Heart to Say

FINAL REPORT - GROUP A

Team Contributors:
Sahid Hasan Rahim
Ifani Pinto Nada
Nan Jiang
Zhao Chen
Mahmoud Elachi

TABLE OF CONTENTS

1. Project Management

1.1. Member Contributions & Roles

Each team member made valuable contributions throughout the project, playing a critical role in ensuring its success. Tasks were distributed according to individual expertise and preferences, which facilitated an equitable workload among all members.

Ifani Pinto was responsible for organizing, time management, and overseeing the overall project structure. She played a key role in maintaining the team's progress, ensuring that the report was well-structured and aligned with the project's objectives and requirements. She initially assumed a leadership role as the other team members were modest and hesitant to step forward. Once the team members gained confidence, she transitioned to a more supportive role after her leadership phase ended. However, whenever the team lost focus or needed guidance, she stepped up to lead them, such as in meeting deadlines.

Mahmoud Elachi contributed by fostering team cohesion. Drawing from his extensive leadership and industry experience, he provided real-world examples and maintained a critical, analytical perspective throughout the project. The most crucial aspect of his role was to ensure that no divisions occurred within the team by actively listening to everyone and encouraging compromise during disagreements. Despite not having a medical background, he contributed valuable critical insights, often questioning the relevance of certain achievements or outcomes and alerting the team when discussions became overly complex.

Sahid Rahim served as the visionary of the team, offering innovative ideas through his "outside-the-box" approach and simplifying complex concepts into relatable examples and anecdotes for others to understand. Moreover, he frequently conducted quality checks to ensure that all reports met high standards and adhered to evidence-based science by cross-referencing the scientific literature on heart failure (the focus of this project). Sahid also monitored the data science part of the project, making necessary changes when it strayed off course, and compared the findings with scientific evidence to strengthen both the programming codes and the reports. His role was that of an editor-in-chief. As the project deadline approached, Sahid took on a more pragmatic leadership role to ensure the team remained on track and met their goals.

Zhao Chen led the programming efforts, ensuring the data analytics and coding components were effectively managed. He conducted thorough research on the dataset containing patients with heart failure in Pakistan, to ensure its relevance to the project's scope and existing scientific knowledge, drawing from his medical background. He also quality-checked the code snippets provided by other team members, making edits as necessary to make them more streamlined. When faced with dilemmas about which programming approach to take, such as in text mining for a separate project, Zhao encouraged discussions within the team to reach a consensus.

Nan Jiang contributed creatively by ensuring that the web dashboard adhered to user-friendly and accessible standards in line with Web Content Accessibility Guidelines (WCAG) [1]. With her proficiency in various software solutions, she helped resolve issues encountered during the project that were on the application level. She also stepped in as the programming lead whenever Zhao encountered problems. Due to her creativity and expertise in various software solutions, she led both the web medical dashboard and portfolios showcasing video efforts.

During the project, all team members took turns leading the team to gain experience in both managing and guiding the team and project. Table 1 outlines the final leadership assignments throughout the project.

Table 1. Member Roles

Activity	Date	Project Leader
Create team rules	04Sep2024 – 08Sep2024	Ifani Pinto Nada
Project charter	08Sep2024 – 22Sep2024	Mahmoud Elachi
Data science	08Sep2024 – 27Oct2024	Zhao Chen
Web medical dashboard	08Sep2024 – 27Oct2024	Nan Jiang
Portfolio showcasing video	08Sep2024 – 27Oct2024	Nan Jiang
Final project report	11Oct2024 – 27Oct2024	Sahid Hasan Rahim
Final presentation slides	20Oct2024 – 27Oct2024	Mahmoud Elachi

1.2. Group Development

The development of the team can best be described using the five stages of Cog's Ladder [2]. This model outlines the progression of team dynamics from initial uncertainty to active and high-level performance. In alignment with Cog's Ladder of group development, the team experienced the following phases:

a. The Polite Phase [2]
   At the start of the project, when the group had just formed, the team entered what is known as "The Polite Phase." During this initial stage, the primary focus was on getting acquainted with one another, recognizing individual strengths, and acknowledging each member's achievements. Team-building efforts began with introductory meetings on-site and on online platforms, where discussions centered on topics unrelated to the project, fostering rapport and mutual understanding of each other.

Subsequently, attention shifted to completing the project proposal, which needed to be finalized within a tight deadline. All members cooperated and avoided conflict, despite the moderate workload. It was later noted that while some individuals felt the pressure of the deadline, they chose to withhold their concerns in order to maintain a harmonious group dynamic. Hence, the challenge in this phase was on conveying honest opinions and thoughts.

b. Why Are We Here Phase [2]
The initial deadline for the project proposal brought the team into alignment, prompting a collective effort to find a solution that would have a meaningful impact in the field of health informatics. Acknowledging each member's strengths and characteristics, as established during the "Polite Phase," facilitated an effective division of tasks. However, challenges arose following feedback on the selection of a dataset for use in the project. It became necessary to reevaluate and reestablish the project, as the selected dataset comprised synthetic data, which was discouraged by the stakeholders for use.

During this stage, the team collaborated to identify a more appropriate and real-life dataset, reorganize tasks, and revise the project charter. This challenge proved instrumental in refining the project’s overall aim and goal. It enhanced communication between each member, which in turn strengthened the team bond. This phase was particularly significant, as it demonstrated the team’s resilience and cohesion, highlighting the capacity to persevere and remain despite facing challenges.

c. The Power Phase [2]
A conflict of ideas emerged in some instances. Regarding the selection of the dataset, the initial dataset was focused on Obsessive Compulsive Disorder, but the data was synthetic. Efforts were made in finding a real-life dataset belonging to the same domain. Some team members proposed switching the focus to heart failure, which was met with reluctance from others. The same issue occurred in a different project assignment when deciding on a text-mining approach. However, all team members finally approached the situation with reason and prioritized the best interests of the team. As long as arguments were supported by sound reasoning, consensus was reached swiftly. This stage was particularly beneficial, as it facilitated critical evaluation of ideas and helped the team narrow the project's scope. The most important lesson from this stage was the value of compromise. While strong passion for certain ideas can sometimes dominate, it was crucial to ensure that such enthusiasm did not divert attention from the primary objectives needed to successfully manage the project.

d. The Cooperation Phase [2]
   The "Cooperation" phase followed the "Why Are We Here" phase and progressed in parallel with other phases throughout the project. After the project charter seminar and the feedback received from the stakeholders, all team members aligned with a shared commitment to cooperate, contribute, and deliver their best efforts toward the successful completion of the project. A key takeaway from this phase was the team's further recognition of individual strengths and limitations, allowing members to focus on enhancing the skills in which they excelled. This approach fostered greater collaboration, as individuals took responsibility for the tasks that matched their strengths and actively sought out others for their expertise, further strengthening team dynamics. Although giving criticism can sometimes be uncomfortable, everyone learned to maintain a critical perspective during this phase to ensure that the project remained feasible and stayed within its intended scope. This collective scrutiny helped keep the team focused and on track.

e. The Esprit de Corps Phase [2]
   As the project neared its final deadline, the team experienced a sense of joy and pride in their collective efforts, drawing inspiration from one another’s contributions and skills throughout the project development. This mutual appreciation helped maintain focus and determination, ensuring that no one wavered from the original vision the team had for the project - to make an impact on the healthcare and health informatics community. This sense of pride is evident in the final prototype of the web dashboard, which was crafted with great care and passion.

1.3. Members’ Feeling Reflection

a. Ifani
   Ifani faced challenges in her role as project leader, particularly in managing the tight deadline while leading a newly formed team. Driven by her desire to meet deadlines, even if it meant pushing herself and expecting the same from others, she set an ambitious timeline and pushed the team to align with it. However, through this experience, she learned the importance of ensuring the members' well-being is also pertinent to ensure the success of the project.

   Organizing the team and ensuring strong connections between members was where Ifani felt most at ease. Her strength lay in assigning and reaffirming roles within the group, which helped maintain clarity and focus. The reassurance she received from team members, recognizing her effective delegation of tasks, further bolstered her confidence in her leadership abilities.

b. Mahmoud

For Mahmoud, tight deadlines were not a big issue, especially when required skill sets existed in the team members. However, the challenge in this project was more about managing and learning new skills. Other challenges were to fit with the pace of the team dynamic; since most of the work was done out of sight. On the other hand, all team members were working together, helping one another to successfully complete the project within the time frame. Additionally, from having experience in leading teams that are more vocal, Mahmoud appreciated the calmness as well as the seriousness each team member possessed for this project. The most significant takeaway for Mahmoud was that active leadership is not always required when team members are able to work freely while understanding their responsibilities.

c. Nan

In the team, each member had a unique experience regarding comfort and challenges. Nas was primarily responsible for creating and refining the dashboard. While Nan felt comfortable in this role, she faced challenges with the Streamlit package. Her desire for a perfect product led Nan to overcomplicate the dashboard by adding options for uploading datasets and models to accommodate various use cases. In the end, she realized that simplicity often results in a better user experience, and her perfectionist mentality made the design harder for users to navigate.

Reflecting on this experience, Nan believed that her steady personality, characterized by patience and attention to detail, pushed her to create structured solutions. However, she learned the importance of listening to others and being open to feedback, rather than sticking stubbornly to her own ideas. Effective collaboration requires valuing different perspectives, which can improve teamwork. This experience made her appreciate the diversity of personalities and how they can affect comfort levels. Nan recognized the value of simplicity and clarity, which ultimately enriched the team dynamics and fostered personal growth.

d. Sahid

Sahid is very ambitious and always wants to deliver excellent end results. Sahid's primary responsibilities included overseeing the overall vision of the project and compiling its various components while adhering to existing evidence-based science, which he felt comfortable with. This included the responsibility of narrowing the scope of the project and defining the analytical questions needed, as well as literature search and writing the reports. He was fortunate to receive support from other talented team members in programming, an area in which he felt less confident. Sometimes, the components for the project did not meet the expected quality. Sahid quietly stepped in to refine or complete those tasks because of being too overly cautious of criticizing others.

Overall, the project proved to be highly challenging and, at times, disheartening due to the criticism received from stakeholders mid-development. This feedback necessitated reshaping the project repeatedly, adding complexity to the process. Sahid is also a very sensitive person who does not take criticism lightly. However, this experience was valuable, as it mirrors the realities of the professional world, where continuous refinement is essential to transform an unpolished idea into a well-crafted, successful outcome. It also made Sahid reflect and understand the meaning of constructive criticism. Just as with an uncut diamond, this criticism was necessary to sharpen and perfect the final product.

e. Zhao

Zhao, with his medical background and programming expertise, felt confident in validating the dataset and driving the project toward its goal of producing a predictive model to generate valuable insights. He was also passionate about sharing his programming knowledge with the team, offering guidance and support whenever needed. However, Zhao found it challenging to narrow down the project’s focus or take initiatives. He preferred listening to others to help refine the project's scope. This approach suited him well, as he thrives on feedback and structured environments. Zhao is most comfortable when given clear instructions and direction, which allows him to focus on executing tasks efficiently and effectively.

1.4. Members’ Leadership Style Reflection

a. Ifani’s Reflection

Upon conducting the Dominance, Influence, Steadiness, and Conscientiousness (DiSC) personality test, the results show that Ifani’s dominant personality trait is Compliance (31%). This aligns with a Deliberate Leadership style, characterized by a focus on accuracy, structure, and ensuring tasks are done correctly [3]. This was evident in her initiative to create project-related documents, facilitate collaboration through various online platforms, and remind the team about deadlines throughout the project. However, this perfectionist tendency can sometimes overlook the emotional well-being of team members, and of herself. When Ifani initially assumed a leadership role, it became apparent that the leadership style did not resonate with everyone, as it was accustomed to the more hierarchical approach commonly practiced in Indonesia [4]. Recognizing this, Ifani adapted by observing and learning from the leadership styles of others, allowing her to foster a more inclusive and collaborative environment within the team.

b. Mahmoud’s Reflection

Over the past ten years, Mahmoud has led teams of varying sizes across different organizations, adapting his leadership style to fit the dynamics of each team and the responsibilities of his role. He believes that there is no absolute right or wrong in leadership, but that open and honest communication between team members and leaders is crucial. According to the DiSC profile, he is Dominant (40%) and aligns with the Pioneering Leadership approach. This leadership approach is categorized as being bold, passionate, and inspirational [3]. While leading his current team, Mahmoud quickly realized that the team members were not only dedicated and determined but also possessed strong leadership skills, enabling them to complete tasks without constant supervision. Recognizing this, he adopted a supportive, hands-off approach, offering help when necessary.

c. Nan’s Reflection

Nan’s DiSC profile showed the dominant trait of Steadiness (34%) and her leaning towards the Affirming Leadership approach. This leadership style is characterized by being respectful and positive [3]. This was evident when Nan oversaw the development of the portfolio showcasing video and the web dashboard. Nan engaged with her team members and gathered creative input throughout the project by fostering a positive environment. While she felt comfortable in this role, there were some challenges, particularly in aligning everyone's ideas and meeting deadlines. Nan encouraged open discussions and sought feedback, which helped the team converge on a cohesive vision and ultimately produce a strong final product. Moreover, during the project, Nan was impressed by the leadership styles of Mahmoud and Ifani. Mahmoud was great at motivating the team and fostering a positive environment, which encouraged everyone to share their thoughts. Ifani provided clear direction and organized tasks effectively, helping the team stay focused on the goals. Nan also noticed growth in Sahid and Zhao's leadership skills. Sahid became more confident in expressing his ideas, while Zhao improved in facilitating team conversations and encouraging collaboration. Their development was encouraging and highlighted how different leadership styles contribute to the team's success. Overall, this experience reinforced the importance of adaptability and learning from each other while working towards objectives.

d. Sahid’s Reflection

Being naturally reserved and unaccustomed to leadership, taking on this new and uncharted role was a significant challenge for Sahid. While he is highly cooperative, he is also very mindful of not offending others, which initially made leadership uncomfortable for him. This is evident in Sahid's DiSC score, which revealed a dominant trait of Steadiness (44%), followed by Compliance (38%). These traits align him with the Humble Leadership category, which is to be modest and fair-minded in leading a team [3]. Moreover, the pursuit of consensus is a characteristic often associated with Swedish culture [5], which aligns with Sahid's background, having been born and raised in Sweden. This cultural influence was reflected in his leadership style, where collaboration and mutual agreement are prioritized. However, the consequence of this leadership style is a tendency to be overly cautious and hesitant to move forward during disagreements. For example, Sahid being overly cautious of others' feelings led him to quietly refine some work/components of the project himself instead of confronting the team members, at the beginning of the project. Still, because the team fostered an environment where no one was overly dominant, collaboration flourished, making it easier for Sahid to step into the leadership role. This was largely due to the team members' compassion and maturity. Even in the midst of passionate discussions, no one objected when ideas were supported by sound reasoning. Ultimately, this project allowed Sahid to explore both his strengths and areas for growth. A key lesson learned was the importance of someone stepping up to take the lead or take responsibility in the absence of others. This is essential for the progress of projects and to ensure they stay on course, rather than diverging into multiple, uncoordinated directions.

e. Zhao's Reflection

Zhao is a highly determined individual who embraces technical challenges with a calm and composed demeanor. His DiSC profile indicates Steadiness as his dominant trait (30%). He adopted an Inclusive Leadership style, which fostered collaboration and an accommodating environment [3]. However, the downside to this approach is that a leader can become passive, particularly when others dominate the discussion. This was evident during the programming sessions with the team, where Zhao allowed others to take charge while he focused on the coding aspect of the project. Nonetheless, he realized that when someone is an expert in a particular area, it is more effective for them to lead the team in that domain, as their comprehensive knowledge can guide the group more effectively. Failing to do so can result in complications later in the development process.

f. Was there any person that was consistently being called on to lead the team, or that was particularly good at it?

Mahmoud Elachi led two phases of the development cycle due to his extensive experience in leadership and team management from prior industrial jobs. Initially expected to lead three phases, Mahmoud recognized the strengths and talents of each team member, allowing him to confidently step back. Nan Jiang demonstrated exceptional skills in developing the web dashboard, which led her to assume leadership during that phase. She also took a leadership role in the final visualization and usability evaluation phase. Overall, this approach of distributing leadership provided the rest of the team with valuable opportunities to learn and gain insights, ultimately contributing to the success of the project.

1.5. Leadership and Group Development Reflection in General

a. What worked well?

Taking on the leadership role within this team was straightforward, thanks to shared personality traits like Steadiness and Compliance, which fostered a harmonious team dynamic. Most of the team members were aware of their responsibilities, understanding that a drop in performance by any individual could impact the entire project and process. Moreover, individual maturity played a pivotal role in the team, enabling constructive collaboration and open communication to facilitate effective teamwork [5,6]. It also enabled the team to go into The Esprit de Corps Phase faster, as per Cog's ladder [2]. Individual maturity is also important for the team itself to mature, particularly in Agile environments [7]. This is particularly important given that the team adopted the Scrum framework (further detailed in section 1.6) [8].

b. What did not?

Cultural differences in leadership initially led to misunderstandings within the team due to varying communication styles and interpretations of English, shaped by each member's background and experience. These differences influenced how instructions were conveyed and understood, resulting in minor misunderstandings. Yet, this challenge ultimately became a valuable opportunity for growth. It allowed the team to develop essential skills in working within a multicultural environment, fostering adaptability and improving cross-cultural communication.

Also, the initial dataset selected introduced further obstacles, as it did not meet the stakeholders' requirements. This revelation placed the project in jeopardy, as the team had to return to the beginning stages and identify a suitable real-life dataset, adding further strain to the workload. <u>How is this relevant to leadership and team growth?</u> These challenges reflect real-world scenarios, where stakeholders or users may change requirements or even cancel projects mid-development. This experience served as a valuable exercise in resilience, adaptability, and maintaining momentum despite setbacks, allowing team members to step up and take on leadership responsibilities.

c. What could have been done differently and how?

A more effective approach could have been for the leader to remain flexible and prepared with alternative solutions when challenges arose, such as the unexpected need to change datasets. Implementing a contingency plan, like having a "plan B" or even additional backup options, would have enabled the team to quickly adapt to new circumstances and minimize disruptions [9]. This could have been integrated into a Gantt Chart to provide a clearer overview of any potential disruptions to the project timeline, allowing for better tracking and adjustment in case delays or issues arose during development [10]. This type of foresight reduces the risks associated with unforeseen circumstances and allows the team to maintain progress without significant delays.

Additionally, clearly defining the requirements for each development stage in detail would have helped reduce confusion and misalignment. At times, tasks were completed without a clear focus on the overall requirements. By providing specific, measurable objectives for each stage, the leader could have ensured that all efforts were directed toward the project’s main outcomes [11]. Furthermore, engaging in multiple discussions with stakeholders would have helped to confirm if the project was progressing in accordance with their requirements. All these would not only keep the team aligned but also enhance accountability and ensure that the work being done was purposeful and in line with the project's vision.

1.6. Project Development Approach

The selected development methodology for the project was Scrum, due to its Agile framework [12,13]. Scrum encourages collaboration through regular meetings and continuous feedback. The roles in the Scrum framework were designated following the guidelines set out in the Scrum Guide [13]. This framework ensured that a product owner was responsible for overseeing and managing the project’s components, while the rest of the team members focused on their respective tasks. Transparency was fostered by regular discussions and the showcasing of project results helped to identify issues. However, some challenges appeared in adhering to this framework. For instance, while preprocessing was conducted, the absence of analytics undermined the original purpose of the project assignment, both affecting the project and the process. There is no point in having a selected dataset if it cannot be used by the target users as envisioned. It is critical to ensure that the dataset, while being technically correct, aligns with the project's goals. Another issue arose in Sprint planning, affecting the process. Although discussions were held, errors were made in the Gantt Chart, like scheduling the video production for the dashboard in parallel with the dashboard's construction, which is not feasible.

