PATIENT CASE SIMILARITY

Batch Number: CAI-G24

Roll Number	Student Name
20211CAI0062	Pranav Ganesh
20211CAI0063	Prerna Kakade
20211CAI0069	Bhuvana V
20211CAI0085	Nida Aiyman

Under the Supervision of,

Dr. Mohammadi Akheela Khanum
Professor
School of Computer Science and Engineering
Presidency University

Name of the Program: B.Tech
Name of the HoD: Dr. Zafar Ali Khan
Name of the Program Project Coordinator: Dr. Afroz Pasha
Name of the School Project Coordinators: Dr. Sampath A K / Dr. Abdul Khadar A / Mr. Md Ziaur Rahman

Introduction

- Patient Case Similarity is used in healthcare systems, particularly in clinical decision support systems, to give similarity scores between the new and old patients.
- In this project, we are comparing a new patient's data with a historical patient. The data is gathered from electronic health records (EHRs) and various research papers.
- This is done to identify similar patterns and predict the similarity score.
- The main goal is to cluster patients based on heart diseases.
- After which we will improve the diagnostic accuracy, predictive models and optimize the models.

Hardware/Software components

Hardware:-

Laptop (11th Gen Intel(R) Core(TM) i5-11320H @ 3.20GHz 2.50 GHz, 16 GB RAM)

Software:-

i. Operating System: Windows

ii. Programming Language: Python 3.7 or higher - The primary language for machine learning and deep learning development.

iii. Machine Learning Libraries:

- NumPy, Pandas: For data manipulation and analysis.
- Matplotlib, Seaborn: For data visualization.

iv. Google colab notebook and Visual Studio Code: For running interactive Python code and sharing results with the research team.

v. Git: For version control and collaboration.

Literature Review

Research Paper	Year	Advantages	Disadvantages
i. Use abstracted patient-specific features to assist an information-theoretic measurement to assess similarity between medical cases	2008	• Improve case similarity measurement • Integrate natural language processing (NLP) for feature abstraction	• The study admits limitations and focuses on only four feature types and the lack of contextual information in weighing features. • Future research should examine the impact of additional features and contextual factors on similarity measures to upgrade their suitability.
ii) Patient Similarity: Emerging Concepts in Systems and Precision Medicine	2016	• The goal is to set the foundation for the integration of computational tools and data analytics to enhance personalized healthcare. • The main focus is how patient similarity algorithms can improve medical decisions by grouping similar patients.	Data Difficulties Growth Challenges

Research Paper	Year	Advantages	Disadvantages
iii) Patient similarity for precision medicine: A systematic review	2018	• The primary aim of this line of research is to improve clinical outcomes for individual patients through more precise treatment targeting by leveraging on genetic, biomarker, phenotypic, or psychosocial characteristics. • That distinguish a given patient from others with similar clinical presentations	• Limited Database Scope • Lack of Real-World Application
iv) A patient-similarity-based model for diagnostic prediction	2019	• To simulate the clinical reasoning of doctors, retrieve analogous patients of an index patient automatically and predict diagnoses by the similar/dissimilar patients. • The main goal is to predict patient diagnoses by comparing the similarities between the clinical features of current patients and historical patient data.	• Limited Dataset Size • High Computational Costs

Research Paper	Year	Advantages	Disadvantages
v) Measurement and application of patient similarity in personalized predictive, modelling based on electronic medical records	2019	• The main goal of this research was to create a new way to measure how similar patients are, based on the data from electronic medical records (EMRs). • By measuring patient similarity, the study aimed to improve how we predict a patient's health outcomes, specifically focusing on diabetes.	• The study didn't fully use all the available data when calculating patient similarity • The models didn't include specific exclusion criteria when choosing patients for the study.
vi) Measuring Patient Similarities via a Deep Architecture with Medical Concept Embedding	2019	• Develop a framework to measure clinical similarities between patients based on EHRs. • Preserve temporal information in patient data, which is often lost in existing models.	• Loss of temporal information in existing methods: • High dimensionality and sparsity

Research Paper	Year	Advantages	Disadvantages
vii) Patient-Case Similarity	2020	• Develop a system to identify patients with similar medical histories. • Improve decision-making processes in clinical settings using patient data.	• Data Quality Dependency • Complex Medical Cases
viii) Patient similarity: methods and applications	2020	• Analyze and compute similarities between patients using electronic health records (EHRs), genetic, and other data. • Improve predictive models in healthcare by integrating patient-specific data from various sources.	• Information Loss • Complexity in Implementation

Research Paper	Year	Advantages	Disadvantages
ix)Patient similarity analytics for explainable clinical risk prediction	2021	• To develop an explainable and interpretable Clinical Risk Prediction Model (CRPM) by leveraging patient similarity analytics, specifically to improve explainability and interpretability. • To use real-world data from electronic medical records of patients with type-2 diabetes, hypertension, and dyslipidemia in Singapore to develop and validate the patient similarity model	• Incomplete Variable Set • Static Data Usage
x)A Novel Patient Similarity Network (PSN) Framework Based on Multi-Model Deep Learning for Precision Medicine	2022	• Utilizes multi-model deep learning to identify similarities among patients. • This patient similarity network (PSN) approach aims to enhance precision medicine by combining multiple data types, such as clinical records, genetic information, and imaging data.	• Data Heterogeneity and Dimensionality • Limited Availability of Open Datasets

