University of New Haven
TAGLIATELA COLLEGE OF ENGINEERING
Electrical & Computer Engineering and Computer Science

Remote Patient Care: Health Monitoring and Adherence with AWS and IoT.

Submitted by (Team #12)
Mounika Sri Sai Apurva Nemalipuri
Kiran Varma Ceemala Shankar
Akhil Pingali
Dinesh Kothandaraman

Submitted for DSCI-6007-03 (Spring 2024)
DR. ARDIANA SULA

Submitted on: 04/24/2024

CONTENTS

Executive Summary

This project utilizes AWS IoT Analytics, AWS S3, and AWS Glue to create a robust remote patient monitoring system. By leveraging IoT devices for real-time health data capture, this system aims to enhance patient adherence to prescribed treatments and enable proactive healthcare interventions.

Abstract

The "Remote Patient Care: Health Monitoring and Adherence with AWS and IoT" project is poised at the nexus of technological innovation and healthcare, offering a comprehensive solution that leverages the Internet of Things (IoT) and Amazon Web Services (AWS) to transform patient monitoring and care management. By integrating advanced IoT devices with AWS's robust cloud infrastructure, the project enables the real-time collection and analysis of critical health data, empowering healthcare providers to deliver personalized, proactive care. This initiative stands as a testament to the potential of digital health technologies to not only enhance patient outcomes but also to extend the reach of quality healthcare to individuals regardless of their location.

This initiative implements an advanced IoT-based health monitoring system that integrates IoT devices with AWS cloud services, focusing on AWS IoT Core, AWS S3, and AWS Glue to ensure real-time data collection and analysis, improving patient outcomes and treatment adherence.

Highlights of Project

This Project encapsulates groundbreaking advancements in telehealth and personalized patient management. At the forefront, the integration of real-time health data collection through IoT devices presents a transformative approach to monitoring vital patient metrics outside traditional healthcare settings. The robust AWS-powered backend architecture creates a seamless pipeline for data ingestion, storage, processing, and querying, which underpins a proactive healthcare model. A suite of analytics tools allows for the immediate interpretation of complex datasets, empowering healthcare providers with the ability to make swift, data-

driven decisions. This innovative system not only promises to elevate the standard of patient care but also represents a significant step towards the future of healthcare, where continuous monitoring and timely interventions can lead to better health outcomes and potentially save lives.

Introductory Section

This project implements an advanced IoT-based health monitoring system that integrates IoT devices with AWS cloud services, focusing on AWS IoT Core, AWS S3, and AWS Glue to ensure real-time data collection and analysis, improving patient outcomes and treatment adherence. This presents significant opportunities for enhancing patient care. This project tackles challenges such as data security, real-time data processing, and patient engagement by using a combination of IoT devices and essential AWS services.

Review of available research

The field of remote patient monitoring through IoT devices is well-documented across several research studies. For instance, Uddin and Koo (2024) provide a comprehensive review of biosensors integrated with multi-hop IoT systems, highlighting the importance of real-time data transmission and cloud connectivity in remote patient monitoring (MDPI).

Research by Alasmary (2024) introduces the ScalableDigitalHealth (SDH) framework, which emphasizes the integration of smart digital health solutions with latency-aware edge computing autoscaling to enhance remote patient monitoring. This framework particularly underscores the use of IoT for real-time tracking of vital health metrics and the importance of scalability in healthcare applications (MDPI).

Other studies, such as those referenced in the Journal of Big Data, explore various IoT-based health monitoring systems. These range from wearable sensors for autistic patients (SpringerOpen), to cloud-based frameworks that ensure the secure transmission of ECG and other health data (SpringerOpen), to the utilization of smart architecture for in-home healthcare (SpringerOpen), and intensive care unit monitoring frameworks (SpringerOpen). Each study

contributes unique insights into the effective deployment of IoT technologies for enhanced patient care and the pivotal role of IoT in transforming the healthcare industry.

Heart Rate and Blood Pressure Monitoring: Data is captured through AWS IoT Core from wearable devices, enabling continuous monitoring of these vital signs.

Activity Levels: IoT sensors provide insights into the patient's physical activity, crucial for managing and monitoring physical health conditions.

Patient Feedback Visualization:

Adherence Reports: IoT devices connected to medication dispensers report on patient adherence, providing key data on how patients comply with their treatment regimens.

Symptom Tracking: Patients report symptoms through a user interface that logs and tracks these inputs over time, providing a timeline of patient health that can be used for treatment adjustments.

Methodology

Data Sources: Data is collected from IOT device and stored in S3 bucket. Patient feedback is collected through web application and stored in S3 Bucket. Patient Past Data is stored in Database and retrieved using Ec2 instance and AWS Athena.

Overview of the Data Pipeline

1. **Streaming Collection Layer:**

AWS IoT: This is where data collection begins. IoT devices connected to AWS IoT collect real-time health data from patients.
Amazon Kinesis Data Firehose: Captured data is streamed in real-time through Amazon Kinesis Data Firehose, which reliably loads massive streams of data into AWS storage services.

2. **Batch Processing Layer:**

Amazon S3: The data streamed by Kinesis Data Firehose is stored in Amazon S3 buckets. Amazon S3 is a highly durable storage service that allows you to store and retrieve any amount of data at any time.

AWS Glue: A managed ETL (extract, transform, load) service that prepares and transforms the data. AWS Glue cleans, normalizes, and structures raw data into a format that is more conducive to analysis.

3. **Stream Processing Layer:**

Amazon Kinesis Data Analytics: Processes the streaming data in real-time, enabling immediate analysis and insight into the health data as it arrives.