Following stakeholder feedback throughout the “Sprint Reviews,” the team was tested on its adherence to Scrum values. Initially, this feedback disrupted the process, as the team had to replace its primary choice of the dataset, which had already been deployed for preprocessing. Furthermore, the use of Scrum in the project highlighted some issues, particularly in the delivery of each team member’s work. Some of the work outputs lacked depth or quality during each sprint. Feedback was sometimes provided during the “Sprint Retrospective,” but there was also a tendency to hold back criticism to avoid offending others. This cautious approach impacted the overall internal feedback process. Team members recognized this issue and have reflected upon it for future projects. The benefit was that, despite incomplete work, evolving requirements could still be met due to the flexibility of the methodology, allowing for easy adjustments.

It is important to recognize that the distinct roles within Scrum were not fully adhered to, as many team members were experiencing these positions for the first time. Additionally, the team did not fully implement the “Daily Scrum” meetings as prescribed due to other commitments, which hindered the ability to hold meetings on a regular basis. Despite the team encountering these difficulties, the iterative nature of Scrum allowed the team to continuously learn and adapt in real time. Needless to say, the flexibility of Scrum made it easier to leverage each team member's strengths, fostering a more collaborative and productive environment as the project progressed.

1.7. Project Development Methodology Choices in Hypothetical Project Restart

Unlike Scrum, which operates in iterative cycles, Waterfall follows a sequential, stage-by-stage approach [14]. Each stage is completed in full before moving on to the next, ensuring a more linear and controlled development process. This sequential approach in Waterfall would likely have enabled more qualitative work output, as each stage would be thoroughly reviewed and completed before the team advanced to the next stages. By moving step-by-step, the team would have progressed through each phase with more precision, making the project more detailed, convincing, and with less errors. Additionally, Waterfall typically requires fewer meetings, which might have been an advantage in this context, given that team members had other commitments.

It could also be argued that Scrum remains a strong option for the development process. Scrum allows for incremental project development and the deployment of updates in short timeframes, which aligns well with the relatively limited timeline set by stakeholders [12]. Additionally, Scrum enables the team to effectively respond to evolving requirements, allowing for prompt adjustments based on feedback and changes [13]. This approach can fit well for the creation of a web-based medical dashboard such as for this project.

Upon reflection, the optimal approach would have been to use a combination of both development strategies. A hybrid approach, integrating Waterfall’s structured planning and staged process with Agile’s flexibility in adapting to evolving requirements, is particularly suitable for complex projects [15]. Given the challenges faced in providing qualitative output for each Sprint, this method would have been ideal. By applying Waterfall at a macro level for overall project management and using Scrum within each phase for iterative development, the team could have achieved a balanced outcome. This hybrid approach would have ensured that strict deadlines were met while maintaining a high standard of quality at each stage of development.

1.8. Lessons Learned from The Chosen Project Development Approach

a. What worked well and why?

Learning throughout the development process and adapting to Scrum, despite initial unfamiliarity, was what was important for the team. It made each team member explore Scrum to understand their benefits and limitations. This experience will help in future project management by knowing the most suitable methodologies to choose from (Waterfall vs Agile/Scrum).

Scrum allowed the commission of tasks in parallel, which helped save time. Ideas were implemented simultaneously, allowing for faster realization without delay. Additionally, consistent discussions and a commitment to transparency helped ensure that ideas were thoroughly evaluated. When errors or inconsistencies occurred, they were quickly identified and addressed. For instance, inaccurate statements in the reports were quickly refined. Redundant programming codes were quickly rewritten to prevent it from accumulating and creating bottlenecks.

b. What could/should be done differently and how?

Inadequate quality in work from team members led to issues because Scrum sprints required fast iterations. With Scrum, it would have been essential to define the requirements more clearly to avoid generating irrelevant or lackluster outcomes. This could have been achieved by using the criteria for the "Definition of Done" [13].

As noted in Section 1.7, a hybrid approach would have been more beneficial. This method would allow the team to focus on one step at a time, while also adapting to time constraints. For instance, the team could have first clearly defined analytical questions, studied the literature on the dataset and its features, and assessed their relevance to the problem domain. This would then have guided the feature engineering and preprocessing of the dataset based on established knowledge at the next stage.

Another beneficial approach would have been the implementation of a Unified Modeling Language (UML) flow chart to visualize the workflow for each sprint. This would have provided a clear representation of the project's progress and helped document the flow of tasks and responsibilities. Using UML diagrams, such as activity or sequence diagrams, could have helped the team better understand each Sprint's structure, identify potential bottlenecks, and ensure alignment with the project requirements. This visual approach would also have served as a reference point throughout the development, enhancing communication and reducing confusion

1.9. Conflict Management

According to the Thomas-Kilmann model [16], the team predominantly employed a collaborative approach to resolve disagreements; however, closer to the final project deadline, an accommodating approach was observed. Most members prioritized finding solutions that benefited the entire group, with a strong emphasis on teamwork and open communication to address conflicts. Whenever there were instances when some team members felt compelled to voice their concerns about alternative approaches and project directions, a meeting was promptly organized, with all team members participating. After discussing differing viewpoints and considering what would most benefit the project, the meetings concluded with a unanimous agreement on the next steps forward. This approach was time-consuming, as it required additional meetings and discussions, but it was essential to delivering a high-quality project by ensuring that everyone remained aligned. As the final deadline approached, it caused the team to shift into an accommodating mode to preserve harmony and avoid conflict, to deliver the project. This was due to the lack of experience in utilizing Scrum methodology as well as other individual factors, described in Section 1.10.

1.10. Conflict or Disagreement Occurrences

A disagreement emerged when some team members made decisions without consulting the entire group during a separate project assignment focused on text mining. This issue arose from miscommunication and a lack of leadership during that phase, leading to anxiety among members and prompting members to manage the project independently. The conflict was resolved through an immediate meeting where all members shared their perspectives. A collaborative discussion followed, allowing each side to present their reasoning, and the team ultimately reached a consensus on the tokenizer selection for text mining. This resolution process underscored the importance of flexibility and adaptability within the team, as well as the value of open dialogue in overcoming setbacks and preventing stagnation. However, this conflict led to a loss of time from this particular project. Additionally, the internal deadline for producing a prediction model was missed, creating conflict and anxiety among team members. The issue stemmed from hesitation to fully address the root cause, as everyone wanted to avoid further conflict. As a result, the problem was only partially resolved, leading to an accommodating approach aimed at quickly fixing the immediate issues. However, the underlying problem remained unaddressed.

1.11. Adherence to Project Timeline Plans

In this project, the team used Jira and followed the Scrum methodology for managing the tasks. This setup helped the team to stay organized and keep track of the progress throughout the project. A Work Breakdown Structure

(WBS) was established to outline the tasks with regular team meetings to discuss progress and any obstacles encountered.

Using Jira, the project was broken down into smaller, manageable tasks and organized into Sprints. Each Sprint had specific goals, which were aimed to make the team focus on completing tasks within set timeframes. The visual boards in Jira made it easy to see progress at any point and allowed us to quickly spot any issues that needed addressing.

Some challenges occurred, particularly at the beginning. Some team members found it tough to get used to the Scrum framework and Jira, which caused a few delays in updating their progress. Also, during the regular meetings, it was discovered that some tasks were taking longer than anticipated, which unfortunately pushed back the overall timeline and required the team to adjust future sprints. The Gantt diagram also wrongfully shows video production in parallel with the final development stage, which was fixed later. Overall, while the team managed to stick to most aspects of the WBS and the schedules in Jira (after readjustment), the inexperience of the team to adhere to Scrum brought difficulties.

1.12. Significant Deviation

Externally, the stakeholders' requirements for dataset selection caused significant deviations from the original timeline, as it required rescheduling and restarting from scratch. Internally, the final Scrum Sprints were not upheld due to complications in delivering key outputs, such as the prediction model, on time, leading the team to require additional time and resources. These factors were not due to any fault of the chosen project management tools, but rather a result of other circumstances.

1.13. Project Management Tools

Throughout the project, the team utilized the following tools to manage tasks and organize workflows efficiently:

a. Jira

Jira served as the primary platform for managing tasks, tracking project progress, and organizing workflows. Jira was chosen due to its strong support for the Scrum development framework, which the team tried to adhere to [17]. Key features like the Jira Scrum board enabled the team to facilitate iterative delivery, enhance team communication, promote transparency, and improve sprint planning. Figure 1-3 illustrate various views within the Jira workspace.

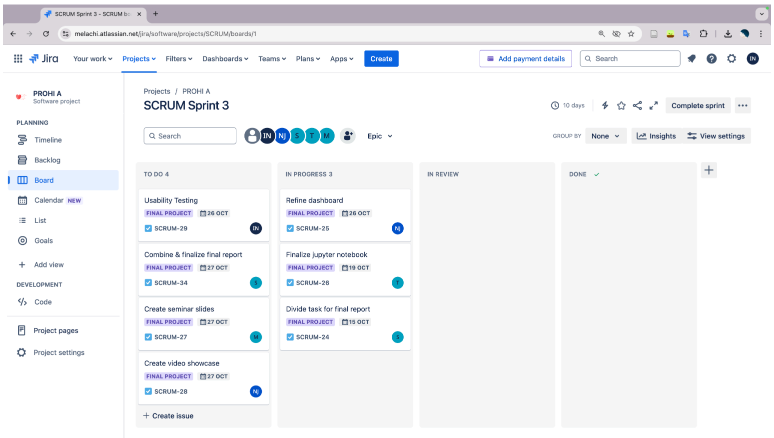


Figure 1. Scrum board displaying task status and assignee information

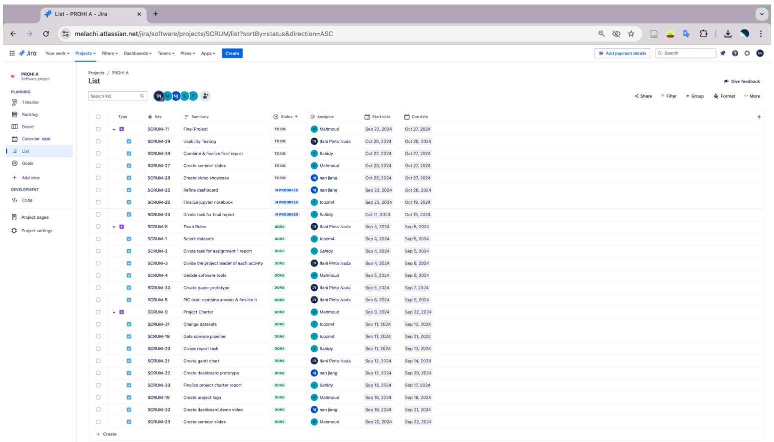


Figure 2. List view showing task details, assignee, start date, and due date

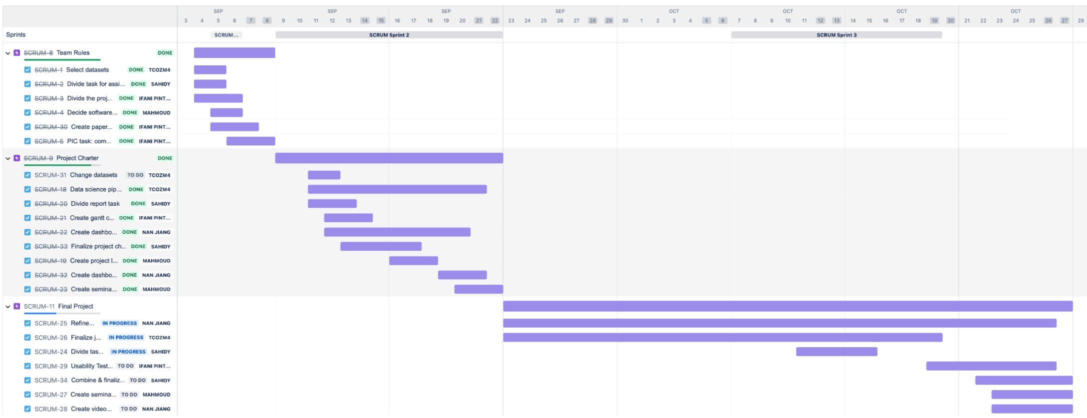


Figure 3. Timeline view for visualizing tasks over a specific time period

b. Google Workspace

Google Workspace was utilized as it facilitates collaborative document editing and file sharing [18]. A dedicated Google Drive folder (Figure 4) has been set up to store and organize project files. This includes documents created with Google Docs, Google Sheets, and Jupyter Notebook, allowing team members to easily collaborate and access the resources needed for the project. The programming was conducted using Jupyter Notebook through Google Colab for easy access and collaborative editing.

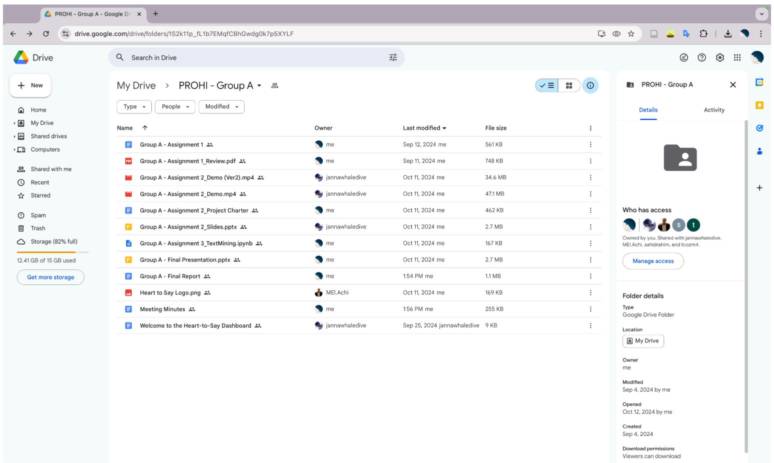


Figure 4. Google Drive folder containing project-related files

c. GitHub

GitHub was used for source code management and version control [19]. It enabled team members to collaborate on data science tasks, such as preparing and analyzing datasets, and to work on the coding necessary to build the Streamlit dashboard. Figure 5 provides an overview of the GitHub account. The GitHub repository can be accessed here: https://github.com/Jiangnanwhale/Heart-to-Say-Medical-Dashboard

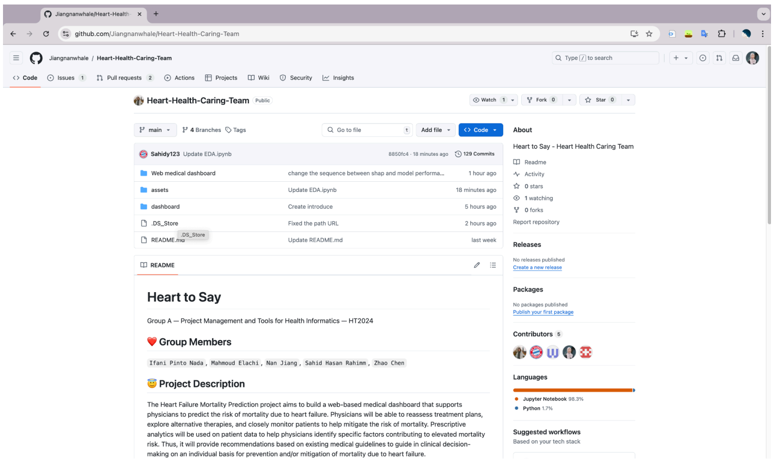


Figure 5. Overview of the GitHub repository

d. Zoom

Zoom was used for online daily meetings [20], to align with the Scrum methodology [21] when face-to-face meetings were not possible. These recurring meetings were scheduled every day from 5:00 to 6:00 PM, deployed from the beginning of the project and continuing until its completion. Although the meetings were initially set for one hour, the actual duration varied. The time and length were adjusted based on the team's availability. Figure 6 depicts the Zoom meeting agenda, showing the scheduled time.

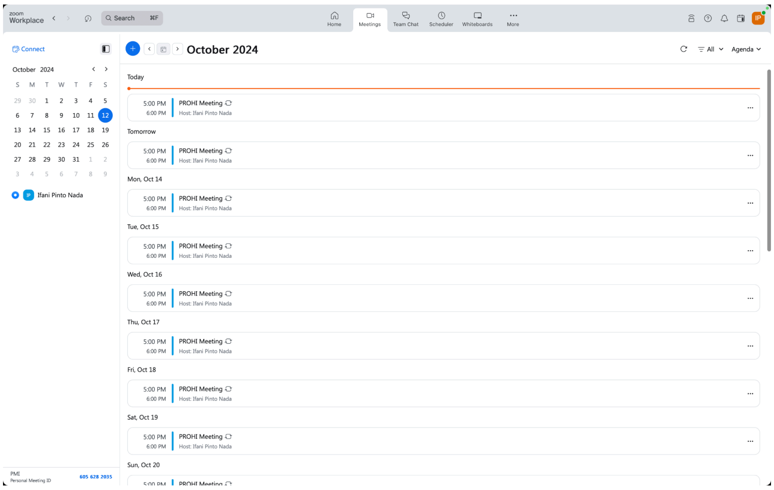


Figure 6. Zoom meeting agenda

e. WhatsApp

WhatsApp served as the primary communication tool for quick updates and discussions about the project [22]. It allowed the team to stay informed and connected instantly. Figure 7 shows a screenshot of the WhatsApp group, where regular updates and discussions took place.

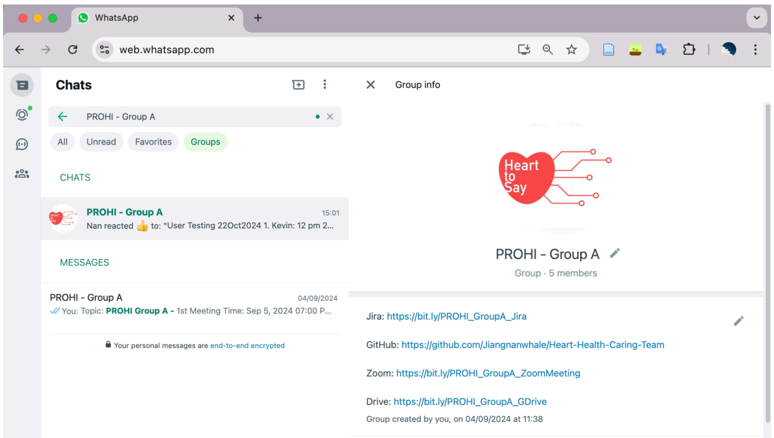


Figure 7. WhatsApp Group

f. Streamlit

Streamlit was used to create interactive web dashboards as it facilitates quick deployment of interactive web applications without requiring advanced programming knowledge [23]. For photographic evidence, see Section 6.2. for the final working dashboard.