Research Paper	Year	Advantages	Disadvantages
xi) Deep Dynamic Patient Similarity Analysis: Model Development and Validation in ICU	2022	• Develop a Novel Dynamic Patient Similarity Model • Validate Model Using Clinical Tasks	• Limited Clinical Application • Computational Complexity
xii) Patient Case Similarity	2024	• Improve healthcare analytics by leveraging data science techniques to enhance diagnostics, treatment recommendations, and patient care outcomes • The system utilizes machine learning (ML) and natural language processing (NLP) to identify similarities between patient cases, enabling more personalized, data-driven medical interventions.	• Prescription Recommendation Accuracy • Scalability Concerns

Existing method Drawback

i. Use abstracted patient-specific features to assist an information-theoretic measurement to assess similarity between medical cases – 2008:-

The study only looks at four types of features and doesn't consider the context when weighing them. Future research should look at how adding more features and context can improve similarity measures

ii. Patient Similarity: Emerging Concepts in Systems and Precision Medicine – 2016:-

Complexity of algorithms – many of the proposed algorithms are not yet optimized for real-world clinical use due to their complexity and reliance on high-end computing infrastructure.

iii. Patient similarity for precision medicine: A systematic review – 2018:-

Lack of Deep Learning Exploration - The paper talks very little about deep learning, which is now important for analyzing complex medical data and finding patient similarities. This might be a missed chance to use better methods.

iv. A patient-similarity-based model for diagnostic prediction – 2019:-

Low Success Percentage - The model's success percentage (the percentage of patients for whom diagnoses were correctly predicted) is low (19%).

v. Measurement and application of patient similarity in personalized predictive, modeling based on electronic medical records – 2019 :-

The models didn't include specific exclusion criteria when choosing patients for the study. This could affect the accuracy of the predictions because not all patients may be equally relevant for the predictive task

vi. Measuring Patient Similarities via a Deep Architecture with Medical Concept Embedding – 2019:

Electronic Health Records are complex, and patient records contain sparse and high-dimensional data.

vii. Patient-Case Similarity – 2020:

The use of machine learning models leads to over fitting, where the model performs well on training data but fails to generalize new or unseen data.

viii. Patient similarity: methods and applications – 2020

During data transformation and integration, particularly in early integration strategies, there is a risk of losing valuable patient information.

ix. Patient Similarity Analytics for Explainable Clinical Risk Prediction – 2021

The model doesn't include important factors like gender, race, diet, and lifestyle, which are linked to complications of diabetes, hypertension, and high cholesterol. Missing this data makes the model less complete.

x. A Novel Patient Similarity Network (PSN) Framework Based on Multi-Model Deep Learning for Precision Medicine – 2022

Data Heterogeneity and Dimensionality - A Novel Patient Similarity Network (PSN) Framework Based on Multi-Model Deep Learning for Precision Medicine – 2022

The mix of different types of clinical data, both structured and unstructured, makes it hard to create accurate models. Handling this complex data that can cause information to be lost during processes like auto encoders, which reduce the data's size. Using auto encoders for this can lead to a loss in accuracy.

xi. Deep Dynamic Patient Similarity Analysis: Model Development and Validation in ICU – 2022

Need for Clinical Protocol - The model requires a clinical protocol for practical implementation, which hasn't been discussed in this research

xii. Patient Case Similarity – 2024: Prescription Recommendation Accuracy: The system shows lower accuracy in prescription recommendations (61.27%), indicating room for improvement in this area.

Objectives

- Make use of historical patient data in order to make accurate and useful predictions about heart diseases.
- The system analyzes trends, patterns, and outcomes from the previous data and then recognize similarities between patients.
- By improving diagnostic accuracy, the model aims to reduce healthcare costs.
- The system helps healthcare professionals to identify the health risks early, which allows them to develop a tailored preventive plan.
- To improve healthcare efficiency by automating the identification of patient similarities, saving the time and resources of healthcare providers.
- Enhancement of clinical research by giving the researchers a tool that helps them in their studies about heart diseases.
- The project is committed to providing advancements in medical research and promoting a deeper level of understanding of heart diseases.
- A user friendly website makes sure that patients and healthcare professionals can access data easily and understand the insights predicted by the system.

Proposed Methodology

1. Loading data
2. Exploring data
3. Missing values
4. Preprocessing
5. Data visualization
6. Training and testing
7. Models
8. Clustering
9. Silhouette score

Loading data:

• Pandas library in Python, used for data manipulation and analysis.

•Loaded two datasets : test data and train data. These two DataFrames are concatenated into a single DataFrame using the concat() function.

Exploring:

•Data.info() is used in the exploration step of the algorithm.

•Data.describe() provides summary of statistics of all the numerical data in the dataframe such as count, mean, std, min, max and the percentiles.

Missing values:

• This method lets us know how many missing values are present in each column of the dataframe.

