BREAST CANCER PREDICTION USING AWS SAGEMAKER

Name : Tijo Thomas

INTRODUCTION

1.1 Overview

Breast cancer is the second most leading cancer in women compared to all other cancers . Around 1.1 million cases were recorded in 2004.Obeserved rates of this cancer increased with industrialization and urbanization and also with facilities for early detection. It remains much more common in high-income countries but now is increasing rapidly in middle and low income countries as well. Breast cancer is fatal in under half of all cases and is the leading cause of death from cancer in women , accounting to 16% of the total world population

1.2 Purpose

The purpose of this project is to predict whether the patient has breast cancer or not using ml algorithm and give maximum accuracy for the same. For this we have taken the dataset from the Wisconsin breast cancer database (kaggle) which is the benchmark for comparing the results.

LITERATURE SURVEY

2.1 Existing problem

To identify which machine learning classifier gives the best accuracy . To count the number of patients having benign and malignant diseases for identifying the tumor.Breast cancer is one of the main causes of cancer death worldwide. Early diagnostics significantly increases the chances of correct treatment and survival, but this process is tedious and often leads to

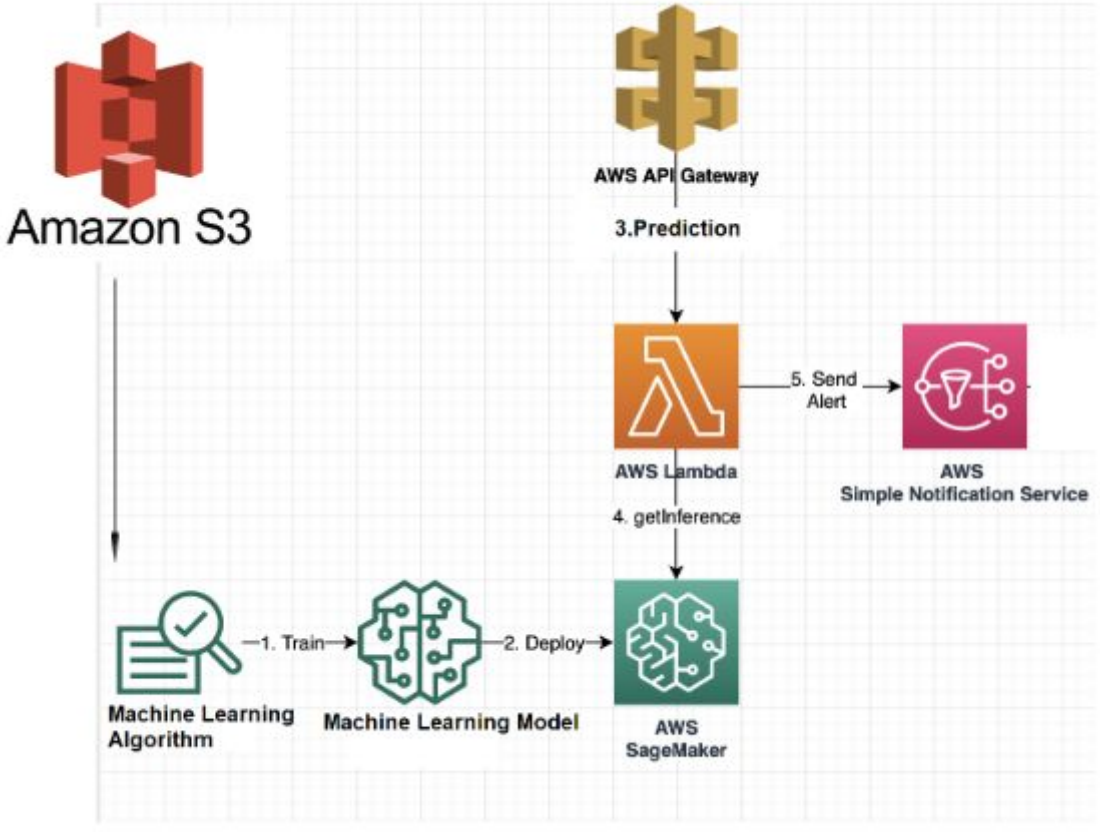
a disagreement between pathologists. Computer-aided diagnosis systems showed the potential for improving diagnostic accuracy. But early detection and prevention can significantly reduce the chances of death. It is important to detect breast cancer as early as possible.

2.2 Proposed solution

We acquired the breast cancer dataset from kaggle dataset and used amazon sagemaker's jupyter notebook for the purpose of coding . The method uses Xgboost classifier and amazon web services like Amazon S3,AWS API Gateway,AWS Lambda,AWS SNS,Amazon SageMaker for completing the project. We will be building and deploying the ml model in aws sagemaker and use sns service to generate alert about the risk

THEORETICAL ANALYSIS

3.1 Block diagram



3.2 Hardware/Software designing ;

Amazon S3,AWS API Gateway,AWS Lambda,AWS SNS,Amazon SageMaker

EXPERIMENTAL INVESTIGATIONS

Attribute Information:

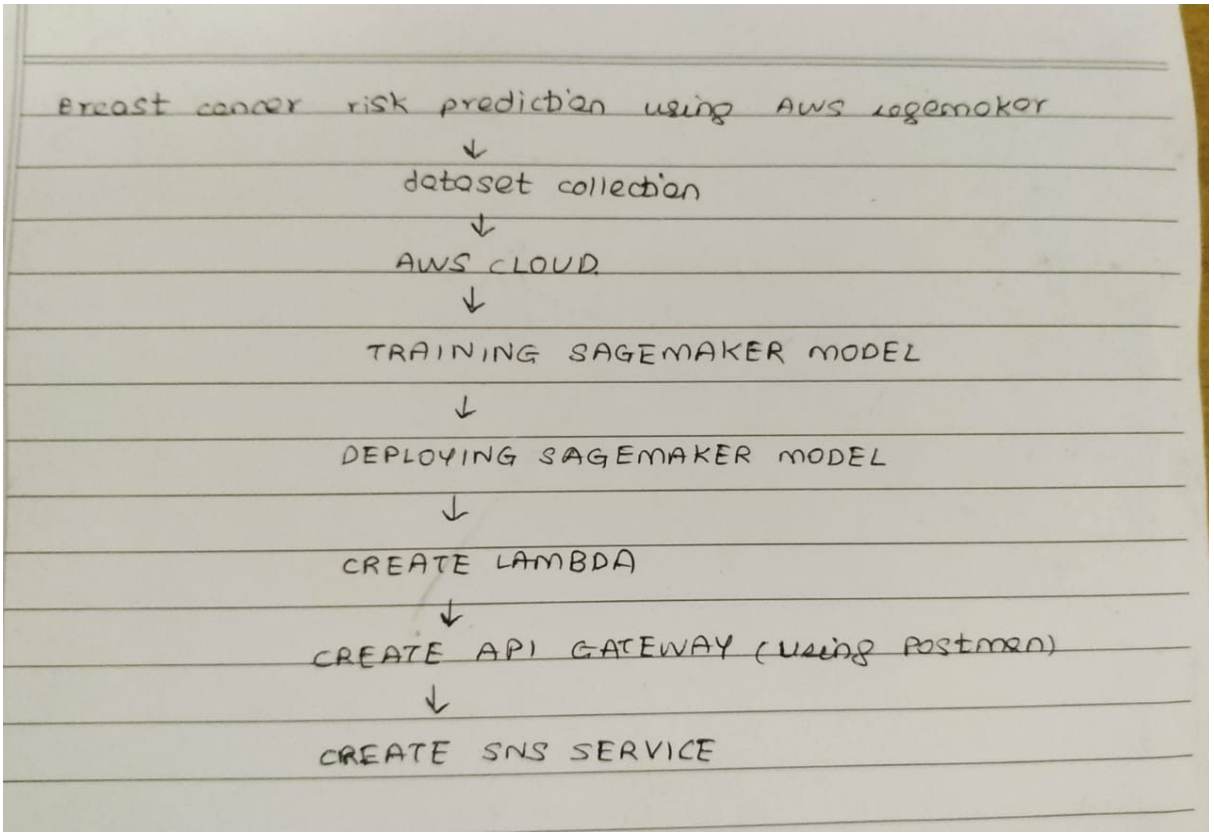
1. ID number 2) Diagnosis (M = malignant, B = benign) 3-32)

Ten real-valued features are computed for each cell nucleus:

a. radius (mean of distances from center to points on the perimeter)
b. texture (standard deviation of gray-scale values)
c. perimeter
d. area
e. smoothness (local variation in radius lengths)
f. compactness (perimeter^2 / area - 1.0)
g. concavity (severity of concave portions of the contour)
h. concave points (number of concave portions of the contour)
i. symmetry
j. fractal dimension (“coastline approximation” - 1)

The mean, standard error and “worst” or largest (mean of the three largest values) of these features were computed for each image, resulting in 30 features. For instance, field 3 is Mean Radius, field 13 is Radius SE, field 23 is Worst Radius.