2. Data Science Pipeline

2.1. Dataset Brief Description

Cardiovascular disease (CVD) is the leading cause of death worldwide, with an estimated 20.5 million deaths reported in 2021 [24,25]. Among CVD conditions, heart failure often presents a survival prognosis comparable to that of severe cancer types [24,26]. Heart failure is a condition where the heart cannot pump or fill with blood appropriately due to structural or functional issues in the heart [25]. Risk factors for heart failure include smoking [27], anemia [28–31], diabetes [32–34], ejection fraction and hypertension [27,35,36]. Furthermore, studies indicate that elderly patients [27,37] and men [27] are the most represented among those who die from heart failure. The severity of heart failure is also influenced by certain biomarkers falling below or rising above specific levels [38–41], as well as the presence of existing comorbidities. Moreover, it has been reported that a substantial increase in hospital costs is associated with heart failure management, highlighting the need for enhanced intervention strategies [27]. All these raise the need for accurate forecasting in heart failure patients to enable efficient time management and the timely implementation of appropriate treatment strategies to prevent mortality [27,42].

The Heart to Say project aims to develop a prediction model to estimate the likelihood of mortality in heart failure patients in Pakistan, using a dataset adopted from Chicco D et al. that focuses on the aforementioned subgroup [43,44]. Focusing on Pakistan is important, as heart failure patients in lower-income countries are 3-5 times more likely to die within 30 days of the first hospital admission compared to high-income countries, even after accounting for patient differences and long-term treatments [45]. Pakistan is categorized as a low-middle-income country, according to the Organisation for Economic Co-operation and Development [46]. Hence, the goal is to deliver a web dashboard for general practitioners, cardiologists, and cardiac nurses in Pakistan to:

- Predict the likelihood of mortality due to heart failure.
- Display the risk factors contributing to it.
- Aid and address the significant factors contributing to the risk of mortality.

The stakeholders are:

- Karolinska Institutet, Sweden in collaboration with Aga Khan University, Pakistan
- Ministry of National Health Services, Regulations and Coordination, Pakistan
- Pakistan Medical Commission, Pakistan
- Pakistan Cardiac Society, Pakistan
- The Swedish International Development Cooperation Agency, Sweden

2.2. Data Processing Pipeline

As listed in Table 2, the dataset includes demographic information (age and sex), smoking status, and the presence of comorbidities such as anemia, diabetes, and hypertension, all of which are factors with a strong correlation to heart failure [28–34,47]. It also includes indicators for heart failure, including creatinine phosphokinase (CPK) levels, platelet counts, serum creatinine levels, serum sodium levels, and ejection fraction percentages [27,38–41]. Death indicates the number of days before the deceased patient's follow-up could be carried out. Hence, DEATH_EVENT will be the binary target class for the prediction model.

Table 2. The features in the dataset, what they mean, and how they are measured.

Feature	Description	Measurements
Age	Age of the patient	Years
Anaemia	Decrease in red blood cells or hemoglobin	Yes/No
Creatinine Phosphokinase	Level of the Creatine Phosphokinase enzyme in the blood	mcg/L
Diabetes	Whether the patient has diabetes	Yes/No
Ejection Fraction	Percentage of blood leaving the heart at each contraction	Percentage (%)
High Blood Pressure	Whether the patient has hypertension	Yes/No
Platelets	Platelet count in the blood	Kiloplatelets/mL
Serum Creatinine	Level of serum creatinine in the blood	mg/dL
Serum Sodium	Level of serum sodium in the blood	mEq/L
Sex	Sex of the patient	Male/Female
Smoking	Whether the patient smokes	Yes/No
Time	Follow-up period	Days
DEATH_EVENT	Whether the patient died before the follow-up	Yes/No

The data processing pipeline is described below:

1) Loading and Initial Exploration: The dataset was loaded using Pandas, followed by an initial examination of the data structure, including column names, data types, and basic statistics. Pandas help to use and manipulate data in a coding environment [48]. Missing values were checked to identify any data gaps that required attention.
2) Data Scaling: To normalize the features, 'StandardScaler' was applied, which ensured that all variables had a mean of 0 and a standard deviation of

1 [49]. This is particularly important for machine learning algorithms sensitive to the scale of input data.

3) Feature Engineering: All the features were relatable to the problem domain since they are important factors in heart failure patients [24–34,47,50–53]. The follow-up period had a significant correlation to DEATH_EVENT. Hence, no features were removed.
4) Data Visualization: Boxen plots were utilized to visualize the distribution of the scaled features. It allowed for the identification of any outliers or skewness that might need to be addressed [54].

2.3. Descriptive Analytics Tab

Descriptive analytics addresses the fundamental question, “What happened?” This type of analysis can be gathered using the web dashboard that utilizes the dataset from Chicco D et al. [43,44]. The following analytical questions serve as the foundation for visualizing insights on the dashboard, along with figures to demonstrate the answers:

a. *What proportion of heart failure patients died compared to those who survived in the dataset?*

As shown in Figure 8, the results show that there are 96 recorded deaths compared to 203 still alive. In percentages, the dataset has almost 68% of heart failure patients still alive, while 32% are dead before any follow-up period.

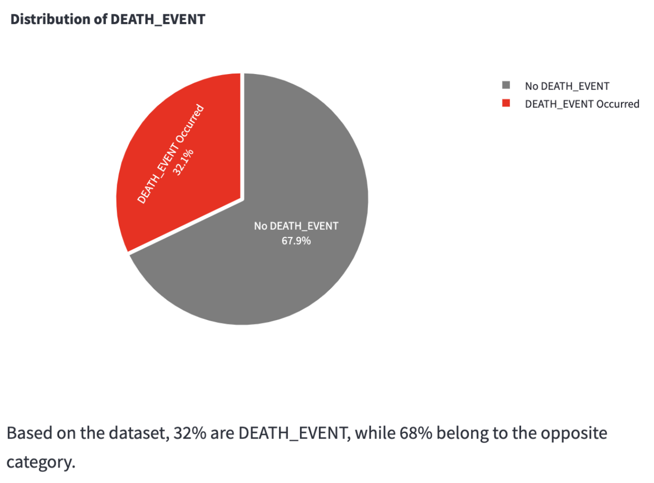


Figure 8. Death events among heart failure patients

b. *How does the age distribution vary, and which age group has the highest number of heart failure patients in the dataset*?

The results show that heart failure patients range from 40 to 95 years old, with the age group close to 60 having the most of this subgroup (Figure 9). This is close to the global average [45].

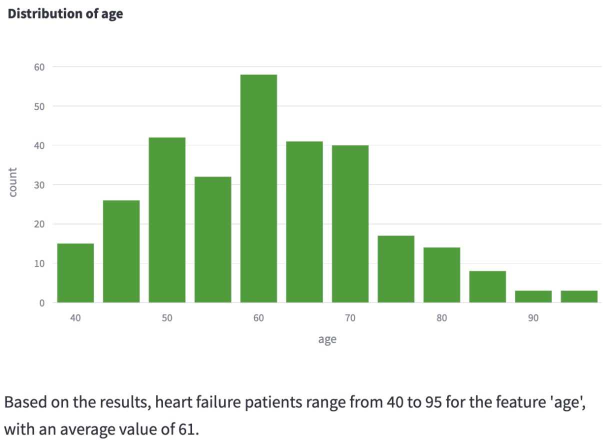


Figure 9. The age range of heart failure patients

c. *What is the age distribution among deceased and surviving heart failure patients?*
The dataset indicates that both death and survival are more prevalent among patients around 60 years of age (Figure 10). The death rate was significantly lower among patients around 40 years old, while it was higher in those aged 80 and above compared to the survival rate. Essentially, death events become more common starting from the age of 60 and onwards, though survival rates still tend to be higher until older age groups like 80-85 and beyond, where death events become more frequent or equal. These findings align with scientific evidence [27,37].

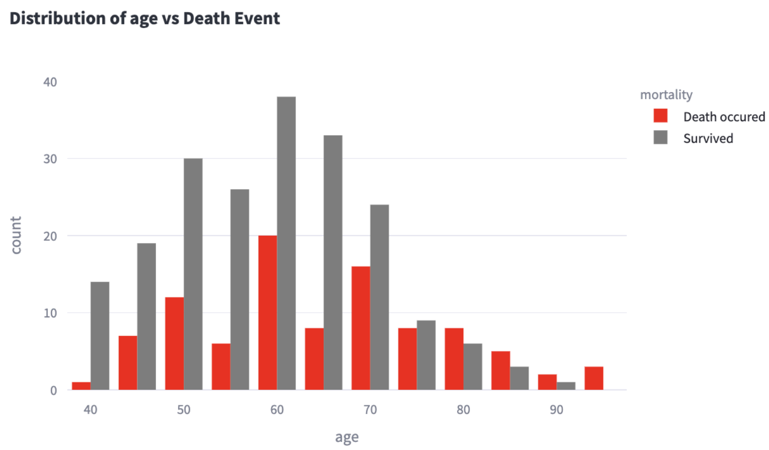


Figure 10. Age distribution and death event

d. *What is the proportion difference between male and female among heart failure patients*?
Based on the results, 65% are male, and 35% are female, indicating that heart failure is more prevalent among males (Figure 11). Scientific literature suggests that females with heart failure tend to have better survival rates and a lower risk of mortality compared to males [50].

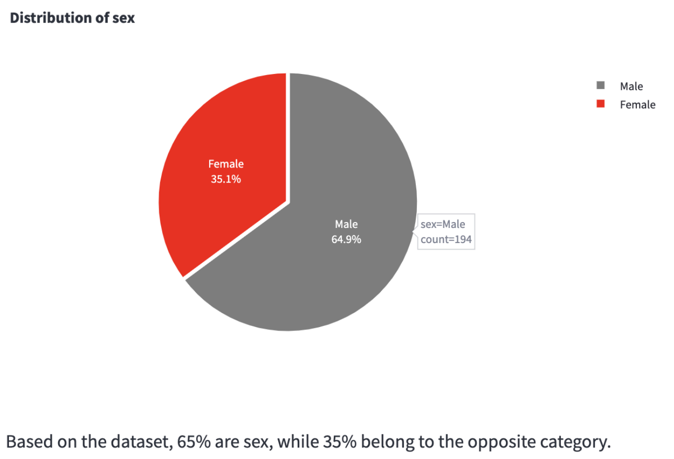


Figure 11. The distribution between both sexes

e. *What is the distribution of both sexes among deceased and surviving heart failure patients?*

   The results show a significant difference in mortality between the sexes (Figure 12), though it could be argued that males are overrepresented in the dataset. The key takeaway from the graph is that there are proportionally more survivors than deaths among the patient cohort in this dataset. However, it is important to note that existing evidence suggests that males have a higher mortality risk among heart failure patients [27,50].

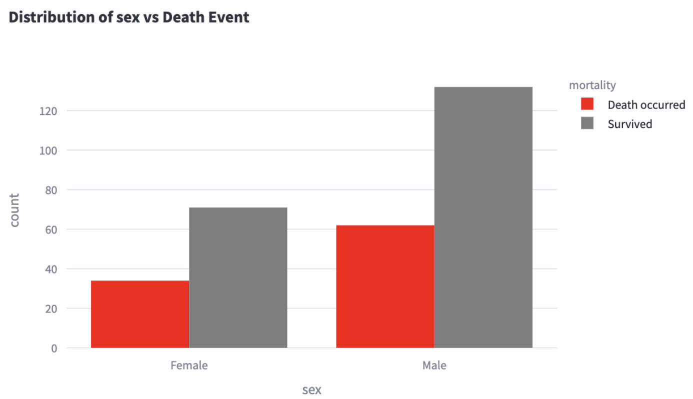


Figure 12. Distribution of sex and death event

2.4. Diagnostic Analytics Tab

Diagnostic analytics focus on answering the question, "Why did this happen?" in the dataset from Chicco D et al. [43,44]. The answers require a data science-oriented approach, as this domain-specific knowledge is essential. The following analytical questions lay the groundwork for visualizing insights on the dashboard, supported by figures that demonstrate the corresponding answers:

a. *Why are certain heart failure patients at risk of death?*

   First, the correlation coefficient needs to be defined to understand its relevancy to the question. Correlation coefficients range from -1 to 1. A positive correlation occurs as the variable increases, approaching 1, while a negative correlation occurs when the coefficient is less than 0 as the variable increases. A coefficient near 0 indicates little to no correlation. Also, different feature types need to be calculated differently (e.g., Binary-Binary, Binary-Numerical, Numerical-Numerical). Based on the heatmap, the following factors demonstrate either a relatively positive or negative correlation with death, compared to other factors (Figure 13):

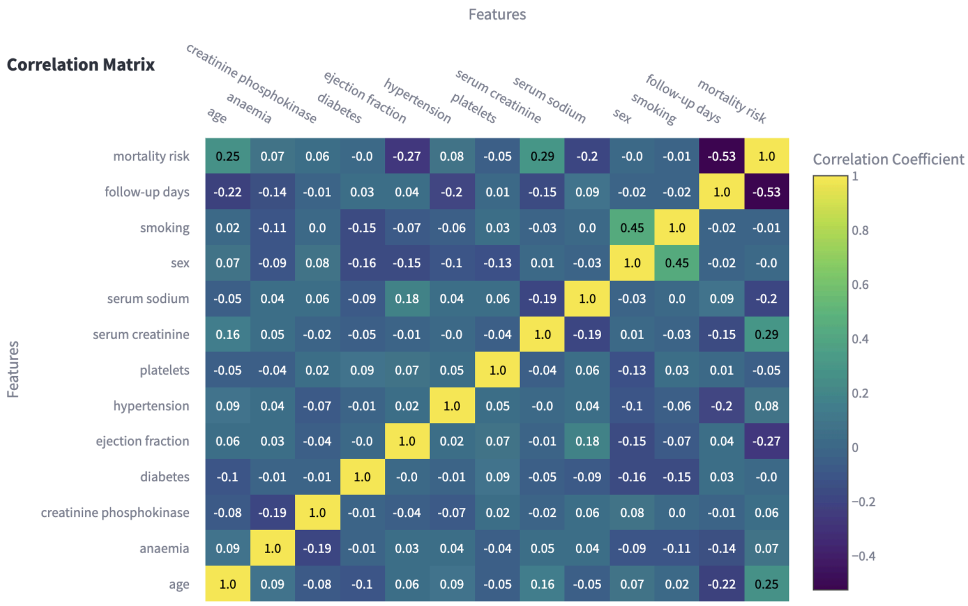


Figure 13. The heatmap from the web dashboard shows the correlation between each factor presented in the dataset.

1. Serum creatinine shows a positive correlation with death (0.29):
2. Ejection fraction shows a negative correlation with death (-0.27):
3. Follow-up days show a negative correlation with death (-0.53):
4. Age shows a positive correlation with death (0.25):
5. Serum sodium shows a negative correlation with death (-0.20):

Certain heart failure patients are more likely to die due to factors such as advanced age, elevated serum creatinine levels, lower ejection fraction, and decreased serum sodium. These factors are supported by scientific literature [27,37,51–53], further validating the correlations identified here. The follow-up days (time) variable also plays a critical role, although its significance is subject to interpretation based on context. It will, however, have a significant factor in the prediction modeling later. Additionally, it is worth noting that smoking, diabetes, anemia, high blood pressure, and sex did not show strong correlations with mortality in this dataset, suggesting that these factors alone are not sufficient to predict mortality risk. Finally, it is important to emphasize that correlation does not imply causation.

b. *Do heart failure patients with multiple conditions (as in anaemia, diabetes, and high blood pressure) and smoking status have higher mortality rates than those without these conditions?*
This question is based on the findings from the previous question, which showed that smoking, diabetes, anaemia, and high blood pressure do not have individual correlations with mortality. Consequently, the analysis will explore whether the combination of these conditions, including smoking, increases the risk of mortality. Sex is excluded from this analysis, as it is discussed separately in “Question e”, under descriptive analytics. Logistic regression was used to model the probability of a binary outcome (death=yes/no). Figure 14 illustrates the mortality rates for heart failure patients, categorized by the number of risk factors as specified by the question:
   - **Patients with 0 Risk Factors**: The mortality rate is close to 40%, indicating that about 40% of patients with no conditions and non-smokers died.
   - **Patients with 1 Risk Factor**: The mortality rate is in between 20-30%, suggesting that having one condition or being a smoker alone does not significantly increase the risk of death.
   - **Patients with 2 Risk Factors**: The mortality rate is just above 30%, suggesting a higher risk of death for patients with two conditions (or one condition in combination with smoking).
   - **Patients with 3 Risk Factors**: The mortality rate rises to around 50%, indicating that patients with three conditions (or two conditions and smoking) face a substantially higher risk of mortality.
   - **Patients with 4 Risk Factors**: The dataset contains no patients with all four risk factors.

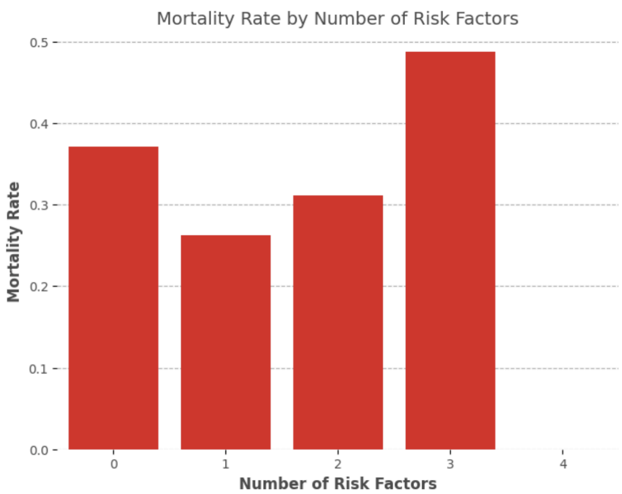


Figure 14. The bar plot explains the multiple risk factors and how they affect the mortality rate.

c. *Can the coefficients of the features in the dataset be considered statistically significant in predicting the outcome of death?*
   This question serves as a continuation of "question a," diagnostic analytics, with the objective of determining whether the features exhibiting the strongest positive or negative correlations possess statistical significance in relation to the outcome of death. Logistic regression was once again employed for this analysis. To maintain a 95% confidence level, the established p-value threshold of 0.05 was also applied [55]. Figure 15 shows that serum creatinine emerged as the most significant predictor of death in this dataset. Age, time, and ejection fraction, though important, had less pronounced effects, indicating that other factors should also be considered.

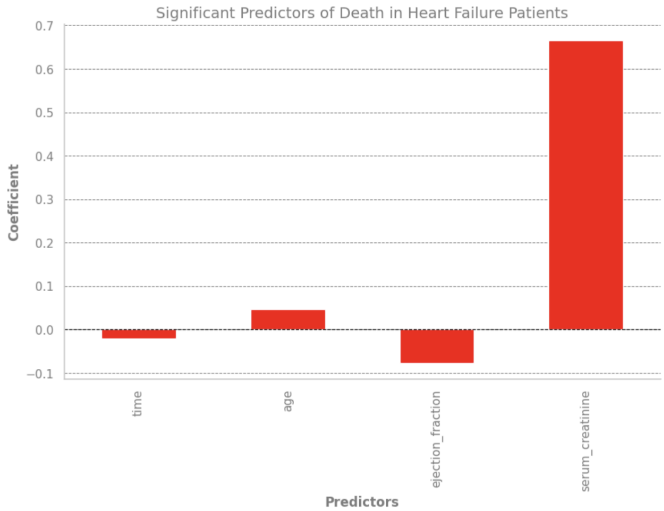


Figure 15. The bar plot shows the features that possessed the most statistical significance in predicting death outcomes.

d. *How does having diabetes, anemia, hypertension, and smoking status seem to impact the likelihood of death in this cohort of heart failure patients?*
   The focus was on testing the categorical features in the dataset with death outcomes. The Chi-Squared test is a viable option for that [56]. Figure 16 shows, however, that no statistically significant association was found between the categorical variables of diabetes, anemia, hypertension, and smoking with death events in heart failure patients. This implies that these conditions, in isolation, do not appear to correlate strongly with the likelihood of death within the dataset. The p-values exceeding 0.05 for the categorical features suggest that the differences observed between patients who experienced death events and those who did not are likely due to random chance or other factors.

