University of Applied Sciences

THESIS – BACHELOR'S DEGREE PROGRAMME
TECHNOLOGY, COMMUNICATION AND TRANSPORT

PRESCRIPTIVE ANALYTICS IN HEALTHCARE:

Advanced Decision Making for Optimal Treatment

AUTHOR/S Daniel Fernandes

SAVONIA UNIVERSITY OF APPLIED SCIENCES

THESIS
Abstract

Field of Study Technology, Communication and Transport			
Degree Programme Degree Programme in Information Technology, Internet of Things			
Author(s) Daniel Fernandes			
Title of Thesis Prescriptive Analysis in Healthcare: Advanced Decision Making For Optimal Treatment			
Date	30-05-2024	Pages/Number of appendices	34
Client Organisation /Partners			

Abstract

This thesis explored the potential of prescriptive analytics to revolutionize healthcare decision-making. It began by highlighting the limitations of traditional, one-size-fits-all treatment approaches and the growing emphasis on personalized medicine based on patient data. The research addressed the challenges and opportunities of implementing prescriptive analytics in healthcare.

A proof-of-concept prescriptive model for drug classification was developed using a decision tree algorithm trained on a sample dataset. The model considered five patient characteristics: age, sex, blood pressure, cholesterol level, and sodium-to- potassium ratio. It demonstrated the potential for prescriptive analytics to provide tailored treatment recommendations based on individual patient data.

The thesis also addressed the ethical considerations and data privacy concerns surrounding prescriptive analytics in healthcare. Mitigating data bias was identified as a critical factor for ensuring the fairness and effectiveness of these models.

The research concluded by emphasizing the transformative potential of prescriptive analytics to improve diagnostic accuracy, optimize treatment plans, and ultimately enhance patient outcomes. It highlighted the need for further research to address implementation challenges, ensure data privacy, and continuously improve the accuracy and precision of prescriptive models. Overall, the thesis suggests that prescriptive analytics has the potential to revolutionize healthcare delivery by enabling more personalized, efficient, and effective patient care.

Keywords

Healthcare, AI, Prescriptive Analytics, Decision-Making, Optimal Treatment, Algorithms, Simulation, Decision Tree, Machine Learning, Techniques, Ethics, Bias

CONTENTS

LIST OF FIGURES

1 INTRODUCTION

This chapter introduces the evolution of medicine prescription, from traditional methods to modern, data-driven approaches. It discusses how healthcare decision-making has shifted from reliance on clinical experience to leveraging advanced analytics, enhancing personalized patient care. The chapter also outlines the thesis objectives, focusing on the development and implications of prescriptive analytics in healthcare.

1.1 Traditional Medicine Prescription

Throughout the decades, the way medical practitioners approached prescription of medicine has seen an impressive evolution, progressing from a mainly reactive approach to a more proactive and personalized one. Often the method of traditional medicine prescription was a one-size-fits-all formula simply based on the patient symptoms rather than on personalized patient attributes. (Naithani et al. 2021.)

Medical practitioners completely depended on academical learnings and their clinical experience making prescriptions just based on common symptoms and population trends which obviously is not sufficient for optimal treatment. As healthcare technology improves and we gather more information for analytics about patients, doctors are changing how they decide what medicine to prescribe. Now, they rely heavily on facts and data to figure out the best treatment for each person. (Naithani et al. 2021.)

This helps them predict specific health problems depending on a lot of analyzed data and data patterns, choose the right treatments, and improve how patients feel. Looking forward, this approach could mean even more personalized care, with treatments tailored specifically to each individual, ultimately making healthcare better for everyone. (Naithani et al. 2021.)

1.2 Background of Decision Making in Healthcare

In the past, doctors relied heavily on their clinical experience and intuition to predict, diagnose, and prescribe treatments for patients. Diagnostic tools were limited, often relying on physical examinations and basic laboratory tests. Treatment decisions were largely based on general medical knowledge and trial-and-error approaches. However, as modern medicine emerged and technological advancements revolutionized healthcare, doctors gradually gained access to sophisticated medical equipment and diagnostic techniques. This ushered in an era of more precise diagnosis and targeted treatments. (Alexopulos et al. 2021.)

In recent years, the integration of healthcare analytics has further transformed medical practice. With the vast amount of patient data now available, medical practitioners are leveraging predictive analytics algorithms to anticipate health conditions, refine diagnostic accuracy, and personalize treatment plans. This shift towards data-driven decision-making in healthcare represents a significant evolution in medical practice, empowering clinicians to optimize patient care and improve overall health outcomes. (Alexopulos et al. 2021.)

1.3 Research Objectives

This thesis aims to address three key objectives within the realm of prescriptive analytics in healthcare. Firstly, it seeks to investigate the challenges and opportunities associated with implementing prescriptive analytics in healthcare settings. Secondly, it endeavors to develop and assess a prescriptive analytics model specifically designed for healthcare decision-making. This model will utilize patient data encompassing profiles, medical histories, and treatment responses to demonstrate its feasibility and effectiveness. Lastly, the thesis aims to explore the ethical considerations and data privacy concerns inherent in the deployment of prescriptive analytics in healthcare.

By addressing these objectives, the research endeavors to contribute to a deeper understanding of the implications and potential benefits of prescriptive analytics in healthcare.

Additionally, the following questions will guide the investigation:

a) What does the evolution of prescriptive analytics look like in healthcare and why is it becoming increasingly important for our future?
b) How can prescriptive analytics models be developed for optimal treatment decision-making in healthcare?
c) How can prescriptive analytics be utilized to predict optimal treatment pathways for various medical conditions, and what factors influence the accuracy and efficiency of these predictions?
d) What ethical considerations and data privacy concerns arise from deploying prescriptive analytics in healthcare?

These questions will serve as guiding principles in exploring the practical implications and ethical implications of implementing prescriptive analytics in healthcare settings.

1.4 Thesis Structure

The structure of this thesis is designed to provide a comprehensive exploration of prescriptive analytics in healthcare decision-making.

In the Chapter 1, traditional medicine prescription practices are examined alongside the background of healthcare decision-making to establish a foundational understanding of the topic. Subsequently, the research objectives are outlined to delineate the specific aims and focus areas of this study.

The subsequent sections delve into the fundamental concepts and applications of prescriptive analytics in healthcare decision-making. "Understanding Healthcare Decision Making Analytics" pro-vides an overview of healthcare analytics, emphasizing its role in decision-making processes and highlighting challenges in traditional approaches. This sets the stage for a deeper exploration of prescriptive analytics in "Fundamentals of Prescriptive Analytics," where concepts, types, techniques, and applications specific to healthcare are elucidated.

The core of the thesis lies in the development and evaluation of a prescriptive model for drug prescription to present a proof of concept. In "Building a Prescriptive Model for Drug Prescription," the

process of model development is detailed, encompassing dataset selection, preparation, methodological approaches, and the implementation of decision tree algorithms for drug prescription.

Following the model development, the integration of prescriptive analytics into clinical practice is examined in "Integration of Prescriptive Model into Clinical Practice." This section explores strategies for integrating the model with existing healthcare systems, addresses challenges in implementation, and emphasizes the importance of ethical considerations and compliance.

A comparative analysis between prescriptive models and traditional treatment decision-making is presented in "Comparison with Traditional Treatment Decision Making," discussing the ad-vantages, limitations, and potential impact on patient outcomes of each approach.

Finally, "Future Directions and Implications" extrapolates on the advancements, potential applications, and implications of prescriptive analytics beyond drug prescription, highlighting the implications for healthcare providers and patients.

The thesis concludes with a summary of findings and contributions to healthcare decision-making, encapsulating the insights gained throughout the study and providing avenues for future research and application in the field.

2 UNDERSTANDING HEALTHCARE DECISION MAKING ANALYTICS

This chapter explores healthcare decision-making analytics, emphasizing the industry's digital transformation through Electronic Health Records (EHR) and big data (large and complex datasets that are difficult to analyze using traditional methods) (Tiao 2024). It examines the four types of analytics—descriptive, diagnostic, predictive, and prescriptive—and their roles in improving healthcare delivery and patient outcomes. The chapter also addresses the advantages and challenges associated with integrating analytics into healthcare decision-making processes.

2.1 Overview of Healthcare Analytics for Decision Making

The healthcare industry in undergoing a digital transformation with EHRs and big data playing an increasingly important role. Big data analytics, which refers to the analysis of large and complex datasets, can be used to improve healthcare delivery and patient outcomes. There are four main types of analytics used in healthcare: descriptive, diagnostic, predictive, and prescriptive (Cyrus 2023).

2.2 Challenges in Traditional Decision-Making Process

In traditional healthcare, patients relied solely on physicians for decisions, leading to a passive role in their own care. However, empowered e-patients now seek active involvement, challenging this paternalistic model. Shared decision-making has emerged as a crucial aspect, particularly in managing chronic conditions, emphasizing cooperation between physicians and patients. Physicians are transitioning from authoritative figures to guides, collaborating with patients in navigating healthcare. This cultural shift underscores the need for a fundamental transformation in healthcare, accommodating the evolving roles of patients and caregivers in decision-making processes. (Meskó et al. 2017.)

2.3 Role of Analytics in Decision Making

Analytics, the structured investigation of past and current data to gain valuable data to draw conclusions and support decision-making (ThoughtSpot 2024), can play a critical role in improving healthcare delivery and patient outcomes. By using analytics, healthcare organizations can identify trends, diagnose problems, predict future events, and take steps to improve care. For example, analytics can be used to identify patients who are at high risk of developing a chronic disease, allowing for early intervention and prevention. Additionally, analytics can be used to track the effectiveness of different treatments, allowing for the development of more effective care plans. Overall, analytics is a powerful tool that can be used to transform healthcare. Along with all the benefits and advantages that Artifical Intelligence (AI) and analytics in healthcare has to offer, there are also some cons. (Sruthi 2024.)

