RESPAI – Project (Case AI solution proposal)

Preventing admission or readmission to advanced specialist health care

Group 3

Member 1 Sundara Lakshmi Paramasivam. supa3400

Member 2 Shirdi Manjunath Adigarla, shad9520

Member 3 Muthupriya Shankaran,mush8206

1. Chapter 1 – Understand Business Problem

1.1. Identify and select key business initiatives

This paper explains about **"Preventing admission or readmission to advanced specialist health care**":

In general, health care organizations always prioritized lowering of admission or readmission rates. Healthcare costs due to unplanned readmissions are high and negatively affect the patient's health and wellness. Moreover, hospital readmission is an undesirable outcome for older adults/elderly patients. Hence, we present readmission risk using AI approaches and graph analytics for predicting 30-day unplanned readmission for elderly patients who are in the age of above 60 years old and carry out proactive measurements. In addition we identify high risk patients based on sets of patient features. Further, we used graph analytics(ML) to analyze a curated set of variables to pinpoint the risk of readmission and proactive prevention. However, due to the huge amount of data stored in the hospital's electronic health record (EHR), patients' records are easily extracted from databases. At last, we present in-depth insights about ethics, morality, and the state of the art in machine learning explain-ability.

1.2. Brainstorm AI business impact on the selected business initiative

The Electronic Health Record (EHR) data stored in the central data warehouse is fetched by mining data relevant to people admitted in Advanced healthcare and the location of the hospital. The data of in-depth features such as age, medication, disease, location and written personal information such as have they attended follow-up meeting, following treatment plan and proper medication are fetched to the location relevant business unit. Considering, the attended follow-up meeting, following treatment plan properly and proper medication are in text format stored in the central database has to be fetched to the relevant business unit. The business unit uses NLP to convert data into suitable score format to be processed by graph analytics or machine learning algorithms. The data from smartwatches like a patient's blood pressure and other intrinsic features relevant to the disease is collected and can also be fed into the algorithm to proactively prevent admission or readmission in health care. The graph analytics centrality algorithm, NLP, big data i.e. data from real world and AI algorithms are identified to be used for this initiative. The features obtained with respect to the location can be trained globally in the federated learning technique to identify patients with risk factors and also to provide robotic automation tele-services with the phone number to the patient through robotic process tele-services and mobile alert . This could be used to analyse and decrease the time, cost, and increase efficiency by preventing admission or readmission in advanced healthcare.

1.3. Breakdown the business initiative into use cases and focus on one

Following are the uses cases:

Rule Based Risk score:

The risk score can be assigned to each disease and medication based on the severity and risks with the medication. The demographic features can also play a role in admission or readmission in healthcare. It's also considered. The score can be provided for age based on the criteria namely below 40, 50, 60 and above. This can be done with the help of a group of domain specialist doctors.

Proactive Identification of patients risks associated with Admission or readmission in advanced specialist healthcare:

The risks associated with admission or readmission in advanced healthcare can be detected using graph analytics of centrality algorithms.

Robotic process Tele Service Automation Assistance for patients:

The tele-service automation assistance can be provided through AI robotic process automation(RPA) services of AI algorithm(interface) to alert or notify patients through smart watch or mobile to remind about medication, treatment plan and follow-up meeting for the patient related with the identified risks of admission or readmission in special healthcare . The data gathered from smart watch about the patient, that he has been taking medication on-time (e.g. through blood pressure) and personal assistance provided by doctors or nurses about the timing of medicines to be consumed, treatment plan in the EHR to assist patients can be used by AI algorithms to monitor, control and care patient's health in various ways like alerting patient or the nurse(provider responsible for the patient).