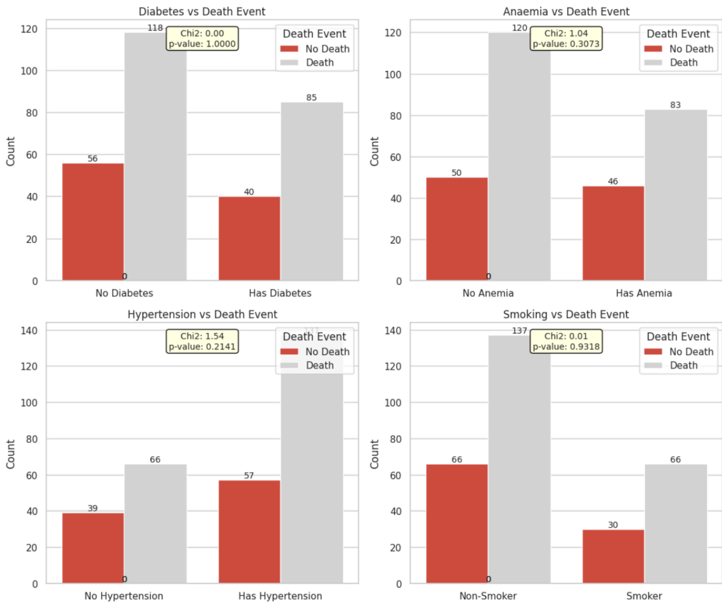


Figure 16. Bar plots for each categorical feature and the corresponding death outcome, along with their statistical significance (p-values) indicated.

e. *How do the different features cluster into groups, and which feature pairs create the most distinct clusters?*

The focus here was to cluster the numerical features of the dataset using K-Means. This did not include the categorical features (Boolean) since the distance-based metrics measure the similarity of data points by distance, which is not meaningful for binary values [57]. It can result in random and unreliable groupings. Moreover, the clustering was decided to divide into two groups according to the silhouette score. The silhouette score was slightly higher for clusters of 9 or more, but creating numerous small clusters would reduce interpretability. For example, a silhouette score of 1 may indicate perfect clustering, but if each data point forms its own cluster, it does not provide meaningful insights.

Figures 17a, 17b, and 17c show the clustering of “age” and “time” (follow-up days), along with a description of each cluster. The following insights were obtained by analyzing each numerical feature in relation to others through K-Means clustering:

CPK Levels: CPK consistently acted as a cluster divider when around 2000 mcg/L, observed in comparisons with age, ejection fraction, serum creatinine, serum sodium, and time.

Age: Age was a distinguishing factor around 60 years, especially in clusters compared with serum creatinine, serum sodium, and ejection fraction.

Platelets: Clusters divided consistently around 30,000 kiloplatelets/mL, unaffected by other features, in comparisons with CPK, ejection fraction, serum creatinine, serum sodium, and time.

Time: Clusters were differentiated around 150 days of follow-up duration across multiple feature comparisons, including ejection fraction, age, serum creatinine, and serum sodium.
Ejection Fraction: A division at 50% in ejection fraction was observed in clusters compared with serum creatinine and serum sodium.
Serum Sodium: Clusters were divided when serum sodium was above 135 mEq/L in comparison with serum creatinine.
Overall Cluster Grouping:
In summary, CPK, platelets, and time emerge as key features for distinct cluster formation across various pairings, while age and ejection fraction also contribute to cluster separations at specific thresholds.

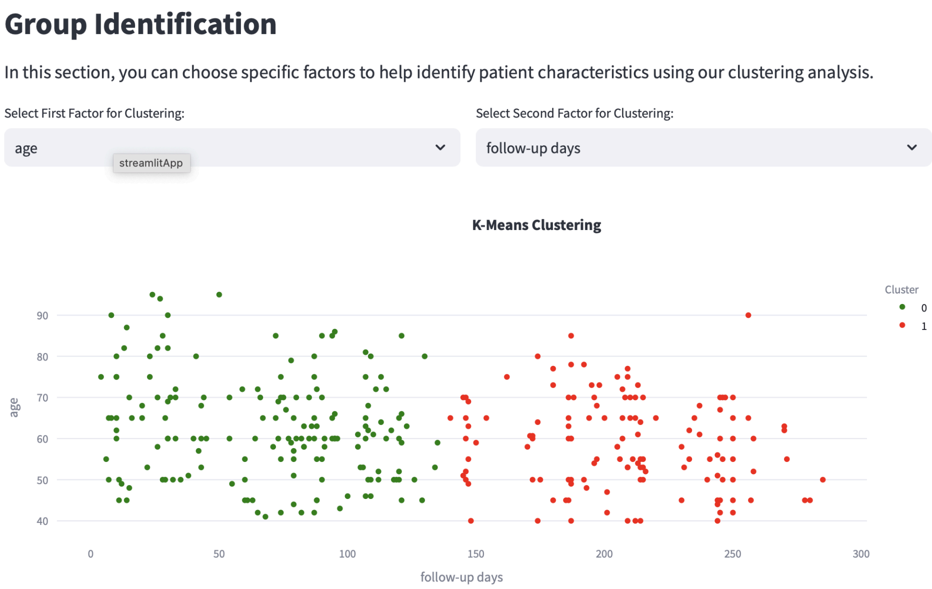


Figure 17a. Clustering of data points by the selected features

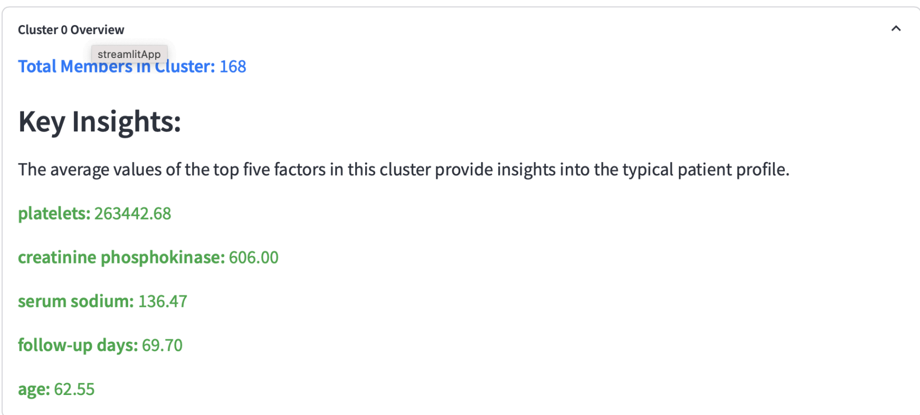


Figure 17b. Details on the Cluster 0

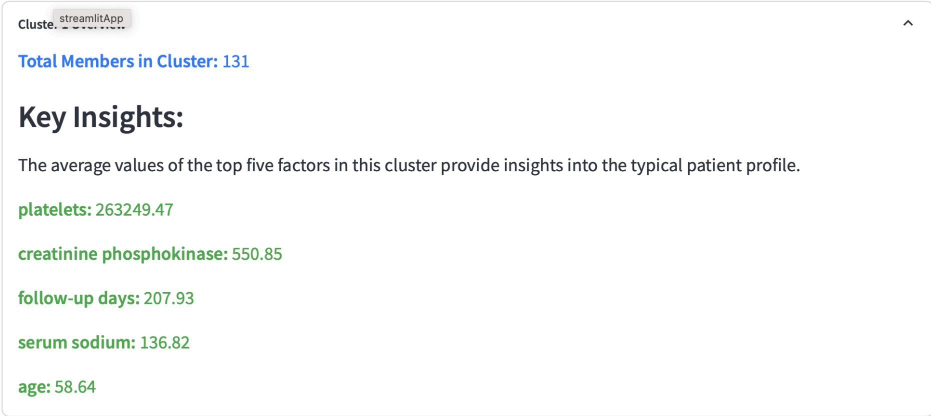


Figure 17c. Details on Cluster 1

2.5. Predictive Analytics Tab

The prediction of fatal heart failure events is a crucial classification problem in healthcare, as it can significantly enhance patient management and care [27,42]. In this context, the prediction model aimed to classify patients based on the likelihood of experiencing a fatal heart failure event (Figure 18), which is represented as a binary outcome. To address the prediction problem, data preprocessing, Machine Learning (ML) model training, and evaluation were performed. All the features in the dataset were relevant in predicting heart failure mortality as they are factors supported by scientific evidence [24–34,47,50–53].

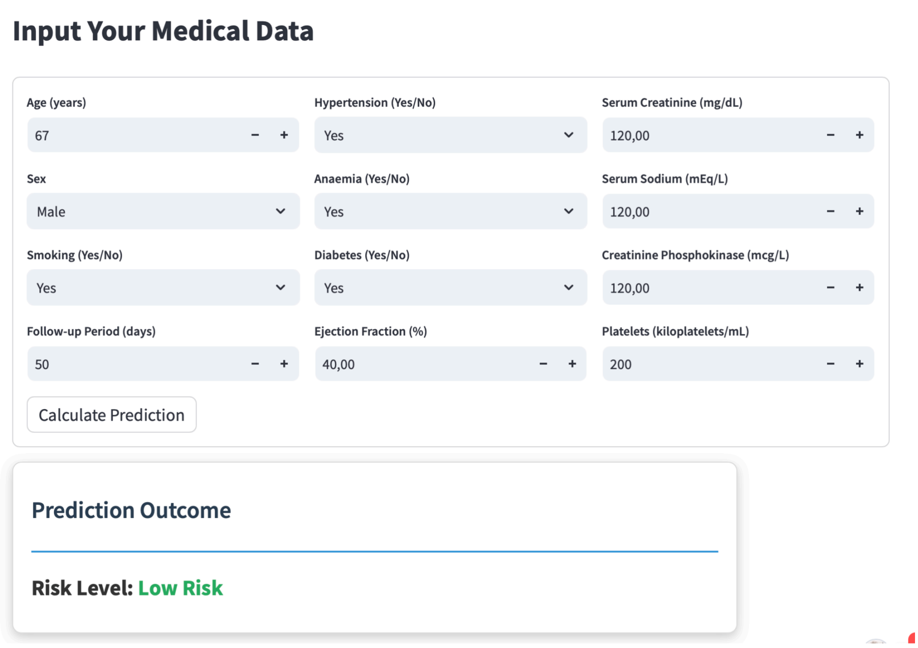


Figure 18. The prediction outcome is derived from the data input.

The dataset was processed by conducting an exploratory data analysis, followed by the separation of the target variable (DEATH_EVENT). 'DEATH_EVENT' is a binary variable indicating whether a patient has experienced a fatal heart failure event. Features were standardized using StandardScaler, and boxen plots visualized their distributions to detect outliers. For the prediction modeling, the dataset was split into training and test sets using a 70/30 ratio, with stratification applied to maintain the proportion of the target variable (DEATH_EVENT). Evidence shows that the 70/30 ratio gives the best outcome [58]. This approach ensured that the model could be evaluated on data it has not seen during training, helping to mitigate overfitting.

A dictionary of 10 different classifiers was created, including Random Forest, K-Nearest Neighbors (KNN), Support Vector Machine (SVM), and XGBoost models, each configured differently. This approach facilitated comparison across multiple algorithms to identify the best-performing model. Additionally, an Artificial Neural Network (ANN) model was created to see if it could capture a pattern to predict the target that is comparable to the ten classifiers utilized. Each model was iteratively trained and evaluated to analyze its performance on the test data, and the results were stored for further analysis.

A modification of the XGBoost model outperformed other models in key performance metrics, indicating its balanced ability to correctly identify positive cases while minimizing false negatives. Compared to the ANN model, it also offered advantages in training and prediction time. ANN models typically require longer training due to the large number of parameters, especially as the network's complexity increases with more layers [59,60]. Furthermore, ANN performance tends to be unstable across training sessions, limiting its reliability for consistent predictions. Additionally, ANN is unnecessary for this relatively small dataset that is being deployed. It is comparable to driving a Lamborghini within 50 km/h speed limit roads.

Lastly, Shapley Additive Explanations (SHAP) values were utilized to interpret model predictions. SHAP helps explain the contribution of each feature to the model's output [61,62]. Visualization of SHAP values for selected instances was performed, showing bar charts that highlighted the impact of individual features. By identifying which features were most influential, SHAP analysis offered valuable insights into key risk factors contributing to mortality risk. See Section 2.6 for further details on SHAP.

Finally, It is important to note that the dataset contained only 299 instances, which may be too limited and could introduce bias, as presented in Sections 2.3 and 2.4. In these sections, certain real-life relevant features did not align with actual predictive outcomes for mortality risk, and some subgroups were overrepresented. Nevertheless, the team proceeded with this dataset to determine if a predictive model could still be developed. For future work, a similar dataset with the same features but a larger number of instances would be beneficial. The primary goal here was to demonstrate that a predictive model could be achieved, even with a smaller dataset, while adhering to scientific evidence on heart failure indicators for mortality.

2.6. Prescriptive Analytics

SHAP is an advanced interpretation technique designed to explain the predictive outcomes of ML models by measuring the individual impact of each feature [61]. For this project, the SHAP Summary Plot and the SHAP Feature Importance were chosen as visualization strategies to elucidate the predictive mechanisms of the XGBoost model deployed to estimate the mortality risk among heart failure patients.

To capture the overall patterns and variations in feature importance across the dataset, the team opted to explain multiple predictions, rather than focusing on a single prediction. Here, the SHAP Summary Plot provided a comprehensive view of how each feature influenced the model's predictions. It displayed the spread of SHAP values by illustrating how each feature contributed to shifts in the model's mortality risk predictions compared to a baseline value. This approach was crucial for understanding how different features influenced the predicted mortality risk across the entire dataset. It is important to state that the SHAP values and importance scores themselves should not be interpreted as a causal link between the features and mortality risk unless supported by additional evidence [61].

From the SHAP analysis, the following key factors and their respective importance scores were identified:

1. Follow-up days: Importance Score = 0.6364
2. Serum creatinine: Importance Score = 0.2538
3. Ejection fraction: Importance Score = 0.2035

The insights garnered from the SHAP plots indicate that the duration of follow-up days exerts the most substantial influence on the model's output, with an importance score that is more than double that of serum creatinine and three times that of ejection fraction (Figure 19). These findings suggest that the length of time patients are followed post-treatment or diagnosis is a critical determinant in the model's estimation of mortality risk, with serum creatinine levels and ejection fraction also playing significant, albeit less pronounced, roles. Age was excluded because SHAP did not consider it as significant compared (Figure 19) to the logistic regression analysis results detailed in Section 2.4, Question C. This can be explained by SHAP considering non-linear features relationship better than logistic regression [63,64], essentially showing age is not as impactful.

While these insights are valuable for understanding the model's behavior and for refining its development, it is imperative to reiterate that the SHAP analysis is primarily a tool for model interpretability, not for inferring causal relationships. Therefore, the insights from SHAP should be treated with caution and not directly used to inform decisions without further causal validation [62]. That being said, scientific evidence supports the correlation between fluctuations in ejection fraction and serum creatinine levels with heart failure and mortality [27,38].

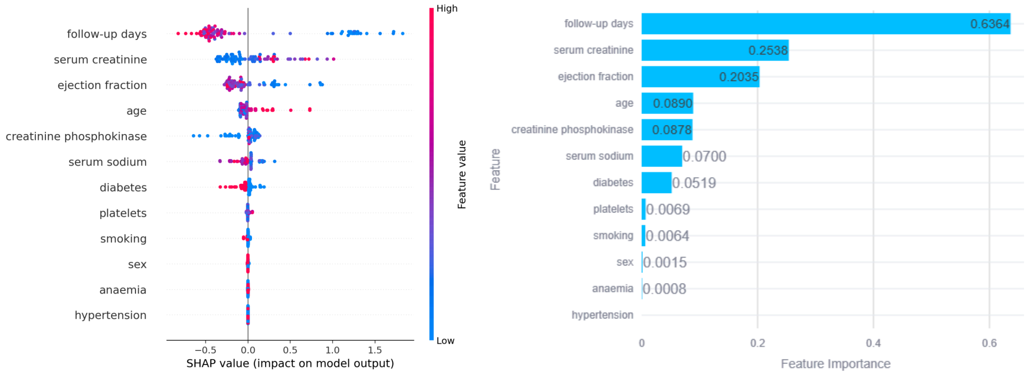


Figure 19. SHAP Analysis Results

This analysis underscores the importance of model selection and interpretability, especially in clinical contexts where understanding prediction rationale is as critical as model performance. As mentioned before, the dataset's limited size of 299 instances presents challenges such as overfitting and limited generalizability, potentially impacting prediction reliability on broader populations. To address this, various classifiers were explored, with SHAP values used to interpret feature contributions to the DEATH_EVENT outcome. Notably, SHAP values identified "follow-up time" as an influential feature, suggesting that patients' observation duration significantly affects predictions. While follow-up time can relate to mortality outcomes, its prominence here may reflect monitoring duration rather than specific health risks, potentially skewing results toward outcomes dependent on observation time rather than inherent patient characteristics. This underscores a limitation in interpretability, as follow-up time may confound relationships between clinical markers and mortality risk. Overall, despite dataset and feature limitations, this study demonstrates that small datasets can yield meaningful insights when coupled with careful model selection and interpretability tools. Future analyses incorporating larger datasets and a reduced focus on time-dependent variables could improve model robustness, supporting more clinically interpretable and objective predictions to aid the intended users in informed, data-driven decision-making.

3. Early-Stage Prototyping

3.1. Early-Stage Prototyping Decisions and Activities

UML and persona are important components in usability design. Hence, these were accounted for during the early stage of prototyping so that the project does not divert from its target users. Specifically, a UML use case diagram is valuable in requirements engineering for web medical dashboards [65], since it illustrates the target users and the proposed solution conceptually. Figure 20 presents the UML use case diagram for this project, with the main use case flow detailed in Table 3 and the alternative flow in Table 4. Meanwhile, personas help define and understand the target users [66–68], which, in this project, included general practitioners, cardiologists, and cardiac nurses. Persona also serves to communicate user characteristics and goals. The advantage of using personas in the development process is that it helps to ensure that systems/solutions are designed with the users' needs at the forefront. For the Heart to Say project, it allowed the team to focus on designing the web dashboard for targeted user groups rather than generic users. Figure 21-23 demonstrates the various personas the team drew inspiration from to design the web dashboard.

Another crucial point was recognizing that healthcare professionals operate in high-stress environments and cannot manage overly complex systems, which could hinder adoption due to low usability. The team, composed of prospective health informaticians, understood the importance of designing socio-technical solutions with an emphasis on user-centered design [69]. Consequently, the team adhered to the WCAG framework, focusing on delivering an accessible dashboard throughout each stage of prototyping [1]. The team identified the following core principles from WCAG that should be prioritized in the design philosophy of the web dashboard:

- **Perceivable**: A streamlined user interface that is logically structured and presented.
- **Operable**: Dashboard functionality should be easily visible and enable straightforward interaction.
- **Understandable**: The dashboard should use healthcare-appropriate terminology, avoiding overly technical jargon that is outside of the healthcare domain.
- **Robust**: The dashboard should be widely accessible without compromising computational performance, accommodating a range of target users.