• The isnull() function generates the missing values in the dataframe

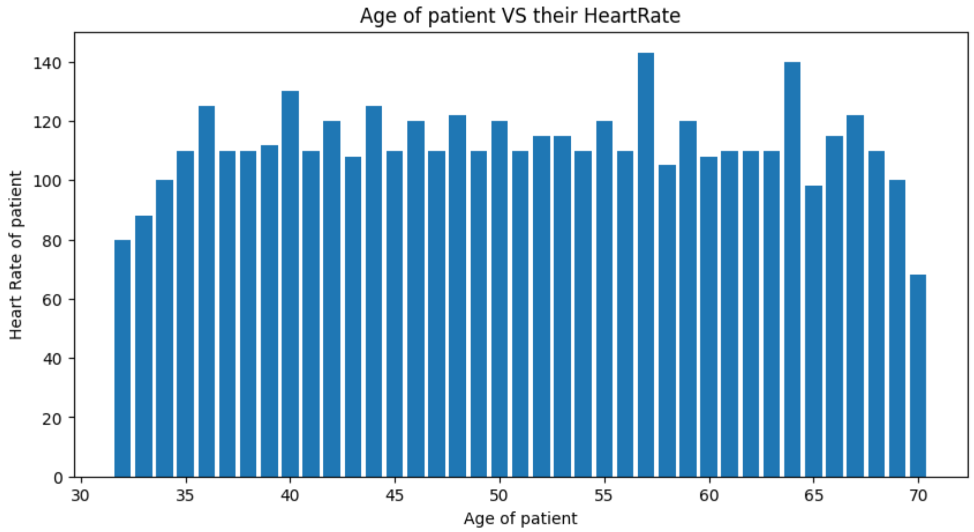
• Columns that show 0 count are the columns that have no missing values.

Preprocessing:

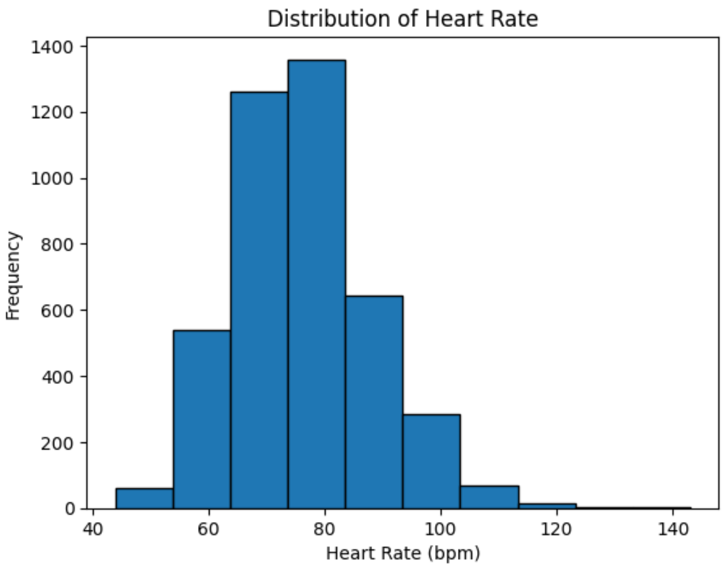
- This is a crucial step in data analysis which prepares raw data for further analysis and modeling.
- An effective way to handle these missing values is by filling them with the median value of each column.
- Certain columns like TenYearCHD, is_smoking, and education are dropped because they provide unnecessary information.