2.3.1 Types of Analytics in Decision-Making

The four main types of analytics used in healthcare:

Descriptive Analytics

This type of analytics helps to identify trends and patterns in healthcare data. For example, a healthcare organization could use descriptive analytics to track the number of patients admitted to the hospital each week. (Cyrus 2023.)

Diagnostic Analytics

This type of analytics goes beyond simply identifying trends and seeks to explain why those trends are occurring. For example, a healthcare organization might use diagnostic analytics to investigate the causes of hospital readmissions. (Cyrus 2023.)

Predictive Analytics

This type of analytics uses historical data to predict future trends. For example, a healthcare organization could use predictive analytics to identify patients who are at high risk of developing a chronic disease. (Cyrus 2023.)

Prescriptive Analytics

This is the most advanced type of analytics, and it uses data to recommend specific treatment plans to help heathcare providers to take better decisions which enchances the quality of care and provides better healthcare outcomes. For example, a healthcare organization might use prescriptive analytics to develop a care plan for a patient who is at high risk of hospitalization. (Cyrus 2023.)

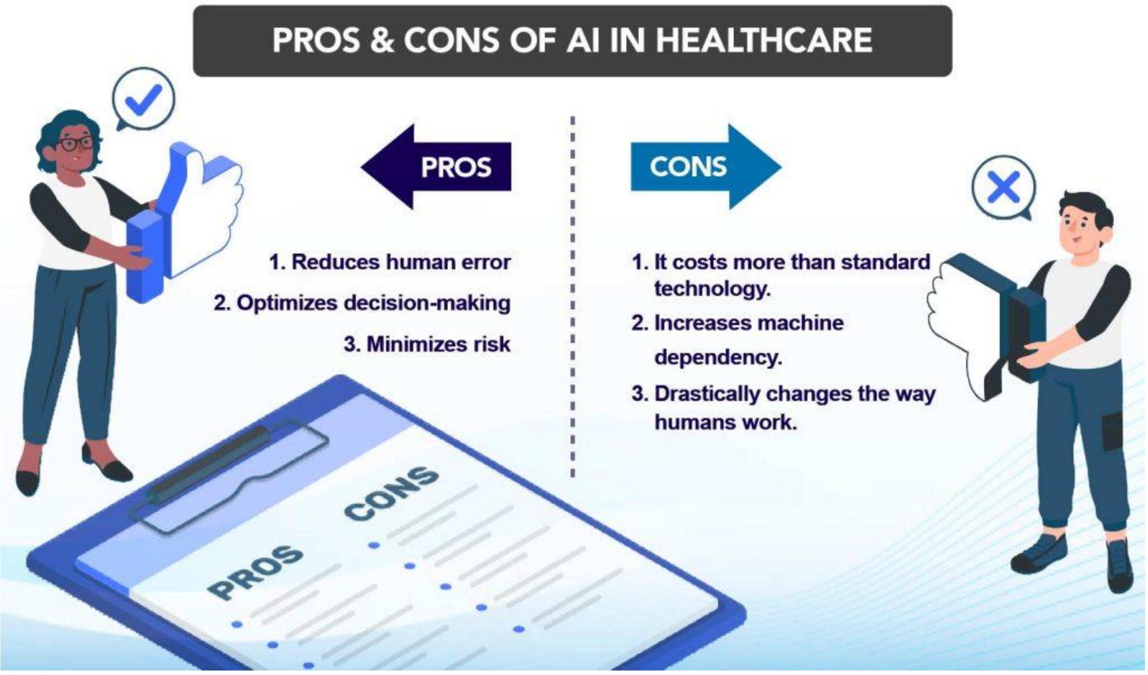

Figure 1. The Pros and Cons of AI in Healthcare (Molakal 2024)

In the context of implementing advanced technologies, it is essential to weigh the advantages and disadvantages. Figure 1 illustrates the key pros and cons associated with these technologies (Molakal 2024). By understanding these pros and cons, stakeholders can make more informed decisions regarding the adoption of advanced technologies. While the benefits of reducing human error, optimizing decision-making, and minimizing risk are substantial, it is important to address the challenges of higher costs, increased machine dependency, and the significant changes to human work processes.

2.3.2 Advantages

Reduces the likelihood of human error: AI technologies are capable of examining extensive volumes of medical data, spotting patterns and trends that may elude human observation. This capability enhances the accuracy of diagnoses and treatment suggestions. (Molakal 2024.)

Enhances decision-making: AI can support healthcare professionals in navigating intricate decisions by furnishing them with pertinent data and insights. Consequently, doctors can allocate more time to delivering direct patient care. (Molakal 2024.)

Mitigates risks: AI-driven surgical robots can aid surgeons in executing intricate procedures with heightened precision and command. This results in improved patient outcomes and diminished complications. (Molakal 2024.)

2.3.3 Disadvantages

Exceeds the costs of conventional technology: The development and integration of AI systems can incur substantial expenses, posing a financial challenge for certain healthcare providers. (Molakal 2024.)

Fosters dependence on machines: With increasing reliance on AI for decision-making, healthcare professionals may risk a decline in critical thinking abilities. (Molakal 2024.)

Alters human work dynamics significantly: The advent of AI may precipitate job displacement within the healthcare sector as automation threatens roles currently fulfilled by humans. (Molakal 2024.)

3 FUNDAMENTALS OF PRESCRIPTIVE ANALYTICS

This chapter delves into the fundamentals of prescriptive analytics, the most advanced form of data analysis. It explores its conceptual basis, techniques, and algorithms, and highlights its transformative applications in healthcare, enabling personalized treatment plans and optimized patient care.

3.1 Concept

Prescriptive analytics is the most advanced form of data analysis, going beyond just identifying trends or predicting future outcomes (Sruthi 2024.).

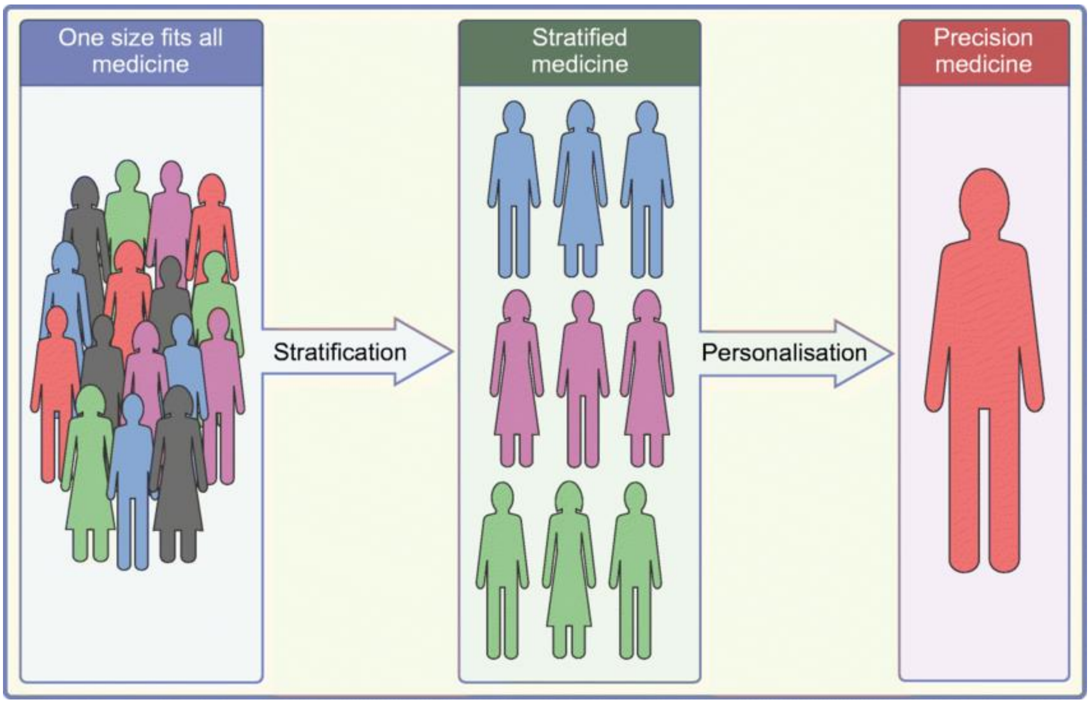

Figure 2. Precision Medicine for Diabetes (Griffin 2022)

As shown in Figure 2, precision medicine surpasses generalized and stratified approaches by considering each patient's unique genetic, environmental, and lifestyle factors. This tailored method enhances treatment efficacy, reduces adverse reactions, and is especially effective for chronic conditions like diabetes, improving management and patient outcomes.

"One size fits all" approaches typically involve generalized treatment protocols that are broadly applied to patient populations without considering individual variability. These approaches often overlook the diverse genetic, environmental, and lifestyle factors influencing disease manifestation and treatment response. Consequently, patients may receive suboptimal or ineffective treatments, leading to poorer outcomes and increased healthcare costs. (Griffin 2022.)

On the other hand, stratified medicine acknowledges patient heterogeneity but categorizes individuals into broad subgroups based on shared characteristics, such as genetic mutations or biomarker profiles. While this approach represents a step towards personalized care, it still lacks the granularity necessary to tailor treatments precisely to individual patients' unique needs and characteristics. (Chung et al. 2020.)

