DOKUZ EYLÜL UNIVERSITY
GRADUATE SCHOOL OF SOCIAL SCIENCES
DEPARTMENT OF BUSINESS ADMINISTRATION
BUSINESS INFORMATION SYSTEMS PROGRAM
MASTER'S THESIS

PREDICTION OF SEPSIS IN INTENSIVE CARE UNIT PATIENTS USING MACHINE LEARNING

Bita AZARI

Supervisor
Prof. Dr. Sabri ERDEM

2024

DOKUZ EYLÜL UNIVERSITY
GRADUATE SCHOOL OF SOCIAL SCIENCES
DEPARTMENT OF BUSINESS ADMINISTRATION
BUSINESS INFORMATION SYSTEMS PROGRAM
MASTER'S THESIS

PREDICTION OF SEPSIS IN INTENSIVE CARE UNIT PATIENTS USING MACHINE LEARNING

Bita AZARI

Supervisor
Prof. Dr. Sabri ERDEM

İZMİR- 2024

THESIS APPROVAL PAGE

DECLARATION

I hereby declare that this master's thesis or non-thesis master's term project titled "Prediction of Sepsis in Intensive Care Unit Patients Using Machine Learning" has been written by myself in compliance with academic guidelines and ethical standards. I further declare that all the materials utilized in this thesis are from the sources listed in the reference list. All of this is verified by my honor.

19/01/2024

Bita AZARI

ABSTRACT

Master's Thesis

Prediction of Sepsis in Intensive Care Unit Patients Using Machine Learning

Bita AZARI

Dokuz Eylül University

Graduate School of Social Sciences

Department of Business Administration

Business Information Systems Program

Sepsis is one of the leading causes of death in intensive care units (ICUs) around the world and detecting sepsis and treating the disease can be particularly challenging for doctors, especially in its early stages. The annual expenditures for hospital treatment for sepsis patients are increasing due to public health issues. According to current literature, early and appropriate antibiotic medication is a major predictor of sepsis outcome, and early detection of sepsis may enable this treatment. As a result, without antibiotic therapy for sepsis, the risk of death increases with each consecutive hour. Finding out if a person has sepsis early and giving them the right treatment quickly can help ensure they have a better chance of getting better. However, relying on traditional rule-based screening methods to identify problems is not always reliable. Since AI-based diagnostic systems are proving useful in many medical fields, machine learning can be used to possibly analyze large patient datasets to uncover previously unknown patterns and trends among features by offering a variety of techniques. Using the Dokuz Eylul Hospital critical care patient dataset, this work aims to develop and verify a machine learning algorithm (MLA) that can predict severe sepsis. Machine learning offers a variety of ways to analyze huge patient datasets, potentially uncovering previously unknown patterns and trends across features. This study used experimental analyses of a random forest, gradient-boosted classifier, k-neighbor's classifier, and neural network to identify patient outcomes.

Keywords: Machine Learning, Sepsis, Early Prediction, Algorithm.

ÖZET

Yüksek Lisans Tezi

Yoğun Bakım Hastalarında Sepsisin Makine Öğrenmesiyle Tahmini

Bita Azari

Dokuz Eylül Üniversitesi

Sosyal Bilimler Enstitüsü

İngilizce İşletme Anabilim Dalı

İngilizce İşletme Bilişim Sistemleri Programı

Sepsis, dünya çapında yoğun bakım ünitelerinde (YBÜ) önde gelen bir ölüm nedenidir ve özellikle hastalığın erken evrelerinde tanınması tıbbi bir zorluk olmaya devam etmektedir. Halk sağlığı sorununa göre, sepsis hastalarının yıllık hastane tedavi maliyetleri artmaktadır. Son literatür, erken ve uygun antibiyotik tedavisinin sepsis sonuçlarını öngören ana faktör olduğunu ve bu tedavinin sepsis erken teşhis edilebilmesi mümkün olduğunu göstermektedir. Bu nedenle, sepsis için antibiyotik tedavisi olmaksızın, ölüm riski her ek saatle birlikte artmaktadır. Doğru erken tanı ve tedavi, olumsuz hasta sonuçları riskini azaltabilir, ancak geleneksel kural tabanlı tarama yöntemlerinin etkinliği sınırlıdır. Yapay zekaya (AI) dayalı bir tanı sisteminin birçok tıbbi alanda etkili olduğu gösterildiğinden; makine öğreniminin, büyük hasta veri kümelerini analiz etmek için bir dizi yaklaşım sağladığı ve potansiyel olarak olmayan özellikler arasında kalıplar ve eğilimler bulduğu açıktır. Bu çalışmanın amacı, Dokuz Eylül Hastanesi yoğun bakım hasta veri kümesini kullanarak sepsis tahmini için bir makine öğrenimi algoritması (MLA) geliştirmek ve doğrulamaktır. Makine öğrenimi, büyük hasta veri kümelerini analiz etmek için bir dizi yaklaşım sunar ve potansiyel olarak özellikler arasında olmayabilecek kalıpları ve eğilimleri bulur. Bu çalışma, klinisyenlere açık hasta sonuçlarını sınıflandırma performanslarını test etmek için rastgele orman, gradyan destekli sınıflandırıcı, k-komşu sınıflandırıcı ve sinir ağının deneysel bir analizini tamamlamıştır.

Anahtar Kelimeler: Makine öğrenmesi, Sepsis, erken tahmin, algoritma.

PREDICTION OF SEPSIS IN INTENSIVE CARE UNIT PATIENTS USING MACHINE LEARNING

CONTENTS

THESIS APPROVAL PAGE	ii
DECLARATION	iii
ABSTRACT	iv
ÖZET	vi
CONTENTS	viii
ABBREVIATIONS	xi
LIST OF FIGURES	xiv
LIST OF TABLES	xv
LIST OF APPENDICES	xvi
INTRODUCTION	1
CHAPTER ONE	
LITERATURE REVIEW	
1.1. SEPSIS-1	5
1.2. SEPSIS-2	6
1.3. SEPSIS-3	6
1.4. MEDICAL LITERATURE FOR SEPSIS WITH AND WITHOUT TECHNOLOGY	9

CHAPTER TWO

MACHINE LEARNING, NEURAL NETWORK, AND DEEP LEARNING

2.1. MACHINE LEARNING 11

2.1.1. Overview 11

2.1.2 Supervised Machine Learning 12

2.1.3. Unsupervised Machine Learning 13

2.1.4. Semi-Supervised Machine Learning 14

2.1.5. Experts' Systems and Knowledge Base 15

2.2. NEURAL NETWORK AND DEEP LEARNING 16

2.2.1. Data Science 17

2.2.2 Data Analyst 18

2.2.3. Data Science Tools 18

2.2.4. Python 19

2.2.5. Knime 19

2.2.6 Predict Sepsis Using Knime and Python 20

2.2.7. Decision Tree 20

2.2.8. Linear Regression 20

2.2.9. Logistic Regression 21

2.2.10. TensorFlow and Keras 22

2.2.11. The Synergy of Keras and TensorFlow 23

2.2.12. Predicted Sepsis using Knime 24

2.3. MODEL EVALUATION'S METRICS 24

2.3.1 Confusion Matrix 25

2.3.2 Accuracy 25

2.3.3. Error Rate 26

2.3.4. Precision 26

2.3.5. Recall 26

2.3.6. F1 Score 27

2.3.7. AUC 27

2.3.8 Sensitivity 27

CHAPTER THREE

METHODOLOGY, MODELING EVALUATION, AND DATA ANALYSIS OF SEPSIS

3.1. APPLICATION 28

3.2. INTRODUCING DOKUZ EYLÜL HOSPITAL 29

3.3 ICU 30

3.5. PREDICTION 32

3.5.1. Prediction Using Decision Tree 32

3.5.2. Prediction Using Logistic Regression 33

3.6. MODEL 45

3.7. PREDICTION USING MACHINE LEARNING ALGORITHMS USING PYTHON 48

3.8. MODEL ARCHITECTURE 48

3.9. FINDINGS 49

3.9.1 Decision Tree Result's 49

3.9.2. Logistic Regression Result's 50

3.9.3. Deep Learning Result's 50

3.9.1. Python Findings 51

CONCLUSION 55

REFERENCES 60

APPENDICES

ABBREVIATIONS

ICU	Intensive Care Unit
AI	Artificial Intelligence
SVM	Support Vector Machine
SOFA	STATUS OF FORCES AGREEMENT
SIRS	Systemic inflammatory response syndrome
SCCM	Society of Critical Care Medicine
ACC	Agenesis of the corpus callosum
SSL	Secure Sockets Layer
ML	Machine Learning
CT	Computed Tomography
Light GBM	Light Gradient-Boosting Machine
MIMIC	Medical Information Mart for Intensive Care
ROC	Receiver operating characteristic
AUC	Area under Curve
RNN	Recurrent neural networks
CCR	Correct Classification Rate
TP, FP, TN, FN	True Positive, False Positive, True Negative, False Negative
TPU	Tensor Processing Units
SSC	Semi-Supervised Classification
MRI	Magnetic Resonance Imaging
WBC	White Blood Cells
HGB	Hemoglobin
RBC	Red Blood Cells
BASO#	Total number of basophils

BASO%	Percentage of basophils
EOS#	Total number of eosinophils
EOS%	Percentage of eosinophils
HCT	Hematocrit
LYM#	Total number of lymphocytes
LYM%	Percentage of lymphocytes
MCH	Mean Corpuscular Hemoglobin
MCHC	Mean Corpuscular Hemoglobin Concentration
MCV	Mean Corpuscular Volume
MONO#	Total number of monocytes
MONO%	Percentage of monocytes
MPV	Mean Platelet Volume
NEU#	Total number of neutrophils
NEU%	Percentage of neutrophils
PCT	Procalcitonin
PLT	Platelet
RDW	Red Cell Distribution Width
SD_C_EGC	Standard Deviations of Conductivity of Early Granulated Cells
SD_C_LY	Standard Deviations of Conductivity of Lymphocytes
SD_C_MO	Standard Deviations of Conductivity of Monocytes
SD_C_NE	Standard Deviations of Conductivity of Neutrophils
SD_MALS_EGC	Standard Deviations of Scattering of Early Granulated Cells
SD_MALS_LY	Standard Deviations of Scattering of Lymphocytes
SD_MALS_MO	Standard Deviations of Scattering of Monocytes
SD_MALS_NE	Standard Deviations of Scattering of Neutrophils
SD_V_EGC	Standard Deviations of Volume of Early Granulated Cells

SD_V_LY	Standard Deviations of Volume of Lymphocytes
SD_V_MO	Standard Deviations of Volume of Monocytes
SD_V_NE	Standard Deviations of Volume of Neutrophils
MDW	Monocyte Volume Distribution Width
MN_C_EGC	Mean Conductivity of Early Granulated Cells
MN_C_LY	Conductivity of Lymphocytes
MN_C_MO	Conductivity of Monocytes
MN_C_NE	Conductivity of Neutrophils
MN_MALS_EGC	Mean Scattering of Early Granulated Cells
MN_MALS_LY	Mean Scattering of Lymphocytes
MN_MALS_MO	Mean Scattering of Monocytes
MN_MALS_NE	Mean Scattering of Neutrophils
MN_V_EGC	Mean Volume of Early Granulated Cells
MN_V_LY	Mean Volume of Lymphocytes
MN_V_MO	Mean Volume of Monocytes
MN_V_NE	Mean Volume of Neutrophils

LIST OF FIGURES

Figure 1: Supervised Learning Illustration p. 12
Figure 2: Python Libraries for Machine Learning and Deep Learning p. 29
Figure 4: Decision Tree Analysis Results p. 34
Figure 5: Decision Tree Analysis Results p. 35
Figure 6: Decision Tree Analysis Results p. 36
Figure 7: Decision Tree Analysis Results p. 37
Figure 8: Decision Tree Analysis Results p. 38
Figure 9: Decision Tree Analysis Results p. 39
Figure 10: Decision Tree Analysis Results p. 40
Figure 11: Decision Tree Analysis Results p. 41
Figure 12: Decision Tree Analysis Results p. 42
Figure 13: Decision Tree Analysis Results p. 43
Figure 14: Decision Tree Analysis Results p. 44
Figure 15: Confusion Matrix Outcome of Sepsis Dataset p. 47

LIST OF TABLES

Table 1: Evolution of Sepsis Definition p. 7
Table 2: Confusion Matrix for Binary Classification p. 25
Table 3: List of Variables in The Study p. 31
Table 4: Partitioned Data Sets p. 32
Table 5: Logistic regression with Knime result's p. 50
Table 6: Model's Evaluation Result p. 52

LIST OF APPENDICES

Appendix 1: Quantitative Description of Dataset with Knime app p. 1
Appendix 2: Parameter References app p. 3
Appendix 3: Ethical Committee Approval for Data Usage app p. 4

INTRODUCTION

Sepsis is one of several complicated disorders where the body responds abnormally to an infection. Sepsis is a severe health concern that can result in-patient fatalities and financial burdens for hospitals (Kang et al., 2018). Sepsis and the body's inflammation reaction can cause multiple organ problems and can result in death. It has been estimated that about one-third of patients receiving treatment in an ICU have sepsis. Even though researchers have studied it for many years, sepsis is still a big reason why people die or get sick in intensive care units (ICUs) globally (Sartelli et al., 2018). Additionally, it poses a significant challenge and burden on healthcare systems. Early recognition of sepsis is crucial, as previous research shows there is an increased mortality rate for every hour that treatment is delayed (Evans et al., 2021; Rachoin & Dellinger, 2014). Early initiation of optimal sepsis care can enhance the chances of survival during hospitalization. Once doctors can identify the early signs of sepsis, they can promptly initiate treatment. A big problem is that many people with sepsis are not diagnosed when they first come to the hospital, so they cannot get treatment right away. Starting treatment early and using inexpensive intervention with antibiotics, giving fluids to help the body recover, source control, and supporting how well the important organs are working can make a big difference in how well the patient gets better. Therefore, it is crucial to predict and diagnose sepsis at an early stage and administer medication to patients promptly. Unfortunately, it is difficult to detect sepsis early due to the variation in symptoms and patient heterogeneity. Besides, current technologies and clinical systems cannot recognize sepsis in its early stages. Furthermore, there is no specialized anti-sepsis treatment; instead, infection control strategies such as source control, antibiotic administration, and supporting organ function are used.

Currently, doctors determine the presence of sepsis by evaluating information obtained from examinations, monitoring devices, and laboratory test findings (i.e., based on empirical clinical decision rules). Additionally, there is no one way to detect sepsis using a single indicator, and laboratory testing is often too slow or not very accurate. So, this process takes a lot of time and relies a lot on the doctor's or nurse's skills and experience. The problem of diagnosing the medical condition has made it

hard for doctors to correctly diagnose it, causing some patients to be over diagnosed and others to be underdiagnosed.

Generally, the hospital will keep track of patients' health regularly after they go to the intensive care unit (ICU). They begin to monitor and manage the information about patients to help detect sepsis as soon as they can. Many academics are focusing on leveraging this data to improve sepsis prediction and get good results using machine learning approaches to tackle the challenge of detecting early indicators of sepsis. Thus, the main objective of this work was to establish a machine learning system for early detection of sepsis and reduce mortality. Both the clinician and the patient are interested in the model's predictions. For the physician, this allows the model to become a clinical tool, aiding in diagnosis. This increases the patient's trust in the decision being made for them. Artificial intelligence (AI), machine learning, deep learning, and semantic computing are a few ideas that have started being used in the medical industry recently. They bring new technology and solutions that are changing healthcare. Artificial intelligence (AI) is a growing field in medicine that has been used to create different innovative clinical decision support systems that assist doctors. It has shown considerable potential for accurately predicting how patients are doing and helping doctors make decisions about their treatment. Therefore, pharmaceutical has seen the development of machine learning as a novel tool to analyze expansive sums of data. After 2012, AI began to explode because of increases in data volume and computing power, as well as advancements in modern machine learning algorithms, resulting in advances in systems for expert machine learning, developing computing, computer vision, natural dialect processing, and other information handling technologies. (Guan et al., 2021). The primary core technique of machine learning is to utilize algorithms to evaluate and learn from data and then to use the outcomes of that learning to generate decisions and foresee approximate events in reality. Machine learning has appeared in the early phases of AI development. Decision trees, Support Vector Machines (SVM), clustering, and other basic methods are included. Machine learning may be categorized based on different learning methodologies. Initially, the algorithms used are supervised learning, unsupervised learning, and semi-supervised learning. Integrated learning, deep learning, and reinforcement learning are examples of recently developed algorithms. Using traditional machine learning algorithms to

manage sepsis has produced some promising outcomes; however, making progress in this area has been incredibly challenging. However, the emergence of deep learning has brought a change. Deep learning employs mathematical models that are supposed to function similarly to the human brain. With the help of computer networks and innovative technology, we can enable extraordinary computational capabilities and uncover significant data from massive volumes of information that may have otherwise been disregarded or gone missing. In the healthcare sector, deep learning networks have the capability to tackle complex issues and extract crucial insights from vast amounts of data. It is a skillset that has not gone unnoticed by the healthcare profession. Machine learning sets itself apart from traditional hard-coded software programs as it acquires knowledge through the examination of vast amounts of data and the application of diverse techniques to comprehend and perform specific tasks. The most popular type of AI algorithm used for various aspects of sepsis is mechanical learning. It can help with early prediction, assessing how the condition will progress, predicting mortality, and finding the best ways to manage sepsis. Machine Learning (ML) provides techniques to analyze large and complex patient datasets. Machine learning interpretation Studies that predict sepsis can be challenging since some of them anticipate the incidence of sepsis right from the start, which may seem unexpected and not very useful in practical terms. Their aim is to determine if a patient meets the criteria for sepsis, which includes looking at indicators like whether they took antibiotics or had samples taken for culture testing to check for infection. During the creation of the model, it can access these proxies. However, when testing the model with a set of data or a new patient, these proxies are not known. A special program has been made to guess if a person has sepsis by looking at different factors. In clinical practice, identifying sepsis may take longer than needed, but finding it early can help with diagnosing and treating it faster. Accordingly, although it is complicated and, in many cases, impossible to diagnose this disease in the early hours, technology and machine learning can be useful in predicting sepsis some hours before it happens and reducing the risk of mortality. Usually, the hospital will keep track of how the patients are doing every hour once they are admitted to the intensive care unit (ICU). Researchers are using this information to make better predictions about sepsis and get good results by using machine learning approaches (Davoudi et al., 2019). They

developed predictive models that predict whether a patient would develop sepsis; Clinical implementation studies, however, are required to understand how these models might be integrated into clinical procedures and adapted to previously unknown data.

