PREDICTIVE ANALYSIS FOR PERSONALIZED PATIENT CARE

NITHIN ISWAR P (2127200502025)
PRAJITH K (2127200502029)
AAKASH A (2127200502301)

AD18711 – MINI PROJECT

Viva Voce Examination

Batch No	: 14
Name of the Supervisor	: Ms. KIRUTHIGA E
Date of Review	: 08/11/2023
Domain	: MACHINE LEARNING

ABSTRACT

- In a large hospital setting, the challenge is to proactively identify patients at a higher risk of critical medical conditions during their stay.
- The aim is to enhance patient outcomes, reduce adverse events, and transform proactive healthcare management through critical condition prevention.
- Significantly advances patient care, efficient resource utilization, and notably reduces adverse events in the hospital environment.
- Multiple machine learning models like ANN, decision tree and XGBoost are implemented to analyze patient data, including age, gender, admission type, lifestyle, and medical test outcomes, to identify patients at high risk of critical illness and finding the optimal algorithm .
- The disease categories that can be predicted include cardiovascular diseases, respiratory diseases, neurological diseases, kidney diseases, and metabolic disorders
- A recommendation system is integrated to suggest appropriate treatments, and medicines for each patient, based on their predicted risk using VADER sentiment analysis and weighted average ratings calculation.

INTRODUCTION TO BASIC CONCEPTS

PREDICTIVE MODELING:

- Predictive modeling is a powerful tool for predicting future medical outcomes based on past data.
- Predictive modeling uses math and computations to predict future outcomes based on past data.
- It can be used to assess the risk of patients developing various medical conditions and inform early diagnosis and personalized treatment plans.
- A model is trained on patient data to learn patterns and make predictions for new patients.
- Multiple machine learning models are used, including decision trees, XGBoost, and artificial neural networks, to get optimal predictions.
- Predictive modeling is a significant advancement in proactive healthcare management, with the potential to improve patient care.

RECOMMENDATION SYSTEM:-

- Recommendation systems use data analysis to suggest relevant items to users based on their past behavior and preferences.
- Leveraging predictive modeling, recommendation system are used to suggest tailored treatment and medication plans based on a patient's individual needs and risks.
- It integrates a comprehensive dataset of disease-medicine mappings, and treatments to inform its recommendations.
- Recommendation system optimizes medical decisions based on accurate predictions, leading to improved outcomes.
- It represents a significant advancement in proactive healthcare management, enabling healthcare professionals to deliver more personalized and effective care.

LITERATURE REVIEW (1/6)

- Md. Imam Hossain, Mehadi Hasan Maruf, Md. Ashikur Rahman Khan, et al. (2022), "**Heart Disease Prediction Using Distinct Artificial Intelligence Techniques: Performance Analysis and Comparison**", Iran Journal of Computer Science, Springer, Vol. 6, No. 4, pp. 2520-8446

Problem Statement:

- The integration of technology and medical data analysis aims to enhance heart disease management. Accurate predictive models are crucial for proactive patient risk identification and improved care.
- The research explores diverse machine learning algorithms, focusing on heart disease prediction using patient data. Random Forest, with 90% accuracy, shows promise for early detection and better patient care.

Methodology:

- The base paper contributes to healthcare research by evaluating machine learning algorithms like decision trees, Gradient Boosting Regression, Naïve Bayes, random forests, support vector machines (SVMs), and neural networks for predictive modeling.
- Various machine learning algorithms, including decision trees and Gradient Boosting Regression, are employed to analyze heart-related data, aiming for higher prediction accuracy. Diverse data sources, including patient records and medical tests, are utilized to improve models and reduce computational costs in healthcare applications.

LITERATURE REVIEW (2/6)

- Deepjyoti Roy, Mala Dutta. (2022), "**A Systematic Review and Research Perspective On Recommender Systems**", Journal of Big Data, Springer, Vol. 9, pp. 2196-3115.

Problem Statement:

- Recommender systems are essential for filtering vast online information, but they face challenges like scalability, cold-start, and sparsity.
- The selection of techniques and understanding their features, advantages, and disadvantages are complex, hindering application-focused recommender system development.

Methodology:

- This paper conducts a systematic review of recent contributions in recommender systems across diverse applications, including books, movies, and products.
- It analyzes the applications of various recommender systems and performs an algorithmic analysis to frame a taxonomy for effective system development.
- The review also evaluates datasets, simulation platforms, and performance metrics used in each contribution, highlighting research gaps and challenges.

LITERATURE REVIEW (3/6)

- Pooja Rani, Rajneesh Kumar, Nada M. O. Sid Ahmed, Anurag Jain. (2022), "**A Decision Support System for Heart Disease Prediction based upon Machine Learning**", Journal of Reliable Intelligent Environments, Springer, Vol. 7, pp. 2199-4668

Problem Statement:

- Timely detection of heart disease is a critical challenge, particularly in areas without easy access to heart specialists.
- Developing an accurate decision support system is vital for early detection in remote and underserved regions.