Amazon Kinesis Data Firehose (again): The processed stream can then be directed back through Kinesis Data Firehose for further actions, such as additional storage or real-time alerting.

4. **Serving Layer:**

Amazon S3 for pre-processed views: Once processed, the data is again stored in S3, but this time in a pre-processed form that makes it quicker and easier to query.

Amazon S3 for filtered data: Another S3 bucket is used to store filtered subsets of the data, which are of particular interest or relevance for specific queries or analysis.

5. **Data Queries:**

Amazon Athena: An interactive query service that makes it easy to analyze data in Amazon S3 using standard SQL queries. Athena is serverless, so there is no infrastructure to manage, and you pay only for the queries you run.

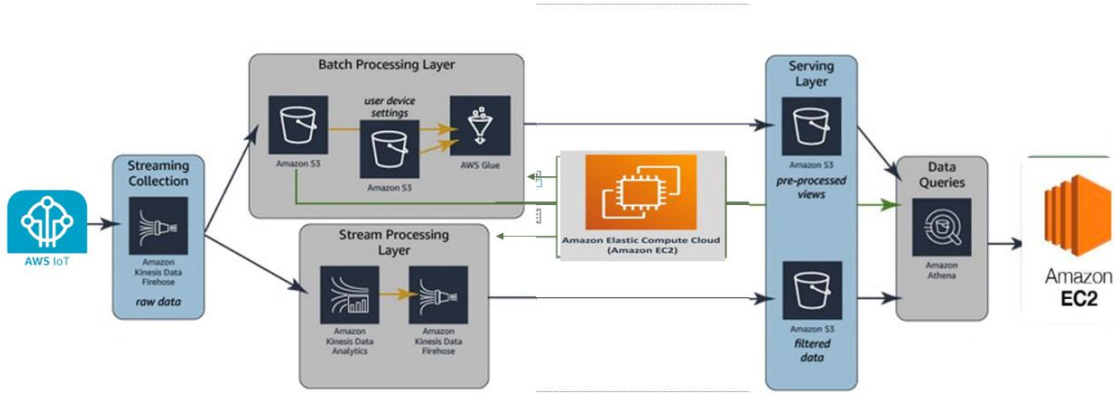
Amazon EC2: Users can run queries from Amazon EC2 instances, which provides scalable compute capacity in the cloud and can be used to run additional analyses or applications based on the query results.

AWS Services and Their Roles in the Pipeline

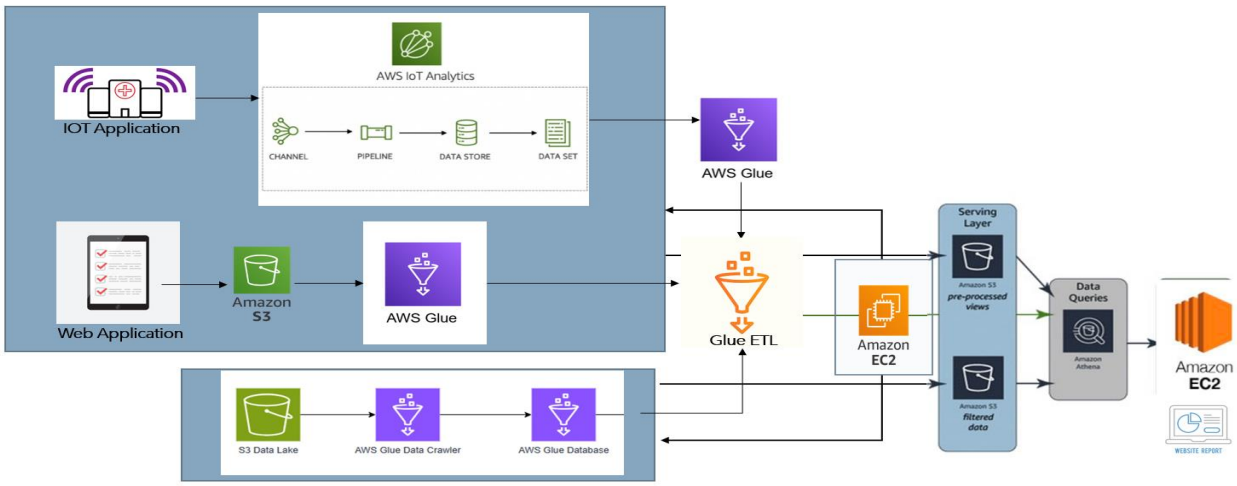
AWS IoT acts as the gateway for IoT devices to securely interact with cloud applications and other devices. Amazon Kinesis Data Firehose is a fully managed service used to load streaming data into data lakes, data stores, and analytics services. Amazon S3 is an object storage service offering industry-leading scalability, data availability, security, and performance. AWS Glue is a serverless data integration service that makes it easy to discover, prepare, and combine data for analytics, machine learning, and application development.

Amazon Kinesis Data Analytics is the easiest way to transform and analyze streaming data in real time with Apache Flink. Amazon Athena is an interactive query service that makes it easy to analyze data directly in Amazon S3 using standard SQL. Amazon EC2 provides resizable compute capacity in the cloud and allows you to scale up or down as your computing requirements change.