The historical Multiparameter Intelligent Monitoring in Intensive Care (MIMIC)-III dataset has been the main resource for predicting sepsis in previous studies. After receiving MIMIC-III and data preprocessing, working with real and primary datasets has been decided. Therefore, ethical permission from Dokuz Eylul University Medical Faculty has been obtained with the help and support of the medical team, and the sepsis patient's dataset has been collected and delivered for this research. Data is collected from 2021 to 2022. Two datasets have been combined for this study. The first one is from 2021, with 556 records and 53 variables. The second one is from 2022, with 785 records and 103 variables. Then, datasets were combined, and some data was eliminated to yield 8 variables and 488 records.

CHAPTER ONE
LITERATURE REVIEW

The literature review of this thesis covers topics and sub-topics related to the scope of the thesis. In this section, medical literature for sepsis without and with technology, variable detection, data analytical techniques, and deep learning are discussed. Compared with previous studies, the Theoretical Framework is explained. Sepsis arises as a hazardous medical situation when the body's immune response fails to function appropriately in fighting an infection (Bone et al., 1992). Another definition of Sepsis is a serious and life-threatening condition that happens when the body's response to an infection harms its own tissues and organs (Bone et al., 1992). Multiple definitions and terminologies have improved throughout time because of the limitations of prior definitions and our understanding of the complicated pathophysiology of sepsis.

1.1. SEPSIS-1

The term sepsis comes from an ancient Greek word "decay" or "decomposition". Around 2700 years ago, it was first referenced in Homer's poems. Hugo Schottmüller, who in 1914 defined the first modern definition, wrote that "sepsis is present if a focus has developed from which pathogenic bacteria, constantly or periodically, invade the bloodstream in such a way that these cause subjective and objective symptoms." The response of the body's immune system plays a crucial role in the development of sepsis symptoms, as indicated by numerous tests and studies. However, the disease’s heterogeneity offered significant challenges in detecting, treating, and researching sepsis. Finally, after two decades, scientists have better understood how cells work, the chemistry of living organisms, how the immune system functions, the study of body structure, and the changes that occur in blood flow and organ function. At an SCCM-ACCP meeting in the year 1991, Roger Bone and his colleagues began the first formal description of sepsis. Sepsis-1 is a kind of SIRS (systemic inflammatory response syndrome). The systemic inflammatory response to a variety of severe clinical insults includes:

The temperature is either above 38°C or below 36°C. The heart rate is above 90 beats per minute. The respiratory rate is above 20 breaths per minute. The PaCO2 level is below 32 mmHg. The white blood cell count is either above 12,000/cu mm, below 4000/cu mm, or above 10% immature (band) forms. Sepsis-1 is a systemic reaction to infection characterized by two or more of the SIRS criteria (Gyawali et al., 2019).

1.2. SEPSIS-2

In 2001, a second consensus attempted to reconsider the SIRS criteria but could not create a simple list of factors to define sepsis; more parameters were introduced, delegates risked making the definition less precise by including terms like inflammatory parameters, hemodynamic variables, and tissue perfusion parameters. Sepsis occurs when an individual is diagnosed with or suspected of having an infection and changes in the mentioned parameters range in Table 1 (Gyawali et al., 2019).

1.3. SEPSIS-3

The Society of Critical Care Medicine and the European Society of Intensive Care Medicine and brought together a team of 19 specialists in January 2014 to review the current definitions. These specialists were experts in critical care, infectious diseases, surgery, and pulmonary medicine. Unrestricted funding support was provided by the societies, allowing the task force to have complete autonomy over their actions. Finally, in 2016, the third international consensus developed a new definition of sepsis as life-threatening organ failure caused by an abnormal host reaction to infection (Table 1) (Gyawali et al., 2019). This new definition focuses on how the body responds to infection, which may be more dangerous than the infection itself. It is important to recognize this quickly.

Table 1: Evolution of Sepsis Definition

Sepsis (1991)	**Sepsis 2(2001)**	**Sepsis 3(2016)**
SIRS (systemic inflammatory response syndrome): a systemic inflammatory response to several severe clinical insults: Temperature >38°C or <36°C; A heart rate of more than 90 beats per minute; a respiratory rate of more than 20 breaths per minute. Sepsis is when your body reacts to an infection and you have two or more symptoms of SIRS because of the infection (Gyawali et al., 2019).	Fever, which has a high body temperature (above 38. °C), Hypothermia means having a low body temperature (below 36°C). If the heart beats more than 90 times per minute or is much faster than what is normal for someone's age, it is considered abnormal. Tachypnea is when you breathe fast—more than 30 breaths per minute. It can be accompanied by a change in mental state, swelling, or having too much fluid in the body (more than 20 mL per kg of body weight in 24 hours).	Sepsis clinical criteria: Suspected or verified infection and a 2 SOFA point rise. The task force thought that if someone has positive qSOFA criteria, it means they might have an infection even if it was not known before. qSOFA criteria: disturbed mental state (GCS score 15); systolic blood pressure more than 100 mmHg; respiratory rate greater than 22 breaths per minute. Septic shock is a kind of sepsis in which there are major difficulties with the circulatory system and cells of the body. These issues can significantly raise the risk of mortality.
	Inflammatory parameters: Leukocytosis (white blood cell count > 12,000/L); leukopenia (white blood cell count 4000/L); normal white blood cell count with greater than 10% immature forms; plasma C-reactive protein greater than 2 SD above the normal value; and plasma procalcitonin > 2 SD above the normal value are all signs of leukocytosis.	Septic shock is a condition where a person has an infection that causes their blood pressure to stay low, even with medication (Gyawali et al., 2019).
	Low blood pressure, specifically when the top number is less than 90 or there	

	is a decrease greater than 40 in adults, or when it is more than 2 standard deviations below the normal range for age. The oxygen saturation in the blood is greater than 70%, and the heart is pumping enough blood with a cardiac index greater than 3. 5 liters per minute per square meter. Organ dysfunction parameters: Acute oliguria (urine output 0.5 mL kg1 h1 or 45 mM L1 for at least 2 h); creatinine rise 0.5 mg dL1; coagulation problems (international normalized ratio > 1.5 or activated partial thromboplastin time > 60 s); ileus (absent bowel sounds); thrombocytopenia (platelet count less than 100,000 L1) Hyperbilirubinemia (plasma total bilirubin more than 4 mg dL1 or 70 mmol L1) (Gyawali et al., 2019).	
GCS: Glasgow Coma Scale; FIO2: percentage of inspired oxygen. SOFA stands for Sequential Organ Failure Assessment.		

Source. ("American College of Chest Physicians/Society of Critical Care Medicine Consensus Conference: Definitions for Sepsis and Organ Failure and Guidelines for the Use of Innovative Therapies in Sepsis.," 1992b)

One important change in the new definitions is that they no longer include any information about SIRS. The Septic-3 Task Force has developed a new bedside index to assist in the identification of patients who are not in critical care but may have infections that can progress to sepsis (Singer et al., 2016). They called this bedside index the qSOFA. The sepsis mortality rate is very high, with a global estimated incidence of 148 per 100,000 people per year and a mortality rate of 26%. It is a big

problem for everyone's health, costing US hospitals a lot of money to treat it. It also causes many people to die. The symptoms of this condition include a high body temperature, a fast heart rate, difficulty breathing, and an increase in certain types of white blood cells outside of the center of the body. Many studies show that the number of sepsis cases is increasing, even though fewer people are dying. It is crucial to identify sepsis in its early stages and give antibiotics immediately to decrease the occurrence of deaths and the elevated cost of it. The Surviving Sepsis Guidelines say that people with sepsis should start taking medicine to fight infection within 1 hour. This is because waiting too long to start treatment can increase the chance of dying. However, it can be hard to recognize sepsis early because the disease is complicated (Kumar et al., 2006) and the heterogeneity of the septic population (De Grooth et al., 2018). Today, medical care is still struggling to diagnose it due to a lack of specific diagnostic and rapid tests.

1.4. MEDICAL LITERATURE FOR SEPSIS WITH AND WITHOUT TECHNOLOGY

In simpler words, sepsis is a medical condition that can have different causes and symptoms. It can be managed, and the chances of recovery are different for each person. Sepsis patients often differ from the triggering organism and concentrate on infection. Furthermore, their approaches to antibiotic selection and infection management varied, as do their approaches to surgical source control. A definite microbiological diagnosis of bacteria or viruses causing infection cannot be made in at least one-third of patients who show signs of sepsis (Sands, 1997; Vincent et al., 2006). Moreover, even when patients with similar types of bacteria or body part infections are compared, the severity of symptoms and the chance of getting better can differ from person to person and from one geographic region to another (Calvano et al., 2005).

Sepsis occurs when the immune system is activated and causes changes in many substances in the body that are responsible for inflammation, blood clotting, and metabolism (Marshall, 2003).

Sepsis has been associated with over 100 different molecules that could potentially be used as helpful signs (Sutton & Barto, 2018). Because it is unclear which of them gives true, valuable information or how the utility is determined to identify sepsis.

CHAPTER TWO
MACHINE LEARNING, NEURAL NETWORK, AND DEEP LEARNING

In this section, the timeline of Machine Learning, AI in the 1940s, Experts' Systems Knowledge Base (Knowledge Management, Neural Networks, and Deep Learning) ML, data mining, data science, and data analytics, as well as the features and benefits of Knime, which is used as an analytical tool in this study, are explained.

2.1. MACHINE LEARNING

Machine learning is a computer science discipline that evolved from studying pattern recognition and data learning in artificial intelligence (Hosch, 2023). Machine learning is an information technology discipline that evolved from studying pattern recognition and data learning in artificial intelligence. These algorithms do not simply follow static program instructions but rather acquire knowledge from example inputs and utilize it to make informed guesses or decisions based on facts. Computational statistics and mathematical optimization, which specialize in prediction and delivery techniques, theory, and application fields, are closely linked **(**Richard Sutton R. S. & Barto A. G.1998). For over a decade, the estimation industry has been plagued with unimpressive statistical estimate model outcomes. Alternative methodologies such as machine learning are rapidly being used in multiple research due to their inability to handle categorical data, deal with missing data points, handle how data is distributed out, and, most importantly, their lack of thinking and reasoning capabilities.

2.1.1. Overview

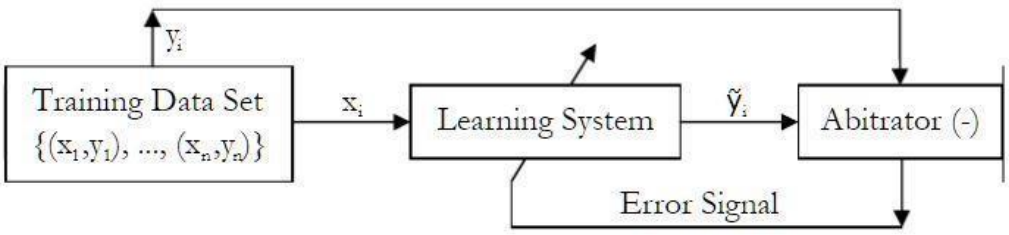
Arthur Samuel described the field of machine learning as a "Field of study that gives computers the ability to learn without being explicitly programmed" in 1959 (Simon, 2015). Machine learning was created through efforts to develop artificial intelligence. Even back when AI was just starting as a field of study, some scientists wanted machines to be able to learn by using data. They experimented with various symbols called "neural networks," which were mostly perceptions and other models

that were later found to be like generalized linear models in statistics. Probabilistic reasoning was also employed, specifically in automated medical diagnosis. Machine learning, which emerged as a distinct field, started to flourish during the 1990s. The industry shifted its focus from trying to create artificial intelligence to solving practical problems that can be solved easily. It changed its attention from using symbolic techniques from AI to using methods and models from statistics and probability theory. Machine learning tasks are usually divided into three categories based on their characteristics: supervised machine learning, unsupervised machine learning, and semi-supervised machine learning.

2.1.2 Supervised Machine Learning

This is a machine learning paradigm to learn how inputs and outputs are related in a system using a set of example pairs as training samples. It includes a model that can forecast using a labeled dataset. The aim of supervised learning is to create an artificial system that can figure out the relationship between input and output and then anticipate the output based on labeled data.

Figure 1: Supervised Learning Illustration



Source: Kotsiantis, 2007: 249–268.

Figure 1 demonstrates supervised learning, where 'x' represents system input, 'y' represents system output (i.e., the supervision or labeling of the input x), and 'i' is the training sample index. During a Supervised Learning process, a training input xi is fed to the Learning System, and the Learning System generates an output $\tilde{y}_i$. The degree of discrepancy between the Learning System output $\tilde{y}_i$ and the ground truth

labeling y_i is assessed by an arbitrator (Seel, 2011). Supervised learning is divided into two types: classification and regression.

Classification is when the outcome is grouped into categories, like differentiating between male and female, yes or no. Regression is a mathematical method used when the output variable is a continuous value. There is a relationship between two or more attitudes or variables that changes in one of them and results in a change in the other as well.

2.1.3. Unsupervised Machine Learning

When a machine learning model searches for a pattern without any label or human supervision in the dataset, this is called unsupervised machine learning. In other words, unsupervised learning is when algorithms find patterns in datasets without any classification or labeling. The goal of this method is to categorize, label, and group the data points in the dataset, whereas the learning algorithm receives no labels. The algorithm is required to recognize patterns and arrangements within the input set. In unsupervised learning, an AI system arranges data that has not been organized according to similarities and differences, even if there are no specific groups or categories given (Hastie et al., 2013). There are four different kinds of unsupervised learning. They are called clustering, Principal component analysis, Anomaly detection, and auto-encoders.

Clustering is an important topic that involves finding patterns in a group of data that has not been sorted. Clustering algorithms in unsupervised learning will analyze the data and discover natural groups. The number of clusters can be defined depending on the algorithm. Clustering is divided into four types of categories: partitioning, hierarchal, overlapping, and probabilistic. Data is grouped in partitioning; therefore, one piece of data can only be part of one cluster, such as K-means. In overlapping fuzzy sets, they are used to cluster and group data. The iterative unions between the two closest clusters make fewer clusters overall. Each point can be part of multiple clusters with different levels of membership.

Association refers to a kind of unsupervised learning in which a computer program or system links various elements or patterns together without any form of

guidance. The association rule is a type of unsupervised learning; this means that the algorithm tries to learn on its own because the data does not have labels.

The association rule is a way to describe relationships in large datasets instead of predicting them. The connection between things is often shown using rules or groups of commonly found items (Cios et al., 2007).

Anomaly Detection is another rule that finds the outlier of a dataset. An anomaly is when the data does not follow the usual patterns. For instance, When the flow of traffic in a network appears abnormal, it indicates that the hacked system is transmitting significant data to an unauthorized host. To find anomalies, we need to spot or forecast the data points that differ from the normal pattern (Staff, 2020).

Auto-encoders are a type of unsupervised learning that uses neural networks to learn how to represent information on their own without needing guidance. This technique is used when the input is the same as the output (Van et al., 2017).

2.1.4. Semi-Supervised Machine Learning

Semi-supervised machine learning is a type of machine learning where a model is trained on a combination of labeled and unlabeled data. This approach can be useful when obtaining large amounts of labeled data is difficult or costly, but there is still some labeled data available to guide the learning process. The main aim of SSL is to fix the issues faced by both supervised and unsupervised learning methods. Supervised learning needs a lot of training data to determine the type of intro. This process can be expensive and takes a long time. On the other hand, unsupervised learning is different because it does not require labeled data. Instead, clusters the data by finding similarities between different data points, either by clustering or using the maximum likelihood method. The main downfall of this approach is that it cannot cluster unknown data accurately (Reddy et al., 2018). To solve these problems, researchers have come up with SSL. SSL can learn with only a little bit of training data and identify unknown (or test) data. SSL creates a model using a small number of labeled patterns as training data, and the remaining patterns are used as test data. This learning method is divided into two types: Semi-supervised classification and Semi-supervised clustering. Semi-Supervised Classification (SSC) is like the Supervised method, but it

needs more training data to classify the test data (Reddy et al., 2018). However, in SSC, we only need a small amount of training data to categorize a large amount of test data. By using this semi-supervised classification, we need less training data. At present, there are more unlabeled data patterns that can be used for research but not enough labeled data. Because training data design is costly, effective, and time-consuming.

2.1.5. Experts' Systems and Knowledge Base

Expert systems are computer systems that try to copy how a human expert in a certain field makes decisions. They use a knowledge base, which is a repository of facts, rules, and heuristics that capture the knowledge of a domain expert. The knowledge base is combined with a reasoning engine, which uses the knowledge to provide advice or solve problems in the domain. Knowledge base is the foundation of an expert system. It is a collection of information that describes the domain of the expert system, including the concepts, rules, procedures, and heuristics used in the decision-making process. The knowledge is often represented using a formal language such as rules, frames, or semantic networks. The knowledge base is built by domain experts who possess extensive knowledge in a particular domain. They provide the knowledge and expertise required for the system to make informed decisions. The experts work with knowledge engineers, who are responsible for capturing, organizing, and encoding the knowledge into a format that can be used by the reasoning engine. The reasoning engine is the software component that processes the knowledge in the knowledge base to provide advice or solve problems. It uses a variety of reasoning methods, including deduction, induction, and abduction, to make inferences and draw conclusions based on the knowledge in the knowledge base. Expert systems have been used in many domains, like pharmaceuticals, engineering, finance, and law. They can provide valuable advice and assistance to users and often make decisions more quickly and accurately than human experts. However, expert systems have some limitations. They are only as good as the knowledge in the knowledge base and may not be able to handle situations outside their domain (Atymtayeva et al., 2014). They also require significant resources to develop and

maintain and may not be cost-effective for all applications. Nevertheless, expert systems remain an important area of research and development in artificial intelligence.

2.2. NEURAL NETWORK AND DEEP LEARNING

Neural networks and deep learning have become very popular in recent years due to advancements in computing power, big data, and the internet. Neural networks are machine learning algorithms inspired by our brain's way of functioning. Conversely, deep learning is a subfield of machine learning that utilizes deep neural networks for solving complex tasks. This report will discuss the concepts, applications, and future of neural networks and deep learning.

