National College of Ireland

Encryption of the Healthcare Data to Protect Against Various Attacks

MSc Research Project
Cloud Computing

Bharat Patil
21204811

School of Computing
National College of Ireland

Supervisor: Sean Heeney

National College of Ireland
Project Submission Sheet
School of Computing

National College of Ireland

Student Name:	Bharat Patil
Student ID:	21204811
Programme:	Cloud Computing
Year:	2023
Module:	MSc Research Project
Supervisor:	Sean Henney
Submission Due Date:	14/08/2023
Project Title:	Encryption of the Healthcare Data to Protect Against Various Attacks
Word Count:	398
Page Count:	7

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.

ALL 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:	Bharat Patil
Date:	14th August 2023

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):	

Encryption of the Healthcare Data to Protect Against Various Attacks

Bharat Patil
21204811

Table of Contents

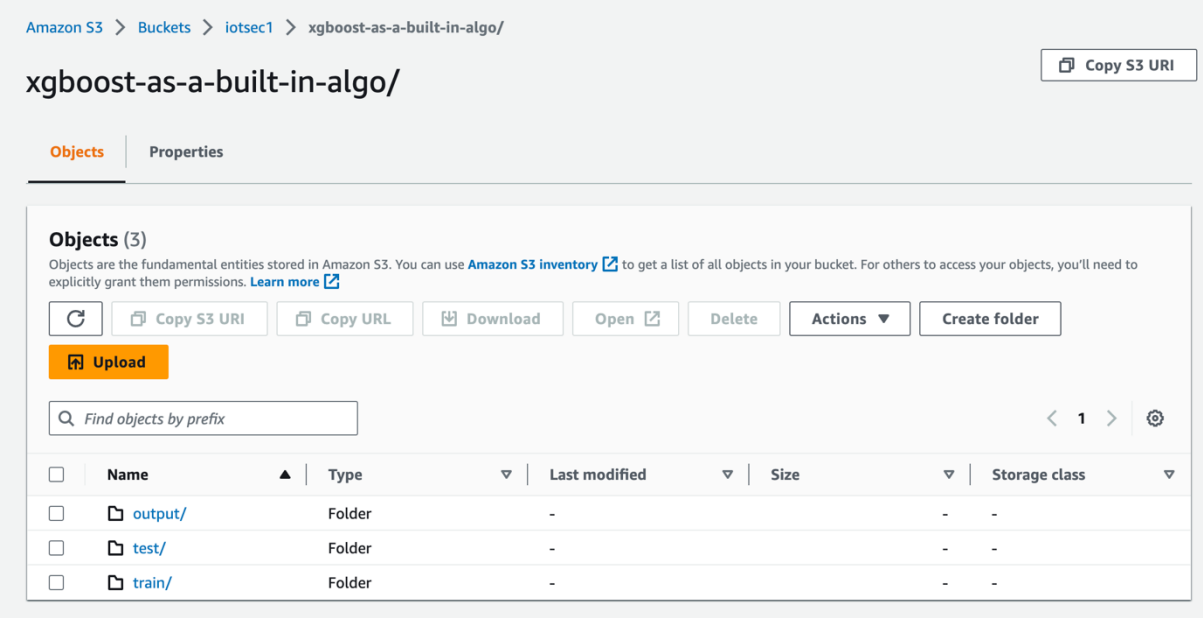
1. Introduction

The healthcare sector is widely recognized as one of the most expansive and rapidly expanding industries on a global scale. The healthcare business has experienced a notable shift from a disease-centered model to a patient-centered one as a result of technological improvements. The increasing need for patient-centric care and value-based healthcare delivery models is motivated by the objectives of enhancing public recognition of healthcare quality and mitigating expenses.

The convergence of the Internet of Things (IoT) with Artificial Intelligence (AI) has presented unprecedented prospects for enhancing clinical and patient services, mitigating expenses, and enhancing community health. The incorporation of Internet of Things (IoT) technology into healthcare software applications has the potential to mitigate some challenges commonly encountered in traditional healthcare systems. The term "medical information system (MIS)" is used to describe a system that is designed to manage healthcare data. This encompasses the operational management of a hospital or a healthcare system that facilitates the development of healthcare policies, along with the implementation of systems responsible for the collection, storage, management, and transmission of a patient's electronic medical record (EMR) Almalawi et al. (2023).

