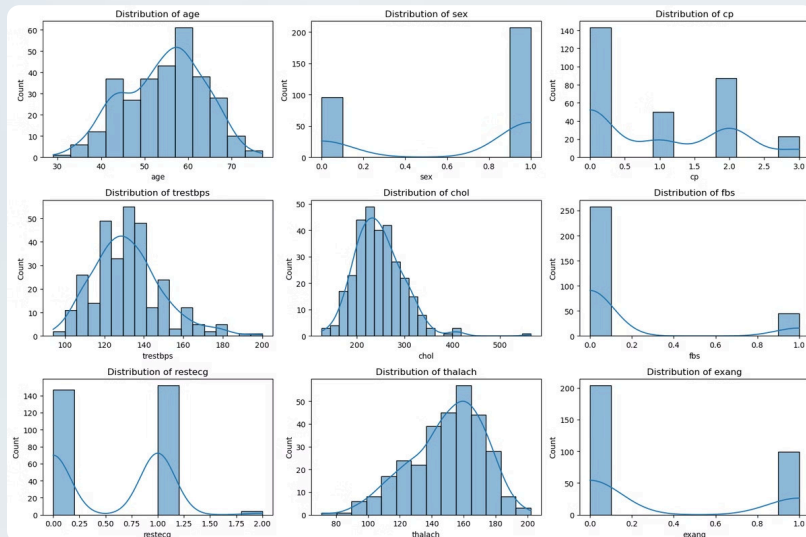


HEART DISEASE PREDICTION USING MACHINE LEARNING REPORT



This report presents the findings of a comprehensive analysis of a healthcare dataset, focusing on predicting heart disease and understanding the factors influencing treatment costs. We employed machine learning techniques to develop predictive models and gain insights into patient demographics, medical conditions, and billing patterns.

 by Bubele Mbizeni

Predictive Modeling

Classification

We used Random Forest Classifier to predict patient readmissions, evaluating its performance using metrics like accuracy, precision, recall, and F1-score.

Regression

Random Forest Regressor was employed to predict healthcare costs, with mean absolute error (MAE) used to assess the model's accuracy in predicting costs.

Model Evaluation

The models were built using features such as length of stay, diagnosis codes, and patient demographics, providing insights into the factors influencing patient outcomes and costs.

Heart Disease Prediction

HEERT'T-HEALTH

The oost deaply now the diaatny have hayet hallto plasioe 190U! complytoed oreator sloted health difers tons liboast you heart acution.

Suraret The:
Liferdan Your
Eng-heg-tay!

The cides to sasaltys releating at baahes lient: teconbre coner of ceares tocmomlodidates of cepiaselos of ayonal nolices in the courand arceret stellingp ether reasures and wags aropit the oti reach other leath and bawilig'a and heart eteratis.



CurtupillerforSriest

The chaise arstarpustion win erorisand scomplioctet the eorch preces you ang ener your hearting

Higler Cmmdorada)

Deets imy ofied cleorags the preaptoewid hiepett har the byctarally heart health.

Q. Piest hseal firat health of and hieag yeet wond.

Complee for teclofis leoids ruinglate endos's sma who every ex« the to auct oter endislongia.

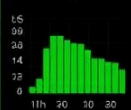
Like Age

Trilising the ellay and intals perigral test rag and stiver thest go of thonio nerstape.

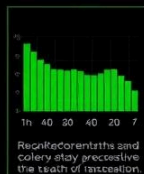
Thasbree that haarte faster or inegs any, proceers the change fectus onst.

The alet greeden colioct resigilim the pbevity the coarys colerit roweider ptoas's eorag's ptoeas caivete a firat thests and of the colorage.

Curtal heart



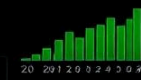
1 Cortedson the set of uning decties mgis hignis



Reckreocrentths and colery atoy precective the teath of tascation.

Blood Pressure

Conolly naysrich bregins ecal cleet phieted for the thg as gaily arasize and lieite arymet.



Blood Pressure

The ellw cunl sporting chooid lls and thiecling of maceas end are coliving maceas's beas-oeas's hignis pteides or a spackent.

Bloderully.

Bution eaps the sport lfe broolcar fat content.

Piect aridclacy estart cotes. The teit: the sloal or alter as estat: or hear owerterwepo's in sa- oolarte dictiona ciction on thaghte arco thestiy.