Methodology:

- This research presents a hybrid decision support system for early heart disease detection by implementing multivariate imputation, hybridized feature selection, and SMOTE preprocessing techniques. it uses support vector machine, naive bayes, logistic regression, random forest, and adaboost classifiers
- Evaluates various machine learning classifiers, achieving 86.6% accuracy on the Cleveland heart disease dataset.

LITERATURE REVIEW (4/6)

- Dibaba Adeba Debal, Tilahun Melak Sitote. (2022), "**Chronic Kidney Disease Prediction Using Machine Learning Techniques**", Journal of Big Data, Springer, Vol. 9, pp. 2211-2218.

Problem Statement:

- Predictive modeling is vital in early disease detection, particularly in resource-constrained healthcare settings. Effective feature selection and robust algorithms are crucial for accurate predictions.
- The paper explores machine learning techniques, including decision trees, Gradient Boosting Regression, Naïve Bayes, random forests, support vector machines (SVMs), and neural networks, to enhance healthcare predictive modeling.

Methodology:

- The research evaluates these machine learning algorithms in analyzing Blood Urea Nitrogen and blood -related data for improved prediction accuracy. Leveraging patient records, medical tests, and lifestyle information enhances disease prediction models' precision.
- The paper also explores hybrid approaches, combining decision trees and Gradient Boosting Regression, to optimize resource allocation and reduce computational costs in healthcare dataset.

LITERATURE REVIEW (5/6)

- M. Chen, Y. Hao, K. Hwang, L. Wang and L. Wang. (2017). “**Disease Prediction by Machine Learning Over Big Data from Healthcare Communities**”, IEEE Access, Vol. 5, pp. 8869-8879.

Problem statement:

- Medical data analysis plays a vital role in early disease detection and patient care, but incomplete data quality can reduce analysis accuracy.
- Regional variations in certain diseases challenge disease outbreak predictions. This paper aims to improve the prediction of chronic disease outbreaks in areas with frequent disease occurrences by utilizing machine learning techniques and addressing incomplete data.

Methodology:

- The study uses machine learning models to predict chronic disease outbreaks in disease-prone communities.
- It employs a novel convolutional neural network (CNN)-based multimodal disease risk prediction algorithm, utilizing both structured and unstructured hospital data, and addresses the issue of incomplete data using a latent factor model for data reconstruction

LITERATURE REVIEW (6/6)

- Purushottam, K. Saxena and R. Sharma. (2015), "**Efficient Heart Disease Prediction System using Decision Tree**“, International Conference on Computing, Communication & Automation, IEEE Xplore, pp. 72-77.

Problem Statement:

- Diagnosing cardiac disease is challenging due to numerous attributes like glucose levels, blood pressure, and cholesterol. Data processing and deep learning algorithms show promise in disease prediction and classification, including heart disease.
- Heart disease is complex and requires cautious treatment. Data mining and medical science are essential for identifying metabolic syndromes and predicting heart disease, highlighting the need for accurate prediction methods.

Methodology:

- Trusted ECG data was collected from Kaggle, followed by preprocessing steps. Attributes affecting the target were selected, outliers removed, and the data split into training and test sets.
- Machine learning algorithms were employed for training and testing. However, specific algorithms used in this process are SVM, random forest and decision trees.

ISSUES AND CHALLENGES

- Combining outputs from diverse predictive models (Decision Trees, XGBoost, ANN) to generate unified predictions can be intricate.

- Sentiment analysis poses intricate challenges like context interpretation and sarcasm detection.

- Ensuring seamless coordination between the disease predictions and the treatment dataset, considering variances in disease names or codes, and aligning them accurately for effective recommendations.

PROBLEM STATEMENT

- In a healthcare setting, the challenge is to develop a robust predictive model that accurately identifies patients at risk of various medical conditions based on their health care medical counts data.
- This involves addressing issues related to user interface complexity, and ensuring consistent accuracy rates in predictive models.
- The ultimate goal is to enhance patient care by providing timely and accurate predictions for improved medical condition management.
- The disease categories that can be predicted include cardiovascular diseases, respiratory diseases, neurological diseases, kidney diseases, and metabolic disorders

PROPOSED WORK

- Cleaning and preprocessing the data, including handling missing values to ensure it's ready for analysis, whether for decision tree, XGBoost, ANN.
- The dataset will be divided into training, validation, and testing sets to support model development and evaluation for machine learning algorithms.
- Creating distinct multioutput decision tree models for various diseases, utilizing relevant features to predict specific disease outcomes.
- Employing the XGBoost algorithm, a robust gradient boosting technique, It fine-tunes model parameters and combines the results with other algorithms for improved disease prediction accuracy.
- Utilizing ANN as a key predictive model for disease identification. ANN is trained on patient data, including diverse features, to enhance prediction accuracy for various medical conditions.
- This system utilizes disease predictions to suggest personalized treatment and medication plans. It incorporates a comprehensive dataset of disease-medicine mappings, and treatments to enhance patient care. weighted average ratings calculation is used for recommendation