Initial Data Pipeline



Final Data Pipeline



CRISP-DM Methodology

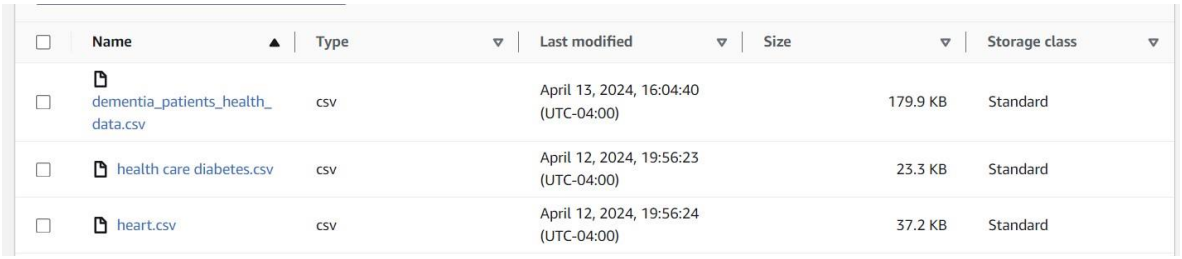
Business Understanding:

Define project objectives and goals, such as improving patient outcomes, enhancing treatment adherence, and extending healthcare accessibility. Identify stakeholders, including healthcare providers, patients, caregivers, and regulatory bodies. Understand the current healthcare challenges and the potential benefits of IoT and AWS integration.

Data Understanding:

Identify the types of data available, such as patient health data from IoT devices, historical patient records, and feedback on treatment adherence. Assess data quality, completeness, and relevance for analysis. Understand data privacy and security requirements, ensuring compliance with healthcare regulations.

S3 Buckets for ML model



Name	Type	Last modified	Size	Storage class
dementia_patients_health_data.csv	csv	April 13, 2024, 16:04:40 (UTC-04:00)	179.9 KB	Standard
health care diabetes.csv	csv	April 12, 2024, 19:56:23 (UTC-04:00)	23.3 KB	Standard
heart.csv	csv	April 12, 2024, 19:56:24 (UTC-04:00)	37.2 KB	Standard

Data Preparation:

Ingest raw data from IoT devices into AWS IoT Core for real-time streaming. Store historical patient data in AWS S3 buckets for data warehousing and analysis. Use AWS Glue for data transformation, cleaning, and integration to create a unified data repository.

Modeling:

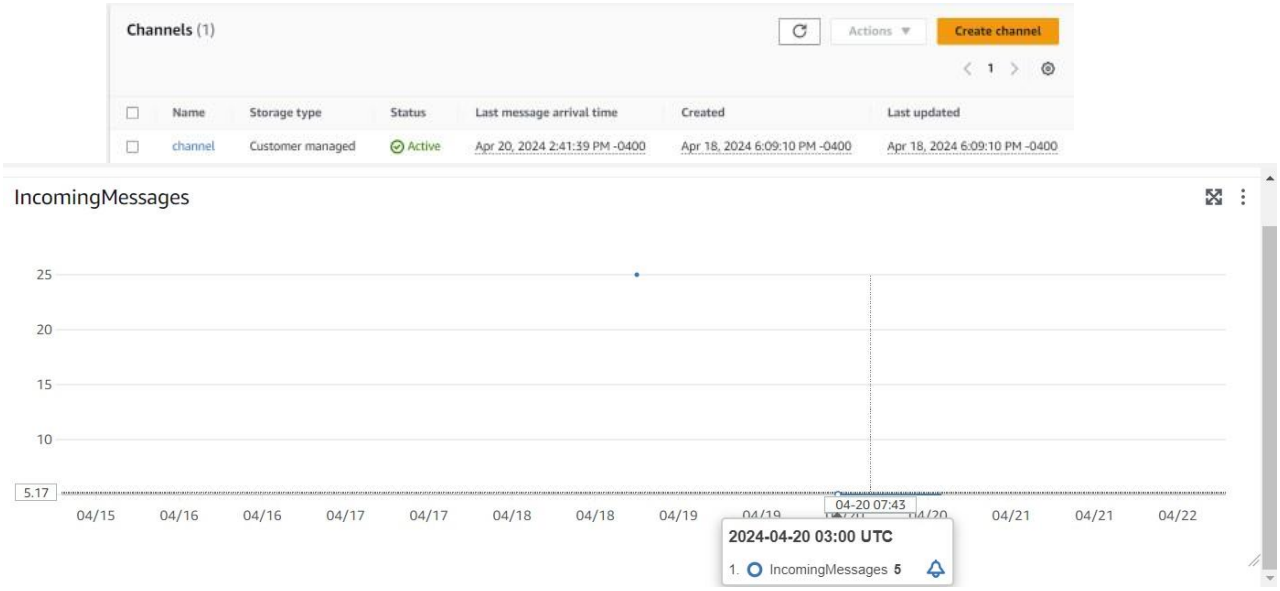
Develop machine learning models on AWS EC2 instances to analyze patient data and predict health outcomes. Utilize advanced analytics techniques to identify patterns, trends, and anomalies in patient health metrics. Implement algorithms for personalized treatment recommendations and adherence monitoring.