Neural networks are computational models designed to learn and generalize from examples (Schmidhuber, 2015). They consist of interconnected processing units called neurons, which process information in a distributed manner. These neurons are organized into layers, where each layer transforms the input information into a more abstract representation. The input layer gets the original information, and the output layer creates the result. The hidden layers between the input and output layers do most of the computation and extract important information from the input.

Deep learning is a part of machine learning that uses deep neural networks with many hidden layers to do difficult tasks. Deep neural networks can learn and represent complex relationships between input and output data. Deep learning algorithms are based on a supervised learning approach, where the network is trained using a large dataset of labeled examples. During the training process, we change the weights and biases of the network so that the predicted output becomes as similar as possible to the actual output (Schmidhuber, 2015).

Neural networks and deep learning can be used in many different industries like medical care, banking, retail, and transportation. Some of the popular applications of neural networks and deep learning are: Image and speech recognition: Deep learning algorithms are used for image and speech recognition tasks. These algorithms can learn and identify patterns in images and sounds, making them ideal for applications such as self-driving cars, facial recognition, and voice assistants.

Deep learning algorithms can understand and analyze human language, making them ideal for applications such as chatbots, sentiment analysis, and language translation. It can be used to predict future trends and patterns based on historical data. This application is popular in finance, where neural networks can be used to predict stock prices, credit scores, and fraud detection.

Medical diagnosis is another ability of Deep learning algorithms to analyze medical images, such as MRI and CT scans, to accurately diagnose diseases and conditions.

The future of neural networks and deep learning looks promising, with new developments and applications being discovered every day. Some of the future advancements in neural networks and deep learning include: Reinforcement learning is a subfield of machine learning that focuses on decision-making in a dynamic environment. It involves learning through trial and error, where the system receives feedback based on its actions. Reinforcement learning has many potential applications, such as robotics and game development (D'Amico et al., 2019).

Explainable AI is an emerging field that centers on creating algorithms to clarify their decisions. This is very important, especially in areas like healthcare, because choices made by AI systems can greatly affect people's lives. Quantum computing: Quantum computing is a new computing paradigm that could potentially revolutionize the field of deep learning. Quantum computing can process vast amounts of data in a fraction of the time it takes classical computers.

Deep learning and neural networks have transformed machine learning and artificial intelligence. They have enabled computers to do complex tasks such as image and audio recognition, processing of natural languages, and analytics for prediction. The future of neural networks and deep learning appears bright, with breakthroughs and applications being found daily.

2.2.1. Data Science

Data Science is a combination of mathematics and statistics, programming, analytics, artificial intelligence, and machine learning with specific subject matter expertise to uncover and get meaningful insight hidden in data. Later, these results and

insights can be used in decision-making, risk analysis and assessment, and strategic planning. The aim of a data scientist is to make recommendations about what sort of data is useful or required. Data scientists and data analysts share similar responsibilities, particularly in the realm of data exploration and visualization. However, a data scientist's skillset is typically broader than that of the average data analyst. Data science consists of five steps: first, capturing the; second, data maintenance; third, processing; analyzing; and communicating (Provost & Fawcett, 2013).

Data scientists determine which questions must be answered and where to get the appropriate data. They are smart in business, can think analytically, and can collect, organize, and show data. Businesses hire data scientists to find, handle, and analyze large amounts of data that do not have a specific format or structure. The outcomes are combined and shared with key stakeholders to help them make important decisions for the organization.

2.2.2 Data Analyst

Data analysts connect data scientists and business analysts together. They receive questions from an organization and then examine and analyze the data to find answers and solutions that match the overall business plan. Data analysts are people who translate technical data into practical insights and share their findings with diverse stakeholders (B & Michael, 2008).

2.2.3. Data Science Tools

There are several tools and platforms available for doing analysis and working with big data. These tools play a crucial role in extracting and getting insight from raw data. There are a variety of tools and techniques that enable data to be processed, visualized, and predicted, like statistics, computer science, predictive modeling, and deep learning. The most popular and famous programming languages for data science are Python and R. Since data scientists also should work with data storage and

databases, SQL, PLSQL, and Hadoop are examples of some tools for getting queries and managing data from databases (Barlas et al., 2015).

2.2.4. Python

Python is a highly popular programming language, and in 1991, it was developed by Guido van Rossum. It is often used to automate tasks through machine learning, AI, and website building. A library in Python is a group of interconnected modules that simplify and facilitate Python programming for programmers. The most important Python library is Pandas, a data analysis and manipulation library. Mostly used in data science for data cleaning, preprocessing, and manipulating data, Pandas might be the most widely used Python library, and Scikit-Learn is mostly used in machine learning. It covers various machine learning algorithms, from linear regression to complex neural networks. Moreover, it gives center capacities for information preprocessing and feature engineering. Also, Scikit-learn provides metrics to evaluate the performance and accuracy of the model. The other library is Matplotlib, which is used in data visualization in data science and helps to discover, interpret, and analyze data. There is a library called TensorFlow that is used for fast numerical computing. Keras created and released it, a deep-learning library that can be used in conjunction with TensorFlow or Theano.

Measures for evaluating the model considered in this study include F1 Score, Precision, Accuracy, Recall, ROC, and AUC.

2.2.5. Knime

The KNIME program Analytics Platform is a useful and complete free platform for dragging and dropping analytics, machine learning, and statistics. Users may do analysis without any coding skills.

There are core features for Knime, such as the intelligent feature of automatically storing data in the background while maximizing throughput performance, the import and export of workflows, and easy drag-and-drop nodes for analysis.

2.2.6 Predict Sepsis Using Knime and Python

This study proposes a deep-learning framework for the detection of Sepsis. Knime has a special place in this study. Logistic regression is used among machine learning models, as are decision trees. However, Keras deep learning is used in both Python and Knime. But the last and better result is from logistic regression.

Different models and libraries are used in Python to increase the accuracy of predicting sepsis and comparing results. Models and methods that are used in this study are provided below.

2.2.7. Decision Tree

Decision trees are a form of supervised machine learning that is used to categorize or predict based on the root node and are often used to represent classifiers. The model is a sort of supervised learning in which it is trained and tested using a dataset that already contains the desired categories. It is composed of nodes that form a rooted tree, with the 'root' node having no incoming edges. The remaining nodes are known as leaves or decision nodes. A decision tree's internal nodes divide the instance space into groups depending on a function that examines the input properties. Scikit-learn is a popular Python framework for generating decision trees (Safavian & Landgrebe, 1991).

2.2.8. Linear Regression

There are two forms of analysis for supervised learning: regression and classification. The output variable to be predicted in regression is continuous in nature (home price, student grade). Linear regression is a statistical analysis used to determine the relationship between two variables. This method believes that there is a linear relationship between the independent variable and the dependent variable being studied, and Its goal is to find the most accurate line that explains this connection **(**Ying Wu, Y. 2012). The line is obtained by adding the sum of the squared differences between the predicted and actual values as small as possible. This model is mainly

used in economic, financial, and social sciences to analyze and predict trends in data (Montgomery et al., 1993).

2.2.9. Logistic Regression

Logistic regression, like linear regression, employs a binomial response variable. The nursing field is increasingly adopting this model for analysis as it effectively demonstrates the correlation between dichotomous dependent variables and one or more independent variables. The binary variable under consideration is known as the response variable or the dependent variable.

We will refer to the variable being modeled as "response" because that is the preferred phrase these days. For a model to fit the data well, it is anticipated that the predictors are either uncorrelated or significantly related to the response, and the last one is when the observations or data elements of a model are also uncorrelated.

Logistic regression is a popular statistical approach for binary classification issues that may also be utilized in medical data analysis, such as predicting sepsis (Hilbe, 2018). Here's a general approach for using logistic regression in the context of predicting sepsis, as this study used logistic regression for the prediction of sepsis.

Data collection: collect a dataset that includes relevant variables or features from patients who have been diagnosed with or treated for sepsis. Demographic information, vital signs, laboratory data, and any other characteristics that may be symptomatic of sepsis might be included.

The First step in analyzing data is Data preprocessing. Preprocess the collected data to ensure it is suitable for logistic regression. This step involves handling missing data, removing irrelevant or redundant features, normalizing, or standardizing numeric features, and encoding categorical variables appropriately. Then, Split the dataset into two subsets: a training set and another set for testing or validation. The training set will be used to train the logistic regression model, while the testing and validation sets will be used to evaluate its performance.

Another important step is featuring selection. Conduct feature selection to identify the most relevant features for predicting sepsis. This can be done using

techniques such as correlation analysis, stepwise selection, or are regularization methods like Lasso or Ridge regression.

Lastly, we train our model using the training dataset. The model will learn the relationships between the selected features and the occurrence of sepsis. Evaluate the trained model using the testing and validation dataset. Common evaluation metrics for binary classification models include accuracy, precision, recall, F1 score, and area under the receiver operating characteristic curve (AUC-ROC).

Model optimization: If the model's performance is not satisfactory, it can iterate and optimize it by trying different feature sets, tuning hyperparameters, or considering other machine learning algorithms.

Once the model has been validated, we may use it to predict sepsis in new patients by inputting their relevant data into the trained logistic regression model. The output of the model will be the predicted probability of sepsis, which can be converted into binary predictions using an appropriate threshold.

It is important to note that building an accurate sepsis prediction model requires high-quality data, domain expertise, and careful consideration of ethical and legal implications. Collaboration with healthcare professionals and adherence to relevant guidelines and regulations are crucial for the successful and responsible application of logistic regression or any other predictive modeling technique in a medical setting.

2.2.10. TensorFlow and Keras

TensorFlow is a popular open-source deep learning framework made by Google that helps to create and use machine learning models and has different tools, libraries, and resources available. TensorFlow supports numerous neural network types, such as convolutional neural networks (CNNs), recurrent neural networks (RNNs), and transformers. It offers a highly optimized backend for numerical computations, allowing efficient execution of complex mathematical operations on CPUs, GPUs, and even specialized hardware like Tensor Processing Units (TPUs). It also provides functionalities for automatic differentiation, which are crucial for training deep learning models using techniques like backpropagation (Haghighat & Juanes, 2021).

Another popular deep-learning framework is Keras, an independent library in Python but now an integral part of Tensorflow. It simplifies building deep learning models, making them accessible to beginners and experts alike. Keras provides a high-level API, enabling users to construct neural networks effortlessly by stacking layers and choosing from a wide range of activation functions, optimizers, and loss functions. Its modular design allows for flexible model architectures, including sequential and functional models. The design philosophy behind Keras revolves around enabling rapid prototyping of deep learning models. The modularity of Keras enables the construction of both sequential and complex network architectures, including multi-input and multi-output models (Haghighat & Juanes, 2021). Overall, Keras serves as a user-friendly interface to deep learning frameworks, enabling researchers and practitioners to efficiently develop and experiment with complex neural networks, accelerating progress in the field of artificial intelligence.

2.2.11. The Synergy of Keras and TensorFlow

Keras and TensorFlow are closely intertwined, with Keras acting as the high-level interface that sits atop the TensorFlow backend. This powerful combination allows users to leverage the performance and flexibility of TensorFlow while benefiting from the simplicity and ease of use provided by Keras. With Keras, developers can rapidly prototype and experiment with different network architectures, seamlessly integrating Tensor Flow's advanced functionalities when needed.

Keras and TensorFlow have emerged as indispensable tools in the deep learning landscape, empowering researchers, developers, and data scientists to tackle complex problems and drive advancements in artificial intelligence. With their ease of use, flexibility, and impressive performance, these frameworks have made deep learning accessible to a wide audience, accelerating the development and deployment of state-of-the-art machine-learning models (Gambús & Shafer, 2018).

Therefore, as a beginner venturing into deep learning or an experienced practitioner pushing the boundaries of AI, Keras, and TensorFlow stand ready to assist in building intelligent systems that can tackle the most challenging tasks.

2.2.12. Predicted Sepsis using Knime

Knime, a powerful data analytics and machine learning platform, can indeed be used for predicting sepsis. By leveraging the data preprocessing, feature engineering, and machine learning capabilities of Knime, it can develop predictive models to identify patients at risk of sepsis. The diagram below is used for predictive sepsis in this study.

Utilize Knime's machine learning capabilities to build predictive models. This can involve using algorithms such as logistic regression, decision trees, random forests, or more advanced techniques like gradient boosting or deep learning. Then, for evaluation, the model should assess the performance of the developed model using appropriate evaluation metrics such as accuracy, sensitivity, specificity, and area under the receiver operating characteristic curve (AUC-ROC). Finally, Deploy the trained model within Knime for real-time prediction or validation on new patient data. This enables healthcare professionals to make informed decisions and take preventive actions when sepsis risks are detected. It should be considered that creating accurate and reliable sepsis prediction models requires careful consideration of data quality, feature selection, model selection, and thorough validation. In this study, data was collected from Dokuz Eylul Hospital.

2.3. MODEL EVALUATION'S METRICS

Model assessment metrics are quantitative measures used to assess a machine learning model's performance. These metrics aid in measuring how well the model performs, recognizing its strengths and limitations, and comparing several models to determine which is best for a particular position. These metrics include Accuracy and Confusion.

2.3.1 Confusion Matrix

As shown in Table 2, the Confusion Matrix compares the model's categorization of different fault groups to the real classification and includes TP, FP, TN, and FN (Bowes, Hall and Gray, 2012).

Table 2: Confusion Matrix for Binary Classification

	Predicted False	**Predicted True**
Actual False	TN	FP
Actual True	FN	TP

Source: Bowes et al., 2012.

To measure accuracy, we need a confusion matrix, the Correct Classification Rate (CCR) is the measure of accuracy

2.3.2 Accuracy

Accuracy, which describes the number of right predictions over all predictions, is frequently used as the basic measure for model evaluation: It calculates the proportion of successfully predicted occurrences to total cases in the dataset. While accuracy is a simple and easy-to-understand statistic, it has limitations, particularly when dealing with unbalanced datasets in which one class dominates the others. Accuracy could fail to provide an accurate depiction of the model's performance in such cases since it may be substantially impacted by the dominant class. As a result, depending on the specific needs and features of the situation at hand, additional assessment metrics like as accuracy, recall, and F1 score should be considered. To assess a model's accuracy, divide the number of correct forecasts by the total number of predictions produced as follows: Accuracy = $(TP+TN)$ / $(TP+TN+FP+FN)$ (Tharwat, 2020).

2.3.3. Error Rate

The error rate is a model evaluation statistic that calculates the proportion of inaccurate predictions generated by a machine learning model. It is the complement of accuracy and reflects the proportion of incorrectly predicted occurrences in relation to the total number of cases in the dataset (Tharwat, 2020). The error rate may be determined mathematically as follows:

Error Rate = **(**$FP+FN$) / ($TP+TN+FP+FN$).

2.3.4. Precision

Precision is a model evaluation metric that quantifies the proportion of accurately predicted positive cases in comparison to all positive instances anticipated. It evaluates the model's ability to anticipate good outcomes accurately (Tharwat, 2020). Precision is computed mathematically as:

Precision = TP / (TP + FP)

Precision focuses on the quality of positive predictions and helps assess the model's ability to avoid false positives.

2.3.5. Recall

Recall, referred to as sensitivity or true positive rate, is a model assessment statistic that evaluates the proportion of properly predicted positive cases in the dataset vs all actual positive instances (Tharwat, 2020). It measures the model's ability to accurately detect positive events. The recall is determined as follows:

Recall = TP / (TP + FN)

A good recall score indicates that the model properly recognizes many positive cases.

2.3.6. F1 Score

The F1 score is a statistic for model evaluation that combines accuracy and recall into a single figure. It works especially well when dealing with skewed datasets or when both false positives and false negatives are common. The F1 score is calculated using the harmonic average of accuracy and recall, which offers a fair assessment of the algorithm's performance (Tharwat, 2020). It is calculated as follows:

F1 Score = 2 * (Precision * Recall) / (Precision + Recall)

The F1 score ranges from 0 to 1, where a higher value indicates better model performance.in other words, a higher F1 score suggests that the model achieved high accuracy and recall simultaneously.

2.3.7. AUC

AUC (Area Under the Curve) is a model evaluation metric commonly used in binary classification tasks to assess the performance of a machine learning model's predictions. This curve compares how often true positive results happen compared to false positive results at different thresholds used for classification. AUC provides a measure of the model's ability to discriminate between positive and negative instances across all possible threshold values. A higher AUC score, indicates better model performance, ranging from 0 to 1, where 1 representing a perfect classifier (Lobo et al., 2008).

2.3.8 Sensitivity

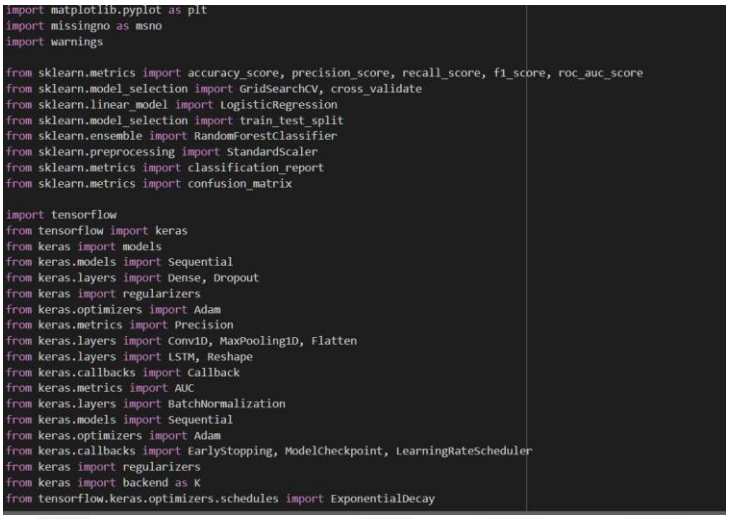
Sensitivity is related to true positive and is defined as the fraction of individuals accurately assigned a positive allocation out of all subjects who have a genuinely positive for the result, reflecting how well a test can categorize subjects who have the desired outcome (Seo et al., 2009). Which can calculate as divided true positives on the sum of true positives and false negatives. Sensitivity = TP/(TP+FN)

CHAPTER THREE
METHODOLOGY, MODELING EVALUATION, AND DATA ANALYSIS OF SEPSIS

3.1. APPLICATION

This section analyzes data using machine learning techniques to develop a reliable and accurate predictive model for sepsis. As previously stated in this study, Sepsis is a complicated and time-sensitive illness, making early identification vital for successful treatment and enhanced patient outcomes. We want to uncover patterns and markers that can effectively forecast the development of sepsis by analyzing a large quantity includes patient information, such as vital signs, test findings, and clinical notes, using machine learning algorithms. The ultimate objective is to develop a strong prediction model that may help healthcare practitioners proactively identify people at risk of sepsis, allowing for prompt treatments and perhaps saving lives. In recent years, combining KNIME, a powerful data analytics platform, and machine learning techniques has shown great promise in improving sepsis prediction. It could be developed robust predictive models with logistic regression by leveraging KNIME's intuitive interface and extensive set of machine learning nodes. For executing Python codes, Google Colab, which is short for Google Colaboratory, is used. This platform provides a free cloud-based platform where it can write and execute Python code. It offers a range of features and benefits for data scientists and machine learning practitioners. Keras-metrics, Tensorflow, Keras-upgrade, are the required libraries for this study.

