TALLINN UNIVERSITY OF TECHNOLOGY
School of Information Technologies

Mihkel Kaalep 206155IAIB

Explainable Artificial Intelligence Application for Clinical Treatment Decision-Making in Rheumatology

Bachelor's Thesis

Supervisor: Gert Kanter
PhD

Co-supervisor: Niels Steen Krogh
MSc

Tallinn 2023

TALLINNA TEHNIKAÜLIKOOL
Infotehnoloogia teaduskond

Mihkel Kaalep 206155IAIB

TEHISINTELLEKTI SELGITAV RAKENDUS KLIINILISTE OTSUSTUSPROTSESSIDE TOETAMISEKS REUMATOLOOGIAS

Bakalaureusetöö

Juhendaja: Gert Kanter
PhD

Kaasjuhendaja: Niels Steen Krogh
MSc

Tallinn 2023

Author’s Declaration of Originality

I hereby certify that I am the sole author of this thesis. All the used materials, references to the literature and the work of others have been referred to. This thesis has not been presented for examination anywhere else.

Author: Mihkel Kaalep

22.05.2023

Abstract

This thesis outlines the design and implementation of an Explainable AI (XAI) app for clinical treatment decision-making in rheumatology. The lack of interpretability and transparency of AI models is a major challenge in implementing AI in healthcare. The proposed study aims to address this challenge by developing an XAI app that provides understandable and transparent explanations for AI-based predictions and recommendations. The study will investigate the benefits and limitations of using XAI in the context of clinical decision-making and patient care and how the app can be designed to maximize these benefits while minimizing the limitations. This study will use mixed-methods, including literature review, stakeholder analysis, proof-of-concept application design and development, user testing, and evaluation. The study will take into account the perspectives and needs of healthcare practitioners to ensure that the app design and functionality align with their requirements. The findings of this study have the potential to improve clinical decision-making and patient outcomes, promote the adoption of AI technologies in precision medicine, and inform future research in this area.

The thesis is written in English and is 40 pages long, including 5 chapters, 20 figures and 3 tables.

Annotatsioon

Tehisintellekti selgitav rakendus kliiniliste otsustusprotsesside toetamiseks reumatoloogias

Käesolevas lõputöös kirjeldatakse tehisintellekti selgitava rakenduse kavandamist ja arendamist kliiniliste otsustusprotsesside toetamiseks reumatoloogias. Tehisintellekti mudelite tõlgendamise keerukus ja mudelite läbipaistmatus takistavad tehisintellekti kasutuselevõttu tervishoius. Käesoleva lõputöö eesmärk on neid puudusi vähendada, töötades välja tehisintellekti selgitav rakendus, mis pakuks tehisintellektil põhinevate prognooside ja soovituste kohta arusaadavaid ja läbipaistvaid selgitusi. Lõputöös uuritakse selgitatava tehisintellekti kasutamise eeliseid ja puuduseid kliinilistes otsustusprotsessides ning seda, kuidas saaks rakendust kujundada selliselt, et ära kasutada eeliseid ja vähendada puuduseid. Selles uuringus kasutatakse kombineeritud uurimismeetodeid, sealhulgas kirjanduse ülevaadet, osapoolte analüüsi ning rakenduste kavandamist ja arendamist. Selleks, et rakendus valmiks vastavalt tervishoiutöötajate vajadustele, viiakse läbi kasutajatestimist ning sellele järgnevalt intervjuid testijatega. Selle lõputöö tulemustel on potentsiaal arendada kliiniliste otsusprotsesside tegemist ja patsientide diagnostikat, edendada tehisintellekti kasutuselevõttu tervishoius ning olla aluseks selle valdkonna tulevastele uuringutele.

Lõputöö on kirjutatud eesti keeles ning sisaldab teksti 40 leheküljel, 5 peatükki, 20 joonist, 3 tabelit.

List of Abbreviations and Terms

Adalimumab	Type of medication used to reduce the signs and symptoms of rheumatoid arthritis
AI	Artificial Intelligence
API	Application Programming Interface
DANBIO	System for collecting and analyzing data on rheumatology patients in Denmark
DAS-28-CRP remission	State of low or absent disease activity in rheumatoid arthritis patients, measured by the Disease Activity Score with 28 joint counts (DAS28) index.
EMR	Electronic Medical Record, used between clinicians for keeping records and exchanging information
FastAPI	Python Web framework for handling fast and concurrent requests through API. It is as much about humans communicating with systems, as systems being able to communicate with each other
FHIR	Short for Fast Healthcare Interoperability Resources. Medical standard maintained by the HL7 consortium
JSON	JavaScript Object Notation
JWT	JSON Web Token
LIME	Local Interpretable Model-agnostic Explanations.
OAuth 2.0	Protocol for authorization.
Precision Medicine	Form of medicine that uses information about a person's own genes or proteins to prevent, diagnose, or treat disease
R	Language for statistical analysis
React	Open-source JavaScript library used for building user interfaces for web and mobile applications.
Rheumatoid Arthritis	Rheumatoid arthritis is a chronic autoimmune disease that causes inflammation and damage to the joints and other parts of the body.
SHAP	SHapley Additive exPlanations.

SMART on FHIR	Open-source, standards-based API that leverages the OAuth 2.0 standard to provide secure, universal access to EHRs. The SMART platform builds on the existing Fast Health Interoperability Resources (FHIR), hence the name "SMART on FHIR"
XAI	Explainable artificial intelligence - a set of methods for understanding and explaining the behaviour of machine learning models.
XGBoost	Machine learning algorithm that uses decision trees to model complex relationships between variables and make predictions.

Table of Contents

List of Figures

List of Tables

1. Introduction

Artificial Intelligence (AI) has emerged as a promising tool for improving clinical decision-making and patient outcomes in the field of healthcare. However, the lack of interpretability and transparency of AI models is a major challenge in implementing AI in healthcare. The development of Explainable AI (XAI) is a potential solution to address this challenge, as it can provide understandable and transparent explanations for AI-based predictions and recommendations. In this thesis, we aim to design and implement an XAI app for clinical treatment decision-making in the field of rheumatology.

1.1 Problem Statement

The use of AI in healthcare has seen significant developments in recent years, with promising results in improving diagnostic accuracy, treatment selection, and disease prediction. However, the lack of interpretability and transparency of AI models has limited the adoption and implementation of AI in clinical settings. This is particularly important in Rheumatology, where complex decision-making is required based on multiple factors such as disease activity, patient preferences, and treatment response. [1]

One solution to address the interpretability challenge is Explainable AI (XAI), which aims to provide explanations for the predictions and recommendations made by AI models. XAI can help clinicians understand the underlying logic of the model and how it arrives at its predictions. This is crucial in the context of clinical treatment decision-making, where transparency and accountability are key factors.