Glue for Data Cleaning

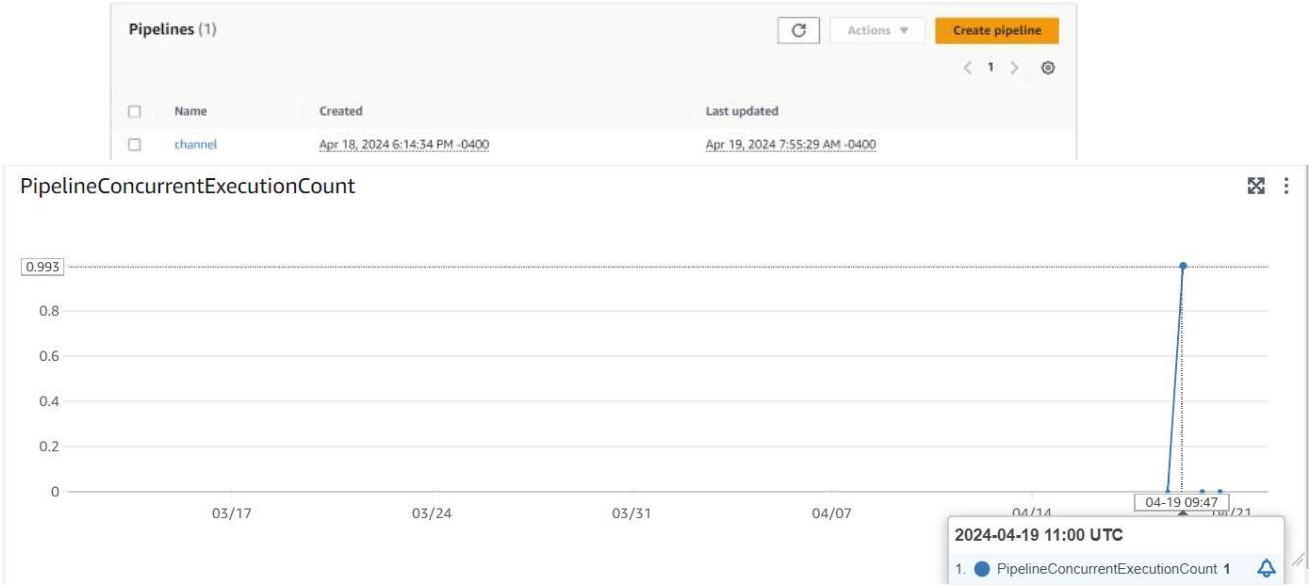
Script Info

```
import sys
from awsglue.transforms import *
from awsglue.utils import getResolvedOptions
from pyspark.context import SparkContext
from awsglue.context import GlueContext
from awsglue.job import Job

args = getResolvedOptions(sys.argv, ['JOB_NAME'])
sc = SparkContext()
glueContext = GlueContext(sc)
spark = glueContext.spark_session
job = Job(glueContext)
job.init(args['JOB_NAME'], args)

# Script generated for node patient_data
patient_data_node1713529687845 = glueContext.create_dynamic_frame.from_catalog(database="project", table_name="patient_data_csv",
    transformation_ctx="patient_data_node1713529687845")

# Script generated for node patient_report
patient_report_node1713529689005 = glueContext.create_dynamic_frame.from_catalog(database="project", table_name="patient_report_csv",
    transformation_ctx="patient_report_node1713529689005")
```