Precision medicine, as defined in the ADA–EASD Consensus Report, represents the pinnacle of personalized care. By integrating complex data from various sources, including genomics, proteomics, metabolomics, and lifestyle factors, precision medicine aims to tailor interventions precisely to each patient. This approach enables healthcare providers to deliver targeted therapies that are most likely to be effective while minimizing adverse effects and optimizing patient outcomes. (Chung et al. 2020.)

In healthcare, it utilizes insights from various analytics stages to recommend specific actions or treatment plans to optimize patient care. Here's the process for prescriptive analytics in healthcare:

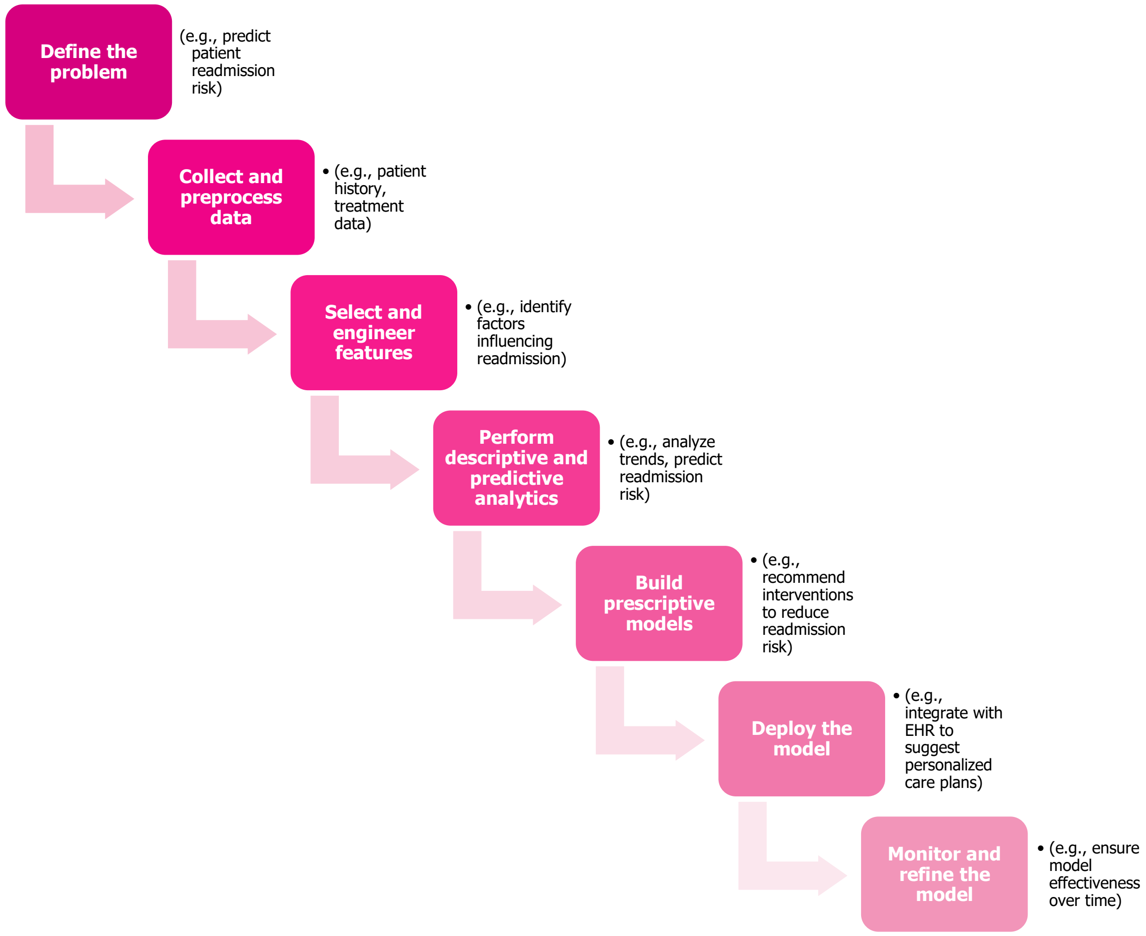

Figure 3. Steps of process for prescriptive analytics in healthcare. (Kuttappa 2020)

Through Figure 3, Kuttappa (2020) outlines a structured approach to implementing prescriptive analytics in healthcare. By adhering to these steps, healthcare providers can harness the full potential of data-driven insights to deliver personalized, efficient, and high-quality care to patients.

3.2 Techniques and Algorithm

Prescriptive analytics employs a range of techniques to guide organizations in making informed decisions based on descriptive and predictive analytics. Below are some commonly utilized techniques:

3.2.1 Optimization

Optimization is a mathematical methodology aimed at identifying the best solution from various alternatives within given constraints. Techniques such as linear programming, integer programming, and

advanced multivariate algorithmic methods are employed. These methods are crucial for resolving complex decision problems involving multiple variables and constraints, ensuring the optimal solution is found. (Elragal et al. 2023.)

3.2.2 Simulation

Simulation techniques enable organizations to create models of situations and simulate different actions to observe potential outcomes. This approach is particularly valuable in scenarios where testing different actions in real life would be impractical or costly. (FasterCapital s.a.)

3.2.3 Heuristics

Heuristic methods involve rule-based techniques designed to expedite the process of finding a satisfactory solution when identifying an optimal solution is challenging. These methods are often applied to complex or large-scale decision problems, providing feasible solutions within a reasonable timeframe. (MindTools s.a.)

3.2.4 Machine Learning

Machine learning algorithms analyses historical data to make predictions about future data. In prescriptive analytics, these predictions inform decision-making processes. Techniques such as regression, classification, and clustering are incorporated to derive actionable insights from the data. (Taiwo 2024.)

3.2.5 Artificial Intelligence (AI)

Advanced prescriptive analytics solutions may leverage AI techniques, including neural networks, deep learning, and reinforcement learning. These techniques model complex relationships and enhance the accuracy of predictions, aiding in the decision-making process. (Jaggi 2023.)

3.2.6 Decision Trees

Decision trees are schematic, tree-like diagrams used to determine courses of action or illustrate statistical probabilities. They lay out potential options, outcomes, and resource costs in a clear, easily understandable format. (Aunalytics 2021.)

3.2.7 Monte Carlo Simulation

Monte Carlo simulation employs probability distributions to model risk or uncertainty. It allows analysts to run numerous 'what-if' scenarios, providing insights into the likelihood and impact of different outcomes. (Biyani 2023.)

3.2.8 Genetic Algorithms

Genetic algorithms are heuristic search algorithms inspired by the process of natural selection. Utilizing concepts such as mutation, crossover, and selection, these algorithms optimize and search for solutions, identifying optimal or near-optimal outcomes for complex problems.

These techniques collectively enhance the capability of prescriptive analytics to provide precise and actionable recommendations for decision-making in various organizational contexts. (Mallawaarachchi 2023.)

3.3 Applications in Healthcare

Prescriptive analytics empowers healthcare professionals with data-driven decisions, resulting in:

Improved Patient Care: Through personalized treatment, preventive strategies, and early interventions, patient outcomes are significantly enhanced.

Enhanced Efficiency: Optimized resource allocation leads to cost reductions and shorter wait times.

Proactive Healthcare: Predicting potential risks allows healthcare systems to take preventive actions, improving overall population health.

As healthcare analytics evolves, prescriptive analytics is poised to revolutionize healthcare delivery and management. (Sarahedwards 2024.)

4 BUILDING A PRESCRIPTIVE MODEL FOR DRUG PRESCRIPTION: DEVELOPING A PROOF OF CONCEPT FOR PRESCRIPTIVE ANALYTICS IN HEALTHCARE

This chapter focuses on the development of a prescriptive model aimed at optimizing drug prescription practices within the healthcare sector. The primary goal is to present a proof of concept for prescriptive analytics by leveraging a data-driven approach. This involves the creation and validation of an algorithm that can accurately recommend appropriate medications based on patient-specific data. By integrating advanced analytical techniques, the model seeks to enhance decision-making processes, improve patient outcomes, and contribute to more efficient healthcare delivery by encouraging more research and development in this field.

4.1 Introduction to model development

Creating a prescriptive analysis model for medical decision support entails a thorough process that blends diverse data collection methods to ensure precision, dependability, and pertinence. The objective is to aid healthcare professionals in making well-informed decisions through actionable insights derived from data-driven analysis. (Coulter et al. 2013.) Here's a condensed overview:

a) Understanding the Problem Domain: Start by defining the problem domain and identifying specific decisions the model will support, such as disease diagnosis, treatment recommendations, or resource allocation. (Coulter et al. 2013.)

   Data Collection Methods:

   1. Structured Data Sources: Utilize organized data sources like Electronic Health Records (EHR), hospital databases, and medical imaging. (Babu 2023.)
   2. Unstructured Data Sources: Integrate unstructured data from clinical notes, research papers, and patient-generated data using Natural Language Processing (NLP) techniques. (Babu 2023.)