In this thesis, the author aims to design and implement an XAI app for clinical decision-making in Rheumatology. The app will be designed to integrate with the Electronic Medical Record (EMR) system and will utilize the SMART on FHIR (Fast Healthcare Interoperability Resources) technology. The app will enable healthcare practitioners to access understandable and transparent explanations for AI-based predictions and recommendations, aiding in clinical decision-making and ultimately improving patient outcomes.

1.2 Research Questions and Hypotheses

Research Question 1: How can an XAI app be designed and integrated within the EMR system to provide transparent and understandable explanations for AI-based predictions and recommendations in rheumatology?

Hypothesis 1: The implementation of an XAI app within the EMR system will enable healthcare practitioners to access transparent and understandable explanations for AI-based predictions and recommendations, resulting in improved clinical decision-making and patient outcomes in Rheumatology.

Research Question 2: What are the benefits and limitations of using an XAI app in the context of clinical decision-making and patient care in Rheumatology?

Hypothesis 2: The use of an XAI app will lead to improved clinical decision-making and patient outcomes by enabling healthcare practitioners to understand the underlying logic of AI models. However, there may be limitations to the use of XAI, such as increased cognitive load for healthcare practitioners and potential challenges in explaining complex AI models to patients.

1.3 Approach and Methodology

This study will utilize a mixed-methods approach, including a literature review, stakeholder analysis, proof-of-concept application design and development, user testing and evaluation. The literature review will identify relevant technologies, standards, and best practices for designing and implementing XAI apps in healthcare. The stakeholder analysis will involve engaging with different healthcare providers to understand their perspectives and needs regarding the use of XAI in rheumatology.

As part of this thesis project, a proof-of-concept XAI app will be designed and developed, which will be integrated into a test EMR system. Integration into the production EMR system will not take place within the timeframe of this thesis. Instead, the focus will be on the planning and architecture for how the XAI app can be integrated into the production system at a future time. The app will be tested and evaluated through user testing to gather feedback and identify usability issues. To explore the benefits and limitations of using XAI in the context of clinical treatment decision-making in rheumatology, user testing will be conducted with three healthcare professionals, selected based on their expertise and experience in rheumatology. The testing process will consist of initially testing the application, followed by conducting comprehensive interviews to collect feedback and insights from users regarding the potential applications and challenges of Explainable Artificial Intelligence (XAI) in this field. The data collected will be carefully analyzed using thematic analysis to extract meaningful conclusions and insights.

1.4 Limitations and Future Directions

This study has several limitations, including the potential for selection bias in the interviews and limited generalizability of the findings. To address these limitations and advance the use of AI in healthcare, future research can involve larger and more diverse samples of participants and more rigorous study designs such as randomized controlled trials.

The successful implementation and widespread adoption of AI algorithms in real-world settings would greatly benefit from the robust design of prospective trials, ideally involving multiple centers and utilizing randomized methodologies to rigorously evaluate the performance and effectiveness of these models.[1] By comparing a control group of healthcare practitioners and patients who follow the standard national guidelines with another control group that uses the AI tool being evaluated, we can determine if the AI tool outperforms the national guidelines. This approach can help address the limitations of smaller, more selective studies and ensure that the results are generalizable to a broader population.

Furthermore, ongoing monitoring and evaluation of AI systems in clinical practice can help identify and address any unforeseen issues or unintended consequences that may arise, such as biases in the data or unexpected patient outcomes.[2] By addressing these limitations and continuing to advance our understanding of AI in healthcare, we can work towards realizing its full potential to improve patient outcomes and enhance the delivery of healthcare services.

Overall, while there are limitations to the current research on AI in healthcare, future studies utilizing larger, more diverse samples and rigorous study designs such as randomized controlled trials, as well as ongoing monitoring and evaluation, can help build trust in the use of AI in clinical practice and further advance the field.

2. Background

2.1 Explainable AI

Explainable AI (XAI) refers to the development of machine learning algorithms that can be transparently understood and interpreted by humans. The goal of XAI is to create models that not only provide accurate predictions but also offer an explanation for how they arrived at those predictions, which is crucial for building trust and confidence in AI systems.

In the context of clinical decision-making in rheumatology, XAI can be used to develop predictive models that help physicians diagnose and treat patients with rheumatologic diseases. For example, XAI could be used to analyze large datasets of patient information and medical records, identifying patterns and correlations that might not be immediately apparent to human clinicians.[1]

XAI could also be used to develop decision support systems that help clinicians make better treatment decisions by providing real-time recommendations based on patient-specific data. These systems could be particularly helpful in cases where there is a high degree of uncertainty or variability in the patient's presentation, such as when diagnosing complex autoimmune diseases like lupus or rheumatoid arthritis.[1]

By providing interpretable and transparent models, XAI can help clinicians understand the rationale behind the predictions and recommendations made by AI systems, enabling them to make informed decisions that are grounded in evidence and clinical expertise. This can ultimately lead to better patient outcomes and improved quality of care.

2.1.1 Choosing the explainability technique

There are four dimensions to consider in the context of machine learning explainability: explanation scope, inputs, access, and stage. [3]

Explanation scope refers to what we are trying to explain and whether it is a global or local explanation. In this case, we want to explain the model both globally across many data points and locally to explain the drivers of a prediction for a single datapoint, which is the patient.

The type of output must also be evaluated, which could be either a classification decision or a probability score. In this case, we are interested in explaining a probability score, specifically the prognosis tool that gives a probability score for achieving DAS-28-CRP remission after six months of Adalimumab treatment. We also know that we are explaining the model's constituent features, which are the baseline characteristics.

The next step is to determine the level of access our explainability technique will have to the model itself. For our purposes, we would like to preserve the option of being able to compare different models on the same dataset and therefore require a limited access technique, such as SHAP[4], LIME[5], or QII[6]. These are model-agnostic techniques that examine the effect that inputs have on model outputs without peering into model internals. Since we are using XGBoost for machine learning, our model is algorithmically interpretable, meaning the rules are derived during the training process itself. However, XGBoost is still considered a "black-box" model, which means explainability techniques can only be applied after the model has finished training, as is the case with our usage, since the trained model is already located in the backend.

Finally, we will describe four popular explainability techniques and choose the most appropriate one.