FLOWCHART



RESULT

We used the Xgboost classifier and it gave an accuracy of approximately 98 %.

APPLICATIONS :

This model can be used for public use for the benefit of health care in the society after integrating it with a UI(Flask/Django) application.

CONCLUSION

The study has shown that thel machine learning algorithm achieved better discriminatory accuracy on identifying women at high risk of breast cancer using the Wisconsin dataset and XGBoost is the best choice for developing a breast cancer prediction model using breast cancer risk factors. We have successfully developed and validated a breast cancer prediction model using aws sagemaker.

FUTURE SCOPE

AI is set to change the medical industry in the coming decades - it wouldn't make sense for pathology to not be disrupted too. Currently ML model and still in testing and experimentation phase for cancer prognosis. As datasets are getting larger and of higher quality, researchers are building accurate models. You perform clinical tests , either at home or clinic . Data is inputted into a pathological ML system. A few minutes later , you receive an email with a detailed report that has an accurate prediction about the development of your cancer. ML is the next step of pathology , and it will disrupt the industry.

BIBLIOGRAPHY

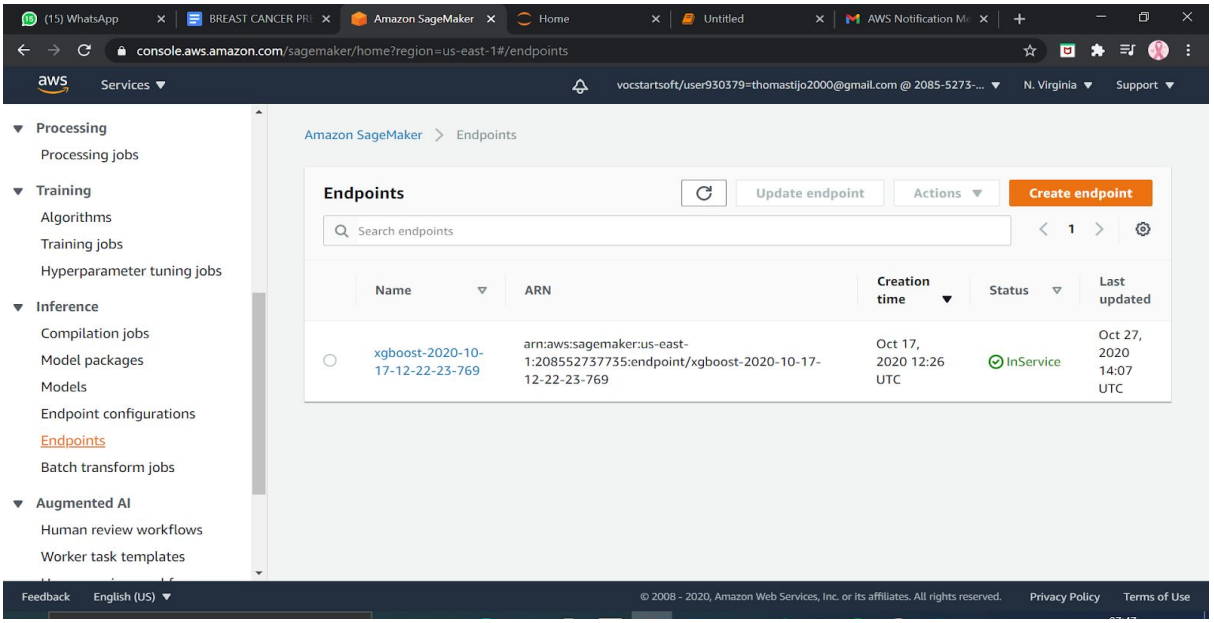
1) https://towardsdatascience.com/machine-learning-is-the-future-of-cancer-prediction-e4d28e7e6dfa

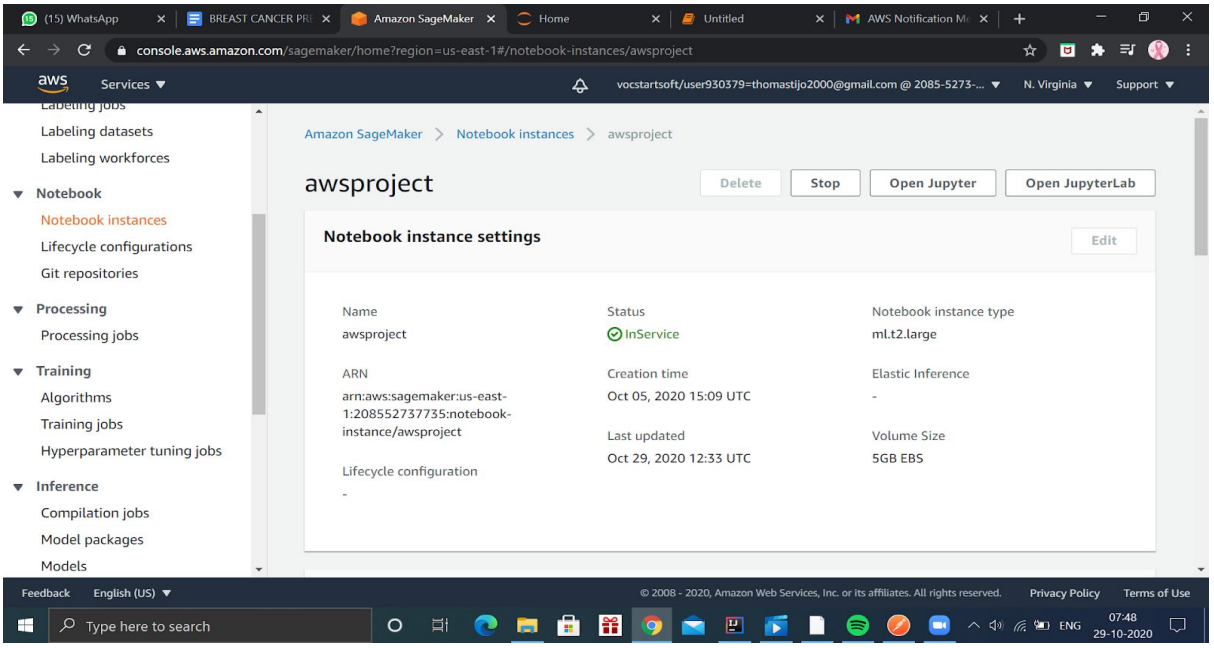
2) https://www.kaggle.com/lbronchal/breast-cancer-dataset-analysis

