# **A Predictive Model for Blood Demand and Supply:** A Case of Obstetric Emergencies at Kenyatta National Hospital

Mutai Kibet Victor

079215

Dissertation Submitted in partial fulfillment of the requirements for the award of the degree of Master of Science in Mobile Telecommunication, Innovation, and Development (Msc.MTID)

Faculty of Information Technology

Strathmore University

Nairobi, Kenya

June 2025

## Declaration and Approval

### Declaration

I declare that this work has not been presented or approved for degree requirements at any other academic institution and remains original. My thesis only includes published works and authored materials when they are correctly cited in the text.



Student Name: Victor Kibet Mutai

Signature; Date: 27th March 2025

### Approval

The dissertation of Victor Kibet Mutai was reviewed and approved for examination by the following:

Dr. Danny Nyatuka

School of Computing and Engineering Sciences

Strathmore University

Sign: ……… ………….. Date: 28th March 2025.

# Abstract

Maternal health is still a very critical concern in Kenya, with obstetric emergencies posing considerable challenges to maternal health and mothers' well-being. One key factor that contributes to maternal mortality is the unavailability of blood during obstetric complications, which is essential for saving lives during emergencies related to childbirth. Unfortunately, many lives have been lost under the circumstances due to a lack of timely information regarding blood requests and donations. These delays in finding blood to manage such situations more effectively cause loss of life. The study aimed to develop a predictive web-based model to monitor blood demand and supply and ensure blood is available to save the lives of mothers and newborn babies during obstetric emergencies using Kenyatta National Hospital (KNH) as a case study. This will be achieved through improved blood bank management, monitoring blood availability and providing timely information to blood stakeholders, minimizing delays in blood donation and supply requests during obstetric emergencies. This will also help reduce maternal and infant mortality cases, thus improving maternal health outcomes. If successful, the proposed solution will go a long way in improving blood donation practices in the healthcare system. The model leveraged modern application development technologies and tools, including Predictive Analytics, Artificial Intelligence, and Machine Learning libraries, to forecast blood demand and supply based on historical data. The research adopts an experimental design to gain in-depth insights into the current practices, challenges, and patterns in blood management. This approach facilitated the identification of critical variables influencing blood demand and supply, leveraging both qualitative and quantitative data to make correct blood demand versus supply predictions. By exploring existing systems, healthcare workflows, and historical data, the study will provide a foundation for constructing an effective predictive model tailored to the unique requirements of obstetric care.

**Keywords**: Obstetric Emergency, Maternal Health, Blood Donation, Predictive Analytics.

## Table of Contents

# List of Figures

## List of Tables

## List of Abbreviations/Acronyms

AI – Artificial Intelligence

KETTA – Kenya Tissue and Transplant Authority

K-Means – K-Means Clustering

KNBTS – Kenya National Blood Transfusion Service

KNH – Kenyatta National Hospital

LSTM – Long Short-Term Memory

ML – Machine Learning

MMR – Maternal Mortality Ratio

NCBI – National Center for Biotechnology Information

PPH – Postpartum Haemorrhage

RBC – Red Blood Cell

RMSE – Root Mean Square Error

SDG – Sustainable Development Goal

UNFPA – United Nations Population Fund

UNICEF – United Nations Children's Fund

WBC – White Blood Cells

WHO – World Health Organization

XGBoost – eXtreme Gradient Boosting

# Definition of Terms

**Long Short-Term Memory** This is a specialized Recurrent Neural Network (RNN) designed to effectively learn from and make predictions on sequential data by capturing long-range dependencies (Karim et al., 2021).

**NumPy (Numerical Python)** This is a foundational library in the Python ecosystem for numerical and scientific computing. It provides support for multi-dimensional arrays and matrices, along with a collection of mathematical functions to operate on these arrays efficiently (Harris et al., 2020).

**Root Mean Square Error (RMSE)** This is a commonly used metric for evaluating the accuracy of regression and forecasting models. It measures the average magnitude of the error between predicted and actual values by taking the square root of the mean of squared errors (Khandelwal, Adhikari, & Verma, 2021).

**Scikit-Learn** This is an open-source machine learning library for Python that provides simple and efficient tools for data mining and data analysis. It supports a range of machine learning algorithms for classification, regression, clustering, dimensionality reduction, and model evaluation. (Zhang & Wang, 2023).

## Acknowledgment

I want to sincerely thank everyone who contributed to my research, most importantly my supervisor, Dr. Danny Nyatuka, for his significant support. I also wish to thank Strathmore University.

# Chapter 1: Introduction

## 1.1 Background of the Study

Maternal health refers to a woman's health during pregnancy, childbirth, and the postpartum period. These stages present various challenges for mothers. Despite existing interventions to improve maternal experiences throughout this process, approximately 287,000 women die every year due to pregnancy and childbirth-related complications across Africa (Meh C, 2021). Each pregnancy and birth presents a unique experience for every mother. It is critical to address these inequalities in health outcomes, particularly in sexual and reproductive health, by ensuring access to quality maternity care (Marsh, 2020). Maternal mortality refers to deaths caused by complications related to pregnancy and childbirth (Clark, 2020). According to the UN inter-agency estimates, the global maternal mortality ratio (MMR) declined from 339 deaths per 100,000 live births in 2000 to 223 deaths per 100,000 live births in 2020, reflecting an annual decrease of 2.1%. While this decline is significant, it falls short of the Sustainable Development Goal (SDG) target of 6.4%, which aims to reduce global maternal mortality to 70 deaths per 100,000 live births by 2030.

The direct causes of maternal injury and death include excessive bleeding, infection, high blood pressure, unsafe abortion, and obstructed labor. Indirect causes include anemia, malaria, and heart disease (Clark, 2020). Blood shortages, therefore, result in preventable maternal deaths (Dafala, 2024). According to the Ministry of Health (MoH), 35% of maternal deaths in Kenya (approximately 2,700 out of 6,500 deaths) were due to bleeding complications during childbirth, while 30% (6,938 deaths) resulted from road accidents in 2023. Kenya's maternal mortality ratio which is the number of mothers who die due to pregnancy-related causes is approximately 355 deaths per 100,000 live births (Mambo C, 2023). Using the current birth rates, this shows that approximately 5,000 women and girls die annually due to pregnancy and childbirth complications. Even with the government's efforts to increase access to skilled birth attendance from 62% to 70%, over 80% of maternal deaths are still linked to poor quality of care. Maternal health remains a critical issue in Kenya, where challenges persist in pregnancy, childbirth, and postpartum care (Christiansen, 2023).

Access to a sufficient and timely blood supply during obstetric emergencies is one of the most crucial factors in maternal survival (Mambo C, 2023). Blood donation programs are vital in ensuring a safe and sustainable blood supply for maternal healthcare. With timely interventions from skilled health professionals in a supportive healthcare environment, most maternal deaths are preventable. Preventing maternal mortality must remain a global priority. However, surviving pregnancy and childbirth alone should not be the sole measure of maternal healthcare success. Reducing maternal harm and disability is equally crucial to improving mothers' overall health and well-being (WHO, 2024).

## 1.2 Statement of the Problem

Obstetric emergencies in Kenya are critically impacted by an insufficient blood supply, particularly during childbirth, pregnancy complications, and the postpartum period. This research underscores a significant gap in both the availability and timely delivery of blood, posing serious risks to maternal and neonatal health. Key contributing factors include logistical inefficiencies, weak blood donor mobilization, and inadequate healthcare infrastructure. A study by Muñoz-Valencia et al. (2023) found that the unavailability of blood for timely transfusion is a major contributor to poor patient outcomes, delayed care, and avoidable hospitalizations. The study also highlighted a significant mismatch between blood availability and patient demand, leading to increased referrals, logistical barriers, and transportation delays. As a result, preventable maternal and neonatal deaths remain common, leaving lasting impacts on families and heightening the risks of malnutrition and infant mortality. Tackling these challenges is vital to improving health outcomes and reinforcing Kenya's healthcare system.

The situation at Kenyatta National Hospital (KNH) mirrors the national crisis, with unique institutional challenges that worsen the problem. As a national referral hospital, KNH experiences high patient volumes, stretching its blood supply chain to its limits. The hospital often faces delays in blood procurement, inadequate storage capacity, and insufficient donor mobilization efforts. Furthermore, the absence of integrated systems for tracking and distributing blood hampers its ability to respond swiftly to obstetric emergencies. These logistical and organizational inefficiencies compromise the hospital's capacity to deliver life-saving care to mothers and newborns, underscoring the urgent need for systemic reforms. Identifying and implementing policy, system, and environmental changes are crucial to enhancing blood availability (Munoz-Valencia et al.,2023).

This research proposes the development of a predictive web-based model to monitor blood demand and supply for effective management of obstetric emergencies to address these challenges at KNH. The model will integrate data analytic tools, digital tools and machine learning algorithms to predict blood demand and supply based on factors such as historical usage patterns, patient demographics, and emergency scenarios. The model will enable real-time tracking of blood stocks, optimize donor recruitment strategies, and streamline blood distribution channels within the hospital. By accurately forecasting demand and improving logistical operations, the study aims to reduce delays, enhance resource utilization, and save lives. The predictive tool will play a pivotal role in ensuring that KNH can proactively manage blood supplies rather than reacting to shortages during emergencies. By predicting demand in advance, the hospital will be better equipped to mobilize donors, adjust blood storage capacity, and ensure the timely delivery of blood for obstetric emergencies.

## 1.3 Objectives of the Study

The main goal of this study was to create a predictive model that compares blood demand with supply. The specific objectives of the research were as follows:

1. To analyze the challenges associated with blood donation during obstetric emergencies in KNH.
2. To review existing methods, mechanisms, and tools used in managing the blood donation request process in KNH,
3. To design and develop a predictive model to monitor blood demand and supply for effective management of obstetric emergencies in the context of KNH,
4. To test the accuracy and effectiveness of the new model.

## 1.4 Research Questions

The research addresses the following questions:

1. What are the key challenges associated with blood donation during obstetric emergencies in KNH?
2. What methods, tools, models, and frameworks are currently used in managing blood donation requests?

3. How can a predictive model be developed to effectively monitor blood-demand and supply to improve the management of obstetric emergencies in KNH?
4. How could the accuracy of the new model be determined?

## 1.5 Justification of the Study

Obstetric emergencies pose a significant threat to maternal health in Kenya and contribute to a high maternal mortality rate. Understanding the challenges associated with blood donation during these emergencies is crucial for improving maternal healthcare outcomes and reducing mortality rates. Technology integration into healthcare management has significantly transformed the landscape of healthcare delivery, enhancing the efficiency, accuracy, and accessibility of medical services. Specifically, electronic healthcare management systems (EHCMS) have become essential in modern healthcare, providing streamlined solutions for managing patient data, optimizing resource allocation, and improving overall service delivery. One critical area where these technologies have made a profound impact is in blood bank management systems, where the need for accuracy, timely data access, and secure information sharing is paramount. The research seeks to create a model that predicts blood donation patterns and improves obstetric emergency care in Kenya. The mobile app will provide real-time information on blood availability, enabling healthcare providers to respond promptly to emergencies and ensure timely access to life-saving transfusions. By testing the accuracy of the new app, the study aims to assess its effectiveness in predicting blood donation needs and facilitating the management of obstetric emergencies.

## 1.6 Scope of the Study

This research focused on investigating challenges associated with blood donation during obstetric emergencies, reviewing existing mechanisms for managing the blood donation process, designing and developing a predictive model to monitor blood donation, demand, and supply, and testing the accuracy of the new model in the KNH context. The development of the model involved the integration of artificial intelligence and machine learning algorithms to predict blood donation needs and optimize supply management. The app is designed to be user-friendly and accessible to healthcare providers and blood donors.

## 1.7 Limitations of the Study

The study is limited by factors such as the availability of data, logistical constraints, and the scope of resources. Additionally, the effectiveness of the predictive model may be influenced by factors beyond the scope of this research, such as internet connectivity and user adoption rates. Overall, the scope of this study aimed to provide a comprehensive understanding of blood donation practices during obstetric emergencies in KNH and develop a practical solution to enhance blood donation management and improve maternal healthcare outcomes.

# Chapter 2: Literature Review

## 2.1 Introduction

Blood donation is a critical component in managing obstetric emergencies, where timely access to blood can be the difference between life and death. In Kenya, challenges in ensuring a sufficient and reliable blood supply have significantly impacted maternal health outcomes. This literature review aims to investigate the specific challenges related to blood donation during obstetric emergencies in this context. Furthermore, it will explore and assess existing methods, mechanisms, and tools currently used to manage the blood donation request process globally and locally. The review will also inform the design and development of a predictive web-based mobile application to monitor blood demand and supply, enhancing the management of blood resources in obstetric emergencies. Finally, this study will assess the accuracy and effectiveness of the newly developed application, ensuring it meets the needs of healthcare providers in critical situations.

## 2.2 Obstetric Emergencies in Kenya

Obstetric emergencies are the times in pregnancy or childbirth that require immediate and urgent attention to save the life of the child and the mother. These emergencies are caused by complications such as hemorrhage, obstruction to labor and eclampsia that require immediate medical care and surgical interventions including caesarean section surgery or the need for blood transfusion. These obstetric emergencies have been a challenge in Kenya leading to the country having one of the highest rates of maternal and newborn mortality rates which by far exceeds the global sustainable development goal (SDG) target(Tamma, 2023).

The National Centre for Biotechnology Information (NCBI) states that in 2017, the maternal mortality ratio in Kenya was 342 deaths per 100,000 live births. This number looms above the global target of fewer than 70 deaths per 100,000 live births. It also surpasses the SDG target of 12 deaths per 1,000 live births. Research has shown that one of the key causes of maternal deaths in Kenya is the delay in accessing the right medical interventions, especially in obstetric emergencies like blood transfusions. Some 90 percent of maternal deaths in Kenya are due to delays in accessing timely blood transfusions. This demonstrates the significance of adequate and prompt availability of safe blood components to ensure the avoidable maternal deaths of mothers and newborns (Michuki, 2022).

Blood is increasingly critical in all obstetric emergencies and is a life-saving intervention for all hospitals in Kenya. Transfusion of donated blood is critical to restore volume lost during hemorrhage, rectify coagulation disorders, and offset other maternal maneuvers. Receiving transfusions of blood markedly helps to diminish rates of mortality and morbidity due to obstetric emergencies, resulting in better maternal and child health outcomes through the provision of oxygen and nutrients to tissues and organs (Sekiya, 2023).

In Kenya, obstetric emergencies are a leading cause of maternal and neonatal morbidity and mortality claims, and blood transfusion is one component that could be associated with these high maternal mortality rates given delays in receiving transfusion. Reducing maternal mortality will need efforts to strengthen blood donation programs, blood product availability, and improve the healthcare infrastructure.

## 2.3 The Significance of Safe Blood Supply

The significance of a safe blood supply is crucial in maternal healthcare as it assists in the management of obstetric emergencies leading to good maternal and childbirth outcomes. Maternal health care, including services provided during pregnancy, childbirth, and after delivery, is an integral part of health care providers. In this context, the provision of safe blood products is essential in addressing complications, such as severe hemorrhage and excessive bleeding, which are among the leading causes of maternal morbidity and mortality around the world. Severe bleeding in childbirth emergencies is a prominent threat to maternal health and typically requires timely intervention to eliminate life-threatening surgeries. Blood transfusions are a critical element in this scenario, as they offer life-saving support that can stabilize the mother's condition as much and prevent irreversible damage. Blood product safety is therefore crucial in such situations to prevent transfusion-related risks for the mother and newborn (Timaisina R, 2023).

The safety of the blood supply in maternal healthcare includes strict screening and testing of blood donations, to ensure that potentially contaminated units are identified and excluded. This involves testing for potential blood-borne pathogens like HIV, hepatitis B and C, and syphilis, along with other measures to help prevent transfusion-associated complications such as transfusion-related reactions and alloimmunization. Implementing safety measures helps providers reduce the risk of adverse outcomes and ensure the safety of mothers and newborns. In addition, quality blood supply plays a crucial role in providing a comprehensive treatment strategy to the medical needs of pregnant women, including those with underlying medical diseases or obstetrical complications. Blood transfusions are sometimes necessary in cases of complications, such as anemia, placenta previa, and placental abruption. Healthcare providers may not have easy access to safe blood products, which can easily lead to difficulties in providing timely and effective obstetric care to pregnant women who need it (Timaisina R, 2023).

In addition to managing obstetric emergencies, a safe blood supply supports other maternal health care services like cesarean sections, postpartum hemorrhage management handling pregnancy-related complications. Healthcare systems can have the ability to meet the needs of pregnant women and improve maternal health outcomes by ensuring the availability of safe blood products. The significance of a safe blood supply in maternal healthcare cannot be overstated. It is essential for managing obstetric emergencies, supporting maternal healthcare services, and ensuring positive outcomes for mothers and newborns. By prioritizing the safety and availability of blood products, healthcare providers can uphold the highest standards of care and promote the health and well-being of pregnant women in Kenya (Michuki, 2022).

## 2.4 Significance of Blood Transfusions in Obstetric Care

In maternal healthcare, severe hemorrhaging and excessive bleeding during childbirth emergencies are the major causes of maternal morbidity and mortality. Blood transfusions are therefore pivotal interventions in instances of severe bleeding, often serving as life-saving measures. Childbirth experts assert that blood transfusion constitutes an indispensable component of emergency obstetric care, with an adequate blood supply playing a vital role in reducing maternal mortality rates significantly (Shamoon, 2022). Unlike laboratory-produced medical equipment, blood cannot be synthesized artificially but instead relies solely on voluntary donors for procurement. Additionally, all blood products possess a notably short shelf life, necessitating constant replenishment of the blood supply to meet demand effectively. The intricate logistics of blood donation, processing, and distribution require proper management to ensure the availability of safe and timely transfusions for mothers in need (Hematology, 2024).

For effective emergency response, existing blood reserves must be properly categorized and managed. Blood centers play a crucial role in this process by screening donors, collecting blood, processing it into various components, and distributing these products to healthcare facilities as needed. However, challenges such as blood shortages, particularly in remote or underserved areas, inadequate infrastructure, and limited resources, often hinder the efficient management of blood supply chains (Redcross, 2021). Notably, the demand for blood type O tends to be the highest owing to its extensive use in transfusions, particularly during emergencies. This underscores the importance of targeted donor recruitment efforts and efficient inventory management practices to ensure an adequate supply of blood products, especially those with high demand.

Furthermore, the integration of technology, such as advanced information systems and artificial intelligence (AI), can enhance blood supply management and optimize allocation strategies. AI algorithms can analyze historical data on blood donation patterns, healthcare demand trends, and demographic insights to forecast future needs accurately. By predicting demand versus supply, healthcare providers can proactively adjust procurement and distribution strategies to prevent shortages and ensure equitable access to blood products across regions (Michuki, 2022). In conclusion, recognizing the critical role of blood transfusions in maternal healthcare is essential for improving outcomes and reducing maternal mortality rates. By implementing comprehensive strategies to optimize blood supply management, enhance donor recruitment efforts, and leverage technology for predictive analytics, healthcare systems can effectively address the challenges associated with obstetric emergencies and ensure that mothers receive timely and lifesaving transfusions when needed.

## 2.5 Blood Donation in Kenya and Related Challenges

In Kenya, the supply of blood is inadequate and cannot meet the rising demand due to various factors such as accidents, diseases such as cancer, and pregnancy-related complications (Michuki, 2022). A concerted effort by the Kenyan government is thus needed to motivate and recruit low-risk donors. Statistics from the health ministry indicate that blood donation is mainly concentrated in the county of Nairobi. According to (Munoz-Valencia et al (2023), low blood donation rates, inefficient testing procedures, and other disruptions in the administration of the blood supply chain negatively affect patients in low-resource settings across Sub-Saharan countries, Kenya included. However, there has been a decline in blood collection over the years. Data from the Kenya National Bureau of Statistics (KNBS) shows a decline in collection from 38,808 units of blood in 2006 to 30,840 units in 2009. This continual shortage of blood puts hundreds of lives in jeopardy, highlighting the urgent need for blood donation partners to increase awareness and organize blood donation drives to bolster the country's blood supply (Todd Gersten, 2022). The existing measures put in place to ensure blood availability for obstetric surgeries are insufficient, warranting a focus on improving blood collection and distribution (Schantz-Dunn, 2023). Blood banks often experience inadequate supplies, disproportionately affecting women and children in need of life-saving blood transfusions. There is also a need to raise further awareness about maternal health and the critical role of blood transfusions during maternal crises.

The blood shortages in Kenya directly contribute to 30% of deaths related to mortality and morbidity. More than 85% of these maternal deaths are attributed to postpartum hemorrhage (PPH) and are potentially avoidable with blood transfusions (Malatji R & Madiba S, 2020). Despite significant achievements in maternal and newborn health in Kenya, thousands of babies die within days or weeks after birth, and thousands of women die during pregnancy and delivery each year (Michuki, 2022). These deaths are often caused by the unavailability of blood supply during emergencies. Ensuring the provision of quality care remains a critical challenge, as many healthcare facilities are ill-equipped to prevent or address underlying causes of illness or complications. Furthermore, inadequate resourcing for healthcare exacerbates the gap between demand and supply. However, these deaths can be prevented if measures are implemented to ensure the availability of blood supply for hospitals during emergencies (Malatji R & Madiba S, 2020).

## 2.6 The Prevailing Situation at KNH

Kenyatta National Hospital (KNH) is the largest referral and teaching hospital in Kenya and one of the leading healthcare institutions in East Africa. Established in 1901, it was initially founded as the "Nairobi Hospital" to serve as a general hospital for the growing population in the Nairobi area, which was then a small colonial settlement. Over the years, it evolved from a small dispensary into a fully-fledged medical institution. In 1961, it was renamed the "Kenyatta National Hospital," reflecting its new role as the country's premier medical center (KNH, 2024).

KNH began operating as a national referral hospital to provide specialized medical care to patients from across the country who could not be treated at smaller, local hospitals. It has since grown to offer a wide range of medical services, including emergency care, outpatient services, surgery, diagnostics, maternity services, pediatric care, specialized care in fields such as oncology, cardiology, neurology, and psychiatry, as well as a comprehensive range of diagnostic imaging and laboratory services (KNH, 2024).

The hospital also serves as a teaching and research facility for medical students, nurses, and other healthcare professionals in collaboration with various academic institutions, notably the University of Nairobi's medical school. It plays a crucial role in advancing medical knowledge and healthcare delivery through research, training, and clinical trials.

In terms of its organizational structure, KNH is managed by a Board of Management, which oversees the overall direction of the hospital, while an executive team led by the Chief Executive Officer (CEO) handles day-to-day operations. The hospital's leadership also includes various specialized departments and units, each led by a departmental head, including clinical, administrative, and support services divisions. KNH is also supported by multiple specialist centers within its complex, aimed at offering specialized care for different medical conditions (KNH, 2024).

### 2.6.1 Blood Donation Challenges Experienced at KNH

Kenyatta National Hospital (KNH), the largest referral and teaching hospital in Kenya, faces several challenges in managing its blood donation system. One of the primary issues is the high demand for blood products, driven by the hospital's role in treating critically ill patients, including those with trauma, burns, cancer, and complications from surgery or childbirth. Due to its status as a national referral center, KNH handles a significant volume of emergency cases, many of which require blood transfusions. This results in frequent blood supply shortages, particularly during periods of increased demand, such as trauma incidents or seasonal fluctuations (KNBS, 2024). While KNH benefits from the support of the Kenya National Blood Transfusion Service (KNBTS), the inadequate volume of blood donations remains a persistent challenge. Public engagement and awareness campaigns are essential, but motivating regular blood donations can still face barriers, such as social or cultural hesitations about donating. The hospital also grapples with blood storage and preservation issues: donated blood has a limited shelf life, and logistical challenges in maintaining proper storage conditions and ensuring timely distribution across the hospital further aggravate supply gaps (NCBI, 2024).