Additionally, the team adopted ISO 9241-210:2019, which emphasizes human-centered design principles [70]. This standard recommends the use of personas in design development, providing recognizable visual cues for users to complete tasks, ensuring a streamlined interface that avoids information overload for users, and prioritizing data privacy and security, particularly given the handling of health data.

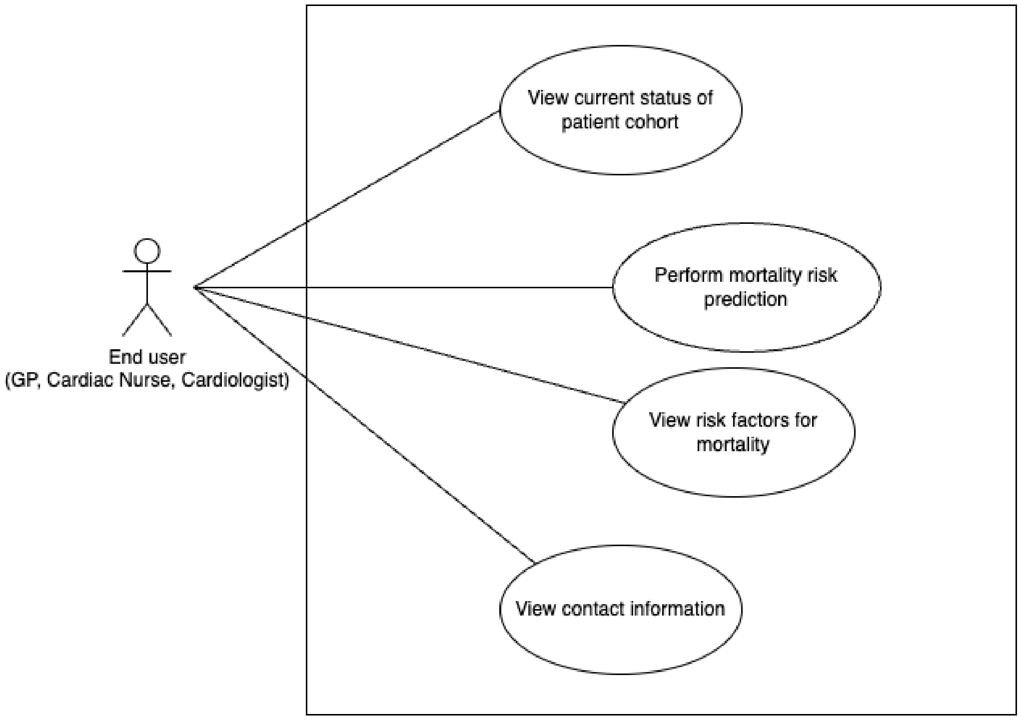


Figure 20. UML use case model with identified requirements for the dashboard

Table 3. Use case flow of events

Step	Description	Comments
1. User logs into the system	The end user (GP, Cardiac Nurse, Cardiologist) logs into the web dashboard using their credentials.	Ensures secure access for healthcare professionals since the dashboard handles sensitive patient data.
2. User welcomed to a home page	Users are welcomed to a home page where the different tabs are visible: View current cohort status, Mortality risk prediction, View risk factors for mortality, Contact information	This step provides quick access to core functionalities, ensuring efficient navigation and ease of use. 2A: If the system home page is inaccessible
3. View current patient status	The user opts to view the current status of the cohort of heart failure patients.	Provides real-time information for informed decision-making.
4. Perform mortality risk prediction	The user chooses to perform a mortality risk prediction on an individual patient by inputting data as required by the dashboard.	Helps assess the patient's risk for mortality and prioritize their care.
5. View risk factors for mortality	The user reviews factors contributing to the patient's risk of mortality, such as	Identifies key risk factors to improve management strategies.

	comorbidities, lifestyle or lab results.	
6. View contact information	The user can view or contact the team behind the web dashboard	For support enquiries.
7.	End of use case	

Table 4. Alternative use case flow derived from step 2 from Table 3.

Step	Description	Comments
2A. System home page is inaccessible	The user attempts to access the home page, but the system encounters an error or technical issue.	N/A
2B. System displays an error message	The system displays an error message with contact information to the team behind the web dashboard.	The error message guides the user on how to proceed.
2C. User contacts support	The user follows the error message instructions and contacts the team for assistance.	The support team has been notified of the issue so they can restore access.
2D. Retry login after the issue is resolved	Once the issue is resolved, the user retries logging into the system to access the home page.	The user can now proceed with the regular flow from Step 2 if access is successfully restored.
	End of use case	

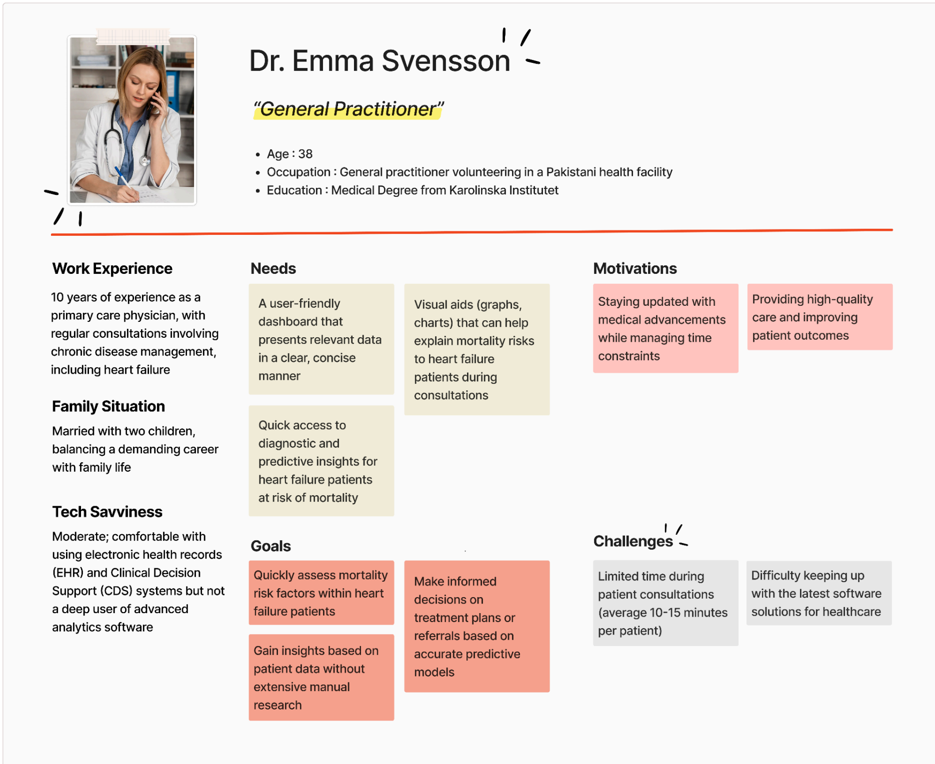


Figure 21. General practitioner's persona

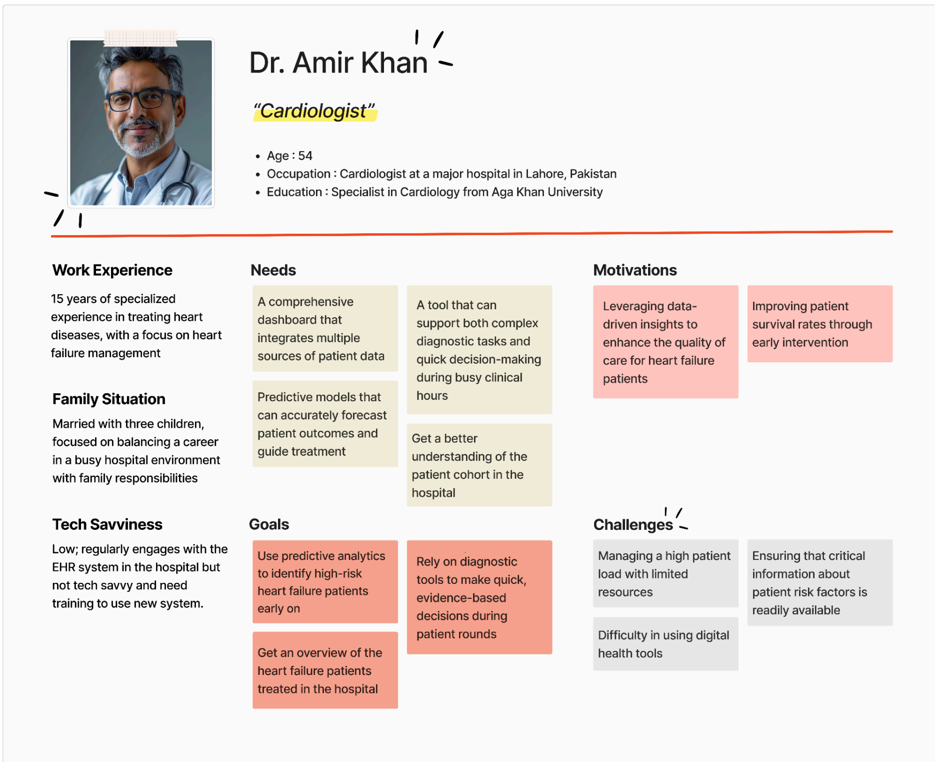


Figure 22. Cardiologist's persona

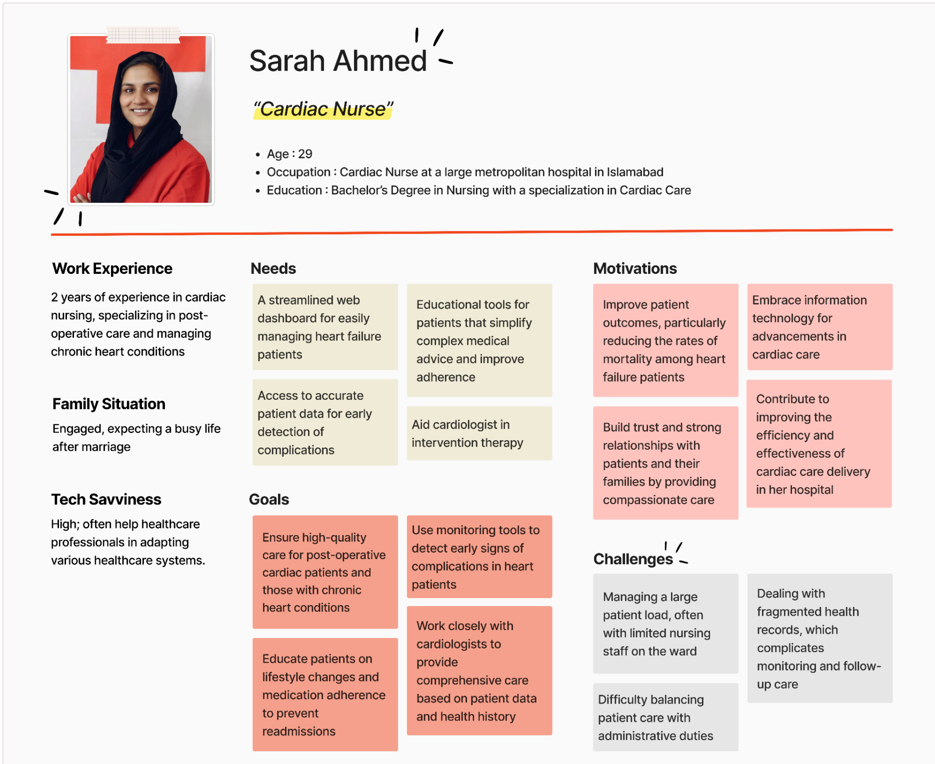


Figure 23. Cardiac nurse's persona

3.2. Low-Fidelity Prototypes Design Process

The following figures illustrate the low-fidelity prototype design process, which began with brainstorming (Figure 24) on the dashboard's visual layout. During this phase, the team conceptualized a general overview of the layout for each tab. Once the layout was outlined, a structured paper prototype (Figures 25-29) was created to provide a clearer vision of the dashboard's appearance. It's important to note that the paper prototype was developed prior to usability testing, so the tab names, layout, and content in the figures differ from those used in the final version.

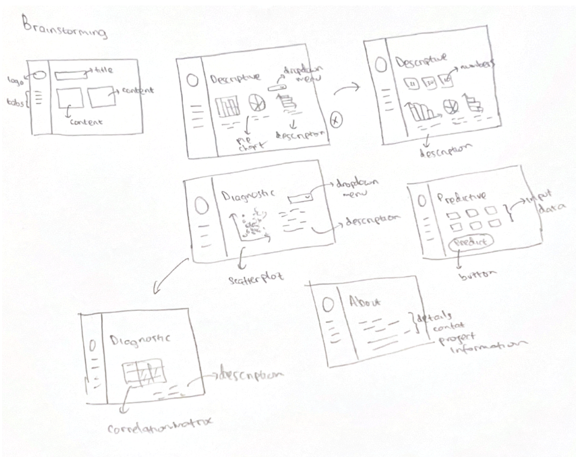


Figure 24. Brainstorming

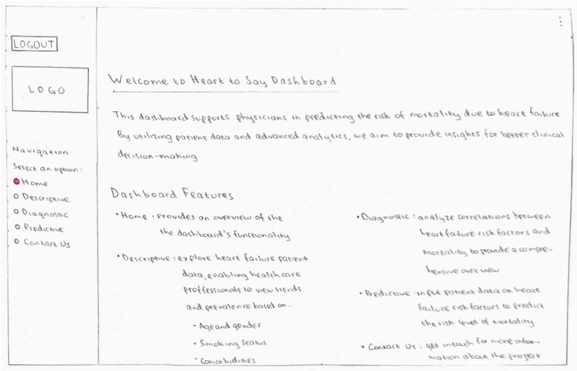


Figure 25. Paper prototype of the dashboard: Home tabs

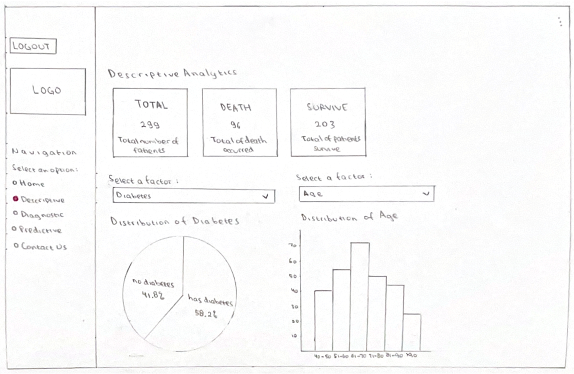


Figure 26. Paper prototype of the dashboard: Descriptive Analytics tabs

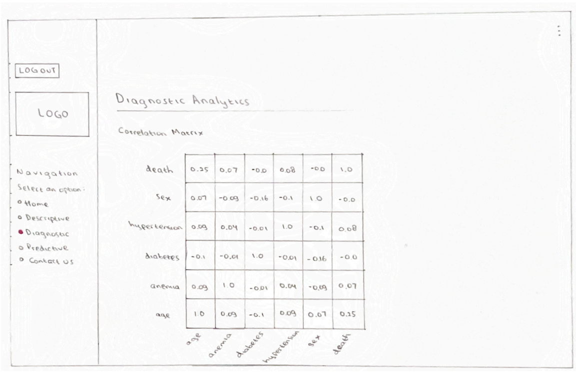


Figure 27. Paper prototype of the dashboard: Diagnostic Analytics tabs

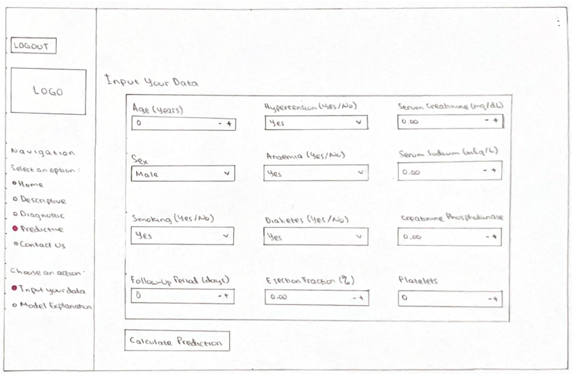


Figure 28. Paper prototype of the dashboard: Predictive Analytics tabs

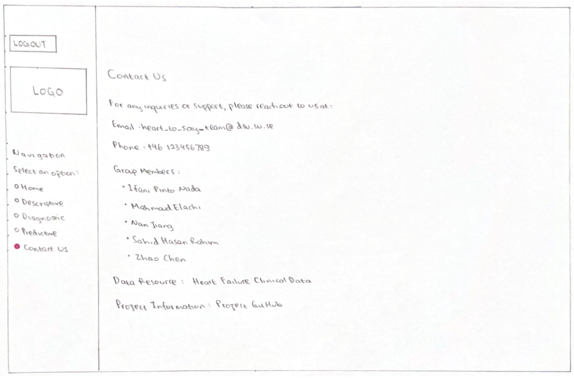


Figure 29. Paper prototype of the dashboard: Contact Us tabs

3.3. Dashboard Layout for Users Getting Answers for Medical Problems

The preliminary dashboard layout was roughly designed, drawing inspiration from Norman's design thoughts [71], while keeping WCAG and ISO standards in mind [1,70], as well as the target users. The preliminary layout of the web dashboard and its user interface components were strategically placed to support the target users in informed decision-making regarding heart failure patients. Here is an overview of the layout and specific features of the dashboard:

1. **Structured Navigation**: The left sidebar offers a concise and intuitive navigation menu, including options such as *Home*, *Descriptive tab*, *Diagnostic tab*, *Prescriptive and Predictive tab*, and *Contact Us*. This structure enables the target users to efficiently locate specific sections based on their immediate needs. Such an organization reduces cognitive load and facilitates seamless navigation through the dashboard [72].
2. **Dashboard Features**:
   - **Home**: This section provides an overview of the dashboard's purpose and key functionalities. It serves as an introductory space that helps users quickly know what features and tools are available.
   - **Descriptive tab**: This feature provides analytical information describing the dataset. On top of the screen, there are color-coded cards to quickly convey the status of the dataset. Using color in the design was crucial as it helps to convey meaningful information to users [73]. This color coding is also represented in the graphs on the left and right sides of the interface, divided into categorical and numerical features. The graphs display the distribution of patients within each feature, along with how subgroups within each feature are associated with mortality outcomes. Headings describing the graphs are provided to inform the users about what the graphs are visualizing.
   - **Diagnostic tab**: This section facilitates the analysis of correlations between various features in the dataset and mortality. The users are given options for diagnostic analytics of the dataset. This design choice aims to prevent cognitive overload by reducing clutter on the interface, enabling

users to focus on relevant functionalities without being overwhelmed [72]. By segmenting the diagnostic options, the dashboard supports a more intuitive and user-friendly experience. Within each option, users are provided a drop-down menu to select between the features. The results are then presented in cards. Colors in the text are also used to highlight significant information.

  - **Prescriptive and Predictive tab**: The layout for the Predictive tab allows users to input values for each parameter using dropdowns for categorical fields and numeric input fields with increment buttons for continuous measures. The "Calculate Prediction" button enables quick risk analysis based on the entered data, generating immediate insights into mortality risk. The “Reset” button allows users to clear the inputs effortlessly, enabling them to conduct multiple predictions without manual data clearing, which is especially useful in fast-paced clinical environments. The SHAP section of the dashboard helps users quickly identify and interpret the most influential factors in predicting mortality risk. Key features are ranked by importance and displayed visually, showing both their impact magnitude and direction on the model’s output. This information is important for the users to validate the prediction output for clinical decision-making. The SHAP option only shows up when the user selects the Prescriptive and Predictive tab to the left. Ideally, the SHAP option is presented on the left sidebar.
  - **Contact us**: This section presents the problem description, project aim, relevant references, and contact information. Essential contact details, such as a telephone number and email address, are provided as clickable hyperlinks, enabling users to initiate direct communication easily. This layout not only enhances accessibility but also ensures that users can promptly reach out for more information or support, enabling effective communication and connection with the team.