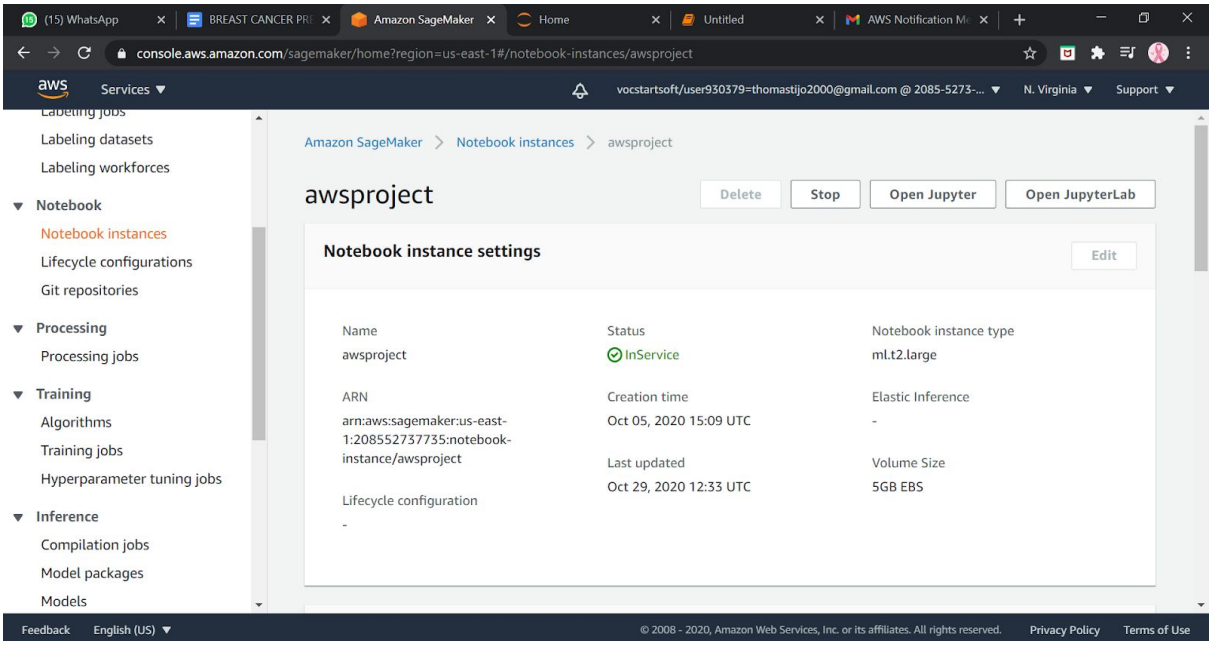
3) https://rstudio-pubs-static.s3.amazonaws.com/344010_1f4d6691092d4544bfbddb092e7223d2.html

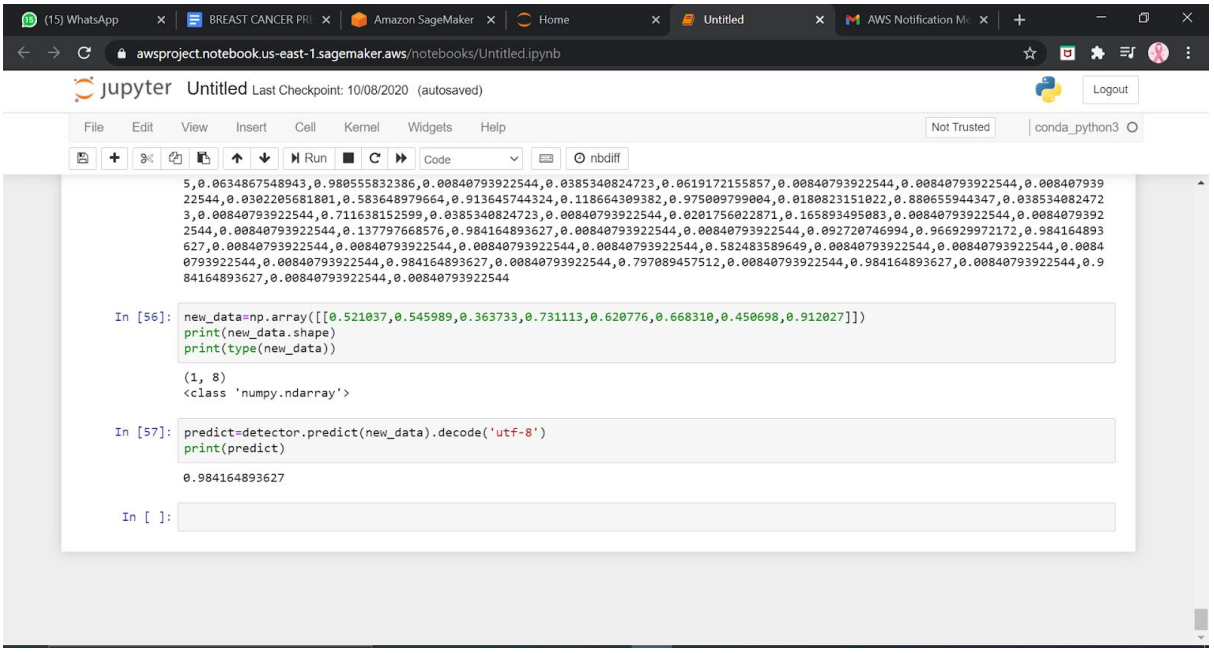
4) https://www.researchgate.net/publication/341508593_BREAST_CANCER_PREDICTION_USING_MACHINE_LEARNING