Coefficients of logistic regression

This technique is purposeful for globally explaining of the model's score. It accepts input features. It's model specific (learned weights), and it suits to self-interpretable models.

Local Interpretable Model-Agnostic Explanations (LIME)

It works for local explanation for model score. It accepts input features. It's model-agnostic. And it's algorithmically interpretable.

SHapley Additive exPlanations (SHAP)

It works for both local and global explanation for any model output. In terms of the type of input it is flexible, but typically input features are used. It is model-agnostic. And it's algorithmically interpretable.

Integrated Gradients

It is designed for local explanations, but can also be aggregated for global explanations. Input features are the recommended input type. It is specifically applicable to neural networks and provides algorithmic interpretability.

Conclusion

After considering the four different explainability techniques, we conclude that SHAP is the best fit for our purposes of explaining the DAS-28-CRP remission probability score from our XGBoost model. SHAP is model-agnostic and algorithmically interpretable, which is important for our black-box model. It also works for both local and global explanation, which is what we require. Although LIME is also model-agnostic and algorithmically interpretable, it is primarily designed for local explanation, whereas we want both local and global explanations. Coefficients of logistic regression, on the other hand, are only suitable for globally explaining self-interpretable models, which is not applicable for our XGBoost model. The Integrated Gradients method requires neural networks, which we are not using. Therefore, we can confidently choose SHAP as our explainability technique to gain insights into the drivers of the DAS-28-CRP remission[1] probability score for our machine learning model.

2.2 The NORA/ScandRA Project

The application developed for this project is a key component of the larger multidisciplinary research initiative with the acronym "ScandRA," which focuses on personalized medicine for rheumatoid arthritis patients in Scandinavia. The project, which commenced in 2021 and spans over three years, aims to integrate genomics, biomarkers, clinical, and patient data from across Scandinavian countries to develop personalized medicine approaches for rheumatoid arthritis. The initial stages of the project involved the collection of clinical data, including blood samples and biomarker analyses. Subsequently, a data enrichment and harmonization phase was carried out, which included bioinformatics and statistical analyses. Stratifiers and prediction algorithms were then developed using the generated knowledge, followed by internal and external validation. The final stage of the project involves the development of tools and applications, including the application being developed as part of this thesis project. These tools and applications will undergo clinical integration to become part of routine clinical practice. [8]

On January 16, 2023, the author of this thesis participated in the NORA/ScandRA consortium conference in Arlanda, Sweden, where they presented the progress of their project titled "Rheumatoid Arthritis Prediction System Embeddable in EMRs." The author received feedback on the application and addressed clarifying questions about the machine learning model and the dataset. It was during this conference that an idea for gamification

[1] DAS-28-CPR remission is a state of low or absent disease activity in patients with rheumatoid arthritis, measured by the Disease Activity Score with 28 joint counts (DAS28) index. It is achieved when the DAS28 score is below a certain threshold, typically less than 2.6, and is an important treatment goal associated with better long-term outcomes.[7]

as a potential approach for improving understandability emerged. This concept was later integrated into the application as a result of the conference discussions.

In March 2023, an abstract titled "Empowering Medical Prognosis in Rheumatoid Arthritis with Explainable AI APP" was submitted to the 39th Scandinavian Congress of Rheumatology [9]. The abstract described the machine learning model and the dataset, together with the internal validation results for accuracy of the model. A follow-up abstract titled "PrognosisAI: A Web-Based Tool for Predicting Six-Month DAS28-Remission in Patients with Rheumatoid Arthritis Initiating Adalimumab as First-Line Biologic Therapy" with a more in-depth discussion of the application and its ability to provide single patient predictions and the explainability methods was submitted at the end of April 2023.

3. Used Technologies

3.1 The Web Application

When selecting technologies for the front-end application, the focus was on utilizing modern, widely-used technologies that are familiar to most developers and are well-documented with a large user base. Technologies that would allow future developers to easily join the project and start writing code without extensive on-boarding, were prioritized. Additionally, significant emphasis was given to UI and UX considerations, ensuring that the front-end technologies utilized would facilitate the creation of an visually appealing and user-friendly application.

3.1.1 React

The front-end of the prognosis app was built using React[10], a widely-used and popular JavaScript library for building user interfaces. React allows for the creation of reusable UI components that efficiently update and render in response to changes in data or user interactions. Its component-based architecture, efficient rendering, and extensive ecosystem make it the go-to choice for building this web application.

After initially developing the application in Angular, a decision was made to transition to React based on several factors. Firstly, the component-based architecture and flexibility of React allowed for more efficient development of reusable UI components. Secondly, React's performance optimizations, such as the virtual DOM, improved the overall performance of the application, resulting in smoother user interactions. Thirdly, React is currently more popular than Angular [11], offering extensive resources, tutorials, and support from React's large and active community. Lastly, the familiarity and preferences of the employees at ZiteLab towards React facilitated a smooth transition, leveraging their existing skills and expertise. Overall, the transition to React proved to be a pragmatic choice, resulting in a more efficient, performant, and maintainable application.

3.1.2 Vite

Vite is a frontend development tool that uses native ES modules and modern browser APIs for fast build times and instant updates in the browser. It eliminates the need for a bundler during development, resulting in improved workflow and productivity. Vite also offers features such as lazy loading of modules and tree-shaking for optimized code sizes and performance. However, it's a relatively new tool with a smaller community and may not be supported by all browsers. Nevertheless, for speed, customization, and an accelerated development experience, Vite is well-suited for this project compared to more widely used alternatives like Create React App.[12]

3.1.3 Typescript

Typescript is a programming language that was chosen for this project over JavaScript due to its strong typing and static checking capabilities. It enhances JavaScript by providing optional static typing, making it easier to catch type-related errors during development and improving code quality. TypeScript also offers modern language features such as classes, interfaces, and modules, which aid in writing maintainable and scalable code. Additionally, TypeScript supports the latest ECMAScript syntax, making it compatible with modern JavaScript frameworks and libraries. Its static typing and type inference features help developers write more robust code, reduce the likelihood of bugs, and improve overall development efficiency.[13]

3.1.4 SASS

SASS, an acronym for Syntactically Awesome Style Sheets, is a preprocessor scripting language that extends CSS (Cascading Style Sheets). It introduces features such as variables, nesting, inheritance, and modularization, making it a versatile tool for writing efficient and maintainable CSS code. SASS enables reusable code, creating dynamic styles, and streamlining development by simplifying repetitive tasks. Its syntax is based on CSS, making it easy for developers to transition. SASS was chosen over alternatives like SCSS and LESS based on personal preferences due to its concise and expressive syntax with indentation-based structure. SASS provides error checking, supports conditional statements and loops, and offers functions for mathematical operations, color manipulation, and more. With its extensive features and flexibility, SASS is a popular choice among web developers for managing stylesheets, improving code organization, and enhancing productivity in web development projects.[14]

3.1.5 Mantine

Mantine is a React component library that offers a set of UI components for building web applications.[15] Mantine is designed to be simple, customizable, and accessible, making it suitable for both beginner and advanced developers. With its collection of components, including forms, modals, notifications, and more, Mantine provides a foundation for building modern and responsive user interfaces. Mantine also includes tools and utilities, such as a built-in dark mode, theming capabilities, and comprehensive documentation with examples and guides for ease of use. Mantine's performance, accessibility, and developer experience make it a reliable choice for building web applications quickly and efficiently, making it a suitable fit for this project.