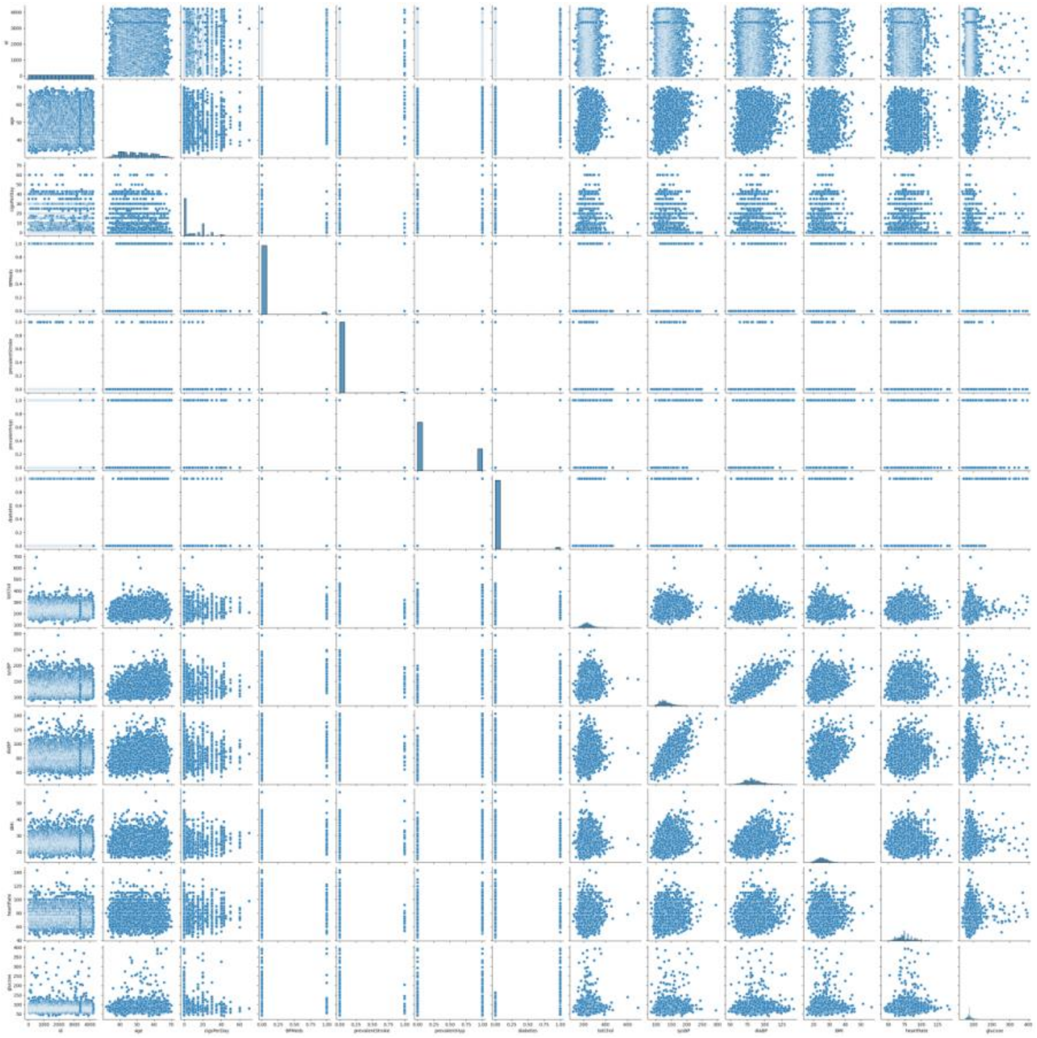
Data visualization:

- Libraries such as matplotlib and seaborn have been used.
- Matplotlib is used for generating charts and plots.
- Seaborn is built on top of matplotlib and is a library that is used to make complicated statistical visuals easier to create and understand.
- A bar plot has been created in order to compare the Age vs. Heart Rate of the patient.



- A Histogram has been created for checking the distribution of Heart Rates across various Frequencies.
- A pair plot has also been created to show the relationships between variables in the dataset.





Training and Testing:

- Before training the model, the data has to be split into two subsets: training data and testing data.
- In most cases, about 70-80% of the data is used for training. And the remaining 20-30% is used for testing. In our project, we have utilized 70% of the data for training and 30% for testing in order to prevent over fitting.
- During training, the model is fed data from which it learns patterns. The input data contains features such as :

1. Age
2. cigsPerDay
3. BPMeds
4. prevalentStroke
5. prevalentHyp
6. Diabetes
7. totChol
8. sysBP
9. diaBP
10. BMI
11. heartrate
12. Glucose.

- Once the training is done, the model is tested on new, unseen data. The model makes predictions and the results are compared to the actual outcomes to evaluate its performance.
- The testing phase is crucial because it makes sure that the model is not just memorizing the training data but also making accurate predictions on new unseen data.

Models:

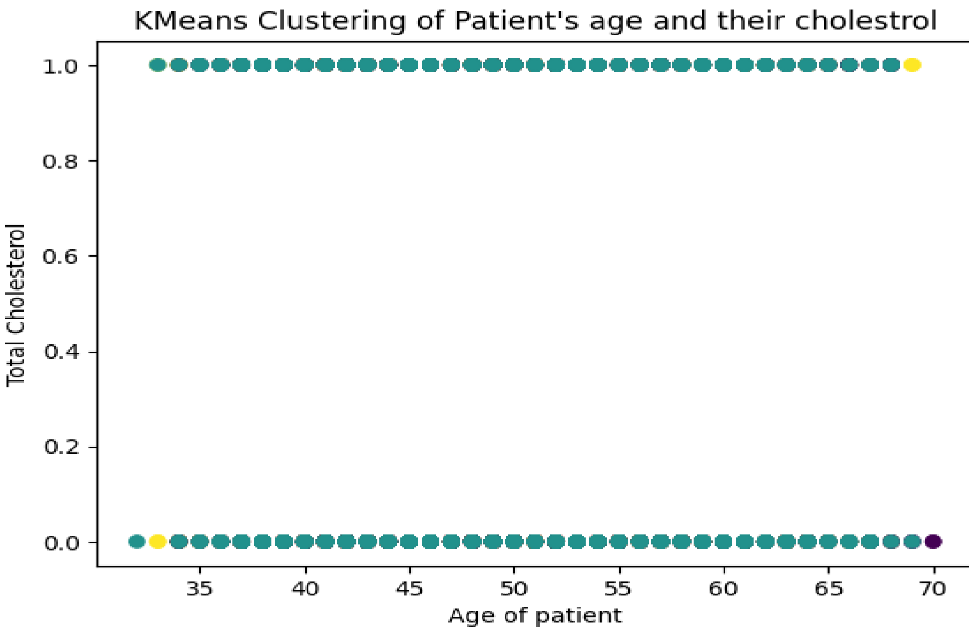
- The models that we have used in our backend are Logistic Regression, Decision Tree and Random Forest.
- Logistic Regression: This is used to predict if the patient has heart diseases or not. ‘1’ if the patient has heart diseases and ‘0’ if the patient does not have heart diseases based on features such as Age and cholesterol.
- Decision Tree: This is a flowchart which is used to make decisions based on different features.
- Random Forest: An ensemble method that combines multiple decision tress instead of just one to make an accurate and stable prediction.