2. Requirements

- Set up Python environment.
- Set up AWS account with access to S3 buckets from the following figure:



3. Installation

Install the required libraries are mentioned in the following figure:

```
[2]: import sagemaker
     import boto3
     from sagemaker.amazon.amazon_estimator import get_image_uri
     from sagemaker.session import s3_input, Session

[3]: import numpy as np # linear algebra
     import pandas as pd # data processing, CSV file I/O (e.g. pd.read_csv)
     import os
     for dirname, _, filenames in os.walk('Attack.csv'):
         for filename in filenames:
             print(os.path.join(dirname, filename))
     import matplotlib.pyplot as plt
     import seaborn as sns
     from sklearn.feature_selection import SelectKBest, f_classif
     from sklearn.feature_selection import chi2
     from sklearn.model_selection import train_test_split
     from sklearn.ensemble import RandomForestClassifier
     from sklearn.tree import DecisionTreeClassifier
     from sklearn.metrics import confusion_matrix, ConfusionMatrixDisplay, accuracy_score, classification_report, accuracy_score, f1_score
     #from sklearn.metrics import accuracy_score,f1_score,confusion_matrix , classification_report
     from sklearn.model_selection import KFold , cross_val_score
     from numpy import mean
     from numpy import std
     from sklearn.model_selection import GridSearchCV
```

1. Python based framework requires some features to be imported. One such example is boto3.

4. Connecting to S3 Bucket

```
[4]: bucket_name = 'iotsec1' # <--- CHANGE THIS VARIABLE TO A UNIQUE NAME FOR YOUR BUCKET
     my_region = boto3.session.Session().region_name # set the region of the instance
     print(my_region)

     us-east-1

[5]: s3 = boto3.resource('s3')
     try:
         if  my_region == 'us-east-1':
             s3.create_bucket(Bucket=bucket_name)
         print('S3 bucket created successfully')
     except Exception as e:
         print('S3 error: ',e)
```

bucket_name = 'iotsec1'
my_region = boto3.session.Session().region_name # set the region of the instance
print(my_region)

s3 = boto3.resource('s3')
try:
  if my_region == 'us-east-1':
    s3.create_bucket(Bucket=bucket_name)
  print('S3 bucket created successfully')
except Exception as e:
  print('S3 error: ',e)

5. Encryption and Decryption

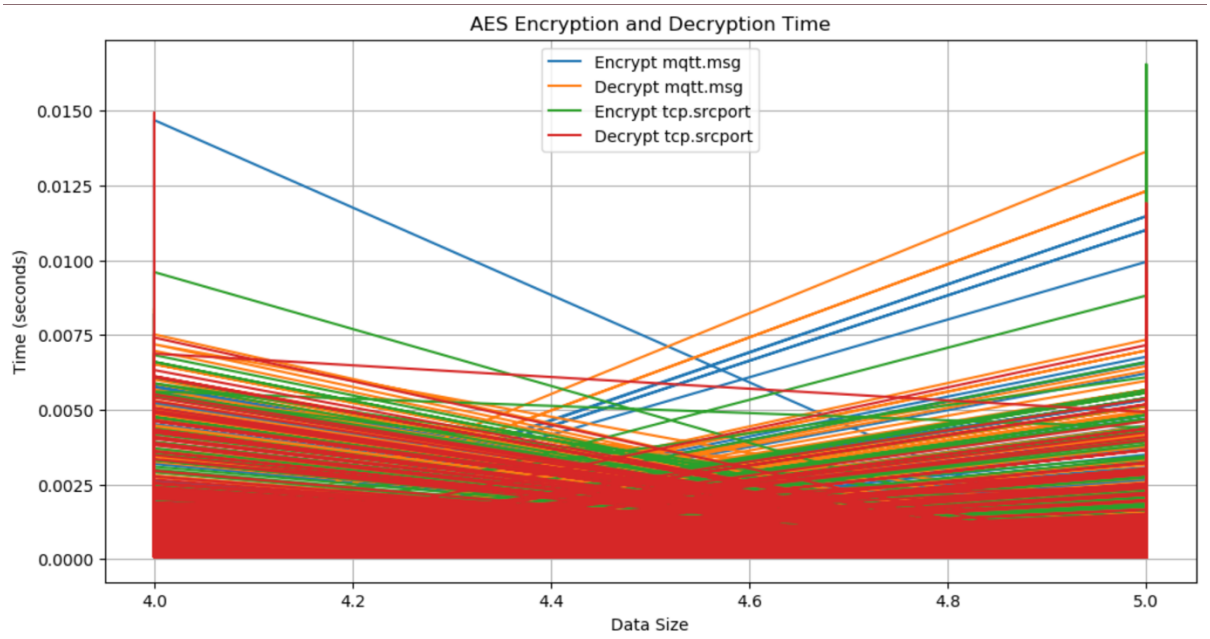
The following code mentioned encrypts and decrypts the data using the variables, 'mqtt.msg', 'tcp.srcport', where in the 80000 columns are processed.