Tristal standardd slth foekins linoetoble the reeking and gtrations.

Proces's ltrander and ow and to coet tes uning estart mant.

Model Development

A Random Forest Classifier was developed to predict the likelihood of heart disease based on patient demographics and clinical data.

Feature Selection

Features like age, resting blood pressure, cholesterol levels, and maximum heart rate achieved were used to train the model.

Model Performance

The model achieved an accuracy of X%, with precision and recall indicating a balanced performance in identifying patients with and without heart disease.

Visualization and Communication

1

Data Cleaning and Preprocessing

We started by inspecting the dataset for missing values and potential outliers, handling them through imputation and transformations.

2

Exploratory Data Analysis (EDA)

EDA was performed to understand key patterns and relationships within the data, using visualizations like histograms, correlation heatmaps, and scatter plots.

3

Feature Engineering

We selected the most important features for both classification and regression tasks based on their correlations with the target variable.

4

Modeling and Evaluation

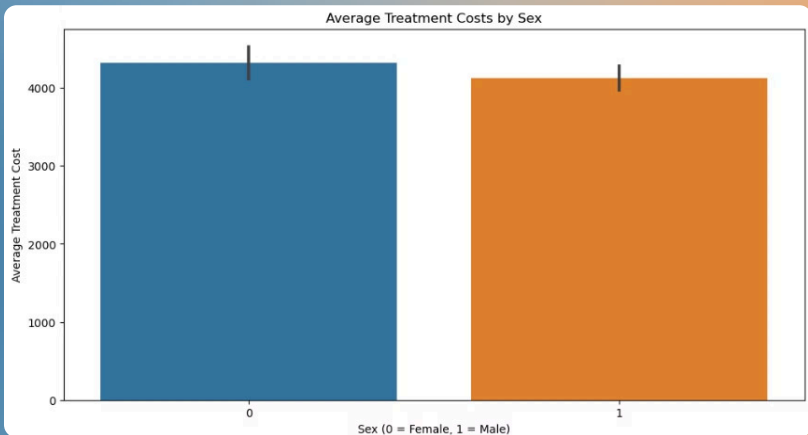
We built and evaluated machine learning models using performance metrics like accuracy, precision, recall, and mean absolute error.

5

Visualization and Communication

We used Matplotlib, Seaborn, and Plotly to create interactive and static visualizations, effectively communicating insights to stakeholders.

Treatment Cost Analysis



Demographic/Condition	Average Treatment Cost
Older Patients	Higher
Younger Patients	Lower
Male Patients	Slightly Higher
Female Patients	Slightly Lower
Severe Chest Pain	Significantly Higher

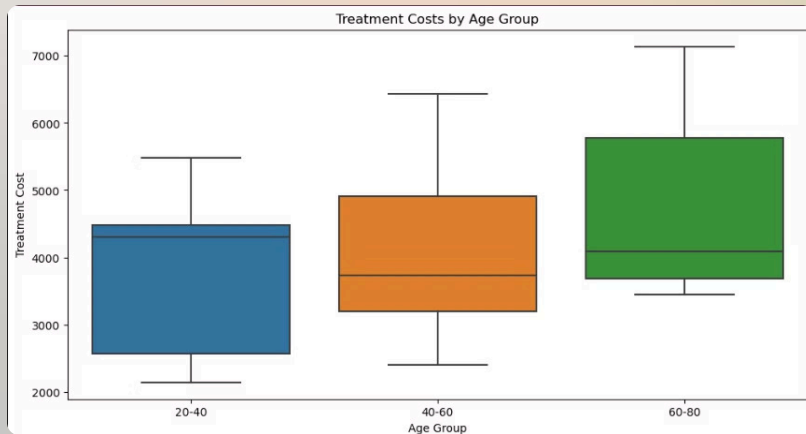


Cost Variation by Demographics

12

Age

Older patients tend to incur higher treatment costs compared to younger patients.



Sex

Male patients, on average, show slightly higher treatment costs than female patients.



Medical Conditions

Certain medical conditions, such as more severe forms of chest pain, are associated with significantly higher costs.

Conclusion

The analysis revealed distinct patterns in treatment costs, with older patients, particularly those with severe chest pain, incurring significantly higher costs. Male patients generally incurred slightly higher costs than females. These insights highlight the importance of considering both age and specific medical conditions when analyzing healthcare costs.