Clustering:

- Clustering is a method in data analysis that groups similar objects or data points together.
- In our project, we have implemented K-means clustering, a well-known ML method used to divide data into a number of groups.
- A scatter plot is made with patient age on the x-axis and cholesterol levels on the y-axis, and the points are colored based on their cluster.

Silhouette score:

- The Silhouette Score checks how well data points are grouped together by calculating the score for different numbers of clusters using K-Means.
- It compares how similar each point is to its own group compared to other groups and prints the score for each number of clusters to help us choose the best one.
- Higher score means the groups are clear and better formed.



Architecture

Data Source Layer:

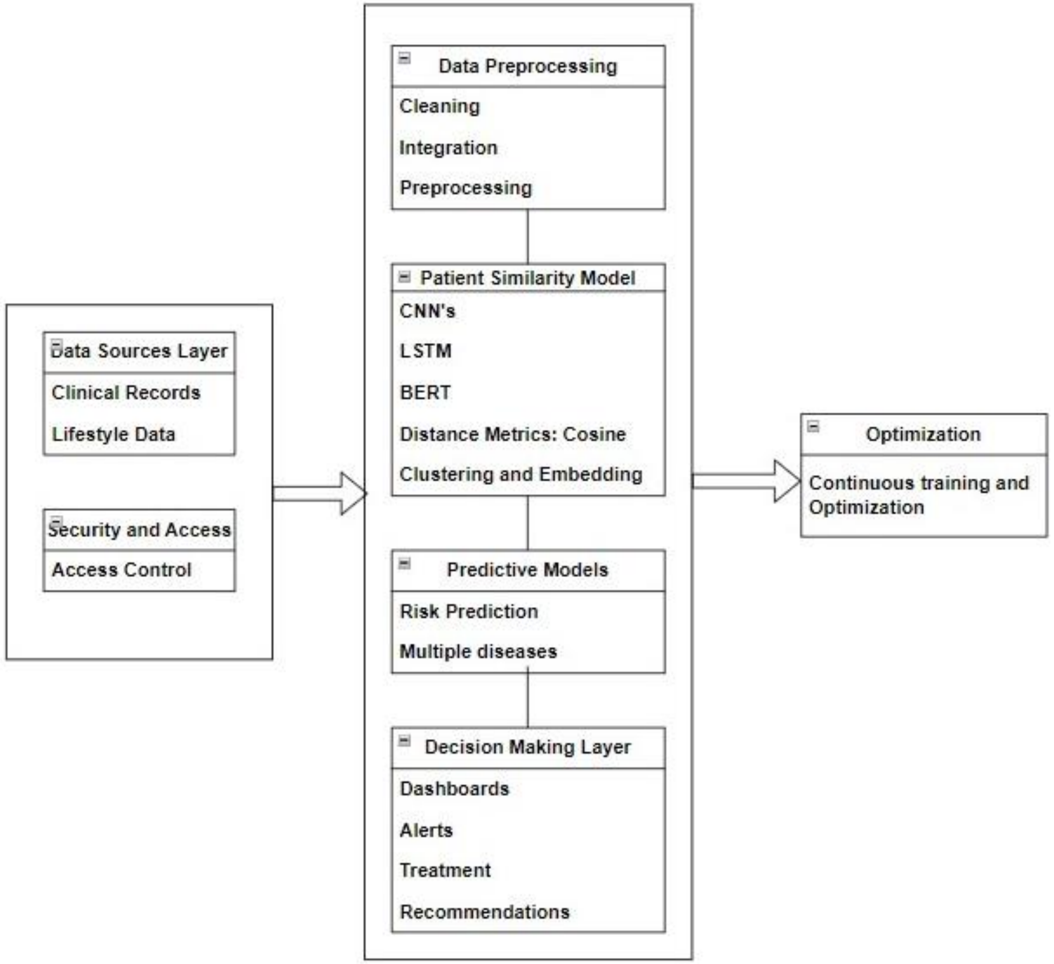
- Collect patient data and split it into two parts: one to train the model and another to test the model.
- Key feature: Keeps data secure and only allows authorized access.

Exploratory Data Analysis (EDA):

- Prepares the data and finds useful insights.
- Fixing problems in the data, like missing information.
- Creating simple charts (like bar charts, histograms, and pair plots) to understand patterns.

Patient Similarity Model:

- It groups patients who are similar to each other.
- Uses tools like:
  - ➢ Neural networks (CNN) for analyzing complex data.
  - ➢ Simple models (Logistic Regression, Decision Trees).
  - ➢ Clustering (K-Means) to create groups.



