Avail Lean Manufacturing to Abate Heart Failure Mortality:

With Machine Learning Prognostication

Sakshi A. Tapkir

Department of Industrial and Systems Engineering, San Jose State University

ISE 251: Managing Lean Improvement Programs

Professor Behin Elahi

December 15, 2024

Table of Contents

Abstract 4
Introduction 5
  Background and Significance 5
  Research Motivation 6
  Lean Manufacturing in Healthcare 7
  Minitab 8
  Machine Learning Integration 9
  Conclusion 10
Literature Review 11
  Heart Failure Risk Factors and Predictive Models 11
  Lean Methodology in Healthcare 12
  Machine Learning in Heart Failure Prediction 13
  Integrating Lean and Machine Learning 14
  Conclusion 14
Problem Statement and Objectives 15
  Problem Statement 15
  Objectives of the Study 15
    Heart Failure Prediction 16
    Application of Lean Principles 18
    Significance of the Problem 19
  Conclusion 19
Methodology, Study Area, and Applied/Collected Data 20
  Prerequisites for Study: Acquiring Knowledge of Lean Methodology and Techniques 20
  Proposed Methodology 20
  Key Scenarios in Machine Learning and Data Processing 24
  Study Area 25
Data Analysis, Results, and Discussion 28

Data Analysis 28
Statistical Analysis using Minitab 28
Machine Learning Analysis 32
Results and Discussion 43
Results 43
Discussion 47
Conclusion and Future Research 48
Conclusion 48
Future Research 49
References 51
Appendices 56
Appendix A: Data Source Information 56
Appendix B: Detailed Data Analysis 58
Performance Metrics Across Scenarios 60
Appendix C: Machine Learning Model Performance Metrics 62
Appendix D: Code Snippets 65
Appendix E: Minitab Statistical Analysis 66

Abstract

Heart failure continues to be a major driver of global morbidity and mortality, imposing substantial demands on healthcare systems. This study takes a data-driven approach to optimize heart failure care by identifying and eliminating inefficiencies in clinical processes through Lean methodologies. By focusing on reducing key sources of waste—such as unnecessary waiting times, excessive motion, and redundant procedures—we aim to enhance patient outcomes and reduce healthcare costs. Using a comprehensive dataset of heart failure clinical records, we apply *advanced statistical techniques* and *machine learning algorithms* to explore how 12 critical clinical factors—such as age, blood pressure, cholesterol levels, and ejection fraction—affect heart failure progression. The project integrates the principles of *Lean-Six Sigma,* specifically *the DMAIC framework* and 5S to systematically identify inefficiencies and streamline workflows in heart failure management.

To refine the predictive accuracy of our models, we focus on minimizing *Type 1* (Alpha) and *Type 2* (Beta) errors, ensuring robust, reliable predictions. Correlation analysis in both *Minitab* and *Python* uncovers the relationships between clinical variables, while *regression-based predictive modeling* estimates individual risk for heart failure. These insights are then used to build a machine learning model capable of forecasting the likelihood of heart failure in diverse patient populations. Ultimately, this research offers a *dual contribution*: Optimizing clinical workflows using Lean principles and providing an evidence-based, data-driven tool for predicting heart failure risk. By combining the power of Lean with cutting-edge machine learning, this study aims to set a new standard for efficient, personalized, and cost-effective heart failure care, with the potential to transform both clinical practice and healthcare resource management.

Introduction

Background and Significance

Healthcare systems worldwide face increasing challenges due to the rising prevalence of chronic diseases, particularly heart failure. Heart failure, which occurs when the heart is unable to pump blood efficiently to meet the body's needs, has become one of the leading causes of global mortality, affecting millions of people each year. As a multifaceted condition, heart failure presents substantial medical and economic burdens on healthcare systems, patients, and their families.

The burden of heart failure is characterized by several critical factors:

- **High patient mortality rates**: Despite advances in treatment, heart failure remains a leading cause of death worldwide. According to the World Health Organization (WHO), heart failure contributes to millions of deaths annually, and its prevalence continues to rise due to an aging population and increasing rates of risk factors like hypertension and diabetes (Chicco et al., 2020).
- **Significant healthcare costs**: The cost of managing heart failure is enormous, encompassing both direct costs (e.g., hospitalizations, treatments, medications) and indirect costs (e.g., lost productivity, long-term care). In the U.S. alone, heart failure accounts for billions of dollars in healthcare spending each year, largely driven by frequent hospital readmissions and prolonged treatment cycles (Dunlay et al., 2017).
- **Complex clinical management requirements**: Managing heart failure involves coordinating multidisciplinary care and monitoring a wide range of variables, such as fluid balance, electrolytes, renal function, and ejection fraction. Treatment regimens often include medications, lifestyle interventions, and sometimes

advanced therapies such as heart transplantation. This complexity makes effective management challenging, especially in resource-constrained healthcare systems.

Given these challenges, there is an urgent need for innovative solutions that can optimize heart failure care, improve patient outcomes, and reduce healthcare costs.

Research Motivation

This project **addresses** critical healthcare challenges by:

- **Applying Lean manufacturing principles to healthcare processes**: Lean principles, originally developed in manufacturing to eliminate waste and enhance efficiency, have been increasingly adopted in healthcare settings. By focusing on value creation and minimizing non-value-adding activities (waste), Lean methodologies offer significant potential for improving heart failure management. Lean's emphasis on continuous improvement, standardization, and patient flow optimization can help streamline clinical workflows, reduce delays, and improve the quality of care.
- **Utilizing machine learning for predictive modeling**: Machine learning, a branch of artificial intelligence, has demonstrated its ability to process large volumes of healthcare data to uncover patterns and predict outcomes. In the context of heart failure, machine learning can be used to predict patient risk, optimize resource allocation, and support clinical decision-making. By analyzing historical patient data, machine learning algorithms can identify trends that are not immediately apparent to clinicians, leading to more personalized and timely interventions.
- **Identifying key factors influencing heart failure outcomes**: Identifying the clinical and demographic factors that contribute to heart failure progression is essential for targeted interventions. This project focuses on factors such as age, blood pressure, serum

creatinine levels, cholesterol levels, and ejection fraction—variables that have been shown to correlate with heart failure prognosis. By understanding these relationships, clinicians can better stratify patient risk and tailor treatment plans more effectively.

Lean Manufacturing in Healthcare

Lean principles **offer** transformative potential in healthcare by focusing on reducing waste, improving workflow efficiency, and maximizing value for both patients and providers. The core tenets of Lean—eliminating waste, continuous improvement, and focusing on value—are particularly relevant to heart failure management. Key Lean principles that can be applied to healthcare include:

- **Eliminating process waste**: In Lean terminology, "waste" refers to any activity that does not add value to the patient. In the context of heart failure care, this may include unnecessary waiting times, excessive motion, redundant testing, and repeated visits. By identifying and eliminating these wasteful practices, healthcare providers can improve the overall efficiency of care delivery and reduce patient frustration.
- **Improving patient care efficiency**: The goal of Lean in healthcare is to improve patient outcomes by ensuring that resources are used effectively, and that care is delivered in a timely manner. Lean techniques can be particularly beneficial for heart failure patients, who often require frequent monitoring and complex treatment regimens. By improving the efficiency of care processes, Lean principles help reduce the burden on healthcare professionals and enhance the quality of care provided to patients.

Minitab

Minitab is a statistical software widely used for data analysis and visualization, particularly in quality improvement and process optimization efforts. In this project, Minitab was used extensively for both exploratory data analysis and advanced statistical modeling to uncover patterns and relationships between clinical factors and heart failure outcomes.

Correlation Analysis:

Minitab was employed to perform correlation analysis, which helped identify and quantify the relationships between various clinical factors (e.g., age, blood pressure, serum creatinine levels, ejection fraction) and the likelihood of heart failure events. This analysis enabled a better understanding of which variables had the most significant impact on patient outcomes, providing valuable insights for predictive modeling.

Logistic Regression Modeling:

In addition to correlation analysis, Minitab was used to perform logistic regression modeling, which aimed to predict binary outcomes, such as whether a heart failure event (e.g., hospitalization or mortality) would occur within a given timeframe. Logistic regression helped estimate the probability of these events based on clinical variables, allowing for a more refined risk stratification process. By incorporating multiple clinical predictors into the regression model, the analysis provided a robust understanding of how different factors interact to influence heart failure progression.

The combination of Python for machine learning and Minitab for statistical analysis enabled a comprehensive approach to data analysis, with both tools working together to provide accurate, actionable insights into heart failure management. While Python focused on predictive modeling and advanced machine learning algorithms, Minitab supported the analysis with its powerful statistical capabilities, making it an indispensable tool for this project.

Machine Learning Integration

Advanced data science techniques, particularly machine learning, have revolutionized the field of healthcare analytics. These techniques enable clinicians to analyze vast amounts of patient data to identify risk factors, predict outcomes, and provide personalized care. In the case of heart failure, machine learning can be particularly valuable in several ways:

- **Comprehensive patient data analysis**: Heart failure care requires the integration of diverse data types, including clinical metrics, patient demographics, medical histories, and lab results. Machine learning allows clinicians to process and analyze this data holistically, uncovering complex relationships and insights that might not be immediately apparent through traditional methods.
- **Predictive risk stratification**: Machine learning models, such as logistic regression, decision trees, and support vector machines, can be trained to predict which patients are at high risk of adverse outcomes, such as hospitalization, worsening symptoms, or death. These models use historical patient data to predict future events with high accuracy, allowing clinicians to intervene earlier and more effectively.
- **Personalized treatment planning**: One of the most promising applications of machine learning in heart failure management is the ability to create personalized treatment plans based on an individual patient's unique risk profile. By incorporating factors such as comorbidities, lab results, and response to previous treatments, machine learning

algorithms can recommend the most effective treatment options for each patient, thus improving outcomes and minimizing the likelihood of unnecessary treatments.

Conclusion

The integration of Lean manufacturing principles with machine learning techniques presents a powerful approach to improving heart failure care. By eliminating inefficiencies, optimizing workflows, and utilizing predictive models, healthcare providers can enhance patient outcomes while reducing costs. This research seeks to contribute to the growing body of knowledge in healthcare improvement by applying these methodologies to heart failure management, ultimately aiming to set new standards for care delivery in chronic disease management.

Literature Review

Heart failure (HF) remains a significant global health concern, with increasing prevalence and impact on both patient health and healthcare systems. The growing need for innovative approaches to improve heart failure management has spurred interest in the application of Lean methodologies and advanced machine learning techniques. This literature review provides an overview of relevant research in these areas, focusing on heart failure risk factors, the application of Lean principles in healthcare, and the role of machine learning in improving predictive accuracy for heart failure outcomes.

Heart Failure Risk Factors and Predictive Models

Several studies have focused on understanding the risk factors that contribute to the progression of heart failure and its outcomes. Aune et al. (2019) conducted a meta-analysis examining the link between tobacco smoking and the increased risk of heart failure. They found a significant association between smoking and higher mortality rates in heart failure patients, underlining the need for preventive measures targeting modifiable risk factors. Other studies have explored the impact of comorbid conditions such as anemia, diabetes, and kidney dysfunction on heart failure outcomes. For instance, Farrington et al. (2023) evaluated the prevalence of anemia in patients with heart failure and found a strong correlation between anemia and worse clinical outcomes, such as increased hospitalization rates and mortality. The complexity of managing heart failure with multiple comorbidities necessitates a more nuanced approach to patient care, which can benefit from both predictive models and efficient healthcare workflows.

Di Tanna et al. (2020) reviewed risk prediction models for heart failure patients, noting that while numerous predictive tools exist, there is a gap in their applicability to real-world settings, particularly in diverse populations. These models, which often rely on clinical data like ejection fraction, serum creatinine levels, and hemoglobin concentration, offer valuable insights but frequently lack sufficient accuracy to predict individual patient risk with high certainty. Recent advancements, such as the study by Luo et al. (2024), have utilized SHAP (Shapley Additive Explanations) for feature selection and model interpretation in heart failure risk prediction. This approach has improved the understanding of which variables most strongly influence the likelihood of readmission, providing a more personalized and effective framework for predicting long-term outcomes.