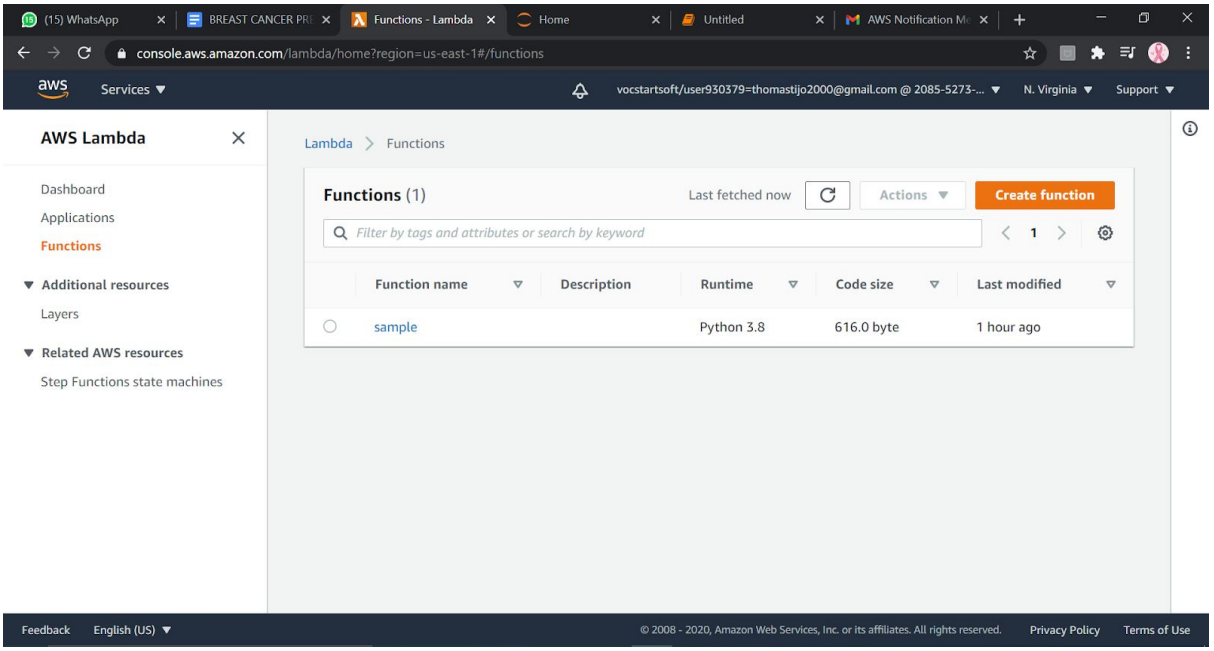
SCREENSHOTS

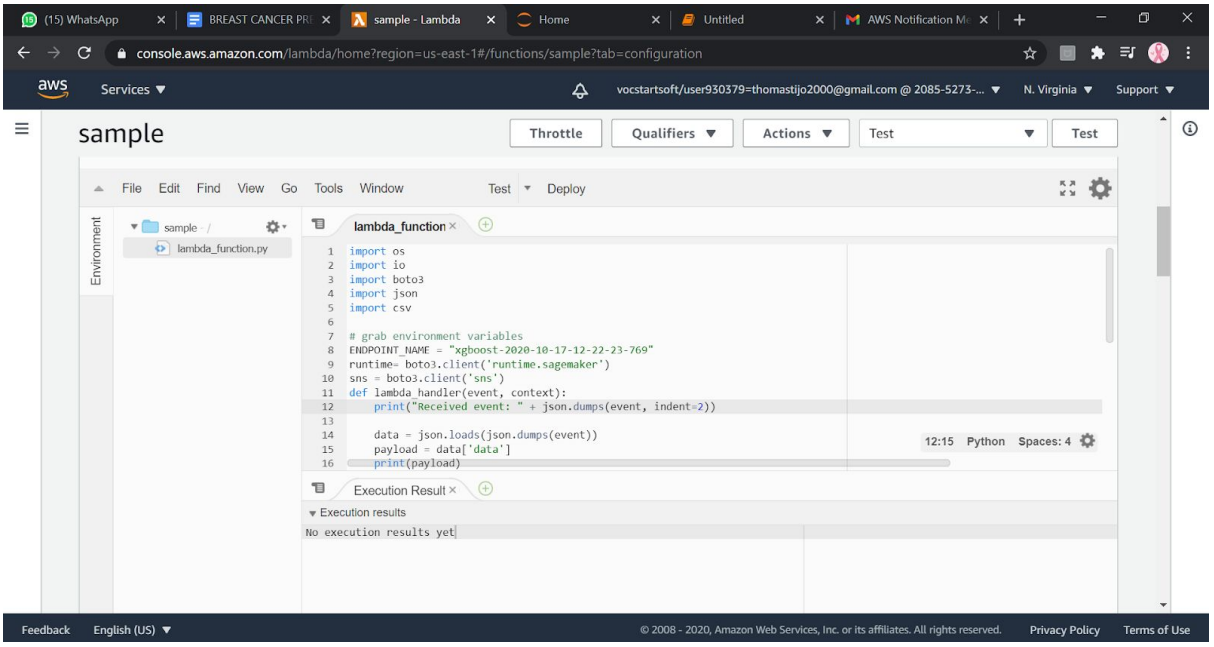

(15) WhatsApp
BREAST CANCER PRI
Amazon SageMaker
Home
Untitled
AWS Notification Ma
console.aws.amazon.com/sagemaker/home?region=us-east-1#/endpoints
aws
Services
vocstartsoft/user930379=thomastijo2000@gmail.com @ 2085-5273-...
N. Virginia
Support
Processing
Processing jobs
Training
Algorithms
Training jobs
Hyperparameter tuning jobs
Inference
Compilation jobs
Model packages
Models
Endpoint configurations
Endpoints
Batch transform jobs
Augmented AI
Human review workflows
Worker task templates
Amazon SageMaker > Endpoints
Endpoints
Update endpoint
Actions
Create endpoint
Search endpoints
1
Name
ARN
Creation time
Status
Last updated
xgboost-2020-10-17-12-22-23-769
arn:aws:sagemaker:us-east-1:208552737735:endpoint/xgboost-2020-10-17-12-22-23-769
Oct 17, 2020 12:26 UTC
InService
Oct 27, 2020 14:07 UTC
Feedback
English (US)
© 2008 - 2020, Amazon Web Services, Inc. or its affiliates. All rights reserved.
Privacy Policy
Terms of Use

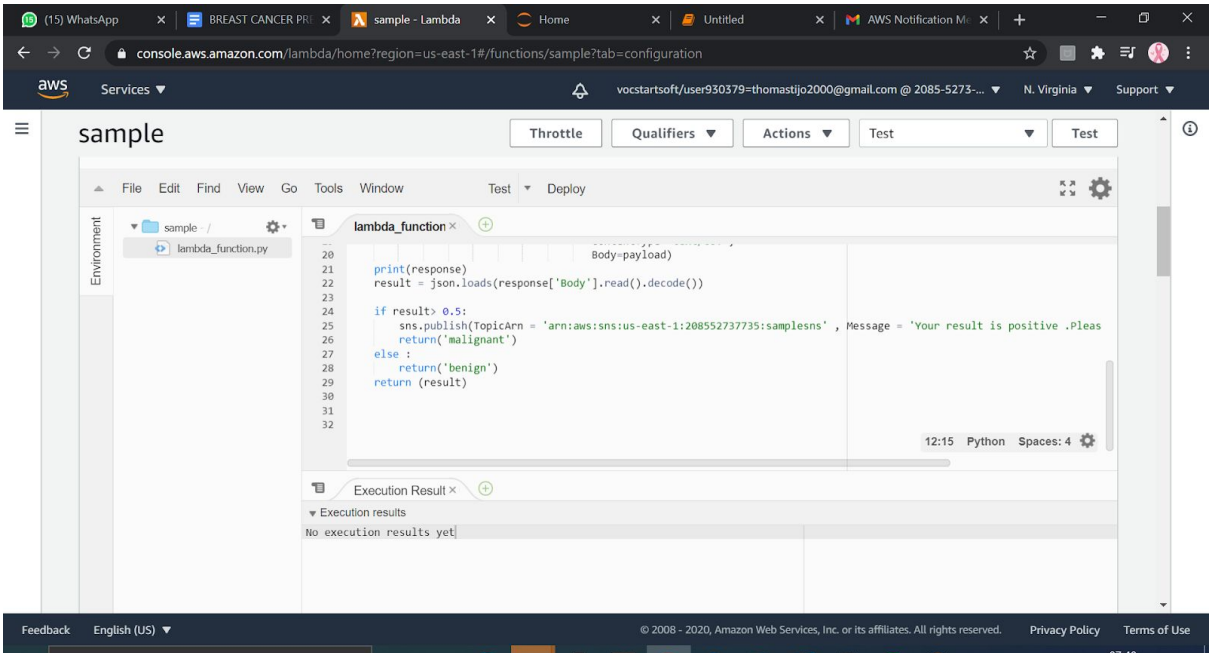

(15) WhatsApp
BREAST CANCER PRI
Amazon SageMaker
Home
Untitled
AWS Notification Ma
console.aws.amazon.com/sagemaker/home?region=us-east-1#/notebook-instances/awsproject
aws
Services
vocstartsoft/user930379=thomastijo2000@gmail.com @ 2085-5273-...
N. Virginia
Support
Labeling datasets
Labeling workforces
Notebook
Notebook instances
Lifecycle configurations
Git repositories
Processing
Processing jobs
Training
Algorithms
Training jobs
Hyperparameter tuning jobs
Inference
Compilation jobs
Model packages
Models
Amazon SageMaker > Notebook instances > awsproject
awsproject
Delete
Stop
Open Jupyter
Open JupyterLab
Notebook instance settings
Edit
Name
awsproject
Status
InService
Notebook instance type
ml.t2.large
ARN
arn:aws:sagemaker:us-east-1:208552737735:notebook-instance/awsproject
Creation time
Oct 05, 2020 15:09 UTC
Elastic Inference
-
Last updated
Oct 29, 2020 12:33 UTC
Volume Size
5GB EBS
Lifecycle configuration
-
Feedback
English (US)
© 2008 - 2020, Amazon Web Services, Inc. or its affiliates. All rights reserved.
Privacy Policy
Terms of Use
Type here to search
ENG
07:48
29-10-2020