3. Real-Time Data Streams: Capture dynamic patient data from IoT devices, sensors, and monitoring systems. (Babu 2023.)
4. Patient Feedback and Surveys: Gather subjective data from patients via surveys or interviews. (Babu 2023.)
5. External Data Sources: Enhance analysis with external sources like medical literature databases and demographic data. (Babu 2023.)

b) Data Preprocessing and Integration: Cleanse, preprocess, and integrate collected data to ensure consistency and accuracy. (Babu 2023.)

c) Feature Engineering: Identify predictive features using statistical analysis and domain-specific knowledge. (Babu 2023.)

d) Model Development:
 1. Algorithm Selection: Choose appropriate algorithms based on data characteristics and desired outcomes. (Babu 2023.)
 2. Model Training and Validation: Train models using historical data and validate their performance. (Babu 2023.)

e) Prescriptive Analysis:
 1. Decision Rules and Guidelines: Develop decision rules based on model insights aligned with clinical best practices. (Babu 2023.)
 2. Scenario Analysis and Simulation: Evaluate decision alternatives through scenario analysis and simulation. (Babu 2023.)

a) Integration with Decision Support System (DSS): Integrate the model into a user-friendly DSS for real-time recommendations. (Babu 2023.)

b) Continuous Monitoring and Improvement: Monitor model performance, gather user feedback, and update the model accordingly to adapt to evolving healthcare needs. (Babu 2023.)

This systematic approach empowers healthcare professionals to enhance decision-making, improve patient outcomes, and optimize healthcare delivery.

While this chapter outlines a comprehensive approach for developing prescriptive healthcare models, this study adopted a more focused and narrower approach. The goal was to demonstrate the feasibility of using a decision tree classifier to predict drug response based on patient characteristics.

We utilized a pre-existing structured dataset (e.g., "drug200.csv", (Kaggle 2021)) for data collection, aligning with the chapter's structured data sources. The code performed data cleaning, handling missing values, and encoding categorical features. A decision tree was chosen for its interpretability and suitability for the data.

Firstly, it aimed to establish the feasibility of this method for drug response prediction. Secondly, readily available data allowed for efficient exploration. Finally, decision trees offer good initial exploration due to their simplicity and efficiency.

This study lays the groundwork for future research that can incorporate more sophisticated techniques and data sources.

4.2 Selection and Preparation of the Dataset

The selected dataset titled "Drugs A, B, C, X, Y for Decision Trees" from the online repository Kaggle (Kaggle 2021) has been chosen. This dataset contains information on 200 patients who all suffered from the same illness and were treated with one of five medications: Drug A, Drug B, Drug C, Drug X, and Drug Y. (Kaggle 2021.)

The purpose of selecting this dataset is to develop a prescriptive analytics model as a proof of concept for the thesis. The model aims to predict which medication would be most suitable for future patients with the same illness based on their demographic and health characteristics. The dataset includes features such as Age, Sex, Blood Pressure, and Cholesterol levels of the patients, with the target variable being the drug to which each patient responded. This dataset offers an opportunity to explore multiclass classification, wherein the goal is to classify patients into one of the five medication categories. By utilizing the training portion of the dataset, a decision tree model will be constructed. This model will enable predictions of the medication class for a new patient with similar characteristics or recommendations of an appropriate drug based on their individual profile.

The significance of this research lies in its potential to enhance personalized medicine by providing tailored treatment recommendations for patients based on their unique attributes. Despite the challenges posed by technological limitations, this study seeks to demonstrate the feasibility and utility of prescriptive analytics in clinical decision-making.

4.3 Model Development

This Python program analyzes and predicts drug prescriptions for 200 patients using a dataset. (Developed using Python 3.10 on PyCharm) Here's a breakdown of the code:

This Python program utilizes several libraries (Zhang 2021), the Python libraries as follows:

- pandas (pd): Used for data manipulation and analysis.
- numpy (np): Used for numerical computations.
- seaborn (sns) and matplotlib.pyplot (plt): Used for data visualization.
- sklearn.model_selection: Used for splitting data into training and testing sets.
- sklearn.preprocessing: Used for label encoding categorical variables.
- sklearn.tree: Used for building and visualizing the decision tree model.
- warnings: Used to suppress warnings.

Check File Path and Existence:

- Prints the current working directory and lists files in that directory.
- This helps ensure the script is in the same directory as the data file (drug200.csv).

Read CSV Data:

- Reads the CSV file named "drug200.csv" into a Pandas DataFrame (df).
- The DataFrame is a tabular data structure with labeled columns.

Data Preprocessing:

```
# Data preprocessing
df.rename(columns={'Na_to_K': 'Sodium_to_Potassium', 'BP': 'Blood_Pressure'}, inplace=True)
df['Sex'].replace({'M': 'Male', 'F': 'Female'}, inplace=True)
df['Sodium_to_Potassium'] = df['Sodium_to_Potassium'].round(0)
df['Sodium_to_Potassium'] = df['Sodium_to_Potassium'].astype('int')

age_counts = df['Age'].value_counts()
top_ages = age_counts.head(5)
df_2 = pd.DataFrame({'Age': top_ages.index, 'Count': top_ages.values})
```

Figure 4. Part of python code that does data preprocessing

Figure 3 displays a code snippet performing several data preprocessing tasks on a DataFrame (df) likely containing information about patients' medical data. Here's a breakdown of the preprocessing steps:

- Renames columns for better readability (e.g., "Na_to_K" to "Sodium_to_Potassium").
- The code replaces 'M' with "Male" and 'F' with "Female" in the "Sex" column, assuming those were the original values.
- Rounds and converts the "Sodium_to_Potassium" column to integers for potential use in the decision tree.
- Analyzes age distribution and creates a separate DataFrame (df_2) to show the top 5 most frequent ages.
- Prints the modified DataFrame (df) to see the changes.
- Creates a copy of the modified DataFrame (df_modified) for potential future use.

Exploratory Data Analysis (EDA):

```
# Plotting
plt.figure(figsize=(10, 6))

plt.subplot( *args: 3, 3, 1)
sns.countplot(x='Sex', data=df, palette='viridis')
plt.title('Countplot of Sex')

plt.subplot( *args: 3, 3, 2)
sns.countplot(x='Blood_Pressure', data=df, palette='viridis')
plt.title('Countplot of Blood_Pressure')

plt.subplot( *args: 3, 3, 3)  # Corrected subplot position
sns.countplot(x='Cholesterol', data=df, palette='viridis')
plt.title('Countplot of Cholesterol')

plt.subplot( *args: 3, 3, 4)
sns.countplot(x='Drug', data=df, palette='viridis')
plt.title('Countplot of Drug')
```

Figure 5. EDA plotting for Sex, blood pressure and cholesterol

The code in Figure 4, is generating the several countplota to explore the distribution of categorical variables in the data. These plots visualize the number of patients in each category.

Here's a breakdown of the code that generates these plots:

- Creates various visualizations using Seaborn to understand the data better.
- Countplots show the distribution of categorical variables like Sex, Blood Pressure, Cholesterol, and Drug.
- Barplots explore Drug frequencies and top 5 frequent ages.
- Boxplot shows the distribution of Sodium_to_Potassium levels across Sex categories.

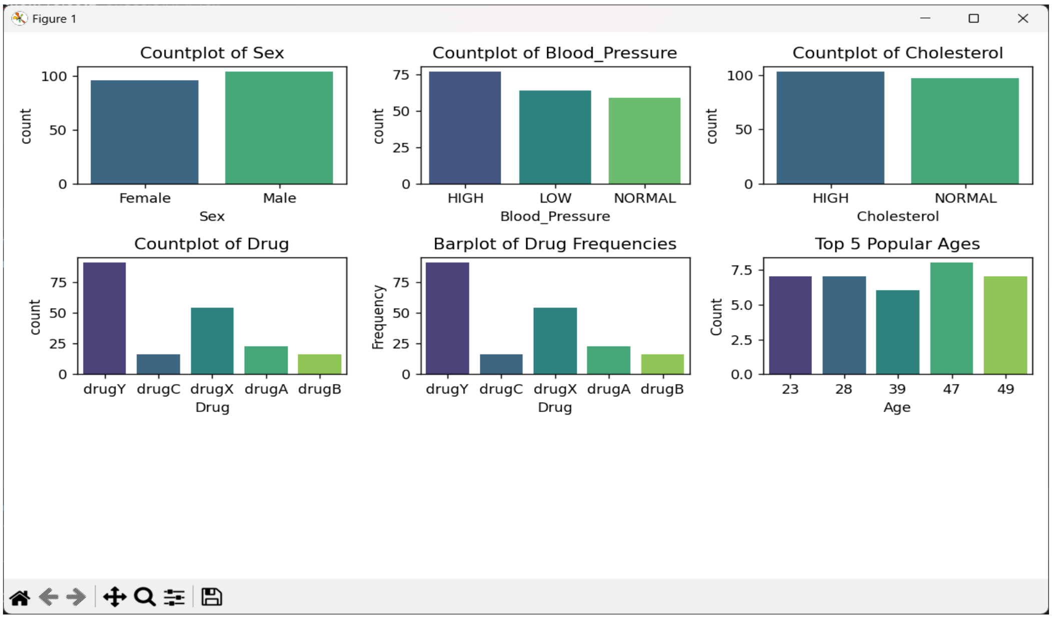

Figure 6. Count plots, frequencies and top 5 frequent ages.

As shown in figure 5, these plots visualizing the number of patients in each category. Additionally, figure 5 is showing the six bar charts for the top five most frequent age groups based on the dataset.