Figure 2: Python Libraries for Machine Learning and Deep Learning



```
import matplotlib.pyplot as plt
import missingno as msno
import warnings

from sklearn.metrics import accuracy_score, precision_score, recall_score, f1_score, roc_auc_score
from sklearn.model_selection import GridSearchCV, cross_validate
from sklearn.linear_model import LogisticRegression
from sklearn.model_selection import train_test_split
from sklearn.ensemble import RandomForestClassifier
from sklearn.preprocessing import StandardScaler
from sklearn.metrics import classification_report
from sklearn.metrics import confusion_matrix

import tensorflow
from tensorflow import keras
from keras import models
from keras.models import Sequential
from keras.layers import Dense, Dropout
from keras import regularizers
from keras.optimizers import Adam
from keras.metrics import Precision
from keras.layers import Conv1D, MaxPooling1D, Flatten
from keras.layers import LSTM, Reshape
from keras.callbacks import Callback
from keras.metrics import AUC
from keras.layers import BatchNormalization
from keras.models import Sequential
from keras.optimizers import Adam
from keras.callbacks import EarlyStopping, ModelCheckpoint, LearningRateScheduler
from keras import regularizers
from keras import backend as K
from tensorflow.keras.optimizers.schedules import ExponentialDecay
```

Source: Created by author

In this study, data was collected from Dokuz Eylul Hospital, and after preprocessing the data, Knime and Python were used to achieve the result. Decision trees, logistic regression, and Keras were the deep learning models used for this study.

3.2. INTRODUCING DOKUZ EYLÜL HOSPITAL

Dokuz Eylül Hospital is a prominent healthcare institution located in İzmir, Turkey. It is affiliated with Dokuz Eylül University, one of the leading universities in the country. The hospital is a prominent referral center, providing comprehensive medical services to a vast population in the region. With its state-of-the-art facilities, advanced medical technologies, and highly skilled medical staff, Dokuz Eylül Hospital offers a wide range of specialized departments and services, including but not limited to cardiology, oncology, orthopedics, neurology, pediatrics, and emergency medicine. The hospital is dedicated to delivering high-quality patient care, promoting medical education and research, and contributing to advancing healthcare in Turkey. It plays a vital role in training future healthcare professionals and is renowned for its commitment to clinical practice improvement, innovation, and patient-centered care.

3.3 ICU

An intensive care unit is a treatment facility for extremely ill patients who need medical attention promptly. An intensive care unit (ICU) is where critically ill or injured patients get specialized treatment and are kept alive. Unless hospitalized in an emergency, admission to the ICU requires a reference from a physician or specialist. Each ICU at a hospital has a distinct atmosphere that reflects the particular surgeries and medical treatments done. Most ICUs, or intensive care units, are comparatively big, sterile facilities with many specialized, technical, and monitoring devices needed to care for critically sick patients. Some visitors and patients might find the ICU setting daunting due to the activity, noises, equipment, tubes, and monitors. In addition, ICUs often feature a greater doctor-to-patient ratio.

3.4. INTRODUCING DATASET AND DATA DESCRIPTION

The dataset used in this study was collected from the ICU of Dokuz Eylül Hospital from 2021–2022, in a CSV file. The dataset is rather small, with 488 patient records from different dates and times after entering the ICU. However, because it was manually collected, the quality of the data is rather good. There are 59 factors in the data gathered on each patient.

Detected from blood tests, whole indicators are from blood test results and repetition since the first admission date in the ICU and several times at different dates. Variables include vital signs including heart rate, RBC, HGB, and WB. The remaining values are laboratory values, as shown in Table 3 We observed the blood test result and the sepsis diagnosis related to it; for instance, on the first admission date, patients were observed as having non-sepsis and corresponding blood test results were recorded. On any day of measurement, if sepsis is recognized, it is accepted as day zero for the sepsis corresponding result for the blood test recorded. Therefore, there are both sepsis and non-sepsis days' blood test results for the same patient. In this study, the data considered depends on a patient's condition two or three days after day one of sepsis. In addition, the data was split into two datasets: training and testing data.

Out of a total 488 records of patients after cleaning data and eliminating ,238 number of patients are considered in the testing dataset and rest in training.

Training dataset: first date of admission with a non-sepsis diagnosis and the last date in the intensive care unit with a sepsis diagnosis Test data includes other repeated measures of patients except for their first and last dates in the ICU. Table 4 illustrates the actual numbers of training and test sets, as well as the class distribution. Also, there are two partitions of datasets: the first one comes during COVID-19 days in 2021 with 556 records, and the second one is after COVID-19 in 2022 with 785 patients. These two datasets have been combined and then split into training and test data. Elimination procedure from 100 to 14.

Table 3: List of Variables in The Study

Vital Signs:	WBC,HGB,RBC,BASO#,BASO%,EOS#,EOS%,HCT,LYM#,LYM%,MCH,MCHC,MCV,MONO#,MONO%,MPV,NEU#,NEU%,PCT,PLT,RDW
Laboratory values:	SD_C_EGC, SD_C_LY, SD_C_MO, SD_C_NE, SD_MALS_EGC, SD_MALS_LY, SD_MALS_MO, SD_MALS_NE, SD_V_EGC, SD_V_LY, SD_V_MO, SD_V_NE, MDW, MN_C_EGC, MN_C_LY, MN_C_MO, MN_C_NE, MN_MALS_EGC, MN_MALS_LY, MN_MALS_MO, MN_MALS_NE, MN_V_EGC, MN_V_LY, MN_V_MO, MN_V_NE
Demographics:	Age, Gender
Outcome:	SepsisLabel
Selected features:	EOS, MDW, MN_C_LY ,MN_MALS_EGC, MN_V_EGC, MN_V_NE, SD_C_NE, SD_MALS_NE, SD_V_EGC, SD_V_LY, SD_V_MO,SD_V_NE

Source: Dokuz Eylül Hospital, 2020-2021

Partitioning is used to create separate subsets of the dataset for our training and testing the model. We used this method for our Deep Learning model- Keras and Logistic regression. The training set is used to train the logistic regression model on labeled data, while the testing set is used to assess the model's performance on unlabeled data. Proper partitioning guarantees that the model generalizes properly and can predict accurately on fresh data. In Decision Tree we didn't split data into training and testing sets.

Table 4: Partitioned Data Sets

Partitioned Dataset Detailed Counts (patients with multiple admissions)			
	Total	0(non – Sepsis)	1(Sepsis)
Training set	288	154	134
Test Set	250	182	68

Source: Created by author

3.5. PREDICTION

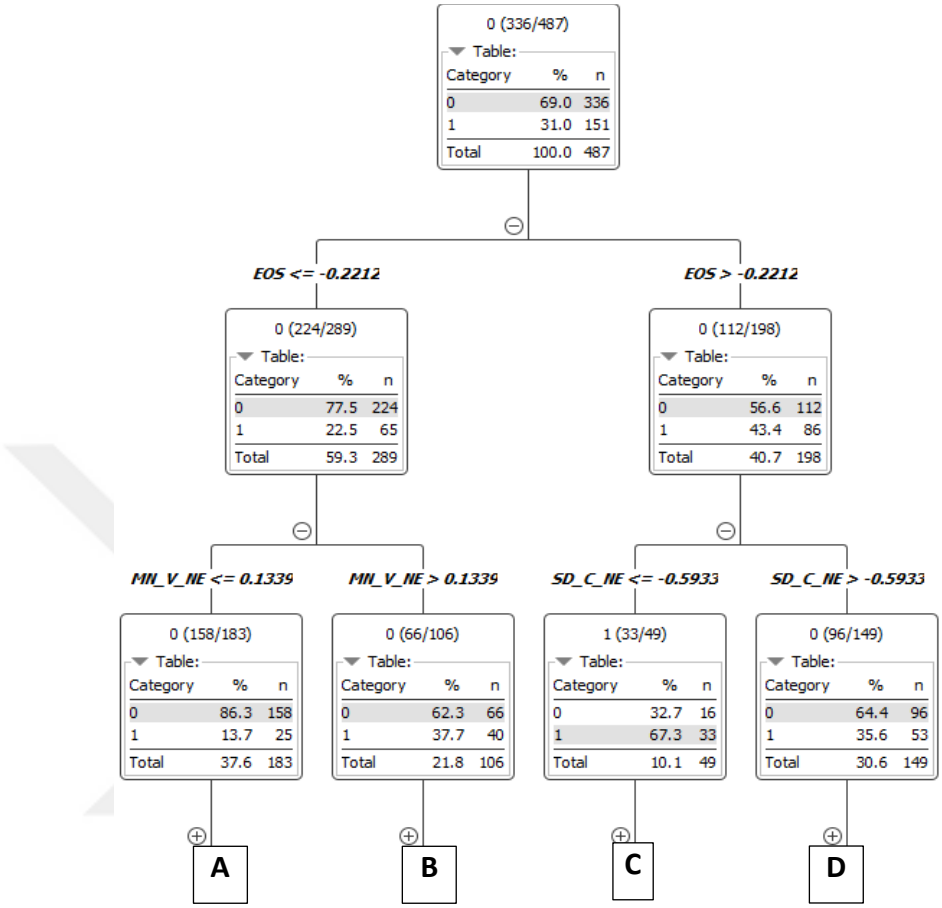
3.5.1. Prediction Using Decision Tree

Based on the characteristics retrieved from our dataset, we created a prediction model for sepsis, which is shown in Table 3 For the logistic regression model, a decision tree had been used for feature selection and classification. Decision trees are supervised learning algorithms that may be used to categorize objects or perform regression tasks. They function by recursively splitting the data based on the supplied feature values. We investigated the value or significance of each feature in the tree construction process while utilizing decision trees for feature selection. The importance of a feature is typically measured by its ability to split the data and reduce impurity or increase information gain at each node of the tree.

3.5.2. Prediction Using Logistic Regression

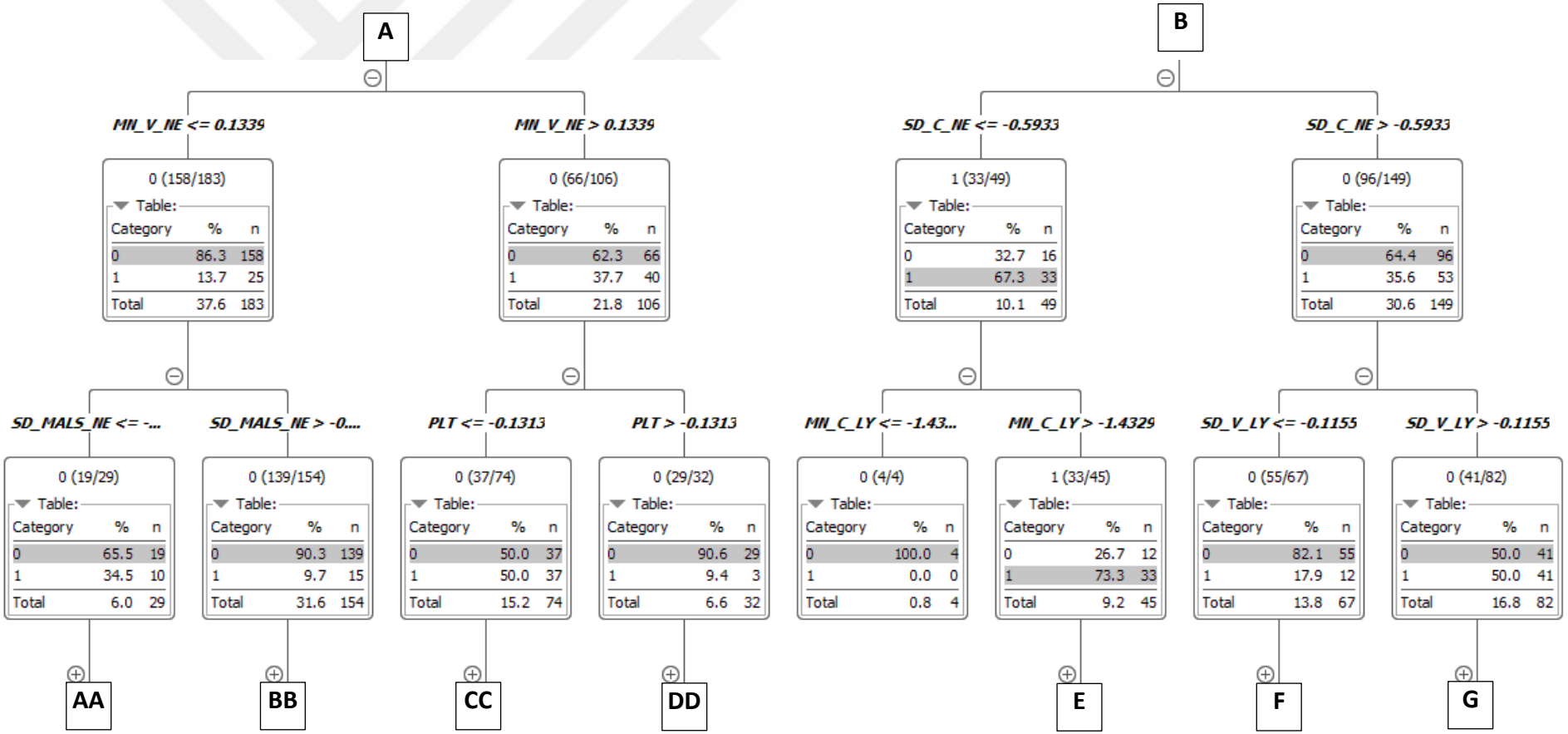
There are different ways to assess feature importance in decision trees, such as Gini importance or mean decrease impurity. Gini importance measures the total reduction in the impurity criterion (e.g., Gini index) achieved by a feature across all nodes in the tree. Mean decrease impurity, on the other hand, calculates the average reduction in impurity for each feature. By examining the feature importance scores obtained from the decision tree, we can rank the features and select the most relevant ones. This process allows us to identify the subset of features that contribute the most to the predictive power of the model, potentially improving both the efficiency and accuracy of the machine learning model. In our first decision tree the first node is EOS then it splitted into two branches MN_V_NE and SD_C_NE, then again, each node splitted in two branches as it has a seen in Figure 3 (a-k).