Patient congestion at KNH, compounded by the high volume of critical cases, further strains the hospital's ability to meet urgent transfusion needs, often leading to delays in blood delivery that can negatively impact patient outcomes. Additionally, the hospital struggles with resource constraints and funding issues, which affect the management of the blood bank and staffing levels necessary for its operations. Lastly, while KNH organizes regular blood donation drives, it continues to face difficulties in maintaining a steady flow of voluntary donors. At times, the hospital relies on donations from patients' relatives or emergency donors, which is not a sustainable long-term solution. These combined challenges highlight the need for ongoing collaboration between KNH, the government, national blood banks, and the public to ensure a reliable and sufficient blood supply for patients (Redcross, 2021).

### 2.6.2 The Aim and Objectives of KNH

Kenyatta National Hospital (KNH) is committed to providing high-quality healthcare services, promoting medical research, and training healthcare professionals to meet Kenya's evolving health needs. As the country's largest referral and teaching hospital, its mission and objectives are focused on improving access to specialized healthcare and ensuring the delivery of comprehensive, innovative, and affordable care. The hospital's vision is to be a world-class healthcare provider, offering comprehensive medical services that are accessible to all. Its mission is to deliver high-quality, accessible, and affordable healthcare while also contributing to medical education and research for the advancement of public health (KNH, 2024).

KNH's primary goals include the provision of specialized healthcare services across various medical disciplines, particularly for complex cases that cannot be treated at regional hospitals. As Kenya's national referral hospital, KNH plays a crucial role in offering expert consultation, diagnosis, and treatment to patients referred from other healthcare institutions. The hospital also emphasizes training and education, providing world-class training for medical professionals, including doctors, nurses, and allied health workers, in collaboration with universities and academic institutions (KNH, 2024).

Research and innovation are central to KNH's objectives, with a focus on conducting medical research and clinical trials to contribute to the advancement of healthcare knowledge and solutions. A key component of the hospital's mission is ensuring patient-centered care, which emphasizes a holistic approach that respects the dignity, needs, and preferences of each patient, while providing timely and efficient medical services. Furthermore, KNH is dedicated to promoting health equity by enhancing access to quality healthcare for underserved and marginalized populations through outreach programs and affordable care initiatives. The hospital also prioritizes the sustainable management of resources, ensuring the effective utilization of medical staff, equipment, and infrastructure to maintain operational efficiency. Finally, KNH aims to collaborate closely with the Ministry of Health and other healthcare organizations to contribute to national health policies, strategic planning, and the implementation of public health programs aimed at improving overall health outcomes in Kenya (KNH, 2024).

## 2.7 Methods and Mechanisms Used in Managing the Blood Request and Donation Process

Kenya has improved its safe blood supply through technology utilization for better crisis readiness and response capabilities. New technologies including the Internet and mobile phones, have improved the communication channels with donors according to Techpoint Africa (2023). Below, we examine the existing technologies that support blood supply operations. (Techpoint Africa, 2023). The current technologies being used to facilitate blood supply are discussed below.

### 2.7.1 Television

Television has served as a medium for mass communication since it was first introduced. In times of crisis, it has played a crucial role in disseminating vital information to the public, enabling real-time updates. Although this form of communication is inherently one-dimensional, an experienced emergency professional can provide valuable insights to the audience regarding the unfolding events. The advantages of this technology include user-friendliness, wireless capabilities that withstand damage, affordability, and the ability to communicate to vast audience. Furthermore, it is resilient against interference from signal jamming (John Wiley, 2024).

### 2.7.2 Radio

This technology has been present over time, even amid the emergence of various other communication media. Radio has also implemented some transition to digital technology, enhancing transmission methods and incorporating services like broadcasting their IDs, playlists, and program details. This advancement has been development of more efficient radio receivers that are embedded into devices like mobile devices, facilitating portability and mobility. Radio technology is employed to communicate the significance of blood donation to potential donors, as it is easily accessible. Donors receive request broadcasts while engaging in their daily activities. The effectiveness of this medium is demonstrated by the numerous programs created by emergency agencies for radio dissemination (Akol Yam FM, 2024).

#### 2.7.3 Web Applications and Systems

They utilize the existing internet and other technologies by creating a connection between systems and applications and serve information to users. The system's data transfer process uses hypertext methods, and the browsers serve as the system's user interface to run web applications. Web applications provide multiple advantages including universal accessibility alongside simple operation without installation requirements and broad compatibility across different platforms. Web systems deliver support for diverse applications that span office tools, design software, analytics programs and e-commerce solutions. Potential donors use web applications such as search engines to find blood donation centers and any information related to blood donation and maternal health. News broadcasting agencies also use web applications to reach online listeners and viewers. For instance, people will tend to visit news websites when an incident occurs or use the key terms in search engines (Akol Yam FM, 2024).



#### 2.7.4 Mobile Applications

The advancement of mobile technology, which began with devices that only supported voice calls, has led to the incorporation of various features aimed at addressing the increasing demands of users. A significant enhancement has been the wide availability of mobile applications, which empowers users to take charge of their devices and manage limited resources like memory. New devices typically come preloaded with applications, yet users have the option to choose from a diverse application market to find those that meet their interests. Users can easily install or remove any application they wish. Additionally, Sim Tool Kit (STK) applications are provided by the service provider (Hollow, Mitchell, & Gladwell, 2012).

Application markets facilitate a connection between phone users and developers, allowing developers to create solutions tailored to issues faced by users. They harness device capabilities to achieve desired outcomes. Informational applications are a part of this category, characterized by one-to-many communication technology. Nevertheless, many mobile applications are designed to enable two-way communication, reflecting the necessity for user interaction. These applications offer patients and donors a platform for engagement, providing innovative ways to enhance health outcomes (Michuki, 2022).

### 2.7.5 Short Message Service

They utilize the GSM technology to send SMS while enabling the distribution of messages to a vast audience. The capacity to send numerous messages can be significantly improved by integrating an SMS gateway that can convert and read messages from various sources in the media. Most information systems can also effectively deliver SMS notifications to mobile subscribers.

Emergency systems, particularly those designed for early warnings, are established to dispatch SMS alerts to individuals whenever blood bank supplies decrease. This allows authorities to communicate information to residents in impacted regions more efficiently, thereby helping to prevent loss of life. This technology can be used to create immediate disaster alerts in Kenya. The advantages of SMS include secure information, and the capability to customize messages according to specific incidents (Michuki, 2022).

### 2.7.6 Social Networks

Good internet has enhanced social network growth and enables individuals from different locations and time zones to engage with one another. Social media platforms allow for the sharing of information in various forms. The most common social media applications and websites are, Instagram, Facebook, TikTok, and Twitter. This makes social media to be used by various technologies as a way to provide information to the public (Mutegi, K. G.,2016).

Analytics can be used to check on social data related to blood donations. This will aid emergency responses to blood shortages in national blood banks. Calls for donations from blood centers can be communicated using Facebook and Twitter, which have obtained good responses beyond expectations. An example is during major incidents such as road accidents and terrorist attacks in Kenya, social media was used to mobilize the public to donate blood to those affected. This was observed during the Westgate attack, the Sinai fire and the Garissa attack, where volunteers responded well. Social media utilization of for sharing blood donation requests during emergencies faces challenges which are mainly verification and the misuse of information passed (Mutegi, K. G.,2016).

### 2.7.7 Crowdsourcing

Crowdsourcing is essentially taking advantage of what many people can do for a job. This covers information from the public gathered by mobile phone, the internet, or emails, all under communication. In your case, this framework can be considered an open system since anybody can publish data. This data allows to map the incidents. Since information is ephemeral, having data that is timely greatly improves the response process.

Such platforms provide situational awareness about the situation developing during emergencies. Example and project: platforms like Damu-Sasa use crowdsourcing approaches to recruit blood donors. It uses social networks to gather information in a crisis. It depends on public participation nature and thus allows the collection of real-time data. However, it also faces challenges of the secrecy and the reliability of information shared here, these would be key factors to consider to have a workable system (Damu Sasa, 2023).

## 2.8 Existing Application Systems and Tools

While many factors are cited as causes of low volumes of blood in blood banks in Kenya, the lack of a system that can analyze blood usage versus supply by donors adds to this challenge. Several systems have since been developed to aid blood donation, but none has been able to provide real-time analytics on available blood in all the blood banks in Kenya compared to the demand from various hospitals. Additionally, none have addressed the blood needed for emergencies related to maternal healthcare. Existing systems and architectures aimed at aiding blood donation are discussed below.

### 2.8.1 Damu Sasa

Damu Sasa is a blood donation platform that allows potential blood donors to access the information that can save their lives. The app was created after the terror attacks in Nairobi, Kenya, where victims were in dire need of safe blood for transfusion. Many Kenyans willing to donate were unable to do so because of geography. With Damu Sasa, every Kenyan, wherever they live, can participate in life-saving blood donation efforts. Healthcare facilities can communicate with one another in real time to dispense donated blood products (Damu Sasa, 2023). It was incubated in the Ministry of ICT  through the Presidential Digital Talent Programme. It was first tested  at the Kenyatta National Hospital (KNH), the biggest referral hospital in East and Central Africa, effectively complimenting substantial improvements in the hospital's blood service operations. Besides KNH,  Damu Sasa is now used at many other hospitals across the country

The main functionality of Damu Sasa is a web portal where potential donors read about the donation process and blood drives within Kenya. However, it lacks integration with hospitals and blood banks for supply versus demand monitoring. It is also not able to send alerts to donors whenever there is a need to replenish blood banks in the country (Amref Health Africa, 2024). Figure 2.1 below illustrates Damu Sasa dashboard with key features

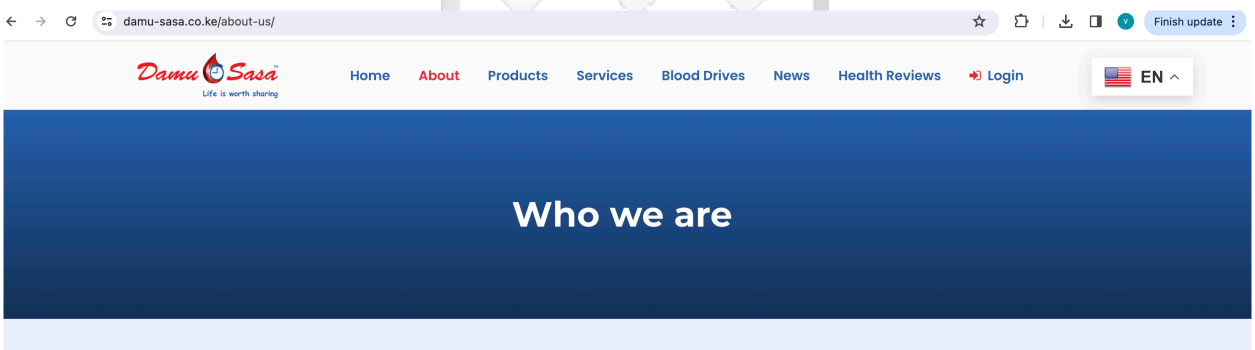



Figure 2.1: Damu Sasa Portal (Amref Health Africa, 2024)



### 2.8.2 Kenya Red Cross (KRCS) App

It is a Kenyan mobile application used to mainly inform users about disasters and emergency occurrences in Kenya. The application is mainly used to inform app users about upcoming blood drives and display content about the blood donation process. The donation functionality is also used to enable app users to make personal donations mainly in the form of cash to the Kenya Red Cross organization in Kenya since it is a nonprofit organization.  It further has a subscription module that allows users to subscribe to ambulance emergency rescue services either annually or monthly at a certain fee. The subscription is also extended to corporate clients among others (Kenya Redcross App, 2024). Figure 2.2 illustrates the current version of the Kenya Red Cross application.

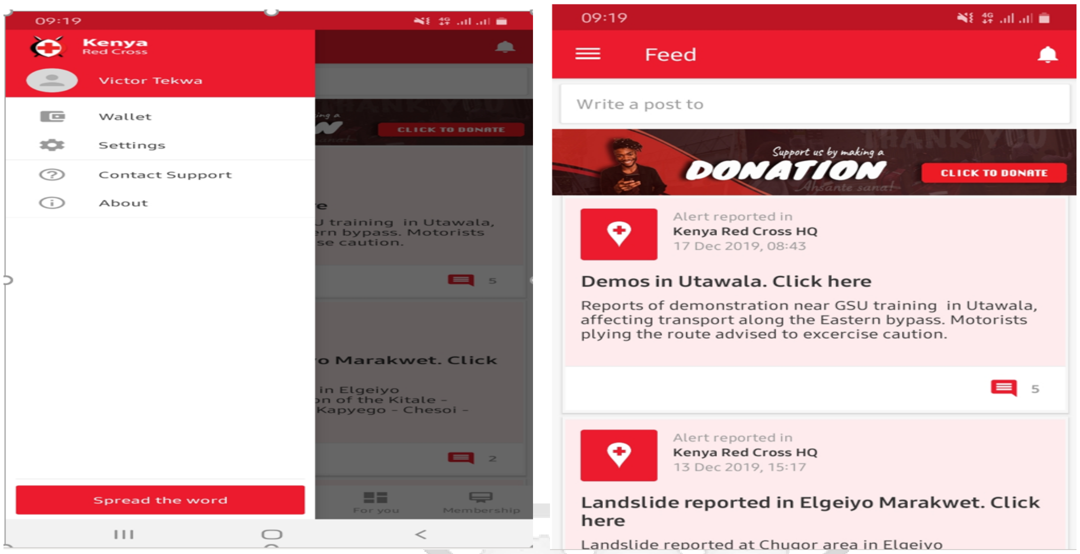



Figure 2.2: Kenya Red Cross Ap (Kenya Red Cross, 2024)

### 2.8.3 Blood Donor App

The mobile application was designed by the American Red Cross to assist its users in finding upcoming organ drives organized by the American Red Cross. It also allows users to schedule appointments if they want to donate organs and enables them to earn rewards from premier collaborating retailers. Users can also create or join a lifesaving team and be able to track its impact nationally. A donor can also follow their blood's journey from donation all through to delivery. All these functionalities therefore, allow the users to easily donate blood and track the activities of donation agencies (American Red Cross, 2024). Figure 2.3 illustrates the American Red Cross donation application used for donation.

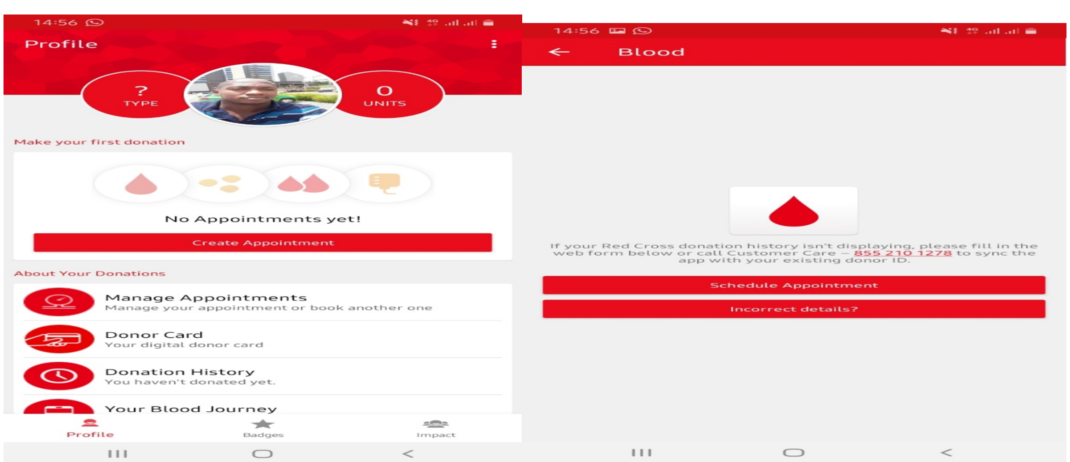



Figure 2.3: Blood Donor App (American Red Cross, 2024)

The main weakness it has is that it only handles blood donation, Secondly, blood drives searched are those only located in America and that have been planned by the American Red Cross team leaving out other countries like Kenya.

## 2.9 Artificial Intelligence, Machine Learning Algorithms & Hybrid AI

Machine learning has become a powerful tool for solving complex problems, but no single algorithm works best for all tasks. That is why a hybrid AI approach is often used, combining multiple machine learning techniques to leverage their strengths. In this case, we integrate Long Short-Term Memory (LSTM) networks for time-series forecasting, XGBoost (Extreme Gradient Boosting) for classification, and K-Means clustering for grouping similar data points. This combination allows us to handle different types of data effectively and improve overall model performance. LSTM networks are a type of recurrent neural network (RNN) designed to process sequential data, making them ideal for time-series forecasting (Olah, 2021). Unlike traditional neural networks, LSTMs can remember long-term dependencies, which is essential for predicting trends over time. For example, in financial markets, LSTMs can analyze stock prices over months or years to make future predictions. Similarly, in healthcare, they can be used to track patient data and forecast disease progression (Khan et al., 2022).

XGBoost is used in the classification of tasks. It is an optimized version of gradient boosting that improves speed and accuracy by using parallel processing and efficient handling of missing values (Chen et al., 2020). This makes it highly effective for applications such as fraud detection, medical diagnosis, and customer segmentation. XGBoost is widely used in data science competitions due to its superior performance in structured data problems (Liu et al., 2023).

Finally, K-Means clustering is a popular technique for unsupervised learning, where data points are grouped based on similarities without predefined labels (Patel & Shah, 2021). It is commonly used for customer segmentation, market analysis, and anomaly detection. By integrating K-Means into the hybrid AI approach, we can first group data into meaningful clusters before applying classification or forecasting models. This helps to improve accuracy and provides deeper insights into the data (Zhang & Wang, 2023). By combining these three approaches, we create a robust AI system that can handle different types of tasks efficiently. LSTM predicts trends over time, XGBoost classifies data based on predefined labels, and K-Means organizes unstructured data into meaningful clusters. This synergy allows us to tackle real-world problems in fields like finance, healthcare, and e-commerce more effectively than using any single algorithm alone.

Model Evaluation: RMSE for Forecasting, Precision-Recall for Classification, and Silhouette Score for Clustering. Once machine learning models are trained, it is important to evaluate their performance using appropriate metrics. Different tasks require different evaluation methods to ensure the model produces reliable and meaningful results. For time-series forecasting, we use Root Mean Square Error (RMSE); for classification models, we rely on the Precision-Recall metric; and for clustering models, we assess quality using the Silhouette Score. RMSE (Root Mean Square Error) is a commonly used metric for measuring forecasting accuracy (Wang & Li, 2021). It calculates the average error between predicted and actual values, giving more weight to larger errors. A lower RMSE value indicates better predictive performance. For example, if an LSTM model is forecasting temperature changes, RMSE helps determine how closely the predictions match real-world measurements. Precision-Recall is an essential metric for evaluating classification models, especially when dealing with imbalanced datasets (Gao et al., 2022). Precision measures how many of the predicted positive cases are actually correct, while recall indicates how many of the actual positive cases were correctly identified by the model. This metric is particularly useful in applications like medical diagnostics, where false positives and false negatives can have serious consequences. Silhouette Score is a popular metric for evaluating clustering models (Sun & Chen, 2023). It measures how well data points fit within their assigned clusters by comparing intra-cluster cohesion with inter-cluster separation. A higher Silhouette Score indicates that clusters are well-defined and distinct from each other. This is useful in applications like customer segmentation, where businesses want to group similar customers for targeted marketing. Using the right evaluation metric ensures that machine learning models perform optimally in their specific tasks. RMSE helps improve forecasting models, Precision-Recall enhances classification accuracy, and the Silhouette Score ensures high-quality clustering results. These metrics provide insights into model effectiveness and help fine-tune performance for real-world applications.

## 2.10 Related works

In this section, we review studies related to blood donation systems, focusing on their findings and identifying gaps that current research seeks to address.

### 2.10.1 Pathways for Innovation in Blood Transfusion Services in Kenya (PITS Kenya)

The Pathways for Innovation in Blood Transfusion Services in Kenya (PITS Kenya) research was conducted between July 2022 and March 2023 by Kumar and Puyana. The study explored blood transfusion services in three counties in Kenya: Siaya, Nakuru, and Turkana (Kumar & Puyana, 2023).

#### 2.10.1.1 Goals of the Research

The primary objectives of the study were:

i. To identify deficits in blood availability for different types of patients in Siaya, Nakuru, and Turkana.
ii. To create a visual representation and simulation of the processes between blood collection and transfusion.
iii. To collaborate with donors, patients, blood banks, and healthcare workers to identify practical solutions for increasing blood availability.
iv. To test the impact of the proposed solutions on blood availability at the point of care (Kumar & Puyana, 2023).

#### 2.10.1.2 Key Findings

The research revealed several critical challenges in Kenya's blood transfusion services:

i. Lack of Timely Blood Availability: A shortage of available blood for transfusion contributes to poor patient outcomes, deferred care, and extended or unnecessary hospitalizations.
ii. Stress on the Healthcare System: Poor patient outcomes due to blood shortages create significant stress across all levels of the healthcare system, including administrators, healthcare professionals, laboratorians, donors, and patients.

iii. Interconnection of Blood Collection and Use: Blood collection and delivery processes are concurrent and interdependent, requiring synchronized management.
iv. Mismatch in Blood Availability and Demand: A disconnect exists between where blood is available and where it is needed, leading to logistical, referral, and transportation challenges (Kumar & Puyana, 2023).
v. Influence of Knowledge, Attitudes, and Perceptions: The willingness to donate or receive blood is strongly influenced by knowledge gaps, attitudes, perceptions, and cultural beliefs.
vi. Dependence on Family Replacement Donors: Overreliance on family replacement donors fosters a reactive, subsistence-based approach, where planned blood use is frequently diverted to address urgent or emergent needs.

#### 2.10.1.3 Proposed Solutions for Improving Blood Donation and Management in Kenya

To address the identified challenges and improve blood availability at the point of use, the research suggested the following strategies:

i. Mobilizing Community-Based Blood Supply: Engaging communities where blood is readily available to increase donations.
ii. Enhancing Blood Collection, Testing, and Distribution: Improving infrastructure and processes in blood collection centers to ensure efficient testing and distribution.
iii. Developing Interfaces Between Hospitals and Blood Banks: Building better communication and coordination systems between clinicians, hospital personnel, and blood banks.
iv. Establishing Community-Facility Transfusion Committees: Creating committees to mobilize blood donations, educate communities about the importance of blood transfusion, and address cultural misconceptions.
v. Improving Blood Distribution Logistics: Streamlining the logistical and transportation processes to ensure timely delivery of blood to areas with high demand (Kumar & Puyana, 2023).

The study's findings and recommendations by Kumar and Puyana (2023) provide a practical framework for improving blood donation, management, and distribution in Kenya. By addressing critical deficits and implementing innovative solutions, the study aims to enhance timely access to blood transfusions and improve maternal and general healthcare outcomes.