PROPOSED WORK

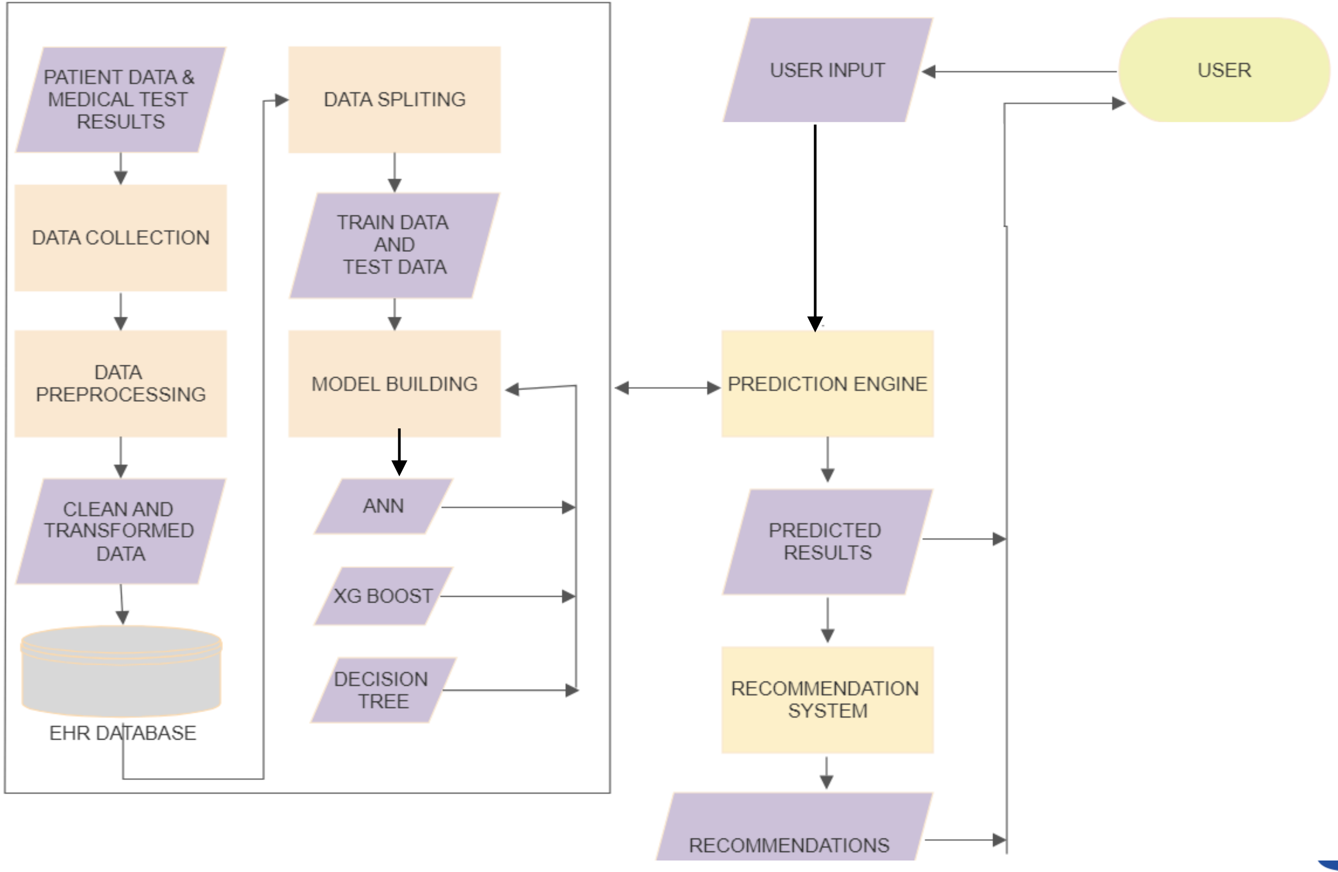
INPUT REPORTS:-

- DM - Diabetes Mellitus
- HTN - Hypertension (High Blood Pressure)
- CAD - Coronary Artery Disease
- PRIOR CM - Prior Cardiac Events
- CKD - Chronic Kidney Disease
- HB - Hemoglobin
- TLC - Total Leukocyte Count
- PLATELETS - Platelet Count
- GLUCOSE - Blood Glucose Level
- UREA - Blood Urea Level
- CREATINNE - Creatinine Level
- BNP - Brain Natriuretic Peptide
- RAISED CARDIAC ENZYMES - Elevated Cardiac Enzymes
- EF - Ejection Fraction

