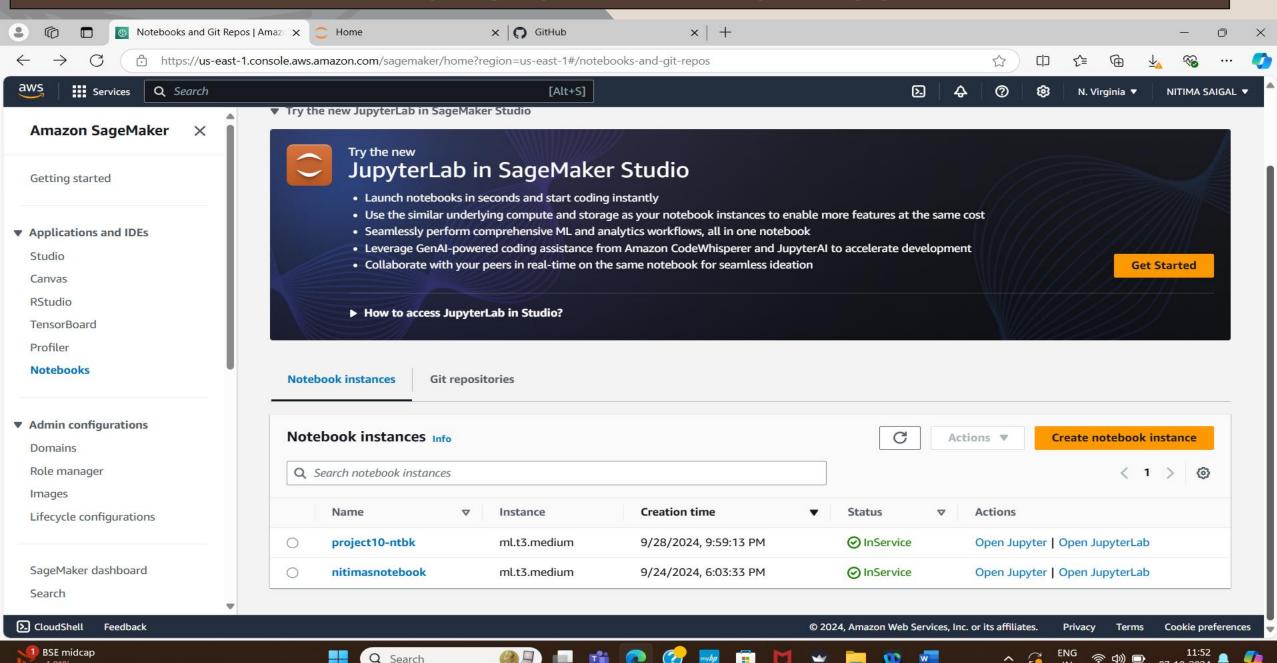# Custom OCR on AWS

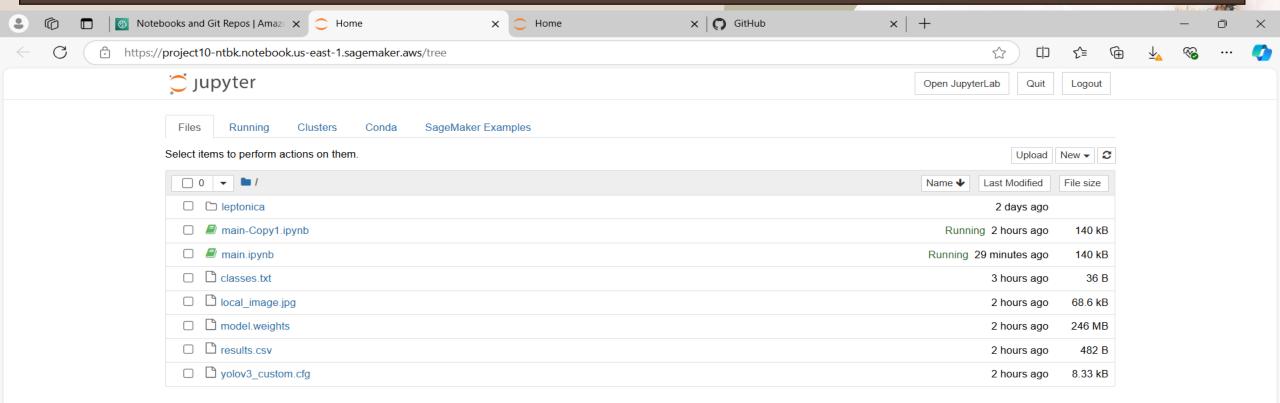**BUILDING A CUSTOM OCR BY COMBINING YOLO AND TESSERACT, TO READ THE SPECIFIC CONTENTS OF A LAB AND TRAINING A CUSTOM YOLO_V3 MODEL USING OUR CUSTOM DATASET.**
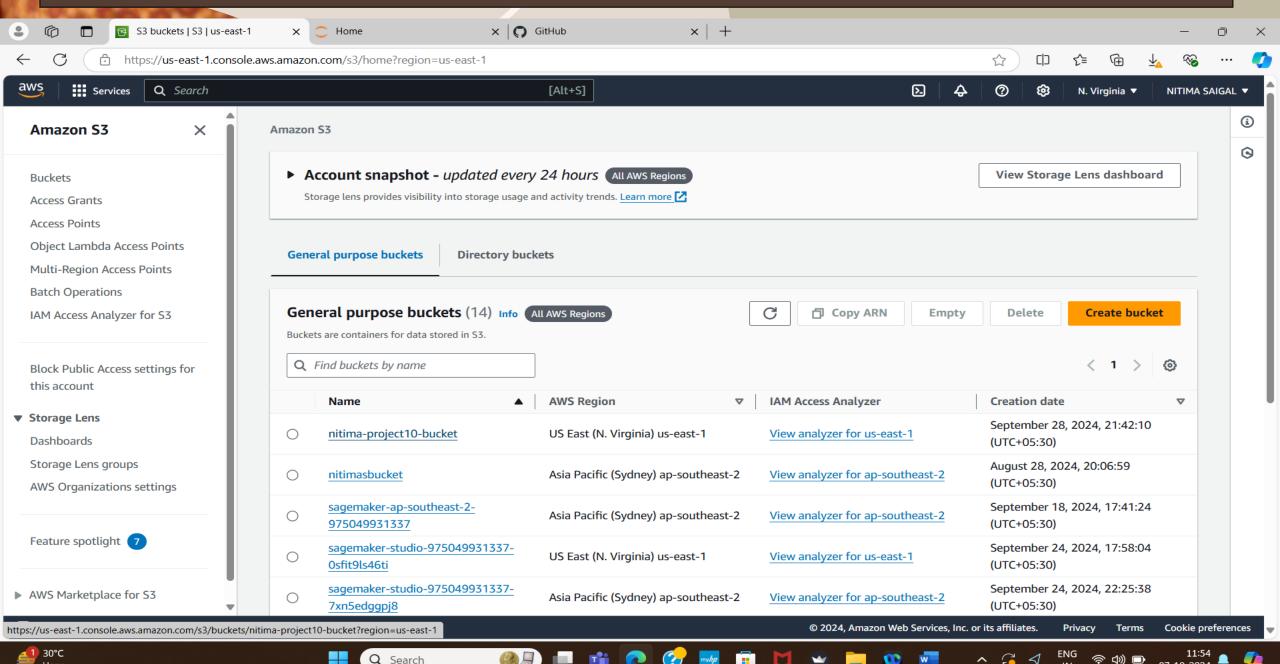
AMAZON SAGEMAKER NOTEBOOK

Notebooks and Git Repos | Amaz... × | Home × | GitHub × | +

https://us-east-1.console.aws.amazon.com/sagemaker/home?region=us-east-1#/notebooks-and-git-repos

aws Services | Q Search [Alt+S] | N. Virginia ▼ | NITIMA SAIGAL ▼

Amazon SageMaker ✕

Getting started

▼ **Applications and IDEs**

Studio

Canvas

RStudio

TensorBoard

Profiler

Notebooks

▼ **Admin configurations**

Domains

Role manager

Images

Lifecycle configurations

SageMaker dashboard

Search

▼ Try the new JupyterLab in SageMaker Studio

Try the new
JupyterLab in SageMaker Studio

- Launch notebooks in seconds and start coding instantly
- Use the similar underlying compute and storage as your notebook instances to enable more features at the same cost
- Seamlessly perform comprehensive ML and analytics workflows, all in one notebook
- Leverage GenAI-powered coding assistance from Amazon CodeWhisperer and JupyterAI to accelerate development
- Collaborate with your peers in real-time on the same notebook for seamless ideation

Get Started

▶ How to access JupyterLab in Studio?