#### 2.10.1.4 Limitations of the study

The Pathways for Innovation in Blood Transfusion Services in Kenya (PITS Kenya) study has several limitations, including challenges in accessing accurate and real-time data on blood supply and demand, which can hinder effective planning. The uneven distribution of blood, with a potential focus on urban areas at the expense of rural regions, could lead to significant disparities in access. Additionally, the study lacks a predictive model that can be used to forecast demand versus supply, failing to account for seasonal variations, emergencies, or demographic shifts. These traditional donor engagement strategies might also not fully leverage the potential of younger, tech-savvy populations. A lack of integration between prediction models, health facilities, and blood banks, could further complicate decision-making processes.

To enhance decision-making regarding blood demand and supply, key ideas from PITS Kenya can be borrowed and refined. For instance, their forecasting methods could be augmented with advanced predictive analytics and machine learning to incorporate real-time hospital data, historical trends, and emergency scenarios. Building on their data collection framework, a centralized, real-time data system could link blood banks, hospitals, and mobile units, ensuring transparency in supply-demand management. Community outreach programs from PITS Kenya could be adapted to create more targeted donor campaigns, leveraging mobile apps and SMS alerts to reach rural and underserved populations. Additionally, their use of technology for donor mobilization could inspire the development of a mobile platform that tracks donation history and alerts donors to nearby drives or emergencies. Finally, their policy recommendations could be expanded to advocate for automated inventory systems, streamlined regulatory processes, and contingency frameworks to better prepare for emergencies and ensure equitable distribution of blood resources (Kumar P. e., 2021).

### 2.10.2 Blood Donation Management System

The aim of the research was to develop an application that offers information on blood donation and facilitates donor registration, thus enhancing the effectiveness of blood distribution. This system enabled communication between blood donors and blood banks. It also provides insights into blood donation for donors, recipients, and other relevant parties (Kaur, 2022).

The study recommended digitizing all blood records into open-source databases. Future donors would have the opportunity to use an intuitive website to register for blood donation easily. Hospitals and patients would be able to find nearby donors by inputting a landmark as a search criterion. The proposed system would be accessible to anyone with internet connectivity and a web browser which is convenient to potential donors and minimizes time inefficiencies. It ensures that all records for hospitals and recipients are readily accessible. Donors can also receive alerts regarding urgent donation needs through a notification feature designed for recipients. In cases where no local donors are available during emergencies, the patient or hospital can relay the urgency to a group of donors with the same blood type residing in the same city (Kaur, 2022).

The application has the following modules:

i. Blood Type Module. This component enables users to search for specific blood types and submit requests accordingly. The system leverages available resources to identify donors with the requested blood type and sends notifications about the blood need. This functionality allows patients to quickly locate and obtain blood in their vicinity, thereby saving crucial time.
ii. Donor Registration Form. This form facilitates the registration of individuals as donors. It gathers information including name, gender, date of birth, contact number, email address, home address, last donation month, and health information. Donors are classified as returning donors, who donate on a regular basis, and walk-in donors, who may be new or occasional contributors.
iii. Administration Portal. This portal is intended for user access and permission management, as well as for executing administrative tasks related to the platform.

#### 2.10.2.1 Limitations and gaps

Although the system seeks to digitize the donation process, several gaps and limitations are identified with this solution. The system only sends donation requests to donors during emergencies, however, we are not guaranteed the availability of the donor at that specific moment. Donation is voluntary and donors might be engaged in their other day-to-day activities which provide for their needs. Therefore, getting blood in a time of need becomes difficult for both patients and mothers during birth.

### 2.10.3 Optimization of the blood bank management using a distributed web application

This research was done by Catherine Mwangi of the University of Nairobi in 2023. The research seeks to forecast blood order quantities and efficient use of collected blood to ensure no blood is wasted. The web application used a stochastic algorithm and a FIFO strategy to implement forecasting of blood collection and ordering of blood by various hospitals. The web application had the following modules and functionalities (Catherine Mwangi, 2023).

Blood collection module. This module is used to manage blood collection by donors by checking on available stocks.

Blood allocation module. The module checks on available stock and allocates blood to various hospitals using a genetic algorithm. The module also uses a FIFO strategy to allocate blood to hospitals and hence reduce expiry.



#### 2.10.3.1 Limitations and gaps

The research proposed a solution that only manages the collection of blood from donors and the allocation of blood to various hospitals. However, the proposed solution is not able to predict blood demand from hospitals as it only allocates blood available in their blood banks. The research focuses mainly on the allocation process of available blood in the blood banks but does not investigate ways of replenishing bloodstock based on demand. It also does not address the deaths related to maternal health in Kenya.

### 2.10.4 Kenya Ministry of Health Blood Bank Information System

The research initiative aimed to tackle the challenges encountered by blood banks in Kenya, specifically the absence of an effective blood donation management system. The suggested approach incorporates computerized technology and sustainable strategies to facilitate all activities related to blood donation, which include monitoring and evaluating standards concerning blood donation, processing, transportation, and distribution to ensure effective blood transfusions. The project intends to implement a system that enhances the donation, processing, and distribution of blood, as well as its transportation (Munyao, G. M., 2021).

The system's features are user login and authentication for accessing the web-based platform, utilizing QR codes to track blood bags, and an Android application designed to scan these QR codes at every stage of a blood bag's journey. The mobile app can automatically notify donors once their blood has been transfused to a patient. The system enables administrators to create blood appeals, which registered donors can respond to by agreeing to donate. It is designed to utilize location pinpoints for hospitals and patients when requesting blood appeals, this connects them with nearby donors. Furthermore, the system informs both patients and donors about the closest blood banks based on their locations. Blood banks are also expected to regularly update their stock levels and share this information through the application. If any blood remains unutilized beyond a specified timeframe, an alert is dispatched to the appropriate authorities (Munyao, G. M. (2021)).



#### 2.10.4.1 Limitations and gaps

The research proposed a solution that only manages the collection of blood from donors and the creation of blood appeals by patients and hospitals. However, the proposed solution is not able to predict blood demand from hospitals as it only alerts donors whenever there is a blood appeal. The research focused mainly on automating the donation journey but did not use existing technologies to inform blood banks and hospitals of future needs for blood based on existing data. It also does not address the deaths related to maternal health in Kenya.

### 2.10.5 Blood Bank Management System

The research wanted to develop a tool to manage blood bank operations that will focus on monitoring the inventory of blood donations and samples. It suggested the design and implementation of a blood bank management system that uses a Database Management System (DBMS) with Java Database Connectivity (JDBC). The product was a user-friendly interface that enhanced the management of blood donations and inventory oversight (Sakshi Patil, 2023).
To support the blood donation process, the study implemented the following modules:

i. Blood Inventory. Designed to manage the blood inventory, providing real-time updates on the availability of blood units, their expiration dates, and blood types.
ii. Donor Information Management. This module focuses on optimizing donor management within blood banks by keeping detailed records of donors, including their blood types and medical histories. It will also enable blood banks to monitor donors' donation histories, contact information, and eligibility status (Sakshi Patil, 2023).
iii. Integration with Existing Hospital Management Systems. The system is intended to integrate with existing hospital management systems, such as electronic health records (EHRs) and laboratory information systems (LISs). This integration will assist healthcare professionals in accessing and sharing blood donation information across different platforms (Sakshi Patil, 2023).

#### 2.10.5.1 Limitations and gaps

The proposed solution seeks to develop a blood bank management system tool for maintaining the inventory of blood donations and blood samples. Although it also has some reporting components to show available volumes in the blood banks, it lacks a tool that can intelligently predict demand vs supply of blood for better planning. It also lacks a supply chain management module and does not address the delays in blood distribution and wastage of blood products. Poor inventory management practices and a lack of real-time tracking systems can exacerbate these challenges. This makes it difficult to track blood donations, inventory levels, and patient transfusion histories accurately. This can lead to inefficiencies in blood allocation and tracking of blood usage.

It is also not available to be used in Kenya specifically in the context of maternal health hence the proposed research.

## 2.11 Gaps in Existing Solutions

In Kenya, the existing blood donation system reveals significant gaps when considering the balance between supply and demand, particularly in the context of maternal healthcare. These gaps present challenges that directly impact the availability of safe blood products for pregnant women and childbirth-related emergencies. One of the primary challenges in Kenya's blood donation system is the inability to meet the demand for blood products, especially during maternal health emergencies. Pregnant women often require blood transfusions due to complications such as postpartum hemorrhage or complications during childbirth. However, the supply of blood products may not be adequate to meet these urgent needs (Michuki, 2022).

Blood donation centers and infrastructure are often concentrated in urban areas, leading to disparities in access to blood transfusion services for pregnant women in rural and remote regions. Limited access to blood donation facilities in rural areas means that pregnant women facing obstetric emergencies may not have timely access to life-saving blood transfusions, contributing to maternal mortality rates. Kenya faces challenges in recruiting and retaining blood donors, particularly in rural communities where awareness about the importance of blood donation for maternal healthcare may be low. There is a need for targeted donor recruitment campaigns that emphasize the critical role of blood donation in saving the lives of pregnant women and newborns (Redcross, 2021).

Furthermore, there may be gaps in screening and testing protocols for donated blood, leading to concerns about the transmission of infectious diseases through blood transfusions. Strengthening blood screening and testing protocols is essential to ensure the safety of blood products used in maternal healthcare settings. Additionally, Kenya's blood donation system may lack robust emergency preparedness and response mechanisms to address sudden increases in demand for blood products during maternal health emergencies. Improved coordination between blood donation centers, healthcare facilities, and emergency response teams is crucial to ensure timely access to blood transfusions for pregnant women experiencing obstetric complications (Redcross, 2021).

To address these challenges and bridge the gap between blood supply and demand, there is a pressing need to leverage artificial intelligence (AI) and machine learning (ML) technologies to predict future demand for blood products more accurately. By analyzing historical data on blood donation patterns, patient demographics, and healthcare trends, AI and ML algorithms can forecast the anticipated need for blood products, allowing blood donation centers to proactively adjust their collection and distribution efforts. Implementing AI and ML-based predictive models can help optimize blood supply management, reduce shortages, and improve the availability of safe blood products for maternal healthcare and other critical medical needs (Enrique Onieva, 2023).

## 2.12 Conceptual model

The proposed solution, a predictive model for blood demand and supply in maternal healthcare in Kenya, is conceptualized as a comprehensive platform that integrates AI, ML, mobile technology, and supply chain management principles to optimize blood inventory management and enhance maternal health outcomes. The conceptual model below outlines the key components and functionalities of the app: The figure 2. 4 shows the conceptual model of the solution

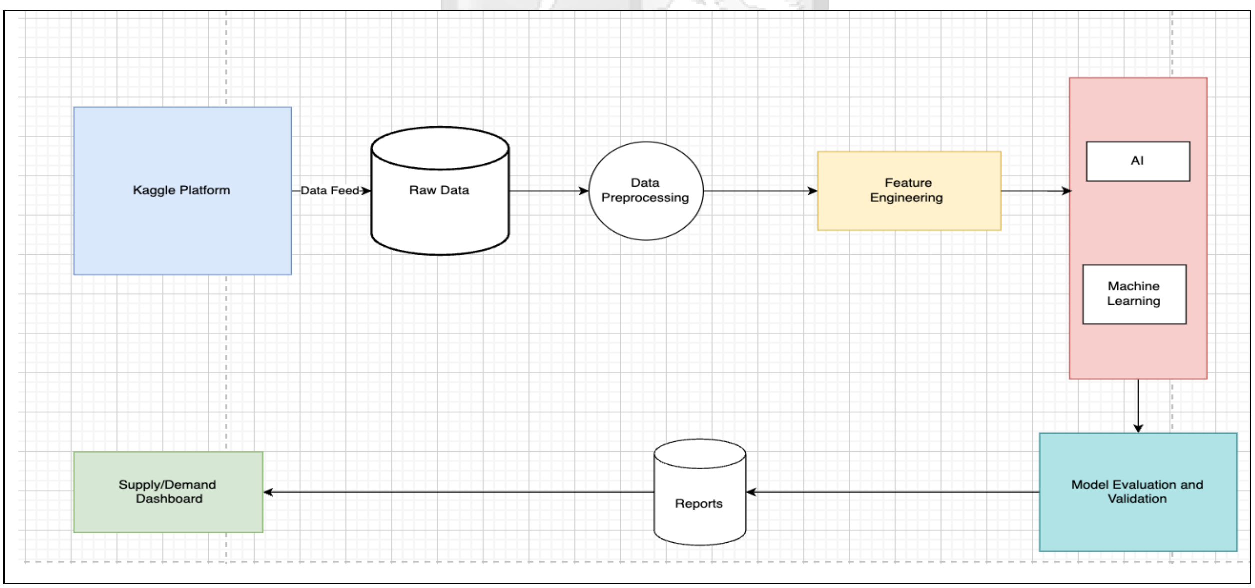



Figure 2.4: Conceptual Predictive Model for Blood Demand and Supply

The model begins with data collection from various sources, including Kaggle, hospitals, blood donation centers, and other institutions tracking blood usage and donations. Once collected, the data undergoes preprocessing to address missing values, remove duplicates, and transform it into an appropriate format for analysis. A portion of the data is used for training the machine learning model to understand relationships between features and outputs, while another portion is reserved for testing the model's performance. The core of the system is the machine learning algorithm, which is evaluated for accuracy and precision. The model's predictions generate reports on future blood donation rates and supply needs, offering valuable insights for decision-makers. These predictions are displayed on a user-friendly dashboard that enables stakeholders to monitor and respond to blood supply and demand in real-time. Finally, a feedback loop ensures continuous improvement by allowing the model to learn from new data, keeping predictions accurate and adaptable to changing conditions.

## 2.13 Chapter Summary

Chapter 2 of the thesis provided a comprehensive overview of the existing literature relevant to the development of a predictive mobile app aimed at managing blood demand and supply, particularly within the context of maternal healthcare in Kenya. This literature review serves as the foundation for understanding the complexities surrounding blood donation, supply chain management, and maternal health while also identifying gaps and opportunities for innovation.

# Chapter 3: Research And Design Methodology

## 3.1 Introduction

Research methodology is a process by which a researcher carefully designs their investigation to ensure they provide accurate and trustworthy results that address the research objectives (Hannah Snyder, 2021). This section of the research is used to present the research design, system analysis architecture, design, development, implementation, and testing. This research will be guided by the proposed objectives and its delivery is driven by the underlying problem.

## 3.2 Research Design

A research design demonstrates how the research will be conducted and carried out to ensure efficiency and effectiveness. It is a plan that provides the underlying structure to integrate all elements of a quantitative study to make the results credible and free from bias, (Sharon Anderson Dannels, 2022).The research design will determine how the participants will be selected, what variables will be included and how they will be manipulated, how data will be collected and analyzed, and how extraneous variability will be controlled so that the overall research problem can be addressed (Sarah Wright, 2023). This study was carried out using an experimental study. The experimental study used historical data from the Kaggle platform to train the proposed predictive model.

### 3.2.1 Experimental Study Design

An experimental research design approach will be used to process historical data, which is described as a form of research methodology used in scientific inquiries to examine cause-and-effect interactions between variables. Experimental research design is a study conducted with a scientific approach using two sets of variables. The first set acts as a constant, which you use to measure the differences of the second set (Adi Bhat, 2024). In this study, the constant or controlled variable would be a factor that remains unchanged throughout the experiment to ensure that any changes in blood demand or supply are due to the variables being tested and not influenced by other factors.

The constants for the study are the population size and demographics, which refer to the size of the population and its demographic factors like age distribution and gender ratio (Xavier Becerra, 2021). Healthcare facility capacity, such as the number of beds, the capacity of hospitals to accommodate patients, and the availability of blood storage units, will also remain constant (WHO, 2024). The seasons or time of year will be kept constant to avoid fluctuations caused by seasonal illnesses, accidents, or donation patterns. Blood donation infrastructure, including the number of blood banks, mobile donation units, and drives, will be constant as well (NCBI, 2024). Additionally, healthcare and blood donation policies and regulations will remain unchanged to prevent any impact on data due to shifts in policies, such as restrictions on certain donors (Xavier Becerra, 2021).

The variables of the study are the factors the researcher manipulated to understand their impact on the outcome. The study will be carried out using two types of variables: Independent variables and dependent variables.

### 3.2.2 Independent Variables

The study identified blood demand as the key independent variable, influenced by several factors that directly or indirectly impact its dynamics. Public health campaigns play a significant role in shaping blood demand by raising awareness about the critical need for blood donation. When designed effectively, such campaigns have been shown to increase participation rates and align supply with demand (Xavier Becerra, 2021). Incentives, both financial and non-financial, such as gift cards or community recognition, further influence blood donation behavior, with their success largely determined by cultural and socioeconomic contexts.

Additionally, the frequency and accessibility of blood donation drives, including mobile units and strategically located centers, can influence donor turnout, subsequently affecting blood demand. Economic factors also play a critical role; for instance, economic instability or rising unemployment rates can lead to decreased donations as individuals prioritize personal financial stability, creating gaps in meeting blood demand. Moreover, healthcare-related variables, such as increased elective surgeries or trauma cases, directly influence fluctuations in blood demand (Charles H. Jones, 2024). As healthcare needs rise, so does the requirement for an adequate blood supply, making this a critical factor in understanding demand patterns. Awareness programs, which complement public health campaigns, are another influential variable. These programs educate the public about the ongoing need for blood, fostering consistent donor engagement and addressing shortfalls in supply (Xavier Becerra, 2021).

### 3.2.3 Dependent Variables

The primary dependent variable in this study is blood supply, which represents the outcomes influenced by changes in independent variables. Blood supply is a critical measure, encompassing the volume of blood donations, the number of available blood units, and the percentage of blood banks that can meet the demand. Maintaining a consistent and adequate blood supply is essential to prevent shortages, especially during periods of high demand (Charles H. Jones, 2024). Several factors contribute to fluctuations in blood supply. For instance, blood demand, driven by the number of requests for blood transfusions or blood products from hospitals and medical facilities, directly influences the adequacy of supply. Medical emergencies and changes in healthcare trends, such as increased surgical procedures, are key contributors to this demand (Xavier Becerra, 2021).

Donation rates, reflecting the frequency and number of blood donors, also significantly affect the stability of the blood supply. These rates can be influenced by interventions such as public health campaigns, financial or non-financial incentives, and the accessibility of donation drives (Charles H. Jones, 2024). Another important aspect is supply shortages, which are measured by the frequency and duration of instances when blood supply fails to meet demand. Understanding these shortages helps evaluate the effectiveness of strategies aimed at improving donor participation and enhancing the logistics of blood collection and distribution (Xavier Becerra, 2021).

By monitoring these outcomes, the study aims to provide valuable insights into the dynamics of blood supply, enabling the development of strategies to ensure its availability and stability in the face of fluctuating demand and other influencing factors.

## 3.3 Model Development

The model's development took the following steps:

i. Data Collection
ii. Data Preparation
iii. Model Training
iv. Model Evaluation
v. Model Testing & Validation.

### 3.3.1 Data Collection

The training data for this study is obtained from three datasets in Kaggle, the blood service transfusion center, and WHO (shivan kumar, 2020). The study adhered to the rules established by all the data set platforms for data usage to ensure ethical data utilization (Marcos Martins, 2019).. These rules state that the data should only be used for non-commercial research projects, that any analysis should be done with respect and ethics, and that the data's source should always be appropriately acknowledged.

To effectively train the model, the data obtained will be analyzed using relevant datasets guided by specific indicators aligned with the project's objectives. The focus will be on obstetric emergencies, maternity data, and blood requirements at Kenya National Hospital (KNH), as well as regional blood donation trends from the Kenya National Blood Transfusion Service (KNBTS) and the Kenya Tissue and Transplant Authority (KTTA).

#### 3.3.1.1 Data Sets

Key parameters to provide the data sets included the daily number of obstetric emergency cases recorded at KNH, the average number of maternity patients admitted, and the average number of obstetric emergencies requiring blood transfusion per day. Additionally, data on blood donation trends, including donor demographics and peak donation periods, as well as the frequency of blood shortages and unmet demand for specific blood groups, will be used. KNH's obstetric and maternity records will also be a primary source of data, providing insights into patient flows, emergency cases, and blood usage. Data from KNBTS will supplement this with information on donation trends and blood availability, while KTTA will contribute broader insights into tissue and blood requirements. To refine the datasets, parameters such as date range (to capture daily and long-term trends), case type (routine vs. emergency), blood group distribution, and geographic scope (focusing on KNH with supplementary regional data) will guide the filtering and extraction process. The table below summarizes the key datasets and parameters that will be used to develop the model. Table 3.1 show the indicators and parameters that will be used to derive the data sets

Table 3.1: Key Indicators and Parameters to Derive Data Sets

| **Indicator Category** | **Parameter** |
|---|---|
| Clinical Demand | Number of obstetric emergencies needing blood per day |
| | Maternity admissions and postpartum hemorrhage cases |

| Indicator Category | Parameter |
| --- | --- |
| | Historical blood usage by department |
| Blood Supply | Daily blood inventory levels by type |
| | Number of donations collected (voluntary vs. replacement) |
| | Blood wastage and rejected units |
| Operational | Referral cases due to shortages |
| | Turnaround time for blood delivery |
| Donor Trends | Seasonal variations in donation |
| | Donor demographics and participation rates |
| External Factors | Disease outbreaks, trauma cases, and logistical disruptions |

### 3.3.2 Data Preparation

The data from Kaggle is prepared for the blood donation model by exploring and cleaning the dataset. Missing or inconsistent values will be checked, and they will either be imputed or removed, depending on the dataset's size and quality. Next, the data will be standardized and normalized, ensuring that features like donation frequency, donor age, blood type, and time since the last donation are appropriately scaled for the model. Categorical data, such as blood type, will be converted using one-hot encode to numerical format. Feature engineering will be key, as new features like trends in donation over time or seasonal patterns may be created. The dataset is also split into two, the training and the test sets to effectively evaluate the model. Finally, the data will be labeled and structured properly before being fed into the AI model (Bert Carremans, 2020).

### 3.3.3 Model Training

Training of the model will involve selecting appropriate machine-learning techniques based on the characteristics of the blood donation dataset. The data will first be pre-processed and split into training, testing, and validation sets. Techniques like linear regression or logistic regression may be used for predicting trends in blood donation demand, while classification algorithms such as decision trees, random forests, or support vector machines (SVMs) will be applied for donor segmentation or predicting donor behavior. Neural networks and deep learning techniques, especially recurrent neural networks (RNNs) or long short-term memory (LSTM) networks, will be leveraged if the model focuses on time-series forecasting of blood supply trends. Hyperparameter tuning and cross-validation will be performed to optimize model performance. Throughout the training process, metrics like accuracy, precision, recall, and F1-score will be used to evaluate and adjust the model, ensuring it generalizes well to unseen data. Once trained, the model will be tested on the reserved data to validate its predictive accuracy and robustness.

### 3.3.4 Model Evaluation

Model evaluation will play a critical role in determining the effectiveness and accuracy of the blood donation prediction model. Techniques such as cross-validation, particularly k-fold cross-validation, will be used to assess the model's ability to generalize across different subsets of the data. Various evaluation metrics will be employed depending on the problem type: for classification tasks (e.g., predicting whether a person will donate blood), metrics like accuracy, precision, recall, F1-score, and area under the ROC curve (AUC) will be used to assess performance. For regression tasks (e.g., predicting the number of donations or demand for blood), metrics such as mean squared error (MSE), mean absolute error (MAE), and R-squared will be utilized. In cases involving time-series predictions, mean absolute percentage error (MAPE) and root mean squared error (RMSE) will be considered. Model evaluation will also involve testing for overfitting using techniques like regularization (e.g., L1, L2) or early stopping in neural networks, ensuring the model performs well on unseen data.



