patient

RECORD MAINTENANCE

Table of Contents

Patient Health Record Maintenance

Project Description

In today's world, managing huge data for healthcare has become really crucial due to the need for accurate and efficient data management in healthcare. Just imagine a system that can handle a large amount of patient data with unparalleled integrity and security but also unlocks profound insights that can revolutionize the healthcare sector.

The main purpose of this project is to create a system that can be useful and improve the health sector in maintaining the health record for the purpose of analyzing and gathering insights. It is really crucial for the healthcare sector to keep accurate data in the hospital to make better patient care and get good accurate diagnosis and treatment plans for the patients. Usually, the patient records are collected in various ways.

Goal and Objective

The main objective of this project is to have a machine learning model that can handle huge amounts of patient data can ensure the integrity and security of data. This model will provide good insights to healthcare professionals such as doctors and nurses.

Data Scientist roles

In this project, we need 6 roles that are needed for this project:

1. **Data Scientist**

   These roles mainly focus on how to design and implement models that can be useful for analyzing patient health records. They would understand how the data can be cleaned and use appropriate statistical modelling techniques to derive insights which can be useful as well as responsible for creating visualisations to help communicate these insights.

2. **Data Engineer**

   This role mainly involves managing the data infrastructure. This can also involve setting up a database to store the patient health records and make sure the data is entered correctly and secure transfer from its source to the database and optimizing the data architecture for efficient analysis.

3. **Healthcare Analyst**

   This role would bring good demand since this project is on patient health records and we need people like healthcare analysts who have expertise in the project and can

interpret the results of the data analysis in the context of the healthcare sector such as hospitals. It would give them good insights into how the data can improve patient care.

4. **Project Manager**

   The main role is to oversee the project coordinating between different roles and ensure the project stays on track. They would also likely be responsible for communicating the results to the stakeholders.

5. **Data Privacy officer**

   The data privacy officer is responsible for ensuring that all data is handled appropriately with the relevant laws and regulations. They would also be responsible for implementing measures to protect the data from unauthorized authorities.

6. **Stakeholders**

   In this project, the stakeholders include healthcare professionals, patients and insurance companies who will be stakeholders. As they will be the end users of the insights derived from the data collected also it will give good insights and questions about the project's aim to the stakeholders.

Business model

In this project, the focus will be on the area of patient care and data record management. How do we store patient data through electronic devices such as Apple Watch, health applications, hospital visits etc? The project focuses more on machine learning models that can store data and predict insightful data for doctors' nurses and other purposes.

Advantage of this project

1. **Give patients personalized treatment plans**

   A comprehensive digital record that will provide a detailed history of a patient's medical background includes the previous treatment that they took, medications, allergies and their family history. From the given detailed history of the patient's medical background the system will recommend an accurate treatment plan to healthcare providers. Personalized treatment plans for the patient can give better outcomes as they are accurately aligned with their conditions and preferences.

2. **Improving the communication between different departments, doctors and nurses when they are checking the patient details.**

   An integrated patient record system should be there since it can enable the sharing of patient information across various departments and between different healthcare providers. This can improve their communication across the healthcare team. It can also reduce the risk of miscommunication and errors.

3. **Real-time sharing of patient records**

Real-time sharing means that any updates to a patient's record such as a new result of the patient or any change on their medication can be available through this to all authorized healthcare providers. This real-time sharing of information of patients can allow doctors to make efficient decisions which are particularly vital in critical care scenarios.

4. **Cost reduction**

Health care can reduce the paperwork and convert it into digital records which can reduce the need for physical storage space and administrative tasks which leads to significant cost savings. Furthermore, decreased errors and increased efficiency translate into fewer unnecessary tests and procedures, which lowers overall healthcare expenses.