Figure 3: Decision Tree Analysis Results



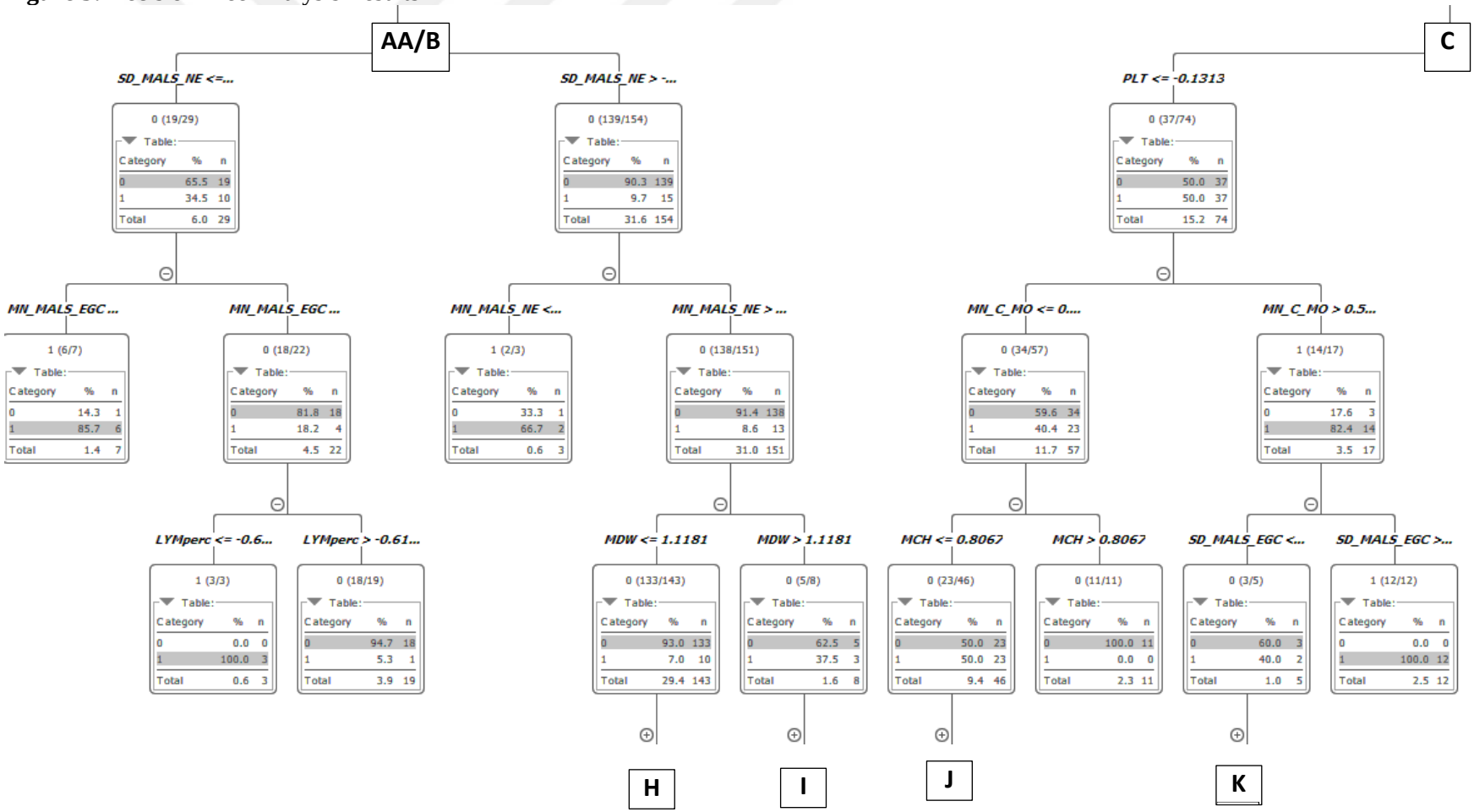
Source: Created by author

Figure 4: Decision Tree Analysis Results



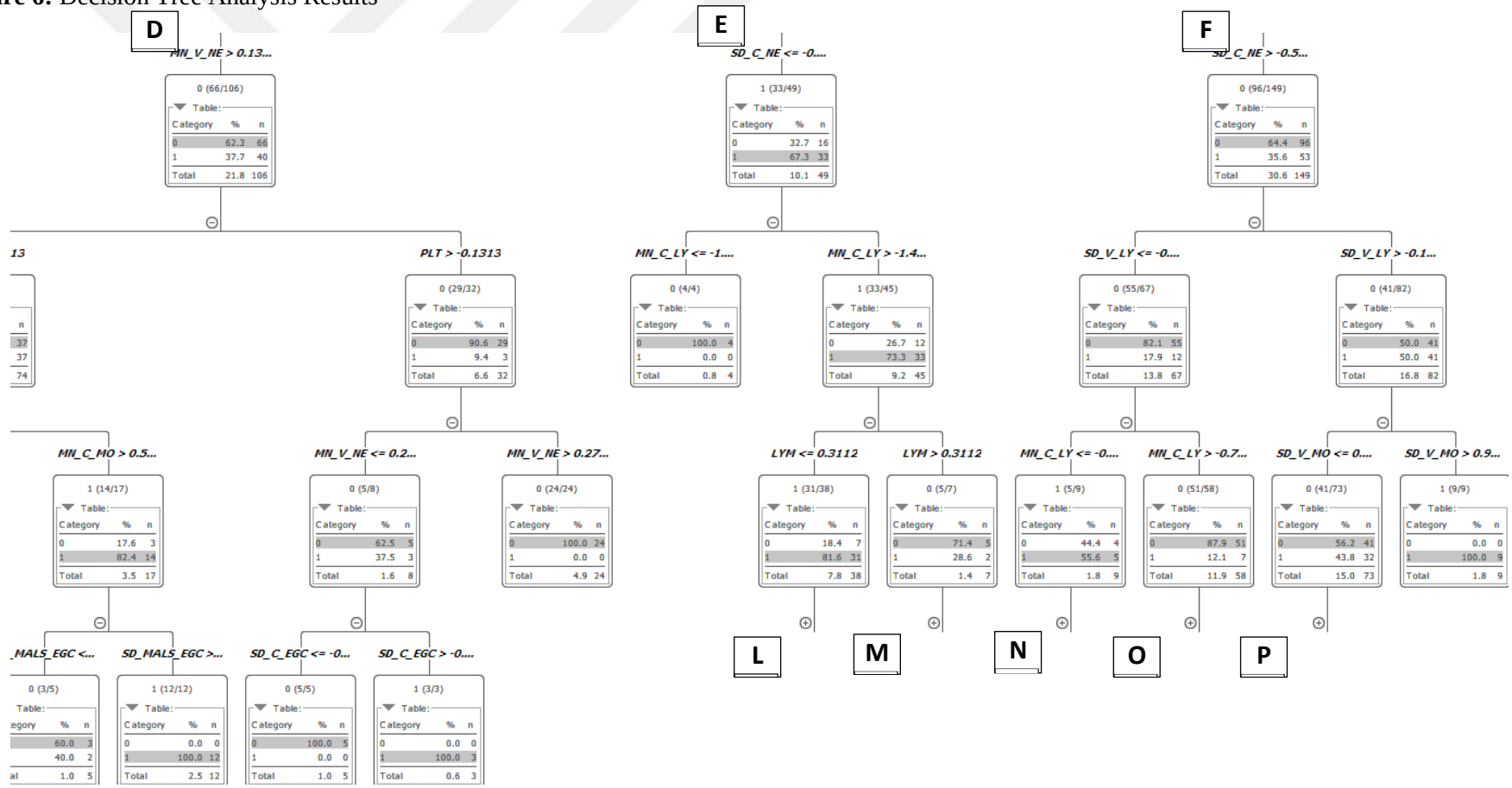
Source: Created by author

Figure 5: Decision Tree Analysis Results



Source: Created by author

Figure 6: Decision Tree Analysis Results



Source: Created by author

Figure 7: Decision Tree Analysis Results

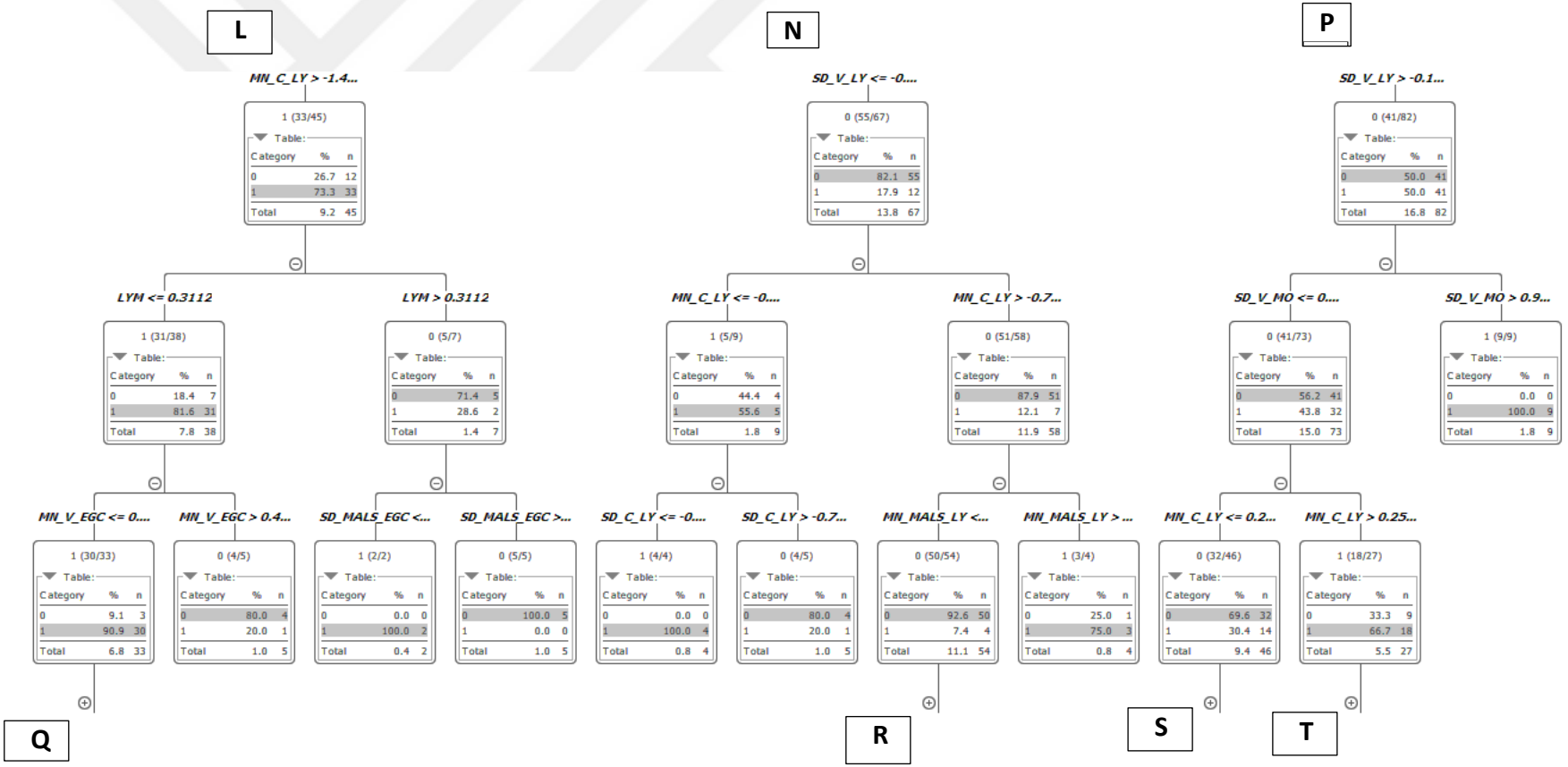
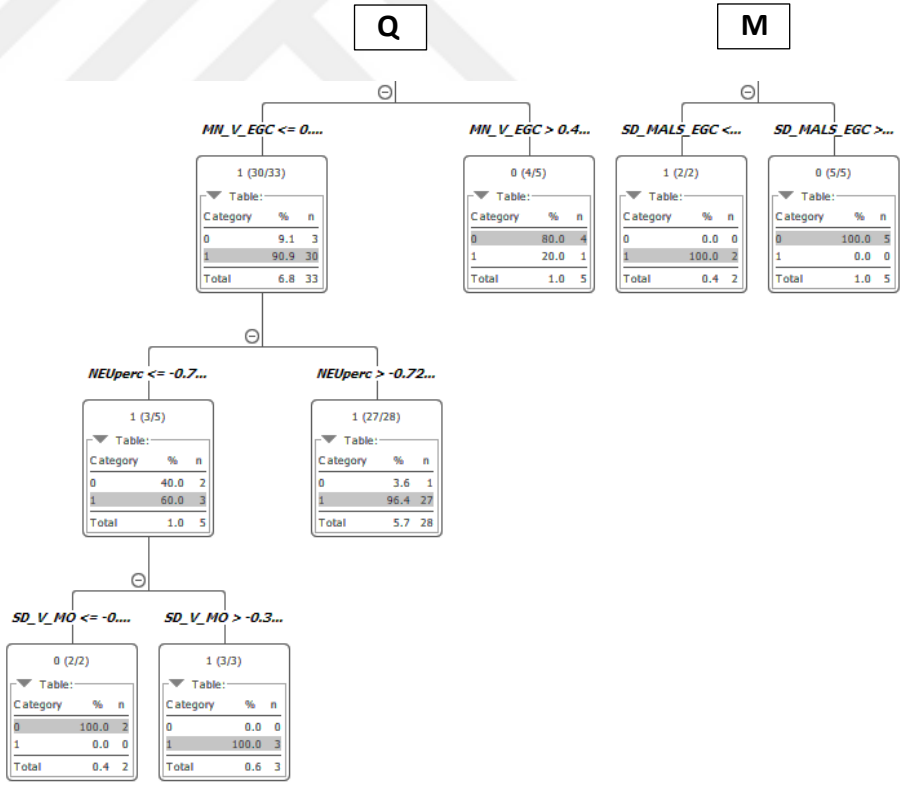
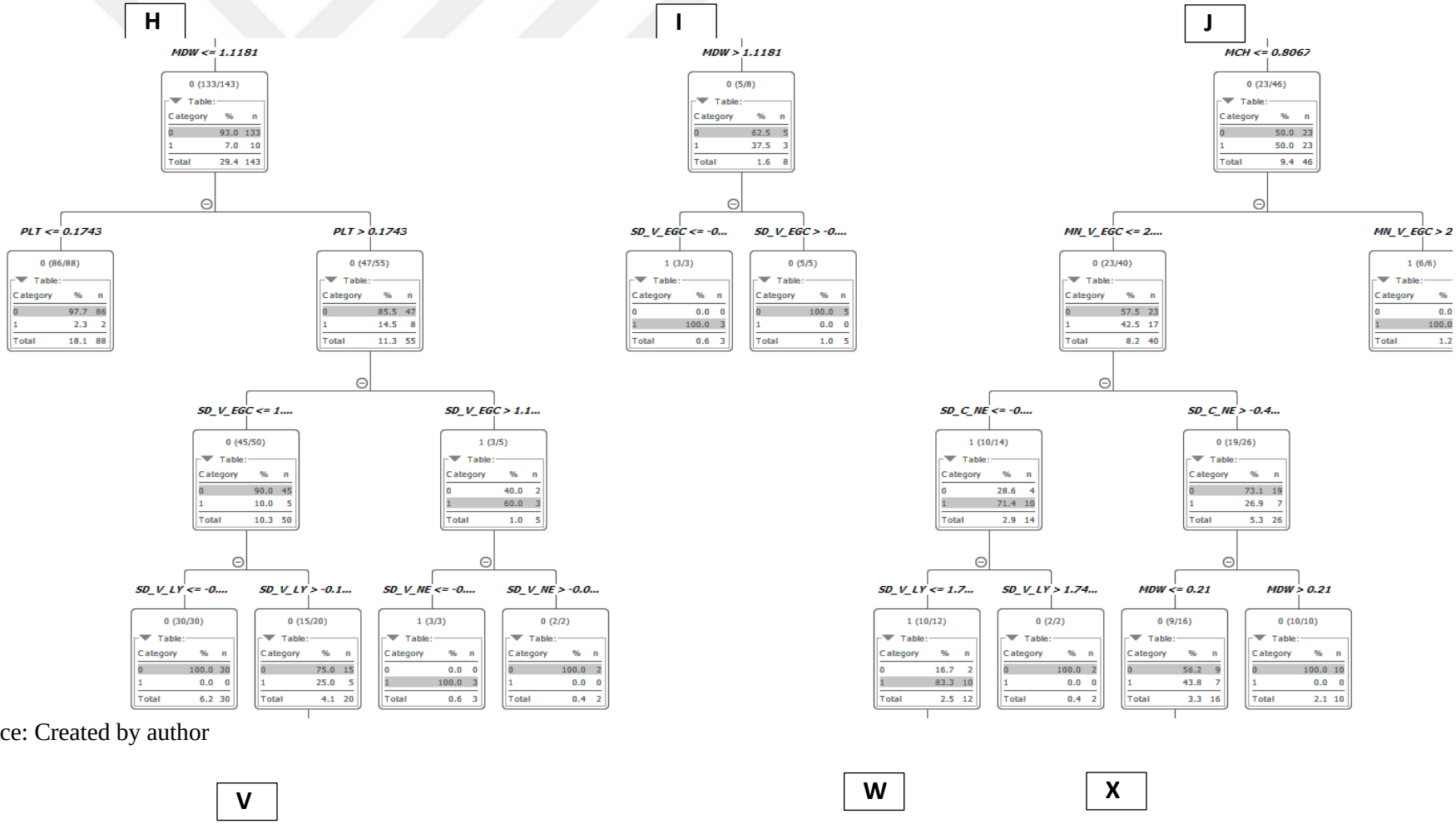