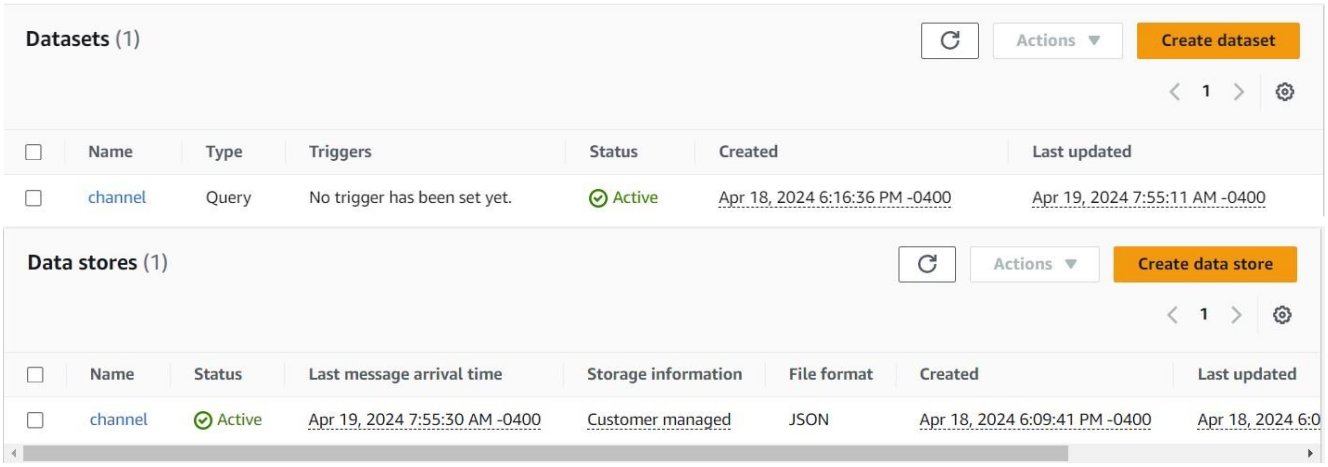
IOT Channel



IOT Pipeline

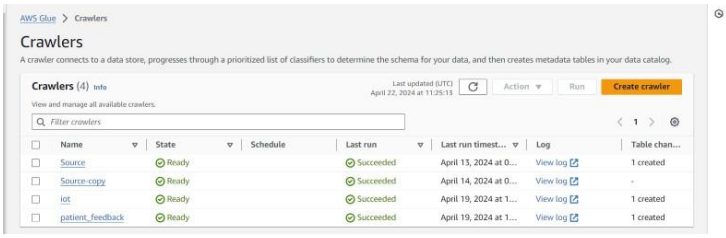


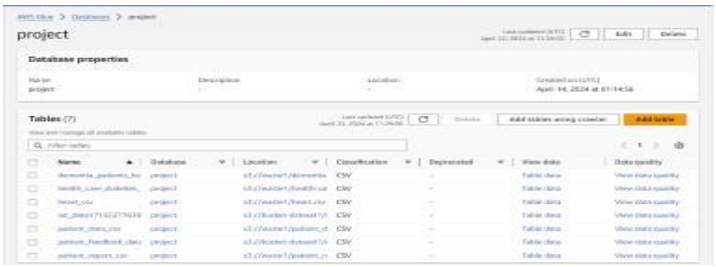
IOT Datasets & Data Stores



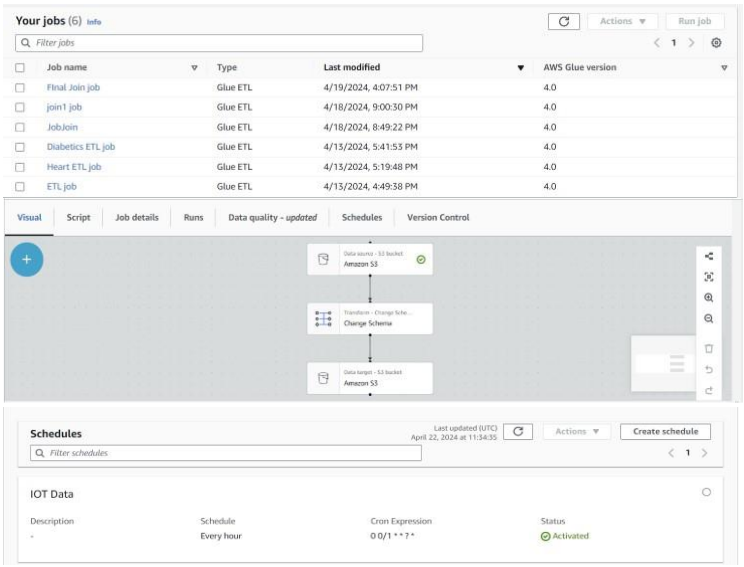
GLUE

This is our AWS Glue Crawler job used for creating Databases.

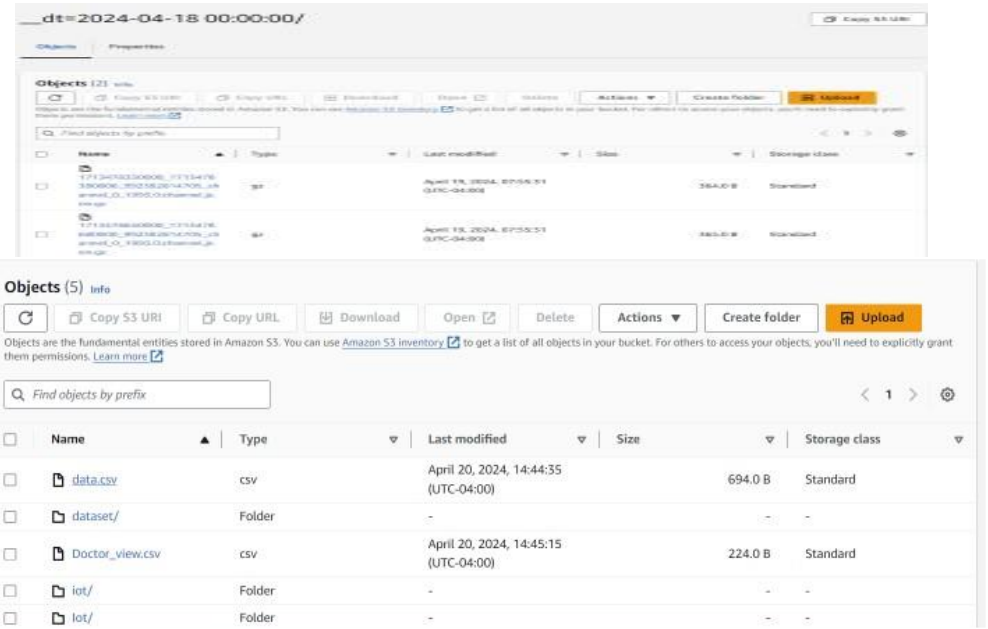




This is our scheduled AWS Glue ETL job.



Processed IOT data on Hourly Schedule by Glue



Evaluation:
Evaluate the performance of machine learning models using metrics such as accuracy, precision, recall, and F1 score. Assess the impact of the IoT-based health monitoring system on patient outcomes, treatment adherence, and healthcare costs. Gather feedback from healthcare providers, patients, and caregivers to validate the effectiveness of the system.

Deployment:
Deploy the finalized machine learning models and analytics solutions on AWS infrastructure for production use. Integrate the system with existing healthcare workflows and electronic health record (EHR) systems. Conduct training sessions and provide documentation for healthcare staff on using the system effectively.

Iterative Improvement:
Monitor system performance and user feedback continuously to identify areas for improvement. Iterate on machine learning models, data processing pipelines, and user interfaces to enhance functionality and usability. Stay updated with advancements in IoT, AWS services, and healthcare technologies to incorporate new features and best practices.

Data Attributes Description

1. **Heart Rate:** Measures the number of heart beats per minute. It is a crucial indicator of cardiac health and can signal potential heart conditions when values are abnormally high or low.

2. **RestECG (Resting Electrocardiographic results):** Shows the heart's electrical activity while at rest. It can help detect heart problems like arrhythmias and heart blockages.

3. **Thalach (Maximum Heart Rate Achieved):** Represents the highest heart rate achieved during stress testing. It is used to assess the cardiac function and to gauge the heart's ability to respond to stress.

4. **Oldpeak (ST depression induced by exercise relative to rest):** Measures the amount of ST-segment depression in the ECG post-exercise. It is an important marker for diagnosing ischemic heart disease.

5. **Slope (the slope of the peak exercise ST segment):** This attribute refers to the rate of change in the ST segment on an ECG during exercise. It helps in assessing the severity and prognosis of coronary artery disease.