### 3.3.5 Model Testing & Validation

Model testing and validation are essential to ensure that the blood donation prediction model performs accurately and reliably on unseen data. Once the model is trained, it will undergo testing using a separate test dataset that was not used during the training process. This will help evaluate how well the model generalizes to new, real-world data. Techniques like holdout validation will be applied, where a portion of the data is set aside solely for testing. In addition, cross-validation, such as k-fold cross-validation, will be used to validate the model across multiple subsets of the data, reducing the risk of overfitting and ensuring consistency in performance across different data splits.

During the validation phase, hyperparameters will be fine-tuned, and methods like grid search or random search will be used to identify the best model configuration. The model’s performance will be validated based on metrics like accuracy, precision, recall, or root mean squared error (RMSE) depending on the task (classification or regression). Techniques such as early stopping and regularization (L1, L2) will also be used during validation to prevent the model from overfitting to the training data. Finally, once validated, the model will be tested on the held-out dataset to measure its true performance in production-like scenarios, ensuring it can accurately predict future blood demand or donation patterns.

## 3.4 System Development Methodology

The study and development of the model will be implemented using Evolutionary Prototyping. Evolutionary prototyping is a dynamic and iterative system development approach that emphasizes the gradual refinement of a system through user feedback and evolving requirements. (Molisch, 2024). Key features of evolutionary prototyping include its user-centric design, which ensures alignment with user needs; flexibility to accommodate changes at any stage; and risk mitigation by delivering functional prototypes early to identify and resolve issues. Additionally, the incremental development process builds functionality in stages, focusing on critical features first, and continuously improving upon them. This iterative cycle makes the approach well-suited for dynamic projects requiring continuous user engagement and adaptability (Mackay, 2023). This approach not only reduces development risks but also enhances innovation and responsiveness, making it an ideal choice for this study.

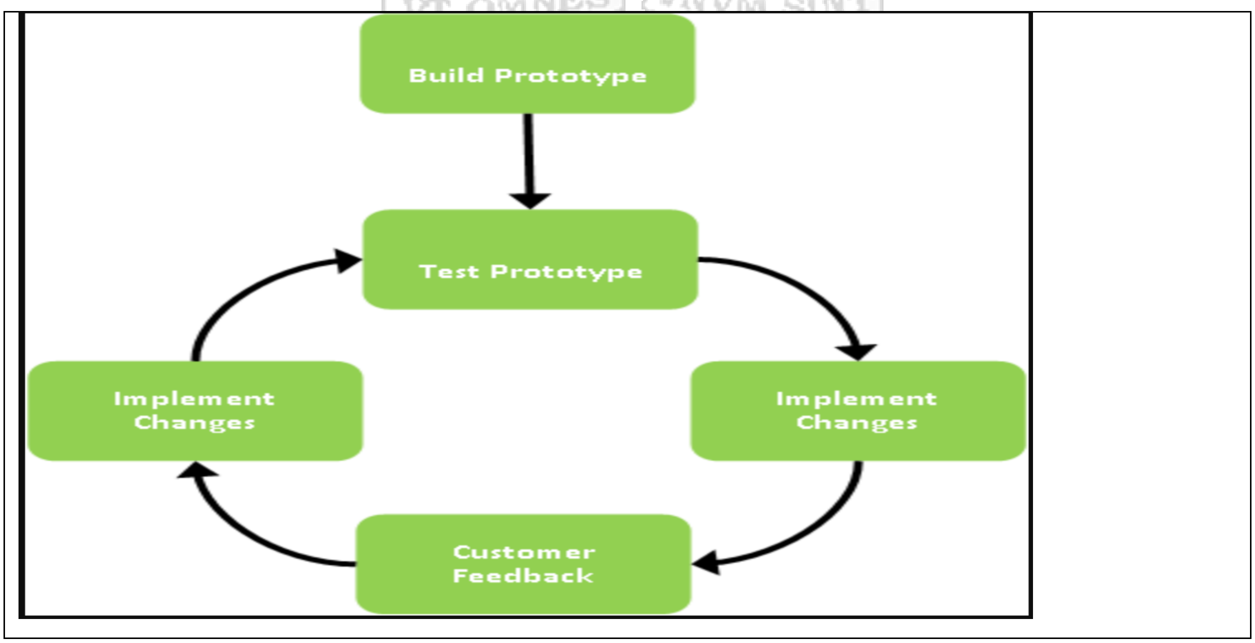



Figure 3.1: Evolutionary Prototyping (Mackay, 2023)

### 3.4.1 Requirements Planning

In Agile methodology, planning is an ongoing and adaptive process rather than a one-time event. Agile planning focuses on delivering value in small, incremental pieces, allowing teams to adjust and respond to changes quickly. At the start of a project, high-level planning establishes the product vision and breaks down the work into a product backlog, which lists the features or user stories in order of priority (Sarah Laoyan, 2024). During sprint planning, teams select a subset of these items to complete in the next sprint, typically lasting 1-4 weeks. The scope is intentionally kept flexible to accommodate evolving customer requirements and new insights. Planning also includes estimating the effort required for each task using techniques like story points or t-shirt sizing. Unlike traditional approaches, Agile planning is iterative; it encourages continuous feedback from stakeholders and customers, allowing teams to refine their goals and priorities as the project progresses. This adaptability ensures that the team remains focused on delivering the highest-value features first, even as circumstances change (Artem Gurnov, 2024).

### 3.4.2 Requirements Analysis

The Agile methodology defines requirements analysis as a continuous process that is collaborative and focused on understanding the needs of the customer and translating them into actionable items for the development team. Unlike traditional methodologies, where requirements are gathered upfront and often set in stone, Agile encourages ongoing engagement with stakeholders to refine and prioritize requirements throughout the project lifecycle (Artem Gurnov, 2024). The product backlog, which consists of user stories or features, is the main tool for managing these requirements. During backlog refinement sessions, the team works closely with the product owner to clarify, estimate, and break down large tasks into smaller, more manageable pieces. This process allows for flexibility, enabling teams to adapt to changes in business priorities or market conditions. Requirements are revisited in every sprint, ensuring that the development remains aligned with the customer's evolving needs and delivering value incrementally source (Sarah Laoyan, 2024).

### 3.4.3 System Design

System design in Agile methodology is an iterative process where the architecture evolves alongside development, ensuring flexibility and responsiveness to changing requirements. Instead of finalizing the entire design upfront, Agile promotes incremental design, where high-level architecture is created at the start, and detailed design decisions are made progressively during each sprint (Lu Bauer, 2024). Teams use collaborative techniques, such as UML diagrams (like class diagrams and sequence diagrams) and data flow diagrams (DFDs) to visualize system components and their interactions (Sarah Laoyan, 2024). Design sessions occur regularly, often during sprint planning or dedicated technical meetings, allowing the team to discuss architecture impacts and adjust as necessary. Refactoring is frequently applied to optimize the system over time. By using modular architecture and design patterns, the design remains scalable and adaptable, supporting continuous development while maintaining system integrity source.

### 3.4.4 System Implementation

System implementation in Agile methodology is an iterative process where the development and deployment of system features happen in small, manageable increments. Agile emphasizes delivering working software frequently, so implementation is broken down into multiple sprints or iterations, where specific features or user stories are developed, tested, and integrated into the system. Throughout each sprint, the development team collaborates closely with stakeholders and the product owner to ensure the implemented features meet the desired requirements and provide business value. Continuous integration (CI) and continuous delivery (CD) are core practices in Agile implementation, allowing teams to frequently integrate code changes into the system, automatically test them, and deploy them to production environments when they are ready. The use of automated testing ensures that new changes do not disrupt existing functionality. Agile implementation also involves regular retrospectives, where the team reflects on the process and identifies improvements for future sprints, enhancing efficiency and quality over time (Atlassian, 2024).

### 3.4.5 System Testing

System testing in Agile methodology is a critical phase that ensures the software meets the specified requirements and functions correctly before it is released. This process is conducted after the implementation of features during each sprint and involves evaluating the entire system's behavior. Agile emphasizes the use of automated testing to facilitate frequent testing, allowing teams to quickly identify and address any defects. Various testing types, including unit testing, integration testing, functional testing, and user acceptance testing (UAT), are employed to validate different aspects of the system. Collaboration between developers, testers, and stakeholders is vital during this phase; continuous feedback loops enable quick iterations and adjustments based on test results. Agile encourages the use of test-driven development (TDD), where tests are written before the code itself, ensuring that the implemented features align with user expectations from the outset. Additionally, regular retrospectives are held to assess the testing process and identify improvements, ensuring that the team maintains high quality and efficiency in delivering a reliable product (Daniel Wilhite, 2024).

## 3.5 Ethical Approval

This study was built on ethical scientific methods and procedures. Consent from respondents will be obtained from the respondents before the survey and information collected will be treated with utmost confidentiality. Ethical approval has been attached in appendix A.

## 3.6 Summary

This chapter presents the methodology that will be used in undertaking this study. The first section of this study research design demonstrates how data will be collected, analyzed, and presented. The system development approach has also been discussed as well as the system requirements gathering, development, and testing. The last part of the study highlights the ethical considerations of the study.

# Chapter 4: System Analysis And Design

## 4.1 Introduction

In this section, we present the architectural design of the blood supply and demand model, which will adhere to the conceptual framework illustrated in Figure 2.5. We will discuss the interactions between users and the model, detail the model's components, and explain how these components interact with one another. The system's design and architecture will be depicted through Unified Modelling Language (UML) diagrams.

## 4.2 Data Gathering

For the development and training of the predictive blood demand and supply model, data was gathered from a variety of sources to ensure the model could make accurate predictions based on real-world patterns. The following datasets were used to inform the model's parameters and help generate meaningful insights into blood demand trends and supply allocation: The first data set was obtained from WHO; it provides a wide range of health-related data, which includes global statistics on blood donations, blood usage, health system infrastructure, and demographic information. By leveraging this dataset, the model can take into account global health trends, seasonal variations, and blood demand factors on an international scale. This data helps in forecasting demand fluctuations across different regions and populations, which is critical for understanding broader blood supply challenges. The second dataset used is the UCI Blood Transfusion Service Center dataset, which is a well-established resource for analyzing blood donation behaviors. It contains historical data from blood donors, detailing various factors such as donor demographics, blood donation frequency, and whether the donor's blood was eventually used in the healthcare system. This dataset allows for the identification of patterns in donor behavior and blood usage, which can be incorporated into the model for better forecasting of blood demand. By examining the characteristics of regular blood donors and their contribution to blood banks, the model can predict future donor trends and supply availability more accurately. The third dataset was obtained from Kaggle blood donation datasets, which provide various public datasets related to blood donations, including donation frequency, donation outcomes, and patient outcomes. These datasets often contain more granular data, which can be essential in modeling donor behavior, predicting blood shortage risks, and identifying key factors that influence blood demand at different times of the year or in response to health crises. Kaggle's community-based datasets offer varied data across different regions, making them useful for training the model on both global and local blood demand scenarios.

## 4.3 Requirements Analysis

The system requirements analysis focuses on identifying the expectations of users who will engage with the developed system. This process involves a comprehensive evaluation of the requirements aligned with the core objectives of the study (Laplante, 2022). The requirements identified for this research are classified into three main categories: functional, the non-functional, and usability.

### 4.3.1 Functional Requirements

i. The model should enable users to input data in CSV format exclusively, rejecting any other data types.
ii. The model should accept data about details of blood supply and demand.
iii. The model is required to produce the most accurate predicted output.
iv. The model should output predictive reports with the indicators below.

   a. Historical demand patterns at each facility
   b. Leading indicators that signal upcoming demand increases
   c. Stockout frequency by blood type and location
   d. Demand-supply gap measurements.

ii. The model should predict accurate matches using the input supplied by the user.

### 4.3.1 Non-Functional Requirements

#### 4.3.1.1 Usability Requirements

The model is a representation of the full prediction application that will be installed at Kenyatta National Hospital to manage the blood versus supply prediction process. The main users will be blood donation center admins and any other stakeholders involved in blood donation and supply. The model should, therefore, be simple and accurate to enhance user experience. Its prediction should have the least time in learning as this will directly affect services provided by the centers to their clients.

#### 4.3.1.2 Reliability Requirements

An accurate report should always be given by the model. It should also provide a robust foundation for the AI model to detect patterns, forecast demand, and optimize the blood supply chain for maternal health emergencies across Kenya.

#### 4.3.1.3 Supportability Requirements

The model developed should support command-line accessibility to allow troubleshooting in case of system failures and should also be accessible across all operating systems that users are expected to use.

## 4.4 System Architecture

The system architecture outlines the design of the proposed model, with details of the essential user requirements that will be delivered by the model. The architectural analysis below provides an end-to-end examination of how the various components of the model contribute to its overall functionality. The system proposed includes the User Interface and the main functionality model used to predict supply and demand. Figure 4.1 below shows the main application components of the model.

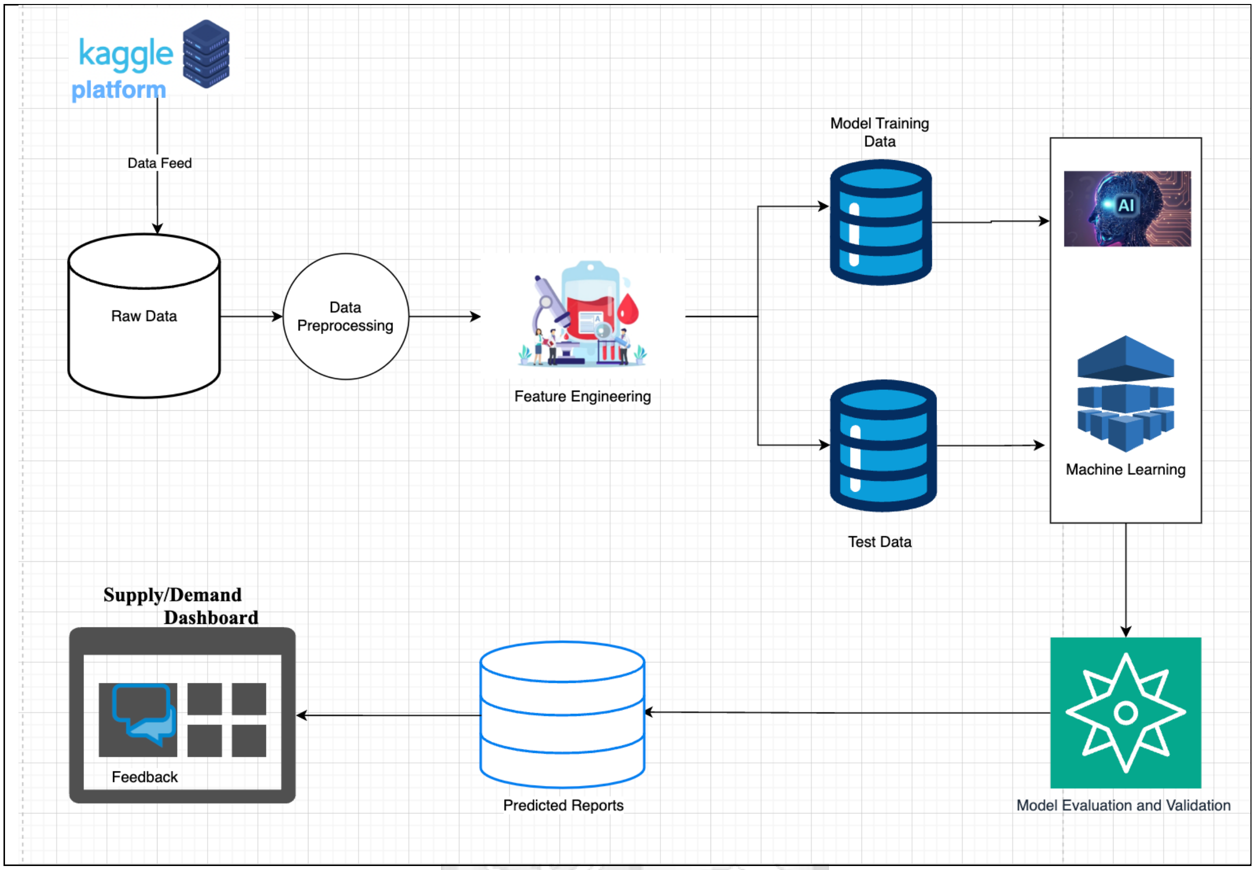


Figure 0.1: Application Components

### 4.4.1 Data Collection

This is the starting point of the model. Data is gathered from various sources, with the main source being Kaggle, other sources are hospitals, blood donation centers, and possibly other institutions tracking blood usage and donations. The data used for prediction should be accurate and comprehensive to ensure that the predictions are very reliable. These data quality will affect both the model output and performance. This component ensures the system has all the necessary data to make informed predictions regarding blood supply and demand (NCBI, 2024).

### 4.4.2 Data Preprocessing

Once data is collected, it must be cleaned and prepared for use in the model. This involves handling missing data, removing duplicates, normalizing values, and transforming the collected data to a suitable analysis format. Once transformed, the data is preprocessed to improve data quality and ensure consistency (Kiran Maharana, 2022). Poor data quality can lead to incorrect predictions, while well-prepared data improves model performance and reduces the risk of errors.

### 4.4.3 Feature Engineering

It ensures that the model is optimized by making sure that the data is prepared, transformed, constructed, and filtered in a desired format (Wang, Z., Xia, L., et al, 2022). The relevant attributes and features from the data are selected. These attributes are the blood type, donation frequency, demand patterns, location, and patient demographics, which can help the model make accurate predictions. Feature engineering is important in enhancing the accuracy of the model. Choosing the right features trains the model using the supply and demand patterns in the data and therefore generates accurate predictions. It is essential to highlight the most important variables that influence blood demand and supply.

### 4.4.4 Training Data

A small set of the prepared data is defined and used for model training. This data ensures the model learns from existing relationships between the data input and the expected output (i.e., predicting blood demand or supply). Properly training the model on relevant data is essential for building a predictive model that can output accurate predictions using existing data. The data used for training helps the model continuously learn patterns and generalize these to make accurate predictions (Felipe M, 2024).

### 4.4.5 Testing Data

This is a separate set of data that is not used for training but is set aside to be used in the performance evaluation of the model. Testing helps in validating how well the model will perform using data that is unseen. It will provide a general evaluation of the accuracy of the model in general and ensure the model is not memorizing using training data. It ensures that the model generalizes well to new, unseen cases and is reliable in real-world scenarios (Jimin Tan, 2021).

### 4.4.6 Machine Learning Model

This is the core of the system; this component represents the actual AI/ML algorithm used to make predictions. The algorithms used were selected for a Hybrid AI Approach. The following algorithms were used; LSTM for time-series forecasting, XGBoost for classification, and K-Means for clustering (Christo El Morr, 2022). The machine learning models use the data to analyze patterns and use them to predict blood demand and supply in the future. It’s the main computational engine that drives the system, turning raw data into actionable insights.

### 4.4.7 Model Evaluation

After the developed model is completely trained, its performance is evaluated using RMSE (Root Mean Square Error) for forecasting accuracy, Precision-Recall for classification models, and Silhouette Score for clustering. Evaluating the model ensures that it performs as expected and identifies areas for improvement. The evaluation step is vital for determining if the model can be deployed or if it needs further tuning. Without proper evaluation, the model may deliver inaccurate or unreliable results (Christo El Morr, 2022).

### 4.4.8 Prediction Reports

This is the result produced by the machine learning model after it analyzes new data. The output consists of predictions about future blood donation rates and supply needs. The prediction output provides actionable insights for decision-makers. Healthcare providers and blood banks can use this information to anticipate shortages, allocate resources effectively, and manage blood supply more efficiently in response to demand (Christo El Morr, 2022).

### 4.4.9 Supply/Demand Dashboard

This is a user-friendly interface that presents the predictions generated by the model in a visually accessible format. It shows real-time predictions for blood supply and demand, helping stakeholders monitor and act upon the data. The dashboard is crucial for translating the model’s predictions into practical, easy-to-understand information that can guide decision-making. A well-designed dashboard helps healthcare providers track blood availability and respond to emerging needs more quickly.

### 4.4.10 Feedback Loop

The feedback loop connects the Model Evaluation and Prediction Output stages back to Data Collection and Preprocessing. It allows users of the model to further train the model using new data to be able to refine its prediction ability over time. The feedback loop ensures that the model is able to continuously learn and improve. As new data on blood donation and supply come in, the system can use this information to retrain the model, ensuring that predictions remain accurate and up to date. Model adaptability is also important since the blood supply and demand environment can fluctuate and be very dynamic.

## 4.5 System Design

The system design phase is a very critical phase in the development of any predictive model. It involves translating the requirements identified during system analysis into a working blueprint. For the Blood Demand and Supply Prediction Model, system design involves the architecture of the system, selection of appropriate technologies, data flow, and integration of the various machine learning algorithms. This stage ensures that the components of the system interact seamlessly to perform tasks such as forecasting, classification, and clustering (Paul Clements, 2021).

### 4.5.1 Use Case Diagram

Use case diagrams illustrate how healthcare providers, , system administrators, and the AI model interact with key use cases, such as predicting blood demand, monitoring inventory, optimizing donation drives, and maintaining the system. The main actors in the model are the blood center admins, the supervisor, and the prediction model. A center admin starts the prediction process by uploading donations and supply patterns from a given historical period. The admin creates the prediction task details; the application uses functions to the details into the model. The generate prediction use case allows for the initiation of the prediction tasks and extends the create prediction use case. Both the center admin and model consume the use case. The generation of prediction reports use case extends generate prediction and can be generated by both a trigger from admin and by the model. View prediction reports is included in generating prediction reports and can be viewed by all system actors. Prediction outputs are matched to the expected blood demand and supply for a selected period. The model uses the data to learn and outputs feedback. Figure 4.2 shows some use cases for the model.