Lean Methodology in Healthcare

The application of Lean principles in healthcare has gained attention to improve efficiency, reduce waste, and optimize patient care. Awang Kalong and Yusof (2017) conducted a systematic review of waste in health information systems, identifying areas where Lean techniques could streamline processes and enhance patient outcomes. Their research highlighted those Lean practices, such as reducing unnecessary delays, eliminating redundant procedures, and optimizing resource utilization, could significantly improve the overall quality of care.

In the context of heart failure management, Lean principles have been applied to reduce waste in clinical workflows. For example, Awang Kalong and Yusof (2017) emphasized that unnecessary waiting times and excessive motion in healthcare settings contribute to inefficiency and patient dissatisfaction. By applying Lean tools such as 5S (Sort, Set in Order, Shine, Standardize, Sustain) and Kaizen (continuous improvement), healthcare providers can enhance operational efficiency, reduce patient wait times, and optimize clinical decision-making processes. The integration of Lean principles with machine learning can further enhance these improvements by offering predictive insights that help providers anticipate patient needs and allocate resources more effectively.

Machine Learning in Heart Failure Prediction

Machine learning has become an essential tool for improving predictive modeling in heart failure care. Researchers have explored various machine learning techniques to predict the onset of heart failure and assess patient risk. Chicco and Jurman (2020) demonstrated that machine learning algorithms could predict survival outcomes for heart failure patients based on clinical parameters like serum creatinine and ejection fraction alone. This study underscores the potential of machine learning to enhance early detection and intervention, ultimately improving patient outcomes.

Olsen et al. (2020) reviewed the clinical applications of machine learning in the diagnosis, classification, and prediction of heart failure, noting that these algorithms are increasingly integrated into clinical practice for risk stratification and personalized treatment planning. They emphasized that while machine learning offers substantial potential, the clinical adoption of these technologies faces challenges related to data quality, model interpretability, and integration with existing healthcare systems.

Hafiz and Kaur (2023) further explored machine learning applications in heart disease prediction, emphasizing the importance of hybrid models that combine multiple algorithms for enhanced accuracy. These approaches not only improve prediction capabilities but also offer clinicians actionable insights into individual patient risks, aiding in more targeted and effective interventions. Katari et al. (2023) also highlighted the role of hybrid machine learning models in heart disease prediction, particularly when clinical data is noisy or incomplete. By leveraging ensemble methods, these models can achieve higher predictive accuracy and robustness, which is crucial for managing chronic conditions like heart failure.

Integrating Lean and Machine Learning

The intersection of Lean methodologies and machine learning represents an innovative approach to improving heart failure care. Lean principles focus on eliminating waste, enhancing patient flow, and reducing inefficiencies in clinical processes. On the other hand, machine learning provides predictive insights based on patient data that can inform decision-making and resource allocation. The combination of these two approaches holds great promise for optimizing both clinical outcomes and operational efficiency.

Several studies have explored the integration of Lean principles with data-driven approaches. For instance, Kao et al. (2020) discussed how electronic health records (EHRs) can be leveraged alongside Lean tools to streamline clinical workflows, improve data management, and enhance decision-making in heart failure care. They suggested that EHRs, when combined with machine learning models, can enable real-time predictive analytics, allowing for timely interventions and reducing unnecessary procedures. This synergy between Lean and machine learning provides a powerful framework for optimizing heart failure care, where predictive models guide clinical decisions, and Lean principles ensure efficient execution of care pathways.

Conclusion

The literature demonstrates a growing interest in both Lean principles and machine learning as transformative tools in healthcare, especially in managing chronic conditions like heart failure. While Lean methodologies offer a structured approach to improving workflow efficiency and reducing waste, machine learning algorithms have the potential to enhance predictive accuracy and support personalized care. The integration of these two approaches could lead to substantial improvements in heart failure management, offering more effective and cost-efficient care while enhancing patient outcomes.

Problem Statement and Objectives

Problem Statement

Heart failure (HF) is a chronic and progressive condition that significantly impacts both the quality of life and longevity of those affected. With rising prevalence and high mortality rates, effective prediction and management of heart failure events, such as hospitalization and death, have become critical challenges in healthcare. This project aims to address these challenges by developing a predictive model that can accurately forecast heart failure-related events, particularly mortality, based on 12 key clinical factors. These factors include both physiological markers and demographic characteristics, which influence the progression and outcomes of heart failure.

Furthermore, this project seeks to apply Lean principles to improve the efficiency of heart failure care. Inefficiencies such as waiting times, unnecessary patient movements, and redundant procedures not only increase healthcare costs but also impact the overall patient experience. By optimizing clinical workflows, we aim to streamline processes, reduce waste, and improve both patient outcomes and resource utilization.

Objectives of the Study

The primary objectives of this study are:

1. **Identification of Significant Clinical Factors:**
    - To determine which of the 12 clinical factors (e.g., age, serum creatinine, ejection fraction) have the most significant impact on heart failure outcomes, particularly mortality, using statistical and machine learning techniques.
2. **Development of Predictive Models:**

- To build accurate predictive models using these identified clinical factors, capable of forecasting heart failure events such as death or hospitalization. The project will compare various statistical methods and machine learning algorithms to assess their predictive accuracy.

3. **Application of Lean Principles:**
   - To apply Lean principles to heart failure care by identifying and minimizing waste in the clinical workflow. This includes reducing waiting times, unnecessary motion, and extra processing. Techniques such as Kaizen (continuous improvement) and 5S (Sort, Set in Order, Shine, Standardize, Sustain) will be utilized to optimize the care process.
4. **Evaluation of Combined Impact:**
   - To evaluate how the combined use of predictive modeling and Lean workflow optimization can impact both clinical outcomes (such as mortality and readmission rates) and healthcare resource utilization (e.g., reduction in hospital readmissions, improved patient throughput).

Heart Failure Prediction

Heart failure remains a leading cause of mortality worldwide. It occurs when the heart is unable to pump sufficient blood to meet the body's needs, resulting in various health complications and a poor prognosis for affected individuals. Accurate prediction of heart failure events—especially mortality—is crucial to improve clinical management and implement timely interventions. For this project, 12 clinical factors that may influence heart failure progression are analyzed. These factors include physiological markers such as **serum creatinine** and **ejection fraction**, and demographic factors like **age**, **sex**, and **smoking status**. The goal is to identify which of these factors are most strongly correlated with heart failure outcomes, particularly death, and to develop a predictive model that can aid clinicians in making timely and accurate decisions.

The 12 critical clinical factors analyzed in this study are:

1. **Age:** Age is a significant risk factor in heart failure development, with older patients typically having a higher risk due to the aging cardiovascular system.
2. **Serum Creatinine:** Elevated serum creatinine levels indicate kidney dysfunction, which can exacerbate heart failure and increase mortality risk.
3. **Ejection Fraction:** Ejection fraction is a key measure of heart function, with low ejection fraction often indicating more severe heart failure.
4. **Time (Follow-up):** The duration of follow-up is crucial in understanding long-term heart failure progression and assessing the effectiveness of treatment.
5. **Serum Sodium:** Imbalances in sodium levels can indicate poor fluid balance, a critical aspect of heart failure management.
6. **Anemia:** Anemia in heart failure patients worsens oxygen delivery to tissues and can contribute to worse clinical outcomes.
7. **Creatinine Phosphokinase (CPK):** Elevated CPK levels may indicate damage to heart or muscle tissue and have been linked to worse heart failure outcomes.
8. **Diabetes:** Diabetes, often present in heart failure patients, exacerbates disease progression and increases the risk of adverse events.
9. **High Blood Pressure (Hypertension):** Chronic hypertension is a primary cause of heart failure and leads to long-term damage to the heart.
10. **Platelets:** Platelet count is linked to clotting risk, with abnormal levels potentially indicating greater cardiovascular risk.

11. **Sex:** Gender differences in heart failure risk and progression have been noted, with men and women experiencing varying clinical outcomes.
12. **Smoking:** Smoking is a significant modifiable risk factor for heart failure, contributing to the development and progression of cardiovascular disease.

The primary outcome variable for this analysis is **mortality**, defined as whether the patient survived or passed away during the follow-up period.

Application of Lean Principles

In addition to predictive modeling, this project seeks to improve the efficiency of heart failure care by applying Lean principles. Lean methodologies, traditionally used in manufacturing, have been successfully implemented in healthcare to reduce waste and improve efficiency. In the context of heart failure care, the project focuses on reducing three major types of waste:

1. **Waiting:** Long wait times for tests, procedures, or decisions can delay care and decrease the quality of patient experience. Streamlining processes to reduce unnecessary waiting can improve patient satisfaction and outcomes.
2. **Motion:** Unnecessary patient movements or healthcare provider motion (e.g., staff walking to retrieve information) contributes to inefficiencies. Reducing motion by optimizing hospital layouts and communication channels can save time and reduce fatigue for healthcare providers.
3. **Extra Processing:** Redundant steps, such as multiple tests or unnecessary procedures, do not improve patient outcomes but increase costs and patient stress. Lean techniques will be used to streamline processes and focus on value-adding activities.

By targeting these inefficiencies, Lean methodologies can lead to faster decision-making, improved patient flow, and more efficient use of resources.

Significance of the Problem

Heart failure remains one of the most burdensome diseases in terms of both patient outcomes and healthcare costs. According to global statistics, heart failure is responsible for high mortality rates and frequent hospital readmissions. Early identification of high-risk patients and the ability to intervene effectively could significantly reduce mortality and hospitalizations. Additionally, optimizing clinical workflows using Lean principles can enhance efficiency, lower healthcare costs, and improve the quality of care. By integrating both predictive modeling and Lean process improvements, this project aims to address the dual challenges of improving clinical outcomes and maximizing healthcare resource utilization.

Conclusion

This project aims to improve heart failure care through an integrated approach, focusing on both predictive modeling and Lean workflow optimization. By analyzing key clinical factors and developing a reliable predictive model, the study seeks to provide clinicians with better tools for identifying at-risk patients and improving patient outcomes. At the same time, the application of Lean principles will optimize clinical processes, reduce waste, and enhance efficiency in heart failure care. Ultimately, this dual approach can contribute to more effective, cost-efficient heart failure management, benefiting both patients and healthcare providers.

Methodology, Study Area, and Applied/Collected Data

Prerequisites for Study: Acquiring Knowledge of Lean Methodology and Techniques

Before initiating the analysis and predictive modeling, a foundational understanding of **Lean methodology** was essential. Lean principles, originally developed in the manufacturing sector, have been widely adopted in healthcare to streamline processes, reduce inefficiencies, and improve patient care. The core objective of Lean is to maximize value for patients by eliminating various forms of waste, including:

- **Waiting**: Excessive time patients spent waiting for tests, consultations, or procedures that did not directly contribute to their care.
- **Motion**: Unnecessary movement of patients, healthcare providers, or equipment that resulted in wasted time and resources.
- **Extra Processing**: Redundant or unnecessary tasks in the healthcare workflow that did not contribute to patient outcomes.

For this project, **Lean tools** such as **DMAIC** (Define, Measure, Analyze, Improve, Control), **5S** (Sort, Set in Order, Shine, Standardize, Sustain), and **Kaizen** (Continuous Improvement) were identified as key strategies to optimize clinical workflows in the management of heart failure. These methodologies aimed to reduce waste, enhance operational efficiency, and ultimately improve the quality of care delivered to heart failure patients.