Below is the AI solution provided for selected use case:

Proactive Identification of patients risks associated with admission or readmission in healthcare

The targeted stakeholders are the hospital's administration and doctors. They are the ones who benefit and are responsible or accountable for this use case to identify risks or features that are associated with admission or readmission in advanced healthcare to save time, clarity using picturization and enhance patient's health by identifying the risks associated with it and helps to find proactive measures. Targeted stakeholders, their roles and responsibilities are tabulated below

Example: -

Tasks	stakeholders	Citizen Domain specialist	Data scientists	Risk/ audit	Data Ops	Production/ exploitation	Resource admin/ architect
Identification	A/R	C		I			
Data preparation	C	A/R	C				
Data modeling	C	A	R				
Model acceptance	I	C	C	A/R			
Productionalization			A/R	I	C		
Capitalization			R		R		A
Integration to external systems					A/R		
Global orchestration		C			R		
User acceptance tests	A/R	R	C		I		
Deployments					R	A	I
Monitoring	I	C				A/R	I

R:Responsible A: Accountable C: Consulted I: Informed

The decision processes can be identified and validated using human validation to find if this algorithm predicted actual outcomes, previous existing methods and can also be identified by checking with the outcome or improvement in this initiative in the future months by checking on reducing admission or readmission in advanced healthcare with the hospitals.

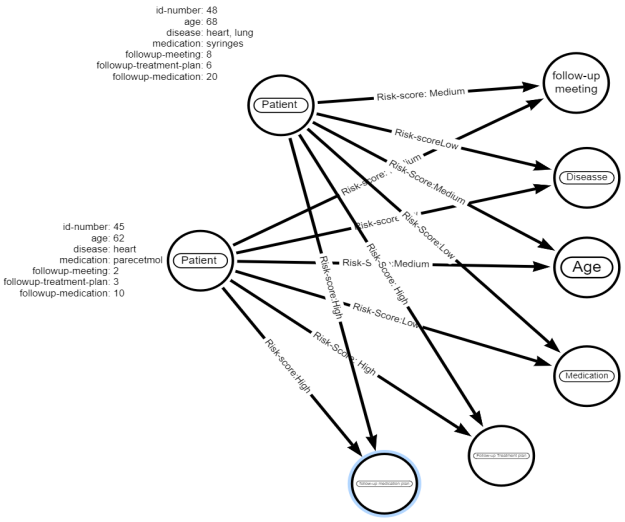
The data used are EHR data of in-depth features namely disease, demographic variables, medication. The other follow-up information from EHR can be converted into scores using NLP techniques and rule based systems with the ideas from domain specialists. The data from smartwatch can also be used, if possible.

The centrality algorithm is used to identify risks that contribute to admission or readmission in advanced healthcare.

1.4. Identify how graph analytics could be used in your use case selected

The patient IDs are used as nodes. The age, medication, diseases, following treatment plan, follow-up meeting, following medication are used as nodes. The relationship with score property is drawn from patient ID to other risk or feature nodes. For instance medication can be categorized as low-risk, medium risk and high-risk medication. The patient ID with low-risk will be assigned low value compared to medium-risk and high-risk medication. A relationship is drawn from patient to other risk feature nodes with the risk score value. Then, the centrality algorithm in graph analytics can be used to identify important risks or features such as more disease(comorbidities), age, medication, proper follow-up, etc, that may contribute to the admission or readmission in advanced healthcare.

Example:



1.5. Elicit requirements related to privacy, potential biases, and ethical issues

1.5.1. Privacy concerns

Data processing activities Check applicable potential data processing activities exposing users to privacy infringements.	**Identified privacy concerns** Describe how **data processing activities** in your new AI solution raise potential privacy infringements.	**Strategy to deal with potential privacy infringements** Suggest strategies to deal with identified **potential privacy concerns**, including but not limited to privacy by design, standards and regulations to deal with possible infringements.
Information collection ● Surveillance □ ● Interrogation □	The surveillance issue may arise, as personal data related to them are collected through EHR or smart devices.	This can be mitigated by giving information that it is used for the benefit of overall people and to improve the personal health and getting consents form signed from them.
Information processing ● Aggregation □ ● Identification □ ● Insecurity □ ● Secondary use □ ● Exclusion □	The aggregation and identification are involved.	The aggregation is done for the benefit of improving personal health of an individual. The positive sum of privacy by Design principle is used. The identification is encrypted while fetching the data and before graph analytics to prevent identification privacy.
Information dissemination ● Breach of confidentiality □ ● Disclosure □ ● Exposure □ ● Increased accessibility □ ● Blackmail □ ● Appropriation □ ● Distortion □	The disclosure, increased accessibility may happen.	The disclosure i.e. personal data are collected to improve their health and doesn't affect the reputation of the customer. The information is collected and gathered with the encrypted ID. The increased accessibility is prevented by mining patient data just relevant to advanced healthcare and the essential features are alone collected from the central data warehouse to the business unit.
I**nvasions** ● Intrusions □ ● Decisional interference □	The intrusion may happen as the data is transferred from EHR to business units.	The encryption is used for patient ID and secured network can be implemented between central data warehouse(EHR), smart networks and hospital business units to avoid interference or

		persuading data by intruders in between the network layers. The model is used globally, the model parameters can alone be passed for training globally through Federated learning to provide robot automation tele services.

1.5.2. Identify potential bias

The source of bias in Entering EHR system: In some situations, certain patients who are under fractured care or the patients with lower health literacy may find difficult or may not be able to access online patient portals and document their problem, in such cases, with respect to clinical decision support if data are missing or not accessible, then the data may eliminate/removed, and results in an inaccurate predictions of algorithm and may not get a clinical care support to people whose data are missing from the dataset. Therefore, to overcome this issue, we must identify the target population and select training and testing sets. Moreover, test algorithms for discrimination behavior throughout data processing should be carried out, also as an added advantage we need to check for validity and develop feedback loops to monitor and verify output. Furthermore, focus on medically improvements in relevant outcomes instead of strict performance metrics.

Conceptual bias: The confirmation bias may occur as the data features are to be selected based on the domain specialist and data scientist preference relevant to admission or readmission in health care. A group of Data scientists, domain specialists namely doctors, data engineers can be assigned and discussed to avoid popping up of bias related to race, etc,.

The **stability bias** may pop-up. As, the risk scores are assigned based on the risk of the medication level, disease level. This must be updated if new medicine in the real world has been found with lower risk. This is used in combination with the smartwatch sensor real time data to identify risks associated with them. It has to reflect in the algorithm. It must be done cyclically to monitor, identify risks and proactively prevent them. The algorithm can be weighted with the cost and benefits of saving time. The use cases for this initiative saves time, cost of the hospital staff, gives graphical representation to understand clearly which is done by automation.

1.5.3. Identify ethical concerns

There are a variety of ethical implications around AI in healthcare. The use of smart machines and algorithms may raise an issue of accountability, trust or transparency, privacy and security to create an ethically responsible approach. The accountability in the algorithm is done by explaining the algorithm and its value is to promote the well being of the patients. The value of identifying risk or factors associated with admission or readmission in advanced health care is found through the AI PageRank centrality algorithm. It also preserves the dignity of the people and organization. The explainability and transparency in the use case is achieved through graph analytics. The data and processes used are specified. The people or the algorithm responsible for this use case has to be addressed. As the chain of people and algorithms are used, the people responsible and accountable are difficult to identify. The quality of the data is ensured by data governance techniques through privacy and security mechanisms by achieving data compliance. Privacy by Design is embedded in the system to have a trade-off between tackling business(defensive) and offensive activities.

1.6. Make a proof-of-concept of selected business use case

The data sources needed for the use case are EHR patient records of people admitted or readmitted in advanced healthcare and it is found from the central data warehouse. The smart watch data of the patient or text data in digital format of EHR are used to identify that the person follows his treatment plan, follows his medication, etc,.