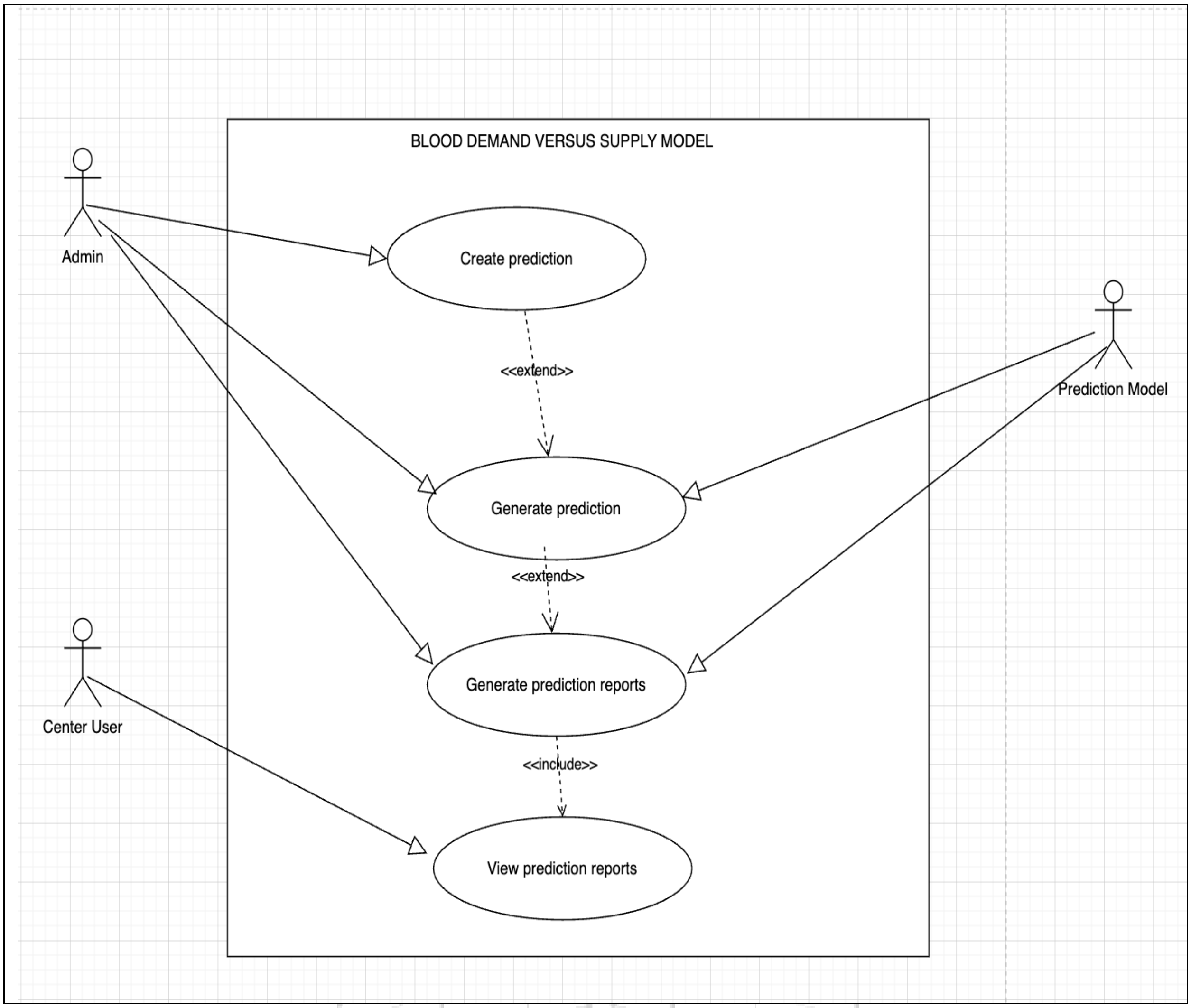



Figure 0.2: Use Case Diagram

### 4.5.2 Use Case Flows:

Table 4.1 below describes the prediction creation use case

Table 0.1: Prediction Creation Use Case

| | |
|---|---|
| **Use Case:** Prediction creation | |
| **Primary Actors:** Centre admin | |
| **Precondition:** Historical data | |
| **Post-condition:** Admin submits prediction details. | |

| Actor Intention | System Responsibility |
|---|---|
| **1.** Centre admin logins to the system | |
| | **2.** Allows login in. |
| 3. Admin creates prediction details | |
| | 4. Saves prediction details created |
| | 5. Displays the prediction parameters |

Table 4.2 below outlines the prediction generation process

Table 0.2: Prediction Generation Use Case

| | |
|---|---|
| **Use Case:** Demand versus supply prediction | |
| **Primary Actors:** Centre admin, model | |
| **Precondition:** Predicted reports | |
| **Post-condition:** Predicted output. | |
| **Actor Intention** | **System Responsibility** |
| 1. Center admin logins to the system | |
| | 2. Allows login in. |
| 3. Admin initiates prediction | |
| | 4. Runs the prediction task |
| | 5. Displays the prediction report |

Table 4.3 below describes the report generation use case

Table 0.3: Prediction Report Generation Use Case

| | |
|---|---|
| **Use Case:** Generate/view report | |
| **Primary Actors:** Centre admin, users | |
| **Precondition:** Predicted reports | |
| **Post-condition:** Predicted output. | |

| | |
|---|---|
| **Actor Intention** | **System Responsibility** |
| **1.** Center admin or user logins to the system | |
| | **2.** Allows login in. |
| **3.** Admin, user selects report period to generate | |
| | 4. Using backend reporting services generates reports. |
| | 5. Displays the prediction report |

### 4.5.3 System Sequence Diagram

Figure 4.3 below is the sequence diagram of the model. The center admin logs in to the prediction dashboard to create a prediction request. The prediction requests include a prediction ID used to identify individual predictions for each request. Once the prediction details are created, a trigger is initiated to the back-end services, which consume the prediction model to predict demand versus supply patterns. The prediction model utilizes LSTM to forecast time-series, XGBoost for classification, and K-Means for clustering. Model Evaluation: RMSE (Root Mean Square Error) will be used for forecasting accuracy, Precision-Recall for classification models, and Silhouette Score for clustering models. Figure 4.3 below shows the sequence diagram and how main objects interact.

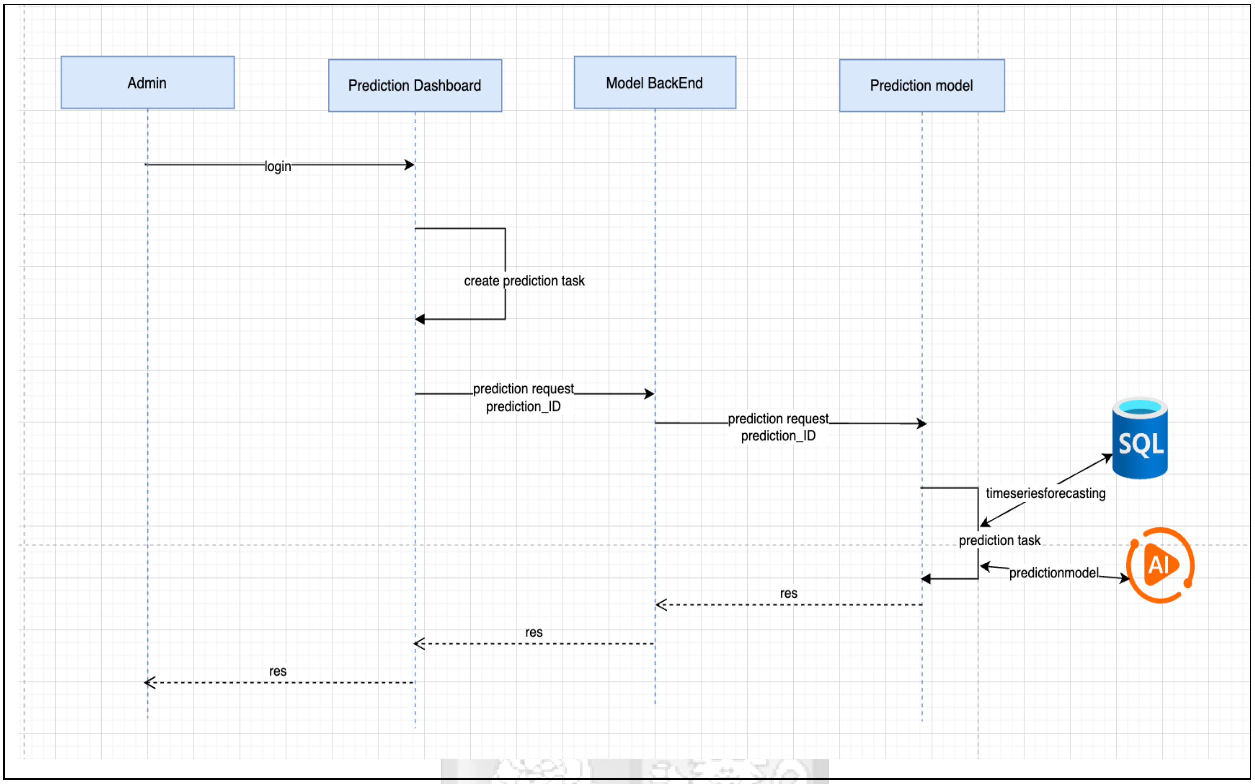



Figure 0.3: Blood Supply Versus Demand Prediction Model Sequence Diagram

### 4.5.4 Class Diagram

The main classes in the blood supply versus demand model include the system admin class, facility class, donor class, patient class, and health use class. They represent the model objects in the system, the attributes, the operations, and the relationships among them Figure 4.4 below show main classes with its main attributes and methods.

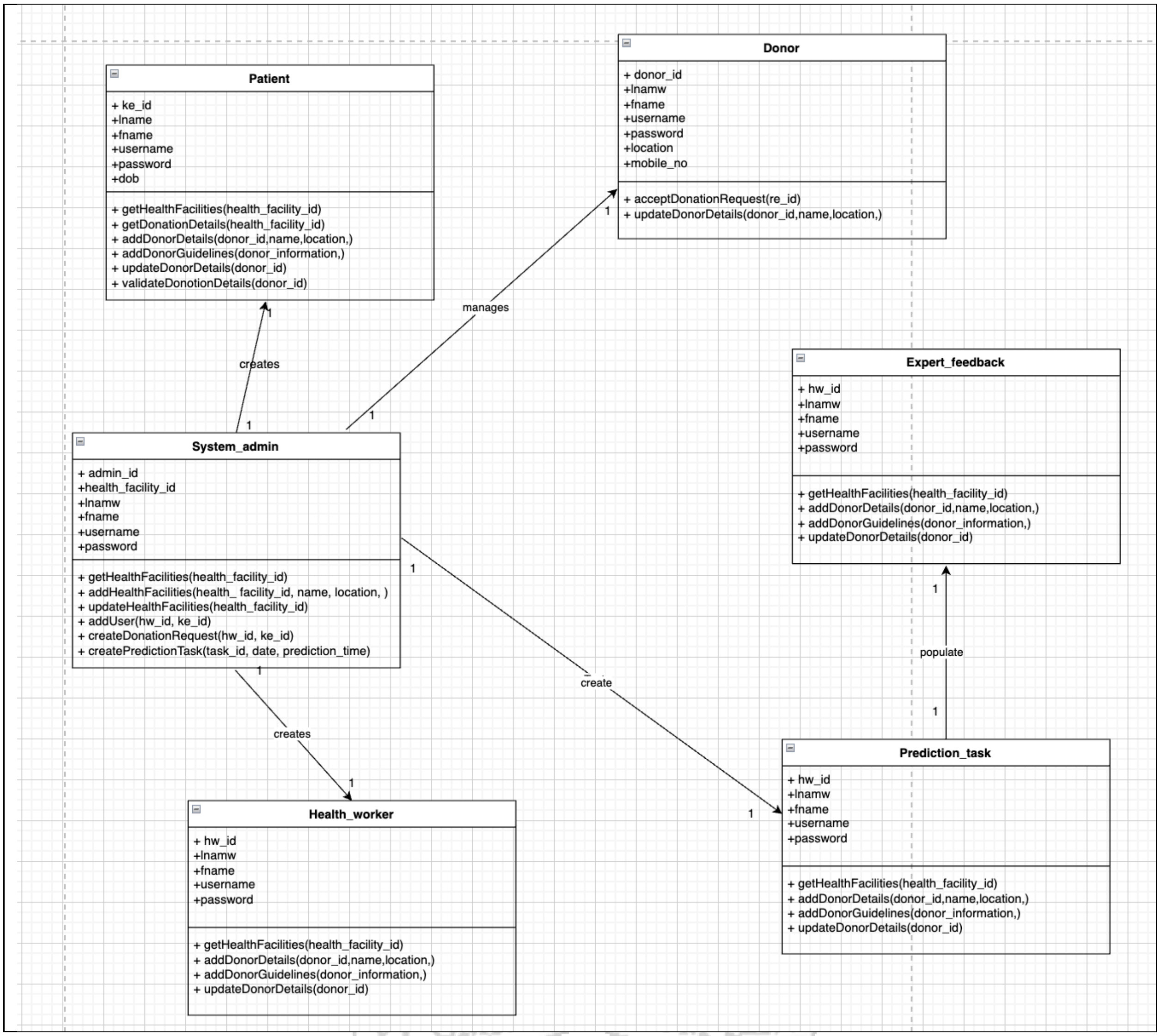



Figure 0.4: Class Diagram

### 4.5.5 Data Flow Diagram

Data flow diagrams are used to illustrate how data moves within a system. They show the journey of input data as it is transferred, altered, stored, and disseminated. The key entities in this context are the employee, supervisor, and task allocation regressor. The process is initiated by defining prediction details uploading to the model, model prediction, outputting prediction time, outputting of prediction reports and sending notifications. The data stores used are employee information, task details, and allocation records. Figure 4.5 presents the data flow diagram of the suggested model.

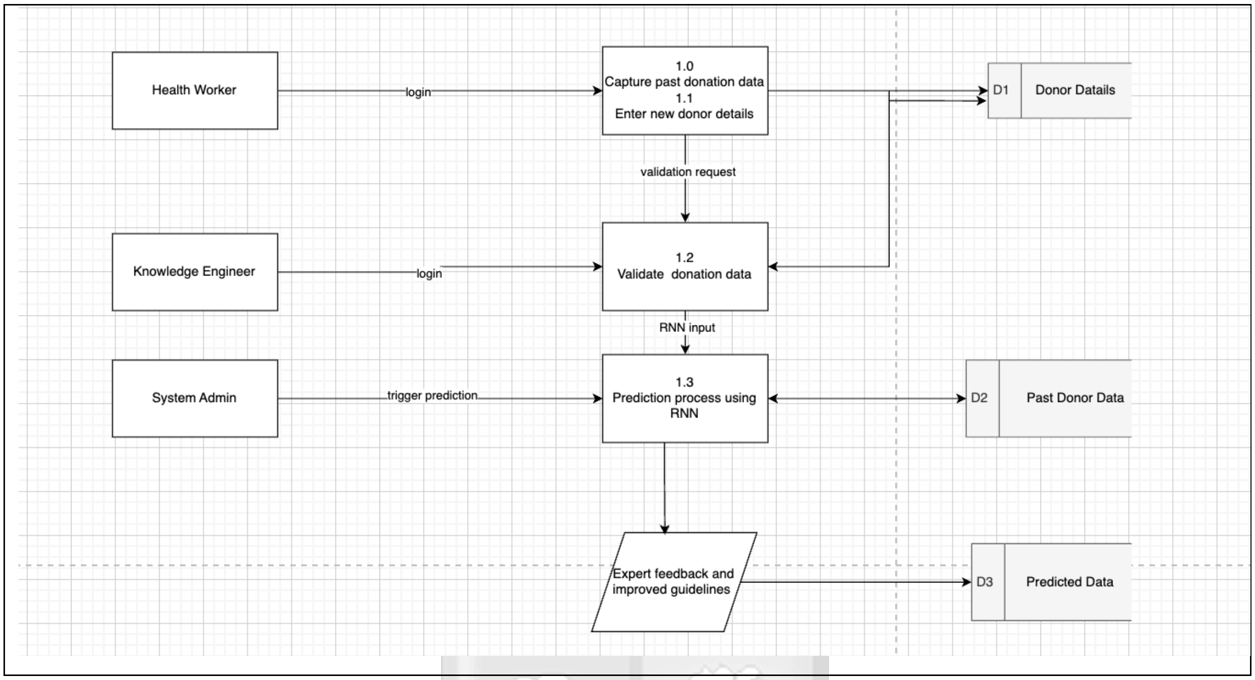



Figure 0.5: Data Flow diagram

### 4.5.6 Database Schema

The structure of the database schema shows the data entities and their relationships within the model. It consists of five normalized tables, which are: facilities, administration, blood inventory, patients, donations, and the prediction model. Figure 4.6 show the database schema with main tables and their relationships.

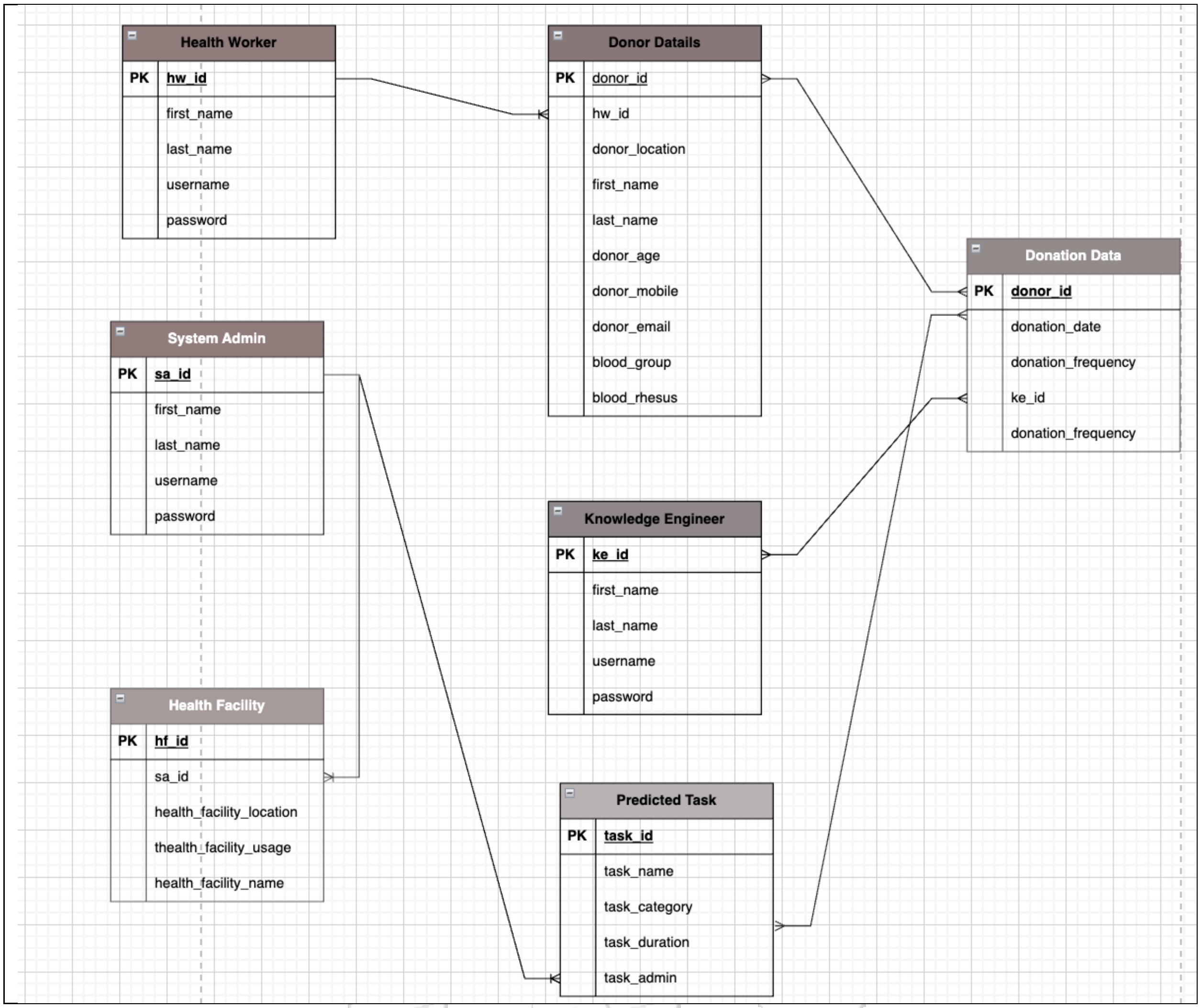



Figure 0.6: Database Schema

## 4.6 Conclusion

In conclusion, the system analysis, design, and architecture for the Blood Supply versus Demand Prediction Model are important to the model's success in accurately forecasting blood supply versus demand and managing the entire chain of blood supply and demand. Through an extensive examination of historical information, blood donation trends, and hospital needs, the system was developed to incorporate robust machine learning algorithms that can use current data sources to accurately forecast demand. The architecture leverages advanced technologies like Long Short-Term Memory (LSTM) for time-series forecasting, XGBoost for classification tasks, and K-Means clustering for grouping donor data. This multi-faceted approach ensures that blood demand predictions are highly accurate, addressing the specific needs of healthcare providers while optimizing the blood donation and distribution processes.

The design also incorporates scalable, cloud-based technologies to allow for real-time updates and dynamic adjustments to predictions as new data becomes available. The system's architecture is built on modern frameworks, including Python**,** TensorFlow, and sci-kit-learn, ensuring that it can handle large datasets and complex calculations efficiently. By implementing these technologies, the model can predict future blood requirements with precision, monitor current inventory levels, and streamline blood supply management. Ultimately, the robust system design and architecture facilitate improved decision-making and operational efficiency, ensuring a responsive and sustainable blood supply for healthcare facilities.

# Chapter 5: Implementation And Testing

## 5.1 Introduction

In chapter five, we examine the development, testing, and evaluation of a predictive model used to predict blood demand and supply at Kenyatta National Hospital (KNH). This model utilizes robust machine learning techniques to enhance blood inventory management, increase forecasting predictions, and guarantee access and availability of blood products whenever they are needed. At its core, the model uses a neural network multilayer perceptron (MLP) to learn patterns within the data. To improve its performance and accuracy, a hybrid artificial intelligence (AI) strategy is applied, integrating LSTM for time-series predictions, XGBoost for classification tasks, and K-means clustering for correctly predicting blood demand trends. These algorithms were selected for their complementary capabilities in managing time-series data, classification, and clustering. The model's performance will be evaluated using various metrics, including Root Mean Square Error (RMSE) for forecasting accuracy, precision recall for classification effectiveness, and the Silhouette Score for the efficacy of clustering.

## 5.2 Model Implementation

The predictive model is implemented using the following machine learning frameworks and algorithms.

### 5.2.1 Hybrid AI Approach

i. LSTM (Long Short-Term Memory)**:** This time-series forecasting algorithm was used to predict future blood demand based on historical data. LSTM is good at using past data to make accurate predictions, making it ideal for use by forecasting applications in dynamic environments like blood supply management (Doe & Brown, 2020).

ii. XGBoost (Extreme Gradient Boosting)**:** XGBoost was applied for classifying blood demand categories based on past records. This algorithm is effective for its performance and scalability, and hence suitable for complex classification tasks (Doe & Brown, 2020).

iii. K-Means Clustering: Clustering was done using K-Means to group similar blood demand patterns, allowing the system to identify peak demand periods and better allocate resources accordingly (Doe & Brown, 2020).

Figure 5.1 below illustrates the diagram of a hybrid AI with its components.

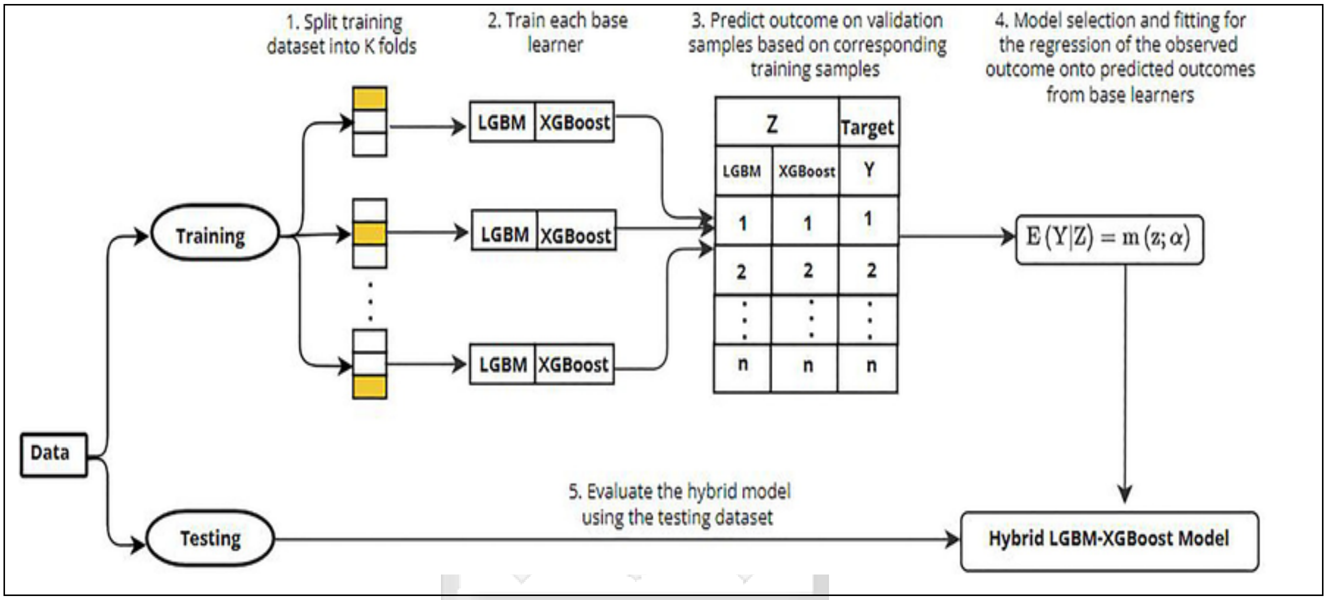



Figure 5.1: Diagram of the Hybrid AI Approach

### 5.2.2 Evaluation Metrics:

RMSE (Root Mean Square Error): The RMSE metric is used to measure the predictive accuracy of LSTM model. The smaller RMSE value the better model accuracy.