3. **Enhanced Accessibility and Usability**: The dashboard’s organized structure, accompanied by visual cues, promotes intuitive navigation, ensuring that users can quickly access and interpret relevant information. Each section is designed to provide both high-level overviews and detailed analyses, allowing users to delve deeper into the specifics of the dataset.

3.4. Advantages of Early-Stage Low-Fidelity Prototyping

The team used paper-based prototypes as a low-fidelity approach for the web dashboard to provide an early conceptualization of the layout, including the placement of navigation tabs, logos, and content areas. Paper-based prototypes are especially beneficial in the early design stages, as they allow for rapid idea generation and easy visualization of concepts [67]. Additionally, they are quick to produce, low cost, and can be easily modified, making them an ideal tool for iterative design. The team learned that paper-based prototypes are a good starting point for conceptualizing ideas before putting them into practice.

4. First High-Fidelity Prototype

4.1. User Testing

User testing was conducted following the first iteration to gather participant feedback on the web dashboard. This input was essential for refining the overall user experience and guiding improvements for the future iterations of the fully functional dashboard. The input collected from potential users provided guidance in making adjustments that align with the users' needs and expectations.

The decision to conduct user testing after the first iteration of the web dashboard was to ensure that the entire web dashboard, with all its key features, could be evaluated comprehensively. If certain elements, such as the predictive tab, had been omitted, the feedback could have been biased. For instance, if participants could not use the predictive tab, they might not have understood the main purpose of the dashboard. This would lead to feedback that does not assess the complete vision of the dashboard. Additionally, any changes made based on this feedback would have only addressed the visible parts of the dashboard. Thus, adding the predictive tab later would have resulted in a feature that might not have fit with the overall dashboard, as it would have been designed without any input.

By presenting the complete dashboard to the user test participants, the team ensured that the feedback received was holistic. This will guide future iterations toward a more functional, user-centered, and cohesive design. Without this full-scale testing approach, there was a real risk that essential components could have been overlooked, leading to a product that was disconnected and only partially usable.

4.2. Status of The First High-Fidelity Prototype Before User Testing

Figures 30-34 show the status of the first high-fidelity prototype before user testing.

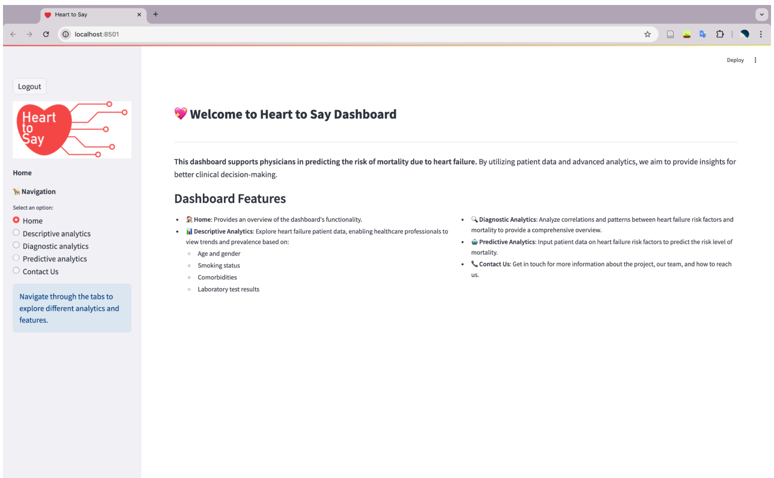


Figure 30. Dashboard before user testing: Home tabs

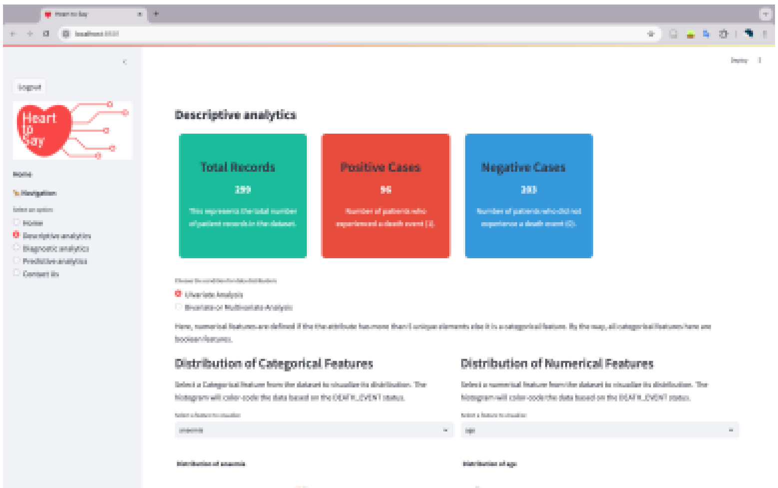


Figure 31. Dashboard before user testing: Descriptive Analytic tabs

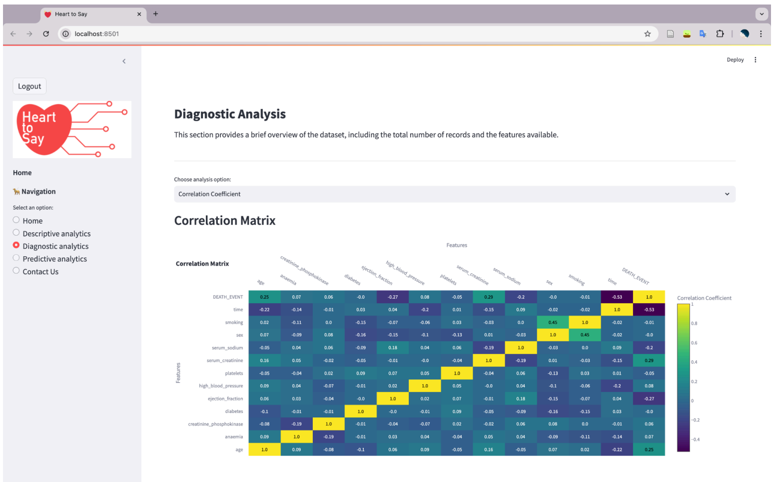


Figure 32. Dashboard before user testing: Diagnostic Analytic tabs

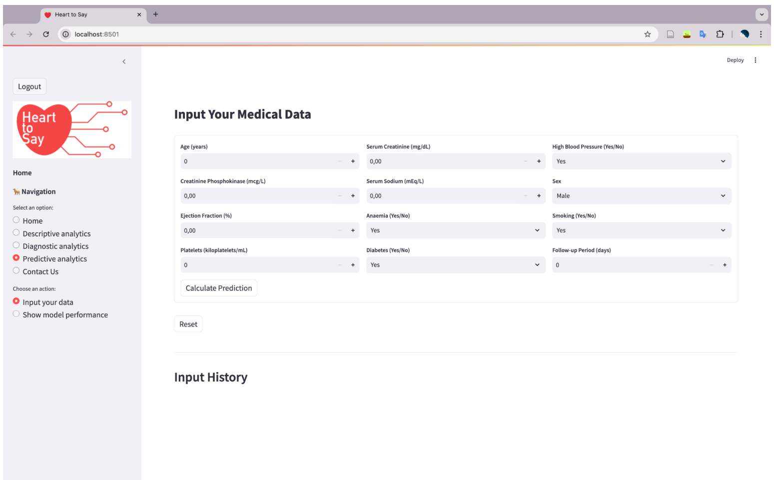


Figure 33. Dashboard before user testing: Predictive Analytic tabs

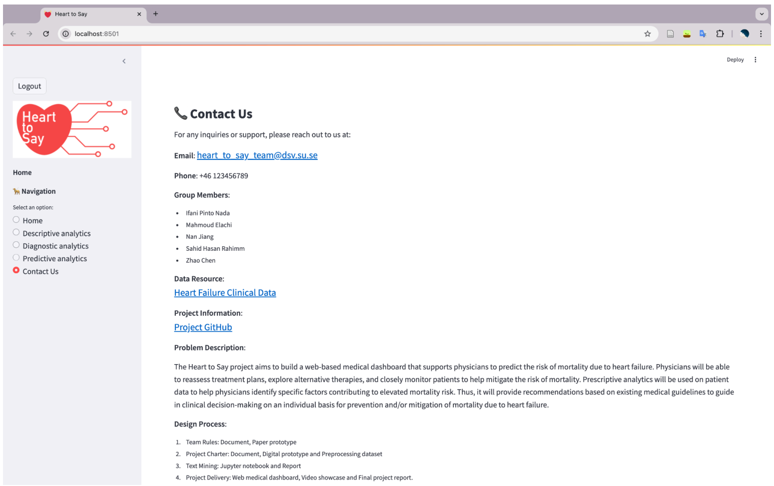


Figure 34. Dashboard before user testing: Contact Us tabs

5. User Testing and Evaluation

5.1. Goals Set For User Testing

User testing is critical to ensure that systems are designed appropriately for the end-user [74]. In this study, usability testing was performed to evaluate the User Interface (UI) and User Experience (UX) of the whole web dashboard, including all of its tabs. The primary goal was to identify potential design issues, learn user behaviors and preferences, and uncover opportunities for improvement. The usability testing was conducted in a controlled environment. In a room where noise and disturbance were minimized, participants were given a laptop with the web dashboard installed, and the page was loaded and logged in. Then, the participants were given brief information about the aim of the study and what they were allowed to do and not to do during the testing. During the 10-20 minute usability testing, at least two members of the team observed and documented what the participants were thinking and doing and any possible issues.

Two hypotheses were constructed:

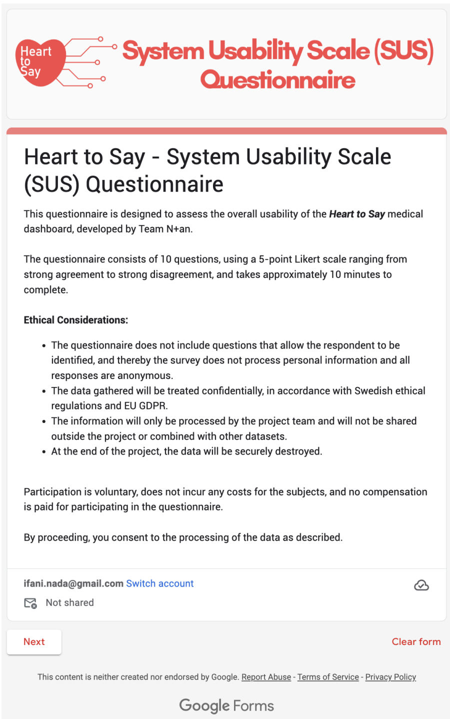
- **Primary Hypothesis:** Participants will be able to navigate the web dashboard intuitively and complete assigned tasks efficiently without significant issues.
- **Secondary Hypothesis:** Usability testing will reveal specific design issues or areas of improvement that, when addressed, will enhance the overall user experience and usability of the dashboard.

The expected outcome of the evaluation is that the participants can navigate within the dashboard and complete the assigned tasks effectively, hence the primary hypothesis. The expected outcome of the second hypothesis is that the evaluation will equip the team with quantitative and qualitative data to guide future UI enhancements, improving overall user experience and usability.

The usability evaluation also addressed ethical considerations by informing and obtaining consent from participants prior to their participation. During the documentation, no personal information was stored. In addition, participants' names were anonymized so that no one could identify them. All photo documentation was captured in a manner that ensured participants could not be identified.

5.2. Quantitative Usability Method

To assess usability quantitatively, the team used the widely-recognized System Usability Scale (SUS) questionnaire. SUS includes ten statements related to user experience, with responses captured on a 5-point Likert scale, ranging from strong agreement to strong disagreement [75–78]. SUS has been extensively validated as a reliable tool for measuring perceived usability [78] and will help to provide insights into users’ overall experience with the web dashboard [75]. The questions were designed to assess users' perceptions of usability by balancing positive and negative statements about the ease of use of the dashboard, functionality integration, complexity, and confidence in usage. Figure 35 and Figure 36 show the SUS questionnaire that was utilized for the quantitative assessment of the Heart to Say dashboard.



Heart to Say

System Usability Scale (SUS) Questionnaire

Heart to Say - System Usability Scale (SUS) Questionnaire

This questionnaire is designed to assess the overall usability of the ***Heart to Say*** medical dashboard, developed by Team N+an.

The questionnaire consists of 10 questions, using a 5-point Likert scale ranging from strong agreement to strong disagreement, and takes approximately 10 minutes to complete.

Ethical Considerations:

- The questionnaire does not include questions that allow the respondent to be identified, and thereby the survey does not process personal information and all responses are anonymous.
- The data gathered will be treated confidentially, in accordance with Swedish ethical regulations and EU GDPR.
- The information will only be processed by the project team and will not be shared outside the project or combined with other datasets.
- At the end of the project, the data will be securely destroyed.

Participation is voluntary, does not incur any costs for the subjects, and no compensation is paid for participating in the questionnaire.

By proceeding, you consent to the processing of the data as described.

ifani.nada@gmail.com Switch account

Not shared

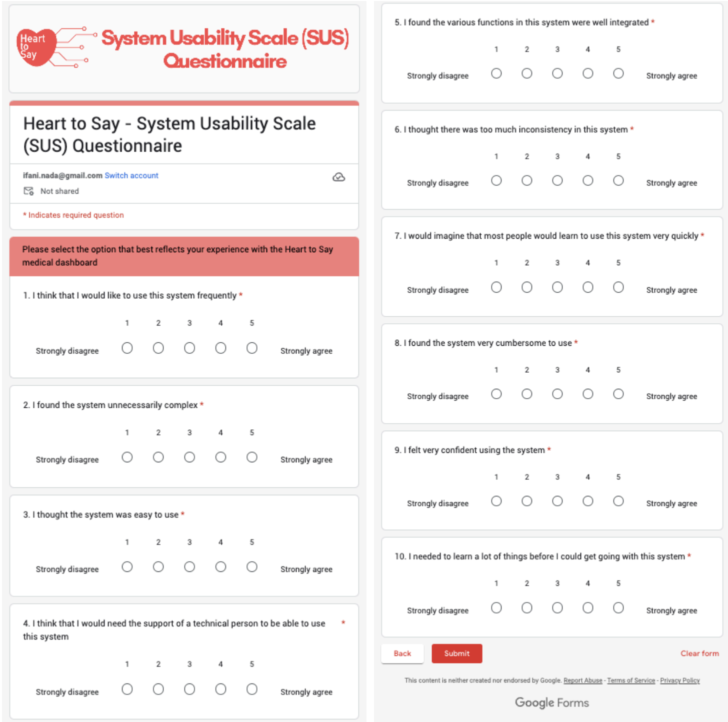
Next

Clear form

This content is neither created nor endorsed by Google. Report Abuse · Terms of Service · Privacy Policy

Google Forms

Figure 35. SUS questionnaire: informed consent page


System Usability Scale (SUS) Questionnaire

Heart to Say - System Usability Scale (SUS) Questionnaire

* Indicates required question

Please select the option that best reflects your experience with the Heart to Say medical dashboard

1. I think that I would like to use this system frequently *

Strongly disagree 1 2 3 4 5 Strongly agree

2. I found the system unnecessarily complex *

Strongly disagree 1 2 3 4 5 Strongly agree

3. I thought the system was easy to use *

Strongly disagree 1 2 3 4 5 Strongly agree

4. I think that I would need the support of a technical person to be able to use this system *

Strongly disagree 1 2 3 4 5 Strongly agree

5. I found the various functions in this system were well integrated *

Strongly disagree 1 2 3 4 5 Strongly agree

6. I thought there was too much inconsistency in this system *

Strongly disagree 1 2 3 4 5 Strongly agree

7. I would imagine that most people would learn to use this system very quickly *

Strongly disagree 1 2 3 4 5 Strongly agree

8. I found the system very cumbersome to use *

Strongly disagree 1 2 3 4 5 Strongly agree

9. I felt very confident using the system *

Strongly disagree 1 2 3 4 5 Strongly agree

10. I needed to learn a lot of things before I could get going with this system *

Strongly disagree 1 2 3 4 5 Strongly agree

Back Submit Clear form

Google Forms

Figure 36. SUS questionnaire: questionnaire page

5.3. Qualitative Usability Method

For the qualitative evaluation, the Think-Aloud method was employed, allowing participants to verbalize their thoughts, motives, and perceptions during their interaction with the dashboard [79,80]. This method is particularly useful for identifying where users may misunderstand the interface, which helps pinpoint areas for design improvement [79–81].

The following tasks were assigned to the participants to be completed during the Think-Aloud method (Figure 37-41):

1. Descriptive Tab

Heart to Say
Dashboard Usability Testing

Task #1	
Task Goal	Find the total number of heart failure patients who have died
Test Scenario	A physician asks you for the total number of heart failure patients who have died. Use the Heart to Say dashboard to locate this information within the descriptive analytics tab. How many death events occurred?

Figure 37. Task number 1

What will be observed:

- How easily participants locate the Descriptive Analytics tab.
- Whether they can efficiently find the relevant statistics showing the number of heart failure patient deaths.
- Any challenges they face navigating or interpreting the presented data.

2. Diagnostic Tab

Heart to Say
Dashboard Usability Testing

Task #2	
Task Goal	Determine the correlation between diabetes and mortality risk
Test Scenario	A researcher wants to know if there is a correlation between diabetes and increased mortality risk in heart failure patients. Use the Heart to Say dashboard to find this information using the diagnostic analytics tab. Is there any correlation between diabetes and mortality risk in heart failure patients?

Figure 38. Task number 2

What will be observed:

- How easily participants locate the Diagnostic Analytics tab.
- Whether participants are able to find the correlation graph between diabetes and mortality among heart failure patients.
- Any challenges they face navigating or interpreting the presented data.

3. Prediction Tab

Heart to Say
Dashboard Qualitative Usability Testing

Task #3	
Task Goal	Predict the risk of mortality for a newly admitted heart failure patient
Test Scenario	A newly admitted heart failure patient has arrived at your hospital. Your task is to predict whether this patient has a high risk of mortality using the predictive analytics tools in the Heart to Say dashboard. The patient's information includes 63 years old, male, who does not have anemia, has both hypertension and diabetes, and is a smoker. The patient has the following lab results: - Creatinine phosphokinase: 124 mcg/L - Ejection fraction: 40% - Platelets: 162000 kiloplatelets/mL - Serum creatinine: 1,2 mg/dl - Serum sodium: 137 mEq/L - Follow-up period: 11 days Do the newly admitted patients have a high risk of mortality?

Figure 39. Task number 3

What will be observed:

- How easily participants navigate to the Predictive Analytics tab.
- Whether they are able to input the patient's clinical data and run the prediction model.

4. Prediction and Prescriptive Tab

Heart to Say
Dashboard Usability Testing

Task #4	
Task Goal	Analyze and interpret risk factors for heart failure patients using SHAP explanations
Test Scenario	Imagine you are a healthcare provider reviewing a heart failure patient's data to understand their risk of death event. On the Heart to Say dashboard, explore the patient's clinical details (e.g., age, gender, anemia, hypertension, diabetes, and smoking status), and review the SHAP explanations to understand how each factor contributes to the predicted likelihood of a death event. Based on these insights, which factors pose the greatest risk and suggest potential focus areas for managing the patient's condition?