5. **Predictive modelling with machine learning algorithms**

Large volumes of patient data can be analysed by sophisticated algorithms to find patterns and forecast future trends in health. Predictive modelling, for instance, can assist in identifying individuals who are highly susceptible to chronic illnesses, allowing for early intervention and preventative measures. Additionally, machine learning can help with faster and more accurate disease diagnosis, treatment recommendations based on case studies, and hospital operations optimization through the prediction of resource requirements and patient admission rates.

The challenge of this project

1. **Data Privacy**

The patient records contain sensitive and private information, and maintaining their security and privacy is a big task. Cyberattacks, unauthorized access, and data breaches can have serious repercussions, such as identity theft, legal problems, and a decline in public confidence in the healthcare system.

2. **Data Quality**

Keeping patient data accurate and comprehensive is essential to delivering high-quality healthcare. Data that is inaccurate, out-of-date, or incomplete can result in misdiagnosis, inappropriate treatments, and generally worse care standards. It's a constant struggle to keep data accurate across time and across many sources.

3. **Integration with Existing Systems**

It is difficult to integrate recently created health record systems with those that healthcare providers currently use. The presence of disparate data formats, standards, and protocols in

legacy systems might pose difficulties for smooth data interchange and interoperability. Data silos and inefficiencies in patient care may result from this.

4. **User Acceptance**

It is quite difficult to overcome healthcare providers' resistance to change, particularly if they are used to the current processes and systems. Scepticism about the advantages, worries about the learning curve, and possible interruptions to processes can all be reasons for resistance to implementing new technology.

Characteristics and Analysis Data

Potential Sources of Data Collection

- **Electronic Health Records (EHRs)**

Electronic Health Records (EHRs) are digitally detailed patient files that hold a wealth of information about the health and care of the patient. EHRs offer a comprehensive picture of a patient's medical background, facilitating precise diagnosis and efficient care.

- **Wearable Devices**

Fitness trackers, smartwatches, and medical wearables are examples of wearable technology that continuously and instantly gathers data on a range of health parameters. By giving users and healthcare practitioners up-to-date information about their physiological and physical status, these gadgets facilitate proactive health management.

- **Health Applications**

Mobile health applications provide a convenient interface for people to track and control several facets of their well-being. These programmes gather and examine information about physical activity, nutrition, mental health, and other topics.

- **Hospital Information Systems (HIS)**

Hospital Information Systems keep track of intricate patient data and oversee daily hospital activities. HIS ensures coordinated and effective care by offering a thorough record of patient contacts with the healthcare system.

- **Medical Imaging**

Medical imaging data is crucial for diagnosing and monitoring various medical conditions. Imaging technologies provide detailed visual representations of the inside of the body, aiding

in accurate and timely diagnosis.

- **Genomic Data**

Analysing a person's genetic makeup yields insights into hereditary predispositions and enables the development of customized treatment plans. Precision medicine is an emerging field in which this data is crucial.

Analysis of data characteristics

1. **Volume**

The volume of patient health records is huge to maintain as well as it has more detailed records to maintain for each patient over time. This includes even the medical tests, treatments as well as interaction with healthcare professionals.

Since the project focuses on machine learning we can use it for the storage and processing of the data like handling and storing of large datasets which require robust infrastructure as well as we can focus on scalability where the model can scale to manage and increase the data volume.

2. Velocity

The velocity is the speed at which the health data can be generated and needs to be processed equally and rapidly increasing. Real-time data from wearable devices, mobile health applications and continuous monitoring. Coming to the project goal and objective, we will be focusing on machine learning techniques to improve the patient record maintenance that we need for real-time processing (including real-time heart rate monitoring and analysis from fitness trackers) and streaming analytics.

3. Variety

Health data comes in different formats and structures as structured data includes electronic health records, and lab results, unstructured data such as medical imaging, and clinical notes and semi-structured data includes data from wearable devices. The project will use machine learning using data integration and model complexity for combining various data types as well as for preprocessing steps and model architectures.

4. Veracity

Ensuring the accuracy, completeness and reasonableness of health data is critical. Inaccurate or incomplete data can lead to incorrect diagnoses and treatment plans and undermine patient care. In this project, we will be using data quality management and bias and variability.