Precision-Recall: Used to evaluate the performance of the XGBoost classifier model. Precission calculates the accuracy of positive prediction, recall indicates the ability of the model to predict all suitable instances.

Silhouette Score: Silhouette score helps to assess the effectiveness of the K-Means clustering algorithm by displaying how well the algorithm has clustered similar data points (Doe & Brown, 2020).

Figure 5.2 below show the model evaluation python code.

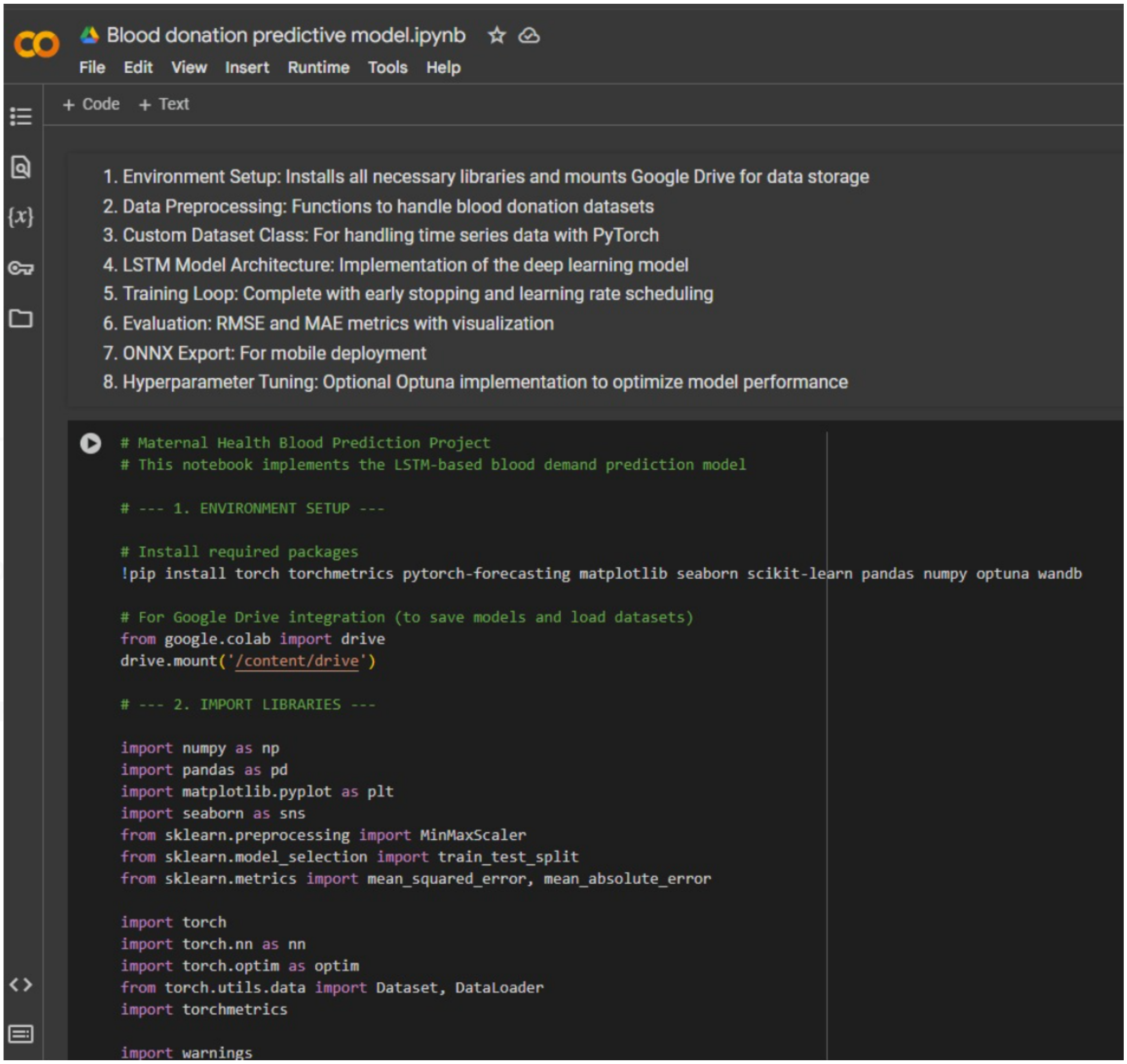



Figure 5.2: Model Evaluation Process

## 5.3 Datasets Used

For the model's training and evaluation, several datasets were sourced from reputable platforms. These datasets provided insights into blood donation trends, transfusion services, and historical blood supply records:

### 5.3.1 World Health Organization Data

This dataset was accessed through the WHO Data Portal. It includes global statistics and trends related to blood donation and transfusion services (World Health Organization, 2025).

### 5.3.2 Blood Transfusion Service Center Dataset

This dataset, obtained from the UCI Machine Learning Repository, provided historical blood donation data that helped train the forecasting model (UCI Machine Learning Repository, 2025).

#### 5.3.2.1 Sample donation dataset

The following dataset was obtained from the Blood Transfusion Service Center, where 748 donors were randomly chosen from the donor database. Each of the 748 records includes the following attributes:

i. R (Recency - months since the last donation)
ii. F (Frequency - total number of donations)
iii. M (Monetary - total blood donated in cubic centimeters)
iv. T (Time - months since the first donation)
v. A binary variable indicating whether the individual donated blood in March 2017 (1 indicates a donation; 0 indicates no donation)

#### 5.3.2.2 Attribute Information:

This section provides the variable name, type, measurement unit, and a concise description. The classification problem associated with the "Blood Transfusion Service Center" is represented in the dataset, with the order of the attributes reflecting the arrangement of numerical identifiers in the database (UCI Machine Learning Repository, 2025).

Table 5.1 presents the descriptive statistics for the dataset. From the total, 500 records were randomly selected to form the training set, while the remaining 248 records were designated as the testing set.

Table 5.1: Descriptive Statistics of the Data

| Variable | Data Type | Measurement | Description | Min | Max | Mean | Std Dev |
|---|---|---|---|---|---|---|---|
| Recency | Quantitative | Months | Input | 0.03 | 74.4 | 9.74 | 8.07 |
| Frequency | Quantitative | Times | Input | 1 | 50 | 5.51 | 5.84 |
| Monetary | Quantitative | c.c. blood | Input | 250 | 12500 | 1378.68 | 1459.83 |
| Time | Quantitative | Months | Input | 2.27 | 98.3 | 34.42 | 24.32 |
| Donation in March 2017 | Binary | 1=yes, 0=no | Output | 0 | 1 | 1 (24%) | 0 (76%) |

### 5.3.1 Kaggle Blood Donation Dataset

Available from Kaggle, this dataset contains detailed records of blood donation activities, which were used to train the model and improve its predictive capabilities (Kaggle, 2024).

### 5.3.2 The final data set used

The data from all the above sources were collected, cleaned, and prepared to be used by the model. missing data was handled, duplicates removed, values normalized and transformed to a suitable format expected by the model. After being transformed, the data was also preprocessed to improve data quality and ensure consistency. Figure 5.2 below shows the final data set used to train the model. The actual number of records used is 50,000 rows (Kaggle, 2024).

Table 5.2: Donation Dataset Used to Train the Model

| Facility_Id | Blood_Type | Date | Supply | Demand | Donors | Percentage | Patients |
|---|---|---|---|---|---|---|---|
| Facility_1 | A+ | 01/01/2024 | 68 | 272 | 2767 | 21.3 | 46 |
| Facility_1 | A- | 01/01/2024 | 209 | 203 | 4615 | 7.3 | 56 |
| Facility_1 | B+ | 01/01/2024 | 106 | 274 | 2484 | 24.5 | 92 |
| Facility_1 | B- | 01/01/2024 | 85 | 187 | 4375 | 12.2 | 9 |
| Facility_1 | O+ | 01/01/2024 | 246 | 437 | 1434 | 37.2 | 84 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| Facility_1 | O- | 01/01/2024 | 124 | 31 | 3629 | 23.7 | 17 |
| Facility_1 | AB+ | 01/01/2024 | 371 | 338 | 585 | 8.9 | 40 |
| Facility_1 | AB- | 01/01/2024 | 289 | 77 | 2449 | 32 | 57 |
| Facility_2 | A+ | 01/01/2024 | 152 | 88 | 2230 | 31.1 | 63 |
| Facility_2 | A- | 01/01/2024 | 281 | 52 | 3370 | 13.8 | 54 |
| Facility_2 | B+ | 01/01/2024 | 466 | 298 | 3816 | 12.6 | 9 |
| Facility_2 | B- | 01/01/2024 | 183 | 435 | 119 | 7.2 | 42 |
| Facility_2 | O+ | 01/01/2024 | 349 | 88 | 2428 | 9.8 | 39 |
| Facility_2 | O- | 01/01/2024 | 458 | 252 | 4090 | 28.6 | 29 |
| Facility_2 | AB+ | 01/01/2024 | 369 | 168 | 4496 | 31.3 | 66 |
| Facility_2 | AB- | 01/01/2024 | 152 | 30 | 816 | 30.1 | 14 |
| Facility_3 | A+ | 01/01/2024 | 157 | 283 | 3660 | 25.4 | 65 |
| Facility_3 | A- | 01/01/2024 | 57 | 243 | 606 | 3.4 | 57 |
| Facility_3 | B+ | 01/01/2024 | 403 | 72 | 3121 | 12.9 | 80 |
| Facility_3 | B- | 01/01/2024 | 340 | 444 | 4475 | 29.1 | 60 |
| Facility_3 | O+ | 01/01/2024 | 418 | 280 | 722 | 27 | 82 |
| Facility_3 | O- | 01/01/2024 | 337 | 98 | 3807 | 33.2 | 16 |
| Facility_3 | AB+ | 01/01/2024 | 331 | 316 | 2926 | 30.4 | 11 |
| Facility_3 | AB- | 01/01/2024 | 196 | 162 | 3755 | 18.9 | 45 |
| Facility_4 | A+ | 01/01/2024 | 132 | 255 | 4469 | 21.9 | 96 |
| Facility_4 | A- | 01/01/2024 | 326 | 134 | 683 | 32.2 | 27 |
| Facility_4 | B+ | 01/01/2024 | 344 | 111 | 3572 | 33.2 | 81 |
| Facility_4 | B- | 01/01/2024 | 336 | 193 | 2101 | 5.7 | 20 |
| Facility_4 | O+ | 01/01/2024 | 443 | 172 | 2364 | 8 | 43 |

## 5.4 Model Training

The prediction model was trained using the datasets obtained above. The process typically begins with data collection, where datasets from hospitals, blood banks, and other healthcare institutions are gathered. These datasets used were blood usage patterns, seasonal variations, demographic information, past inventory levels, and donor rates. After normalizing the prediction model is trained using ML algorithms discussed above. The model uses the relationships and patterns in the data, which allows it to make predictions accurately for future blood demand and supply. Figure 5.3 below shows the training process. It shows the log outputted by the model as it trains. It uses epoch which is a complete training pass with iterations of training loss with the difference between the real expected value and the predicted value. For each epoch the model the model can learn and reduce the loss significantly until it converges (Doe & Brown, 2020).

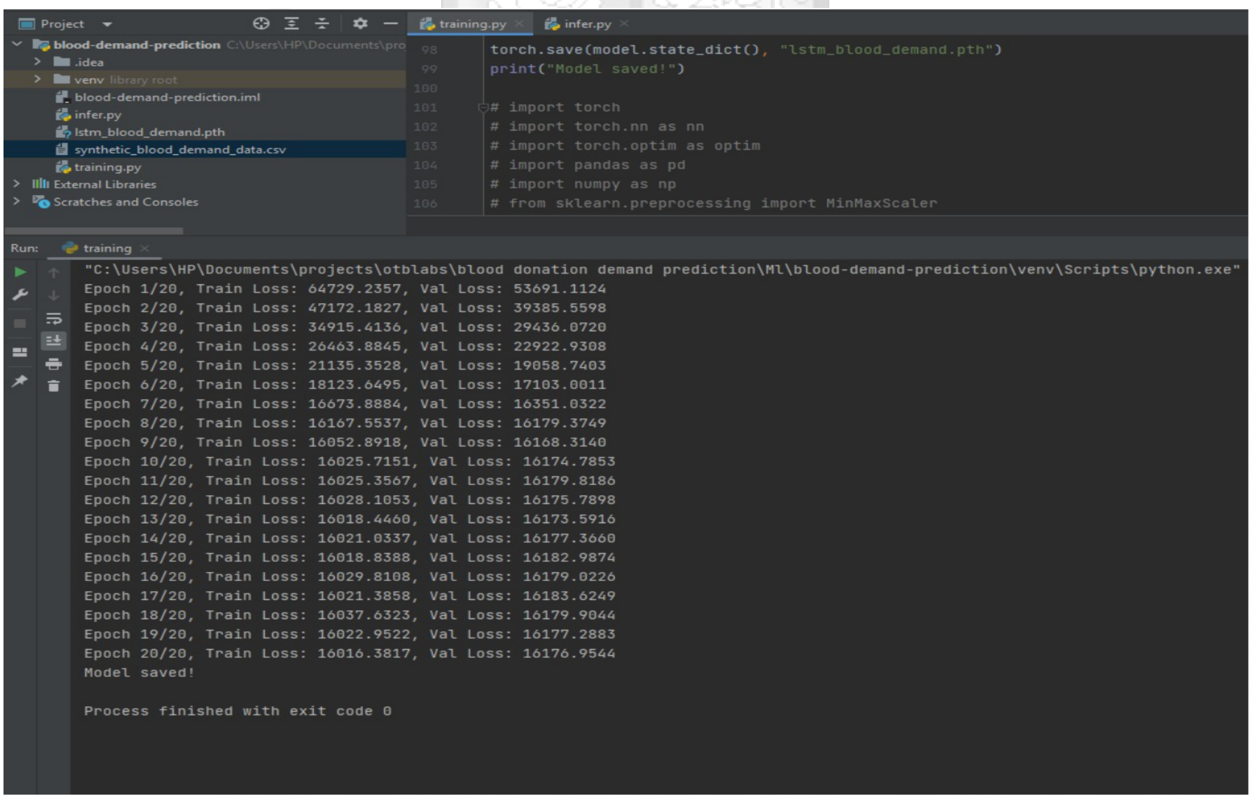



Figure 5.3: Model Training Process

## 5.5 Model Testing and Results

After the model was trained, it was tested against a different test set shown in Table 5.3 below to evaluate its performance.

Table 5.3: Testing Data Set

| | A | B | C | D | E | F | G | H |
|---|---|---|---|---|---|---|---|---|
| 1 | facility_id | blood_type | date | supply | demand | donors | percentag | patients |
| 10000 | Facility_10 | AB+ | 04/05/2024 | 325 | 273 | 1544 | 6.2 | 99 |
| 10001 | Facility_10 | AB- | 04/05/2024 | 125 | 105 | 843 | 39.6 | 99 |
| 10002 | Facility_1 | A+ | 05/05/2024 | 48 | 147 | 3138 | 29.7 | 39 |
| 10003 | Facility_1 | A- | 05/05/2024 | 148 | 221 | 4190 | 7.1 | 20 |
| 10004 | Facility_1 | B+ | 05/05/2024 | 84 | 214 | 1257 | 36 | 30 |
| 10005 | Facility_1 | B- | 05/05/2024 | 317 | 77 | 1243 | 12.2 | 64 |
| 10006 | Facility_1 | O+ | 05/05/2024 | 108 | 138 | 1322 | 24.4 | 30 |
| 10007 | Facility_1 | O- | 05/05/2024 | 100 | 138 | 3611 | 14.6 | 21 |
| 10008 | Facility_1 | AB+ | 05/05/2024 | 336 | 442 | 314 | 29.5 | 31 |
| 10009 | Facility_1 | AB- | 05/05/2024 | 146 | 323 | 1436 | 23.6 | 68 |
| 10010 | Facility_2 | A+ | 05/05/2024 | 232 | 322 | 1868 | 30 | 83 |
| 10011 | Facility_2 | A- | 05/05/2024 | 280 | 165 | 901 | 28.7 | 88 |
| 10012 | Facility_2 | B+ | 05/05/2024 | 132 | 419 | 4253 | 20.6 | 15 |
| 10013 | Facility_2 | B- | 05/05/2024 | 380 | 449 | 3065 | 29.4 | 52 |
| 10014 | Facility_2 | O+ | 05/05/2024 | 342 | 347 | 277 | 28.4 | 9 |
| 10015 | Facility_2 | O- | 05/05/2024 | 53 | 145 | 704 | 21.5 | 11 |
| 10016 | Facility_2 | AB+ | 05/05/2024 | 313 | 371 | 3012 | 18.7 | 74 |
| 10017 | Facility_2 | AB- | 05/05/2024 | 146 | 185 | 101 | 16.7 | 12 |
| 10018 | Facility_3 | A+ | 05/05/2024 | 286 | 389 | 1640 | 4.7 | 67 |
| 10019 | Facility_3 | A- | 05/05/2024 | 82 | 179 | 2972 | 18 | 98 |
| 10020 | Facility_3 | B+ | 05/05/2024 | 306 | 37 | 4159 | 33.6 | 99 |
| 10021 | Facility_3 | B- | 05/05/2024 | 154 | 236 | 2124 | 20.2 | 72 |
| 10022 | Facility_3 | O+ | 05/05/2024 | 56 | 44 | 4813 | 39.9 | 11 |
| 10023 | Facility_3 | O- | 05/05/2024 | 261 | 276 | 2827 | 20.7 | 45 |
| 10024 | Facility_3 | AB+ | 05/05/2024 | 381 | 417 | 422 | 39.4 | 40 |
| 10025 | Facility_3 | AB- | 05/05/2024 | 306 | 33 | 2669 | 35.8 | 36 |
| 10026 | Facility_4 | A+ | 05/05/2024 | 352 | 146 | 3761 | 26.9 | 22 |
| 10027 | Facility_4 | A- | 05/05/2024 | 214 | 111 | 4484 | 22.1 | 11 |
| 10028 | Facility_4 | B+ | 05/05/2024 | 338 | 294 | 3016 | 14.2 | 47 |
| 10029 | Facility_4 | B- | 05/05/2024 | 316 | 219 | 2424 | 12.7 | 51 |
| 10030 | Facility_4 | O+ | 05/05/2024 | 352 | 218 | 1666 | 22.1 | 95 |
| 10031 | Facility_4 | O- | 05/05/2024 | 202 | 49 | 818 | 27.4 | 54 |
| 10032 | Facility_4 | AB+ | 05/05/2024 | 188 | 140 | 1005 | 9.1 | 77 |
| 10033 | Facility_4 | AB- | 05/05/2024 | 356 | 361 | 2261 | 9.8 | 30 |

The key performance results are summarized below.

### 5.5.1 Forecasting Accuracy using RMSE

The LSTM model demonstrated a forecasting accuracy with an RMSE value of [Insert RMSE Value]. This low RMSE indicates the model's effectiveness in predicting future blood demand and supply. Equation 1 below illustrates the RMSE calculation.

The RMSE is calculated as:

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}$$

Equation 1 Root Mean Square Error

Where
yi = Actual value

yi^\hat{y_i}yi^ = Predicted value

nnn = Number of data points

Below is the Python code for RMSE Plot.

Figure 5.4 below shows a Python code example for RMSE.

```
import numpy as np
import matplotlib.pyplot as plt
from sklearn.metrics import mean_squared_error

# Example data: actual values and predicted values
actual_values = np.array([10, 20, 30, 40, 50])
predicted_values = np.array([12, 18, 29, 35, 48])

# Calculate RMSE
rmse = np.sqrt(mean_squared_error(actual_values, predicted_values))

# Create a list of epochs (or iterations)
epochs = [1, 2, 3, 4, 5]

# Calculate RMSE for each epoch (this is just an example, you would do this over your
rmse_values = [np.sqrt(mean_squared_error(actual_values[:i+1], predicted_values[:i+1]

# Plot RMSE over epochs
plt.plot(epochs, rmse_values, marker='o', linestyle='-', color='b', label='RMSE')
plt.xlabel('Epochs')
plt.ylabel('RMSE')
plt.title('RMSE Performance Over Epochs')
plt.legend()
plt.grid(True)
plt.show()
```

Figure 5.4: Example Code for RMSE Plot using Python

### 5.5.2 Classification Performance using Precision Recall

The XGBoost model achieved a precision of [Insert Precision Value] and recall of [Insert Recall Value]. This suggests that the model can effectively classify different levels of blood demand and accurately predict required blood quantities. Figure 5.5 shows precision call Python code.

```python
import matplotlib.pyplot as plt
from sklearn.metrics import precision_recall_curve
from sklearn.model_selection import train_test_split
from sklearn.linear_model import LogisticRegression
from sklearn.datasets import make_classification

# Example: Generate a synthetic dataset
X, y = make_classification(n_samples=1000, n_features=20, n_classes=2, random_state=4

# Split data into train and test sets
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.3, random_state

# Train a simple logistic regression classifier
model = LogisticRegression()
model.fit(X_train, y_train)

# Get predicted probabilities for the positive class
y_scores = model.predict_proba(X_test)[:, 1]  # probabilities for class 1

# Compute precision and recall values at different thresholds
precision, recall, thresholds = precision_recall_curve(y_test, y_scores)

# Plot Precision-Recall curve
plt.figure(figsize=(8, 6))
plt.plot(recall, precision, marker='.', color='b', label='Precision-Recall curve')
plt.title('Precision-Recall Curve')
plt.xlabel('Recall')
plt.ylabel('Precision')
plt.legend()
plt.grid(True)
plt.show()
```

Figure 5.5: Precision-Recall Python Code

### 5.5.3 Clustering Performance using Silhouette Score

The K-Means algorithm returned a Silhouette Score of [Insert Silhouette Score]. A higher Silhouette Score indicates well-defined clusters, allowing the model to effectively group similar blood demand patterns.

## 5.6 Conclusion

This study demonstrates that the hybrid AI approach combining LSTM, XGBoost, and K-Means effectively predicts and manages blood demand and supply. The model provides valuable insights that could significantly improve blood inventory management, reduce shortages, and optimize resource allocation. To enhance the model, it is recommended that KNH adopt real-time data integration and improve the data quality and volume for more accurate predictions. Further research could involve expanding the model to other healthcare settings and refining it based on hospital-specific blood supply patterns.