Figure 8: Decision Tree Analysis Results



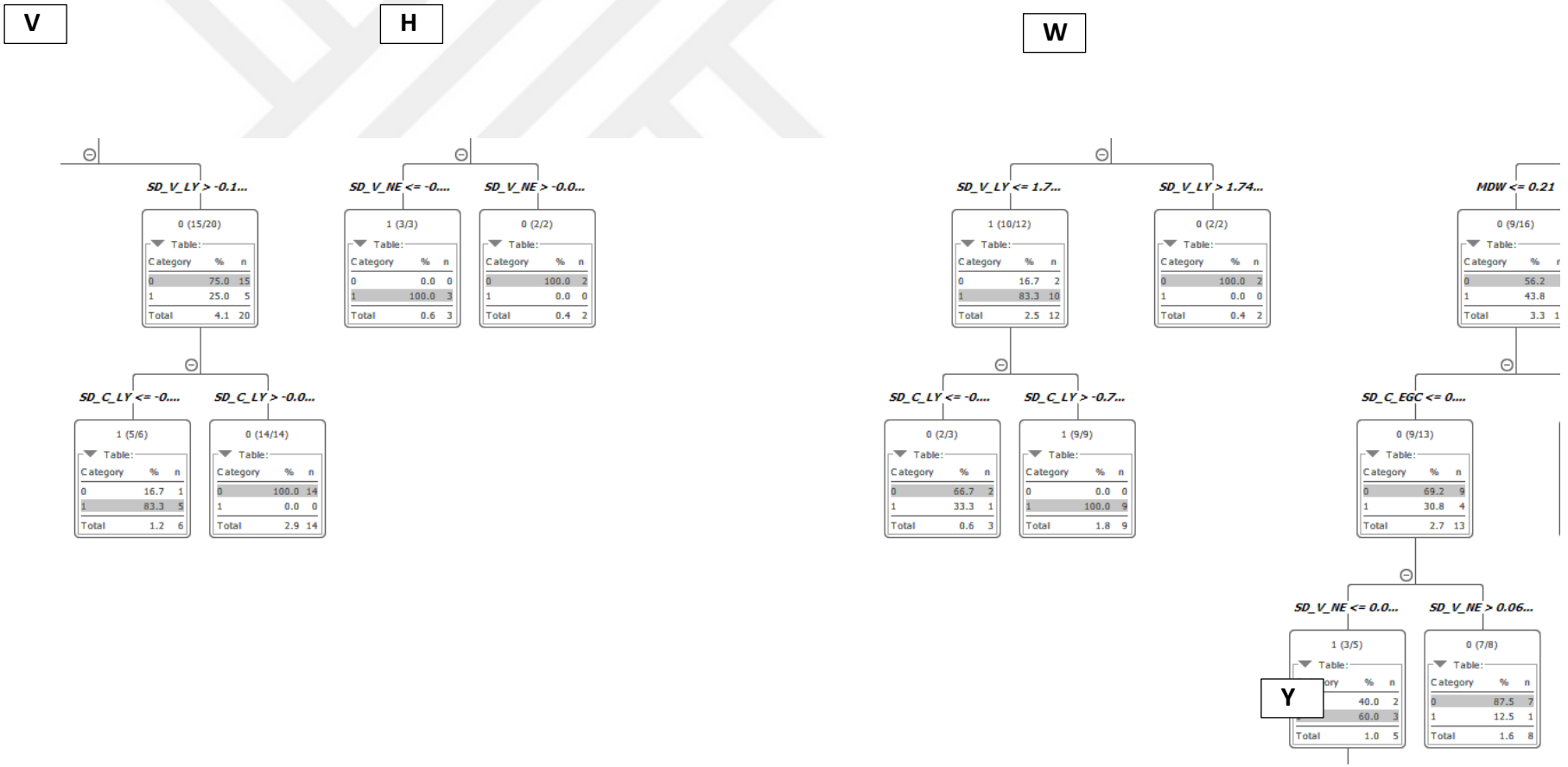
Source: Created by author

Figure 9: Decision Tree Analysis Results



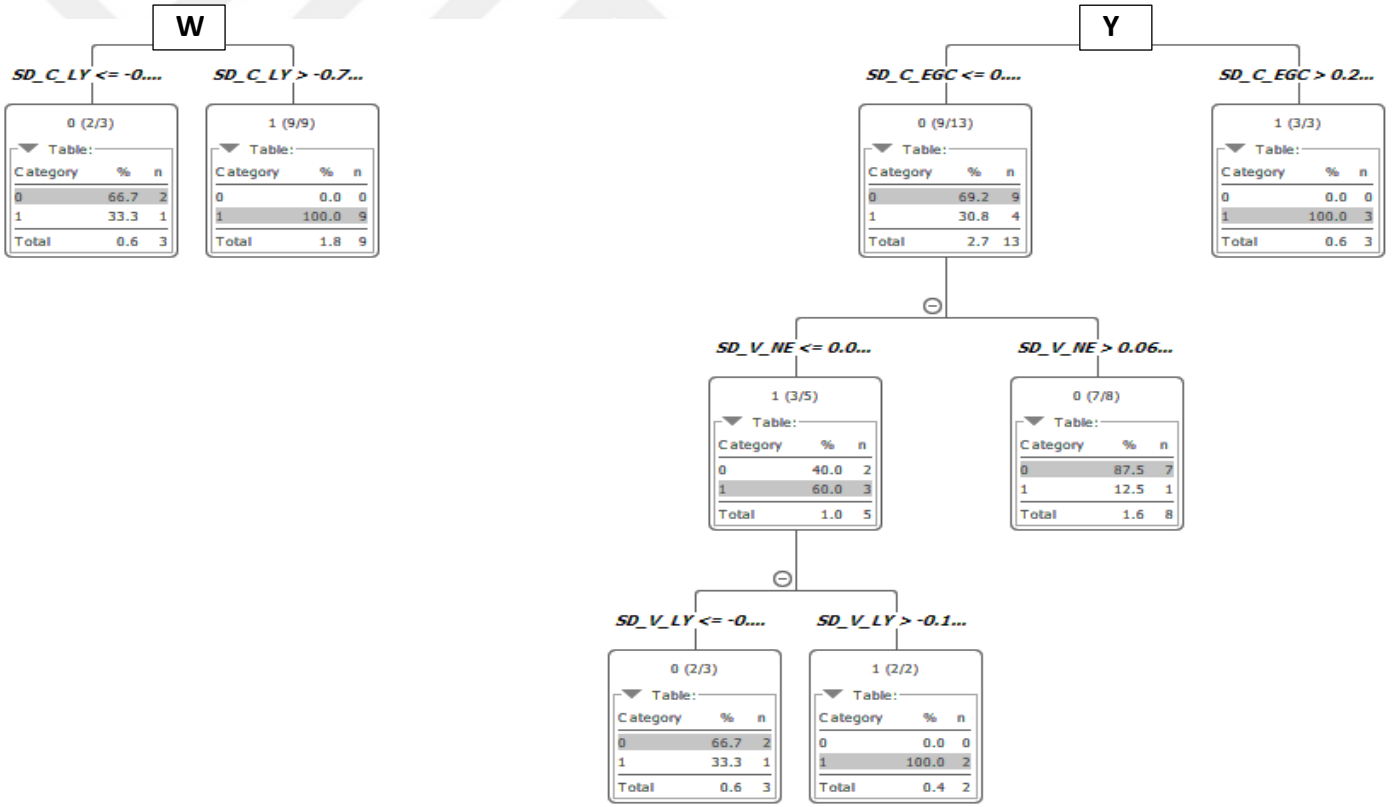
Source: Created by author

Figure 10: Decision Tree Analysis Results



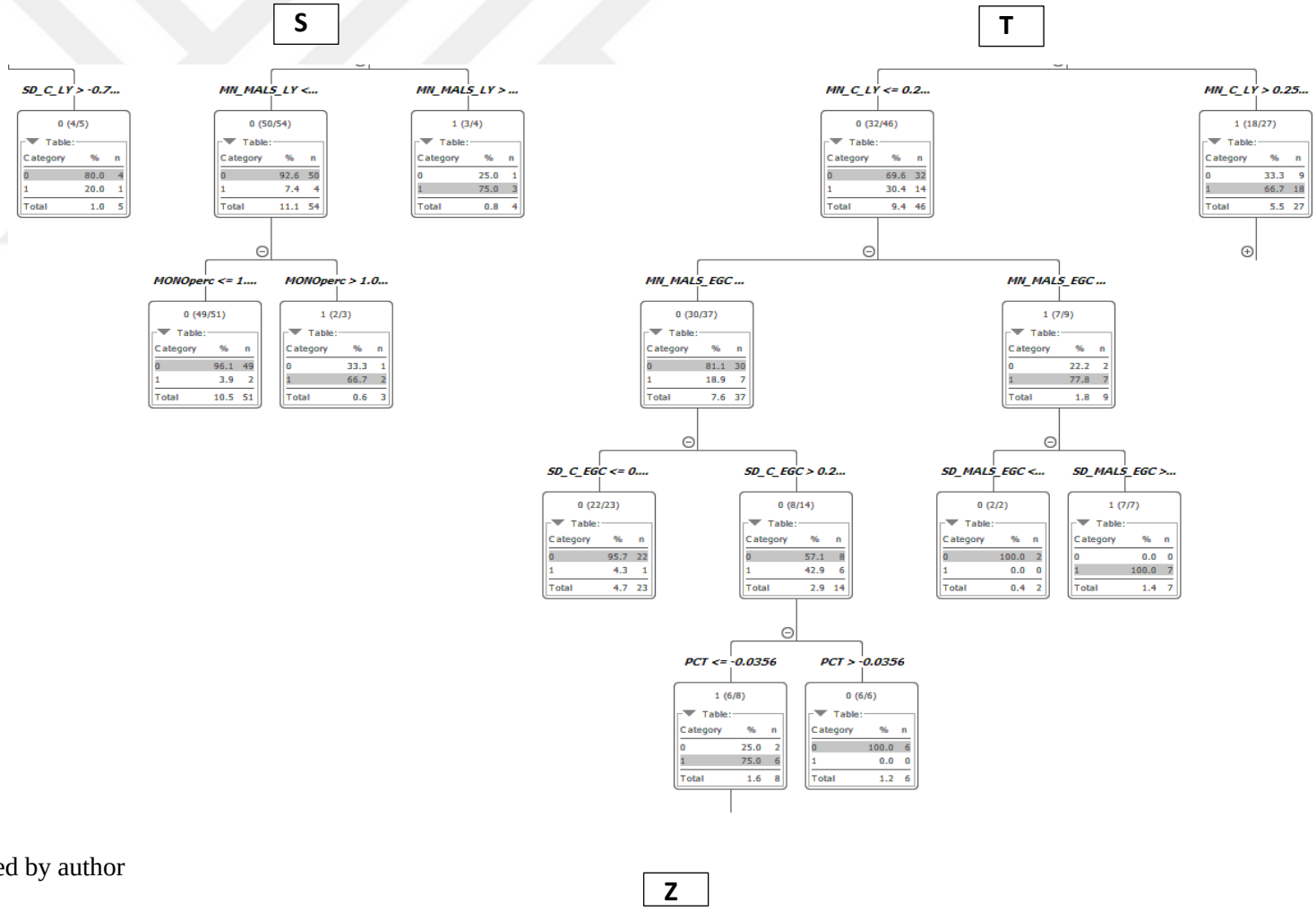
Source: Created by author

Figure 11: Decision Tree Analysis Results



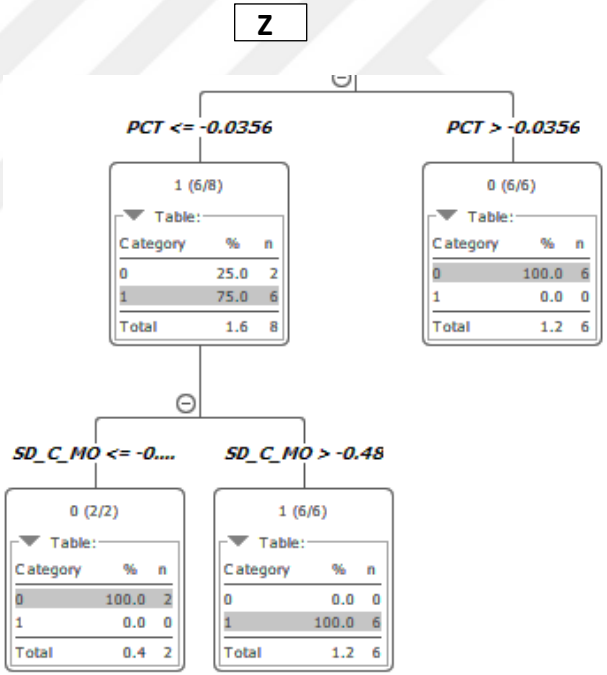
Source: Created by author

Figure 12: Decision Tree Analysis Results



Source: Created by author

Figure 13: Decision Tree Analysis Results



Source: Created by author

After selecting features and cleaning the dataset, logistic regression applied to our dataset. This model assumes a linear connection between the independent factors and the dependent variable's log-odds. The logistic function, often known as the sigmoid function, is used to convert log-odds to probabilities.

The logistic function, denoted as σ(z), Peng, So, Stage, and St. John (2002) describe it as:

$$\sigma(z) = 1 / (1 + e^{(-z)})$$

where z represents the linear combination of the independent variables:

$$z = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + ... + \beta_p x_p$$

In the equation above, β_0, β_1, β_2, ..., β_p are the coefficients or weights associated with each independent variable, and x_1, x_2, ..., x_p are the values of the independent variables. The coefficients β_0, β_1, β_2, ..., β_p are estimated using a method called maximum likelihood estimation, which aims to find the values that maximize the likelihood of observing the given data. Once the coefficients are estimated, the logistic regression model may predict the likelihood of the dependent variable falling into a particular class. The anticipated probabilities can be turned into class labels (e.g., 0 or 1) by selecting a threshold, typically 0.5.

Logistic regression may be extended to tackle issues with numerous classes by employing one-vs-rest or softmax regression approaches. In these cases, the model calculates separate binary logistic regressions for each class, and the final prediction is based on the class with the highest probability. One needs to realize that logistic regression is predicated on some assumptions, such as linearity, independence of errors, and absence of multicollinearity. When applying logistic regression to a specific problem, these assumptions should be considered and evaluated (Fleuren et al., 2020).

3.6. MODEL

In the first phase, data should be cleaned and prepared for modeling. In this thesis, two CSV files from different dates were integrated into one CSV file. The entire dataset is then divided into two parts: training and test sets. First, missing values

are filled with median for numeric variables and the most frequent value for string values.

The original data is then normalized using the Z-score normalization approach, except for the sepsis label. If the feature ranges diverge, the data should be normalized before being included in the training model. (Schellenberger et al., 2019).

After normalizing data, using the partitioning node, the given input table is divided into two different partitions row-wise. The first partition is termed train data and the second is test data. The results of these two partitions are made available through the partitioning node's two output ports. The input table from normalizer is split into two partitions, one is training datasets with 238 rows and the rest is used for testing data in two output tables.

Logistic Regression has been used after preparing data for modeling. There is a Logistic regression learner node in Knime which the inputs are tables from partitioning nodes with. The model includes 8 features and patients who have sepsis are targeted in the model.

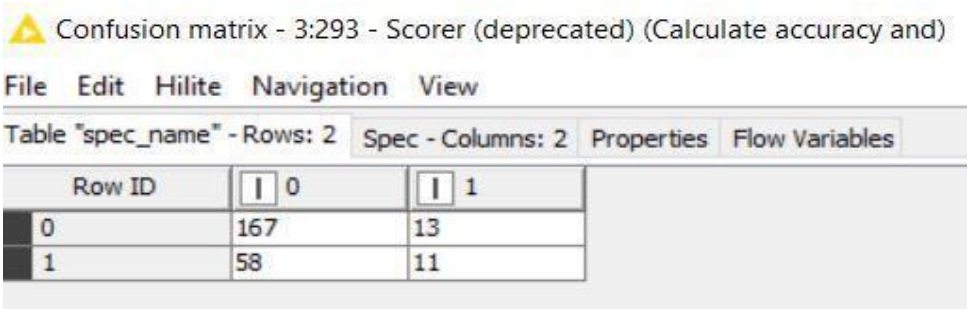
To predict the logistic regression model, link the predictor node to the learner node using test data. The model can only be implemented if the chosen modeling characteristics are present in the test data, which contains the columns needed by the learner model. The forecast for each row is appended to the input table by this node.

To see the confusion matrix there is another node, Scorer. This node examines two columns based on the attribute value pairs in each column and generates a confusion matrix that shows how many rows of each attribute and classification match. This matrix may also be highlighted to reveal the underlying rows. The dialog allows us to compare two columns; the first column's values are represented in the confusion matrix's rows, while the second column's values are represented in the confusion matrix's columns, as shown in Figure 15.

The result of the node is the confusion matrix, which shows how many correct matches are in each cell. Furthermore, the second output gives accuracy data such as True-Positives, False-Positives, True-Negatives, False-Negatives, Recall, Precision, Sensitivity, Specificity, F-measure, and overall accuracy and Cohen's kappa.

After applying the logistic regression model and from the scorer node we could see the confusion matrix as shown in Figure 14.

Figure 14: Confusion Matrix Outcome of Sepsis Dataset



Row ID	0	1
0	167	13
1	58	11

Source: Created by author

The model correctly predicted 167 samples as the positive class. These are instances where the model's prediction and the actual ground truth label were positive. While the model incorrectly predicted 13 samples as the positive class when they belong to the negative class. In other words, the model predicted positive, but the actual ground truth label was negative.

Our model incorrectly predicted 58 samples as the negative class when they belong to the positive class. The model predicted negative in this case, but the actual ground truth label was positive. If we increase the size of our datasets for patients with Sepsis, we might get a better result, and our model will perform better. The True Negative, model correctly predicted 11 samples as the negative class.

The model had an accuracy of 0.71 and a specificity of 0.92, but it did not have a high sensitivity. As a result, the model failed to discover enough true positives. Despite using random forest and examining the confusion matrix to see whether the number of true positives improved, this model performed the best in accuracy and sensitivity.

Python has also been used for Keras deep learning model; however, the result was not as satisfying as Knime. Keras results compared to logistic regression in Knime.

3.7. PREDICTION USING MACHINE LEARNING ALGORITHMS USING PYTHON

Deep learning models can understand complicated patterns and connections in large datasets, enabling them to make accurate predictions. In this study, we used Keras deep learning model, and almost a hundred models have been tried to find the best result out of our dataset. Google Colab is an online platform provided by Google used for writing, executing, and sharing Python codes in the Jupyter notebook environment. After importing the data and libraries that were earlier mentioned in this study, Keras model is used, and data partitioning is the same with Knime.

3.8. MODEL ARCHITECTURE

The code begins by importing the necessary dependencies, including the clear_session function from the Keras backend, optimizers, and callbacks. The clear_session function is called to clear any existing models or layers from previous runs. Then, The Sequential model is created, which represents a linear stack of layers. The first hidden layer is added to the model using the Dense function with 45 units and the ReLU activation function. The input shape is specified as (45), indicating that the input data has 45 features.

```
1 model = Sequential()
2 model.add(Dense(12, input_shape=(45,), activation='relu'))
3 model.add(Dense(12, activation='relu'))
```

Batch Normalization is applied after the first hidden layer. Also, Dropout regularization with a rate of 0.1 is added using Dropout to prevent overfitting. The final output layer is added with 1 unit and the sigmoid activation function to obtain the sepsis prediction probability. The Adam optimizer with a learning rate of 0.001 and decay is defined, and the learning rate schedule is defined using exponential decay, which exponentially decays the learning rate over time.

```
4 model.add(Dense(1, activation='sigmoid'))
5 model.compile(loss='binary_crossentropy',optimizer='adam',metrics=
['accuracy'])
```

```
6  model.fit(X_train, y_train, epochs=5000, batch_size=100)
```

The model is created with the binary cross-entropy loss function and numerous evaluation measures for Receiver Operating Characteristic (ROC) curve analysis, including accuracy, recall, precision, F1 score, and AUC (Area Under the Curve).

In the Early Stopping and Model Checkpoint section, the code defines an early_stop callback using MultiMetricEarlyStopping, which monitors multiple metrics (accuracy, precision, recall, and loss) and stops training if no improvement is seen within a specified period. In this case, we specified thresholds so that whenever accuracy decreased to 0.5, we stopped the model**.** Additionally, a checkpoint callback is defined using MultipliedMetricsCheckpoint. It saves the model's weights only when any of the multiplied metrics (val_f1_score, val_accuracy, val_recall, val_precision, val_auc) improve.

Lastly, training the model: the fit function is called to train the model. The X and y represent the input features and sepsis labels, respectively. Our model is trained for 50000 epochs with a batch size of 32. The validation data (X_val, y_val) is provided to evaluate the model's performance during training.

3.9. FINDINGS

3.9.1 Decision Tree Result’s

Decision tree has used for selecting most relevant features, our model uses 46 variables to make predictions.12 variables has been selected in our model including: EOS, MDW, MN_C_LY, MN_ MALS_EGC, MN_V_EGC, MN_V_NE, SD_C_NE, SD_MALS_NE, SD_V_EGC, SD_V_LY, SD_V_MO,SD_V_NE.these are most relevant from first node which is EOS and then splitted leaf nodes as its illustrated in Table 5.