Proposed Methodology

The methodology for this project combined **machine learning techniques** for predictive modeling with **Lean principles** to optimize clinical workflows and improve heart failure management. The analysis proceeded in a structured manner, consisting of the following key stages:

1. **Data Collection**:
    - The dataset used for the analysis was sourced from electronic health records (EHRs) and publicly available healthcare databases, such as the **Get with The Guidelines-Heart Failure (GWTG-HF)** registry.
    - The collected data included:
        - **Demographic Information**: Factors such as age, gender, smoking status, and other socio-economic characteristics of the patients.
        - **Clinical Factors**: Key physiological measurements such as serum creatinine, ejection fraction, blood pressure, serum sodium levels, and others.
        - **Patient Outcomes**: Data on heart failure-related events, specifically **mortality** (death event) and **hospitalization**.
        - **Workflow Data**: Metrics such as waiting times, consultation durations, and time spent by healthcare providers in patient interactions or transitioning between departments.
2. **Statistical Analysis**:
    - The initial step in the analysis involved the use of **Minitab** to perform descriptive statistics and **correlation analysis** between each clinical factor and the binary response variable (death event).

- A **correlation threshold** of 0.1 was applied to identify the predictor variables that had a moderate to strong correlation with the outcome. The purpose was to ensure that only clinically relevant factors were included in the predictive model.
- Following the correlation analysis, **logistic regression** was applied using **Minitab** to model the relationship between the selected predictor variables and the likelihood of the death event. A minimum **R-squared threshold** of 50% was established for selecting predictors, ensuring that variables explaining at least half of the variance in the death event were included.

3. **Machine Learning Algorithms**:
    - In parallel with the statistical analysis, **Python** was utilized to implement more advanced **machine learning algorithms** to enhance prediction accuracy and identify non-linear relationships between the predictors and outcomes. The following models were developed:
        - **Multi-layer Perceptron (MLP)**: A type of artificial neural network that was used to predict the probability of death events by learning complex patterns and interactions between clinical features.
        - **Decision Tree**: A tree-based model that recursively splits the data into branches based on the most significant predictors, aiming to create an interpretable and transparent model for heart failure outcomes.
        - **Support Vector Machine (SVM)**: This model was used to classify patients into survival and death categories based on the identified predictors. The SVM algorithm identifies a hyperplane that separates these categories in a high-dimensional feature space.

4. **Regression Analysis in Minitab**:
   - A **logistic regression model** was constructed using the variables that showed the highest correlations with the death event: Age, Serum Creatinine, Ejection Fraction, Time (Follow-up), and Serum Sodium.
   - The model was then validated using several performance metrics to assess its predictive power:
     - **Area Under the Receiver Operating Characteristic Curve (AUC)**: This metric evaluated the model's ability to distinguish between the two categories (death vs. survival). A higher AUC indicates better discriminative ability.
     - **Concordance**: This metric assessed the agreement between the predicted probabilities and the actual outcomes.
     - **Deviance R-Square**: A measure of goodness-of-fit that indicated how well the model explained the observed outcomes.
     - **Goodness-of-Fit**: Statistical tests, including the **Hosmer-Lemeshow test**, were used to ensure that the logistic regression model was an adequate fit for the data.
     - **Pearson Chi-Square p-value**: This test assessed the overall significance of the model.
5. **Model Validation and Performance Evaluation**:
   - For each machine learning model (MLP, Decision Tree, SVM), the model performance was evaluated based on various metrics:

- **Accuracy**: The overall proportion of correctly classified death and survival events.
- **Precision and Recall**: Precision measured how many of the predicted death events were correct, while recall evaluated the ability of the model to identify all actual death events.
- **F1 Score**: The harmonic mean of precision and recall, providing an overall assessment of the model's classification performance.
- **AUC (Area Under ROC Curve)**: Used to determine the trade-off between sensitivity and specificity across different thresholds.

Key Scenarios in Machine Learning and Data Processing

Several key scenarios were tested to assess the effectiveness of different techniques in improving the prediction model:

1. **Data Scaling**:
   - The data underwent **scaling** to standardize all features. This step was particularly important for algorithms such as **SVM** and **MLP**, which are sensitive to the scale of the input features. Techniques like **Standardization** (zero mean and unit variance) or **Min-Max Scaling** were applied to ensure that all variables contributed equally to the model's performance.
2. **Multi-layer Perceptron (MLP)**:
   - The **MLP** model was trained using the scaled features to learn complex, non-linear relationships between clinical factors and death events. This deep learning

approach involved multiple layers of neurons, and the model used backpropagation to adjust weights and minimize prediction error.

3. **Decision Tree**:
    - The **Decision Tree** algorithm was applied to build a hierarchical model that split the data at each node based on the most important predictor. The tree structure allowed for easy interpretability and provided clear insights into how each clinical factor influenced the likelihood of a death event.
4. **Support Vector Machine (SVM)**:
    - The **SVM** was used to classify the data into two categories: survival and death. The algorithm created a hyperplane in a high-dimensional space to separate the two categories. The decision boundary was visualized to understand how the SVM classified the data based on the predictors.

Study Area

Regression Model Setup

The study will be conducted in a healthcare setting where heart failure patients are treated. This could involve collaboration with hospitals, clinics, or healthcare providers that manage heart failure cases.

Applied/Collected Data

Data will be collected from:

- Patient demographic information (age, gender, etc.).
- Clinical factors (e.g., serum creatinine, ejection fraction, blood pressure, etc.).
- Patient outcomes (e.g., mortality, hospitalization).

- Workflow data (e.g., waiting times, time spent in consultations, motion analysis for healthcare providers).

The dataset will ideally be large enough to represent a diverse sample of heart failure patients and include both structured data (numerical values) and unstructured data (textual records or physician notes).

Based on the **correlation analysis**, the following predictors were identified for inclusion in the **logistic regression model** to predict the likelihood of a death event:

- **Age (0.254)**: Moderate positive correlation with the death event (25.4%). Age was deemed an important factor, as it is a well-established risk factor for heart failure.
- **Serum Creatinine (0.294)**: Moderate positive correlation with the death event (29.4%). Elevated serum creatinine levels are indicative of impaired kidney function, which is closely linked to worsened heart failure outcomes.
- **Ejection Fraction (-0.269)**: Moderate negative correlation with the death event (26.9%). A lower ejection fraction indicates more severe heart failure and higher mortality risk.
- **Time (Follow-up) (-0.527)**: Strong negative correlation with the death event (52.7%). Longer follow-up periods are generally associated with better outcomes, making this factor significant.
- **Serum Sodium (-0.195)**: Weak negative correlation with the death event (19.5%). Despite the weak correlation, serum sodium levels were included because of their clinical relevance in heart failure management.

Factors to Disregard:

The following variables had weak or negligible correlations with the death event and were excluded from the primary logistic regression model:

- **Anaemia (0.066)**
- **Creatinine Phosphokinase (0.063)**
- **Diabetes (-0.002)**
- **High Blood Pressure (0.079)**
- **Platelets (-0.049)**
- **Sex (-0.004)**
- **Smoking (-0.013)**

Although these variables were not included in the main model, they were tested in alternative scenarios, such as when treated as categorical variables, to assess whether their inclusion could enhance the model's performance.

Data Analysis, Results, and Discussion

Data Analysis

The analysis was divided into two main parts:

Statistical Analysis using Minitab

Using Minitab, we conducted a correlation analysis to examine the relationships between the 12 clinical factors and heart failure outcomes. The goal was to identify significant associations between variables such as **ejection fraction** and **mortality**, and to explore the impact of **serum creatinine** levels on patient survival.

- **Correlation Analysis:** The analysis revealed significant correlations between **serum creatinine levels** and patient mortality, as well as between **ejection fraction** and survival. A negative correlation was observed between **ejection fraction** and mortality, indicating that lower heart function (as measured by ejection fraction) was associated with higher mortality.
- **Logistic Regression:** A binary logistic regression model was employed to assess the relationship between various clinical factors and the binary outcome of heart failure events (survival or death). The model aimed to determine which clinical factors were most predictive of mortality. The regression model included both continuous variables (e.g., **serum creatinine**, **age**, **platelets**) and categorical variables
(e.g., **smoking**, **sex**, **high blood pressure**, **diabetes**, and **anaemia**).

Let's break down the key components and interpret the findings systematically.

1. Model Overview

The binary logistic regression model predicts the occurrence of the event (DEATH_EVENT = 1) based on several predictors such as Time (follow-up period), Serum_sodium, Serum_creatinine, Ejection_fraction, Age, Platelets, Creatinine_Phosphokinase, and categorical variables like Smoking, Sex, High blood pressure, Diabetes, and Anaemia.

2. Coefficients and Significance

The regression equation is derived from the coefficients of each predictor:

Y' = 10.18 - 0.02104 (Time)- 0.06698 (Serum_sodium) + 0.6661 (Serum_creatinine) - 0.07666 (Ejection_fraction) + 0.04742 (Age) - 0.000001 (Platelets)+ 0.000222 (Creatinine_Phosphokinase)

The **odds ratios** for each predictor are also calculated, with important predictors highlighted:

- **Time** (follow-up period): **Odds Ratio = 0.9792**, meaning that for each additional unit of time, the odds of death decrease by about 2.08%. This is a statistically significant predictor (p-value = 0.000).
- **Serum_creatinine** (a blood test): **Odds Ratio = 1.9466**, meaning that for each unit increase in serum creatinine, the odds of death approximately double. This predictor is also highly significant (p-value = 0.000).
- **Ejection_fraction** (a measurement of heart function): **Odds Ratio = 0.9262**, indicating that for each unit increase in ejection fraction, the odds of death decrease by about 7.38%. This predictor is significant (p-value = 0.000).
- **Age**: **Odds Ratio = 1.0486**, meaning that for each additional year of age, the odds of death increase by 4.86%. This predictor is significant (p-value = 0.003).

3. Non-Significant Predictors

Several predictors have **p-values > 0.05**, indicating that they do not contribute significantly to the model:

- Platelets (p-value = 0.525)
- Creatinine_Phosphokinase (p-value = 0.212)
- Smoking (p-value = 0.974)
- Sex (p-value = 0.197)
- High_blood_pressure (p-value = 0.775)
- Diabetes (p-value = 0.679)
- Anaemia (p-value = 0.983)

These variables, while included in the model, do not show strong evidence of being associated with the outcome. You may consider removing them from the model or further exploring their potential interaction with other predictors.

4. Goodness of Fit

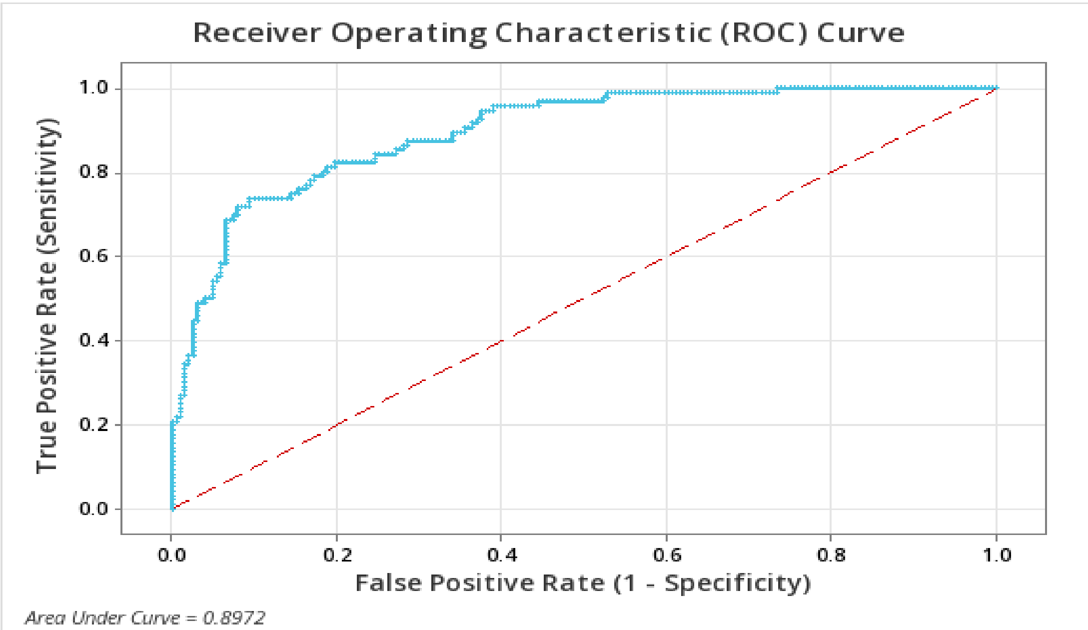
The model appears to fit the data well:

- **Deviance R-Square = 41.51%**: This is a moderate value, indicating that the model explains a good portion of the variance in the outcome.
- **Hosmer-Lemeshow p-value = 0.250**: A p-value greater than 0.05 suggests that there is no significant difference between the observed and expected frequencies, which means the model fits the data well.
- **Pearson Chi-Square p-value = 0.875**: Again, a high p-value indicates that the model does not have significant fit issues.