Emotion

Emotion, a CSS-in-JS library, is the default styling solution in Mantine. It enables dynamic, scoped, and themable styles for Mantine components, ensuring customization and responsiveness. Emotion's scoped styles prevent global style collisions, and its performance optimizations enhance web application performance. Its versatility and community support make it a trusted choice for styling Mantine components. [16]

3.1.6 HTTP Client

Axios

Axios is a versatile JavaScript library used for making HTTP requests from web browsers. With its user-friendly interface, Axios allows for seamless asynchronous HTTP requests, boasting advanced functionalities such as request and response manipulation, cancellation, and error handling. [17] Its concise syntax, flexibility, scalability, and reliability make it the preferred choice for this project as the HTTP client.

3.1.7 State management

Zustand

Zustand is a lightweight state management solution designed for React projects, offering efficiency and simplicity in how data is managed.[18] It allows for the creation of reusable "stores" to store and update data, such as the JWT tokens in this project. Zustand facilitates easy organization and management of state without unnecessary technical complexity. Additionally, it offers features like subscriptions and middleware to handle side effects and track state changes. Zustand's lightweight and efficient approach makes it an ideal choice for this project.

3.1.8 Routing

Wouter

Wouter is a lightweight, declarative, and flexible routing library for React applications, designed to handle navigation and routing in single-page applications (SPAs) efficiently. [19] It was chosen over the default React router due to its lightweight size and ease of use. Since this application does not require any exotic routing solutions, Wouter makes sense here as it reduces unnecessary overhead.

3.2 The Backend Application

The following subsections provide details about the underlying technologies used for the backend. The technologies were chosen specifically for deploying machine learning models and handling API queries.

3.2.1 FastAPI

FastAPI is a modern, high-performance web framework for building APIs with Python. It provides a user-friendly and Pythonic syntax, leverages the latest features of Python, such as type hints and asynchronous programming, and includes automatic API documentation, data validation, and serialization. Built on top of Starlette, FastAPI is known for its high performance and scalability, and offers flexibility in application architecture. FastAPI also boosts productivity with code generation features and reduces the risk of bugs with built-in data validation. [20] Additionally, it has a large and active community, making it a compelling choice for backend development with Python, offering extensive support and resources. FastAPI is a powerful and efficient framework that provides a seamless development experience and is well-suited for a wide range of backend applications.

3.2.2 Python

Python is a popular, high-level, interpreted programming language known for its simplicity, readability, and versatility. It was chosen as the language for the backend since FastAPI requires it and is well-suited for building the backend of machine learning applications. Python has extensive libraries for machine learning, making it easy to prototype, experiment, and deploy models. It has a large community and support, interoperability with other languages, flexibility in programming paradigms, and scalability. Python provides a robust and efficient environment for implementing, integrating, and deploying machine learning models in the backend of an application.

3.2.3 R

R is a versatile programming language widely utilized for data analysis, statistical modeling, and machine learning. In the FastAPI backend, R has been employed for a range of statistical analysis tasks, including hypothesis testing, regression analysis, and time series analysis, on data received from API endpoints and the datasets. Moreover, R's XGBoost packages have been utilized for the development and deployment of machine learning models within the FastAPI backend, facilitating tasks such as classification, regression, clustering, and recommendation. These models have provided predictive and analytical capabilities to the backend API. By harnessing the strengths of R's statistical computing and machine learning capabilities in conjunction with the flexibility and ease of use of FastAPI, a data-driven backend with advanced statistical and machine learning features was developed.

Tidymodels

The app utilizes machine-learning models for making predictions, which are based on the Tidymodels framework. Tidymodels is an open-source R package that provides a unified and streamlined framework for machine learning and statistical modeling. With a focus on tidy data principles and a consistent API design, Tidymodels offers a wide range of modeling techniques, tools for feature engineering, model tuning, and ensemble modeling, and integrates seamlessly with other popular R packages. [21]

SHAPviz

In this project, SHAPviz is used for rendering SHAP plots for patient-level predictions. SHAPviz is an R package that provides visualizations for interpreting machine learning models using Shapley values. Shapley values, a concept from cooperative game theory, can be used to attribute a value to each feature in a prediction, indicating its contribution to the prediction. SHAPviz enables generating various visualizations, such as summary plots, dependency plots, and force plots, to gain insights into the important features driving the model's predictions, identify interactions among features, and explain the model's behavior in an interpretable way. SHAPviz is a powerful tool for gaining a deeper understanding of their models' predictions and interpreting their outputs in a transparent and explainable manner.[22]

4. Application Design

4.1 Backend Design

The artificial intelligence (AI) that powers the web application is provided by the backend, which was developed by ZiteLab's employees using FastAPI. Because the backend contains sensitive data, it will remain located on GitLab's ZiteLab server. Refer to Table 3 in the appendix for the backend project link. The backend was specifically created to deploy machine learning models and handle API queries. A schematic of the backend structure can be found in Figure 1, and detailed information is provided in Table 1. When API calls are initiated from the web application, FastAPI responds accordingly. Subsequently, the backend employs an internal R package to process the patient data received through microservices. This process involves loading a pre-trained machine learning model and generating predictions. The responses are returned in JSON or JPEG format, depending on whether they are textual or graphical in nature, respectively.