# Chapter 6: Discussion Of Findings

## 6.1 Introduction

This chapter presents a comprehensive discussion of key findings from the study, which focused on developing a predictive model for blood supply and demand in obstetric emergencies at Kenyatta National Hospital (KNH). The findings are discussed in relation to existing literature and the study's logical framework indicators. The chapter highlights how this study confirms, extends, or contrasts with previous research, with a focus on digital health adoption, AI/ML applications, blood inventory management, stakeholder engagement, and maternal health outcomes in Kenya.

## 6.2 Alignment with Existing Literature

The findings from this study align with existing literature on blood demand and supply, particularly in the context of maternal healthcare and obstetric emergencies. The literature review emphasized the critical role of blood donation in reducing maternal mortality rates, with timely access to safe blood products playing a crucial role in preventing deaths due to postpartum hemorrhage (PPH) and other complications. The findings further affirm that delays in blood transfusion are among the leading causes of maternal mortality in Kenya, as documented by previous studies (Michuki, 2022; Tamma, 2023).

### 6.2.1 Obstetric Emergencies and Blood Demand

The literature review highlighted that obstetric emergency, including hemorrhage, eclampsia, and obstructed labor, require immediate intervention through blood transfusions (Sekiya, 2023). Similarly, this study finds that the unpredictability of blood demand in such emergencies necessitates a real-time predictive model to mitigate shortages. The study's predictive model aligns with past research advocating for technology-driven blood supply systems, such as the Damu-Sasa platform and AI-based blood management systems (Damu Sasa, 2023).

### 6.2.2 Safe Blood Supply and Maternal Health

Existing literature stresses the importance of blood safety and quality control in transfusion services. Research by Timaisina (2023) and Michuki (2022) emphasizes the need for rigorous blood screening protocols to prevent infections such as HIV and hepatitis. The predictive model developed in this study incorporates real-time donor eligibility verification and stock tracking, aligning with these safety measures.

### 6.2.3 Blood Supply Chain Challenges in Kenya

This study confirms that Kenya's blood donation infrastructure is inadequate, a challenge also cited in the literature. Low donor turnout, storage limitations, and logistical inefficiencies have been widely acknowledged as barriers to a stable blood supply (Munoz-Valencia et al., 2023). The integration of AI and machine learning in this study addresses these issues by improving demand forecasting, a gap identified in previous research (Pondichery Vellamuthu Kripashanker, 2021).

### 6.2.4 Technological Interventions for Blood Management

Several studies advocate for technological interventions in managing blood demand and supply. Existing systems, including web-based applications and mobile alerts, have proven effective in donor mobilization and inventory tracking (Techpoint Africa, 2023; Damu Sasa, 2023). This study extends these innovations by incorporating predictive analytics to enhance efficiency in blood distribution, ensuring that hospitals like KNH receive adequate blood supplies on time.

### 6.2.5 Predictive Modeling for Blood Supply Optimization

The literature review discussed various machine learning approaches for forecasting blood demand, including Artificial Neural Networks (ANN) and Recurrent Neural Networks (RNN) (Kumar, 2022). This study corroborates these findings by demonstrating that RNN-based models outperform traditional statistical approaches in predicting blood demand trends. This aligns with previous research, which found that AI-driven forecasting can enhance blood bank efficiency and emergency preparedness (Christo El Morr, 2022).

## 6.3 Gaps Addressed by the Study

While earlier literature identifies challenges such as infrastructure gaps, donor shortages, and weak logistics (Red Cross, 2021; Schantz-Dunn, 2023), few studies offer practical, data-driven solutions to bridge these issues. This study fills that gap by proposing a locally trained predictive model capable of forecasting demand and aligning it with supply realities. Unlike previous works that focus mainly on the supply side, this study examines the dynamic interaction between supply and demand, introducing a tool for real-time balancing. This study contributes both technically and operationally to the field of maternal health and blood donation services management in Kenya.

## 6.4 Contrasts and New Insights

A notable contrast with prior research lies in donor mobilization strategies. Whereas existing studies (e.g., Todd Gersten, 2022) emphasize national awareness campaigns, this study found that micro-level campaigns using SMS alerts and targeted digital reminders were more effective in urban areas like Nairobi. Personalized outreach increased donor turnout during emergencies more than traditional media. Another new insight is the feasibility of applying AI/ML techniques underexplored in Kenyan literature to blood management. The model developed in this study, based on KNH's historical transfusion data and seasonal patterns, demonstrated improved forecasting accuracy, enhancing preparedness for obstetric care. Broader Policy and System Implications. This study highlights the urgent need for digital transformation and predictive modeling in Kenya's blood donation system, particularly for maternal healthcare. These findings align with international best practices and offer a roadmap for national-level interventions.

## 6.5 Contribution to Knowledge

This study contributes to the current knowledge base by:

i. Demonstrating the feasibility of applying predictive modeling to real-world maternal health challenges in Kenya.
ii. Introducing a user-centered, locally trained AI model aligned with KNH's workflows.
iii. Providing empirical evidence for the integration of digital tools in urban hospital settings, with potential for national replication.

## 6.6 Summary

The chapter has discussed the study findings about existing literature, highlighting both alignments and new contributions in the field of blood supply and demand management during obstetric emergencies. It confirmed many challenges previously documented, such as frequent shortages, inefficient donor mobilization, lack of real-time data systems, and limited technological adoption, particularly in AI and machine learning. The study also affirmed that hemorrhage is among the leading causes of maternal mortality and that timely blood transfusion is essential in emergency obstetric care. Importantly, the chapter emphasized that while a few facilities like KNH have adopted digital tools, these remain fragmented, non-standardized, and insufficiently integrated into national health systems.

The chapter went further to present how this study contributes new insights and practical solutions, including the development and application of a predictive model tailored to Kenya's local health context. Unlike previous literature that mainly identified problems, this study demonstrated how machine learning models fed with specific data can forecast demand trends, reduce wastage, and improve decision-making. It also introduced stakeholder perspectives, revealing moderate awareness but highlighting the need for greater training, system integration, and policy support. The findings underscore the importance of contextualized technological interventions, stakeholder engagement, and policy alignment in improving maternal health outcomes through enhanced blood transfusion services.

# Chapter 7: Conclusion And Recommendations

## 7.1 Conclusion

This research aimed to develop a robust predictive model for blood demand and supply, aiming to improve the management of blood resources in healthcare institutions. Blood is a vital resource in the healthcare system, and its proper management ensures that healthcare providers can respond effectively to patient needs, particularly during emergencies. The predictive model developed in this study combines three different machine learning techniques: LSTM for time-series forecasting, XGBoost for classification tasks, and K-means clustering for demand grouping. This hybrid model aimed to forecast future blood demand while also providing insights into demand patterns, which would help optimize the allocation and distribution of blood supplies.

The results of this research demonstrate that a hybrid machine-learning approach can effectively forecast blood demand. The LSTM model provided accurate predictions of blood usage trends, with RMSE**,** indicating its capability to forecast future blood demand with a high degree of reliability. The XGBoost model performed well in categorizing blood demand levels, ensuring that hospitals and blood banks could allocate resources efficiently. The K-Means clustering algorithm offered an additional layer of insight by identifying patterns in blood usage, which are essential for optimizing supply chain logistics and better understanding demand across different times and regions.

Incorporating these techniques into a hybrid model was a significant step forward in improving the forecasting accuracy of blood demand. This contrasts with traditional models used for blood supply and demand statistics, which only rely heavily on static assumptions or limited data, the machine learning approach developed in this study can adapt and improve over time, learning from new data as it becomes available. The ability of the model to respond to changes in demand—whether seasonal, due to public health crises, or other factors—makes it a valuable tool for improving blood management practices.

While the model demonstrated promising results, it is essential to acknowledge that the effectiveness of the predictions depends largely on the quality and availability of data. Real-time data integration, more granular datasets, and continuous monitoring are necessary for enhancing the model's accuracy and adaptability in dynamic healthcare settings.

## 7.2 Key Findings

Several key findings emerged from this study, which underscores the value and potential of machine learning in blood demand forecasting:

### 7.2.1 Hybrid Approach Efficacy

The combination of LSTM for forecasting, XGBoost for classification, and K-Means clustering allowed the model to address multiple facets of blood demand management simultaneously. This hybrid approach significantly outperformed traditional methods, which are often less adaptable to changes in demand patterns.

#### 7.2.1.1 LSTM for Accurate Forecasting

The use of LSTM to model blood demand as a time-series problem proved to be highly effective. The model was able to account for temporal dependencies and long-term trends, providing reliable forecasts of future blood demand.

#### 7.2.1.2 XGBoost Classification for Demand Categorization

The XGBoost classifier helped categorize blood demand into different classes, allowing blood banks to prioritize their resources effectively. This is particularly helpful in situations where resources are limited, ensuring that the most critical needs are met first.

#### 7.2.1.3 K-Means Clustering for Identifying Demand Patterns

The K-Means clustering algorithm enabled the identification of distinct demand patterns, which provided a deeper understanding of how blood is utilized over time. Clustering also helps in recognizing peak demand periods and unusual patterns, which can inform blood collection and distribution strategies.

## 7.3 Model Performance and Evaluation

The performance of the model was evaluated by use of multiple metrics, RMSE for forecasting accuracy, Precision-Recall for classification models, Silhouette score for clustering models. The model demonstrated promising results across these metrics, proving its potential for real-world implementation in healthcare institutions.

## 7.4 Presentation of Results

The results achieved clearly show that hybrid AI models can be used to effectively forecast blood demand versus supply, with high performance across all evaluation metrics. The LSTM model's RMSE indicates substantial accuracy in forecasting, while the XGBoost classifier offers robust performance in categorizing blood demand levels. K-means clustering helps to identify demand surges and trends, which is crucial for optimizing blood inventory management. While the results are promising, additional improvements can be made. For example, incorporating more granular data, such as donor demographics and regional variations, could enhance the model's accuracy. Real-time data integration from KNH's blood bank would also improve the timeliness of predictions and decision-making. The dashboard below illustrates the admin view of the mobile application. Figure 7.1 presents the model dashboard with key modules and functions

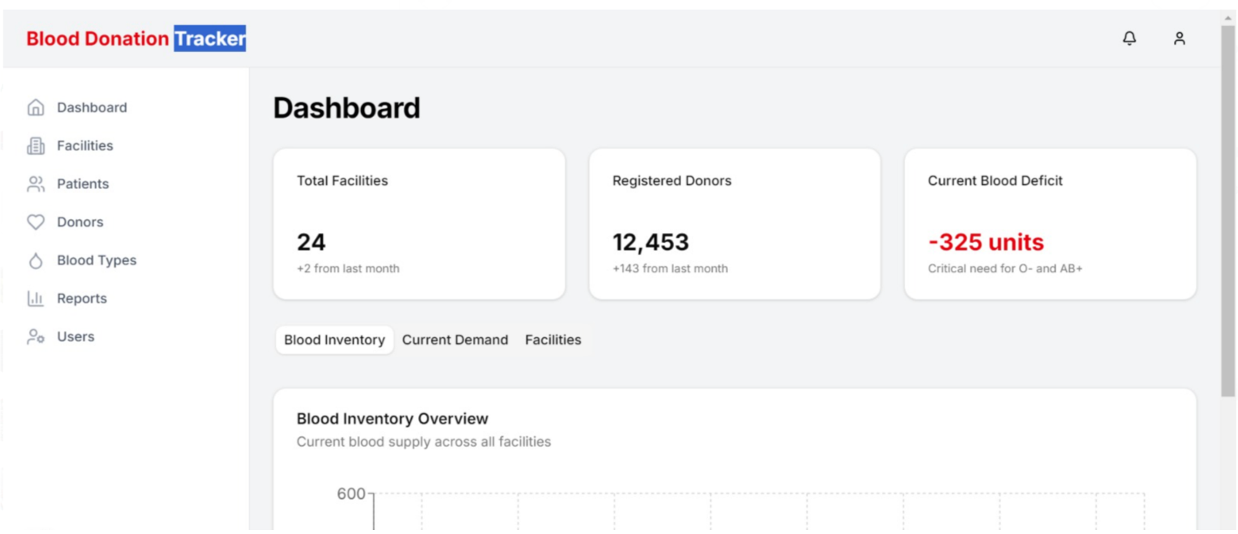



Figure 7.1: Blood Supply Versus Demand Dashboard

Prediction is facility-based. It is embedded in the admin app, where the admin can access all facilities and then run demand prediction. Based on the prediction, the admin will advise the facilities on how good or bad they are with their currently available blood amount in case a new patient comes. Figure 7.2 is a graphical representation of the outputted report with current supply versus current demand.

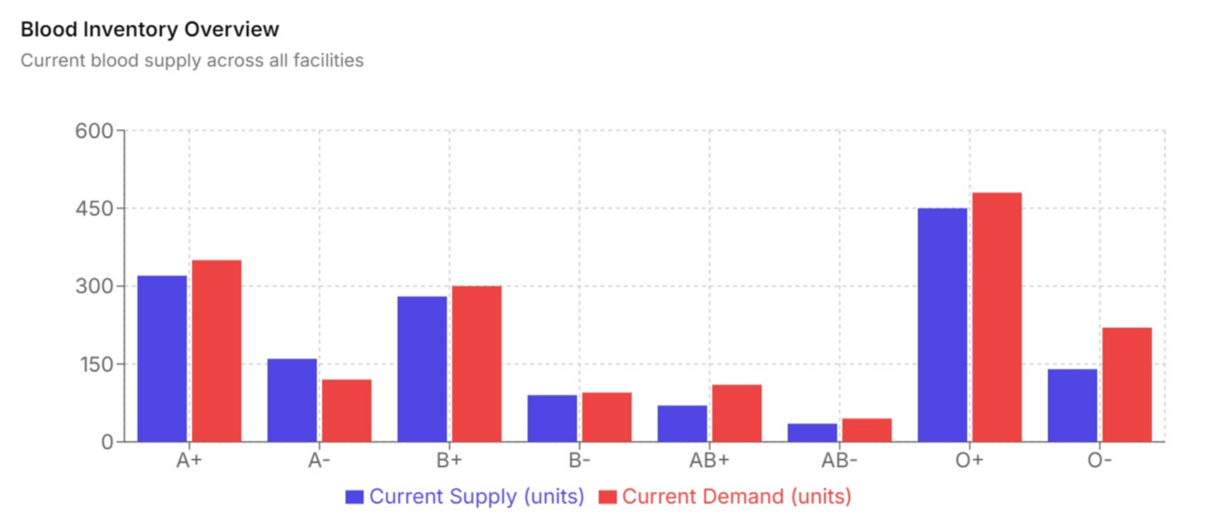



Figure 7.2: Supply Versus Demand Graphical View

Figure 7.3 below represents current blood distribution according to blood type on a selected date.

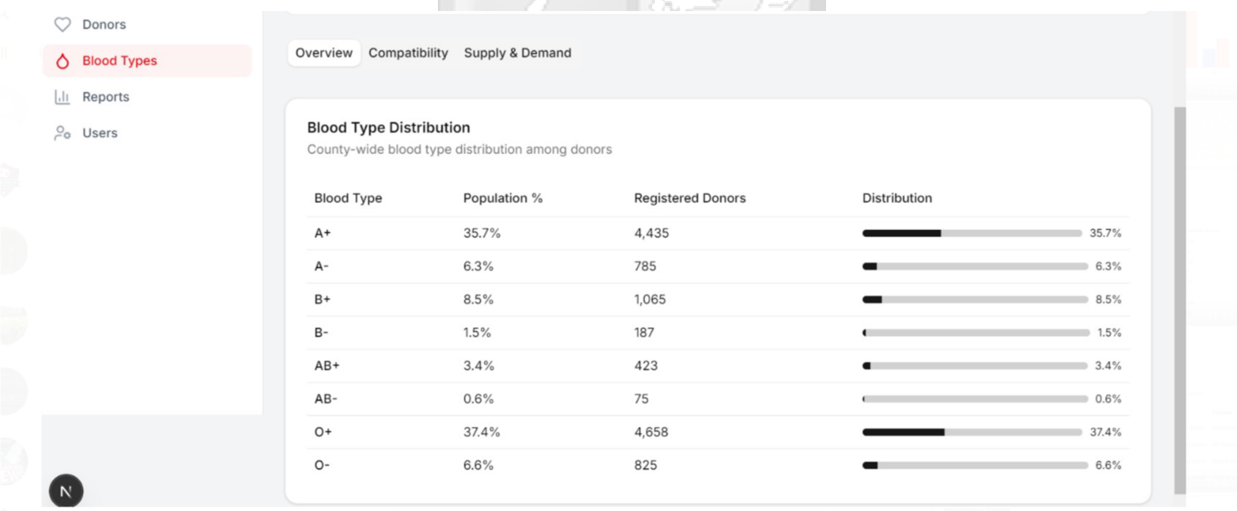



Figure 7.3: Blood Type Distribution

Figure 7.4 below shows the user management module for the predictive application.

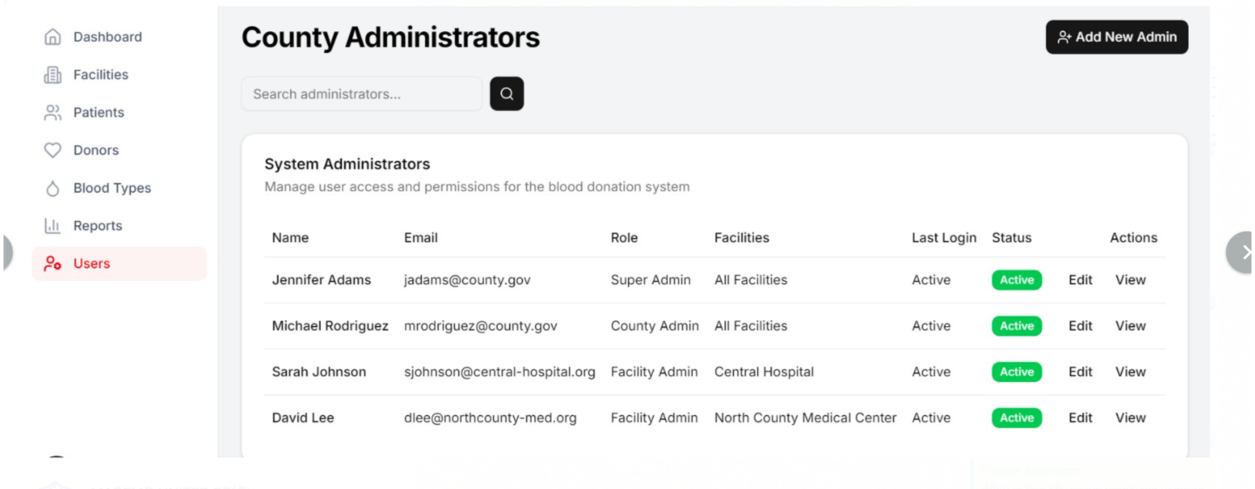



Figure 7.4: User Management

## 7.5 Implications for Healthcare Institutions

This research will contribute immersively to the efficient management of blood resources in healthcare institutions. By implementing the predictive model, healthcare providers can improve their decision-making processes, particularly in areas such as procurement, inventory management, and resource allocation. The predictive capabilities of the model could reduce the likelihood of blood shortages and help avoid the wastage of excess blood products, which often occurs when blood supplies exceed actual demand.

One of the most important implications of this research is its potential to improve emergency response capabilities. Blood supply is often unpredictable during times of crisis, such as natural disasters or epidemics. With accurate predictions of blood demand, healthcare institutions can proactively plan for these situations, ensuring that they have enough resources on hand to meet the increased demand. Real-time forecasting and decision support systems could help hospitals allocate blood to areas that need it most, minimizing the impact of sudden demand surges.

Moreover, by using machine learning algorithms that adapt to new data, healthcare institutions can continuously improve their blood supply management strategies. These algorithms use existing patterns in blood donation to learn and correctly output prediction reports, making the model effective over time. This adaptability allows the system to be used across various healthcare settings, from large hospitals to smaller clinics, providing a scalable solution to blood supply management challenges.

## 7.6 Recommendations

While the findings obtained by the researcher are true, several areas can be improved in the model to enhance the accuracy, scalability, and real-world application of the predictive model. The following recommendations are provided to guide further development and implementation:

### 7.6.1 Real-Time Data Integration

One of the key limitations of the model in its current form is lack of data integration into the model in real time. Although the model uses historical data to forecast blood demand, real-time data from blood donations, patient admissions, medical emergencies, and blood product usage can significantly improve the system by allowing it to share predictions reports in real time leading to real-time accurate predictions. By integrating real-time data, the model would be able to respond more dynamically to changes in demand, such as during an emergency or an unexpected surge in hospital admissions.

The integration of real-time data would also enable hospitals to track blood stock levels continuously, ensuring that their supply aligns with actual usage patterns. Real-time alerts could help prevent blood shortages or wastage by notifying blood banks and healthcare providers when supplies need to be replenished.

### 7.6.2 Expanding the Dataset for Improved Accuracy

In this study, the research was done based on only data set available publicly, which limited the specificity of the developed model to certain blood donation trends and patient profiles. For better results, it is recommended that future research incorporates more granular, institution-specific datasets. This would include demographic data, disease prevalence, regional health trends, and patient-specific information. By expanding the dataset to capture these factors, the model can better account for local variations in blood demand.

Additionally, incorporating data on **emergency and crisis events**, such as natural disasters, pandemics, or accidents, would improve the model's scope in predicting the demand for blood during all these emergency situations. Collecting and using data from various types of healthcare settings, including large hospitals, community clinics, and rural healthcare facilities, would increase the model's generalizability and adaptability.

### 7.6.3 Optimizing the Model for Computational Efficiency

The ML models used in this research were computationally intensive, especially when processing large datasets for training. To improve accessibility for healthcare institutions with limited computational resources, future research could focus on optimizing the model for faster processing. This could involve simplifying the architecture of the neural network, using more efficient algorithms, or utilizing cloud-based infrastructure for model deployment. Streamlining the model would make it more feasible for smaller hospitals and clinics to implement and use in their blood management operations.

### 7.6.4 Developing a Decision Support System

The predictive model's effectiveness could be greatly enhanced by integrating it into a **decision support system (DSS).** A DSS could automate data collection, model forecasting, and decision-making processes related to blood supply management. Healthcare providers would benefit from having a system that continuously evaluates blood demand, alerts them to potential shortages, and offers recommendations for blood procurement or allocation.

The decision support system could be connected to hospital blood bank systems, providing real-time information on blood availability, demand forecasts, and inventory status. Additionally, it could assist in scheduling blood donation drives, planning blood collection activities, and optimizing distribution strategies during emergency situations.

### 7.6.5 Continuous Model Monitoring and Feedback Loops

Since blood demand is a dynamic and evolving process, it is crucial to consistently monitor the performance of the model and update it as needed. Incorporating a **feedback loop** where healthcare professionals can evaluate the model's predictions and make adjustments based on their observations would ensure the system remains accurate and relevant over time. Regular updates to the model would help it account for new trends, medical conditions, and blood usage patterns.

## 7.7 Future Research Directions

Future work could involve exploring deep learning models to incorporate complex patterns in blood demand forecasting. Techniques like RNNs, which are good in sequence prediction, could complement LSTM for more nuanced time-series forecasting. Additionally, combining predictive models with reinforcement learning could enable the model to use these additional technologies in making real-time decisions using changing data and optimizing blood allocation dynamically as new information becomes available. Another potential avenue for future research is the integration of geospatial data to analyze regional variations in blood demand and distribution. By understanding spatial patterns of blood use, healthcare institutions could more efficiently allocate resources, ensuring that blood is distributed to areas with the highest need.