PREDICTABLE DISEASE

- Anemia
- Stable angina
- Acute Coronary Syndrome (ACS)
- ST-Elevation Myocardial Infarction (STEMI)
- Atypical chest pain
- Heart failure
- Heart failure with reduced ejection fraction (HEREF)
- Heart failure with preserved ejection fraction (HENEF)
- Valvular heart disease
- Complete heart block (CHB)
- Sick Sinus Syndrome (SSS)
- Acute Kidney Injury (AKI)
- Cerebrovascular Accident (CVA) - Ischemic Stroke
- Cerebrovascular Accident (CVA) - Hemorrhagic Stroke
- Atrial Fibrillation (AF)
- Ventricular Tachycardia (VT)
- Paroxysmal Supraventricular Tachycardia (PSVT)
- Congenital heart disease
- Urinary Tract Infection (UTI)
- Neurocardiogenic syncope

ARCHITECTURE DIAGRAM



SYSTEM REQUIREMENTS AND TOOLS

HARDWARE COMPONENTS:

- CPU – Intel i5 13th gen
- 16 GB RAM
- 512 GB SSD
- Windows 10 (64-bit)

SOFTWARE COMPONENTS:

- Python
- Machine Learning Framework
- IDE: Visual studio code

MODULES

- Data Extraction & Processing
- Data Visualization & Splitting
- Prediction Engine
- Recommendation System

DATA EXTRACTION & PREPROCESSING

Data Extraction:

- This phase involves gathering essential information from diverse healthcare sources, compiling a comprehensive dataset for analysis. For disease prediction the data is extracted from Kaggle and comprises 15,000 records and for recommendation system the data is also extracted from Kaggle and comprises of 18000 records

Data Preprocessing:

- In this phase, we ensure that the data is prepared for analysis, including compatibility with our machine learning models. This involves tasks like data cleaning and feature transformation.
- Identifying and removing rows with missing values doesn't significantly impact the prediction process because the dataset is reduced by only 400 values
- Features like EF and BNP does not provide sufficient data to contribute effectively to the predictive models, removing those columns simplifies the predictive modeling process.
- In the recommendation system dataset, it is necessary to adjust the data formatting and matching according to the predicted output

DATA VISUALIZATION & SPLITTING

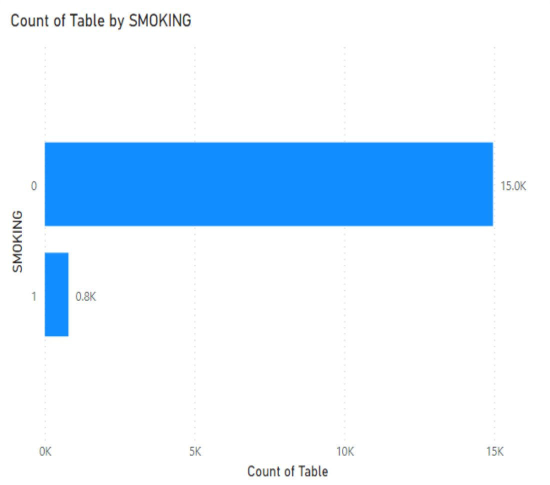
Data Visualization:

- Data visualization serves as a pivotal tool, providing graphical representations of healthcare data to facilitate:
- Visualizations help in recognizing recurring patterns within patient data, aiding in the identification of disease risk factors
- Visualizing patient health data over time allows you to discern trends that might not be apparent in raw data
- It helps in gaining insights into the effectiveness of different treatments, medication plans, and doctor recommendations.

Data Splitting:

- In healthcare predictive modeling, dividing the dataset into training, validation, and testing subsets is vital, presenting challenges and considerations.
- **Training Data:** Selecting the right training subset for the three models . It involves ensuring diversity in patient cases and handling class imbalances.
- **Testing Data:** It must be independent from training data and accurate model evaluation metrics.