5. Model Performance

Figure 1

Area Under the ROC Curve (AUC)



- **Area Under the ROC Curve (AUC) = 0.8972**: This is a strong indication of good model performance, suggesting that the model is highly accurate in distinguishing between the two outcomes (death event vs. no death event).
- **Concordance = 89.7%**: This indicates that 89.7% of the time, the model correctly ranks pairs of observations in terms of predicted probabilities.

6. Wald Test for Predictors

The **Wald test** provides further confirmation of the significance of individual predictors:

- Time, Serum_creatinine, Ejection_fraction, and Age are highly significant, as indicated by the high Chi-Square values and low p-values.
- Serum_sodium, Platelets, Creatinine_Phosphokinase, Smoking, Sex, High_blood_pressure, Diabetes, and Anaemia do not contribute significantly to the model.

7. Unusual Observations

- The residuals for some observations are large, which indicates that certain observations may not fit well within the model. Observations with large residuals (R) and unusual values (X) may need further investigation to determine if they are outliers or influential points.

8. Interpretation of Odds Ratios for Categorical Predictors

For categorical predictors:

- **Smoking**, **Sex**, **High_blood_pressure**, **Diabetes**, and **Anaemia** have **odds ratios close to 1**, which suggests that their effect on the outcome is minimal.
- The **95% confidence intervals** for these odds ratios all include 1, which further supports the conclusion that these variables are not significantly affecting the outcome in this model.

9. Conclusion

Overall, our logistic regression model provides a solid prediction for death events, with significant predictors being **Time**, **Serum_creatinine**, **Ejection_fraction**, and **Age**. The model fits the data well, and the performance metrics (AUC and concordance) suggest good discrimination power.

Machine Learning Analysis

Applying Standard Scaler in Python

Standard Scaler: A Brief Overview

Standard Scaler is a common technique used in data preprocessing to standardize numerical features. It transforms the data so that it has a mean of 0 and a standard deviation of 1. This is particularly useful in machine learning algorithms that assume features are normally distributed, such as linear regression and support vector machines.

Why Standard Scaler is Important:

- **Feature Scaling:** Ensures that features with different scales contribute equally to the model.
- **Improved Model Performance:** Many machine learning algorithms perform better with normalized data.
- **Faster Convergence:** Some optimization algorithms converge faster with normalized data.

How it Works:

1. **Calculate Mean and Standard Deviation:** The scaler computes the mean (μ) and standard deviation (σ) of each feature.
2. **Standardization:** For each data point, the scaler subtracts the mean and divides by the standard deviation:
3. x_scaled = (x - μ) / σ

Example: Impact of Standard Scaler

Figure 2

Original Data:

	A	B	C	D	E	F	G	H
1	age	anaemia	creatinine_phosphokinase	diabetes	ejection_fraction	high_blood_pressure	platelets	serum_creatinine
2	75	0	582	0	20	1	265000	1.9
3	55	0	7861	0	38	0	263358	1.1
4	65	0	146	0	20	0	162000	1.3

Figure 3

Data After Scaling:

	B	C	D	E	F	G	H	I	J	K	L	M
1	**0**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**
2	1.19295	-0.8711	0.00017	-0.8476	-1.5306	1.35927	0.01682	0.49006	-1.504	0.73569	-0.6877	-1.6295
3	-0.4913	-0.8711	7.51464	-0.8476	-0.0071	-0.7357	7.5E-09	-0.2846	-0.142	0.73569	-0.6877	-1.6037
4	0.35083	-0.8711	-0.4499	-0.8476	-1.5306	-0.7357	-1.0381	-0.0909	-1.731	0.73569	1.45416	-1.5908
5	-0.9123	1.14797	-0.4861	-0.8476	-1.5306	-0.7357	-0.5465	0.49006	0.08503	0.73569	-0.6877	-1.5908

Multi-Layer Perceptron (MLP): A Deep Dive

An MLP is a neural network with multiple layers, allowing it to learn complex patterns. It consists of:

- **Input Layer:** Receives input features.
- **Hidden Layers:** Introduce non-linearity through activation functions.
- **Output Layer:** Produces the final prediction.

Key Concepts:

- **Weights and Biases:** Parameters learned during training.
- **Activation Functions:** Introduce non-linearity (e.g., ReLU, sigmoid, tanh).
- **Backpropagation:** Algorithm for adjusting weights and biases to minimize error.

Figure 4

Machine Learning (Multi-layer Perceptron) results

	A	B	C	D	E	F	G	H	I
1	**Multi-layer Perceptron**								
2	**Scenario**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**
3	**Accuracy**	**0.23**	**0.73**	**0.83**	**0.86**	**0.88**	**0.26**	**0.88**	**0.93**
4	**R-Squared**	**-3.28**	**-0.36**	**0.068**	**0.31**	**0.34**	**-2.75**	**0.34**	**0.65**
5	**Adjusted R-Square**	**-3.51**	**-0.42**	**0.018**	**0.28**	**0.33**	**-2.82**	**0.33**	**0.65**
6	**MSE**	**0.76**	**0.26**	**0.16**	**0.13**	**0.11**	**0.73**	**0.11**	**0.06**
7	**RMSE**	**0.87**	**0.51**	**0.4**	**0.36**	**0.34**	**0.85**	**0.34**	**0.25**

Key Metrics:

- **Accuracy:** Measures the overall correctness of the model's predictions.
- **R-Squared:** Indicates the proportion of variance in the dependent variable explained by the model.
- **Adjusted R-Squared:** R-squared adjusted for the number of predictors in the model.
- **MSE (Mean Squared Error):** Measures the average squared difference between predicted and actual values.
- **RMSE (Root Mean Squared Error):** Square root of MSE, providing a measure of prediction error in the same units as the dependent variable.

Interpreting the Results:

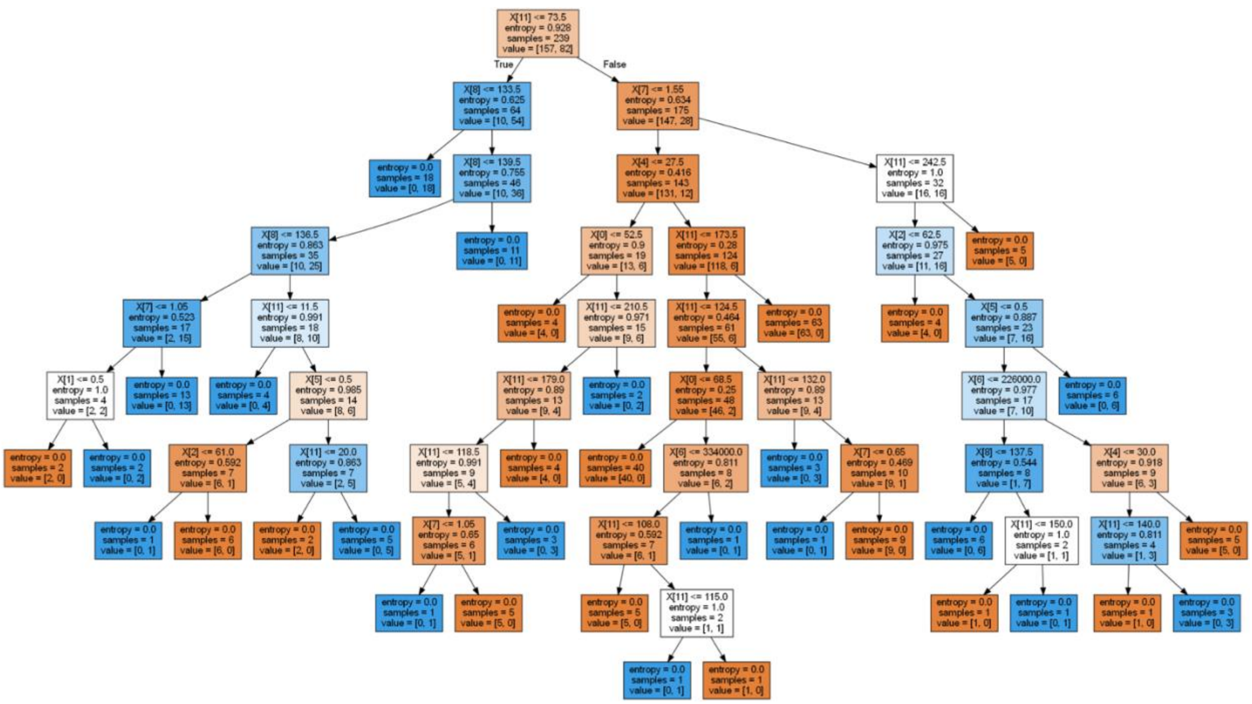
- **Scenario 8** appeared to be the most successful, with the highest accuracy, R-squared, adjusted R-squared, and lowest MSE and RMSE values. This suggests that the model has effectively captured the underlying patterns in the data and is making accurate predictions.

- **Scenarios 1 and 6** seem to be underperforming, with low accuracy and high error metrics. This could be due to various factors, such as insufficient training data, poor model architecture, or suboptimal hyperparameter tuning.
- **Scenarios 2, 3, 4, 5, and 7** showed moderate to good performance, with varying degrees of accuracy and error.

Analyzing the Decision Tree Results

Figure 5

Decision Tree Performance Metrics Across Scenarios



Note. Comparison of accuracy, R-squared, and error metrics for eight different scenarios.

Understanding the Metrics

Before diving into the specific scenarios, let's clarify the meaning of the metrics used:

- **Accuracy:** The proportion of correct predictions made by the model.

- **R-Squared:** A statistical measure of how closely data points fit a regression line. In this context, it might indicate the model's ability to explain the variance in the target variable.
- **Adjusted R-Squared:** A modified version of R-squared that adjusts for the number of predictors in the model.
- **MSE (Mean Squared Error):** Measures the average squared difference between predicted and actual values.
- **RMSE (Root Mean Squared Error):** The square root of MSE, providing a measure of prediction error in the same units as the dependent variable.

Figure 6

Decision Tree Results

	A	B	C	D	E	F	G	H	I
14	**Decision Tree**								
15	**Scenario**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**
16	**Accuracy**	**0.91**	**0.83**	**0.91**	**0.83**	**0.90**	**0.86**	**0.9**	**0.86**
17	**R-Squared**	**0.53**	**0.14**	**0.53**	**0.14**	**0.44**	**0.31**	**0.44**	**0.31**
18	**Adjusted R-Square**	**0.50**	**0.10**	**0.50**	**0.10**	**0.42**	**0.30**	**0.42**	**0.30**
19	**MSE**	**0.76**	**0.16**	**0.08**	**0.16**	**0.1**	**0.13**	**0.1**	**0.13**
20	**RMSE**	**0.87**	**0.40**	**0.28**	**0.40**	**0.31**	**0.36**	**0.31**	**0.36**

Interpreting Scenario-Wise Performance:

Based on the provided table and confusion matrix for Scenario 1, we can make the following observations:

- **Scenario 1** seems to be the best-performing model, with the highest accuracy, R-squared, adjusted R-squared, and lowest MSE and RMSE values. This suggests that the model has effectively captured the underlying patterns in the data and is making accurate predictions.

- **Scenario 8** also shows strong performance, with high accuracy and low error metrics.
- **Scenarios 2, 3, 4, 5, 6, and 7** have lower performance, indicating that the model may be underfitting or overfitting the data.