3.9.2. Logistic Regression Result's

Table 5: Logistic regression with Knime result's

Predicting Sepsis /non-Sepsis	Recall	Precision	Sensitivity	Specificity	F-means
0	0.92	0.74	0.74	0.15	0.82
1	0.15	0.45	0.45	0.92	0.23

Source: Created by author

The table shows the recall of 0.92 or 92% - this means that 92% of the actual positive instances for class "0" were correctly identified by the model, while only 15% of the actual positive instances for class "1" were correctly identified by the model. Precision is 0.74 or 74%. Only 15% of the actual negative instances for class "0" were correctly classified as negative by the model. The F1 score considers precision and recall, and provides a combined performance measure. An F1-score of 0.82 suggests a reasonably balanced model performance for class zero, while the F1-score for class "1" is relatively low, indicating poorer performance than class "0."

In summary, the model seems to perform well for class "0" with high recall and decent precision. However, the performance for class "1" is not as good, with low recall and precision.

3.9.3. Deep Learning Result's

Based on the provided metrics, the model results are as follows: for class 0 (no sepsis), the precision is 0.68, indicating that out of all the samples predicted as class 0, 68% of them are correct. The precision for class 1 (sepsis) is 0.64, which means that 64% of the samples predicted as class 1 are right. whereas the recall for class 0 (no sepsis) is 0.16, meaning that only 16% of the actual class 0 samples were properly recognized. The recall for class 1 (sepsis) is 0.95, indicating that 95% of the actual class 1 samples were correctly identified as class 1. The F1 score for non-sepsis groups is 0.26. Even measurement of precision and recall can be obtained through the F1-

score. The F1-score for class 1 (sepsis), is 0.77, indicating a high level of the model's overall accuracy is 0.65, which means that the model successfully classified 65% of the samples in the dataset. For precision, recall, and F1 score, the macro average is 0.66, 0.56, and 0.51, respectively.

3.9.1. Python Findings

We tested 84 models in Python using Keras deep learning, with indicators like Accuracy, Recall, Precision, and F1score being used to evaluate our model's performance. The number of layers, epochs, and other factors have altered in each model. early Stopping training occurs when a monitored parameter, in this example 0.62 accuracy, ceases to improve. The model's weights are then saved at intervals or depending on certain situations by Model Checkpoint. As a result, the lowest accuracy is 0.62 and the highest accuracy is 0.66. We can consider and select the model depending on our preferences for assessment measures. When we compare the models, we can see that their accuracies and error rates are generally stable, hovering around 0.62 and 0.38, respectively. All models have the same accuracy of 0.62, meaning they correctly predict the target class approximately 62% of the time. The error rate is the inverse of accuracy, reflecting how frequently the model predicts incorrectly. All models have an error rate of 0.38, meaning that they are wrong about 38% of the time.in terms of precision model 6 has the highest rate of 77% among other models which indicate make fewer false positive predictions. Model 83,84 reached the highest recall of 0.72 It represents the model's ability to find all positive instances. There are differences amongst the models in terms of accuracy, recall, F1-score, and sensitivity. Some models, such as the Model 3, Model 4, and Model 5, routinely outperform others in terms of accuracy, recall, F1-score, and sensitivity. This suggests that they are better at generating accurate positive predictions and identifying real positive situations. Some models, such as the Model 6 and Model 18, have lower accuracy and recall values. These models have a greater rate of false positives or sensitivity, which implies they may overlook some positive examples. If True positive is important to us, we should select the model with the highest Recall, which is 0.72 among our 84 models. In nutshell, the metrics can be used to assess the model's capacity to distinguish

between sepsis and non-sepsis patients. We need more real data to enhance the model's results, which will be addressed in future studies.

Table 6: Model's Evaluation Result

Model	Error Rate	Accuracy	Precision	Recall	F1-Score	Sensitivity
1	0.38	0.62	0.75	0.58	0.65	0.57
2	0.38	0.62	0.72	0.61	0.66	0.60
3	0.38	0.62	0.70	0.67	0.68	0.66
4	0.37	0.63	0.70	0.68	0.69	0.68
5	0.36	0.64	0.71	0.70	0.70	0.69
6	0.38	0.62	0.77	0.53	0.63	0.53
7	0.38	0.62	0.76	0.56	0.64	0.56
8	0.38	0.62	0.76	0.57	0.65	0.56
9	0.38	0.62	0.75	0.57	0.65	0.57
10	0.38	0.62	0.74	0.59	0.66	0.59
11	0.38	0.62	0.73	0.61	0.66	0.60
12	0.38	0.62	0.70	0.66	0.68	0.66
13	0.38	0.62	0.69	0.68	0.68	0.67
14	0.37	0.63	0.70	0.69	0.69	0.68
15	0.38	0.62	0.69	0.69	0.69	0.68
16	0.38	0.62	0.71	0.64	0.67	0.64
17	0.38	0.62	0.70	0.67	0.68	0.66
18	0.38	0.62	**0.78**	0.52	0.63	0.52
19	0.38	0.62	0.77	0.53	0.63	0.52
20	0.37	0.63	0.76	0.57	0.65	0.57
21	0.38	0.62	0.74	0.58	0.65	0.57
22	0.38	0.62	0.74	0.58	0.65	0.57
23	0.38	0.62	0.74	0.58	0.65	0.58
24	0.36	0.64	0.76	0.59	0.66	0.58
25	0.38	0.62	0.73	0.59	0.65	0.58
26	0.36	0.64	0.77	0.59	0.67	0.59
27	0.37	0.63	0.74	0.59	0.66	0.59
28	0.37	0.63	0.74	0.60	0.66	0.59
29	0.38	0.62	0.74	0.60	0.66	0.59
30	0.38	0.62	0.73	0.60	0.66	0.59
31	0.36	0.64	0.76	0.60	0.67	0.60
32	0.38	0.62	0.73	0.60	0.66	0.60
33	0.38	0.62	0.72	0.61	0.66	0.60
34	0.37	0.63	0.74	0.61	0.67	0.60

35	0.38	0.62	0.73	0.61	0.66	0.60
36	0.36	0.64	0.75	0.61	0.67	0.61
37	0.38	0.62	0.73	0.61	0.67	0.61
38	0.44	0.56	0.60	0.62	0.61	0.62
39	0.38	0.62	0.73	0.62	0.67	0.62
40	0.38	0.62	0.72	0.62	0.66	0.62
41	0.38	0.62	0.72	0.63	0.67	0.62
42	0.38	0.62	0.71	0.63	0.67	0.62
43	0.36	0.64	0.74	0.63	0.68	0.63
44	0.38	0.62	0.72	0.63	0.67	0.62
45	**0.34**	**0.66**	0.71	0.71	**0.71**	0.71
46	0.37	0.63	0.72	0.64	0.68	0.64
47	0.38	0.62	0.71	0.64	0.68	0.64
48	0.38	0.62	0.71	0.64	0.67	0.64
49	0.38	0.62	0.71	0.64	0.67	0.64
50	0.38	0.62	0.71	0.65	0.68	0.64
51	0.38	0.62	0.70	0.65	0.67	0.64
52	0.38	0.62	0.71	0.65	0.68	0.64
53	0.38	0.62	0.70	0.65	0.67	0.64
54	0.38	0.62	0.71	0.65	0.68	0.64
55	0.38	0.62	0.70	0.65	0.67	0.64
56	0.38	0.62	0.71	0.65	0.68	0.65
57	0.38	0.62	0.70	0.65	0.68	0.65
58	0.37	0.63	0.72	0.66	0.69	0.65
59	0.38	0.62	0.71	0.66	0.68	0.65
60	0.38	0.62	0.70	0.66	0.68	0.65
61	0.38	0.62	0.70	0.66	0.68	0.65
62	0.38	0.62	0.70	0.66	0.68	0.66
63	0.36	0.64	0.72	0.67	0.69	0.66
64	0.37	0.63	0.71	0.67	0.69	0.66
65	0.37	0.63	0.71	0.67	0.69	0.66
66	0.38	0.62	0.70	0.67	0.68	0.66
67	0.36	0.64	0.72	0.67	0.70	0.67
68	0.38	0.62	0.70	0.67	0.69	0.67
69	0.38	0.62	0.70	0.68	0.69	0.67
70	0.36	0.64	0.71	0.68	0.70	0.67
71	0.38	0.62	0.72	0.65	0.68	0.64
72	0.37	0.63	0.71	0.68	0.70	0.67
73	0.38	0.62	0.69	0.68	0.69	0.68
74	0.36	0.64	0.71	0.68	0.70	0.68
75	0.37	0.63	0.70	0.68	0.69	0.68
76	0.38	0.62	0.69	0.68	0.69	0.68
77	0.37	0.63	0.70	0.69	0.70	0.68
78	0.38	0.62	0.69	0.69	0.69	0.68
79	0.37	0.63	0.70	0.70	0.70	0.69

80	0.36	0.64	0.71	0.70	0.70	0.69
81	0.37	0.63	0.69	0.71	0.70	0.70
82	0.36	0.64	0.70	0.71	**0.71**	0.71
83	0.35	0.65	0.71	**0.72**	**0.71**	**0.72**
84	0.36	0.64	0.70	**0.72**	**0.71**	**0.72**

Source: Created by author

In this table bold values show the best scores of the relevant measurements among all models. Therefore, there is not any model of which all the performance indicators are the best at all.

CONCLUSION

Machine learning (ML) provides a variety of methods for evaluating large patient datasets and detecting links between elements that individuals may not perceive, as well as lowering death rates by allowing for early treatment and mitigation efforts. Even though the deep models' performance indicators suggest a promising level of prediction ability, the complexity of the challenges of multidimensional space makes basic prediction probabilities impossible to utilize or trust. In general, machine learning algorithms are well adapted to detecting early sepsis, and positive results from studies utilizing various models have been published. Using a machine learning system can reduce the number of doctors required to monitor sepsis patients, thereby mitigating a substantial social impact.

The presented models for predicting sepsis offer a range of pros and cons. Models 83 and 84 exhibit high recall, making them adept at capturing true positive cases early, while Model 6 excels in high precision, minimizing false positive predictions. Models 3, 4, and 5 strike a balance between precision and recall with solid F1-scores. However, the trade-offs between precision and recall need careful consideration, potentially leading to missed true positives or increased false positives. While the evaluation provides informed decision-making insights, the plateaued accuracy and error rates around 0.62 and 0.38 respectively signal potential limitations in the dataset and features. The need for increasing real-world data for improved generalizability and resource allocation for different models further shape the practitioner's approach to sepsis prediction.

The key advantage of machine learning models lies in their interpretability, making them suitable for medical applications where understanding the factors driving predictions is crucial. our models can provide practitioners with insights into which variables contribute most to sepsis risk, like aiding clinical decision-making, from our decision tree we realized EOS, MDW, MN_C_LY ,MN_MALS_EGC, MN_V_EGC, MN_V_NE, SD_C_NE, SD_MALS_NE, SD_V_EGC, SD_V_LY, SD_V_MO,SD_V_NE are the most important variable that have effect on sepsis detection. Additionally, they are relatively simple to implement and computationally efficient, making them viable even with limited resources. However, regression

models might struggle to capture the complex and nonlinear relationships inherent in medical data, potentially leading to suboptimal predictive performance. They also heavily rely on well-engineered features, and if critical features are not identified, the model's predictive capabilities could be compromised. we also used Keras deep learning model which offer a promising avenue for sepsis prediction, particularly due to its ability to capture intricate patterns and relationships within our complex medical data. Deep learning excels at automatic feature learning, which reduces the dependence on extensive feature engineering and could uncover previously unseen predictive factors. This is especially valuable in medical contexts where the relationships between variables might not be well understood. Additionally, deep learning models have the potential for high performance, achieving state-of-the-art results when properly trained with ample data. However, these advantages come with certain drawbacks. Deep learning models demand substantial amounts of data for effective training and significant computational resources, which might be a limitation for some practitioners. Furthermore, the black-box nature of deep neural networks can hinder their interpretability, potentially raising concerns in medical settings where understanding the rationale behind predictions is vital. Careful validation and regularization strategies are also required to mitigate the risk of overfitting, particularly when working with smaller datasets.

After creating our initial project strategy and goals, we assessed how we might best align with the values and ideals behind responsible innovation. We may use prior studies to guarantee that our concept is original. In terms of dataset size and restrictions, the first research is nearly identical to Chicco and Luca (Leiva et al., 2019) used machine learning to predict survival, septic shock, and numerical SOFA values from 364 patient electronic health data. However, because their labels were significantly imbalanced, only a multilayer perceptron for guessing if someone survives was able to correctly predict more than 70% of the times when someone survived and more than 70% of the times when someone did not survive. They then used a model based on random forests to anticipate the relevance of septic shock traits and compared the findings to statistical approaches including the Pearson correlation coefficient, student's t-test, and p-values. They discovered that the random forest revealed several critical characteristics that represent contemporary scientific

achievements that statistical approaches could not. Their research, however, was limited by their poor model performance and small dataset (Guan et al, 2021b). The approach worked well in a range of settings, age groups, genders, and races. We did a study called SHAP analysis and found that the latest data on heart rate and breathing rate have a big impact on predicting the future, even when looking far ahead in time, this was accomplished using a technique known as game-theoretic Shapley approach. One of the important findings is that the model may detect early indicators of sepsis before the SIRS criteria can be applied to establish a diagnosis.

There are several facets of intensivist practice that might benefit from AI technologies, beginning with the diagnosis/monitoring/prediction. In our study we used supervised machine learning with primary data from Intensive care unit of Dokuz Eylul Hospital. We used a decision tree as a classification and feature selection approach, logistic regression, and Keras deep learning to model our dataset in this article. We explored what features in each model were considered relevant after evaluating the performance of the 488 records in our dataset. After using decision tree we prioritized and selected the most relevant features for our logistic regression model as EOS, MDW, MN_C_LY, MN_ MALS_EGC, MN_V_EGC, MN_V_NE, SD_C_NE,SD_MALS_NE, SD_V_EGC, SD_V_LY, SD_V_MO,SD_V_NE. The logistic regression model performed the best across nearly all assessment measures, with an overall precision of 45% and an accuracy of 71%. It means 45% true positive predictions out of all positive predictions made by our model. After testing over 80 models in Keras deep learning, we achieved 65% precision and accuracy. As a result, the statistic of average precision is critical. However, the accuracy of our models in a medical environment, it would be insufficient since one out of every five cases would be misclassified.

The prediction of sepsis using machine learning presents transformative managerial impacts in healthcare institutions. By enabling timely interventions, it enhances patient outcomes, reduces treatment costs, and optimizes resource allocation. Streamlined workflows driven by predictive insights lead to operational efficiency gains, while data-driven decision-making informs strategic planning and process refinement. Collaboration among interdisciplinary teams is fostered, and risk management is fortified, bolstering patient safety and liability mitigation.

Furthermore, this innovative approach showcases the institution's technological prowess, compliance with regulations, and commitment to quality care, thereby positioning it as a leader in healthcare delivery.

The social impacts of early sepsis prediction through machine learning are profound, encompassing improved patient well-being and reduced mortality rates, promoting healthcare equity by ensuring equal access to timely interventions, and enhancing public health management through targeted interventions and disease prevention strategies. These impacts extend to economic savings by alleviating the burden on healthcare systems, reducing the emotional and financial strain on families, fostering community resilience, and driving technological progress in healthcare, thus contributing to a healthier, more equitable, and technologically advanced society.

For executives and policymakers, embracing the implementation of machine learning for Sepsis prediction requires prioritizing data security, collaborating closely with healthcare professionals, investing in necessary infrastructure, and promoting interdisciplinary training. Developing clear clinical guidelines, addressing ethical considerations, and fostering a culture of innovation are vital. Continuously monitoring outcomes, engaging with regulatory bodies, communicating transparently, and staying informed about technological advancements will collectively ensure the responsible integration of machine learning, positively impacting patient care, healthcare systems, and society at large.

The methods and insights discussed in this thesis for early sepsis prediction using machine learning logistic regression model, decision tree and Keras deep learning can serve as a foundational framework for application in various other areas of healthcare. The principles of timely intervention, data-driven decision-making, resource optimization, and collaboration between domain experts and data scientists can be extended to predict and manage a wide range of medical conditions. By adapting the strategies outlined here to diverse healthcare challenges, practitioners and researchers can harness the potential of machine learning to improve patient outcomes, enhance healthcare delivery, and drive innovation across the entire healthcare landscape.

There are certain limitations of the present study, Unfortunately, we could not analyze each patient separately because the dataset was too small. The approach we

applied converged to the point where all patients had the same features. We are certain that the small size of dataset was a significant restriction, limiting how effectively the models generalized to unseen data.

Firstly, we would like to have more data to explore modifying our neural network model to improve performance, as indicated by the feature significance outputs. Attempting to obtain a better and more reliable grasp of important aspects and potential new features. Furthermore, Histogram-based gradient-boosted classifiers, such as lightGBM, have support for missing values, which may improve performance given the significant number of missing values in our dataset. To produce more consistent outcomes, further adjustments may be necessary. For instance, adding additional parameters based on those that are significantly linked to sepsis may produce better outcomes. Furthermore, it is crucial to remember that sepsis is a new growing area with many constraints and no standard approach to performing such research.

REFERENCES

Atymtayeva, L., Kozhakhmet, K. and Bortsova, G. (2014). Building a knowledge base for expert system in information security. *In Soft computing in artificial intelligence* (pp. 57-76). Springer International Publishing.

Barlas, P., Lanning, I. and Heavey, C. (2015). A survey of open source data science tools. *International Journal of Intelligent Computing and Cybernetics*, 8(3): 232-261.

Bone, R. C., Balk, R. A., Cerra, F. B., Dellinger, R. P., Fein, A. M., Knaus, W. A., ... & Sibbald, W. J. (1992). Definitions for sepsis and organ failure and guidelines for the use of innovative therapies in sepsis. *Chest*, 101(6): 1644-1655.

Bowes, D., Hall, T. and Gray, D. (2012, September). Comparing the performance of fault prediction models which report multiple performance measures: recomputing the confusion matrix. *In Proceedings of the 8th international conference on predictive models in software engineering* (pp. 109-118).

Calvano, S. E., Xiao, W., Richards, D. R., Felciano, R. M., Baker, H. V., Cho, R. J., ... & Inflammation and Host Response to Injury Large Scale Collaborative Research Program. (2005). A network-based analysis of systemic inflammation in humans. *Nature*, 437(7061): 1032-1037.