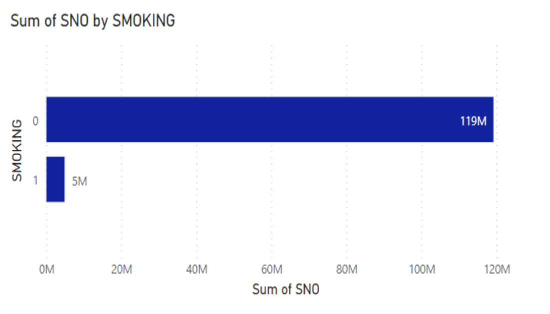
DATA VISUALIZATION & SPLITTING

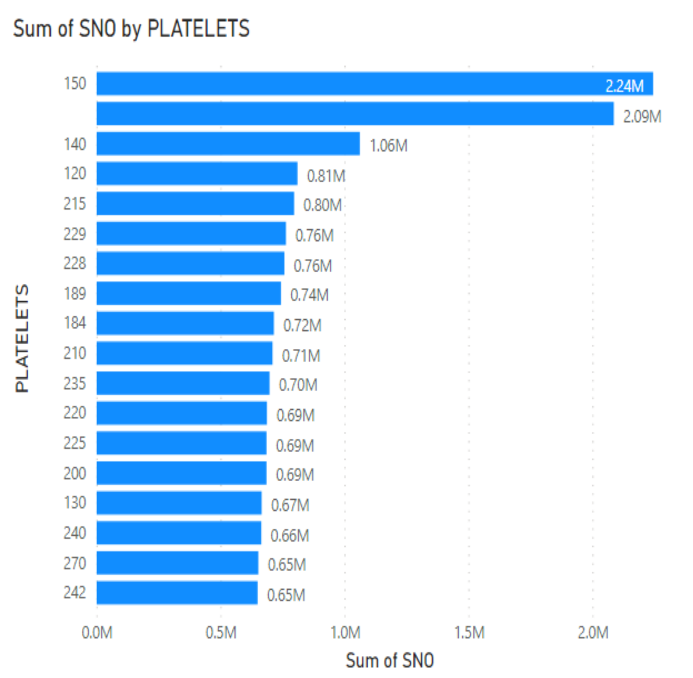


Count of Table for 0 (14,964) was higher than 1 (793).

0 accounted for 94.97% of Count of Table.

0 had 14,964 Count of Table and 1 had 793.





PREDICTION ENGINE

Decision Tree:

- Importing the Decision Tree Classifier, a powerful tool for creating predictive models based on the training data's patterns and relationships.
- For handling multi-label classification tasks, we are importing a MultiOutput Classifier, which extends the Decision Tree's capabilities to predict multiple output columns simultaneously.
- Training the Decision Tree model using the training data. The classifier learns from the input columns and their corresponding output columns to make predictions.
- Utilizing the MultiOutput Decision Tree Classifier to predict the output columns based on the provided input columns. This extends the Decision Tree's capabilities to handle multiple output variables.
- Applying the trained model to the test dataset to assess its predictive accuracy. This evaluation step ensures the model's reliability in making predictions on new, unseen data.

PREDICTION ENGINE

XGBoost:

- Importing the XGBoost algorithm, known for its robustness and effectiveness in predictive modeling.
- Employing the XGBoost algorithm to train the model using the training data. XGBoost is an ensemble learning method that sequentially builds multiple decision trees to refine predictions.
- Fine-tuning the model by adjusting hyperparameters to optimize its predictive performance. Parameters like max_depth, learning_rate , n_estimators and subsample are tuned with the help of GridSearchCV
- Utilizing the trained XGBoost model to predict the output columns based on the input features, effectively handling multiple output variables.
- Applying the model to the test dataset to assess its predictive accuracy, ensuring reliable predictions on unseen data.

PREDICTION ENGINE

Artificial Neural Network:

- Designing the architecture of the Artificial Neural Network. Specify the number of layers, neurons in each layer, and activation functions. Tailor the network structure to your specific disease prediction task.
- Training the ANN using the training dataset. ANN is a deep learning model that learns complex patterns from data through iterative training.
- Fine-tuning the ANN by adjusting hyperparameters like learning rates, batch sizes, and regularization techniques to enhance its predictive performance.
- Leveraging the trained ANN to make predictions on the output variables based on the input features. ANN can handle multiple output variables simultaneously.
- Evaluating the ANN's predictive accuracy using the test dataset, ensuring it performs reliably on unseen data.
- **Optimal Algorithm**: Upon comparing the performance of Artificial Neural Networks (ANN) with that of Decision Tree and XG Boost, it becomes evident that ANN outperforms the other two in terms of accuracy. Consequently, ANN emerges as the most optimal algorithm for this project, exhibiting superior predictive capabilities.

RECOMMENDATION SYSTEM

- The VADER sentiment analysis tool is applied to the 'review' column, resulting in the computation of compound, positive, negative, and neutral sentiment scores for each review. The sentiment scores are stored in separate lists, and reviews are categorized into 'Positive,' 'Neutral,' or 'Negative' based on these scores.
- The data originating from the sentiment analysis and the original dataset are merged into a unified dataset for streamlined analysis. Columns associated with sentiment analysis are dropped from the dataset. The dataset is sorted, cleaned, and a new column 'Review Sentiment' is introduced for the purpose of categorizing reviews.
- Weighted average ratings for drugs are computed, with a focus on the 'useful Count' column as a key factor. The calculated weighted ratings are integrated with the drug names in the dataset.
- The code groups the data by condition (disease) and identifies the top-rated drugs for each condition based on weighted average ratings and 'useful Count.’ For a given disease, it suggests the three drugs with the highest weighted average ratings.