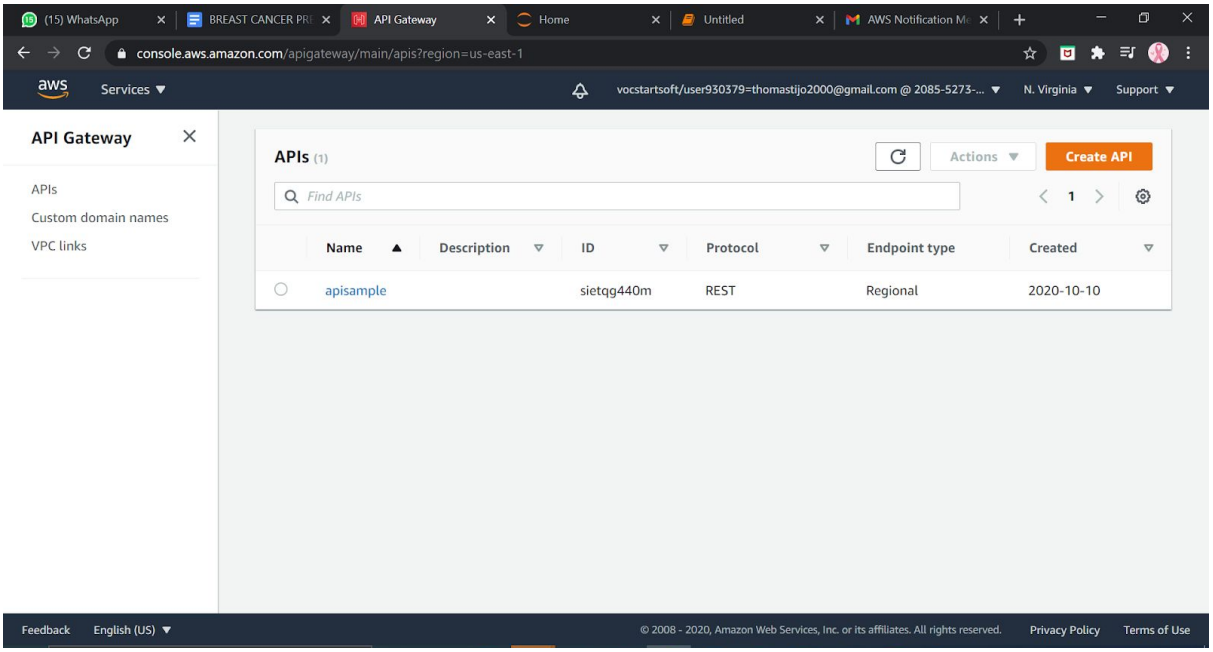

(15) WhatsApp
BREAST CANCER PRE
Amazon SageMaker
Home
Untitled
AWS Notification Me
console.aws.amazon.com/sagemaker/home?region=us-east-1#/notebook-instances/awsproject
aws
Services
vocstartsoft/user930379=thomastijo2000@gmail.com @ 2085-5273-...
N. Virginia
Support
Labeling jobs
Labeling datasets
Labeling workforces
Notebook
Notebook instances
Lifecycle configurations
Git repositories
Processing
Processing jobs
Training
Algorithms
Training jobs
Hyperparameter tuning jobs
Inference
Compilation jobs
Model packages
Models
Amazon SageMaker > Notebook instances > awsproject
awsproject
Delete
Stop
Open Jupyter
Open JupyterLab
Notebook instance settings
Edit
Name
awsproject
Status
InService
Notebook instance type
ml.t2.large
ARN
arn:aws:sagemaker:us-east-1:208552737735:notebook-instance/awsproject
Creation time
Oct 05, 2020 15:09 UTC
Elastic Inference
-
Last updated
Oct 29, 2020 12:33 UTC
Volume Size
5GB EBS
Lifecycle configuration
-
Feedback
English (US)
© 2008 - 2020, Amazon Web Services, Inc. or its affiliates. All rights reserved.
Privacy Policy
Terms of Use

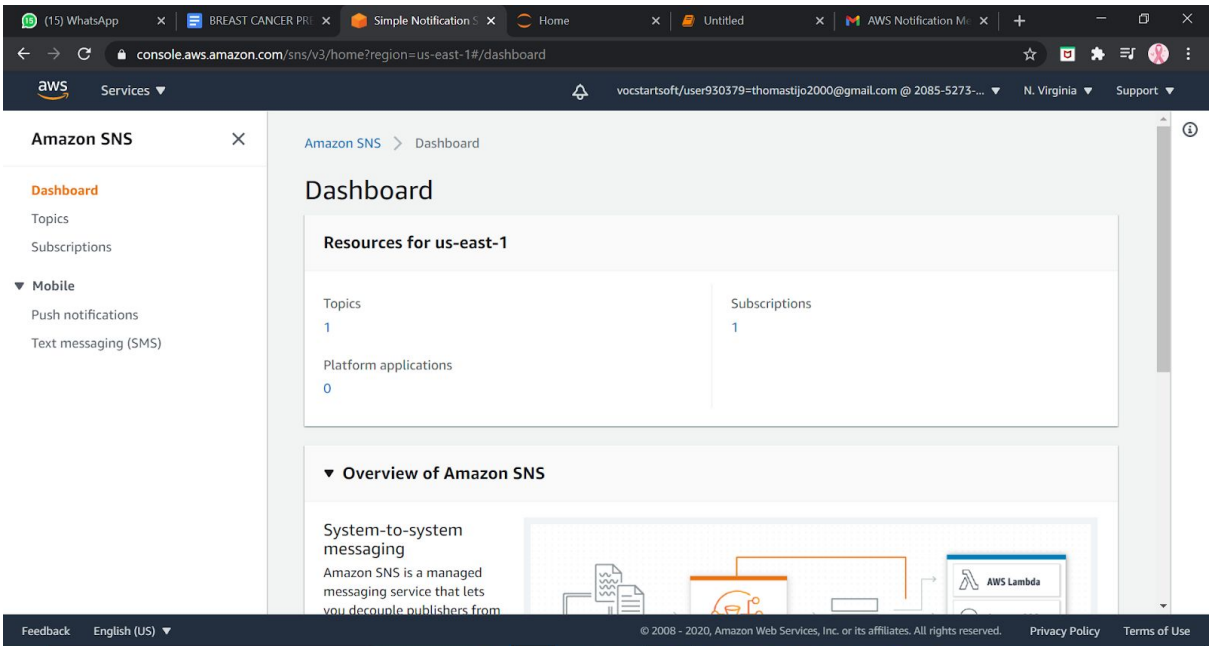

(15) WhatsApp
BREAST CANCER PRE
Amazon SageMaker
Home
Untitled
AWS Notification Me
awsproject.notebook.us-east-1.sagemaker.aws/notebooks/Untitled.ipynb
jupyter Untitled Last Checkpoint: 10/08/2020 (autosaved)
Logout
File Edit View Insert Cell Kernel Widgets Help
Not Trusted
conda_python3
Run
Code
nbdiff
5,0.0634867548943,0.980555832386,0.00840793922544,0.0385340824723,0.0619172155857,0.00840793922544,0.00840793922544,0.008407939
22544,0.030220568180l,0.583648979664,0.913645744324,0.118664309382,0.975009799004,0.0180823151022,0.880655944347,0.038534082472
3,0.00840793922544,0.711638152599,0.0385340824723,0.00840793922544,0.0201756022871,0.165893495083,0.00840793922544,0.0084079392
2544,0.00840793922544,0.137797668576,0.984164893627,0.00840793922544,0.00840793922544,0.092720746994,0.966929972172,0.984164893
627,0.00840793922544,0.00840793922544,0.00840793922544,0.00840793922544,0.582483589649,0.00840793922544,0.00840793922544,0.0084
0793922544,0.00840793922544,0.984164893627,0.00840793922544,0.797089457512,0.00840793922544,0.984164893627,0.00840793922544,0.9
84164893627,0.00840793922544,0.00840793922544
In [56]:
new_data=np.array([[0.521037,0.545989,0.363733,0.731113,0.620776,0.668310,0.450698,0.912027]])
print(new_data.shape)
print(type(new_data))
(1, 8)
<class 'numpy.ndarray'>
In [57]:
predict=detector.predict(new_data).decode('utf-8')
print(predict)
0.984164893627
In [ ]:

