'sodd': jon.dwarg('successfully started pipeline execution: ' + r 21 except teoption as e:</pre>	esponse['PipelineExecutionArn'])
22 print(0) 23 return (succode'; 500, 34 'statucode'; 500, 35 'sody'; json.demps('sailed to start pipeline execution: ' + str(e) 36 })

Fig 3: Lambda Function and Code

- 13. Add a trigger as S3 (PUT action only)
- 14. Insert dataset files into S3 bucket.
- 15. Re enter into Sagemaker and click on Pipelines from left panel.

16. There will be a pipeline executing due to new dataset uploaded in the S3 bucket.

PIPELINE sagemaker-mlops-train-pipeline						
OVERVIEW Executions	~	Executions	T			m
Graph		Name	Status	Elapsed Time	Modified On	Created On
SETTINGS	~	ExecutionFromLambda	Succeeded	2m 33s	2 hours ago	2 hours ago
Parameters	~	execution-1723456133880	Succeeded	2m 34s	2 hours ago	2 hours ago
Information	Ť	ExecutionFromLambda	Failed	2s	2 hours ago	2 hours ago
		ExecutionFromLambda	Succeeded	2m 58s	4 days ago	4 days ago
		ExecutionFromLambda	Succeeded	2m 34s	5 days ago	5 days ago
		ExecutionFromLambda	Succeeded	2m 33s	5 days ago	5 days ago
		execution-1723062237773	Succeeded	2m 34s	5 days ago	5 days ago

Fig 4: SageMaker Pipeline

4. Starting CI/CD Pipeline

- 1. Start an Cloud9 instance.
- 2. Git clone: https://github.com/priyalpatil98/Movie-Recommendation-System-MLOps.git

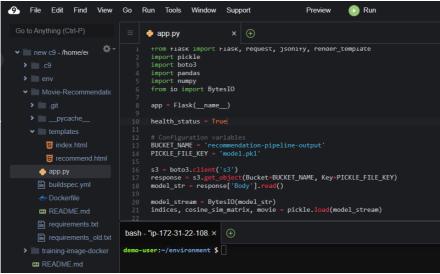


Fig 5: Cloud9 IDE

3. Create CodePipeline.

- 4. Add Source as Github.
- 5. Create CodeBuild project.
- 6. Use Buildspec.yml file to build project.
- 7. Create Deploy stage and add ECS.
- 8. Create a ECR and build an Image on Cloud9 using DockerFile and push to the ECR.

Amazon	Amazon ECR > Private registry > Repositories					
Priva	Private repositories					
Rep	ositories (2)	C				
Q	Filter status					
	Repository name	URI				
0	mlops-repository- aws	590183699263.dkr.ecr.eu-west- 1.amazonaws.com/mlops-repository-aws				
0	my-sagemaker- training-image	590183699263.dkr.ecr.eu-west- 1.amazonaws.com/my-sagemaker-training- image				

Fig 6: AWS ECR Repo

- 9. Trigger the pipeline by a git push to the main branch.
- 10. Wait until all three phases are implemented.

💌 🖸 ChatGi 🕲 Recomi 📓 AWS C 📓 Re. X 🔯 Bastici 💆 Senice 📓 new Ci +	×
← → C 🕾 eu-west-1.console.aws.amazon.com/codesuite/codepipelin Q ☆ 🖸 🖆) 1
Voluble Imported Holl The Provide Statement Sound Statement Sound The Imported Login	9263 🗸
😇 IAM 🥝 Elastic Beanstalk 👩 Amazon SageMaker 😨 53 🔯 RDS 🔞 Secrets Manager 📆 CodePipeline 💦 Lambda 🕅 API G	atewa >
Developer Tools > CodePipeline > Pipelines > Hask-app-pipeline	0
flask-app-pipeline A Notify * Edit Storp execution Clone pipeline Pelease change	
Pjeeline type; V2 Execution mode; QUEUED Source Succeeded Pjeeline execution ID; addba/3-5081-609-bab/-56851;782ada	
Source GitHub/Hansion.11 02 ⊘ Succeeded - <u>Immutes.age</u> Oblacided 02 View details	
Histophillis (If Source: Final Commit	
Disable transition O Build O Succeeded Start rollback Image: Control Contron Control Control Control Control	
Build MXS CodeBuild Succeeded - Eminutes.app View details	
(ISBA2205) ID Source: Final Commit	
Disable transition	
Deployment Succeeded Start rollback Pipeline execution ID: arbite3:001-4698-bite0-66851r742ada	
ECS_Deployment Amain ECS 18 Succeeded - Iminat ago View details	
858a2855 🖸 Source: Final Commit	
-	
CloudShell Feedback Privacy Terms Cookie pre	ferences
0 2024, Amazon Web Services, Inc. or its affiliates. ENG 🗇 🗇 边	

Fig 7: CI/CD CodePipeline

References

What is Amazon SageMaker? - Amazon SageMaker [WWW Document], n.d. URL https://docs.aws.amazon.com/sagemaker/latest/dg/whatis.html (accessed 8.12.24).