IMPLEMENTATION SCREENSHOTS

- **DATA PREPROCESSING & CLEANING:-**

```
valid_columns = df.columns[null_pct < .06]
```

```
valid_columns
```

```
Index(['SNO', 'MRD No.', 'D.O.A', 'D.O.D', 'AGE', 'GENDER', 'RURAL',
       'TYPE OF ADMISSION-EMERGENCY/OPD', 'month year', 'DURATION OF STAY',
       'duration of intensive unit stay', 'OUTCOME', 'SMOKING ', 'ALCOHOL',
       'DM', 'HTN', 'CAD', 'PRIOR CMP', 'CKD', 'HB', 'TLC', 'PLATELETS',
       'GLUCOSE', 'UREA', 'CREATININE', 'RAISED CARDIAC ENZYMES',
       'SEVERE ANAEMIA', 'ANAEMIA', 'STABLE ANGINA', 'ACS', 'STEMI',
       'ATYPICAL CHEST PAIN', 'HEART FAILURE', 'HFREF', 'HFNEF', 'VALVULAR',
       'CHB', 'SSS', 'AKI', 'CVA INFRACT', 'CVA BLEED', 'AF', 'VT', 'PSVT',
       'CONGENITAL', 'UTI', 'NEURO CARDIOGENIC SYNCOPE', 'ORTHOSTATIC',
       'INFECTIVE ENDOCARDITIS', 'DVT', 'CARDIOGENIC SHOCK', 'SHOCK',
       'PULMONARY EMBOLISM', 'CHEST INFECTION'],
      dtype='object')
```

IMPLEMENTATION SCREENSHOTS

- **DEFINING INPUT AND OUTPUT COLUMNS:-**

```
In [74]: indep
```

Out[74]:

	AGE	GENDER	SMOKING	ALCOHOL	DM	HTN	CAD	PRIOR CMP	CKD	HB	TLC	PLATELETS	GLUCOSE	UREA	CREATININE	RAISED CARDIAC ENZYMES
0	81	M	0	0	1	0	0	0	0	9.5	16.1	337	80	34	0.9	1
1	65	M	0	1	0	1	1	0	0	13.7	9	149	112	18	0.9	0
2	53	M	0	0	1	0	1	0	0	10.6	14.7	329	187	93	2.3	0
3	67	F	0	0	0	1	1	0	0	12.8	9.9	286	130	27	0.6	0
4	60	F	0	0	0	1	0	1	0	13.6	9.1	26	144	55	1.25	0
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
15751	86	F	0	0	1	1	1	0	0	8.8	13.7	361	131	57	1.4	1
15752	50	M	0	0	1	1	0	1	0	13.2	15.6	142	248	94	1.8	0
15753	82	M	0	0	0	1	1	0	0	9.3	11.7	372	210	67	1.9	0
15754	59	F	0	0	0	1	1	0	0	13.1	12.5	431	153	29	0.8	0
15755	59	F	0	0	0	1	1	0	0	13.1	12.5	431	153	29	0.8	0

15756 rows × 16 columns

- **DEFINING INPUT AND OUTPUT COLUMNS:-**

```
dept = df[['SEVERE ANAEMIA', 'ANAEMIA', 'STABLE ANGINA', 'ACS', 'STEMI', 'ATYPICAL CHEST PAIN', 'HEART FAILURE', 'HFREF', 'HFNEF'
```

```
dept
```

	SEVERE ANAEMIA	ANAEMIA	STABLE ANGINA	ACS	STEMI	ATYPICAL CHEST PAIN	HEART FAILURE	HFREF	HFNEF	VALVULAR	...	CONGENITAL	UTI	NEURO CARDIOGENIC SYNCOPE	ORTHOSTATIC	E
0	0	1	0	1	0	0	1	1	0	0	...	0	0	0	0	
1	0	0	0	0	0	0	0	0	0	0	...	0	0	0	0	
2	0	0	0	0	0	0	1	1	0	0	...	0	0	0	0	
3	0	0	0	0	0	0	0	0	0	0	...	0	0	0	0	
4	0	0	0	0	0	0	0	0	0	0	...	0	0	0	0	
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	
15751	0	1	0	1	0	0	0	0	0	0	...	0	0	0	0	
15752	0	0	0	0	0	0	0	0	0	0	...	0	0	0	0	
15753	0	1	0	0	0	0	1	0	1	0	...	0	0	0	0	
15754	0	0	0	0	0	0	0	0	0	0	...	0	0	0	0	
15755	0	0	0	0	0	0	0	0	0	0	...	0	0	0	0	