Notebook instances | **Git repositories**

Notebook instances Info

[Actions ▼] [**Create notebook instance**]

Q Search notebook instances

< 1 >

	Name ▽	Instance	Creation time ▽	Status ▽	Actions
○	**project10-ntbk**	ml.t3.medium	9/28/2024, 9:59:13 PM	⊘ InService	Open Jupyter | Open JupyterLab
○	**nitimasnotebook**	ml.t3.medium	9/24/2024, 6:03:33 PM	⊘ InService	Open Jupyter | Open JupyterLab

CloudShell   Feedback

© 2024, Amazon Web Services, Inc. or its affiliates.   Privacy   Terms   Cookie preferences

Notebooks and Git Repos | Amaz... Home Home GitHub

https://project10-ntbk.notebook.us-east-1.sagemaker.aws/tree

jupyter

Files   Running   Clusters   Conda   SageMaker Examples

Select items to perform actions on them.

Upload   New ▾   ↻

	▢ 0 ▾ ▣ /		Name ↓	Last Modified	File size
▢	📁 leptonica			2 days ago	
▢	📗 main-Copy1.ipynb		Running	2 hours ago	140 kB
▢	📗 main.ipynb		Running	29 minutes ago	140 kB
▢	📄 classes.txt			3 hours ago	36 B
▢	📄 local_image.jpg			2 hours ago	68.6 kB
▢	📄 model.weights			2 hours ago	246 MB
▢	📄 results.csv			2 hours ago	482 B
▢	📄 yolov3_custom.cfg			2 hours ago	8.33 kB

Amazon S3

Buckets
Access Grants
Access Points
Object Lambda Access Points
Multi-Region Access Points
Batch Operations
IAM Access Analyzer for S3

Block Public Access settings for this account

▼ Storage Lens
Dashboards
Storage Lens groups
AWS Organizations settings

Feature spotlight 7

▸ AWS Marketplace for S3

Amazon S3

▸ **Account snapshot** - *updated every 24 hours* [All AWS Regions]   [View Storage Lens dashboard]

Storage lens provides visibility into storage usage and activity trends. Learn more ☒

| General purpose buckets | Directory buckets |

General purpose buckets (14) Info [All AWS Regions]

Buckets are containers for data stored in S3.

[↻] [⧉ Copy ARN] [Empty] [Delete] [**Create bucket**]

🔍 Find buckets by name   ‹ 1 › ⚙

○	Name ▲	AWS Region ▽	IAM Access Analyzer	Creation date ▽
○	nitima-project10-bucket	US East (N. Virginia) us-east-1	View analyzer for us-east-1	September 28, 2024, 21:42:10 (UTC+05:30)
○	nitimasbucket	Asia Pacific (Sydney) ap-southeast-2	View analyzer for ap-southeast-2	August 28, 2024, 20:06:59 (UTC+05:30)
○	sagemaker-ap-southeast-2-975049931337	Asia Pacific (Sydney) ap-southeast-2	View analyzer for ap-southeast-2	September 18, 2024, 17:41:24 (UTC+05:30)
○	sagemaker-studio-975049931337-0sfit9ls46ti	US East (N. Virginia) us-east-1	View analyzer for us-east-1	September 24, 2024, 17:58:04 (UTC+05:30)
○	sagemaker-studio-975049931337-7xn5edggpj8	Asia Pacific (Sydney) ap-southeast-2	View analyzer for ap-southeast-2	September 24, 2024, 22:25:38 (UTC+05:30)

AMAZON S3 BUCKET FILES - DATASETS, MODELS & RESULTS

aws Services | Search [Alt+S] | N. Virginia ▾ | NITIMA SAIGAL ▾

Amazon S3

Buckets
Access Grants
Access Points
Object Lambda Access Points
Multi-Region Access Points
Batch Operations
IAM Access Analyzer for S3