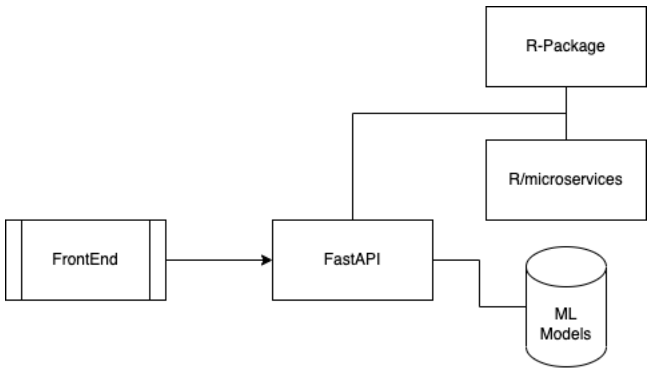

Figure 1. Schematic of the backend.

Folder	**Purpose**
models	Machine Learning Model Snapshots
R-Package	R code for function and class definition.
R/microservices	R function calls executed from API calls
data	Data files used for training/testing/explanations.
apis	FastAPI endpoint functionality.
core	FastAPI configuration.

Table 1. Prognosis-AI backend directory structure.

4.1.1 The Dataset

This subsection provides a description of the dataset used for modelling the predictor. The dataset consists of Rheumatoid Arthritis (RA) patients who underwent first-line Adalimumab treatment. The patients were selected based on specific criteria as described below.

Patient Selection Criteria

The patients included in the dataset were filtered for Adalimumab-only treatment, patients after 2015, and DAS28 score available at month 6-12. This resulted in a total of 692 patients, who met the inclusion criteria for the study.

Training and Test Split

The dataset was randomly split into training and test sets with a 75:25 ratio. This resulted in 518 patients in the training set and 174 patients in the test set.

Patient Demographics and Clinical Characteristics

All patients included in the dataset were from medical centers in Denmark. The patients' ages, sexes, and smoking status were recorded. Additionally, physiological measures such as swollen joints, tender joints, and X-Ray Erosion were also recorded. Responses to patient surveys (VAS and HAQ) and blood tests (CRP, Rheumatoid Factor or AntiCCP if available) were also documented. The DAS28 score at baseline (month 0) was also included.

The dataset includes a comprehensive list of features, which can be seen in section 5.2.1. Some patients may have had a history of taking Prednisolone, an inflammation-suppressing steroid. It is important to note that Prednisolone does not cure RA but can help control the symptoms.

Treatment History

All patients included in the dataset underwent first-line Adalimumab treatment. This means that prior to Adalimumab treatment, they had no drug history related to treating RA.

In summary, the dataset used for modelling our predictor includes 692 RA patients who underwent first-line Adalimumab treatment. The patients were selected based on specific criteria, including Adalimumab-only treatment, patients after 2015, and DAS28 score available at month 6-12. The dataset includes patient demographics, clinical characteristics, treatment history, and a comprehensive list of features.

4.1.2 XGBoost

The prediction tool uses XGBoost (eXtreme Gradient Boosting), a machine learning algorithm commonly employed for tasks such as prediction and classification. XGBoost is specifically designed to handle regression problems and is particularly good at modeling complex and non-linear relationships between variables, which is often the case in RA disease activity. It also has a regularization technique that helps to prevent overfitting and improve model generalization performance, and can handle missing data and automatically select the most important features for prediction, which can help to improve model accuracy and efficiency.[23]

While there are several other machine learning algorithms that could be used for predicting DAS-28-CRP remission probability, such as Random Forest, Support Vector Machines (SVMs), and Neural Networks, XGBoost has several advantages that make it a good choice for this particular task. XGBoost's ability to handle complex and non-linear relationships between variables is particularly important in the context of RA disease activity, where there are often multiple factors that contribute to disease activity. Additionally, XGBoost's regularization technique and ability to handle missing data can help to improve model performance and reduce the risk of overfitting.

Overall, XGBoost is a powerful and versatile algorithm that can handle the complex and non-linear relationships between variables in predicting DAS28 remission probability. It is a good choice for this particular task due to its ability to handle missing data, automatically select important features, and model complex relationships between variables, all of which are important factors in predicting RA disease activity.

4.2 Frontend Design

A frontend is a software program or website that provides a user interface for a particular application or system. In the context of this project, the frontend refers to a website that enables clinicians to make predictions about DAS-28-CRP remission in rheumatoid arthritis patients using patient data from DANBIO. The website serves as the platform for clinicians to input patient data and receive predicted outcomes based on the machine learning algorithms integrated into the system. Refer to Table 3 in the appendix for the frontend project link.

The website's frontend design follows the principles of user-centered design [24]. The user is placed at the center of the experience, and the interface is designed to meet their needs and preferences. Consideration is given to the specific requirements and expectations of the target users, such as rheumatologists or healthcare practitioners, and ensuring that the interface is intuitive, easy to use, and provides relevant information.

A consistent visual language, including typography, color schemes, and icons, is employed to create a cohesive and visually appealing interface, ensuring clarity and consistency in the frontend design. Simplicity and minimalism are prioritized, avoiding clutter and unnecessary elements that can distract users from the core functionality of the website. This streamlined approach allows users to focus on the essential information and actions, thereby improving the overall user experience.

The frontend design is optimized for responsiveness, ensuring that users can access the website seamlessly from desktops, tablets, and mobile devices, regardless of the device used, to ensure a consistent user experience.

An iterative design approach is followed, with usability testing conducted with practitioners, feedback gathered, and continuous improvements made to the frontend design. This allows for addressing any usability issues, refining the interface, and ensuring that it meets the needs and expectations of the intended users.

In conclusion, the frontend design of the website for predicting DAS-28-CRP remission in rheumatoid arthritis patients adheres to user-centered design principles, including clarity and consistency, responsiveness, minimalism and simplicity, and iterative design through usability testing. The interface is designed to provide a seamless and intuitive experience for healthcare practitioners, allowing them to effectively predict and manage remission in their patients.

4.2.1 Code Design

The frontend code is designed following the best practices and principles promoted by the React community and documentation. Component names are written in PascalCase, helper files follow camelCase naming conventions, and folders also use camelCase. The code is structured into multiple files/components to prevent code repetition and improve maintainability. Inline CSS is avoided, and instead, external CSS files are utilized to keep the code organized and easy to maintain. The use of linters helps in maintaining code consistency and readability. Service calls are separated into their own files to keep the code modular and manageable. Moreover, TypeScript is chosen to enhance code quality and readability, as it is a strongly typed language, unlike JavaScript, which adds an additional layer of type safety and helps in catching potential bugs early in the development process. Overall, the frontend code is designed with a focus on code quality, maintainability, and readability to ensure a smooth and efficient development process. Frontend folder structure is displayed in Table 2.