Splitting Data:

```
# Splitting the data
input_features = df.iloc[:, :-1]  # Renamed from 'input' to 'input_features'
Target = df.iloc[:, -1]  # Selecting the last column as the target variable

# Encoding the categorical variables
le_sex = LabelEncoder()  # Creating a LabelEncoder instance for 'Sex'
le_BP = LabelEncoder()   # Creating a LabelEncoder instance for 'Blood_Pressure'
le_cholesterol = LabelEncoder()  # Creating a LabelEncoder instance for 'Cholesterol'

input_features['Sex'] = le_sex.fit_transform(input_features['Sex'])
input_features['Blood_Pressure'] = le_BP.fit_transform(input_features['Blood_Pressure'])
input_features['Cholesterol'] = le_cholesterol.fit_transform(input_features['Cholesterol'])

print(input_features)

X_train, X_test, y_train, y_test = train_test_split( *arrays: input_features, Target, test_size=0.3, random_state=42)
clf = DecisionTreeClassifier()
clf.fit(X_train, y_train)
accuracy = clf.score(X_test, y_test)
print("Model Accuracy:", round(accuracy*100, 1))
```

Figure 7. Shows the splitting of data for decision tree and testing

The figure 6 accurately represents the code for splitting data. Splitting data into training and testing sets is a crucial step in machine learning because it allows you to train the model on a portion of the data (training set) and then evaluate its performance on unseen data (testing set). This helps to prevent overfitting, which occurs when a model performs well on the training data but poorly on new data. Here's a breakdown:

- Separates the features (all columns except the last) into input_features and the target variable (drug name) into Target.
- Splits the data into training and testing sets using train_test_split.
- Training data (70%) is used to build the decision tree model.
- Testing data (30%) is used to evaluate the model's performance.

Encoding Categorical Variables:

The code focuses on encoding categorical variables in the data before feeding them into a decision tree model. This is an important step because decision tree models typically work better with numerical features rather than textual labels.

LabelEncoder: This is a class from the scikit-learn library used for encoding categorical variables. It works by:

- Identifying all unique categories (textual labels) present in a feature.
- Assigning a unique integer value to each category.
- Replacing the original textual labels in the data with the corresponding integer values.

For example, if the "Sex" feature has categories "Male" and "Female", the LabelEncoder might assign "Male" the value 0 and "Female" the value 1. This allows the decision tree model to understand the relationship between the features and the target variable more effectively.

Code Implementation:

- Creates separate LabelEncoder instances for "Sex", "Blood_Pressure", and "Cholesterol".
- Label encoders convert textual labels (e.g., "Male", "LOW") into numerical values for the decision tree to understand.
- Encodes the categorical features in input_features using the respective LabelEncoders.

Building and Evaluating the Decision Tree Model:

Here's a code breakdown:

- Creates a DecisionTreeClassifier object (clf).
- Trains the model using the training data (X_train and y_train).
- Evaluates the model's accuracy on the testing data (X_test and y_test).
- Prints the model accuracy as a percentage.

A decision tree model was constructed to predict drug classifications based on patient characteristics using scikit-learn's DecisionTreeClassifier. The model was trained on 70% of the data to learn relationships between features and the target variable. To assess generalizability and avoid overfitting,

the model's accuracy (proportion of correct predictions) was evaluated on the remaining 30% of the data. This evaluation ensures the model performs well on unseen data, which is crucial for real-world application.

Visualizing the Decision Tree:

```
# Plotting the decision tree
unique_classes = np.unique(y_train)
plt.figure(figsize=(20, 10))
plot_tree(clf, class_names=unique_classes, filled=True)
plt.show()
```

Figure 8. Uses plot_tree to create a visual representation of the decision tree model.

The figure 8, depicts a visual representation of the decision tree model that is built to classify drugs based on patient data. This visualization, potentially generated using plot_tree from scikit-learn, helps understand the decision-making process of the model.

Printing LabelEncoder Classes:

```
# Check and print the classes for each LabelEncoder
print("Sex classes: ", le_sex.classes_)
print("Blood_Pressure classes: ", le_BP.classes_)
print("Cholesterol classes: ", le_cholesterol.classes_)
```

Figure 9. Prints the class labels for each LabelEncoder used for reference.

In Figure 8, the code prints the class labels for each LabelEncoder used for categorical features. First it confirms that the LabelEncoders have correctly learned and assigned unique integer values to each category. Second, it provides a reference to which integer value corresponds to which category. It ensures the data is prepared correctly for the decision tree model.

Prediction with User Input:

a) Prompts the user to enter values for Age, Sex, Blood Pressure, Cholesterol, and Sodium_to_Potassium.
b) Creates a new list (X_new) to store the user-provided data.
c) Handles unseen labels (categories not present in the training data) for "Sex", "Blood_Pressure", and "Cholesterol".
d) If an unseen label is encountered, a warning message is printed, and a default value (-1) is assigned.
e) Converts the list X_new to a NumPy array for compatibility with the model.
f) Uses the trained model (clf) to predict the drug class (represented by a number) for the user's input.
g) Prints the predicted drug class based on the model's prediction.

4.4 Decision Tree

The decision tree generated from the code functions as a model to predict appropriate drug types based on a patient's characteristics. It considers five features: age, sex (encoded numerically), blood pressure (encoded numerically), cholesterol (encoded numerically), and the sodium-to-potassium ratio (rounded to the nearest integer). The model starts at the root node, representing the entire training dataset. At each subsequent node, a decision rule based on a single feature splits the data. The left and right branches represent "yes" and "no" answers to the question, respectively.

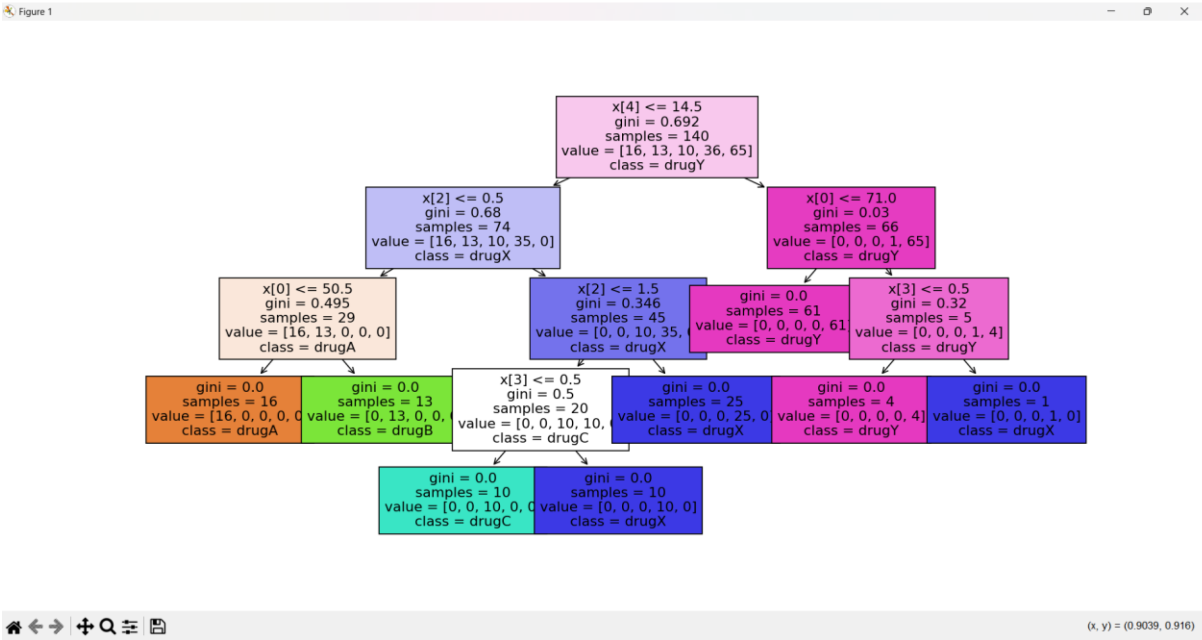

Figure 10. Shows a visualization of a decision tree for the prescriptive model

As shown in the figure 10, the visualization of the decision tree is making a prediction for a new patient, the model starts at the root node. If the patient's age is less than or equal to 14.5 years, they are classified as needing drugY. Otherwise, the decision tree follows the right branch, considering blood pressure next. Patients older than 14.5 years with low blood pressure are further divided based on their sex. Those with low blood pressure and a sex code of 0 (likely female) or 1 (likely male) are both predicted to need drugX. For patients older than 14.5 years with normal or high blood pressure, the sodium-to-potassium ratio becomes the next decision factor. Patients in this category with a ratio less than or equal to 0.5 are predicted to need drugY, while those with a higher ratio are predicted to need drugC.

4.5 Drug Prescription

This study implemented a decision tree model to predict drug prescriptions for a dataset of 200 patients. The model considered five features of each patient: age, sex (encoded numerically), blood pressure (encoded numerically), cholesterol level (encoded numerically), and the sodium-to-potassium ratio (rounded to the nearest integer). The decision tree functioned by iteratively splitting the data based on a single feature at each node. The left and right branches of the tree represented "yes" and "no" answers, respectively, to the question posed by the decision rule at that node.

A very important consideration while conducting this study is the fact that is data set as well as the model is intended to provide a proof of concept for the topic of prescriptive analysis. Real world data and algorithm would be extremely complex and definitely not relying just on 5 variables as seen in the model development in this study. Additionally, the accuracy of the model is contingent on the quality and completeness of the training data.

4.6 Implementation and Testing

This chapter describes how we built and tested a decision tree model to predict drug classifications from patient data. Here's a testing of the model's performance:

```
Enter age: 27
Enter sex (Male/Female): Female
Enter blood pressure (LOW/NORMAL/HIGH): HIGH
Enter cholesterol (NORMAL/HIGH): NORMAL
Enter sodium to potassium ratio: 8
Predicted drug: drugA
```

Figure 11. User Input for Prediction (Test 1).