Platforms, software and tools for data processing and storage

- **Platforms:** Cloud computing platforms like AWD, Azure or Google Cloud for scalable storage and processing capabilities.
- **Software:** Database Management System (e.g. MySQL, PostgreSQL) for structured data storage, Apache Hadoop for big data and Apache Spark for real-time analytics.
- **Tools:** Python, TensorFlow or PyTorch for machine learning and Tableau or PowerBI.

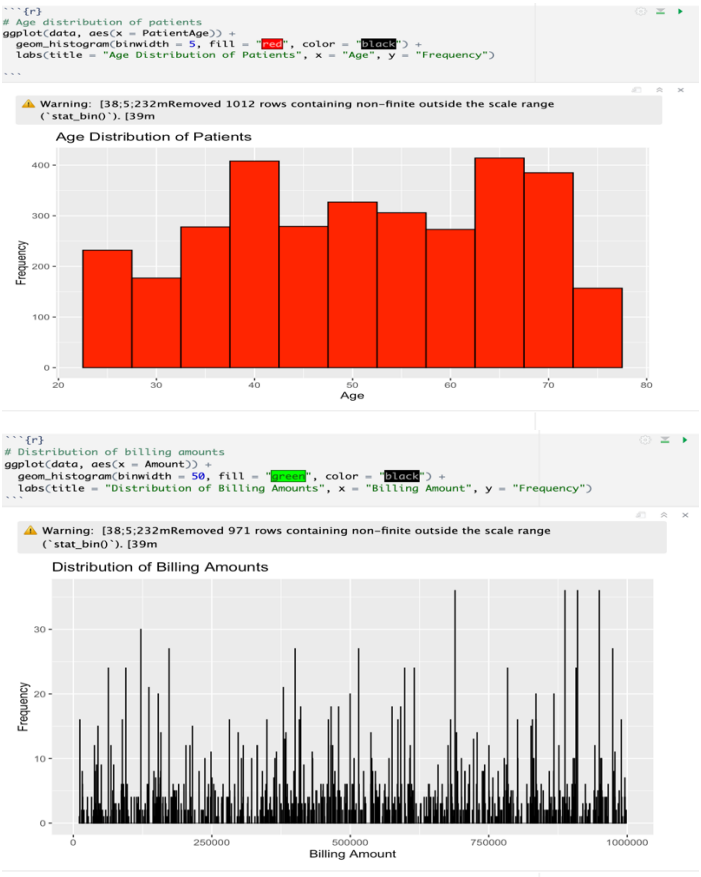
Data Analysis and Statistical Methods

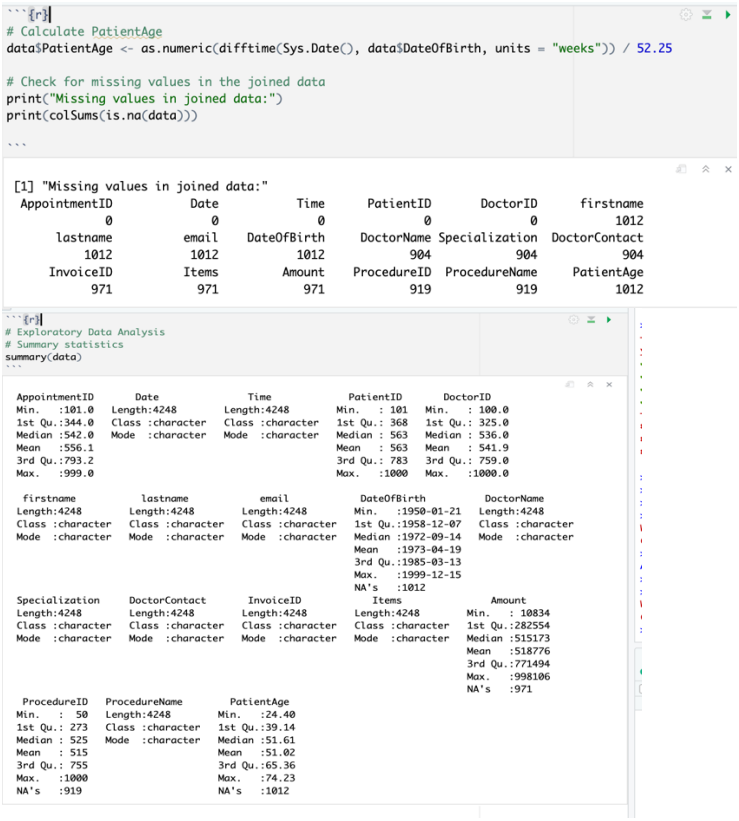
- **Decision Trees (Classification and regression)**
  - **Numerical and Categorical data**
    - Making them suitable for predicting patient outcomes or treatment effectiveness
- **Regression Analysis**
  - Model relationship between patient characteristics and health outcomes.