Block Public Access settings for this account

▾ **Storage Lens**
Dashboards
Storage Lens groups
AWS Organizations settings

▸ Feature spotlight 7

▸ AWS Marketplace for S3

Amazon S3 > Buckets > nitima-project10-bucket > datasets/

datasets/

[Copy S3 URI]

Objects | Properties

Objects (1) Info

[⟳] [Copy S3 URI] [Copy URL] [⬇ Download] [Open ⧉] [Delete] [Actions ▾] [Create folder] [⬆ Upload]

Objects are the fundamental entities stored in Amazon S3. You can use **Amazon S3 inventory** ⧉ to get a list of all objects in your bucket. For others to access your objects, you'll need to explicitly grant them permissions. **Learn more** ⧉

[🔍 Find objects by prefix]   ‹ 1 › ⚙

☐	Name ▲	Type ▽	Last modified ▽	Size ▽	Storage class ▽
☐	🗋 Dataset_OCR/	Folder	-	-	-

CloudShell Feedback   © 2024, Amazon Web Services, Inc. or its affiliates. Privacy Terms Cookie preferences
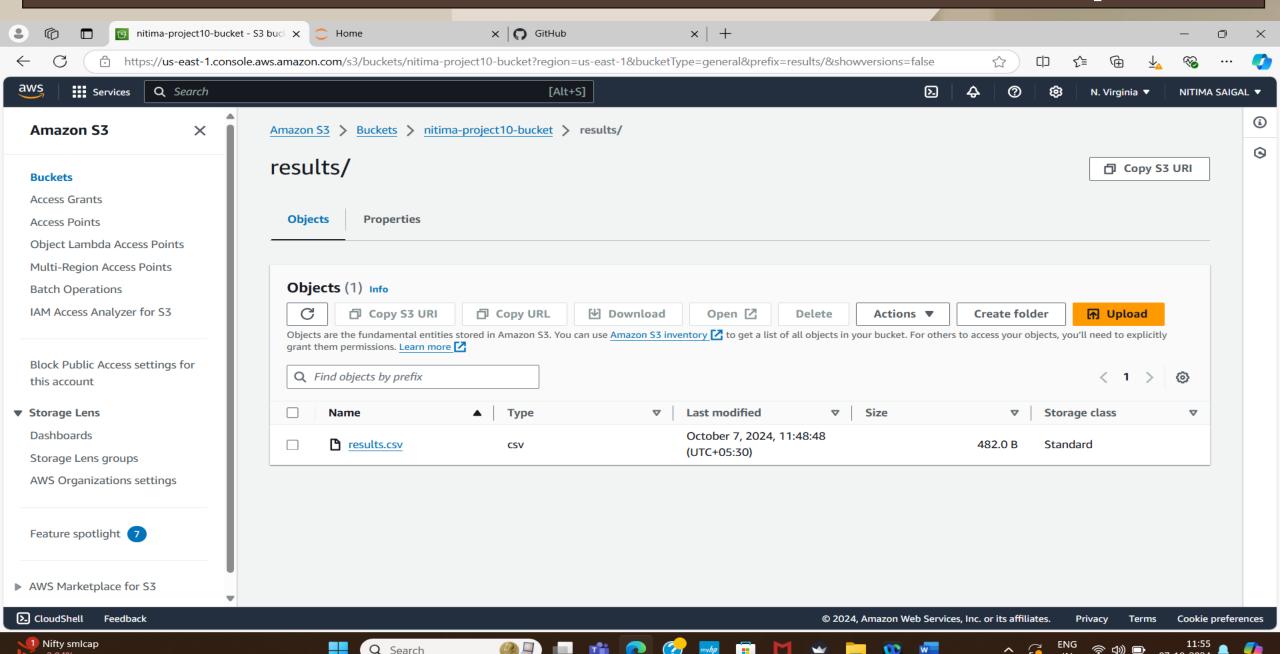
Dataset_OCR contains all preprocessed images for training & testing