Predictive Models:

- It predicts what might happen to a patient (e.g., risks or outcomes).
- Uses advanced models like:

➢ Neural networks and Random Forests for predictions.

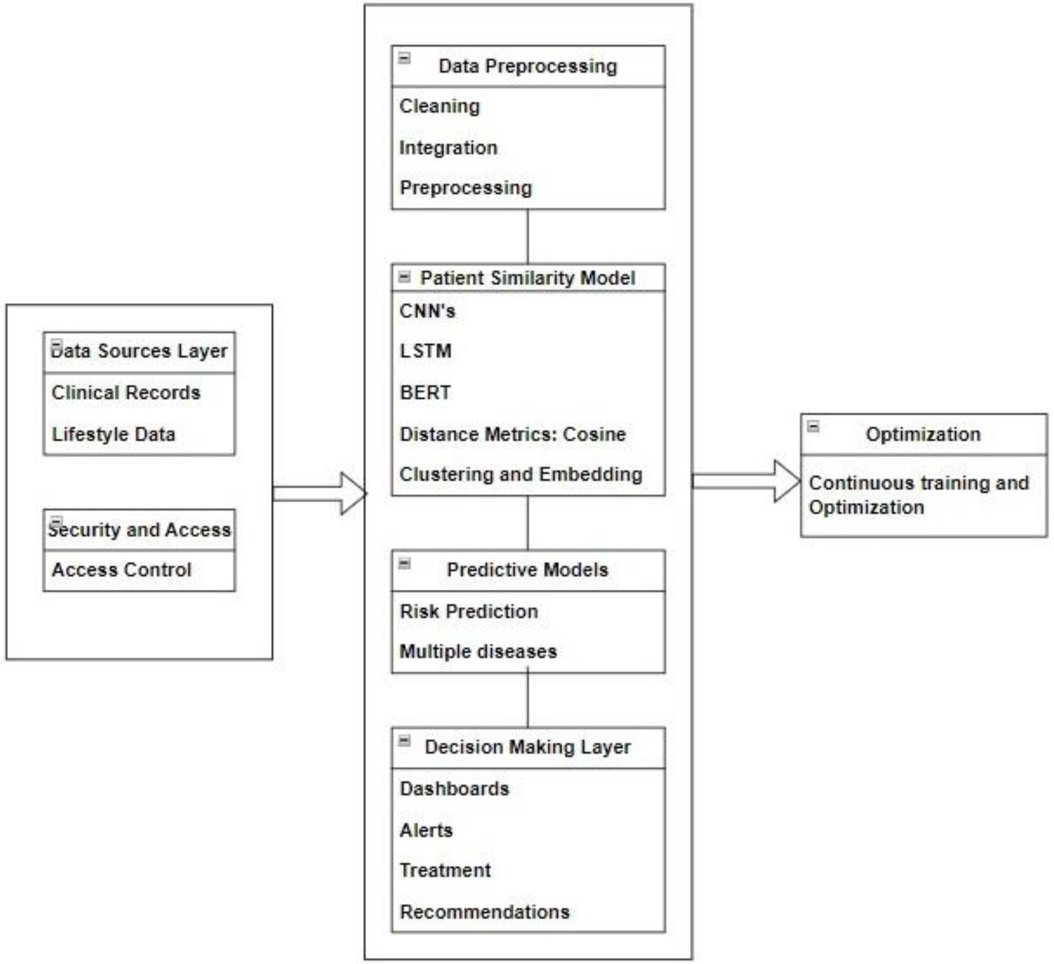
➢ Scores (Silhouette and Similarity) to measure accuracy and patient matching.

Decision-Making Layer:

- Helps doctors with clear, actionable outputs.
- Easy-to-read visual summaries.
- It notifies about critical events.

Optimization:

- It continuously improves the models by retraining them with new data to make better predictions over time.



Timeline of Project

G24 PATIENT CASE SIMILARITY			
Project Start Date:	04-09-2024		Time allocated
Current Date:	14-10-2024		Progress
Weeks:	Wk 6		

Task	Start Date	End Date	Days	Progress
Review 0				
Title Selection & PPT	04-09-2024	13-09-2024	9	100%
GitHub Repository	05-09-2024	13-09-2024	8	100%
Finalizing Objectives	08-09-2024	13-09-2024	5	100%
Methodology	08-09-2024	13-09-2024	5	100%
Review 1				
Abstract	12-09-2024	15-10-2024	33	100%
Software Details	12-09-2024	15-10-2024	33	100%
Exploring Research papers	16-09-2024	15-10-2024	29	100%
Architecture Diagram	18-09-2024	15-10-2024	27	100%
Source Code Details	18-09-2024	15-10-2024	27	75%
Review 1 Report	18-09-2024	15-10-2024	27	100%
Review 2				
Algorithm	15-10-2024	19-11-2024	35	0%
50% implementation of Code	15-10-2024	19-11-2024	35	0%
50% Report Completion	15-10-2024	19-11-2024	35	0%
Review 3				
100% implementation of Code	19-11-2024	17-12-2024	28	0%
100% Report Completion	19-11-2024	17-12-2024	28	0%
Live Demo	19-11-2024	17-12-2024	28	0%
Final Viva				
Live Demo	19-11-2024	17-12-2024	28	0%
Plagiarism Check	19-11-2024	17-12-2024	28	0%
Hard Copy Report Submission	19-11-2024	17-12-2024	28	0%
Publication of Research Paper	19-11-2024	17-12-2024	28	0%