6. **Thal (Thallium Stress Test):** Reflects how well the heart distributes blood during periods of stress compared to rest. Abnormalities can indicate coronary artery blockages.

7. **Trestbps (Resting Blood Pressure):** The blood pressure while at rest. High resting blood pressure can indicate cardiovascular risks and other health issues.

8. **Blood Pressure:** Encompasses both systolic and diastolic blood pressure measurements, indicating the force of blood against artery walls. It is crucial for diagnosing hypertension.

9. **High BP:** A binary attribute indicating whether a patient has hypertension, which is a major risk factor for heart disease and stroke.

10. **FBS (Fasting Blood Sugar):** Measures blood sugar levels after a period of fasting. High levels can indicate diabetes, which is a risk factor for heart disease.

11. **Body Temperature:** Reflects the body's thermal regulation. Persistent abnormal temperatures can signal infection or other health issues.

12. **Physical Activity:** Tracks the level of physical activity, which is crucial for overall health and particularly for cardiovascular and metabolic health.

13. **Weight:** Body weight can affect various health aspects, including heart health, and is used to calculate body mass index (BMI).

14. **Sleep Quality:** Poor sleep quality is associated with various health problems, including cardiovascular disease and decreased immune function.

15. **Exang (Exercise-induced angina):** Indicates whether angina (chest pain or discomfort) was induced by exercise, reflecting coronary artery condition. This is critical for diagnosing and managing ischemic heart disease.

16. **Patient ID:** A unique identifier for each patient, ensuring that the data collected is accurately associated with the correct patient file.

Importance of the above attributes in Healthcare.

The above attributes are critical for a remote patient monitoring system as they collectively provide a comprehensive picture of a patient's health. They allow healthcare providers to:

Detect early signs of potential health issues: Attributes like heart rate, blood pressure, and oldpeak can help in early detection of cardiovascular issues.

Monitor chronic conditions: Regularly tracking values like FBS and High BP helps in managing diabetes and hypertension.

Adjust treatments based on real-time data: Continuous monitoring and data collection provide insights that can lead to adjustments in medication or lifestyle recommendations.

Improve patient outcomes: By analyzing trends and changes in these attributes, healthcare providers can intervene proactively, potentially preventing severe health events.

Each of these attributes contributes to a detailed health profile that can be utilized not only for ongoing monitoring and immediate interventions but also for long-term health planning and

management. This comprehensive data collection is pivotal for enhancing the efficacy of remote patient care systems.

Patient feedback

Patient Feedback Data Attributes Description:

1. **Username:** A unique identifier used to log into the system, maintaining patient privacy and security while ensuring that data collected is attributed to the correct patient record.
2. **Patient Name:** The full name of the patient, which helps personalize care and ensures that healthcare providers can accurately associate health records with individual patients.
3. **Weight:** Regular monitoring of the patient's weight is critical for assessing health trends, managing conditions like obesity, diabetes, or heart disease, and evaluating the effectiveness of prescribed treatment plans.
4. **Smoking Status**: Indicates whether the patient currently smokes, has quit, or has never smoked. Smoking status is a vital factor in assessing risk for numerous conditions, including cardiovascular diseases and lung disorders.
5. **Depression Status**: Mental health is integral to overall health. Monitoring depression helps in understanding the patient's mental well-being and can influence the management of various physical health conditions, as mental stress can exacerbate issues like hypertension and heart disease.
6. **Nutrition - Diet:** Information about the patient's dietary habits provides insights into their nutritional intake, which is crucial for managing conditions such as diabetes, obesity, and heart disease. A well-balanced diet is often a core component of treatment plans.
7. **CP (Chest Pain):** This attribute records the occurrence and severity of chest pain, which can indicate underlying conditions such as angina or myocardial infarction. Understanding the patterns of chest pain can help in early detection and prevention of more severe cardiac events.

8. **Sleep Quality:** Tracks how well the patient sleeps, including duration and disturbances. Sleep quality is closely linked to various health outcomes, influencing everything from mental health to metabolic efficiency and cardiovascular risk.

Results Section

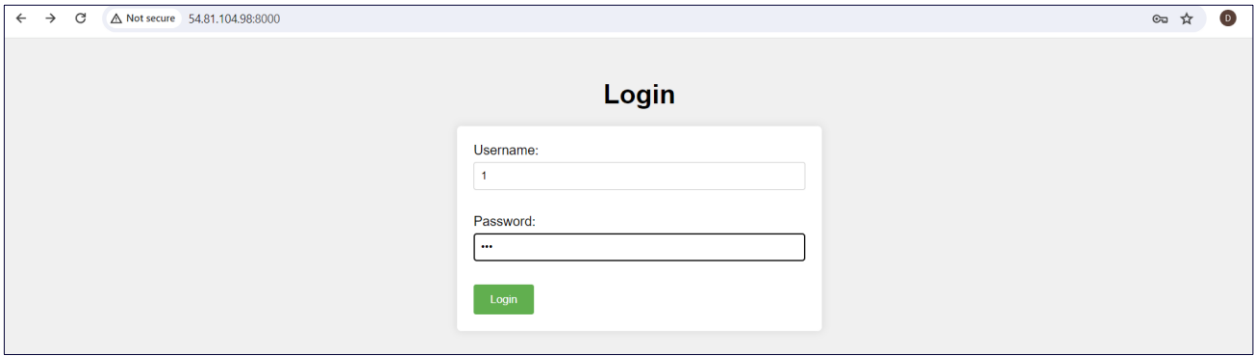
Initial implementation has shown promising results in improving patient adherence patterns and enabling healthcare providers to make informed decisions quickly based on real-time data. First the IOT Data will be uploaded using web application. After uploading the data will be passed to AWS IOT analytics and from there the data will be stored in S3 bucket.