Models contains both cfg and weights models

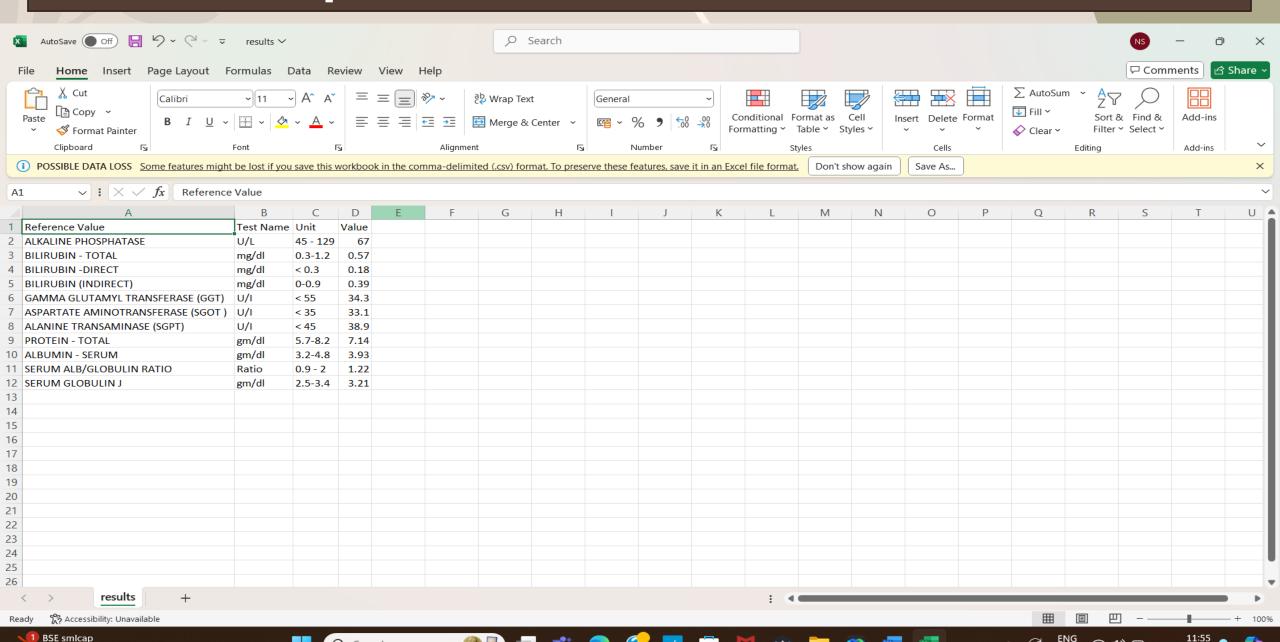# results folder contains results.csv file which is the output

The output results.csv file which shows all the text

	Reference Value	Test Name	Unit	Value
1	Reference Value	Test Name	Unit	Value
2	ALKALINE PHOSPHATASE	U/L	45 - 129	67
3	BILIRUBIN - TOTAL	mg/dl	0.3-1.2	0.57
4	BILIRUBIN -DIRECT	mg/dl	< 0.3	0.18
5	BILIRUBIN (INDIRECT)	mg/dl	0-0.9	0.39
6	GAMMA GLUTAMYL TRANSFERASE (GGT)	U/l	< 55	34.3
7	ASPARTATE AMINOTRANSFERASE (SGOT)	U/l	< 35	33.1
8	ALANINE TRANSAMINASE (SGPT)	U/l	< 45	38.9
9	PROTEIN - TOTAL	gm/dl	5.7-8.2	7.14
10	ALBUMIN - SERUM	gm/dl	3.2-4.8	3.93
11	SERUM ALB/GLOBULIN RATIO	Ratio	0.9 - 2	1.22
12	SERUM GLOBULIN J	gm/dl	2.5-3.4	3.21

○ **Created detailed documentation of the entire process, including model training, inference, and post-processing steps.**

○ **Stored documentation files in S3.**

○ **Generated Final Reports**

○ **Compiled a final report summarizing the project, including model performance, challenges, and next steps.**

○ **Uploaded the report to S3.**

thank you

NITIMA SAIGAL

9205262297

NITIMASAIGAL@GMAIL.COM