# Encryption
    start_time = time.time()
    encrypted_data = cipher_suite.encrypt(data)
    encryption_column_times.append(time.time() - start_time)

```
# Decryption
start_time = time.time()
decrypted_data = cipher_suite.decrypt(encrypted_data)
decryption_column_times.append(time.time() - start_time)
```

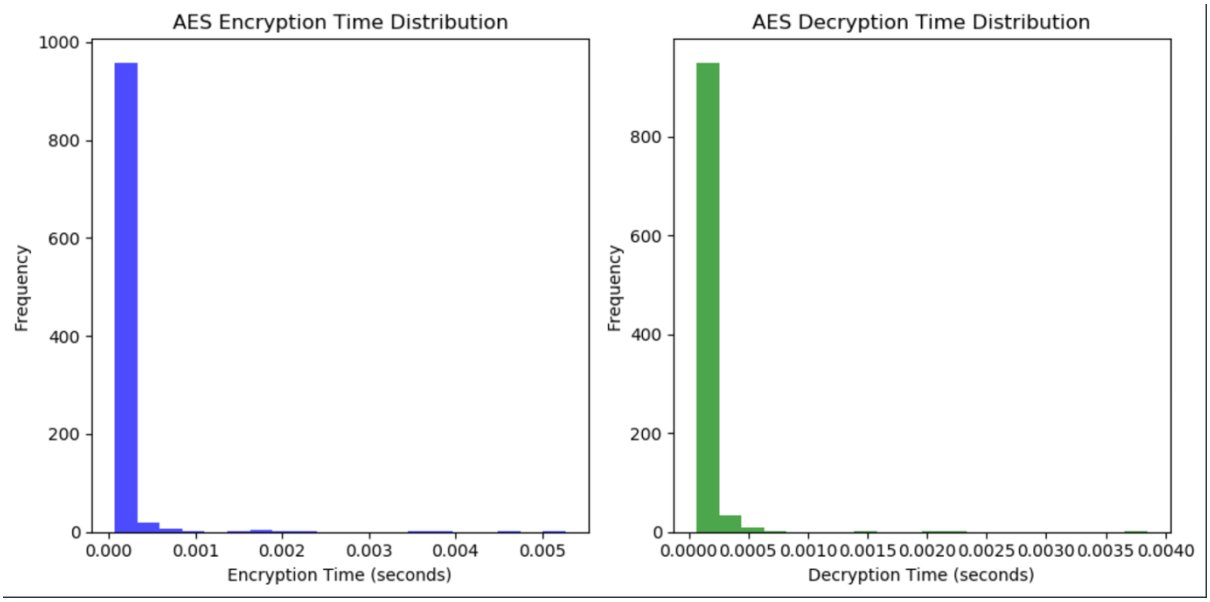
6. Evaluating the Results

As the data is processed, the output graph is plotted containing the Encrypted and Decrypted data.

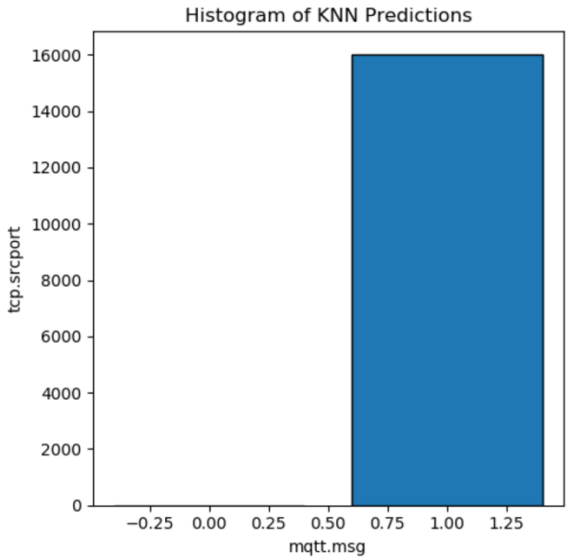


7. Final Results

The histogram results using the AES and KNN has been plotted below. The encryption and decryption time for most of the data is between 0.000 to 0.001.



Using the KNN predictions, we see the graph for message against the source port.



Also while running the code we see that we have saved around 34% of billable seconds.

References

Almalawi, Abdulmohsen & Khan, Asif & Alsolami, Fawaz & Abushark, Yoosef & Alfakeeh, Ahmed. (2023). Managing Security of Healthcare Data for a Modern Healthcare System. Sensors. 23. 3612. 10.3390/s23073612.