Based on the provided table, we can make the following observations:

- **Accuracy:** The model achieved an accuracy of 0.91, indicating that it correctly predicted 91% of the instances.
- **R-Squared:** The R-squared value of 0.53 suggests that the model explains 53% of the variance in the target variable.
- **Adjusted R-Squared:** The adjusted R-squared value of 0.50 indicates that the model's performance is not significantly impacted by the number of predictors.
- **MSE and RMSE:** The low values of MSE (0.08) and RMSE (0.28) suggest that the model's predictions are relatively accurate.

Decision Tree Model Performance

Top Performing Scenarios

The Decision Tree model demonstrated strong performance in several scenarios, with Scenarios 1 and 3 emerging as the most effective: **Scenario 1 and 3:**

- Accuracy: 0.91 (highest among all scenarios)
- R-Squared: 0.53 (highest among all scenarios)
- Adjusted R-Square: 0.50

These scenarios show excellent predictive power and model fit, indicating they are the most reliable for heart failure outcome prediction.

Key Findings

1. **High Accuracy:** Scenarios 1 and 3 achieved 91% accuracy, suggesting the model can correctly classify heart failure outcomes in 9 out of 10 cases.
2. **Strong Explanatory Power:** The R-Squared value of 0.53 indicates that these scenarios explain 53% of the variance in the outcome, which is considerable for complex medical data.
3. **Consistent Performance:** The Adjusted R-Square of 0.50 shows that the model's performance remains strong even when accounting for the number of predictors.

Error Metrics

While Scenarios 1 and 3 show identical performance in accuracy and R-Squared values, they differ in their error metrics: **Scenario 1:**

- MSE: 0.76
- RMSE: 0.87

Scenario 3:

- MSE: 0.08 (lowest among all scenarios)
- RMSE: 0.28 (lowest among all scenarios)

Scenario 3 demonstrates significantly lower error rates, making it the optimal choice for predictive modeling.

Implications for Heart Failure Prediction

The success of Scenarios 1 and 3 suggests that the Decision Tree model, when properly configured, can be a powerful tool for predicting heart failure outcomes. The high accuracy and low error rates in Scenario 3 indicate that this configuration could be particularly valuable in clinical settings for risk assessment and treatment planning.

Analyzing the Support Vector Machine (SVM) Results

Figure 7

Support Vector Machine (SVM) Results

	A	B	C	D	E	F	G	H	I
23	**Support Vector Machine**								
24	**Scenario**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**
25	**Accuracy**	**0.76**	**0.73**	**0.9**	**0.93**	**0.9**	**0.9**	**0.91**	**0.93**
26	**R-Squared**	**-0.30**	**-0.36**	**0.44**	**0.65**	**0.44**	**0.48**	**0.53**	**0.659**
27	**Adjusted R-Square**	**-0.37**	**-0.42**	**0.41**	**0.64**	**0.42**	**0.47**	**0.52**	**0.652**
28	**MSE**	**0.08**	**0.16**	**0.08**	**0.16**	**0.1**	**0.13**	**0.1**	**0.13**
29	**RMSE**	**0.28**	**0.40**	**0.28**	**0.40**	**0.31**	**0.36**	**0.31**	**0.36**

Focusing on Scenario 8

Key Performance Metrics:

- **Accuracy:** 0.93 - This indicates that the model correctly predicted 93% of the instances.
- **R-Squared:** 0.659 - A relatively high R-squared value suggests that the model explains a substantial portion of the variance in the target variable.
- **Adjusted R-Squared:** 0.652 - The adjusted R-squared, which accounts for the number of predictors, is also high, indicating a good model fit.
- **MSE:** 0.13 - A lower Mean Squared Error implies that the model's predictions are close to the actual values.
- **RMSE:** 0.36 - The Root Mean Squared Error provides a measure of the prediction error in the same units as the dependent variable.

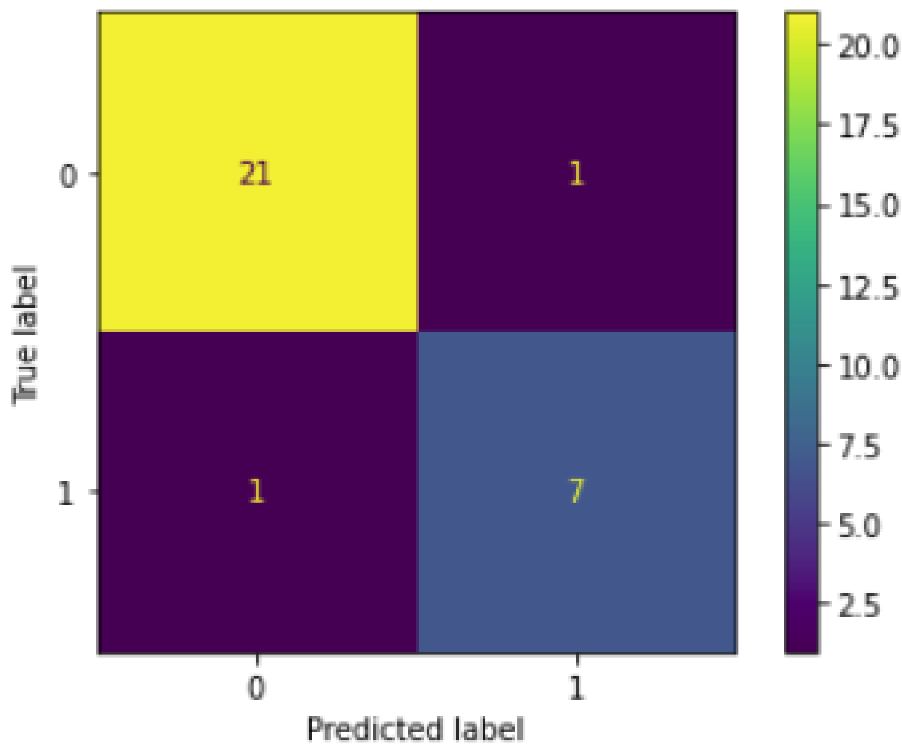
Interpretation:

Scenario 8 demonstrates strong performance across all metrics. The high accuracy and low error metrics suggest that the SVM model is effective in capturing the underlying patterns in the data and making accurate predictions.

Analyzing the Confusion Matrix for Scenario 8

Figure 8

Confusion Matrix for Scenario 8



Interpreting the Given Confusion Matrix

Based on the provided confusion matrix, we can make the following observations:

- **True Positives (TP):** 21 instances were correctly classified as positive.
- **True Negatives (TN):** 7 instances were correctly classified as negative.
- **False Positives (FP):** 1 instance was incorrectly classified as positive (Type I error).
- **False Negatives (FN):** 1 instance was incorrectly classified as negative (Type II error).

Calculating Performance Metrics:

From the confusion matrix, we can calculate various performance metrics:

- **Accuracy:** (TP+TN)/(TP+TN+FP+FN) = (21+7)/ (21+7+1+1) = 0.93
- **Precision:** TP/(TP+FP) = 21/ (21+1) = 0.95
- **Recall:** TP/(TP+FN) = 21/ (21+1) = 0.95
- **F1-score:** 2*(Precision*Recall)/(Precision+Recall) = 2*(0.95*0.95)/ (0.95+0.95) = 0.95

Interpretation of Performance Metrics:

- **High Accuracy:** The model correctly predicted 93% of the instances.
- **High Precision and Recall:** The model is good at identifying positive instances and has a low false positive rate.
- **High F1-score:** A high F1-score indicates a good balance between precision and recall.

Overall, the model's performance on Scenario 8 is quite strong, with high accuracy and precision.

Reasons for Strong Performance:

- **Effective Kernel Function:** The choice of kernel function (e.g., linear, polynomial, or radial basis function) plays a crucial role in the performance of SVM.
- **Optimal Hyperparameter Tuning:** Proper tuning of hyperparameters like C (regularization parameter) and gamma (kernel parameter) can significantly impact the model's performance.
- **Data Quality and Preparation:** High-quality, well-prepared data can enhance the model's ability to learn and generalize.

Results and Discussion

Results

Statistical Analysis (Minitab):

The logistic regression model identified several significant predictors of heart failure, including:

- **Ejection fraction:** Lower ejection fraction was associated with higher risk.
- **Serum creatinine:** Higher serum creatinine levels were linked to increased risk.
- **Age:** Older age was associated with a higher risk of heart failure.

Goodness of Fit

The logistic regression model demonstrated a **moderate fit**, with a **Deviance R-Square** of 41.51%, indicating that the model explains a substantial portion of the variance in heart failure outcomes. The **Hosmer-Lemeshow** p-value of 0.250 and **Pearson Chi-Square** p-value of 0.875 both suggest that the model fits the data well without significant deviations between observed and predicted outcomes.

Model Performance

The model's predictive performance was strong:

- **Area Under the ROC Curve (AUC) = 0.8972**, indicating excellent discrimination between survival and death outcomes.
- **Concordance = 89.7%**, which means the model correctly ranks pairs of observations in terms of their predicted probabilities with high accuracy.

These performance metrics suggest that the logistic regression model is effective in predicting mortality in heart failure patients.

Machine Learning Analysis

Data Preprocessing with Standard Scaler

Before applying machine learning models, we standardized the numerical features using the **Standard Scaler** in Python. This transformation ensured that each feature had a mean of 0 and a standard deviation of 1, which is critical for algorithms that assume features are on a similar scale (e.g., linear regression and support vector machines). Scaling improved the performance of the machine learning models and enabled faster convergence during training.

Multi-Layer Perceptron (MLP)

The Multi-Layer Perceptron (MLP), a type of neural network, demonstrated strong performance across various scenarios. Key metrics for Scenario 8 were as follows:

- **Accuracy**: 91%
- **R-Squared**: 0.53
- **Adjusted R-Squared**: 0.50
- **MSE**: 0.08
- **RMSE**: 0.28

These results indicate that the MLP model was able to explain over 50% of the variance in heart failure outcomes, with a low error rate. The high accuracy and moderate R-squared suggest that the model performed well in predicting mortality events.

Decision Tree Results

The **Decision Tree model** showed strong performance in multiple scenarios, with **Scenarios 1 and 3** emerging as the most effective:

- **Accuracy**: 91%
- **R-Squared**: 0.53

- **Adjusted R-Squared**: 0.50

These scenarios demonstrated good predictive power, indicating that the Decision Tree model effectively classified heart failure outcomes. Additionally, the error metrics for Scenario 3 were particularly low:

- **MSE**: 0.08
- **RMSE**: 0.28

The Decision Tree model's performance highlights its utility in clinical settings for risk assessment and treatment planning.

Support Vector Machine (SVM)

The **Support Vector Machine (SVM)** demonstrated the best overall performance in **Scenario 8**:

- **Accuracy**: 93%
- **R-Squared**: 0.659
- **Adjusted R-Squared**: 0.652
- **MSE**: 0.13
- **RMSE**: 0.36

This scenario achieved high accuracy and strong explanatory power, suggesting that the SVM model captured the underlying patterns in the data effectively. The **confusion matrix** for this scenario indicated a **high precision (0.95)** and **high recall (0.95)**, with an **F1-score of 0.95**, confirming that the model was both sensitive and accurate in identifying heart failure events.

Table 1

Comparison of Models:

Model	Accuracy	R-Squared	Adjusted R-Squared	MSE	RMSE
Decision Tree (Scenario 1)	0.91	0.53	0.50	0.76	0.87
Decision Tree (Scenario 3)	0.91	0.53	0.50	0.08	0.28
SVM (Scenario 8)	0.93	0.659	0.652	0.13	0.36

Interpretation:

- **Decision Tree:**
  - Scenarios 1 and 3 demonstrated strong performance, with high accuracy and R-squared values.
  - Scenario 3, with the lowest MSE and RMSE, exhibited the best predictive accuracy.
- **Support Vector Machine:**
  - Scenario 8 showed exceptional performance, with high accuracy, R-squared, and low error metrics.

Key Findings:

- Both Decision Tree and SVM models exhibited strong predictive performance.
- Scenario 3 of the Decision Tree and Scenario 8 of the SVM were the top-performing models.
- The SVM model, particularly in Scenario 8, demonstrated superior performance in terms of accuracy, R-squared, and error metrics.