Folder	**Purpose**
dist	Build files
node_modules	Dependencies
src/assets	Images and icons
src/api	Service calls
src/views	Views that are used once
src/routes	Routes
<root>	TypeScript configuration, environment configuration, Pretier configuration, Vite configuration, package.json, index.html, README

Table 2. Prognosis-AI frontend folder structure.

4.2.2 Authentication and Authorization

In the Prognosis-AI app, authentication and authorization are two distinct but interconnected processes that work together to ensure secure user access. Currently, authentication is performed within the Prognosis-AI system using HTTP authentication, which involves verifying a user's identity through a username and password. However, the goal is to shift the authentication process to DANBIO, an external system.

Once a user is successfully authenticated in DANBIO, they will be granted access to the Prognosis-AI system via DANBIO. This approach aims to prevent duplicate logins and streamline the authentication process for users. By leveraging DANBIO for authentication, Prognosis-AI can enhance security measures and ensure that only authorized users are granted access to the system. This separation of authentication and authorization processes adds an additional layer of security to the Prognosis-AI app, safeguarding user data and system integrity.

To establish a secure connection between DANBIO and the Prognosis-AI app, the Prognosis-AI system can generate a JSON Web Token (JWT) when an authorized user arrives from DANBIO. This JWT serves as a unique token that acts as proof of access and is stored in cache within the Prognosis-AI app. The JWT contains information that validates the authorization of the user, indicating that the access was initiated from DANBIO.

As the user navigates between pages within the Prognosis-AI webpage, the system consistently checks for the presence of the correct JWT to ensure that the access is legitimate and authorized. This token-based approach helps verify that the user's access to the Prognosis-AI system is authenticated through DANBIO and that the user has the proper authorization to use the app.

By separating authentication and authorization, the Prognosis-AI app ensures that only authenticated users from DANBIO, with appropriate authorization, can access the Prognosis-AI system. This robust security measure, using JWTs for token-based authorization, helps safeguard patient data and protects against unauthorized access, ensuring that only authorized users can utilize the app and its functionalities. Proper security measures, such as signing and encrypting the JWTs, setting appropriate expiration times, and securely storing the secret key used for signing, should be implemented to ensure the integrity and confidentiality of the authorization process.

4.2.3 Data Visualization

In the context of predicting DAS-28-CRP remission probability, explanations for the model's output at both global and local levels are provided. Global explanations help understand the overall drivers of the model's predictions across all data points. One example of a global explanation is the SHAP beeswarm summary plot, which can be found in the Frequently Asked Questions (FAQ) section of the application. Refer to Figure 2 for an illustration of this plot.

This beeswarm summary plot provides a visual summary of the SHAP values, which represent the contribution of each feature to the model's prediction for each data point. The x-axis represents the SHAP values, while the y-axis represents the baseline characteristics. The plot displays the distribution of SHAP values for each feature, allowing users to gain insights into the impact of different features on the model's output across the entire dataset.

The variables are sorted by their mean absolute SHAP value in descending order, with the most important variables positioned at the top. Points in the plot are colored in orange for high feature values and purple for low feature values. The further away a point is from the vertical line at x=0, the larger the impact on the model's prediction. Negative SHAP values indicate a contribution towards a decreased chance of achieving DAS-28-CRP remission, while positive SHAP values push the model's prediction towards a higher probability of achieving remission.

For example, high values of fatigue scores (indicated by orange color) suggest a change

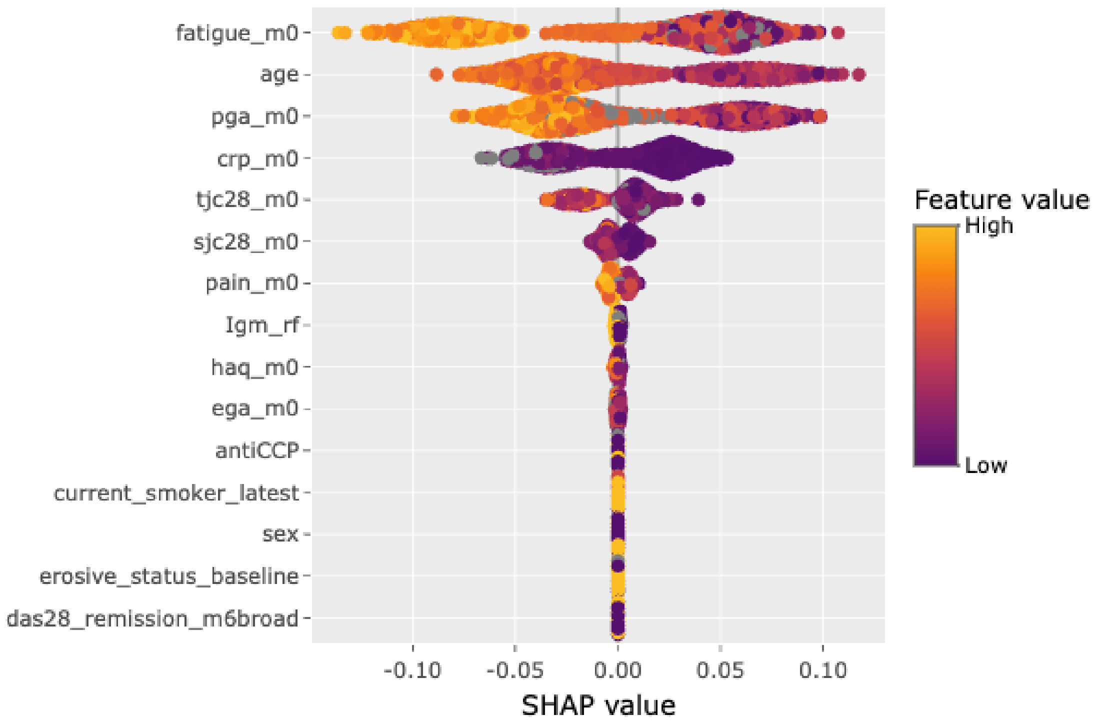

Figure 2. *Dataset summary.*

towards not achieving DAS-28-CRP remission, while low values of CRP (mg/L) (indicated by purple color) suggest a change towards achieving remission. Variables with little impact are centered around the y-axis, such as Anti-CCP and sex. Missing data points are marked in grey for reference.

An example of a local explanation is the SHAP waterfall plot, which was utilized in the prognosis tool to explain the output for a single datapoint (patient). Refer to Figure 3.