The source data can be extracted from the central EHR record to the business unit. Here, the cleaning of data like missing values and distribution of the values are analysed through visualization graphs. The data from smart watches is given higher priority over EHR records as it gives more accurate data through measurement of the patient's health data in the real time. The EHRs existing data are converted into table or usable format to convert into homogenous data, making it suitable for the algorithm. They are fetched in the local business unit to be processed by the centrality algorithm in NEO4j. The important features extracted from the Centrality algorithm identified can be trained globally, through federated learning to provide robot process automation tele and alert services for the patient with risk factors to be admitted or readmitted in advanced special healthcare, as a way of reminder and to help patients. This is done as a proactive measure using RPA to provide robotic automation tele services on the medication and disease complexities. In non-availability of answers to queries, it can be diverted to the respective nurse or doctors. The proposed AI solution can be integrated as an app.

1.7. MLOps and challenges in organizations

Challenges: Cost factor for smart watch, designing app etc are initially high but the returns of reduced readmission overcomes this challenge. Integrating with the legacy EHR system is a time consuming challenge which can be overcomed by communication with different departments, understanding and initial migration to improve the scalability and efficiency. Training or educating employees and users overcomes the challenge of usage of these systems to the hospital and users. Planning the deployment by investigating and talking with different experiments and cannibal deployment, A/b testing rolling slowly is employed to deploy the model in the right way.. Hiring the specialized people with different experience like a data science unicorn, software engineers, specialists finding is also a challenge. Hiring large consultants initially will overcome this problem.. We also use simple algorithms where the features etc provide better explainability etc. While designing different standards are considered like GDPR, privacy by design, openness etc to avoid compliance issues.

To solve the problem of inter communication, risk mitigation between different departments like business people, doctors and users, visualization analysis(KPI) tools like tableau are provided,meetings are held weekly to discuss the adoption progress etc. For programmers cloud technology is used to provide AI tools, licensing with proper scalability, continuous model integration/improvement/deployment. Responsible AI, Scalable AI(clusters in cloud using neo4j databases etc) deployed across various regions to reduce the lag, provide large bandwidth .When deploying A/B test methods are used that is to not deploy full model at once or data.But deploying slowly to some servers or parts and if no bugs are found then full deployment is issued. Objectives:Feedback from users, concept and data drift are considered and integrated into the process continuously after deployment in batches after severe testing. The environments are separated into Development:same as prod env for testing.. Risk evaluation.QA model, severe test case writing etc are employed. .Each phase in ML application is designed such that it satisfies requirements like reproduction,documentation and security. Dockers are also used to reduce dependencies, isolate environments, multiple models deployment and provide faster scalability apart from clusters. Monitoring is provided in the form of dashboards like tableau etc. The data is separated into training, validation and testing. The testing dataset is strictly not made available, used while designing the model.

References

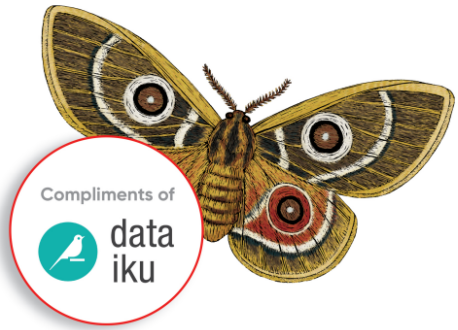
- Lecture slides
- Machine learning for predicting readmission risk among the frail: Explainable AI for healthcare,https://www.sciencedirect.com/science/article/pii/S2666389921002622
- Cavoukian, A. (2011). Privacy by design in law, policy and practice. A white paper for regulators, decision-makers and policy-makers.
- Solove, D. J. (2006). A Taxonomy of Privacy. University of Pennsylvania Law Review, 154(3), 477.
- https://wiki.openrightsgroup.org/wiki/A_Taxonomy_of_Privacy
- https://databricks.com/wp-content/uploads/2021/09/ML-Engineering-Ebook-Final.pdf

Introducing MLOps

O'REILLY®

Introducing MLOps

How to Scale Machine Learning in the Enterprise



Compliments of

data iku

Mark Treveil
& the Dataiku Team