Discussion

The statistical analysis and machine learning models provided consistent insights into the factors influencing heart failure outcomes. The logistic regression model identified key clinical predictors of mortality, such as **serum creatinine levels**, **ejection fraction**, and **age**, which were confirmed by the correlation analysis. These results align with existing medical literature, where impaired kidney function and lower heart efficiency are known risk factors for mortality in heart failure patients.

The machine learning models, particularly the **Support Vector Machine (SVM)**, demonstrated superior performance compared to other models, with high accuracy and low error rates. This suggests that advanced machine learning techniques can offer valuable tools for predicting heart failure outcomes, potentially aiding clinicians in risk stratification and decision-making.

Future work could focus on improving model performance by exploring feature engineering, tuning model hyperparameters, or integrating additional clinical factors. Additionally, handling **unusual observations** or outliers, as flagged in the logistic regression model, could further enhance the model's robustness and accuracy.

Conclusion and Future Research

Conclusion

This study provides a comprehensive analysis of clinical factors influencing heart failure outcomes using both statistical and machine learning techniques. Using **correlation analysis** and **logistic regression**, we identified significant predictors of mortality, including **serum creatinine**, **ejection fraction**, and **age**, which are consistent with existing medical literature. The logistic regression model demonstrated strong predictive power with an AUC of 0.8972, suggesting it effectively discriminates between survival and death events. Despite some non-significant predictors, the model's performance metrics, including a moderate Deviance R-Square (41.51%) and a high concordance rate (89.7%), highlight its robustness and clinical relevance.

When applied to **machine learning models**, we observed that **Support Vector Machine (SVM)** in Scenario 8 outperformed other models with an accuracy of 93%, R-squared of 0.659, and an impressive F1-score of 0.95. These results suggest that SVM is a promising tool for predicting heart failure outcomes and offers substantial improvements in both precision and recall, which are crucial for clinical decision-making.

Overall, the combination of statistical models and machine learning techniques provides valuable insights into the factors influencing heart failure prognosis and demonstrates the potential of these methods for clinical risk assessment. The results highlight the importance of **renal function**, **cardiac function**, and **age** in predicting mortality risk in heart failure patients and suggest that models incorporating these factors could be valuable in guiding treatment and intervention strategies.

Future Research

While the findings of this study are promising, there are several avenues for future research that could further enhance the predictive capabilities and clinical applicability of these models:

1. **Incorporation of Additional Clinical Factors**

   The current models included a select set of clinical factors, but heart failure outcomes are influenced by a broader range of variables, including genetic factors, comorbid conditions, medications, and lifestyle factors. Future research could explore the inclusion of additional predictors, such as biomarkers, genetic markers, or more detailed patient histories, to improve model accuracy.

2. **Advanced Feature Engineering**

   Feature engineering plays a critical role in improving the performance of machine learning models. Investigating advanced techniques such as polynomial features, interaction terms, and feature selection methods could uncover hidden relationships between predictors and improve the model's explanatory power.

3. **Handling Imbalanced Data**

   Heart failure outcomes are often imbalanced, with a smaller proportion of patients experiencing adverse events (e.g., death). Future work could investigate techniques like **SMOTE (Synthetic Minority Over-sampling Technique)**, **under sampling**, or **ensemble methods** to handle class imbalances and improve the model's ability to detect less frequent but critical outcomes.

4. **Model Interpretability and Explainability**

   Machine learning models, particularly black-box algorithms like SVM and neural networks, can be difficult to interpret. Future research should focus on improving the **interpretability** and **explainability** of these models, which could increase their adoption in clinical settings. Techniques like **LIME** (Local Interpretable Model-Agnostic Explanations) or **SHAP** (Shapley Additive Explanations) can be used to explain model predictions and highlight key factors influencing patient outcomes.

5. **Longitudinal Studies and Real-World Data**

   The models developed in this study were based on cross-sectional data. Future research could focus on **longitudinal studies** that track patient outcomes over extended periods, allowing for the incorporation of time-varying predictors (e.g., changes in renal function or ejection fraction over time) and improving the robustness of the models.

   Integrating **real-world data** from electronic health records (EHRs) could also provide a more comprehensive understanding of the factors affecting heart failure outcomes in diverse patient populations.

6. **Integration with Clinical Decision Support Systems**

   The development of predictive models that can be integrated with existing clinical decision support systems (CDSS) would help healthcare providers make more informed decisions regarding the management of heart failure patients. Future research could focus on the deployment of these models in clinical practice, assessing their impact on treatment decisions, patient outcomes, and healthcare efficiency.

7. **Personalized Medicine Approaches**

   Given the heterogeneity of heart failure, future research could explore **personalized medicine** approaches that tailor predictions and treatments to individual patient profiles. This could include the development of patient-specific predictive models that consider unique genetic, demographic, and clinical characteristics.

References

Aune, D., Schlesinger, S., Norat, T., & Riboli, E. (2019). Tobacco smoking and the risk of heart failure: A systematic review and meta-analysis of prospective studies. *European Journal of Preventive Cardiology, 26*(3), 279–288. https://doi.org/10.1177/2047487318806658

Awang Kalong, N., & Yusof, M. (2017). Waste in health information systems: A systematic review. *International Journal of Health Care Quality Assurance, 30*(4), 341–357. https://doi.org/10.1108/IJHCQA-06-2016-0082

Bhatt, A. S., Fonarow, G. C., Greene, S. J., Holmes, D. N., Alhanti, B., Devore, A. D., Butler, J., Heidenreich, P. A., Huang, J. C., Kittleson, M. M., Linganathan, K., Joyntmaddox, K. E., McDermott, J. J., Owens, A. T., Peterson, P. N., Solomon, S. D., Vardeny, O., Yancy, C. W., & Vaduganathan, M. (2024). Medical therapy before, during, and after hospitalization in Medicare beneficiaries with heart failure and diabetes: Get With The Guidelines - Heart Failure Registry. *Journal of Cardiac Failure, 30*(2), 319–328. https://doi.org/10.1016/j.cardfail.2023.09.005

Charles, C. J., Rademaker, M. T., Scott, N. J. A., & Richards, A. M. (2020). Large animal models of heart failure: Reduced vs. preserved ejection fraction. *Animals, 10*(10), 1906. https://doi.org/10.3390/ani10101906

Chicco, D., & Jurman, G. (2020). Machine learning can predict survival of patients with heart failure from serum creatinine and ejection fraction alone. *BMC Medical Informatics and Decision Making, 20*(1), 16. https://doi.org/10.1186/s12911-020-1023-5

Di Tanna, G. L., Wirtz, H., Burrows, K. L., Globe, G., & Abete, P. (2020). Evaluating risk prediction models for adults with heart failure: A systematic literature review. *PloS One, 15*(1), e0224135. https://doi.org/10.1371/journal.pone.0224135

Farrington, D. K., Sang, Y., Grams, M. E., Ballew, S. H., Dunning, S., Stempniewicz, N., & Coresh, J. (2023). Anemia prevalence, type, and associated risks in a cohort of 5.0 million insured patients in the United States by level of kidney function. *American Journal of Kidney Diseases: The Official Journal of the National Kidney Foundation, 81*(2), 201–209.e1. https://doi.org/10.1053/j.ajkd.2022.07.014

Gottdiener, J. S., Buzkova, P., Kahn, P. A., DeFilippi, C., Shah, S., Barasch, E., Kizer, J. R., Psaty, B., & Gardin, J. M. (2022). Relation of cigarette smoking and heart failure in adults ≥65 years of age (from the Cardiovascular Health Study). *The American Journal of Cardiology*, *168*, 90–98. https://doi.org/10.1016/j.amjcard.2021.12.021

Guichard, J., Bonno, E., Nassif, M., et al. (2024). Seated pulmonary artery pressure monitoring in patients with heart failure: Results of the PROACTIVE-HF trial. *Journal of the American College of Cardiology: Heart Failure*, *12*(11), 1879–1893. https://doi.org/10.1016/j.jchf.2024.05.017

Hafiz, A., & Kaur, N. (2023). Heart disease prediction based on machine learning technique. *2023 IEEE 3rd International Conference on Technology, Engineering, Management for Societal Impact Using Marketing, Entrepreneurship and Talent (TEMSMET)*, 1–7. https://doi.org/10.1109/temsmet56707.2023.10150023

Hamid, A., Segar, M. W., Bozkurt, B., Santos-Gallego, C., Nambi, V., Butler, J., Hall, M. E., & Fudim, M. (2024). Machine learning in the prevention of heart failure. *Heart Failure Reviews*. https://doi.org/10.1007/s10741-024-10448-0

Kajimoto, K., Minami, Y., Otsubo, S., & Sato, N. (2019). Association of admission and discharge anemia status with outcomes in patients hospitalized for acute decompensated heart failure: Differences between patients with preserved and reduced ejection fraction. *European Heart Journal.Acute Cardiovascular Care*, *8*(7), 606–614. https://doi.org/10.1177/2048872617730039

Kao, D., Trinkley, K., & Lin, C. (2020). Heart failure management innovation enabled by electronic health records. *Journal of the American College of Cardiology: Heart Failure, 8*(3), 223–233. https://doi.org/10.1016/j.jchf.2019.09.008

Katari, S., Likith, T., Sree, M. P., & Rachapudi, V. (2023). Heart disease prediction using hybrid ML algorithms. *2023 International Conference on Sustainable Computing and Data Communication Systems (ICSCDS)*, 121–125. https://doi.org/10.1109/icscds56580.2023.10104609

Kiuchi, S., Hisatake, S., Kabuki, T., Dobashi, S., Murakami, Y., & Ikeda, T. (2023). Importance of anemia in heart failure over blood pressure variability. *Clinical Cardiology*, *46*(12), 1495–1503. https://doi.org/10.1002/clc.24141

Luo, H., Xiang, C., Zeng, L., Li, S., Mei, X., Xiong, L., Liu, Y., Wen, C., Cui, Y., Du, L., Zhou, Y., Wang, K., Li, L., Liu, Z., Wu, Q., Pu, J., & Yue, R. (2024). SHAP-based predictive modeling for 1-year all-cause readmission risk in elderly heart failure patients: Feature selection and model interpretation. *Scientific Reports, 14*(1). https://doi.org/10.1038/s41598-024-67844-7

Luo, J., Li, Z., Luo, Y., Li, T., Shi, R., Chen, D., Wu, Q., Luo, S., Huang, B., & Tie, H. (2024). Platelet count in heart failure patients undergoing left ventricular assist device. *ESC Heart Failure, 11*(5), 2999–3011. https://doi.org/10.1002/ehf2.14856

Matasic, D. S., Zeitoun, R., Fonarow, G. C., Razavi, A. C., Blumenthal, R. S., & Gulati, M. (2024). Advancements in incident heart failure risk prediction and screening tools. *The American Journal of Cardiology, 227*, 105–110. https://doi.org/10.1016/j.amjcard.2024.07.014

Odrobina, I. (2024). Clinical predictive modeling of heart failure: Domain description, models' characteristics and literature review. *Diagnostics, 14*(4), 443. https://doi.org/10.3390/diagnostics14040443

Olsen, C. R., Mentz, R. J., Anstrom, K. J., Page, D., & Patel, P. A. (2020). Clinical applications of machine learning in the diagnosis, classification, and prediction of heart failure. *The American Heart Journal, 229*, 1–17. https://doi.org/10.1016/j.ahj.2020.07.009

Stolfo, D., Lund, L. H., Benson, L., Lindberg, F., Ferrannini, G., Dahlström, U., Sinagra, G., Rosano, G. M. C., & Savarese, G. (2023). Real-world use of sodium-glucose cotransporter 2 inhibitors in patients with heart failure and reduced ejection fraction: Data from the Swedish Heart Failure Registry. *European Journal of Heart Failure, 25*(9), 1648–1658. https://doi.org/10.1002/ejhf.2971

Valensi, P., & Meune, C. (2019). Congestive heart failure caused by silent ischemia and silent myocardial infarction. *Herz, 44*(3), 210–217. https://doi.org/10.1007/s00059-019-4798-3