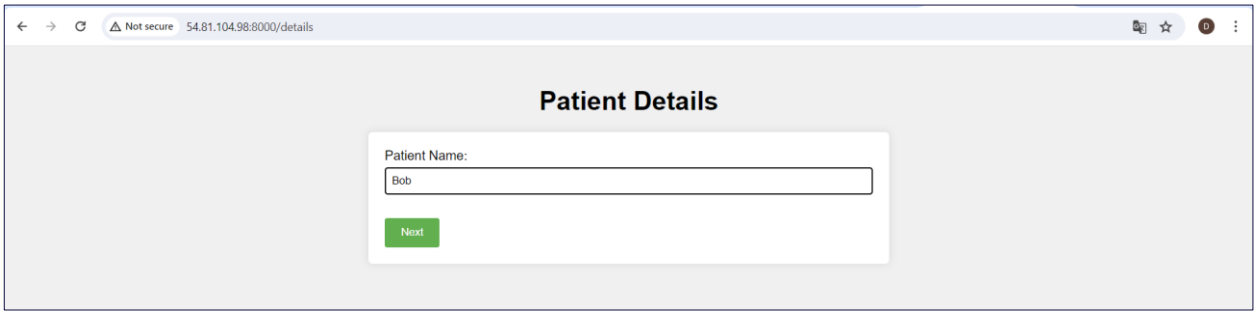
Sample IOT Data stored in S3 Bucket:

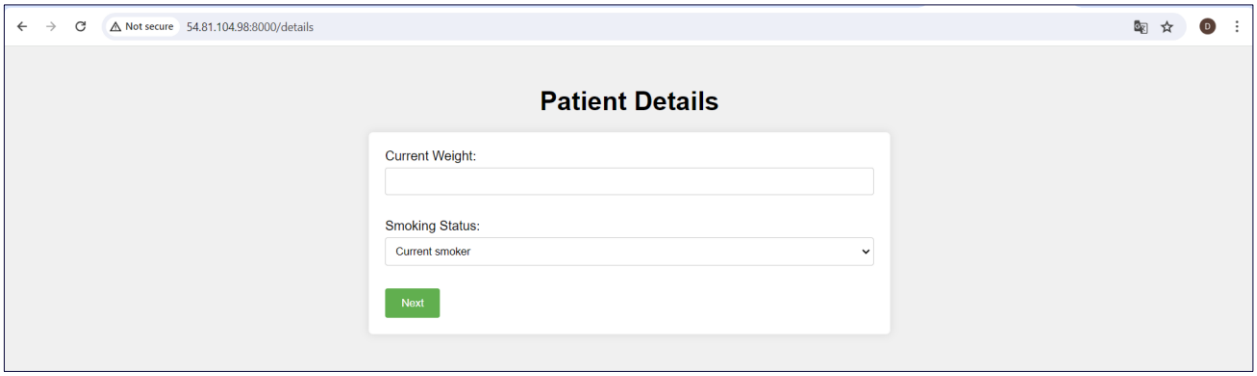
	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
1	heartrate	restecg	thalach	oldpeak	slope	thal	trestbps	bloodpressure	highbp	fbs	bodytemperature	physactivity	exang	weight	sleep_quality	patient_id
2	160	2	160	1	3	1	140	75	Yes	0	36.5	Moderate Activity	1	80	Good	3
3	180	0	180	3.2	2	2	120	72	No	0	36	Sedentary	0	65	Poor	2
4	155	0	155	1.8	2	2	125	70	No	1	36.8	Mild Activity	1	72	Good	5
5	170	1	170	0.5	1	3	130	62	No	0	37.2	Moderate Activity	0	75	Poor	4
6	150	1	150	2.3	1	3	145	0	Yes	1	37	Mild Activity	0	70	Good	1