Figure 40. Task number 4

What will be observed:

- Whether users experience any difficulties navigating to the intended tab containing SHAP visualizations.
- How easily participants interpret SHAP visualizations and connect them to the patient's risk of a death event.
- Whether users can identify the most critical risk factors based on the SHAP values.

5. Contact Us Tab

Heart to Say
Dashboard Usability Testing

Task #5	
Task Goal	Find how to contact support for dashboard-related questions.
Test Scenario	You have a suggestion for the dashboard. Use the Contact Us tab to find out how you can contact the support team. What is their email or phone number?

Figure 41. Task number 5

What will be observed:

- How quickly and easily participants locate the Contact Us tab.
- Whether they can find the relevant contact information without confusion.

Additionally, a brief post-task interview was conducted to gather follow-up information. This will clarify observations made during the session and provide further feedback on the overall user experience [82], enriching the qualitative data collected. The following questions were asked after task completion:

1. How would you describe your overall experience using the dashboard?
2. Did you find it easy to navigate the dashboard and locate the information you needed?
3. What did you like most about the dashboard?
4. Were there any particular tasks or features that you found difficult or confusing?
5. Are there any features or changes you would suggest to improve the dashboard?

5.4. Prototype Evaluation

Both quantitative and qualitative usability evaluations were conducted with careful ethical considerations. For the quantitative evaluation, informed consent was provided on the first page (Figure 35, Section 5.2) before participants proceeded with the questionnaire. For the qualitative evaluation, at the start of the evaluation, participants were first informed about what type of evaluation they would take part in and how the team ensured their anonymity and privacy, followed by verbal consent. Additionally, participants were also encouraged to raise any questions or concerns at any point before, during, or after the evaluation.

Photos were taken during the evaluation, and verbal consent was obtained from each participant beforehand. To protect participant privacy, all photos were taken from behind and any visible faces were blurred (Figure 42). In total, five participants took part, all of whom were general practitioners and a nurse, and were aligned with the intended user group. However, due to an error in the photo taking, one participant was not captured during the session, so only four participants appear in the following figure (Figure 42).

Figure 42. Prototype evaluation

To ensure the evaluations were conducted with integrity, a controlled environment was reserved where participants were placed in a quiet room free from outside noise, mobile notifications, and other potential distractions. During the first part of the usability testing, participants were highly focused and engaged. However, an unforeseen incident, a gas leak in the building, required relocating to a common area near a café in a neighboring building. This location introduced background noise that might have distracted participants. Despite this event, participants remained cooperative and understanding of the unexpected situation. This experience reinforced the importance of using a controlled environment for usability testing to allow participants to fully concentrate.

After completing the usability testing, each participant took part in a follow-up interview, where they were asked questions outlined in Section 5.3. This interview aimed to capture additional insights into their experiences, allowing participants to share specific feedback on the dashboard's usability and any challenges they encountered. Once the qualitative usability testing was concluded, participants were provided with a QR code that directed them to an online questionnaire (Figure 43). This questionnaire served as a structured method to gather additional quantitative data, ensuring that both qualitative and quantitative perspectives were captured to inform a comprehensive evaluation of the dashboard's usability. Figure 44 displays the results from the questionnaire.

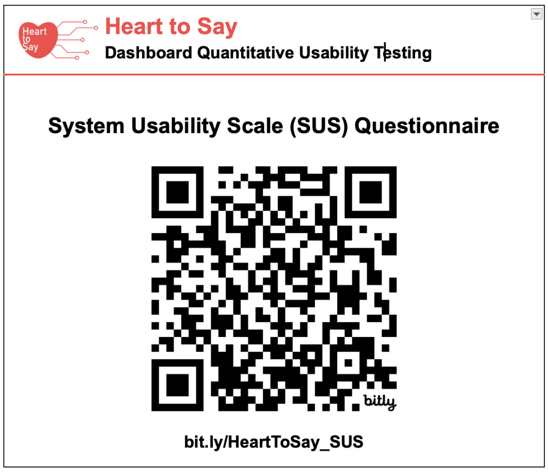


Figure 43. QR code for the questionnaire

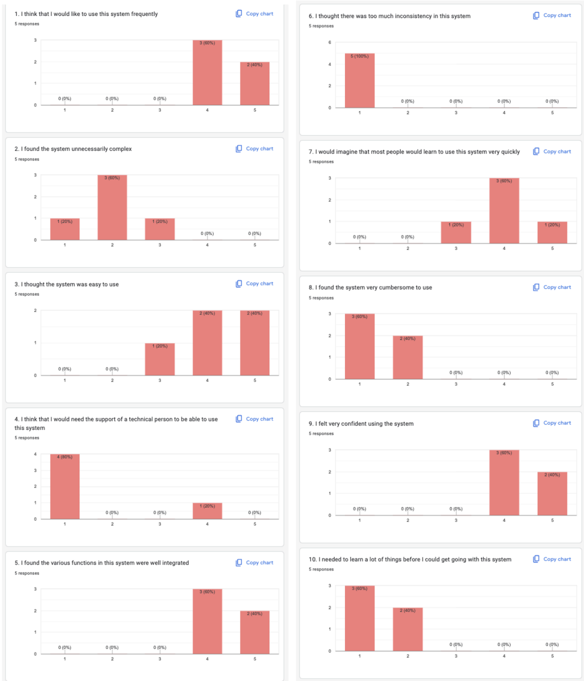


Figure 44. SUS questionnaire results

6. Final Working Dashboard

6.1. Evaluation Stage Lessons

For the quantitative usability testing using the SUS questionnaire, the average final score was 85 (grade A+), indicating excellent usability and learnability [83]. However, it is important to note that only five participants completed the questionnaire, which limits the representativeness of these results. Quantitative methods generally require a larger sample size, typically 20-30 users, for more reliable and generalizable outcomes [77]. This small participant pool may therefore impact the reliability of the findings.

For the qualitative data, a thematic analysis was first conducted as it is useful in identifying patterns from feedback to gain insights into users' experiences, preferences, and perceptions [84]. The same limitation applies to the data gathered here since it only included data from five participants. Table 5 outlines the themes identified with their corresponding description.

Table 5. Themes derived from the thematic analysis along with description

Theme	Description
Design Simplicity	Positive feedback on the dashboard’s clean, simple, and clear design.
Content Density	Some pages felt too lengthy; preference for more concise content.

Information Arrangement	Suggestions to rearrange information for better visibility and accessibility.
Unnecessary Content	Some options in the diagnostic tab felt too overwhelming, with a preference to keep clustering and correlation.
Label Clarity	Need for clearer labels, especially in distinguishing between death and survival cases, as well as providing more understandable tab descriptions.
Terminology Simplification	Preference for less technical language, making terms more user-friendly.
Understanding Complex Analysis	Difficulty in understanding complex analysis like SHAP due to unfamiliarity.

From the thematic analysis, it was derived that participants demonstrated strong proficiency in using the dashboard. Participants overwhelmingly commented that its design was clean, simple, and clear. However, they noted that some pages felt lengthy. Initially, scrollable content was provided to include all relevant information, but based on feedback, the content was reduced and replaced with concise summaries for clarity. For example, in the Diagnostic tab, the summary for the correlation matrix initially appeared below the graph, leading some participants to overlook it. This was adjusted by moving the summary above the graph and removing irrelevant information. Some content was also organized within a menu, with different diagnostic analytics segmented as options under the Diagnostic tab for users to select. This removed unnecessary scrolling and cluttering of the interface. Additionally, participants suggested rearranging information and adding summaries to plots and figures.

Participants also noted that the Diagnostic tab included an excessive number of highly technical analytics, expressing a preference for more informative options targeted to the user, such as clustering and correlation. Consequently, other analytics, such as logistic regression, were removed. This adjustment was considered appropriate, as certain diagnostic insights were already provided through the SHAP analysis, and including logistic regression information could create confusion for users when compared to the SHAP analysis.

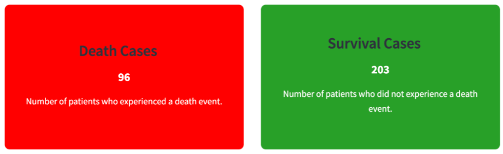
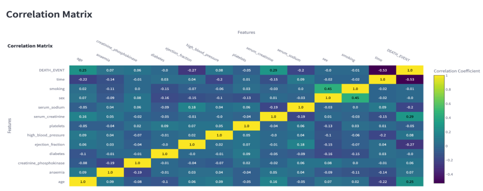
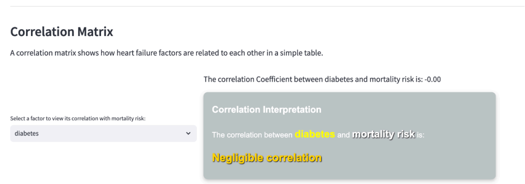
Additional feedback focused on label clarity and color used to distinguish between death and survival cases. For example, the initial label for death cases, “positive cases,” was confusing; it was revised to accurately reflect the data. The dashboard’s color scheme was also updated to align with agreed color codes [73]. Moreover, there were suggestions for changes to tab labels and reducing technical terminology. Consequently, tab labels and content were revised to more user-friendly terms for the intended users. Moreover, in the Predictive and Prescriptive tab, there were concerns about the difficulty in understanding SHAP analysis, as it was unfamiliar to the participants. This was changed by adding a clear summary of SHAP results below the graph, along with a short description of what it entails. Examples of adjustments made are shown in Table 6.

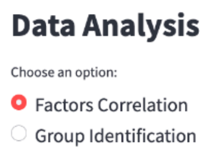
Table 6. Example of adjustments made based on usability testing feedback

<table>
<tr><th>Before</th><th>After</th></tr>
<tr><td>Tabs labeled with technical terms, potentially confusing intended users

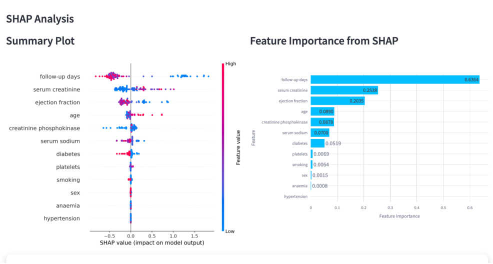
</td><td>Tabs changed to more user-friendly terms: “Descriptive Analytics,” “Diagnostic Analytics,” and “Predictive Analytics” replaced with simpler labels

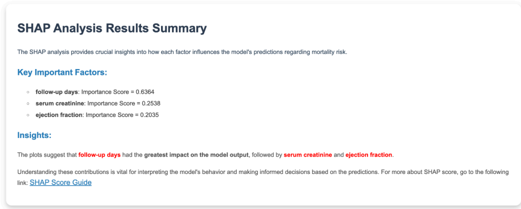
</td></tr>
<tr><td>Ambiguity around “Positive Cases” and “Negative Cases” labels

</td><td>Terms revised to “Death Cases” and “Survival Cases” for clarity

</td></tr>
<tr><td>Plot summary placed at the bottom of graphs

</td><td>The plot summary is moved to the top of the graphs for better visibility

</td></tr>
<tr><td>No option for diagnostic analytics

</td><td>Option for the preferable diagnostic analytics

</td></tr>
</table>

Before	After
No SHAP analysis summary provided	A summary of the SHAP analysis is added below the graph





Through usability testing, the team gathered extensive feedback that has been essential in refining the dashboard to meet user needs and improve its functionality for the target users. While feedback can be subjective, the team reviewed all suggestions carefully and implemented the changes most beneficial to this project. Nevertheless, the team recognized the importance of conducting further evaluations with a larger group of participants to obtain more precise feedback. This will ensure that future iterations stay in line with the goal of the project.

6.2. Final Dashboard

The final working Heart to Say dashboard can be accessed using this hyperlink: https://heart-to-say-medical-dashboard-rsgsxcheqo438vawrmyct2.streamlit.app/. Figures 45-51 show the final working web dashboard:

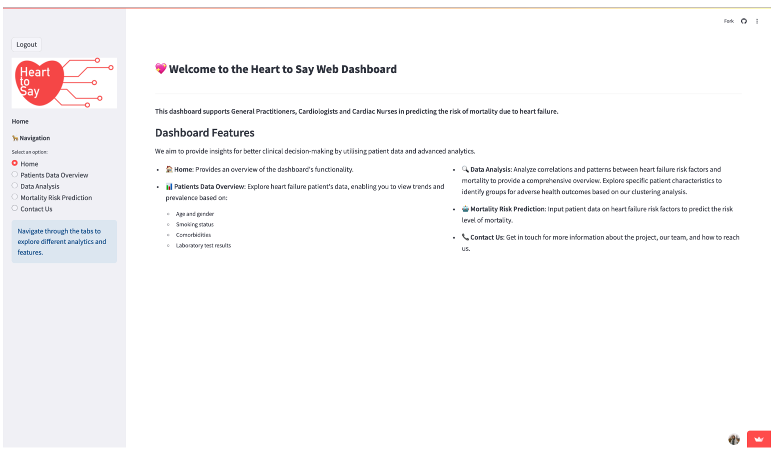


Figure 45. Final working dashboard: Home tabs

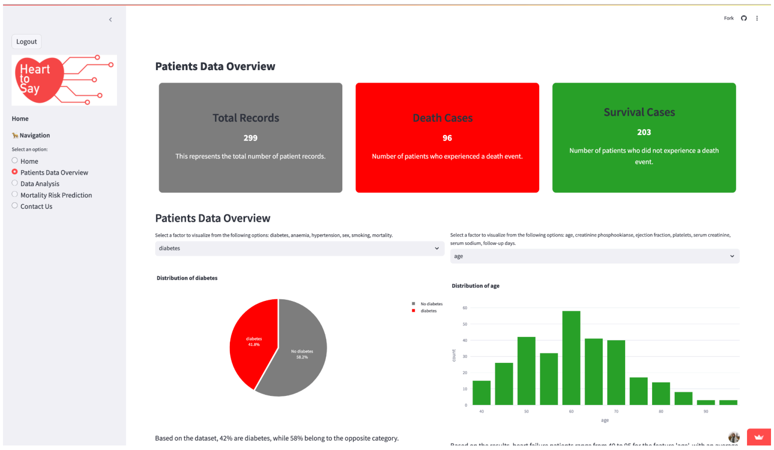


Figure 46. Final working dashboard: Patients Data Overview tabs

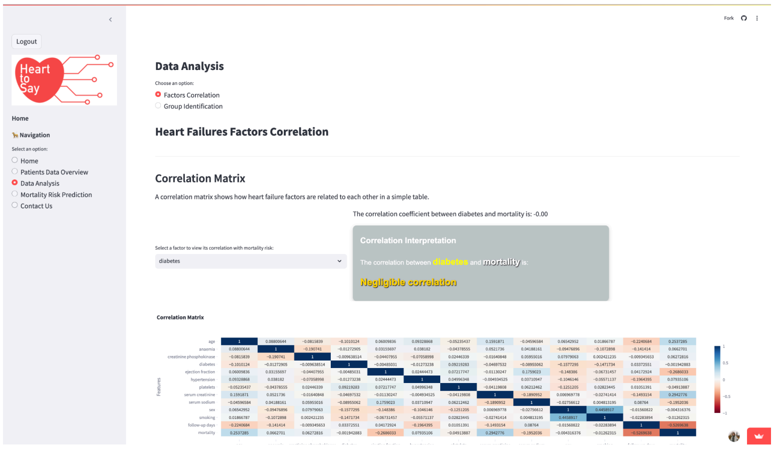


Figure 47. Final working dashboard: Data Analysis tabs-Correlations

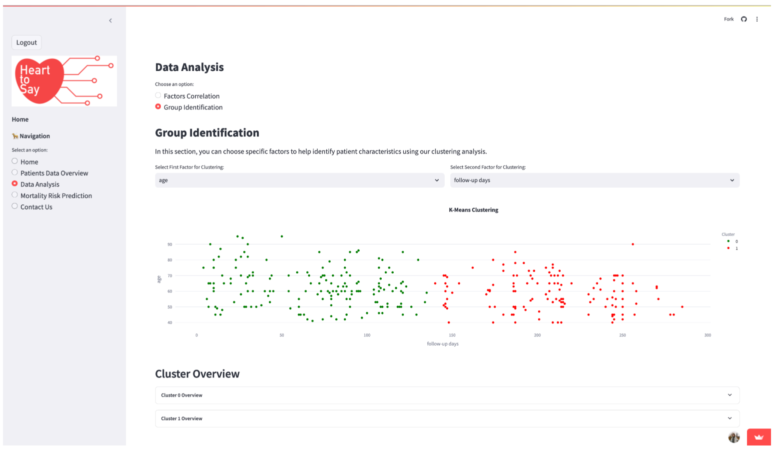


Figure 48. Final working dashboard: Data Analysis tabs-Clustering

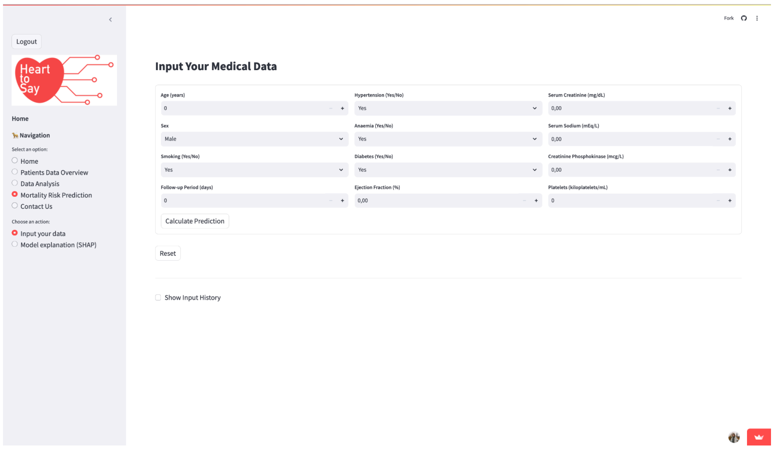


Figure 49. Final working dashboard: Mortality Risk Prediction tabs-Input Your Data

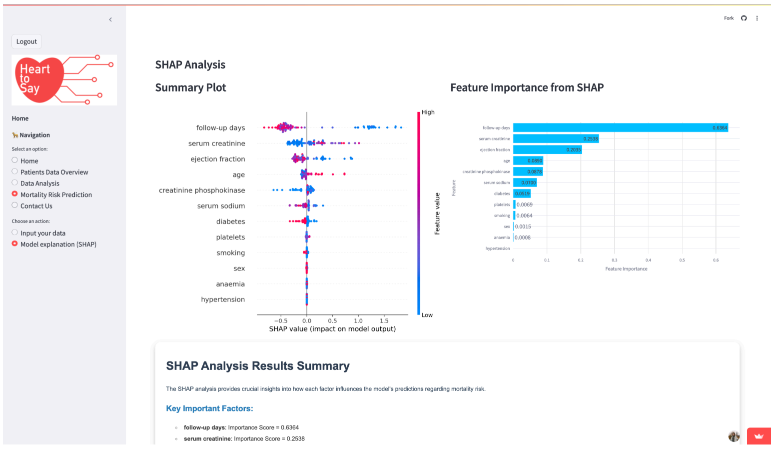


Figure 50. Final working dashboard: Mortality Risk Prediction tabs-Model Explanation (SHAP)