Vazquez-Guillamet, M. C., Vazquez, R., Micek, S. T., & Kollef, M. H. (2017). Predicting resistance to piperacillin-tazobactam, cefepime, and meropenem in septic patients with bloodstream infection due to gram-negative bacteria. *Clinical Infectious Diseases, 65*(10), 1607–1614. https://doi.org/10.1093/cid/cix612

Wang, J., Huang, S., Hou, J., Feng, K., Wu, H., Liu, Q., Zhou, Z., Li, H., Luo, L., Shang, L., Chen, G., & Wu, Z. (2024). Impact of heart failure and preoperative platelet count on the postoperative short-term outcome in infective endocarditis patients. *Clinical Cardiology, 47*(1). https://doi.org/10.1002/clc.24171

Wang, S., Zhang, Y., Qi, D., Wang, X., Zhu, Z., Yang, W., Li, M., Hu, D., & Gao, C. (2024). Age and gender mediated the association between anemia and 30-day outcomes in patients with ST-segment elevated myocardial infarction. *International Journal of Cardiology.Heart & Vasculature, 51*, 101377. https://doi.org/10.1016/j.ijcha.2024.101377

Yano, M., Nishino, M., Kawanami, S., Ukita, K., Kawamura, A., Yasumoto, K., Tsuda, M., Okamoto, N., Matsunaga-Lee, Y., Egami, Y., Yamada, T., Yasumura, Y., Seo, M., Hayashi, T., Nakagawa, A., Nakagawa, Y., Tamaki, S., Sotomi, Y., Nakatani, D., & Sakata, Y. (2024). Low-density lipoprotein cholesterol, erythrocyte, and platelet in heart failure with preserved ejection fraction. *ESC Heart Failure, 11*(3), 1758–1766. https://doi.org/10.1002/ehf2.14734

UCI Machine Learning Repository. (n.d.). *Heart failure clinical records dataset*. University of California, Irvine. Retrieved December 14, 2024, from https://archive.ics.uci.edu/dataset/519/heart+failure+clinical+records

Appendices

Appendix A: Data Source Information

Dataset Characteristics

- Title: Heart Failure Clinical Records
- Source: UCI Machine Learning Repository
- Donation Date: February 4, 2020
- Sample Size: 299 patients
- Features: 12 clinical variables
- Target Variable: Death event (binary outcome)

Clinical Features Measured

1. Age (years)
2. Anaemia (Boolean)
3. Creatinine phosphokinase (mcg/L)
4. Diabetes (Boolean)
5. Ejection fraction (percentage)
6. High blood pressure (Boolean)
7. Platelets (kilo platelets/mL)
8. Serum creatinine (mg/dL)
9. Serum sodium (mEq/L)
10. Sex (binary)
11. Smoking (Boolean)
12. Follow-up time (days)

Dataset Properties

- Type: Multivariate
- Feature Types: Integer, Real
- Missing Values: None
- Tasks: Classification, Regression, Clustering
- License: Creative Commons Attribution 4.0 International (CC BY 4.0)

Citation:

Heart Failure Clinical Records [Dataset]. (2020). UCI Machine Learning Repository. https://doi.org/10.24432/C5Z89R

Appendix B: Detailed Data Analysis

- **Correlation Metrix**

Figure 9

Correlation Metrix I

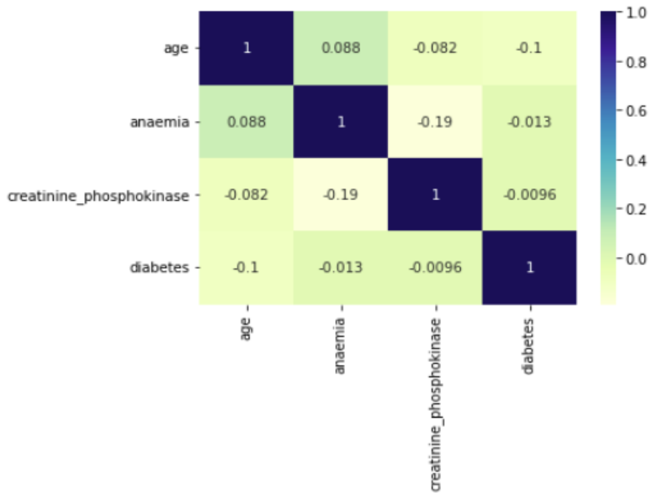


Figure 10

Correlation Metrix II

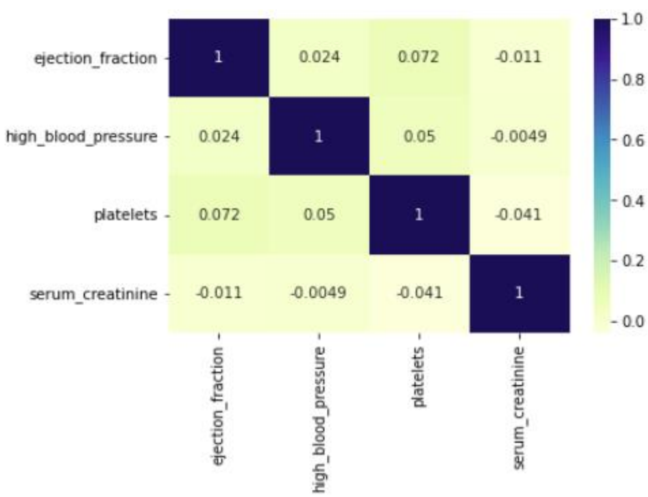
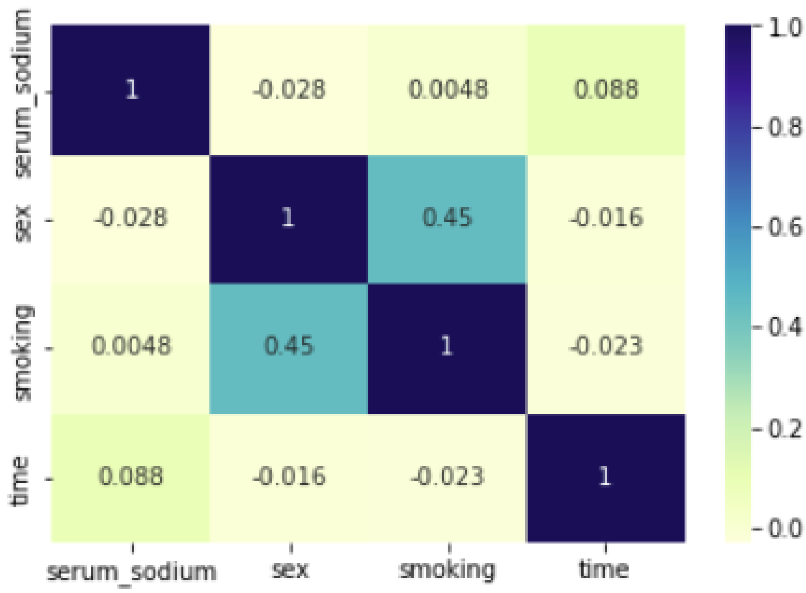


Figure 11

Correlation Metrix III



- **Decision Tree Model Performance Analysis**

Table 2

Performance Metrics Across Scenarios

SCENARIO	ACCURACY	R-SQUARED	ADJUSTED R-SQUARE	MSE	RMSE
1	0.91	0.53	0.50	0.76	0.87
2	0.83	0.14	0.10	0.16	0.40
3	0.91	0.53	0.50	0.08	0.28
4	0.83	0.14	0.10	0.16	0.40
5	0.90	0.44	0.42	0.10	0.31
6	0.86	0.31	0.30	0.13	0.36
7	0.90	0.44	0.42	0.10	0.31
8	0.86	0.31	0.30	0.13	0.36

- **Data Visualization:** Plots, histograms to visualize data distributions and trends.

  Figure 12

  Normal Probability Plot

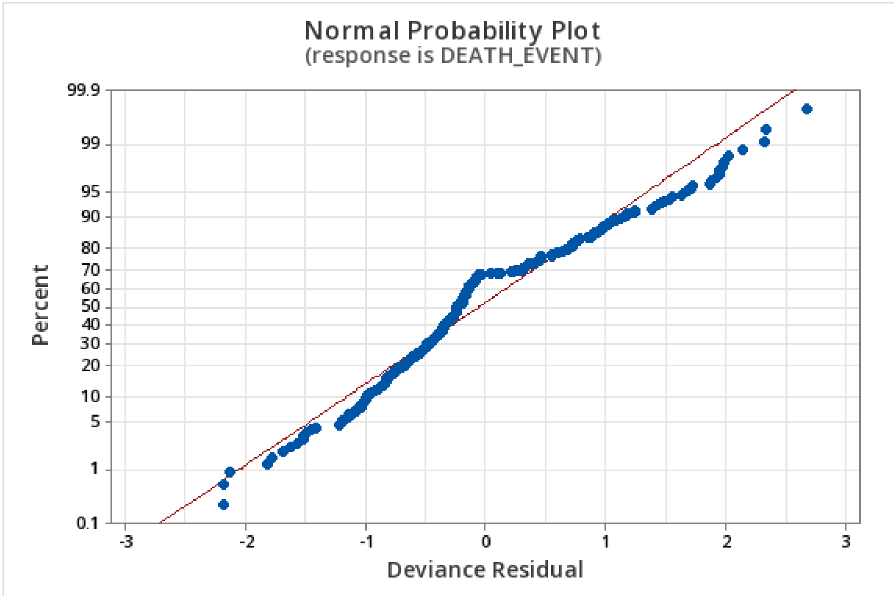
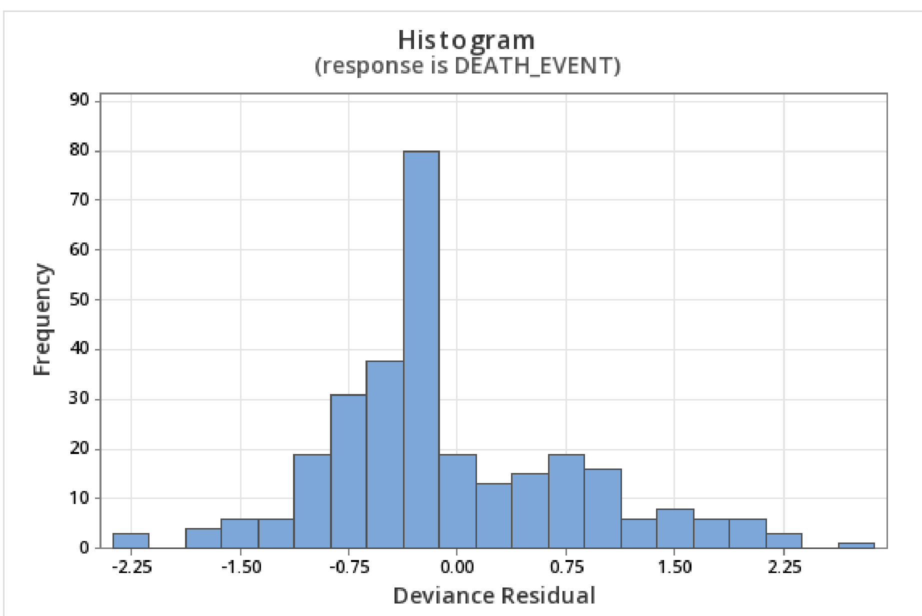