04-09	08-09	12-09	16-09	20-09	24-09	28-09	02-10	06-10	10-10	14-10	18-10	22-10	26-10	30-10	03-11	07-11	11-11	15-11	19-11	23-11	27-11	01-12	05-12	09-12	13-12	17-12

Expected Outcomes

- Patient Case Similarity Model gives a similarity score of a new patient by referring to the historical data of an old or existing patient.
- In our project, we have determined the similarity score between a new patient and the five most similar patients in the historical data.
- This outcome is to bring more substantial improvements in the healthcare management.
- The similarity score determines how closely their cases align, with values approaching 1 representing higher similarity.

Most Similar Patients:

Patient ID: 10, Similarity: 0.9894803221396218

Patient ID: 11, Similarity: 0.9896825252977868

Patient ID: 14, Similarity: 0.9972892831540058

Patient ID: 21, Similarity: 0.9921180983850746

Patient ID: 23, Similarity: 0.9886220303454638

Results

- The dataset “test.csv” and “train.csv” were merged for data analysis and modeling.
- Missing values in columns such as cigsPerDay, BPMeds, totChol, BMI, heartRate, glucose, were filled with their median values.
- Valuable insights were gained from visualization. A bar chart showed the relationship between age of patients and their heart rates. Histogram showed a normal distribution of heart rate values. Pairwise relationships between numeric variables were visualized using pair plot.
- Logistic Regression gave an accuracy of 85.84%, Decision Tree Classifier gave an accuracy of 75.42% and Random Forest Classifier gave an accuracy of 85.05%.
- K-Means clustering grouped patients together based on their features and for these clusters a Silhouette score was determined.
- We have successfully determined the similarity score of a new patient in comparison to five patients from the historical data.

Conclusion

- The project "Patient Case Similarity" has successfully been used to demonstrate the power of Machine Learning in order to enhance clinical decision support systems.
- It shows how ML systems can assist doctors in making more accurate diagnoses and predicting health outcomes.
- Using historical data, the project has created an easy-to-use web application that looks at patient information, finds patterns, and predicts results by comparing current cases with similar past ones.
- The system employs powerful Machine Learning models, including Logistic Regression, Decision Tree, and Random Forest Classifiers. Among all these, the Logistic Regression model was the most reliable for prediction.
- The K-Means clustering method grouped patients with similar characteristics, and the results were supported by strong silhouette scores.
- The project also used visual tools like bar charts, histograms, and pair plots to provide useful insights into how different factors are related, improving the analysis.
- Furthermore, the system demonstrated its precision and adaptability by accurately calculating similarity scores for new patients compared to historical cases.

Github Link

[Patient Case Similarity]

References

i. Use abstracted patient-specific features to assist an information-theoretic measurement to assess similarity between medical cases – 2008: https://www.sciencedirect.com/science/article/pii/S1532046408000440
ii. Patient Similarity: Emerging Concepts in Systems and Precision Medicine – 2016: https://www.frontiersin.org/journals/physiology/articles/10.3389/fphys.2016.00561/full
iii. Patient similarity for precision medicine: A systematic review – 2018: https://doi.org/10.1016/j.jbi.2018.06.001
iv. A patient-similarity-based model for diagnostic prediction – 2019: https://doi.org/10.1016/j.ijmedinf.2019.104073
v. Measurement and application of patient similarity in personalized predictive, modelling based on electronic medical records – 2019: https://doi.org/10.1186/s12938-019-0718-2
vi. Measuring Patient Similarities via a Deep Architecture with Medical Concept Embedding – 2019: https://arxiv.org/pdf/1902.03376
vii. Patient-Case Similarity – 2020: https://www.researchpublish.com/upload/book/Patient%20Case%20Similarity-8606.pdf
viii. Patient similarity: methods and applications – 2020: https://arxiv.org/pdf/2012.01976
ix. Patient similarity analytics for explainable clinical risk prediction – 2021: https://bmcmedinformdecismak.biomedcentral.com/articles/10.1186/s12911-021-01566-y
x. A Novel Patient Similarity Network (PSN) Framework Based on Multi-Model Deep Learning for Precision Medicine – 2022: https://doi.org/10.3390/jpm12050768
xi. Deep Dynamic Patient Similarity Analysis: Model Development and Validation in ICU – 2022: https://www.sciencedirect.com/science/article/pii/S0169260722004151
xii. Patient Case Similarity - 2024: https://www.doi.org/10.56726/IRJMETS48246

Publication Proof


International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Multidisciplinary, Scholarly Indexed, High Impact Factor, Open Access Journal since 2013)