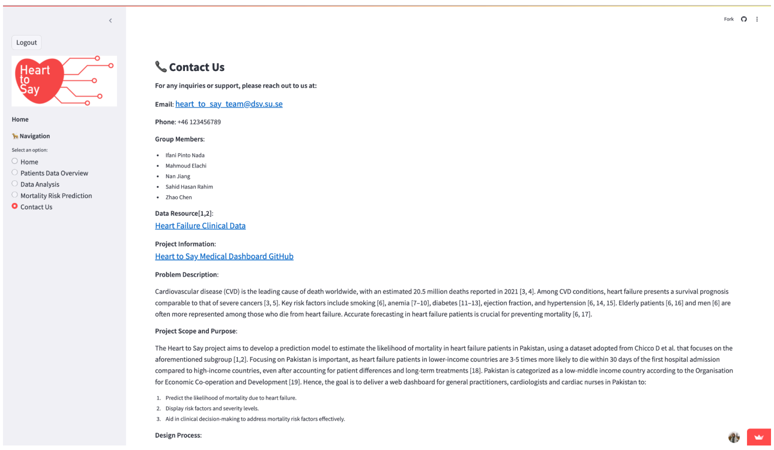


Figure 51. Final working dashboard: Contact Us tabs

Reference

1. World Wide Web Consortium (W3C). How to Meet WCAG (Quickref Reference) [Internet]. [cited 2024 Oct 14]. Available from: https://www.w3.org/WAI/WCAG22/quickref/?versions=2.1#parsing
2. Vaida S, Șerban D. Group Development Stages. A Brief Comparative Analysis of Various Models. Studia Universitatis Babeș-Bolyai Psychologia-Paedagogia. 2021 Jun 1;LXVI:91–110.
3. Bullwinkle K. DiSC leadership types [Internet]. discprofiles.com. 2022 [cited 2024 Oct 11]. Available from: https://www.discprofiles.com/disc-leadership-types/
4. Irawanto DW, Ramsey PL, Ryan JC. Tailoring Leadership Theory to Indonesian Culture. Global Business Review. 2011 Oct 1;12(3):355–66.
5. Babcock J. Great Leadership Requires Emotional Maturity: Here's Why [Internet]. Leaders.com. 2022 [cited 2024 Oct 12]. Available from: https://leaders.com/articles/personal-growth/emotional-maturity/
6. Babiker A, El Husseini M, Al Nemri A, Al Frayh A, Al Juryyan N, Faki MO, et al. Health care professional development: Working as a team to improve patient care. Sudan J Paediatr. 2014;14(2):9–16.
7. McKinsey. The journey to an agile organization [Internet]. McKinsey & Company. [cited 2024 Oct 12]. Available from: https://www.mckinsey.com/capabilities/people-and-organizational-performance/our-insights/the-journey-to-an-agile-organization
8. Kadenic MD, Koumaditis K, Junker-Jensen L. Mastering scrum with a focus on team maturity and key components of scrum. Inf Softw Technol. 2023 Jan 1;153:107079.
9. Fischer RJ, Halibozek EP, Walters DC. Contingency Planning Emergency Response and Safety. Introduction to Security. 2019;249–68.
10. Zarghami SA. Resilience to disruptions: a missing piece of contingency planning in projects. Int J Prod Res. 2024;62(17):6029–45.
11. Voler Systems. Project Management Series: The Importance of Defined Requirements in Project Management [Internet]. [cited 2024 Oct 12]. Available from: https://www.volersystems.com/blog/the-importance-of-defined-requirements-in-project-management
12. Scrum.org. What is Scrum? [Internet]. [cited 2024 Oct 11]. Available from: https://www.scrum.org/resources/what-scrum-module
13. Schwaber K, Sutherland J. The Scrum Guide [Internet]. Scrum Guides. 2020 [cited 2024 Oct 12]. Available from: https://scrumguides.org/scrum-guide.html
14. Gallagher A, Dunleavy J, minutes PRA 2019Time to read: 5. The Waterfall Model: Advantages, disadvantages, and when you should use it [Internet]. IBM Developer. 2019 [cited 2024 Oct 12]. Available from: https://developer.ibm.com/articles/waterfall-model-advantages-disadvantages/
15. Nieto-Rodriguez A. It's Time to End the Battle Between Waterfall and Agile. Harvard Business Review [Internet]. 2023 Oct 10 [cited 2024 Oct 12]; Available from: https://hbr.org/2023/10/its-time-to-end-the-battle-between-waterfall-and-agile
16. Mossanen M, Johnston SS, Green J, Joyner BD. A Practical Approach to Conflict Management for Program Directors. J Grad Med Educ. 2014 Jun;6(2):345–6.
17. Scrum.org. Using Jira to Practice Professional Scrum [Internet]. 2021 [cited 2024 Oct 12]. Available from: https://www.scrum.org/resources/using-jira-practice-professional-scrum
18. Google. Google Workspace: Secure Online Productivity & Collaboration Tools [Internet]. Google Workspace. [cited 2024 Oct 20]. Available from: https://workspace.google.com/
19. GitHub. GitHub: Build software better, together [Internet]. GitHub. [cited 2024 Oct 20]. Available from: https://github.com
20. Zoom. Zoom Video Communications: One platform to connect [Internet]. Zoom. [cited 2024 Oct 20]. Available from: https://zoom.us
21. Scrum.org. What is a Daily Scrum? [Internet]. [cited 2024 Oct 12]. Available from: https://www.scrum.org/resources/what-is-a-daily-scrum
22. WhatsApp. About supported operating systems | WhatsApp Help Center [Internet]. [cited 2024 Oct 20]. Available from: https://faq.whatsapp.com/1150261202542208#
23. Streamlit. Streamlit: A faster way to build and share data apps [Internet]. Streamlit. 2021 [cited 2024 Oct 20]. Available from: https://streamlit.io/
24. Spitaleri G, Zamora E, Cediel G, Codina P, Santiago-Vacas E, Domingo M, et al. Cause of Death in Heart Failure Based on Etiology: Long-Term Cohort Study of All-Cause and Cardiovascular Mortality. J Clin Med. 2022 Jan 31;11(3):784.
25. Savarese G, Becher PM, Lund LH, Seferovic P, Rosano GMC, Coats AJS. Global burden of heart failure: a comprehensive and updated review of epidemiology. Cardiovasc Res. 2023 Jan 18;118(17):3272–87.

26. Mamas MA, Sperrin M, Watson MC, Coutts A, Wilde K, Burton C, et al. Do patients have worse outcomes in heart failure than in cancer? A primary care-based cohort study with 10-year follow-up in Scotland. Eur J Heart Fail. 2017 Sep;19(9):1095–104.
27. Bozkurt B, Ahmad T, Alexander KM, Baker WL, Bosak K, Breathett K, et al. Heart Failure Epidemiology and Outcomes Statistics: A Report of the Heart Failure Society of America. J Card Fail. 2023 Oct;29(10):1412–51.
28. Köseoğlu FD, Özlek B. Anemia and Iron Deficiency Predict All-Cause Mortality in Patients with Heart Failure and Preserved Ejection Fraction: 6-Year Follow-Up Study. Diagnostics. 2024 Jan;14(2):209.
29. Siddiqui SW, Ashok T, Patni N, Fatima M, Lamis A, Anne KK. Anemia and Heart Failure: A Narrative Review. Cureus. 2022;14(7):e27167.
30. Groenveld HF, Januzzi JL, Damman K, van Wijngaarden J, Hillege HL, van Veldhuisen DJ, et al. Anemia and Mortality in Heart Failure Patients: A Systematic Review and Meta-Analysis. Journal of the American College of Cardiology. 2008 Sep 2;52(10):818–27.
31. Xia H, Shen H, Cha W, Lu Q. The Prognostic Significance of Anemia in Patients With Heart Failure: A Meta-Analysis of Studies From the Last Decade. Front Cardiovasc Med. 2021 May 13;8:632318.
32. Dunlay SM, Givertz MM, Aguilar D, Allen LA, Chan M, Desai AS, et al. Type 2 Diabetes Mellitus and Heart Failure: A Scientific Statement From the American Heart Association and the Heart Failure Society of America: This statement does not represent an update of the 2017 ACC/AHA/HFSA heart failure guideline update. Circulation. 2019 Aug 13;140(7):e294–324.
33. Petrie JR, Guzik TJ, Touyz RM. Diabetes, Hypertension, and Cardiovascular Disease: Clinical Insights and Vascular Mechanisms. Can J Cardiol. 2018 May;34(5):575–84.
34. Siao WZ, Chen YH, Tsai CF, Lee CM, Jong GP. Diabetes Mellitus and Heart Failure. J Pers Med. 2022 Oct 11;12(10):1698.
35. Oh GC, Cho HJ. Blood pressure and heart failure. Clin Hypertens. 2020 Jan 2;26:1.
36. Triposkiadis F, Sarafidis P, Briasoulis A, Magouliotis DE, Athanasiou T, Skoularigis J, et al. Hypertensive Heart Failure. Journal of Clinical Medicine. 2023 Jan;12(15):5090.
37. Krittayaphong R, Karaketklang K, Yindeengam A, Janwanishstaporn S. Heart failure mortality compared between elderly and non-elderly Thai patients. J Geriatr Cardiol. 2018 Dec;15(12):718–24.
38. Metra M, Cotter G, Gheorghiade M, Dei Cas L, Voors AA. The role of the kidney in heart failure. Eur Heart J. 2012 Sep;33(17):2135–42.
39. Franco AD, Ambrosio G, Baroncelli L, Pizzorusso T, Barison A, Olivotto I, et al. Creatine deficiency and heart failure. Heart Fail Rev. 2021 Oct 7;27(5):1605.
40. Peng S, Peng J, Yang L, Ke W. Relationship between serum sodium levels and all-cause mortality in congestive heart failure patients: A retrospective cohort study based on the Mimic-III database. Front Cardiovasc Med. 2023 Jan 13;9:1082845.
41. Chung I, Lip GYH. Platelets and heart failure. Eur Heart J. 2006 Nov;27(22):2623–31.
42. Goff DC, Brass L, Braun LT, Croft JB, Flesch JD, Fowkes FGR, et al. Essential Features of a Surveillance System to Support the Prevention and Management of Heart Disease and Stroke. Circulation. 2007 Jan 2;115(1):127–55.
43. Chicco D, Jurman G. Machine learning can predict survival of patients with heart failure from serum creatinine and ejection fraction alone. BMC Med Inform Decis Mak. 2020 Mar 2;20(1):16.
44. Kaggle. Heart Failure Prediction [Internet]. Kaggle. [cited 2024 Sep 11]. Available from: https://www.kaggle.com/datasets/andrewmvd/heart-failure-clinical-data
45. Joseph P, Roy A, Lonn E, Störk S, Floras J, Mielniczuk L, et al. Global Variations in Heart Failure Etiology, Management, and Outcomes. JAMA. 2023 May 16;329(19):1650–61.
46. Organisation for Economic Co-operation and Development. DAC List of ODA Recipients | Effective for reporting on 2024 and 2025 flows [Internet]. OECD. 2023 [cited 2024 Oct 17]. Available from: https://www.oecd.org/content/dam/oecd/en/topics/policy-sub-issues/oda-eligibility-and-conditions/DAC-List-of-ODA-Recipients-for-reporting-2024-25-flows.pdf
47. Palazzuoli A, Iacoviello M. Diabetes leading to heart failure and heart failure leading to diabetes: epidemiological and clinical evidence. Heart Fail Rev. 2023;28(3):585–96.
48. Pandas Development Team. About pandas [Internet]. Pandas. [cited 2024 Oct 24]. Available from: https://pandas.pydata.org/about/
49. Scikit-learn Developers. sklearn.preprocessing.StandardScaler [Internet]. Scikit-learn. [cited 2024 Oct 24]. Available from: https://scikit-learn/stable/modules/generated/sklearn.preprocessing.StandardScaler.html
50. Cooper AJ, Gupta SR, Moustafa AF, Chao AM. Sex/Gender Differences in Obesity Prevalence, Comorbidities, and Treatment. Curr Obes Rep. 2021 Dec;10(4):458–66.
51. Donzé JD, Beeler PE, Bates DW. Impact of Hyponatremia Correction on the Risk for 30-Day Readmission and Death in Patients with Congestive Heart Failure. Am J Med. 2016 Aug

1;129(8):836–42.
52. Stewart S, Playford D, Scalia GM, Currie P, Celermajer DS, Prior D, et al. Ejection fraction and mortality: a nationwide register-based cohort study of 499 153 women and men. Eur J Heart Fail. 2021 Mar;23(3):406–16.
53. Zhong J, Gao J, Luo JC, Zheng JL, Tu GW, Xue Y. Serum creatinine as a predictor of mortality in patients readmitted to the intensive care unit after cardiac surgery: a retrospective cohort study in China. J Thorac Dis. 2021 Mar;13(3):1728–36.
54. GeeksforGeeks. Python - seaborn.boxenplot() method [Internet]. GeeksforGeeks. 2020 [cited 2024 Oct 24]. Available from: https://www.geeksforgeeks.org/python-seaborn-boxenplot-method/
55. Das D, Das T. The “P”-Value: The Primary Alphabet of Research Revisited. Int J Prev Med. 2023;14:41.
56. Ugoni A, Walker BF. The Chi square test: an introduction. COMSIG Rev. 1995 Nov 1;4(3):61–4.
57. IBM. Clustering binary data with K-Means (should be avoided) [Internet]. IBM. 2020 [cited 2024 Oct 26]. Available from: https://www.ibm.com/support/pages/clustering-binary-data-k-means-should-be-avoided
58. Gholamy A, Kreinovich V, Kosheleva O. Why 70/30 or 80/20 Relation Between Training and Testing Sets: A Pedagogical Explanation. Int J Intell Technol Appl Stat. 2018 Feb 1;11(2):105–11.
59. Abiodun OI, Jantan A, Omolara AE, Dada KV, Mohamed NA, Arshad H. State-of-the-art in artificial neural network applications: A survey. Heliyon. 2018 Nov 1;4(11):e00938.
60. Portillo Juan N, Matutano C, Negro Valdecantos V. Uncertainties in the application of artificial neural networks in ocean engineering. Ocean Eng. 2023 Sep 15;284:115193.
61. PiML Toolbox. 4.7. SHAP (SHapley Additive exPlanations) [Internet]. [cited 2024 Oct 24]. Available from: https://selfexplainml.github.io/PiML-Toolbox/_build/html/guides/explain/shap.html
62. Lundberg S. Be careful when interpreting predictive models in search of causal insights — SHAP latest documentation [Internet]. 2018 [cited 2024 Oct 24]. Available from: https://shap.readthedocs.io/en/latest/example_notebooks/overviews/Be%20careful%20when%20interpreting%20predictive%20models%20in%20search%20of%20causal%20insights.html
63. Yadav A. SHAP Values for Logistic Regression [Internet]. Biased-Algorithms. 2024 [cited 2024 Oct 26]. Available from: https://medium.com/biased-algorithms/shap-values-for-logistic-regression-51ffa5bdcedc
64. Orsini N, Moore A, Wolk A. Interaction Analysis Based on Shapley Values and Extreme Gradient Boosting: A Realistic Simulation and Application to a Large Epidemiological Prospective Study. Front Nutr. 2022 Jul 18;9:871768.
65. Suriya DS, S N. Design of UML Diagrams for WEBMED - Healthcare Service System Services. EAI Endorsed Transactions on e-Learning. 2022;8(1):e5–e5.
66. Rogers Y, Sharp H, Preece J. Interaction design: beyond human-computer interaction. Sixth edition. Hoboken, New Jersey: Wiley; 2023. 689 p.
67. Arnowitz J, Arent M, Berger N. The Effective Prototyping Process: Define Prototype Content and Fidelity. In: Effective Prototyping for Software Makers. Morgan Kaufmann; 2010. p. 87–8.
68. Saffer D. Designing for interaction: creating innovative applications and devices. 2nd ed. Berkeley, CA: New Riders; 2010. 223 p. (Voices that matter).
69. Baxter G, Sommerville I. Socio-technical systems: From design methods to systems engineering. Interacting with Computers. 2011 Jan 1;23(1):4–17.
70. ISO. ISO 9241-210:2019. Ergonomics of human-system interaction — Part 210: Human-centered design for interactive systems [Internet]. Geneva; 2022 [cited 2024 Oct 16]. Available from: https://www.iso.org/standard/77520.html
71. Norman DA. The Design of Everyday Things [Internet]. London: The MIT Press; 1988 [cited 2024 Oct 24]. Available from: https://mitpress.mit.edu/9780262640374/the-design-of-everyday-things/
72. Whitenton K. Minimize Cognitive Load to Maximize Usability [Internet]. Nielsen Norman Group. 2013 [cited 2024 Oct 20]. Available from: https://www.nngroup.com/articles/minimize-cognitive-load/
73. Fitschen T. Dashboard Design Best Practices: Color - The Data School [Internet]. The Data School. [cited 2024 Oct 20]. Available from: https://www.thedataschool.co.uk/tobias-fitschen/dashboard-design-best-practices-color/
74. Almasi S, Bahaadinbeigy K, Ahmadi H, Sohrabei S, Rabiei R. Usability Evaluation of Dashboards: A Systematic Literature Review of Tools. Biomed Res Int. 2023 Feb 22;2023:9990933.
75. Soegaard M. System Usability Scale for Data-Driven UX [Internet]. The Interaction Design Foundation. 2023 [cited 2024 Oct 15]. Available from: https://www.interaction-design.org/literature/article/system-usability-scale
76. Hyzy M, Bond R, Mulvenna M, Bai L, Dix A, Leigh S, et al. System Usability Scale Benchmarking for Digital Health Apps: Meta-analysis. JMIR Mhealth Uhealth. 2022 Aug 18;10(8):e37290.

77. Laubheimer P. Beyond the NPS: Measuring Perceived Usability with the SUS, NASA-TLX, and the Single Ease Question After Tasks and Usability Tests [Internet]. Nielsen Norman Group. 2018 [cited 2024 Oct 15]. Available from: https://www.nngroup.com/articles/measuring-perceived-usability/
78. Brooke J. SUS: A quick and dirty usability scale. Usability Eval Ind. 1995 Nov 30;189.
79. Nielsen J. Thinking Aloud: The #1 Usability Tool [Internet]. Nielsen Norman Group. 2012 [cited 2024 Oct 15]. Available from: https://www.nngroup.com/articles/thinking-aloud-the-1-usability-tool/
80. Hartson R, Pyla PS. Chapter 12 - UX Evaluation Introduction. In: Hartson R, Pyla PS, editors. The UX Book [Internet]. Boston: Morgan Kaufmann; 2012 [cited 2024 Oct 15]. p. 427–65. Available from: https://www.sciencedirect.com/science/article/pii/B9780123852410000129
81. Office for Health Improvement and Disparities. Think aloud study: qualitative studies [Internet]. GOV.UK. 2021 [cited 2024 Oct 15]. Available from: https://www.gov.uk/guidance/think-aloud-study-qualitative-studies
82. User research community. Using moderated usability testing - Service Manual - GOV.UK [Internet]. GOV.UK. 2017 [cited 2024 Oct 15]. Available from: https://www.gov.uk/service-manual/user-research/using-moderated-usability-testing
83. Lewis JR, Sauro J. Item Benchmarks for the System Usability Scale. JUX. 2018 May 25;13(3):158–67.
84. Kiger ME, Varpio L. Thematic analysis of qualitative data: AMEE Guide No. 131. Med Teach. 2020 Aug;42(8):846–54.