15756 rows × 28 columns

IMPLEMENTATION SCREENSHOTS

DECISION TREE:-

```
Enter your age (eg: 19,56): 20

Mention whether you are smoker are not (eg: no,yes): yes

Mention whether you are alcoholic or not (eg: no,yes): yes

Mention whether you have diabetes or not (eg: no,yes): yes

Mention whether you have hypertension or not (eg: no,yes): no

Mention whether you have coronary artery disease or not (eg: no,yes): yes

Mention whether you have prior cardiac events or not (eg: no,yes): yes

Mention whether you have kidney disease or not (eg: no,yes): yes

Mention your hemoglobin count (eg: 13.7): 5.6

Mention your leukocyte count (eg: 14.7): 12.3

Mention your platelets count(eg: 334): 253

Mention your glucose count (eg: 144): 134

Mention your urea count (eg: 55): 46

Mention your creatinine count (eg: 1.4): 1.4

Mention whether you have raised cardiac enzymes or not (eg: no,yes): yes

Mention your gender (eg: male,female): male
```

```
You don't have Severe_Anemia.
You don't have Anaemia.
You don't have Stable_Angina.
You don't have Acute_Coronary_Syndrome.
You don't have STEMI.
You don't have Atypical_chest_pain.
You don't have Heart_Failure.
You don't have Heart_Failure_Reduced_EF.
You don't have Heart_Failure_preserved_EF.
You don't have Valvular_Heart_Disease.
You don't have Complete_Heart_Block.
You don't have Sick_Sinus_Syndrome.
You don't have Acute_Kidney_Injury.
You don't have Ischemic_Stroke.
You don't have Hemorrhagic_Stroke.
You don't have Atrial_Fibrillation.
You don't have Ventricular_Tachycardia.
You don't have PSVT.
You don't have Congential_heart_Disease.
You don't have Urinary_Tract.
You don't have Neuro_Cardiogenic_Syncope.
You don't have Orthostatic.
You don't have Infective_Endocarditis.
You don't have DVT.
You don't have Cardiogenic_Shock.
You don't have Shock.
You don't have Pulmonary_Embolism.
You don't have Chest_Infection.
```

```
[92]  model.score(Indep_n,dept)

...   0.8274308200050774
```

IMPLEMENTATION SCREENSHOTS

XG BOOST:-

```
Enter your age (eg: 19,56): 20
Mention whether you are smoker are not (eg: no,yes): yes
Mention whether you are alcoholic or not (eg: no,yes): yes
Mention whether you have diabetes or not (eg: no,yes): yes
Mention whether you have hypertension or not (eg: no,yes): no
Mention whether you have coronary artery disease or not (eg: no,yes): yes
Mention whether you have prior cardiac events or not (eg: no,yes): yes
Mention whether you have kidney disease or not (eg: no,yes): yes
Mention your hemoglobin count (eg: 13.7): 5.6
Mention your leukocyte count (eg: 14.7): 12.3
Mention your platelets count(eg: 334): 253
Mention your glucose count (eg: 144): 134
Mention your urea count (eg: 55): 46
Mention your creatinine count (eg: 1.4): 1.4
Mention whether you have raised cardiac enzymes or not (eg: no,yes): yes
Mention your gender (eg: male,female): male
```

```
You don't have Severe_Anemia.
You don't have Anaemia.
You don't have Stable_Angina.
You don't have Acute_Coronary_Syndrome.
You don't have STEMI.
You don't have Atypical_chest_pain.
You don't have Heart_Failure.
You don't have Heart_Failure_Reduced_EF.
You don't have Heart_Failure_preserved_EF.
You don't have Valvular_Heart_Disease.
You don't have Complete_Heart_Block.
You don't have Sick_Sinus_Syndrome.
You don't have Acute_Kidney_Injury.
You don't have Ischemic_Stroke.
You don't have Hemorrhagic_Stroke.
You don't have Atrial_Fibrillation.
You don't have Ventricular_Tachycardia.
You don't have PSVT.
You don't have Congential_heart_Disease.
You don't have Urinary_Tract.
You don't have Neuro_Cardiogenic_Syncope.
You don't have Orthostatic.
You don't have Infective_Endocarditis.
You don't have DVT.
You don't have Cardiogenic_Shock.
You don't have Shock.
You don't have Pulmonary_Embolism.
You don't have Chest_Infection.
```

```
[107] model.score(Indep_n,dept)
... 0.5632140137090632
```