IJIRCCE

Impact Factor 8.625

CERTIFICATE

OF PUBLICATION

The Board of IJIRCCE is hereby awarding this certificate to

BHUVANA V

Student, Department of Computer Science and Engineering, Presidency University, Bengaluru, India

in Recognition of Publication of the Paper Entitled

"Patient Case Similarity"

in IJIRCCE, Volume 12, Issue 12, December 2024

e-ISSN: 2320-9801
p-ISSN: 2320-9798

Editor-in-Chief

www.ijircce.com ijircce@gmail.com


International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Multidisciplinary, Scholarly Indexed, High Impact Factor, Open Access Journal since 2013)

IJIRCCE

Impact Factor 8.625

CERTIFICATE

OF PUBLICATION

The Board of IJIRCCE is hereby awarding this certificate to

PRANAV GANESH

Student, Department of Computer Science and Engineering, Presidency University, Bengaluru, India

in Recognition of Publication of the Paper Entitled

"Patient Case Similarity"

in IJIRCCE, Volume 12, Issue 12, December 2024

e-ISSN: 2320-9801
p-ISSN: 2320-9798

Editor-in-Chief

www.ijircce.com ijircce@gmail.com

Plagiarism Check

Mohammadi Akheela Khanum - Final Report Capstone G24 (1)

ORIGINALITY REPORT

19%	16%	12%	13%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	Submitted to Presidency University	Student Paper	11%
2	www2.mdpi.com	Internet Source	1%
3	eitca.org	Internet Source	1%
4	fastercapital.com	Internet Source	1%
5	experts.umn.edu	Internet Source	<1%
6	Submitted to Coventry University	Student Paper	<1%
7	medinform.jmir.org	Internet Source	<1%
8	E. Parimbelli, S. Marini, L. Sacchi, R. Bellazzi. "Patient similarity for precision medicine: a systematic review", Journal of Biomedical Informatics, 2018	Publication	<1%
9	scholarworks.uaeu.ac.ae	Internet Source	<1%
10	Submitted to University of North Texas	Student Paper	<1%
11	www.irjmets.com	Internet Source	<1%
12	www.scribd.com	Internet Source	<1%
13	amds123.github.io	Internet Source	<1%
14	Dinesh Goyal, Bhanu Pratap, Sandeep Gupta, Saurabh Raj, Rekha Rani Agrawal, Indra Kishor. "Recent Advances in Sciences, Engineering, Information Technology & Management - Proceedings of the 6th International Conference "Convergence2024" Recent Advances in Sciences, Engineering, Information Technology & Management, April 24–25, 2024, Jaipur, India", CRC Press, 2025	Publication	<1%
15	business.uc.edu	Internet Source	<1%
16	donut.topology.rocks	Internet Source	<1%
17	www.kluniversity.in	Internet Source	<1%
18	bmcmedinformdecismak.biomedcentral.com	Internet Source	<1%
19	coek.info	Internet Source	<1%
20	assets-eu.researchsquare.com	Internet Source	<1%
21	arxiv.org	Internet Source	<1%
22	www.collegesinstitutes.ca	Internet Source	<1%
23	Poonam Tanwar, Tapas Kumar, K. Kalaiselvi, Haider Raza, Seema Rawat. "Predictive Data Modelling for Biomedical Data and", River Publishers, 2024	Publication	<1%
24	Submitted to The University of Manchester	Student Paper	<1%
25	Submitted to University of Hertfordshire	Student Paper	<1%
26	Submitted to University of Stirling	Student Paper	<1%
27	towardsdatascience.com	Internet Source	<1%
28	www.openpr.com	Internet Source	

		<1 %
29	lawyersnjurists.com Internet Source	<1 %
30	Suresh Pokharel, Guido Zuccon, Xue Li, Chandra Prasetyo Utomo, Yu Li. "Temporal Tree Representation for Similarity Computation between Medical Patients", Artificial Intelligence in Medicine, 2020 Publication	<1 %
31	Hui Cao, Genevieve B. Melton, Marianthi Markatou, George Hripcsak. "Use abstracted patient-specific features to assist an information-theoretic measurement to assess similarity between medical cases", Journal of Biomedical Informatics, 2008 Publication	<1 %
32	Rama Devi K, Syed Siraj Ahmed N, Ramchandran P, Parvathi S, Althaf Ali A. "Proactive Detection of Malicious Webpages Using Hybrid Natural Language Processing and Ensemble Learning Techniques", Journal of information and organizational sciences, 2024 Publication	<1 %
33	aquatera.ca Internet Source	<1 %
34	ebin.pub Internet Source	<1 %
35	epub.lib.aalto.fi Internet Source	<1 %
36	mdpi-res.com Internet Source	<1 %
37	"Software Engineering Methods in Systems and Network Systems", Springer Science and Business Media LLC, 2024 Publication	<1 %
38	Syed Nisar Hussain Bukhari. "Data-Driven Farming - Harnessing the Power of AI and Machine Learning in Agriculture", CRC Press, 2024 Publication	<1 %

Exclude quotes	Off	Exclude matches	Off
Exclude bibliography	On		

Project work mapping with SDG

- Our project is mapped to SDG-3,that is, Good health and Well Being.
- Our main goal is to cluster patients based on heart diseases, diabetes and other diseases and give similarity scores between historical and new patients.
- This will help in improving the diagnostic accuracy, predictive models and optimize the models.

Thank You