Cios, K. J., Swiniarski, R. W., Pedrycz, W., Kurgan, L. A., Cios, K. J., Swiniarski, R. W., ... & Kurgan, L. A. (2007). Unsupervised learning: association rules. *Data Mining: A Knowledge Discovery Approach*, 289-306.

de Grooth, H. J., Postema, J., Loer, S. A., Parienti, J. J., Oudemans-van Straaten, H. M. and Girbes, A. R. (2018). Unexplained mortality differences between septic shock trials: a systematic analysis of population characteristics and control-group mortality rates. *Intensive care medicine*, 44: 311-322.

Ferrer, R., Martin-Loeches, I., Phillips, G., Osborn, T. M., Townsend, S., Dellinger, R. P., ... & Levy, M. M. (2014). Empiric antibiotic treatment reduces mortality in severe sepsis and septic shock from the first hour: results from a guideline-based performance improvement program. *Critical care Medicine*, 42(8): 1749-1755.

Fleuren, L. M., Klausch, T. L., Zwager, C. L., Schoonmade, L. J., Guo, T., Roggeveen, L. F., ... & Elbers, P. W. (2020). Machine learning for the prediction of sepsis: a systematic review and meta-analysis of diagnostic test accuracy. *Intensive care medicine*, 46, 383-400.

Gambus, P. and Shafer, S. L. (2018). Artificial intelligence for everyone. *Anesthesiology*, 128(3): 431-433.

Greco, M., Caruso, P. F. and Cecconi, M. (2020, November). Artificial intelligence in the intensive care unit. In Seminars in Respiratory and Critical Care Medicine (Vol. 42, No. 01, pp. 002-009). 333 Seventh Avenue, 18th Floor, New York, NY 10001, USA: Thieme Medical Publishers, Inc.

Gyawali, B., Ramakrishna, K. and Dhamoon, A. S. (2019). Sepsis: The evolution in definition, pathophysiology, and management. SAGE open medicine, 7, 2050312119835043.

Haghighat, E. and Juanes, R. (2021). SciANN: A Keras/TensorFlow wrapper for scientific computations and physics-informed deep learning using artificial neural networks. *Computer Methods in Applied Mechanics and Engineering*, 373: 113552.

Kumar, A., Roberts, D., Wood, K. E., Light, B., Parrillo, J. E., Sharma, S., ... & Cheang, M. (2006). Duration of hypotension before initiation of effective antimicrobial therapy is the critical determinant of survival in human septic shock. *Critical care medicine*, 34(6): 1589-1596.

Lobo, J. M., Jiménez-Valverde, A. and Real, R. (2008). AUC: a misleading measure of the performance of predictive distribution models. *Global ecology and Biogeography*, 17(2): 145-151.

Marshall, J. C. (2003). Such stuff as dreams are made on: mediator-directed therapy in sepsis. *Nature reviews Drug discovery*, 2(5), 391-405.

Montgomery, D. C., Peck, E. A. and Vining, G. G. (2021). Introduction to linear Regression Analysis. John Wiley & Sons.

Provost, F. and Fawcett, T. (2013). Data science and its relationship to big data and data-driven decision making. *Big Data*. 1(1): 51-59.

Rachoin, J. S. and Dellinger, R. P. (2014). Timing of Norepinephrine in Septic Patients: *NOT too little too late. Critical Care*, 18: 1-2.

Reddy, Y. C. A. P., Viswanath, P. and Reddy, B. E. (2018). Semi-supervised learning: A brief Review. *Int. J. Eng. Technol*, 7(1.8): 81.

Sakr, Y., Jaschinski, U., Wittebole, X., Szakmany, T., Lipman, J., Ñamendys-Silva, S. A., ... & ICON Investigators. (2018, December). Sepsis in intensive care unit patients: worldwide data from the intensive care over nations audit. *In Open forum infectious diseases* (Vol. 5, No. 12, p. ofy313). US: Oxford University Press.

Safavian, S. and Landgrebe, D. A. (1991). A survey of decision tree classifier methodology. IEEE Transactions on Systems, Man, and Cybernetics, 21(3), 660–674.
Sands, K. E., Bates, D. W., Lanken, P. N., Graman, P. S., Hibberd, P. L., Kahn, K. L., ... & Platt, R. (1997). *Epidemiology of sepsis syndrome in 8 academic medical centers. Jama*, 278(3): 234-240.

Schmidhuber, J. (2015). Deep learning in neural networks: An overview. *Neural Networks*, 61: 85-117.

Seo, B. J., Nissly, C., Angeli, G., Ellerbroek, B., Nelson, J., Sigrist, N. and Troy, M. (2009). Analysis of normalized point source sensitivity as a performance metric for large telescopes. *Applied Optics*, 48(31): 5997-6007.

Simon, P. (2013). Too big to ignore: the business case for big data (Vol. 72). John Wiley & Sons.

Singer, M., Deutschman, C. S., Seymour, C. W., Shankar-Hari, M., Annane, D., Bauer, M., ... & Angus, D. C. (2016). The Third International Consensus Definitions For Sepsis And Septic Shock (Sepsis-3). Jama, 315(8), 801-810.

Sutton, R. S. ve Barto, A. G. (2018). *Reinforcement learning: An introduction*. MIT press.

Talisa, V. B., Yende, S., Seymour, C. W. and Angus, D. C. (2018). Arguing for Adaptive Clinical Trials in Sepsis. *Frontiers in Immunology*, 9, 1502.

Tharwat, A. (2020). Classification assessment methods. *Applied Computing and Informatics*, 17(1): 168-192.

Toshniwal, A., Mahesh, K. ve Jayashree, R. (2020, October). *Overview of anomaly detection techniques in machine learning*. In 2020 fourth international conference on I-SMAC (IoT in social, mobile, analytics and cloud)(I-SMAC) (pp. 808-815). IEEE.

Van, N. T. and Thinh, T. N. (2017, July). *An Anomaly-Based Network Intrusion Detection System Using Deep Learning*. In 2017 International Conference on System Science and Engineering (ICSSE) (pp. 210-214). IEEE.

Vincent, J. L. (2016). The clinical Challenge of Sepsis Identification and Monitoring. *PLoS medicine*, 13(5), e1002022.

Vincent, J. L., Sakr, Y., Sprung, C. L., Ranieri, V. M., Reinhart, K., Gerlach, H., ... & Sepsis Occurrence in Acutely Ill Patients Investigators. (2006). Sepsis in European intensive care units: results of the SOAP study. *Critical care Medicine*, 34(2): 344-353.

Wickham, H. and Wickham, H. (2016). *Data analysis* (pp. 189-201). Springer International Publishing.

APPENDICES

Appendix 1: Quantitative Description of Dataset with Knime

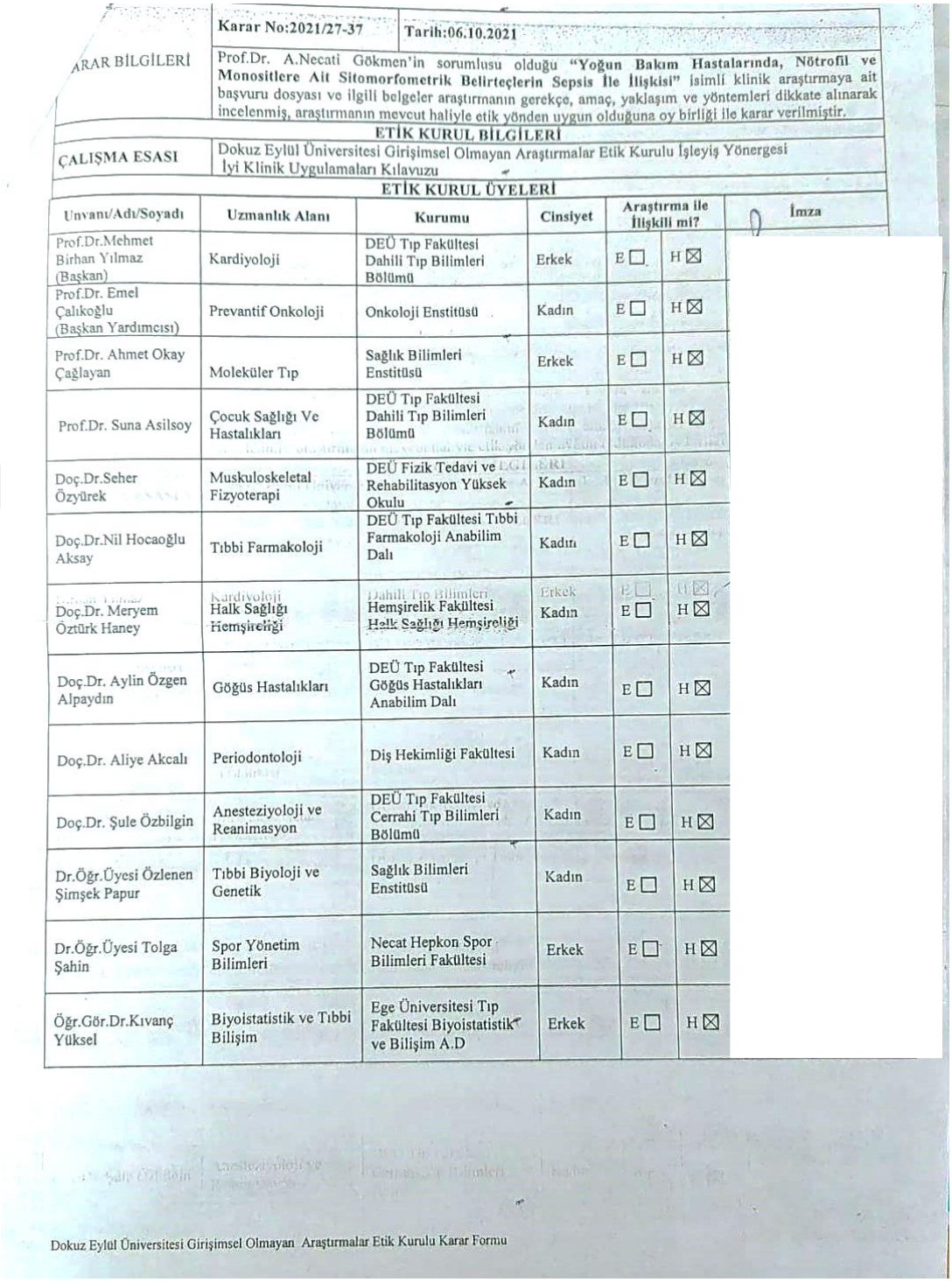
	Min	**Max**	**Mean**	**Std.**	**Variance**	**Skewness**
WBC	0.00	70.20	12.28	8.04	64.59	2.55
RBC	1.68	5.10	3.17	0.70	0.49	0.49
BASO	0.00	0.30	0.03	0.06	0.00	1.60
BASOperc	0.00	1.80	0.38	0.34	0.12	1.57
EOS	0.00	0.90	0.09	0.15	0.02	2.51
EOSperc	0.00	11.30	1.00	1.60	2.56	2.59
HCT	14.70	45.60	27.97	6.27	39.34	0.66
LYM	0.00	66.10	1.42	5.41	29.23	9.82
LYMperc	0.70	94.20	11.25	15.53	241.15	3.76
MCH	21.40	36.90	29.33	1.98	3.93	-0.60
MCHC	29.60	37.60	33.23	1.18	1.40	-0.16
MCV	67.60	102.00	88.25	5.12	26.23	-0.75
MDW	15.31	53.00	26.75	6.15	37.78	1.26
MN_C_EGC	120.00	143.00	131.68	3.27	10.68	-0.27
MN_C_LY	93.00	128.00	114.70	3.83	14.64	-1.13
MN_C_MO	71.00	139.00	121.19	6.40	40.97	-3.56
MN_C_NE	129.00	183.00	141.80	4.07	16.59	3.22
MN_MALS_EGC	100.00	168.00	142.83	9.90	97.94	-0.75
MN_MALS_LY	36.00	114.00	75.91	11.65	135.68	-0.12
MN_MALS_MO	47.00	123.00	90.60	9.76	95.23	-0.63
MN_MALS_NE	101.00	160.00	135.95	10.22	104.48	-0.37
MN_V_EGC	138.00	244.00	178.49	17.91	320.66	0.53
MN_V_LY	52.00	111.00	92.43	7.55	56.96	-0.69
MN_V_MO	155.00	247.00	187.13	13.89	192.86	1.02

MN_V_NE	102.00	203.00	157.89	13.43	180.44	0.14
MONO	0.00	28.90	0.74	1.76	3.09	13.98
MONOperc	0.00	92.40	6.42	9.02	81.38	7.54
MPV	6.80	15.00	9.68	1.48	2.20	0.66
NEU	0.00	47.60	9.98	6.22	38.64	1.35
NEUperc	0.00	98.20	80.94	18.57	345.01	-2.88
PCT	0.01	0.55	0.19	0.12	0.01	0.60
PLT	7.50	706.00	205.22	145.04	21035.35	0.96
RDW	12.20	32.00	16.95	3.28	10.76	1.92
SD_C_EGC	0.58	5.42	2.51	0.73	0.54	0.70
SD_C_LY	3.89	43.50	10.35	4.59	21.05	2.40
SD_C_MO	2.49	49.10	8.60	7.01	49.11	3.18
SD_C_NE	0.00	25.70	5.31	2.07	4.27	5.57
SD_MALS_EGC	1.91	13.60	7.04	1.84	3.38	0.55
SD_MALS_LY	11.34	40.10	19.43	4.82	23.27	1.81
SD_MALS_MO	2.94	36.10	12.41	3.78	14.29	2.47
SD_MALS_NE	0.00	31.30	11.78	3.03	9.16	2.82
SD_V_EGC	9.26	51.00	31.63	7.64	58.41	-0.20
SD_V_LY	9.30	43.50	17.82	4.32	18.70	2.12
SD_V_MO	9.50	50.40	25.70	5.26	27.64	0.78
SD_V_NE	0.00	95.70	23.38	8.26	68.18	5.10

Appendix 2: Parameter References

Parameter	References ARALIK		Male	Female
WBC	4	10,3		
HGB	E/K		13,5-17,5	12-16
RBC	4	5,77		
BASO	0	0,2		
BASO%	0,3	1,5		
EOS	0	0,5		
EOS%	0,9	6		
HCT	E/K		41-53	36-46
LYM	1,3	3,5		
LYM%	19,4	44,9		
MCH	27,2	33,5		
MCHC	32,7	35,6		
MCV	80,7	95,5		
MONO	0,3	0,9		
MONO%	5,1	10,9		
MPV	6,9	10,8		
NEU	2,1	6,1		
NEU%	41	73		
PCT	0,2	0,36		
PLT	156	373		
RDW	11,8	14,3		

Appendix 3: Ethical Committee Approval for Data Usage



KARAR BİLGİLERİ	Karar No:2021/27-37	Tarih:06.10.2021
	Prof.Dr. A.Necati Gökmen'in sorumlusu olduğu **"Yoğun Bakım Hastalarında, Nötrofil ve Monositlere Ait Sitomorfometrik Belirteçlerin Sepsis İle İlişkisi"** isimli klinik araştırmaya ait başvuru dosyası ve ilgili belgeler araştırmanın gerekçe, amaç, yaklaşım ve yöntemleri dikkate alınarak incelenmiş, araştırmanın mevcut haliyle etik yönden uygun olduğuna oy birliği ile karar verilmiştir.	
ETİK KURUL BİLGİLERİ		
ÇALIŞMA ESASI	Dokuz Eylül Üniversitesi Girişimsel Olmayan Araştırmalar Etik Kurulu İşleyiş Yönergesi İyi Klinik Uygulamaları Kılavuzu	

ETİK KURUL ÜYELERİ

Unvanı/Adı/Soyadı	Uzmanlık Alanı	Kurumu	Cinsiyet	Araştırma ile İlişkili mi?		İmza
Prof.Dr.Mehmet Birhan Yılmaz (Başkan)	Kardiyoloji	DEÜ Tıp Fakültesi Dahili Tıp Bilimleri Bölümü	Erkek	E ☐	H ☒	
Prof.Dr. Emel Çalıkoğlu (Başkan Yardımcısı)	Prevantif Onkoloji	Onkoloji Enstitüsü	Kadın	E ☐	H ☒	
Prof.Dr. Ahmet Okay Çağlayan	Moleküler Tıp	Sağlık Bilimleri Enstitüsü	Erkek	E ☐	H ☒	
Prof.Dr. Suna Asilsoy	Çocuk Sağlığı Ve Hastalıkları	DEÜ Tıp Fakültesi Dahili Tıp Bilimleri Bölümü	Kadın	E ☐	H ☒	
Doç.Dr.Seher Özyürek	Muskuloskeletal Fizyoterapi	DEÜ Fizik Tedavi ve Rehabilitasyon Yüksek Okulu	Kadın	E ☐	H ☒	
Doç.Dr.Nil Hocaoğlu Aksay	Tıbbi Farmakoloji	DEÜ Tıp Fakültesi Tıbbi Farmakoloji Anabilim Dalı	Kadın	E ☐	H ☒	
Doç.Dr. Meryem Öztürk Haney	Halk Sağlığı Hemşireliği	Hemşirelik Fakültesi Halk Sağlığı Hemşireliği	Kadın	E ☐	H ☒	
Doç.Dr. Aylin Özgen Alpaydın	Göğüs Hastalıkları	DEÜ Tıp Fakültesi Göğüs Hastalıkları Anabilim Dalı	Kadın	E ☐	H ☒	
Doç.Dr. Aliye Akcalı	Periodontoloji	Diş Hekimliği Fakültesi	Kadın	E ☐	H ☒	
Doç.Dr. Şule Özbilgin	Anesteziyoloji ve Reanimasyon	DEÜ Tıp Fakültesi Cerrahi Tıp Bilimleri Bölümü	Kadın	E ☐	H ☒	
Dr.Öğr.Üyesi Özlenen Şimşek Papur	Tıbbi Biyoloji ve Genetik	Sağlık Bilimleri Enstitüsü	Kadın	E ☐	H ☒	
Dr.Öğr.Üyesi Tolga Şahin	Spor Yönetim Bilimleri	Necat Hepkon Spor Bilimleri Fakültesi	Erkek	E ☐	H ☒	
Öğr.Gör.Dr.Kıvanç Yüksel	Biyoistatistik ve Tıbbi Bilişim	Ege Üniversitesi Tıp Fakültesi Biyoistatistik ve Bilişim A.D	Erkek	E ☐	H ☒	

Dokuz Eylül Üniversitesi Girişimsel Olmayan Araştırmalar Etik Kurulu Karar Formu