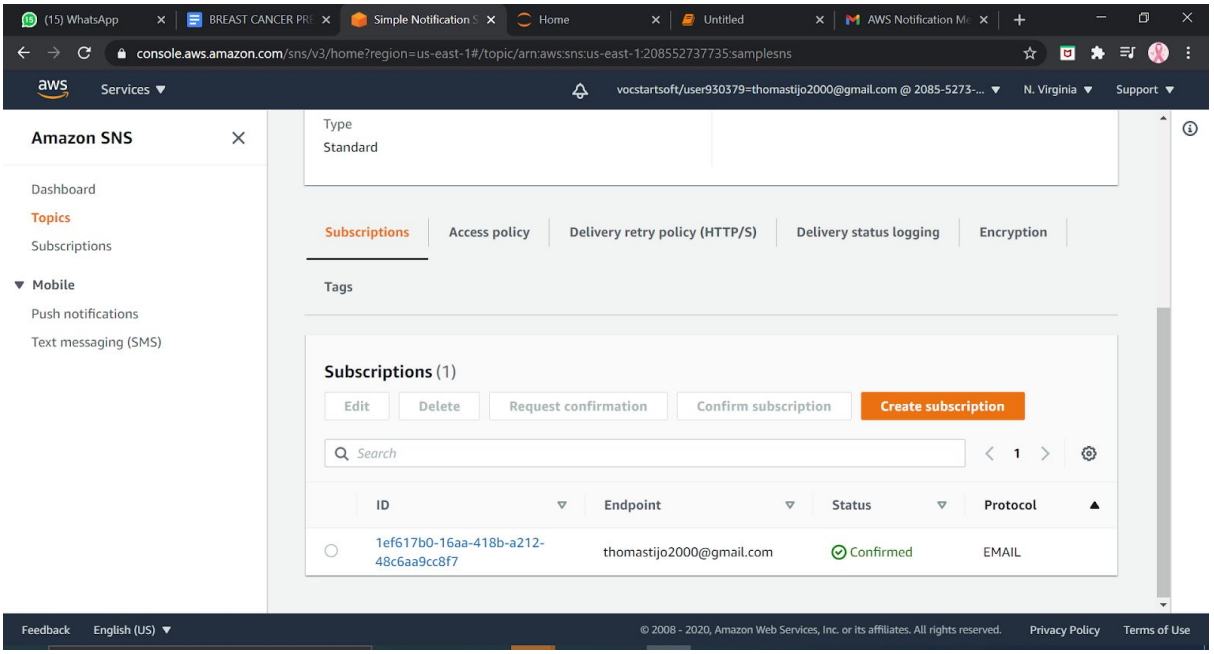

(15) WhatsApp
BREAST CANCER PRE
Functions - Lambda
Home
Untitled
AWS Notification Me
console.aws.amazon.com/lambda/home?region=us-east-1#/functions
aws
Services
vocstartsoft/user930379=thomastijo2000@gmail.com @ 2085-5273-...
N. Virginia
Support
AWS Lambda
Dashboard
Applications
Functions
Additional resources
Layers
Related AWS resources
Step Functions state machines
Lambda > Functions
Functions (1)
Last fetched now
Actions
Create function
Filter by tags and attributes or search by keyword
1
Function name
Description
Runtime
Code size
Last modified
sample
Python 3.8
616.0 byte
1 hour ago
Feedback
English (US)
© 2008 - 2020, Amazon Web Services, Inc. or its affiliates. All rights reserved.
Privacy Policy
Terms of Use

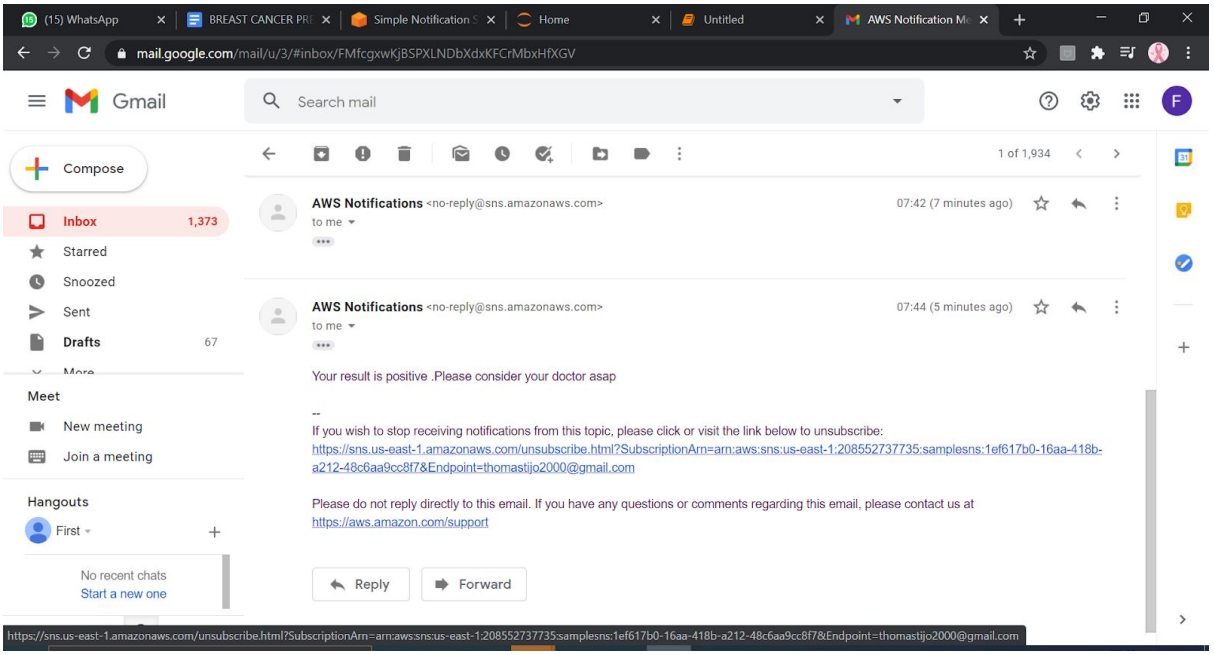

(15) WhatsApp
BREAST CANCER PRE
sample - Lambda
Home
Untitled
AWS Notification Me
console.aws.amazon.com/lambda/home?region=us-east-1#/functions/sample?tab=configuration
aws
Services
vocstartsoft/user930379=thomastijo2000@gmail.com @ 2085-5273-...
N. Virginia
Support
sample
Throttle
Qualifiers
Actions
Test
Test
File Edit Find View Go Tools Window Test Deploy
Environment
sample /
lambda_function.py
lambda_function
import os
import io
import boto3
import json
import csv
# grab environment variables
ENDPOINT_NAME = "xgboost-2020-10-17-12-22-23-769"
runtime= boto3.client('runtime.sagemaker')
sns = boto3.client('sns')
def lambda_handler(event, context):
print("Received event: " + json.dumps(event, indent=2))
data = json.loads(json.dumps(event))
payload = data['data']
print(payload)
12:15 Python Spaces: 4
Execution Result
Execution results
No execution results yet
Feedback
English (US)
© 2008 - 2020, Amazon Web Services, Inc. or its affiliates. All rights reserved.
Privacy Policy
Terms of Use