Figure 11 demonstrates the user interface aspect of the model. After training the decision tree model, this figure depicts the code prompting the user for input. This allows for testing the model on data not included in the training set.

In this example, a fictional patient's data was entered for testing purposes. The model then utilized the trained decision tree and the provided patient characteristics to generate a predicted drug classification, resulting in "drugA".

```
Enter age: 47
Enter sex (Male/Female): Male
Enter blood pressure (LOW/NORMAL/HIGH): LOW
Enter cholesterol (NORMAL/HIGH): HIGH
Enter sodium to potassium ratio: 13
Predicted drug: drugC
```

Figure 12. User Input for Prediction (Test 2).

Figure 12, proves that the features play a significant role in differentiating the predications. It shows how inputting the different data will predict suitable drug to the patient. The decision tree model has a series of decision points based on the features (e.g., age, sex, blood pressure, cholesterol and ratio of sodium to potassium).

```
Enter age: 89
Enter sex (Male/Female): Female
Enter blood pressure (LOW/NORMAL/HIGH): NORMAL
Enter cholesterol (NORMAL/HIGH): NORMAL
Enter sodium to potassium ratio: 40
Predicted drug: drugX
```

Figure 13. User Input for Prediction (Test 3).

In figure 13, it demonstrates that at each decision point, the model considers the value of a specific feature of the patient. The model follows a specific branch in the tree, leading it to predict the drug class based on the input values.

Even though Figures reveal the specific calculations, we can infer that the age, sex, and other features likely had different values in Test 1, Test 2 and Test 3. These different values caused the model to follow different paths through the decision tree in each test. Since the decision paths were different, the model ended up resulting in different predicted drug classifications.

5 INTEGRATION OF PRESCRIPTIVE MODEL INTO CLINICAL PRACTICE

This chapter explores the integration of prescriptive models into clinical practice, highlighting strategies for effective implementation within healthcare systems. It addresses the challenges of real-time data access, data bias, and ensuring patient privacy, emphasizing the need for robust data collection, advanced technological infrastructure, and ethical considerations to enhance healthcare outcomes.

5.1 Integration Strategies with Healthcare Systems

Here Organizations need to prioritize the development of a robust data collection process to ensure they have reliable and up-to-date patient and clinical data, which is crucial for building effective prescriptive models. Accurate data is the cornerstone of prescriptive analytics, significantly influencing the precision and dependability of the insights generated.

An efficient data gathering process encompasses several key components:

a. Integration of Multiple Data Sources: Merging data from various sources such as Electronic Health Records (EHRs), wearable health devices, lab results, and patient surveys provides a holistic view of patient health. This comprehensive data collection enhances the accuracy and personalization of prescriptive models. (Encora 2023.)

b. Real-Time Data Acquisition: Implementing systems that support real-time data collection ensures models are updated with the latest information, which is critical for timely and relevant healthcare decisions. This data can be sourced from IoT devices, continuous monitoring systems, and regular updates to EHRs. (Del Prado s.a.)

c. Ensuring Data Quality and Standardization: It is vital to maintain high data quality and consistency by standardizing data formats and performing data cleaning to address errors and missing values. This ensures the reliability of the data used in prescriptive models. (Encora 2023.)

d. Patient Privacy and Security: Robust data privacy and security measures must be in place to protect sensitive patient information. Compliance with regulations like HIPAA (Health Insurance Portability and Accountability Act) is essential to secure data collection practices and respect patient confidentiality. (Del Prado s.a.)

e. Technological Infrastructure: Investing in advanced technological infrastructure that supports data interoperability, scalability, and integration is crucial. This includes leveraging cloud-based solutions, APIs, and data warehousing technologies to facilitate smooth data exchange and storage. (Encora 2023.)

By focusing on these elements, healthcare organizations can establish a strong data collection process that supports the development of reliable prescriptive analytics models. These models can then provide actionable insights to improve patient care, optimize resource utilization, and enhance overall healthcare outcomes.

Then follows model development and deployment where Application Programming Interfaces (APIs) allow seamless communication between prescriptive analytics models and existing healthcare IT systems. This enables real-time integration of model recommendations into clinical workflows. By integrating prescriptive analytics models with clinical decision support systems (CDSS), clinicians can receive real-time, evidence-based recommendations directly at the point of care. This improves healthcare integration. This empowers them to make more informed decisions while considering patient-specific factors. (Noteboom et al. 2022.)

5.2 Challenges of Implementing Prescriptive Analytics in Healthcare

Accessing real-time data poses a challenge as it's not always readily available within Electronic Health Record (EHR) systems. For improved patient safety, medical devices need to deliver current patient vitals to alert clinicians effectively, without disrupting their workflow or causing alert fatigue. Clinical decision support systems relying on accurate diagnoses can leverage available data, including through health information exchanges. Despite widespread interest among healthcare professionals in predictive analytics, securing sufficient funding for implementing these new tools proves challenging. Technological hurdles persist, hindering the realization of anticipated results. (Encora 2023.)

5.3 Data Bias

Data bias is a critical consideration in the development of prescriptive analysis models for healthcare. Failure to address data bias can significantly undermine the effectiveness and reliability of healthcare analytics, potentially leading to erroneous conclusions and detrimental outcomes for patients.

Bias in healthcare data can manifest in various forms, including demographic bias, selection bias, and algorithmic bias. Demographic bias occurs when certain demographic groups are overrepresented or

underrepresented in the dataset, skewing the analysis and resulting recommendations. Selection bias may arise if the data collected is not representative of the target population, leading to inaccurate insights and recommendations. Algorithmic bias occurs when machine learning algorithms perpetuate or amplify existing biases present in the data, resulting in unfair or discriminatory outcomes for certain groups of patients. (Celi et al. 2022.)

5.3.1 Prevalence of bias in healthcare statistics

Healthcare data is derived from electronic health records, surveys, and insurance claims. The datasets may be distorted by historical disparities in healthcare access, quality, and treatment. Below are few common biases that might occur in healthcare data:

Sample Bias: Healthcare statistics may exhibit underrepresentation of specific demographic groups, leading to a distortion of results. (Edmond 2024.)

Selection bias: occurs when data collected from select healthcare providers or centres does not accurately represent the entire community, leading to biased results. Erroneous or insufficient variable measurements can distort analysis. (Edmond 2024.)

Confirmation bias: refers to the tendency of researchers to evaluate and interpret data in a way that aligns with their pre-existing opinions or prejudices.(Edmond 2024.)

5.3.2 Effects of Bias and Discrimination

Biased healthcare data analysis may contribute to disparities in healthcare delivery and outcomes. Healthcare data bias and prejudice have multiple impacts:

Disparate Treatment: Prejudiced analysis might result in unequal recommendations for different demographic groups, leading to healthcare disparities (Vela et al 2022.).

Diagnosis: The presence of underrepresented groups in healthcare data might lead to the overlooking of their symptoms and issues, leading to delays or incorrect diagnoses (Vela et al 2022.).

Inappropriate data analysis: Bias can result in distorted study findings, impeding the progress of medical research and evidence-based practices (Vela et al 2022.).

Biased data analysis: It might exacerbate preconceived notions and marginalize specific demographic groups, impeding equitable and inclusive healthcare (Vela et al 2022.).

5.3.3 Addressing Prejudice and Inequality

To achieve fair and equitable healthcare outcomes, it is necessary to tackle bias and discrimination in the analysis of healthcare data. The following techniques can resolve these problems:

Comprehensive Data Collection: Include data from a wide range of sources and demographic groups to minimize any bias in the sample and improve the accuracy of the study. In order to ensure precise and dependable analysis, it is crucial to meticulously cleanse and validate healthcare data in order to detect and rectify any biases in measurement. (Babyar 2018.)

Transparent Reporting: Researchers must provide a clear and comprehensive account of their methodology, assumptions, and biases in order to ensure accountability and enable the replication of their work. Implementing ethical protocols and employing conscientious decision-making during the process of data analysis might help mitigate researcher biases and yield impartial conclusions. (Ethical Considerations in Healthcare Data Analysis and Privacy 2024.)

Ongoing Monitoring and Improvement: Consistently monitoring and incorporating feedback aids in identifying and removing biases, leading to enhanced precision and impartiality in data analysis. (Babyar 2018.)

To establish a more equitable healthcare system, we can take proactive measures to address bias and discrimination in the study of healthcare data. Collaboration between researchers, data scientists, and policymakers is necessary to ensure the provision of equitable and accurate healthcare data insights. (Ethical Considerations in Healthcare Data Analysis and Privacy 2024.)

5.4 Ensuring Privacy and Ethical consideration

In order to maintain patient trust and adhere to tight regulatory standards, healthcare institutions must establish strong privacy protections while collecting and analyzing large amounts of sensitive patient data.

5.4.1 The Significance of Maintaining Patient Confidentiality

Ensuring patient privacy is of utmost importance in the healthcare sector due to a multitude of reasons:

1. Trust and Confidentiality: Patients must have confidence in the assurance that their personal information will be maintained in strict confidentiality. Promoting privacy fosters confidence between patients and healthcare providers, incentivizing individuals to pursue essential medical treatment without apprehension about the abuse or disclosure of their data. (Ethical Considerations in Healthcare Data Analysis and Privacy 2024.)
2. Legal Compliance: Healthcare organizations are obligated to adhere to stringent regulatory mandates, including General Data Protection Regulation (GDPR) om European Union and the Health Insurance Portability and Accountability Act (HIPAA) in the United States. Noncompliance with these requirements may lead to significant fines and legal repercussions. (Perry 2019.)
3. Ethical Obligations: It is an ethical obligation to safeguard patient privacy. Healthcare practitioners must adhere to the values of autonomy, beneficence, and non-maleficence, while also maintaining the confidentiality and protection of patient information. (Ethical Considerations in Healthcare Data Analysis and Privacy 2024.)