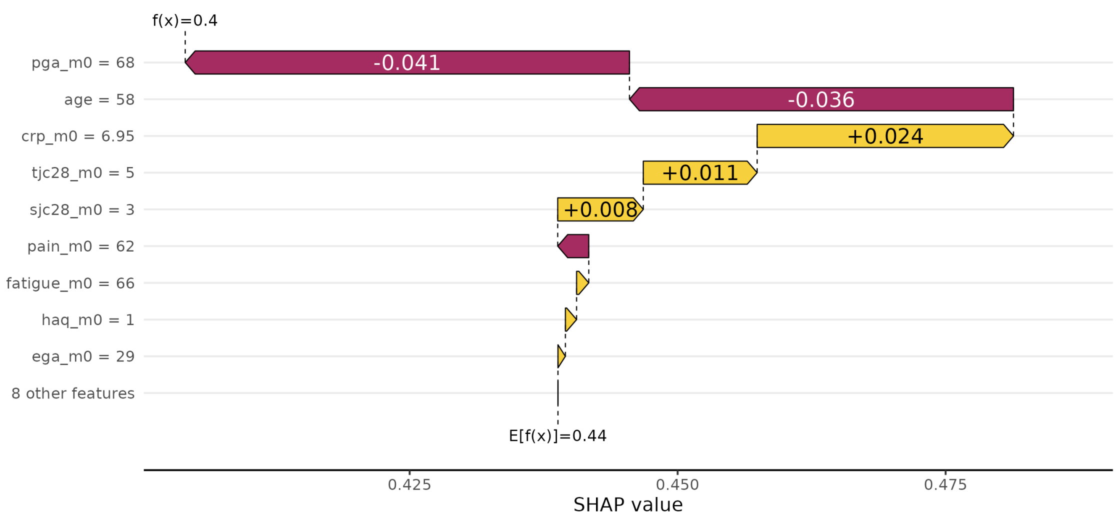

Figure 3. *Single patient prediction.*

This plot, based on SHAP, illustrates how different baseline characteristics impact the probability of achieving DAS-28-CRP remission after 6 months. Each baseline characteristic is represented by a horizontal bar, where the length and color indicate the direction and magnitude of its effect on the predicted probability. Positive values indicate that increasing the clinical variable's value leads to a higher probability of remission, while negative values indicate that increasing the clinical variable's value leads to a lower probability of remission.

4.3 The EMR-System

The EMR-system, DANBIO, is a Danish digital database for rheumatoid arthritis and is developed and managed by ZiteLab ApS. It includes a website called danbio-online.dk that can be accessed nationwide, as well as touchscreen kiosks in waiting rooms for patients. There is also a patient-facing app called DANBIO from home that can be used on smartphones, tablets, or computers with NemId/MitID login. DANBIO also has a feature called Reuma-eCRF for collecting research data and can automatically capture laboratory values. The IT solution is created using Python, R, and the backend uses technologies like Plone and FastAPI. It uses FHIR as the standard for questionnaires and ordering systems. DANBIO is used in various disease areas in Denmark, and it has been regularly updated with new features, such as access from home and MitID login. [25]

4.3.1 Integration With the EMR-System

The integration was achieved by adding a link to Prognosis-AI within the test DANBIO system. The link can be accessed as "PROGNOSIS-AI" in the left navigation column, as shown in Figure 4. The integration into the production system will be carried out at a later stage.

4.3.2 Security

To enhance the security of data exchanged between DANBIO and the Prognosis-AI app, SMART on FHIR can be incorporated as a standard for interoperability. SMART on FHIR is a widely used healthcare IT standard that combines the power of FHIR - a standard for exchanging healthcare data electronically - with the security and authentication capabilities of OAuth2, a widely used authorization framework. This allows for secure and standardized access to health data across different systems, ensuring that only authorized users can access and exchange sensitive patient information.

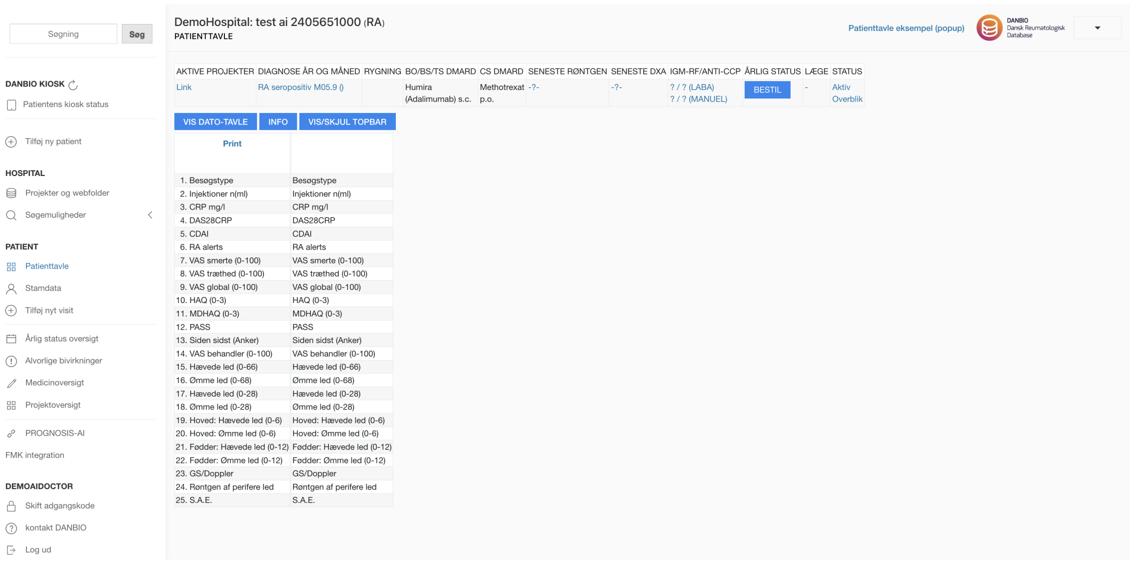

Figure 4. *Access to Prognosis-AI via DANBIO.*

When combined with secure protocols such as HTTPS, SMART on FHIR provides an additional layer of security to protect the confidentiality and integrity of the data transferred between DANBIO and the Prognosis-AI app. For example, when users log in to DANBIO and obtain a JWT token for authorization, SMART on FHIR can be used to securely manage the authentication process, verifying the identity of the user and ensuring that they have the appropriate permissions to access the data. Additionally, when patients' baseline characteristics are transmitted as inputs for the prognosis tool, SMART on FHIR can ensure that the data is exchanged in a secure and standardized manner, following the established healthcare interoperability standards.