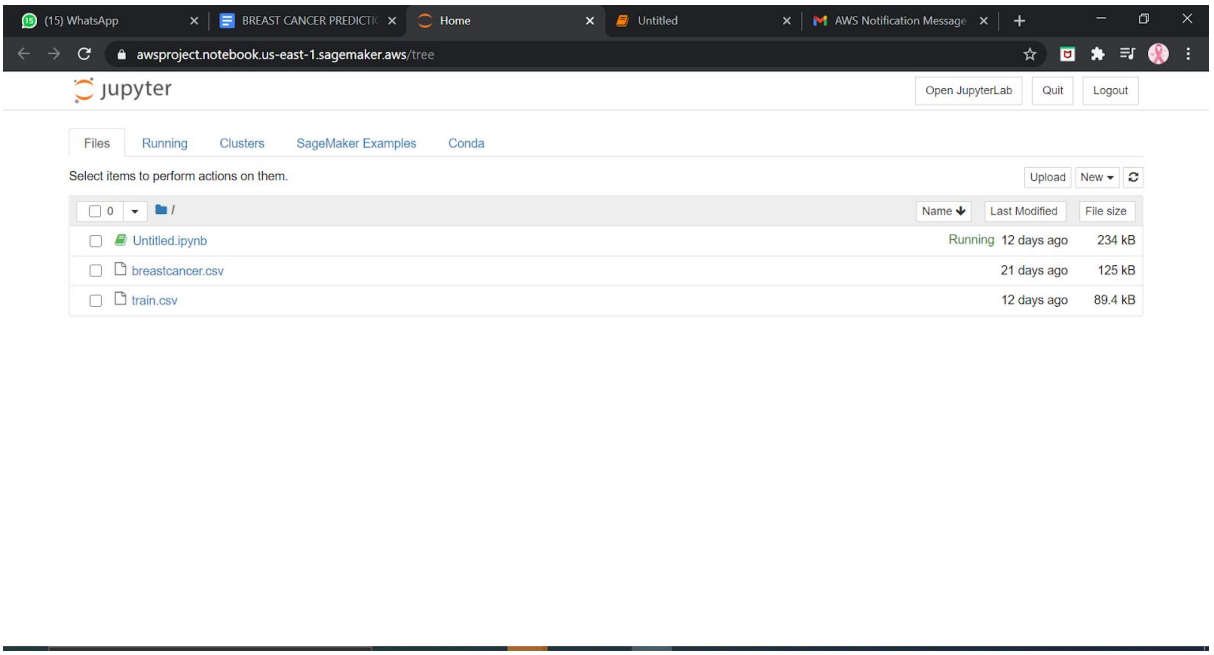

(15) WhatsApp
BREAST CANCER PRE
sample - Lambda
Home
Untitled
AWS Notification Me
console.aws.amazon.com/lambda/home?region=us-east-1#/functions/sample?tab=configuration
Services
vocstartsoft/user930379=thomastijo2000@gmail.com @ 2085-5273-...
N. Virginia
Support
sample
Throttle
Qualifiers
Actions
Test
Test
File Edit Find View Go Tools Window
Test
Deploy
Environment
sample
lambda_function.py
lambda_function
Body=payload)
print(response)
result = json.loads(response['Body'].read().decode())
if result> 0.5:
sns.publish(TopicArn = 'arn:aws:sns:us-east-1:208552737735:samplesns' , Message = 'Your result is positive .Pleas
return('malignant')
else :
return('benign')
return (result)
12:15 Python Spaces: 4
Execution Result
Execution results
No execution results yet
Feedback
English (US)
© 2008 - 2020, Amazon Web Services, Inc. or its affiliates. All rights reserved.
Privacy Policy
Terms of Use

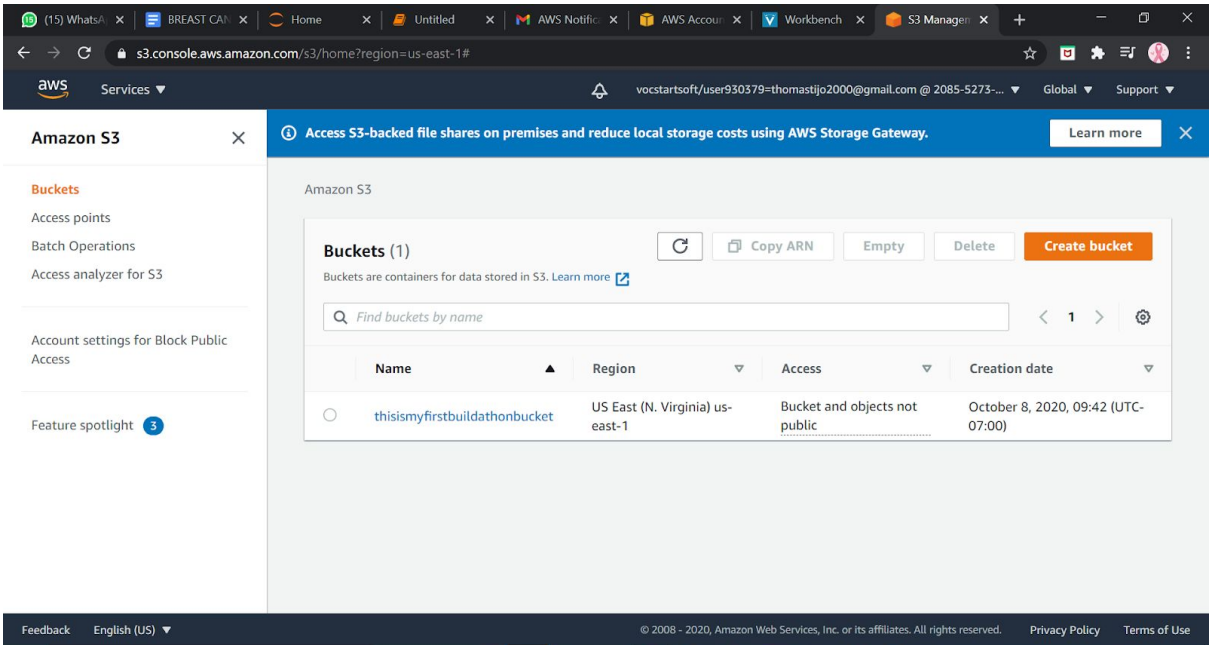

(15) WhatsApp
BREAST CANCER PRE
API Gateway
Home
Untitled
AWS Notification Me
console.aws.amazon.com/apigateway/main/apis?region=us-east-1
Services
vocstartsoft/user930379=thomastijo2000@gmail.com @ 2085-5273-...
N. Virginia
Support
API Gateway
APIs
Custom domain names
VPC links
APIs (1)
Actions
Create API
Find APIs
1
Name
Description
ID
Protocol
Endpoint type
Created
apisample
sietqg440m
REST
Regional
2020-10-10
Feedback
English (US)
© 2008 - 2020, Amazon Web Services, Inc. or its affiliates. All rights reserved.
Privacy Policy
Terms of Use

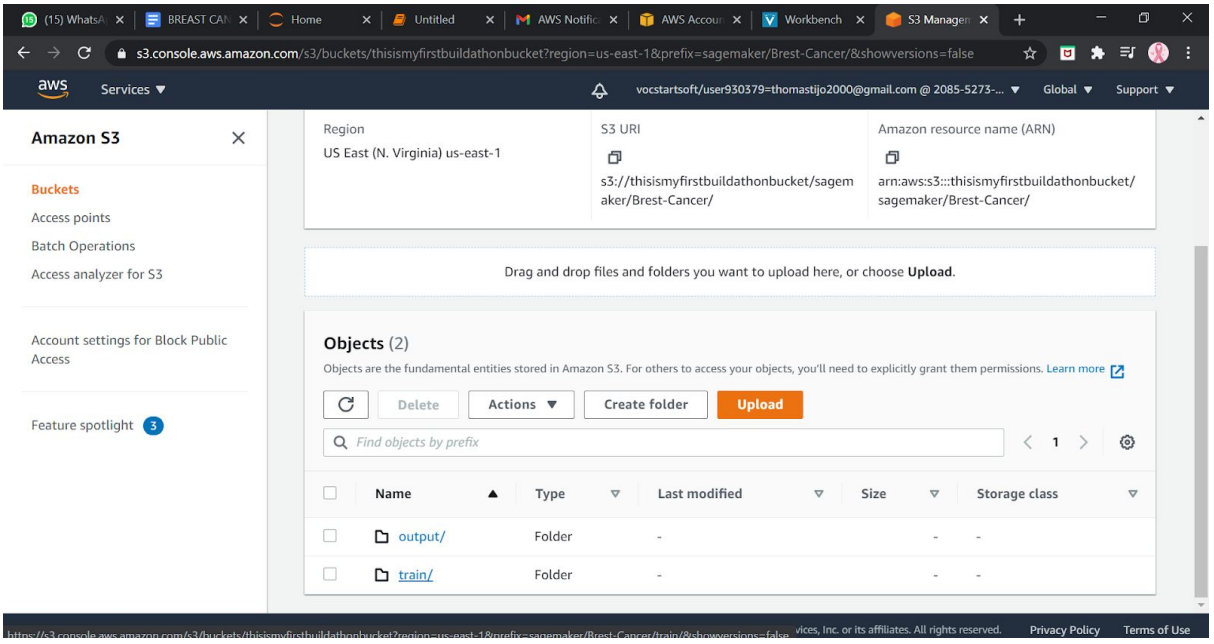

console.aws.amazon.com/sns/v3/home?region=us-east-1#/dashboard
Services
N. Virginia
Support
Amazon SNS
Dashboard
Topics
Subscriptions
Mobile
Push notifications
Text messaging (SMS)
Amazon SNS > Dashboard
Dashboard
Resources for us-east-1
Topics
1
Subscriptions
1
Platform applications
0
Overview of Amazon SNS
System-to-system messaging
Amazon SNS is a managed messaging service that lets
AWS Lambda
Feedback
English (US)
© 2008 - 2020, Amazon Web Services, Inc. or its affiliates. All rights reserved.
Privacy Policy
Terms of Use