5.4.2 Ensuring Patient Privacy using Sophisticated Data Analysis Techniques

In order to protect patient confidentiality while taking advantage of the advantages of sophisticated data analysis, healthcare institutions can adopt the following approaches:

1. Robust Data Encryption

Encrypting confidential patient data is crucial to thwart unauthorized access. Healthcare institutions can safeguard patient information by employing strong encryption techniques, rendering it nearly indecipherable to unauthorized users. This greatly decreases the likelihood of data breaches and improves the confidentiality of patient information. (Amod 2023.)

2. Regulated Data Access

It is essential to restrict access to patient data only to authorized staff. Healthcare institutions should implement stringent rules and role-based access controls to ensure that only authorized persons with valid intentions can access and analyze patient data. Conducting regular access audits and monitoring will assist in detecting any illegal attempts to get access. (De Carvalho & Bandiera-Paiva 2018.)

3. Methods for Anonymization and De-identification

It is crucial to employ anonymization and de-identification methods in order to safeguard patient privacy throughout the execution of sophisticated data analytics. To ensure patient data remains anonymous, healthcare institutions should decrease the risk of re-identification by eliminating direct identifiers such as names, social security numbers, and addresses from datasets. (Lehtipuu 2023.)

4. Periodic security audits and training sessions

Healthcare businesses must to regularly do security audits to detect vulnerabilities in their systems and implement appropriate measures to reduce risks. In addition, offering extensive training to employees on data privacy best practices and the significance of patient confidentiality would increase awareness and mitigate the risk of privacy breaches. (Vicente 2023.)

5. Clear and Unambiguous Privacy Policies

Unambiguous and easily understandable privacy policies are crucial for upholding patient confidence. Healthcare organizations should effectively convey their data privacy policies to patients, ensuring they comprehend the methods by which their data will be gathered, utilized, and safeguarded. This enables patients to make well-informed decisions on their involvement in data-driven projects. (Ethical Considerations in Healthcare Data Analysis and Privacy 2024.)

Ensuring patient privacy in the era of advanced data analytics in healthcare is a significant concern due to data breaches, identity threats, and illegal access. Preserving patient privacy is essential for maintaining trust, adhering to rules, and upholding ethical responsibilities. Methods to safeguard patient confidentiality encompass robust data encryption, regulated data accessibility, anonymization, periodic security assessments, and transparent privacy protocols. (Ethical Considerations in Healthcare Data Analysis and Privacy 2024.)

Ultimately, with the ongoing evolution of complex data analytics in the healthcare sector, safeguarding patient privacy must be a paramount concern for healthcare organizations. Organizations may ensure patient trust, adhere to legislation, and confidently navigate the changing data-driven healthcare environment by establishing strong privacy protections. (Ethical Considerations in Healthcare Data Analysis and Privacy 2024.)

6 FUTURE DIRECTIONS AND IMPLICATIONS

This chapter explores future directions and implications for prescriptive analytics in healthcare. It examines advancements transforming clinical and administrative practices, potential applications beyond drug prescription, and the significant benefits for healthcare providers and patients, emphasizing enhanced decision-making and personalized care.

6.1 Advancements in Prescriptive Analytics for Healthcare

The field of prescriptive analytics in healthcare is rapidly evolving, offering significant advancements that are transforming both clinical and administrative practices. Building on the foundation of predictive analytics, prescriptive analytics goes a step further by not only forecasting future outcomes but also recommending actionable steps to achieve desired results.

In clinical settings, prescriptive analytics plays a crucial role by providing insights into disease spread and patient care. For instance, it can analyze patterns in patient data to suggest optimal treatment plans, anticipate potential complications, and improve overall patient outcomes. This proactive approach ensures that healthcare providers can deliver more accurate and timely care, ultimately enhancing patient health and safety. On the administrative front, prescriptive analytics significantly impacts the management of healthcare organizations. By optimizing resource allocation and operational efficiency, it helps hospitals streamline their processes. For example, prescriptive analytics can suggest the best ways to reduce waiting times, manage hospital beds, and allocate staff more effectively. This not only improves patient satisfaction but also allows hospitals to serve a growing community more efficiently.

The integration of prescriptive analytics as a complement to predictive analytics creates a comprehensive, intelligent system capable of preventing adverse scenarios and recommending the best course of action. By incorporating mathematical models and statistical algorithms, healthcare organizations can achieve complete management oversight. This optimization can lead to shorter waiting times and better resource utilization, ensuring that those in need receive timely and effective treatment. (Lopes et al 2020.)

6.2 Potential Applications Beyond Drug Prescription

Looking ahead, future advancements in prescriptive analytics will likely involve the development of integrated systems that combine various clinical and administrative components. Such systems would support both healthcare providers and administrators, fostering a more cohesive and efficient healthcare environment. The goal is to create a holistic approach that not only anticipates and responds to healthcare challenges but also enhances the overall quality of care provided to patients. The advancements in prescriptive analytics are paving the way for smarter, more efficient healthcare systems. By leveraging these tools, healthcare organizations can improve patient outcomes, optimize operations, and better meet the needs of their communities. (Lopes et al 2020)

6.3 Implications for Healthcare Providers and Patients

Prescriptive analytics enables healthcare decision-makers to optimize outcomes by recommending the best actions for patients and providers. This advanced analytics approach allows for the evaluation of multiple "what if" scenarios, helping to assess the impact of different choices. For healthcare providers, the implications are significant. Prescriptive analytics can enhance decision-making processes by suggesting the most effective treatments and operational strategies. For example, it can help determine the best way to allocate resources, manage patient flow, and schedule procedures. This leads to improved efficiency, reduced costs, and better patient care. Patients also benefit from prescriptive analytics through more personalized and timely care. By comparing different scenarios, healthcare providers can choose interventions that offer the greatest benefit, improving overall patient outcomes. For instance, in managing chronic diseases, prescriptive analytics can recommend the best combination of lifestyle changes and treatments to prevent disease progression, enhancing the patient's quality of life. In summary, prescriptive analytics offers a powerful tool for optimizing healthcare decisions, providing both providers and patients with data-driven insights that lead to better health outcomes and more efficient care delivery. (Kuttappa 2020.)

7 CONCLUSION

The integration of prescriptive analytics in healthcare signifies a transformative shift from traditional, reactive approaches to a more proactive, data-driven methodology. This thesis has examined the evolution of medicine prescription practices, highlighting the limitations of one-size-fits-all models and the progressive move towards personalized treatment plans informed by extensive patient data. Through a detailed exploration of healthcare decision-making processes, it is evident that the reliance on clinical experience and intuition alone is becoming outdated, as modern medicine increasingly embraces technological advancements and sophisticated analytical techniques.

The research objectives set forth in this thesis aimed to address the challenges and opportunities associated with implementing prescriptive analytics, develop a prescriptive model for healthcare decision-making serving as a proof of concept, explore the ethical and privacy concerns surrounding its deployment and also to discuss limitations with its implementation like a major concern called 'Data Bias'. The investigation has revealed that prescriptive analytics holds immense potential in enhancing diagnostic accuracy, optimizing treatment plans, and ultimately improving patient outcomes. By leveraging comprehensive patient data and predictive algorithms, healthcare providers can anticipate health issues, personalize treatments, and deliver care that is both effective and efficient.

The thesis has also underscored the ethical considerations and data privacy issues inherent in prescriptive analytics. As healthcare systems become more data-driven, ensuring patient confidentiality and ethical use of data is paramount. Balancing the benefits of advanced analytics with the need for stringent data protection measures is crucial for gaining and maintaining patient trust. The development and evaluation of a prescriptive model for drug prescription provided a proof of concept for the do ability and effectiveness of such approaches. The model demonstrated how integrating prescriptive analytics into clinical practice can offer tailored treatment recommendations, thereby enhancing the quality of care. The comparative analysis with traditional treatment decision-making highlighted the

superior accuracy and personalized approach offered by prescriptive models, albeit with considerations for implementation challenges and costs.

Looking ahead, the future of healthcare lies in further encouraging such model development and expanding their applications beyond drug prescription or basically to make it much more accurate and precise. The continuous evolution of prescriptive analytics promises to revolutionize not only individual patient care but also overall healthcare delivery systems. Future research should focus on addressing the challenges, such as integration with existing healthcare infrastructures, cost implications, continuous model improvement and mainly data bias.

In summary, this thesis has contributed to a wider understanding of the implications and potential benefits of prescriptive analytics in healthcare. The findings underscore the transformative impact of data-driven decision-making in medical practice, advocating for the broader adoption of prescriptive analytics to achieve optimal treatment outcomes. As healthcare continues to evolve, embracing prescriptive analytics will be key to delivering personalized, efficient, and effective care, ultimately improving health outcomes for patients worldwide.

REFERENCES

Artificial intelligence has been used in the work as follows:

ChatGPT 2023. OpenAI. GPT-3.5. Accessed for language check, May 2024. https://chat.openai.com

Alexopulos, S., Cancelliere, C., Cote, P., & Mior, S. 2021. Reconciling evidence and experience in the context of evidence-based practice. The Journal of the Canadian Chiropractic Association, 65(2), 132. Accessed 25.4.2024.

Amod, F. 2023. Encryption in healthcare: The basics. Paubox. https://www.paubox.com/blog/encryption-in-healthcare-the-basics. Accessed 29.4.2024.