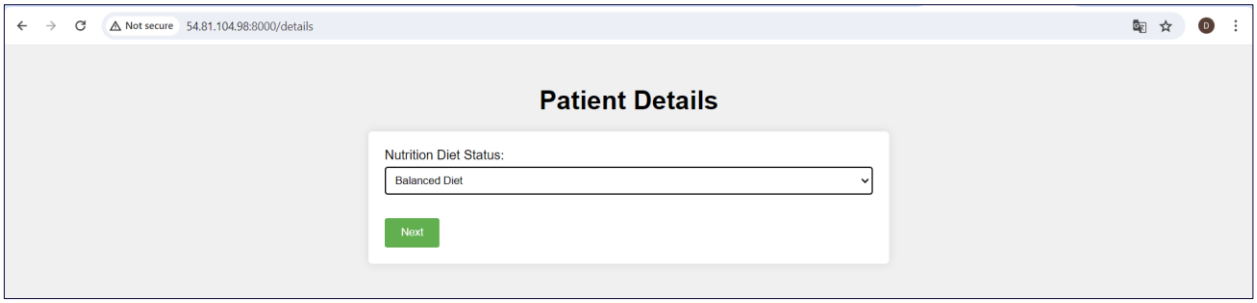
Patient Feedback

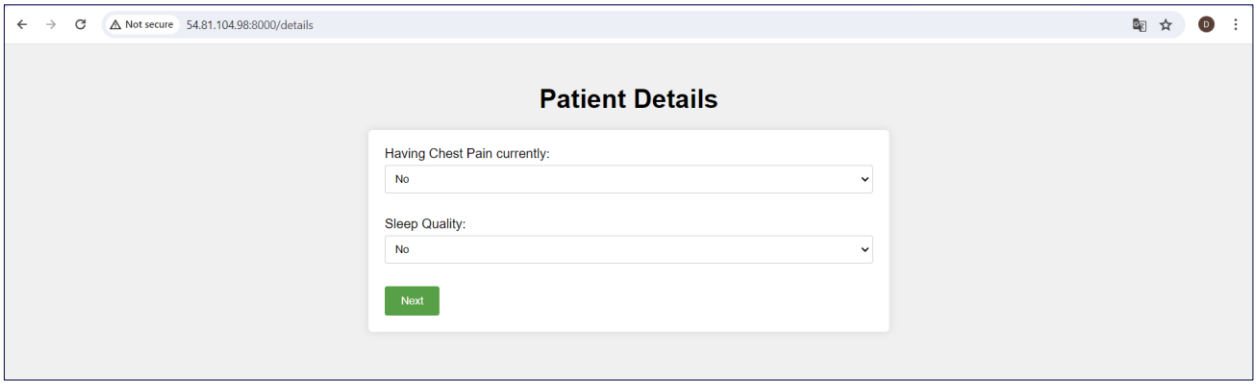
Patient Feedback will be collected through web application and stored in S3 bucket.

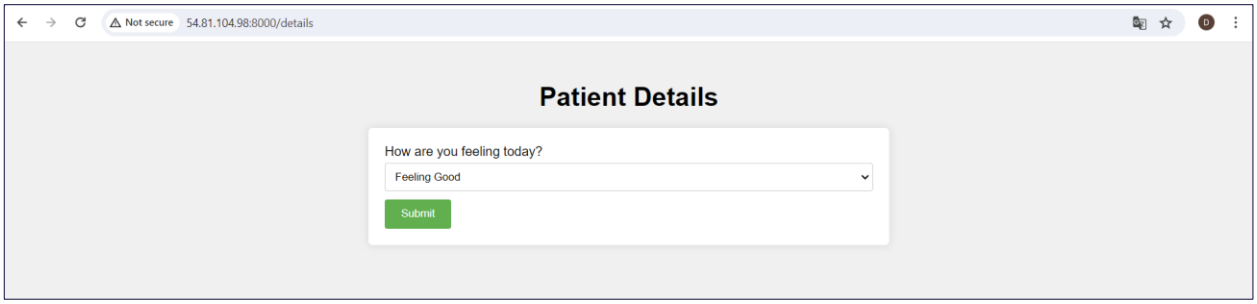




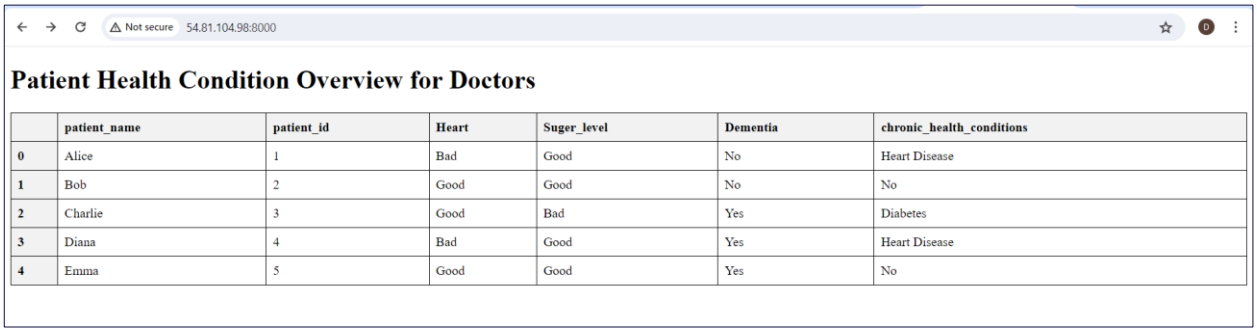








Once the data (IOT and Patient Feedback) are stored in S3 bucket. Using AWS Ec2, Machine learning model is run on that data and also the past data stored in S3 bucket. The model predicts whether the health condition based on the given data and the result is displayed through web application for Doctors to overview the result of the patient.



	patient_name	patient_id	Heart	Suger_level	Dementia	chronic_health_conditions
0	Alice	1	Bad	Good	No	Heart Disease
1	Bob	2	Good	Good	No	No
2	Charlie	3	Good	Bad	Yes	Diabetes
3	Diana	4	Bad	Good	Yes	Heart Disease
4	Emma	5	Good	Good	Yes	No

Discussion

This project demonstrates the potential of IoT in healthcare when combined with robust cloud computing services like AWS IoT Core and AWS S3. These technologies provide a secure and scalable foundation for collecting and processing health data, enhancing the ability to monitor and respond to patient needs proactively.

Conclusion

This IoT-based patient monitoring system represents a significant improvement in remote patient care. By leveraging AWS technologies for secure data handling and efficient processing, the system enhances the capability of healthcare providers to deliver timely and personalized care.

Contributions/References

Raihan Uddin and Insoo Koo, 2024. "Real-Time Remote Patient Monitoring: A Review of Biosensors Integrated with Multi-Hop IoT Systems via Cloud Connectivity." Applied Sciences. https://www.mdpi.com/2076-3417/14/5/1876

Hisham Alasmary, 2024. "ScalableDigitalHealth (SDH): An IoT-Based Scalable Framework for Remote Patient Monitoring." Sensors. https://www.researchgate.net/publication/378328449_ScalableDigitalHealth_SDH_An_IoT-Based_Scalable_Framework_for_Remote_Patient_Monitoring/fulltext/65d4a2fd113afb390513859f/ScalableDigitalHealth-SDH-An-IoT-Based-Scalable-Framework-for-Remote-Patient-Monitoring.pdf