High-Level Output

- Decision trees can provide insights into factors influencing patient health outcomes, identifying key predictors and their relative importance.
- Regression analysis can quantify the impact of patient demographics and clinical variables on health outcomes.

Demonstration





```r
# Calculate PatientAge
data$PatientAge <- as.numeric(difftime(Sys.Date(), data$DateOfBirth, units = "weeks")) / 52.25

# Check for missing values in the joined data
print("Missing values in joined data:")
print(colSums(is.na(data)))
```

```
[1] "Missing values in joined data:"
 AppointmentID           Date           Time      PatientID       DoctorID      firstname
             0              0              0              0              0           1012
      lastname          email    DateOfBirth     DoctorName Specialization  DoctorContact
          1012           1012           1012            904            904            904
     InvoiceID          Items         Amount    ProcedureID  ProcedureName     PatientAge
           971            971            971            919            919           1012
```

```r
# Exploratory Data Analysis
# Summary statistics
summary(data)
```

```
 AppointmentID       Date              Time            PatientID       DoctorID
 Min.   :101.0   Length:4248        Length:4248        Min.   : 101   Min.   : 100.0
 1st Qu.:344.0   Class :character   Class :character   1st Qu.: 368   1st Qu.: 325.0
 Median :542.0   Mode  :character   Mode  :character   Median : 563   Median : 536.0
 Mean   :556.1                                         Mean   : 563   Mean   : 541.9
 3rd Qu.:793.2                                         3rd Qu.: 783   3rd Qu.: 759.0
 Max.   :999.0                                         Max.   :1000   Max.   :1000.0

  firstname          lastname           email           DateOfBirth          DoctorName
 Length:4248        Length:4248        Length:4248        Min.   :1950-01-21   Length:4248
 Class :character   Class :character   Class :character   1st Qu.:1958-12-07   Class :character
 Mode  :character   Mode  :character   Mode  :character   Median :1972-09-14   Mode  :character
                                                          Mean   :1973-04-19
                                                          3rd Qu.:1985-03-13
                                                          Max.   :1999-12-15
                                                          NA's   :1012
 Specialization     DoctorContact       InvoiceID          Items               Amount
 Length:4248        Length:4248        Length:4248        Length:4248        Min.   : 10834
 Class :character   Class :character   Class :character   Class :character   1st Qu.:282554
 Mode  :character   Mode  :character   Mode  :character   Mode  :character   Median :515173
                                                                             Mean   :518776
                                                                             3rd Qu.:771494
                                                                             Max.   :998106
                                                                             NA's   :971

  ProcedureID   ProcedureName        PatientAge
 Min.   :  50   Length:4248        Min.   :24.40
 1st Qu.: 273   Class :character   1st Qu.:39.14
 Median : 525   Mode  :character   Median :51.61
 Mean   : 515                      Mean   :51.02
 3rd Qu.: 755                      3rd Qu.:65.36
 Max.   :1000                      Max.   :74.23
 NA's   :919                       NA's   :1012
```

Standard for Data Science Process, Data Governance and Management

The Health Level Seven (HL7) standard is one of the primary benchmarks utilised in the data science procedure for maintaining patient health records. An international standard called HL7 is used for electronic health information retrieval, sharing, integration, and exchange. It establishes message, document, and interface standards that provide efficient communication between various healthcare systems.