Aunalytics. 2021. Decision Trees: An Overview - Aunalytics. https://www.aunalytics.com/decision-trees-an-overview/. Accessed 31.4.2024.

Babu, A. 2023. Data collection in healthcare: tools, methods, and importance. Emedlogix. https://www.emedlogix.com/post/data-collection-in-healthcare-tools-methods-and-importance. Accessed 2.5.2024.

Babyar, J. 2018. Equitable health: let's stick together as we address global discrimination, prejudice and stigma. Archives of Public Health, 76(1), 44. Accessed 24.42024.

Biyani, A. 2023. What is the Monte Carlo Method? CareerFoundry. https://careerfoundry.com/en/blog/data-analytics/monte-carlo-method/#:~:text=The%20Monte%20Carlo%20simulation%20is,solutions%20for%20complex%2C%20ambiguous%20problems. Accessed 14.4.2024.

Celi, L. A., Cellini, J., Charpignon, M., Dee, E. C., Dernoncourt, F., Eber, R., Mitchell, W. G., Moukheiber, L., Schirmer, J., Situ, J., Paguio, J., Park, J., Wawira, J. G., & Yao, S. 2022. Sources of bias in artificial intelligence that perpetuate healthcare disparities—A global review. PLOS Digital Health, 1(3), e0000022. https://doi.org/10.1371/journal.pdig.0000022. Accessed 10.5.2024.

Chung, W. K., Erion, K., Florez, J. C., Hattersley, A. T., Hivert, M., Lee, C. G., McCarthy, M. I., No-lan, J. J., Norris, J. M., Pearson, E. R., Philipson, L., McElvaine, A. T., Cefalu, W. T., Rich, S. S., & Franks, P. W. 2020. Precision Medicine in Diabetes: A consensus report from the American Dia-betes Association (ADA) and the European Association for the Study of Diabetes (EASD). Diabe-tes Care, 43(7), 1617–1635. https://doi.org/10.2337/dci20-0022. Accessed 11.5.2024.

Coulter, A., Stilwell, D., Kryworuchko, J., Mullen, P. D., Ng, C. J., & Van Der Weijden, T. 2013. A systematic development process for patient decision aids. BMC Medical Informatics and Decision Making, 13(S2). https://doi.org/10.1186/1472-6947-13-s2-s2. Accessed 5.4.2024.

Cyrus, T. 2023. 4 Types of Data Analytics & How They Improve Healthcare. Mi-croHealth, LLC. https://www.microhealthllc.com/blog/4-types-of-data-analytics-and-how-they-can-improve-healthcare/. Accessed 30.4.2024.

De Carvalho, M. A., Junior, & Bandiera-Paiva, P. 2018. Health Information System Role-Based Access Control Current Security Trends and Challenges. Journal of Healthcare Engineering, 2018, 1–8. https://doi.org/10.1155/2018/6510249. Accessed 29.4.2024.

Del Prado, R. s.a. Understanding Prescriptive Analytics in Healthcare Data Management. healthcarein-formation.management. https://www.healthcareinformation.management/data-analysis-prescriptive-analytics. Accessed 2.5.2024.

Edmond, C. 2024. Healthcare suffers from a racial bias. What can we do to change that? World Eco-nomic Forum. https://www.weforum.org/agenda/2024/02/racial-bias-equity-future-of-healthcare-clinical-trial/. Accessed 13.5.2024.

Elragal, R., Elragal, A., & Habibipour, A. 2023. Healthcare analytics—A literature review and proposed research agenda. Frontiers in Big Data, 6. https://doi.org/10.3389/fdata.2023.1277976. Accessed 9.5.2024.

Encora. 2023. Making the case for prescriptive analytics in healthcare - Excellarate. Encora. https://www.encora.com/insights/making-the-case-for-prescriptive-analytics-in-healthcare. Accessed 16.5.2024.

Ethical considerations in healthcare data analysis and privacy. 2024. https://moldstud.com/articles/p-ethical-considerations-in-healthcare-data-analysis-and-privacy. Accessed 25.4.2024

FasterCapital. s.a. Benefits of Using Simulation Techniques for Efficiency Improvement. https://fastercapital.com/topics/benefits-of-using-simulation-techniques-for-efficiency-improvement.html. Accessed 3.5.2024.

Griffin, S. 2022. Diabetes precision medicine: plenty of potential, pitfalls and perils but not yet ready for prime time. Diabetologia, 65(11), 1913–1921. https://doi.org/10.1007/s00125-022-05782-7. Ac-cessed 10.5.2024.

Jaggi, A. 2023. The Power of AI in Data Analytics: Enhancing prescriptive and predictive analysis. https://www.linkedin.com/pulse/power-ai-data-analytics-enhancing-prescriptive-predictive-arjun-jaggi/. Accessed 7.5.2024.

Kaggle. 2021. Drugs A, B, C, X, y for decision trees. https://www.kaggle.com/datasets/pablomgomez21/drugs-a-b-c-x-y-for-decision-trees. Accessed 15.4.2024

Kuttappa, S. 2020. Optimize healthcare delivery and reduce costs with prescriptive an-alytics. IBM Blog. https://www.ibm.com/blog/optimize-healthcare-delivery-and-reduce-costs-with-prescriptive-analytics/. Accessed 12.5.2024.

Lehtipuu, K. 2023. What is the difference between ‘de-identified’ and ‘anonymized’ data? VEIL.AI. https://veil.ai/what-is-the-difference-between-de-identified-and-anonymized-data/. Accessed 29.4.2024.

Lopes, J., Guimarães, T., & Santos, M. F. 2020. Predictive and Prescriptive Analytics in Healthcare: A Survey. Procedia Computer Science, 170, 1029–1034. https://doi.org/10.1016/j.procs.2020.03.078. Accessed 15.5.2024.

Mallawaarachchi, V. 2023. Introduction to genetic algorithms — including example code. Medium. https://towardsdatascience.com/introduction-to-genetic-algorithms-including-example-code-e396e98d8bf3#:~:text=A%20genetic%20algorithm%20is%20a,off-spring%20of%20the%20next%20generation. Accessed 25.4.2024.

Meskó, B., Drobni, Z., Bényei, É., Gergely, B., & Győrffy, Z. 2017. Digital health is a cultural transformation of traditional healthcare. mHealth, 3, 38. https://doi.org/10.21037/mhealth.2017.08.07. Accessed 30.4.2024.

MindTools s.a. Heuristic Methods. https://www.mindtools.com/a01ufjx/heuristic-methods. Accessed 3.5.2024.

Molakal, G. 2024. Prescriptive Analytics in Healthcare: The Benefits of AI - STAMOD. Stamod. https://stamod.com/prescriptive-analytics-in-healthcare-the-benefits-of-ai/. Accessed 12.5.2024.

Naithani, N., Sinha, S., Misra, P., Vasudevan, B., & Sahu, R. 2021. Precision medicine: Concept and tools. medical journal armed forces india, 77(3), 249-257. Accessed 19.4.2024.

Noteboom, C., Zeng, D., Sutrave, K., Behrens, A., Godasu, R., & Chauhan, A. 2022. Data-Driven Clinical Decision Support Systems Theory and research. In IGI Global eBooks (pp. 1373–1390). https://doi.org/10.4018/978-1-7998-9220-5.ch081. Accessed 12.5.2024.

Perry, Y. 2019. Meeting Data Compliance with a Wave of New Privacy Regulations: GDPR, CCPA, PIPEDA, POPI, LGPD, HIPAA, PCI-DSS, and More. https://bluexp.netapp.com/blog/data-compliance-regulations-hipaa-gdpr-and-pci-dss. Accessed 25.4.2024.

Sarahedwards. 2024. What is Prescriptive Analytics? Definition & Uses in Healthcare. USF Health Online. https://www.usfhealthonline.com/resources/healthcare-analytics/prescriptive-analytics/. Accessed 15.5.2024.

Sruthi. 2024. The role of big data analytics in healthcare decision making. Ezovion. Re-trieved May 20, 2024, from https://ezovion.com/the-role-of-big-data-analytics-in-healthcare-decision-making/. Accessed 25.5.2024.

Taiwo, O. J. 2024. The role of machine learning in predictive analytics. https://www.linkedin.com/pulse/role-machine-learning-predictive-analytics-o-johnson-taiwo-mba-1101e/. Accessed 30.5.2024.

ThoughtSpot, T. 2024. What is healthcare analytics? Definition, benefits, and examples. ThoughtSpot. https://www.thoughtspot.com/data-trends/analytics/healthcare-analytics. Accessed 21.4.2024.

Tiao, S. 2024. What is Big Data? https://www.oracle.com/big-data/what-is-big-data/. Accessed 1.4.2024.

Vela, M. B., Erondu, A. I., Smith, N. A., Peek, M. E., Woodruff, J. N., & Chin, M. H. 2022. Eliminating explicit and implicit biases in health care: evidence and research needs. Annual Review of Public Health, 43(1), 477–501. https://doi.org/10.1146/annurev-publhealth-052620-103528. Accessed 15.4.2024.

Vicente, V. 2023. Security Audits: A Comprehensive Overview. AuditBoard. https://www.audit-board.com/blog/what-is-security-audit/. Accessed 27.4.2024.

Zhang, D. 2021. Exploring Classifiers with Python Scikit-learn — Iris Dataset. Medium. https://towardsdatascience.com/exploring-classifiers-with-python-scikit-learn-iris-dataset-2bcb490d2e1b. Accessed 31.3.2024.