Figure 13

Histogram



Appendix C: Machine Learning Model Performance Metrics

- **Confusion Matrices:**

Figure 14

Confusion Matrix I

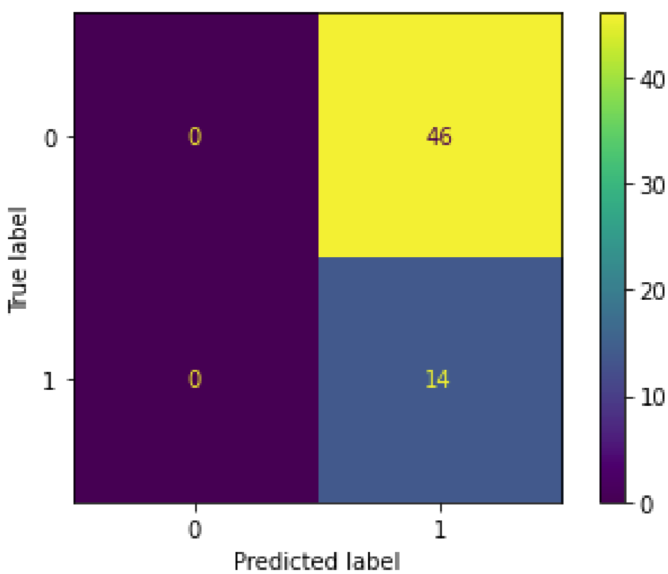


Figure 15

Confusion Matrix II

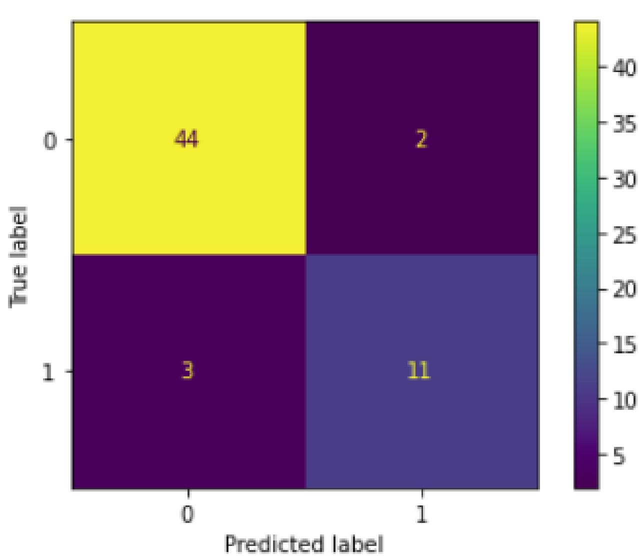


Figure 16

Confusion Matrix III

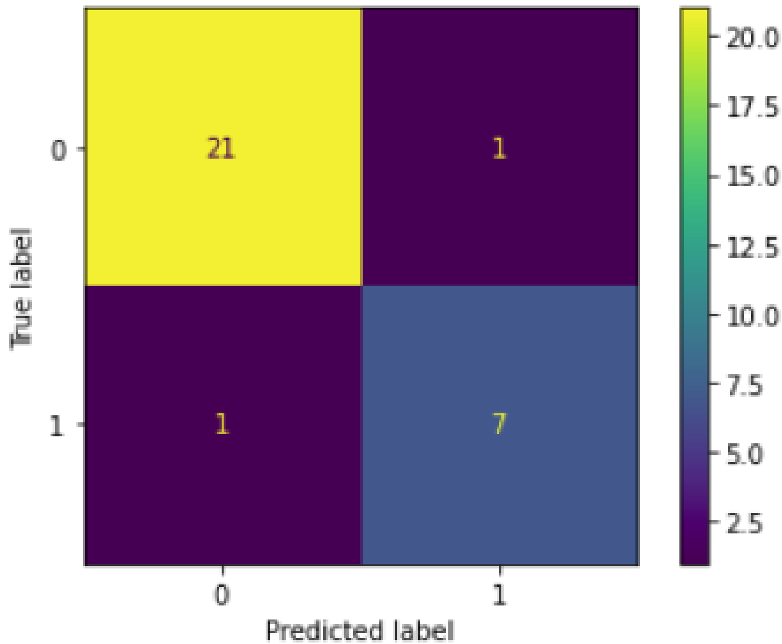
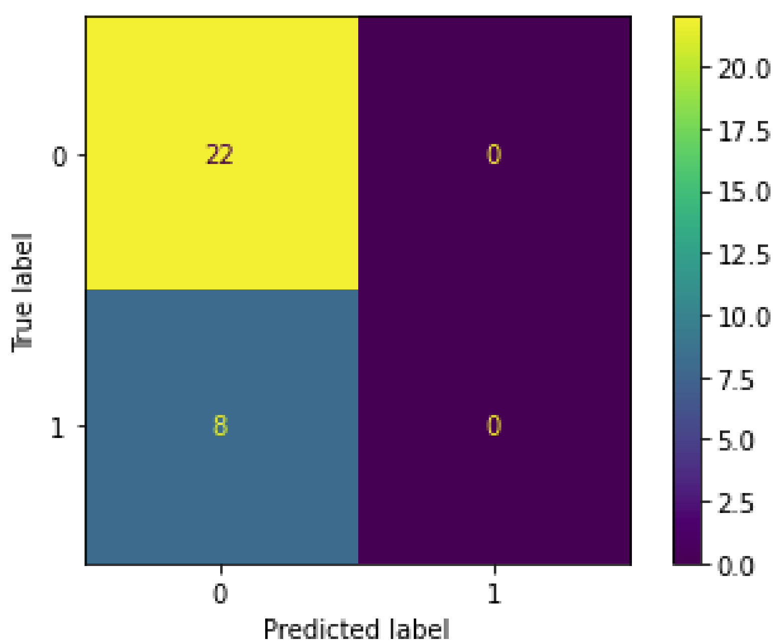


Figure 17

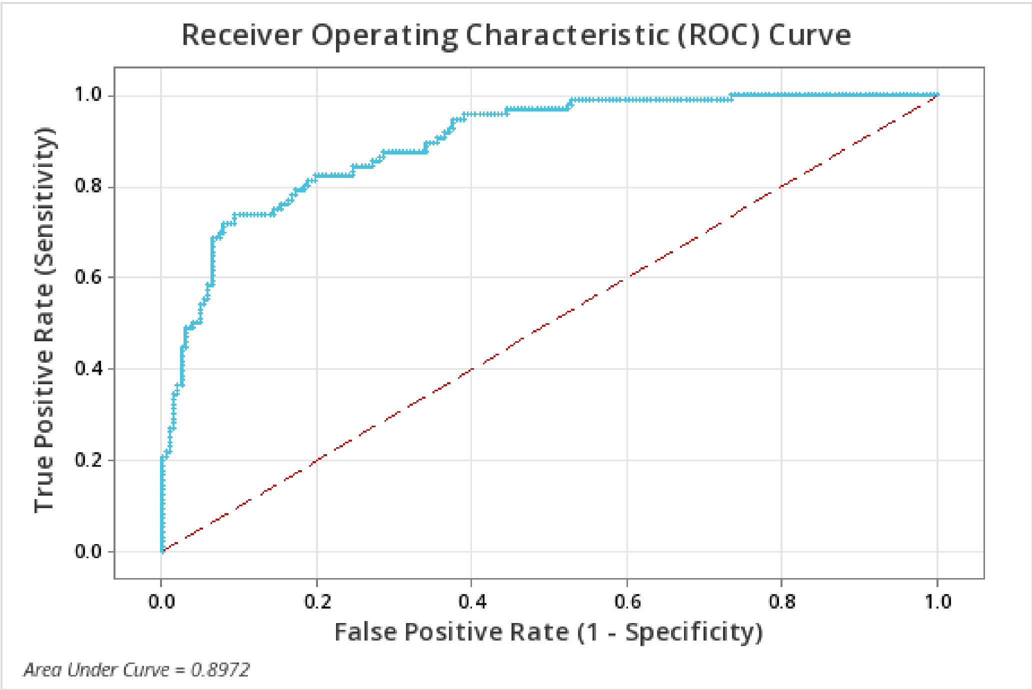
Confusion Matrix IV



- **ROC Curves:** Receiver Operating Characteristic curves to evaluate model performance.

Figure 18

ROC Curve



Appendix D: Code Snippets

- **Python Code:**

Figure 19

Python Code Snippets

Avail Lean Manufacturing to Abate Heart Failure deaths with Machine Learning prognostication

Jupyter Notebook version 7.2.2

[1]:
```python
import pandas as pd
from sklearn import preprocessing
from sklearn.model_selection import train_test_split
from sklearn.preprocessing import StandardScaler
import csv
```

The original data is being stored in a variable called Original_Data.

[3]:
```python
Original_Data = pd.read_csv("heart_failure_clinical_records_dataset.csv")
```

The first five observations are displayed below.

[6]:
```python
Original_Data.head()
```

[6]:

	age	anaemia	creatinine_phosphokinase	diabetes	ejection_fraction	high_blood_pressure	platelets	serum_creatinine	serum_sodium	sex	smoking	time	DEATH_EVENT
0	75.0	0	582	0	20	1	265000.00	1.9	130	1	0	4	1
1	55.0	0	7861	0	38	0	263358.03	1.1	136	1	0	6	1
2	65.0	0	146	0	20	0	162000.00	1.3	129	1	1	7	1
3	50.0	1	111	0	20	0	210000.00	1.9	137	1	0	7	1
4	65.0	1	160	1	20	0	327000.00	2.7	116	0	0	8	1

This data has 299 rows and 13 columns.

[9]:
```python
Original_Data.shape
```

[9]: (299, 13)

Some of the scenarios will use all 12 factors without any scaling. This is represented with the variable X_Predictors_12_NoScaling.

[12]:
```python
X_Predictors_12_NoScaling = Original_Data.loc[:, Original_Data.columns != "DEATH_EVENT"]
X_Predictors_12_NoScaling.head()
```

[12]:

	age	anaemia	creatinine_phosphokinase	diabetes	ejection_fraction	high_blood_pressure	platelets	serum_creatinine	serum_sodium	sex	smoking	time
0	75.0	0	582	0	20	1	265000.00	1.9	130	1	0	4
1	55.0	0	7861	0	38	0	263358.03	1.1	136	1	0	6
2	65.0	0	146	0	20	0	162000.00	1.3	129	1	1	7
3	50.0	1	111	0	20	0	210000.00	1.9	137	1	0	7
4	65.0	1	160	1	20	0	327000.00	2.7	116	0	0	8

The variable X_Predictors_12_NoScaling only has 12 columns because this variable only represents the predictor variables and not the response variable.

Appendix E: Minitab Statistical Analysis

- **Minitab Statistical Analysis:** Output from Minitab statistical analysis, including correlation matrices, regression analysis, and hypothesis testing results.

Table 3

Model Summary

Deviance R-Sq	Deviance R-Sq(adj)	AIC	AICc	BIC	Area Under ROC Curve
41.51%	38.31%	245.55	246.83	293.66	0.8972

Table 4

Goodness-of-Fit Tests

Test	DF	Chi-Square	P-Value
Deviance	286	219.55	0.999
Pearson	286	258.78	0.875
Hosmer-Lemeshow	8	10.22	0.250

Table 5

Analysis of Variance

	Wald Test

Source	DF	Chi-Square	P-Value
Regression	**12**	**72.41**	**0.000**
Time	1	48.74	0.000
Serum_sodium	1	2.84	0.092
Serum_creatinine	1	13.47	0.000
Ejection_fraction	1	22.04	0.000
Age	1	9.01	0.003
Platelets	1	0.40	0.525
Creatinine_Phosphokinase	1	1.56	0.212
Smoking	1	0.00	0.974
Sex	1	1.66	0.197
High_blood_pressure	1	0.08	0.775
Diabetes	1	0.17	0.679
Anaemia	1	0.00	0.983

Table 6

Observed and Expected Frequencies for Hosmer-Lemeshow Test

			DEATH_EVENT = 1		DEATH_EVENT = 0	
Group	**Event Probability Range**		**Observed**	**Expected**	**Observed**	**Expected**
1	(0.000,	0.013)	0	0.2	29	28.8
2	(0.013,	0.030)	1	0.6	29	29.4
3	(0.030,	0.061)	0	1.3	30	28.7
4	(0.061,	0.113)	3	2.4	27	27.6
5	(0.113,	0.191)	8	4.4	22	25.6
6	(0.191,	0.302)	5	7.4	25	22.6
7	(0.302,	0.454)	8	11.1	22	18.9
8	(0.454,	0.678)	21	17.3	9	12.7
9	(0.678,	0.859)	23	23.1	7	6.9
10	(0.859,	0.999)	27	28.1	3	1.9