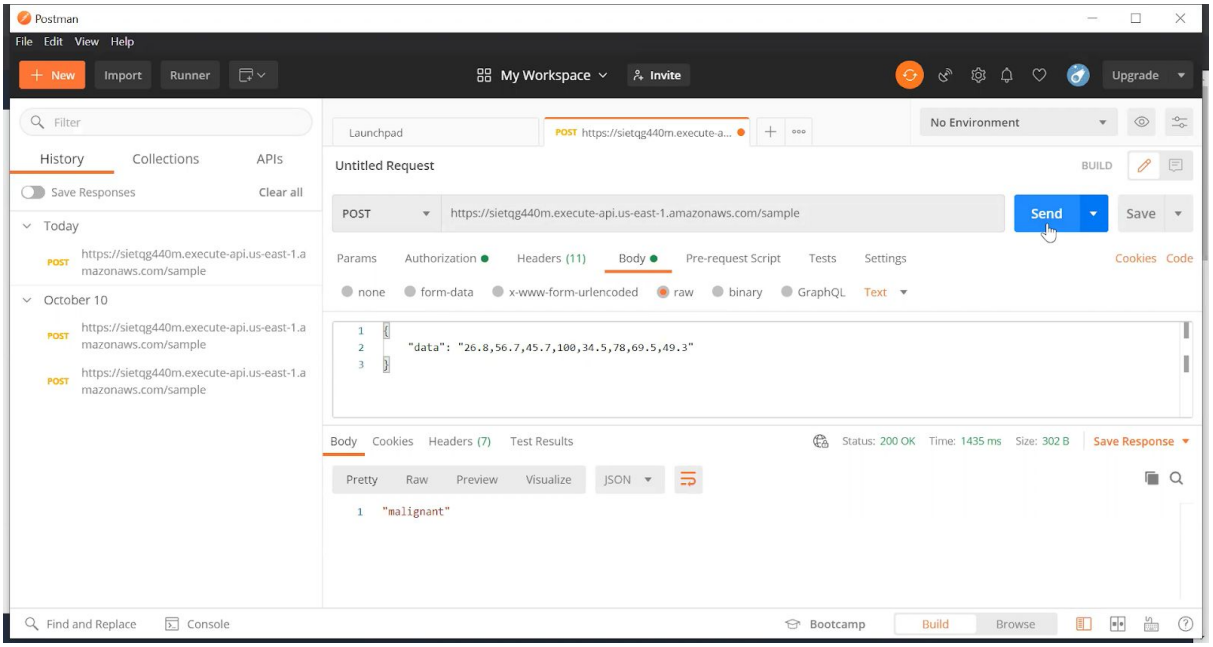

console.aws.amazon.com/sns/v3/home?region=us-east-1#/topic/arn:aws:sns:us-east-1:208552737735:samplesns
Services
N. Virginia
Support
Amazon SNS
Dashboard
Topics
Subscriptions
Mobile
Push notifications
Text messaging (SMS)
Type
Standard
Subscriptions
Access policy
Delivery retry policy (HTTP/S)
Delivery status logging
Encryption
Tags
Subscriptions (1)
Edit
Delete
Request confirmation
Confirm subscription
Create subscription
Search
ID
Endpoint
Status
Protocol
1ef617b0-16aa-418b-a212-48c6aa9cc8f7
thomastijo2000@gmail.com
Confirmed
EMAIL
Feedback
English (US)
© 2008 - 2020, Amazon Web Services, Inc. or its affiliates. All rights reserved.
Privacy Policy
Terms of Use


(15) WhatsApp
BREAST CANCER PRE
Simple Notification S
Home
Untitled
AWS Notification Me
mail.google.com/mail/u/3/#inbox/FMfcgxwKjBSPXLNDbXdxKFCrMbxHfXGV
Gmail
Search mail
Compose
Inbox 1,373
Starred
Snoozed
Sent
Drafts 67
Meet
New meeting
Join a meeting
Hangouts
First
No recent chats
Start a new one
1 of 1,934
AWS Notifications <no-reply@sns.amazonaws.com>
to me
07:42 (7 minutes ago)
AWS Notifications <no-reply@sns.amazonaws.com>
to me
07:44 (5 minutes ago)
Your result is positive .Please consider your doctor asap
--
If you wish to stop receiving notifications from this topic, please click or visit the link below to unsubscribe:
https://sns.us-east-1.amazonaws.com/unsubscribe.html?SubscriptionArn=arn:aws:sns:us-east-1:208552737735:samplesns:1ef617b0-16aa-418b-a212-48c6aa9cc8f7&Endpoint=thomastijo2000@gmail.com
Please do not reply directly to this email. If you have any questions or comments regarding this email, please contact us at
https://aws.amazon.com/support
Reply
Forward
https://sns.us-east-1.amazonaws.com/unsubscribe.html?SubscriptionArn=arn:aws:sns:us-east-1:208552737735:samplesns:1ef617b0-16aa-418b-a212-48c6aa9cc8f7&Endpoint=thomastijo2000@gmail.com


(15) WhatsApp
BREAST CANCER PREDICTI
Home
Untitled
AWS Notification Message
awsproject.notebook.us-east-1.sagemaker.aws/tree
jupyter
Open JupyterLab
Quit
Logout
Files
Running
Clusters
SageMaker Examples
Conda
Select items to perform actions on them.
Upload
New
Name
Last Modified
File size
Untitled.ipynb
Running 12 days ago 234 kB
breastcancer.csv
21 days ago 125 kB
train.csv
12 days ago 89.4 kB


s3.console.aws.amazon.com/s3/home?region=us-east-1#
Services
vocstartsoft/user930379=thomastijo2000@gmail.com @ 2085-5273-...
Global
Support
Amazon S3
Buckets
Access points
Batch Operations
Access analyzer for S3
Account settings for Block Public Access
Feature spotlight 3
Access S3-backed file shares on premises and reduce local storage costs using AWS Storage Gateway.
Learn more
Amazon S3
Buckets (1)
Buckets are containers for data stored in S3. Learn more
Copy ARN
Empty
Delete
Create bucket
Find buckets by name
Name
Region
Access
Creation date
thisismyfirstbuildathonbucket
US East (N. Virginia) us-east-1
Bucket and objects not public
October 8, 2020, 09:42 (UTC-07:00)
Feedback
English (US)
© 2008 - 2020, Amazon Web Services, Inc. or its affiliates. All rights reserved.
Privacy Policy
Terms of Use


s3.console.aws.amazon.com/s3/buckets/thisismyfirstbuildathonbucket?region=us-east-1&prefix=sagemaker/Brest-Cancer/&showversions=false
Services
vocstartsoft/user930379=thomastijo2000@gmail.com @ 2085-5273-...
Global
Support
Amazon S3
Buckets
Access points
Batch Operations
Access analyzer for S3
Account settings for Block Public Access
Feature spotlight 3
Region
US East (N. Virginia) us-east-1
S3 URI
s3://thisismyfirstbuildathonbucket/sagemaker/Brest-Cancer/
Amazon resource name (ARN)
arn:aws:s3:::thisismyfirstbuildathonbucket/sagemaker/Brest-Cancer/
Drag and drop files and folders you want to upload here, or choose Upload.
Objects (2)
Objects are the fundamental entities stored in Amazon S3. For others to access your objects, you'll need to explicitly grant them permissions. Learn more
Delete
Actions
Create folder
Upload
Find objects by prefix
Name
Type
Last modified
Size
Storage class
output/
Folder
train/
Folder
Privacy Policy
Terms of Use



APPENDIX

Source code
https://drive.google.com/drive/folders/11SbTSg9YWS22QGVMbyJOR-c8iuM2iHnv?usp=sharing

TIJO THOMAS