IMPLEMENTATION SCREENSHOTS

ANN:-

Enter your age (eg: 19,56): 20

Mention whether you are smoker are not (eg: no,yes): yes

Mention whether you are alcoholic or not (eg: no,yes): yes

Mention whether you have diabetes or not (eg: no,yes): yes

Mention whether you have hypertension or not (eg: no,yes): no

Mention whether you have coronary artery disease or not (eg: no,yes): yes

Mention whether you have prior cardiac events or not (eg: no,yes): yes

Mention whether you have kidney disease or not (eg: no,yes): yes

Mention your hemoglobin count (eg: 13.7): 5.6

Mention your leukocyte count (eg: 14.7): 12.3

Mention your platelets count(eg: 334): 253

Mention your glucose count (eg: 144): 134

Mention your urea count (eg: 55): 46

Mention your creatinine count (eg: 1.4): 1.4

Mention whether you have raised cardiac enzymes or not (eg: no,yes): yes

Mention your gender (eg: male,female): male

```
You don't have Severe_Anemia.
You don't have Anaemia.
You don't have Stable_Angina.
You don't have Acute_Coronary_Syndrome.
You don't have STEMI.
You don't have Atypical_chest_pain.
You don't have Heart_Failure.
You don't have Heart_Failure_Reduced_EF.
You don't have Heart_Failure_preserved_EF.
You don't have Valvular_Heart_Disease.
You don't have Complete_Heart_Block.
You don't have Sick_Sinus_Syndrome.
You don't have Acute_Kidney_Injury.
You don't have Ischemic_Stroke.
You don't have Hemorrhagic_Stroke.
You don't have Atrial_Fibrillation.
You don't have Ventricular_Tachycardia.
You don't have PSVT.
You don't have Congential_heart_Disease.
You don't have Urinary_Tract.
You don't have Neuro_Cardiogenic_Syncope.
You don't have Orthostatic.
You don't have Infective_Endocarditis.
You don't have DVT.
You don't have Cardiogenic_Shock.
You don't have Shock.
You don't have Pulmonary_Embolism.
You don't have Chest_Infection.
```

```
model.score(indep,dept)
49]
1.0
```

IMPLEMENTATION

You may be affected with Severe_Anemia.
You may be affected with Anaemia.
You may be affected with Heart_Failure.
You may be affected with Heart_Failure_Reduced_EF.
You may be affected with Heart_Failure_preserved_EF.
You may be affected with Cardiogenic_Shock.
You may be affected with Shock.
You may be affected with Chest_Infection.

Recommended drugs for Severe_Anemia, Anaemia:

Integra
13 - found it usefull

Ferralet 90
9 - found it usefull

Integra F
29 - found it usefull

Recommended drugs for Heart_Failure, Heart_Failure_Reduced_EF, Heart_Failure_preserved_EF:

Corlanor
6 - found it usefull

Ivabradine
6 - found it usefull

Warfarin
7 - found it usefull

Recommended drugs for Cardiogenic_Shock:

Nitro-Bid
8 - found it usefull

Streptokinase
2 - found it usefull

Propranolol
11 - found it usefull

Recommended drugs for Shock:

Cisplatin
2 - found it usefull

Platinol
2 - found it usefull

Afinitor
0 - found it usefull

Recommended drugs for Chest_Infection:

Cedax
6 - found it usefull

Ceftibuten
6 - found it usefull

Ceftriaxone
2 - found it usefull

CONCLUSION

- The project successfully achieved its primary objective of developing a robust disease prediction system using machine learning techniques, particularly Artificial Neural Networks (ANN), which demonstrated superior accuracy compared to other algorithms.
- The implementation of a sentiment analysis-driven recommendation system categorizing patient reviews into 'Positive,' 'Neutral,' or 'Negative' sentiments enhances personalized drug recommendations and resource allocation for improved patient care.
- The project's success lays a strong foundation for proactive healthcare management, emphasizing innovation and data-driven insights as key elements for advancing patient well-being and healthcare outcomes.
- Future work includes continuous data collection, real-time patient data integration, and the incorporation of natural language processing techniques to better understand patient sentiment and provide more nuanced treatment suggestions.
- The integration of electronic health records (EHR) and wearable device data, along with collaboration with healthcare providers and institutions, can enable early detection of health issues and the translation of research into real-world healthcare improvements.

REFERENCES

- M. Chen, Y. Hao, K. Hwang, L. Wang and L. Wang (2017). “Disease Prediction by Machine Learning Over Big Data from Healthcare Communities”, IEEE Access, Vol. 5, pp. 8869-8879.
- Deepjyoti Roy, Mala Dutta. (2022), “A Systematic Review and Research Perspective On Recommender Systems”, Journal of Big Data, Springer, Vol. 9, pp. 2196-1115.
- Dibaba Adeba Debal, Tilahun Melak Sitote. (2022), “Chronic Kidney Disease Prediction Using Machine Learning Techniques”, Journal of Big Data, Springer, Vol. 9, pp. 2211-2218.
- S. Garg. (2021), “Drug Recommendation System based on Sentiment Analysis of Drug Reviews using Machine Learning”, 11th International Conference on Cloud Computing, Data Science and Engineering (Confluence), Noida, India, pp. 175-181.
- S. Grampurohit, C. Sagarnal. (2020), “Disease Prediction using Machine Learning Algorithms”, 2020 International Conference for Emerging Technology, Belgaum, India, pp. 1-7.
- Md. Imam Hossain, Mehadi Hasan Maruf, Md. Ashikur Rahman Khan, et al. (2022), “Heart Disease Prediction Using Distinct Artificial Intelligence Techniques: Performance Analysis and Comparison”, Iran Journal of Computer Science, Springer, Vol. 6, No. 4, pp. 2520-8446

THANK YOU