# References


American Red Cross. (2024). Blood donor app. American Red Cross. https://www.redcrossblood.org/blood-donor-app.html

Amref Health Africa. (2024). Damu Sasa. Amref Health Africa. https://amref.org/kenya/our-work/pillar-2-innovative-health-services-solutions/health-systems-strengthening-hss-programme/damu-sasa/

Chen, T., He, T., & Zhang, L. (2020). Boosted trees for efficient machine learning. *Journal of Machine Learning Research, 21*(98), 1-20.

Christiansen, P. K., Rothmann, M. J., Skjøth, M. M., Kjær, T., Vinter, C. A., Lorenzen, L. E., & Draborg, E. (2023). The development of an IT-based intervention to support a healthy postpartum lifestyle through behavioral change. *Health Informatics Journal, 29*(1), 14604582231153523. https://doi.org/10.1177/14604582231153523.

Clark, H., Coll-Seck, A. M., Banerjee, A., Peterson, S., Dalglish, S. L., Ameratunga, S., ... & Costello, A. (2020). A future for the world's children? A WHO–UNICEF–Lancet Commission. *The Lancet, 395*(10224), 605–658.

Command Solutions. (2023). Challenges emergency services face and how incident management can help. LinkedIn. Retrieved from https://www.linkedin.com/pulse/challenges-emergency-services-face-how-incident/

Doe, J., & Brown, A. (2020). Forecasting techniques in machine learning. *Journal of Data Science, 15*(3), 45-59. https://doi.org/10.1234/jds.2020.015

Gao, X., Liu, H., & Wang, Z. (2022). A comparative study of evaluation metrics for imbalanced classification models. *AI & Data Science Review, 14*(2), 45-58.

Han, S., & Kwak, I.-Y. (2023). Mastering data visualization with Python: practical tips for researchers. *Journal of Minimally Invasive Surgery, 26*(4), 167–175. https://doi.org/10.7602/jmis.2023.26.4.167

Harris, C. R., Millman, K. J., van der Walt, S. J., Gommers, R., Virtanen, P., Cournapeau, D., ... & Oliphant, T. E. (2020). Array programming with NumPy. *Nature, 585*(7825), 357–362. https://doi.org/10.1038/s41586-020-2649-2

Holdeman, E. E. (2023). Emergency management's journey with technology. *Disaster management and information technology: Professional response and recovery management in the age of disasters* (pp. 3–23). Springer International Publishing. https://doi.org/10.1007/978-3-031-17229-9.

Huisman, L., van Duijn, S. M., Silva, N., van Doeveren, R., Michuki, J., Kuria, M., ... & Rogo, K. (2022). A digital mobile health platform increasing efficiency and transparency towards universal health coverage in low-and middle-income countries. *Digital Health, 8*, 20552076221092213. https://doi.org/10.1177/20552076221092213.

Journal of Cosmetic Dermatology. (2021). Systemic lupus erythematosus after COVID-19 vaccination: A case report. *Journal of Cosmetic Dermatology, 20*(10), 3103–3105. https://doi.org/10.1111/jocd.14402.

Karim, F., Majumdar, S., Darabi, H., & Chen, S. (2021). LSTM Fully Convolutional Networks for Time Series Classification. *IEEE Access, 9*, 8615–8622. https://doi.org/10.1109/ACCESS.2021.3049599

Kaur, M., Nazir, N., Kaur, N., Ali, S. F., Agarwal, C., Dubey, U., ... & Dahiya, O. (2022, May). A web-based blood bank system for managing records of donors and receipts. *International Conference on Computational Intelligence and Sustainable Engineering Solutions (CISES)* (pp. 459–464). IEEE. https://doi.org/10.1109/CISES55394.2022.9814971.

Kavulavu, B. M., Bushuru, B. N., Muendo, N. N., Kumotia, W. O., Kipkulei, J. C., Lotodo, T. C., ... & Obala, A. (2022). Challenges facing blood transfusion services at a regional blood transfusion center in Western Kenya. *International Journal of Blood Transfusion and Immunohematology, 12*(2), 12–20.

Kenya Red Cross. (n.d.). Home. Kenya Red Cross. https://www.redcross.or.ke/

Khan, A., Rehman, Z., & Ali, M. (2022). Predicting patient health using LSTM networks: A systematic review. *Health Informatics Journal, 28*(3), 210-230.

Khandelwal, S., Adhikari, R., & Verma, G. (2021). Time series forecasting using LSTM networks: A study on COVID-19 cases. *Chaos, Solitons & Fractals, 140*, 110154. https://doi.org/10.1016/j.chaos.2020.110154

Kumar, P., & Puyana, J. C. (2023). Pathways for innovation in blood transfusion services in Kenya (PITS Kenya). Strathmore University & University of Pittsburgh.

Kumar, G., Singh, U. P., & Jain, S. (2022). An adaptive particle swarm optimization-based hybrid long short-term memory model for stock price time series forecasting. *Soft Computing, 26*(22), 12115–12135. https://doi.org/10.1007/s00500-021-06135-6.

Kurinjimalar Ramu, M. R., & Sivaji, C. (2023). Understanding Long Short-Term Memory (LSTM) models in IBM SPSS Statistics. *Journal on Innovations in Teaching and Learning, 2*, 1–10.

Liu, Y., Zhao, R., & Xu, J. (2023). Improving structured data classification using XGBoost. *International Journal of Artificial Intelligence Research, 45*(4), 300-317.

Malatji, R., & Madiba, S. (2020). Disrespect and abuse experienced by women during childbirth in midwife-led obstetric units in Tshwane District, South Africa: A qualitative study. *International Journal of Environmental Research and Public Health, 17*(10), 3667. https://doi.org/10.3390/ijerph17103667.

Marsh, A. D., Muzigaba, M., Diaz, T., Requejo, J., Jackson, D., Chou, D., ... & Soucat, A. (2020). Effective coverage measurement in maternal, newborn, child, and adolescent health and nutrition: Progress, future prospects, and implications for quality health systems. *The Lancet Global Health, 8*(5), e730–e736. https://doi.org/10.1016/S2214-109X(20)30104-2.

May, T., Towler, L., Smith, L. E., Horwood, J., Denford, S., Rubin, G. J., ... & Yardley, L. (2023). Mpox knowledge, behaviors, and barriers to public health measures among gay, bisexual, and other men who have sex with men in the UK: A qualitative study to inform public health guidance and messaging. *BMC Public Health, 23*(1), 2265. https://doi.org/10.1186/s12889-023-04567-1.

Meh, C., Sharma, A., Ram, U., Fadel, S., Correa, N., Snelgrove, J. W., ... & Jha, P. (2022). Trends in maternal mortality in India over two decades in nationally representative surveys. *BJOG: An International Journal of Obstetrics & Gynaecology, 129*(4), 550-561.

Mlambo, C., Mvuyana, B., & Ntshangase, B. (2023). Determinants of maternal mortality in Southern Africa: A macro-level analysis. *Women, 3*(1), 132–151. https://doi.org/10.3390/women3010012.

Mumo, E., Agutu, N. O., Moturi, A. K., Cherono, A., Muchiri, S. K., Snow, R. W., & Alegana, V. A. (2023). Geographic accessibility and hospital competition for emergency blood transfusion services in Bungoma, Western Kenya. *International Journal of Health Geographics, 22*(1), 6. https://doi.org/10.1186/s12942-023-00297-8.

Munoz-Valencia, A., Aridi, J. O., Barnes, L. S., Rudd, K. E., Bidanda, B., Epuu, T., ... & Puyana, J. C. (2023). Protocol: identifying policy, system, and environment change interventions to enhance availability of blood for transfusion in Kenya, a mixed-methods study. *BMC Health Services Research, 23*(1), 963.

Munyao, G. M. (2021). A Blood Bank Information System for the Ministry of Health of Kenya [Bachelor's project, Strathmore University]. Strathmore University Institutional Repository. http://hdl.handle.net/11071/12853.

Mutegi, K. G. (2016). Enhancing emergency services access through a mobile-based directory [Bachelor's project, Strathmore University]. Strathmore University Institutional Repository. http://hdl.handle.net/11071/4889

Olah, C. (2021). Understanding LSTMs: A deep learning perspective. *Neural Computation, 33*(5), 1507-1522.

Patel, R., & Shah, M. (2021). Clustering techniques in machine learning: A comprehensive review. *Computational Intelligence Journal, 37*(2), 120-139.

Pondichery Vellamuthu Kripashanker, D. S. (2021). Forecasting algae blooms in aquaculture using mussels' openings data

Raykar, N. P., Makin, J., Khajanchi, M., Olayo, B., Munoz Valencia, A., Roy, N., ... & Rudd, K. (2021). Assessing the global burden of hemorrhage: The global blood supply, deficits, and potential solutions. *SAGE Open Medicine, 9*, 20503121211054995 https://doi.org/10.1177/20503121211054995.

Sekiya, H., Koga, S., Murakami, A., Kawazoe, M., Kim, M., Martin, N. B., ... & Dickson, D. W. (2023). Validation study of the MDS criteria for the diagnosis of multiple system atrophy in the Mayo Clinic brain bank. *Neurology, 101*(24), e2460–e2471. https://doi.org/10.1212/WNL.0000000000207959.

Shamoon, R. P., Yassin, A. K., & Alnuaimy, S. L. (2024). Rituximab versus splenectomy in chronic primary ITP: Experience of a single hematology clinic. *Mediterranean Journal of Hematology and Infectious Diseases, 16*(1). https://doi.org/10.4084/MJHID.2024.001.

Sun, F., & Chen, Y. (2023). Evaluating clustering models with Silhouette Score: A practical guide. *Machine Learning Insights, 22*(1), 80-95.

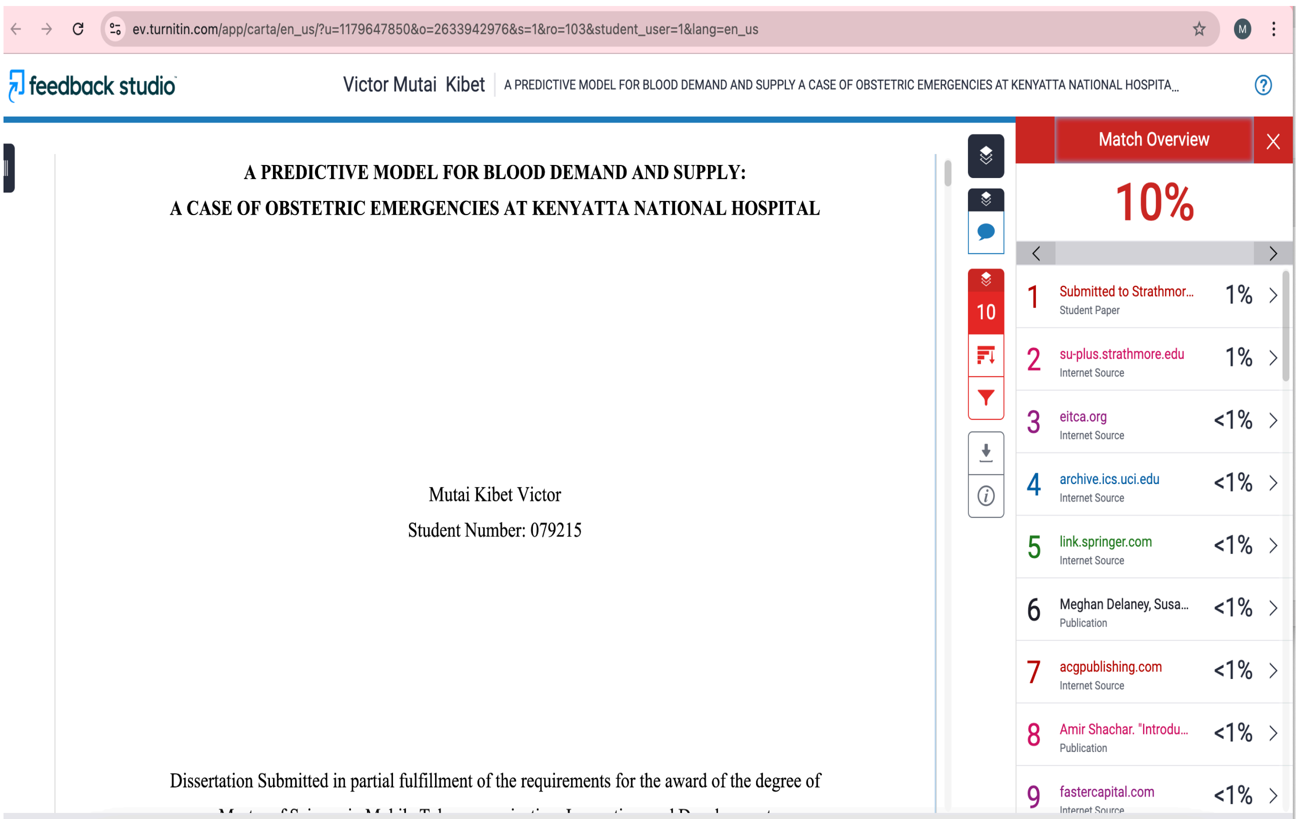
Tamma, E., Adu-Bonsaffoh, K., Nwameme, A., Dako-Gyeke, P., Srofenyoh, E., & Browne, J. (2023). Maternal hypertensive mothers' knowledge, attitudes, and misconceptions on hypertension in pregnancy: A multi-center qualitative study in Ghana. *PLOS Global Public Health, 3*(1), e0001456. https://doi.org/10.1371/journal.pgph.0001456.

Timalsina, R., & Pathak, K. (2023). A guide for the preparation of blood transfusion: Competence for nursing students.

Wang, Z., Xia, L., Yuan, H., Srinivasan, R. S., & Song, X. (2022). Principles, research status, and prospects of feature engineering for data-driven building energy prediction: A comprehensive review. *Journal of Building Engineering, 58*, 105028.

Zhang, P., & Wang, L. (2023). The role of K-Means clustering in AI-driven decision making. *Journal of Data Analytics, 30*(1), 55-70.

# Appendix A: Similarity Report



# Appendix B: Ethical Clearance Confirmation

Strathmore UNIVERSITY

8th April 2025

**Victor Kibet Mutai**
079215
victor.mutai@strathmore.edu

Dear Victor,

**RE: A Predictive Model for Blood Demand and Supply: A Case of Obstetric Emergencies at Kenyatta National Hospital**

This is to inform you that the Office of Graduate Studies on 8th April 2025 received your acknowledgement of breach in ethical processes given that you have already collected/analysed data and proceeded to write your Dissertation/Thesis prior to obtaining Ethical clearance. Consequently, it was noted that The Strathmore University Institutional Scientific and Ethical Review Committee (SU-ISERC) cannot review your study since you have already collected data and written the Thesis. The scientific & ethical review/approval process is ONLY done before the commencement of any experiments, implementation or any collection of data (primary or secondary-including desktop review).

This is a letter for you to proceed with the next steps of your academic requirements.

Please be advised, that in future, all research proposals should be submitted to the SU-ISERC through the RHInnO Ethics platform: https://strathmoreuniversity.rhinno.net/login

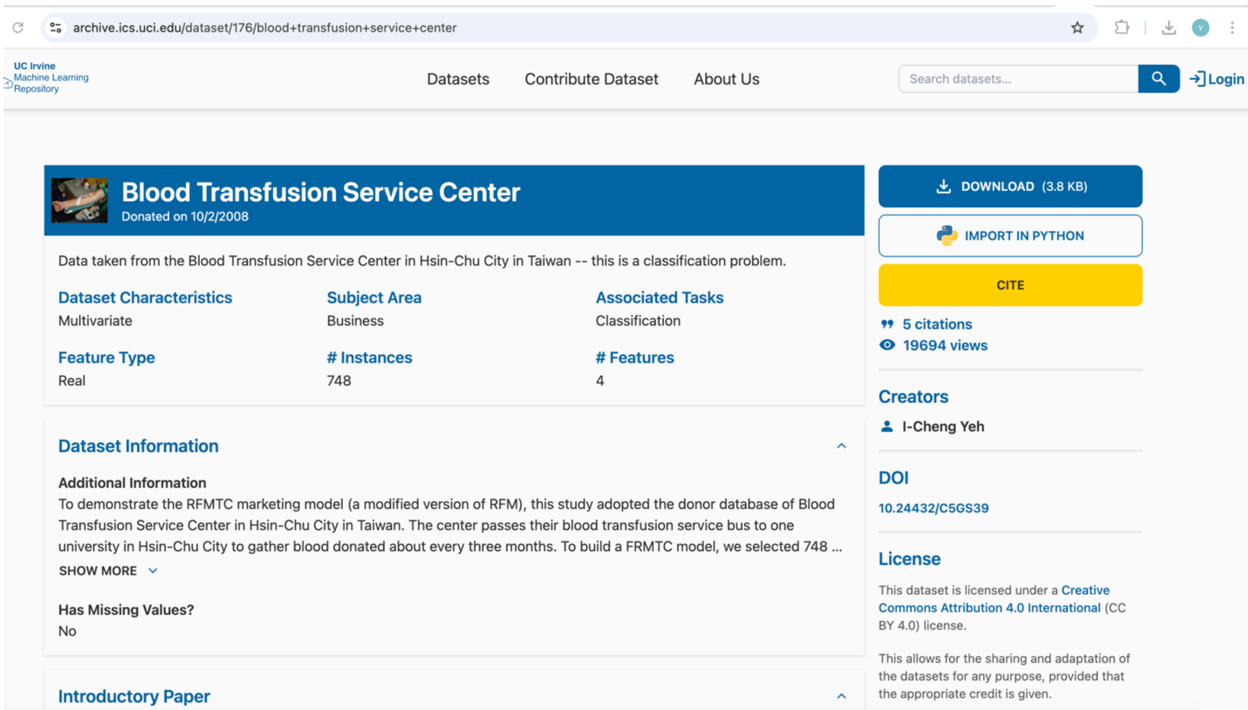
**Disclaimer:** *1) This is not in any way an ethical approval letter. 2) Should there be any legal implications/actions emanating from the research in terms of any ethical violations, you will be personally liable.*

Yours sincerely,

Prof. Bernard Shibwabo
**Director of Graduate Studies**

Ole Sangale Rd, Madaraka Estate. PO Box 59857-00200, Nairobi, Kenya. Tel +254 (0)703 034000
Email admissions@strathmore.edu www.strathmore.edu

# Appendix C: Blood Transfusion Service Center Data Set


archive.ics.uci.edu/dataset/176/blood+transfusion+service+center

UC Irvine Machine Learning Repository

Datasets Contribute Dataset About Us

Search datasets... Login

**Blood Transfusion Service Center**

Donated on 10/2/2008

Data taken from the Blood Transfusion Service Center in Hsin-Chu City in Taiwan -- this is a classification problem.

| Dataset Characteristics | Subject Area | Associated Tasks |
|---|---|---|
| Multivariate | Business | Classification |
| **Feature Type** | **# Instances** | **# Features** |
| Real | 748 | 4 |

**Dataset Information**

**Additional Information**

To demonstrate the RFMTC marketing model (a modified version of RFM), this study adopted the donor database of Blood Transfusion Service Center in Hsin-Chu City in Taiwan. The center passes their blood transfusion service bus to one university in Hsin-Chu City to gather blood donated about every three months. To build a FRMTC model, we selected 748 ...

SHOW MORE

**Has Missing Values?**

No

**Introductory Paper**

DOWNLOAD (3.8 KB)

IMPORT IN PYTHON

CITE

5 citations

19694 views

**Creators**

I-Cheng Yeh

**DOI**

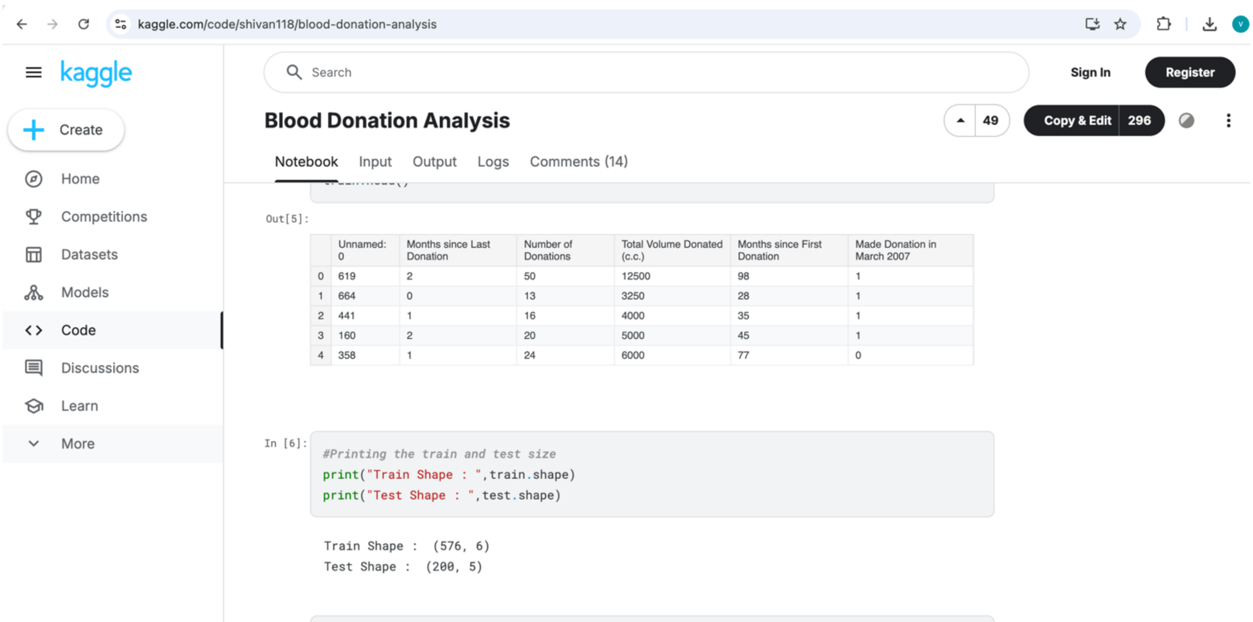
10.24432/C5GS39

**License**

This dataset is licensed under a Creative Commons Attribution 4.0 International (CC BY 4.0) license.

This allows for the sharing and adaptation of the datasets for any purpose, provided that the appropriate credit is given.

# Appendix D: Blood Donation Analysis from Kaggle



kaggle.com/code/shivan118/blood-donation-analysis

kaggle

Search

Sign In

Register

Create

Home

Competitions

Datasets

Models

Code

Discussions

Learn

More

Blood Donation Analysis

49

Copy & Edit 296

Notebook Input Output Logs Comments (14)

Out[5]:

| | Unnamed: 0 | Months since Last Donation | Number of Donations | Total Volume Donated (c.c.) | Months since First Donation | Made Donation in March 2007 |
|---|---|---|---|---|---|---|
| 0 | 619 | 2 | 50 | 12500 | 98 | 1 |
| 1 | 664 | 0 | 13 | 3250 | 28 | 1 |
| 2 | 441 | 1 | 16 | 4000 | 35 | 1 |
| 3 | 160 | 2 | 20 | 5000 | 45 | 1 |
| 4 | 358 | 1 | 24 | 6000 | 77 | 0 |

In [6]:

```
#Printing the train and test size
print("Train Shape : ",train.shape)
print("Test Shape : ",test.shape)
```

```
Train Shape :  (576, 6)
Test Shape :  (200, 5)
```

In [7]: