**Task 1:Healthcare – Patient Risk Prediction System**

|  |  |
| --- | --- |
| Section | Description |
| Intern Name | Viraj V. Gurbade |
| Problem Title | Healthcare – Patient Readmission Risk Prediction System |
| Problem Description | **Target Users:** Doctors, clinicians, hospital management, and healthcare analysts.**Pain Points:** Unplanned readmissions increase healthcare costs and strain hospital capacity. Manual prediction methods are time-consuming and inaccurate. Lack of automation delays post-discharge follow-up and risk alerts. |
| Proposed AI/ML Solution | A **machine learning–based predictive system** that estimates a patient’s 30-day readmission risk using clinical and demographic data. The model leverages **XGBoost or Random Forest**, trained on historical patient data (age, BMI, vitals, lab results, comorbidities). The model is deployed via **Flask / Streamlit** and hosted on **AWS SageMaker or Azure ML** for real-time risk scoring. **Workflow:** Data Source (S3/Blob) → ETL & Preprocessing (Glue/Data Factory) → Model Training (XGBoost/Random Forest on SageMaker/Azure ML) → Model Deployment (Flask/Streamlit API) → Doctor Dashboard (Risk Alerts) → Monitoring (CloudWatch/Stackdriver). **Preprocessing:** Data cleaning, encoding, normalization, outlier removal, feature engineering (vitals, lab trends). |
| Cloud & Tech Stack | * **Cloud Provider:** AWS / Azure * **Services:** - AWS S3 / Azure Blob (data storage)- AWS Glue / Data Factory (ETL)- AWS SageMaker / Azure ML (model training)- Flask / Streamlit (API and dashboard)- DynamoDB / Cosmos DB (risk records)- CloudWatch / Stackdriver (monitoring)- Quick Sight / Power BI (analytics) * **Deployment:** Web-based dashboard hosted on EC2 / Azure App Service with secure IAM authentication. |
| Expected Outcome / Impact | - **Accuracy:** >80% AUC score for predicting 30-day readmission.  - **Efficiency:** Early alerts reduce manual risk analysis by ~40%.  - **Automation Benefits:** Continuous monitoring and risk flagging for clinicians.  - **Operational Impact:** Better bed management and reduced avoidable readmissions.  - **Clinical Outcome:** Improved patient follow-up and quality of care. |
| Architecture Diagram | **Digital End-to-End Flow:**Patient Data (S3/Blob) → ETL (Glue/Data Factory) → ML Model (SageMaker/Azure ML) → Flask/Streamlit API → Doctor Dashboard → Alerts & Monitoring (CloudWatch/Stackdriver/QuickSight). |
| Optional Notes | **Dataset Schema:** patient\_id, age, gender, blood\_pressure, cholesterol, bmi, diabetes, hypertension, medication\_count, length\_of\_stay, discharge\_destination, readmitted\_30\_days.  **Sample Input:** { "age": 65, "bmi": 29.5, "diabetes": "Yes", "hypertension": "No", "length\_of\_stay": 5 }  **Sample Output:** "Readmission Risk Score": 0.87" |