Suitable Procedures for Data Management and Governance

1. Accessibility
- Use role-based access control (RBAC) to make sure that, in accordance with their duties and responsibilities, only workers who are authorised can access certain patient health records.
- Use safe authentication procedures and encryption techniques to grant regulated and secure access to data.

2. Security
- Protect data while it's in transit and at rest by using encryption techniques to guarantee its confidentiality and integrity.
- To protect against unauthorized access and cyber threats, put strong cybersecurity measures in place, such as firewalls, intrusion detection systems, and frequent security audits.

- Conduct regular security training and awareness programs for staff to mitigate risks associated with human error and social engineering attacks.

3. Confidentiality

- Adhere to privacy regulations such as HIPAA (Health Insurance Portability and Accountability Act) to protect the confidentiality of patient health information.
- Anonymize or de-identify sensitive data whenever possible to minimise the risk of re-identification.

4. Ethical Concerns

- Respect patient autonomy and rights by obtaining informed consent for data collection, usage, and sharing whenever applicable.
- Maintain transparency regarding data handling practices, including data storage, processing, and sharing, to build trust with patients and stakeholders.

Reference

Africa, E. (2023, July 20). *Key advantages of adopting a centralised patient medical record system*. https://www.linkedin.com/pulse/key-advantages-adopting-centralised-patient-medical-record/

SecureScan. (2024, January 23). *Medical Records Management: challenges and best practices*. SecureScan .https://www.securescan.com/articles/records-management/medical-records-management-challenges-and-best-practices/

Saiyad, T. (2023, December 25). *What is the Patient Record Management System: and its role in healthcare*. Bigscal - Software Development Company. https://www.bigscal.com/blogs/healthcare-industry/patient-record-management-system/

Prince. (2023, May 16). Who can benefit from AI in hospital management systems. Medium. https://medium.com/@prince_11681/who-can-benefit-from-ai-in-hospital-management-systems-6216c80217f4#:~:text=AI%20algorithms%20can%20be%20used,thus%2C%20maintains%20total%20efficient%20operations

Skedulo. (n.d.). AI in healthcare operations management. Skedulo. https://www.skedulo.com/blog/ai-healthcare-operations-management/

Bigscal. (2023, March 10). AI act: Best tool for hospital management system software. Bigscal. https://www.bigscal.com/blogs/healthcare-industry/ai-act-best-tool-hospital-management-system-software/

Deb, R., & Lichtenstein, S. (2019). Big data in healthcare: Investigating the most critical areas. Journal of Big Data, 6(44). https://doi.org/10.1186/s40537-019-0217-0

Nearshore IT. (n.d.). Big data in healthcare: How technology is improving patient care. https://www.nearshore-it.eu/articles/technologies/big-data-in-healthcare/

Kaur, H., & Sharma, N. (2018). A review on big data analytics in healthcare: Applications and challenges. Journal of Biomedical Informatics, 84, 66-79. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6032968/

datapine. (2020, January 3). 10 Powerful examples of big data in healthcare that can save people. https://www.datapine.com/blog/big-data-examples-in-healthcare/

VinBrain. (2023, March 14). Top 10 big data applications in healthcare (Part 1). https://vinbrain.net/top-10-big-data-application-in-healthcare-part-1

Banerjee, A., & Biswas, M. (2021). Applications of big data analytics in healthcare management systems. ResearchGate. https://www.researchgate.net/publication/352545013_Applications_of_Big_Data_Analytics_in_Healthcare_Management_Systems