By incorporating SMART on FHIR along with HTTPS, the data exchanged between DANBIO and the Prognosis-AI app is protected from unauthorized access, interception, tampering, and eavesdropping during transit. This helps to ensure the confidentiality and integrity of the data, safeguarding sensitive patient information and maintaining the privacy and security of the healthcare ecosystem.

5. Use of the App in Clinical Practice

5.1 Accessing the app via DANBIO

DANBIO is The Danish nationwide clinical register for patients with rheumatoid arthritis, accessible at https://danbio-online.dk/. The Prognosis Tool is integrated into the test environment, which can be accessed at https://db6-test.danbio.dk/. Once logged in, the user will be directed to the dashboard where they can search for the rheumatoid arthritis patient using the search bar. Upon navigating to the patient's page, 25 different clinical parameters about the patient will be displayed, part of which will serve as the baseline characteristics for the Prognosis Tool. The left navigation column contains a link for "PROGNOSIS-AI" which leads to the prognosis tool.

5.2 Using the Prognosis Tool

The Concept

The prognosis tool is a medical tool that predicts the likelihood of achieving DAS-28-CRP remission in patients with rheumatoid arthritis starting Adalimumab as their first-line biologic therapy. Adalimumab is a type of medication used to reduce the signs and symptoms of rheumatoid arthritis. The DAS-28-CRP score is a measure used in rheumatology to assess disease activity in patients with rheumatoid arthritis. The score ranges from 0 to 10, with higher scores indicating higher disease activity. The threshold for remission in DAS-28-CRP score is typically considered to be a score of less than 2.6. The tool aims to predict the probability of achieving a score of less than or equal to 2.6.

Benefits of using the tool

The tool can help rheumatologists make informed treatment decisions by providing them with a reliable estimate of the probability of achieving remission with Adalimumab treatment in a given patient. This information can be used to guide treatment decisions and help determine which patients are most likely to benefit from Adalimumab therapy. Usually there is a bit of trial-and-error associated with which drug to administer. This results in increased costs from trying out different drugs and increased suffering from the patient's end. The tool is a step forward in precision medicine, and it has the potential to reduce the aforementioned limitations.

Navigating inside the application

The Prognosis Tool is located on the following domain: https://prognosisai.zitelab.eu/. This link will first prompt a browser HTTP login. After authenticating, the homepage view of the app will be opened, which includes a brief description of the app's main purpose and background, including information about the wider project (NORA/ScandRA), associated congresses, and abstracts. The Prognosis Tool can be accessed via the left navbar or the "Get started" button. On the prognosis tool page, there is a form for filling out the baseline characteristics, which will be used as input for the machine learning model. Descriptions of the parameters are provided in the section 5.2.1. Additionally, there is a button to select a file, which is intended for future use when the baseline characteristics can be automatically provided from DANBIO in the form of a JSON file using SMART on FHIR. The button is included to illustrate that the model can accept a JSON file in the correct format. Following the form, there is an input for gamification. This idea was suggested by one of the clients and is intended to initiate a thought process about the probability of remission based on the given parameters, drawing from the clinician's past experiences. The user can adjust a slider to set the probability rating for remission, and there is also a free text input field located below the slider for additional comments.

5.2.1 The baseline characteristics used in the prognosis tool

The prognosis tool that utilizes machine learning incorporates a comprehensive set of baseline characteristics that are standard clinical features currently used by healthcare practitioners to assess the condition of rheumatoid arthritis patients:

Age: The age of the patient is an important factor in predicting the likelihood of remission because age can influence disease severity and progression.

Sex: Gender can also be a significant factor in predicting remission rates, as women are more likely to develop rheumatoid arthritis and may experience different symptoms and responses to treatment than men.

Smoking status: Smoking is a well-established risk factor for the development of rheumatoid arthritis and can also affect disease severity and response to treatment.

Symptom duration in months: The duration of symptoms is an important predictor of disease severity and response to treatment, with longer symptom duration generally indicating a more severe disease course.

Global score (0-100): This is a patient-reported assessment of their overall well-being and is used to assess disease activity and response to treatment.

Prednisolone prescription: Prednisolone is a corticosteroid medication commonly used to reduce inflammation and pain in rheumatoid arthritis patients, and its use can affect disease activity and response to treatment.

VAS Global: This is a patient-reported assessment of their overall disease activity and is used to assess disease severity and response to treatment.

VAS Doctor: This is a physician-reported assessment of the patient's disease activity and is used to assess disease severity and response to treatment.

VAS Pain: This is a patient-reported assessment of their pain levels and is used to assess disease activity and response to treatment.

VAS Fatigue: This is a patient-reported assessment of their fatigue levels and is used to assess disease activity and response to treatment.

HAQ: The Health Assessment Questionnaire (HAQ) is a patient-reported assessment of their physical function and disability, and is used to assess disease severity and response to treatment.

CRP (mg/L): C-reactive protein (CRP) is a blood marker of inflammation and can be used to assess disease activity and response to treatment.

Number of swollen joints (0-28): This is a physician-reported assessment of the number of joints that are swollen and is used to assess disease severity and response to treatment.

Number of tender joints (0-28): This is a physician-reported assessment of the number of joints that are tender and is used to assess disease severity and response to treatment.

X-ray erosions: X-ray images can be used to assess joint damage and predict disease progression and response to treatment.

Rheumatoid factor: Rheumatoid factor is an antibody that is present in the blood of many rheumatoid arthritis patients and can be used to help diagnose the disease.

Anti-ccp: Anti-cyclic citrullinated peptide (anti-CCP) antibodies are another blood marker that can help diagnose rheumatoid arthritis and predict disease

5.3 Feedback From Healthcare Professionals

Feedback was collected through interviews with three key stakeholders who specialize in rheumatology. The results were analyzed using thematic analysis.

5.3.1 Thematic Analysis

Thematic analysis is a qualitative research method that involves identifying and analyzing recurring themes or patterns within data, such as interviews or textual data. [26] Thematic analysis allows for a systematic and in-depth exploration of the underlying meanings, perspectives, and experiences expressed by participants. Alternative methods to thematic analysis include content analysis, narrative analysis, and grounded theory. Content analysis focuses on quantifying and categorizing data, whereas narrative analysis emphasizes the structure and storytelling elements of data, and grounded theory aims to develop theoretical frameworks from data.

However, these alternatives may not provide the same level of interpretive richness and flexibility as thematic analysis. Thematic analysis allows for a more nuanced understanding of complex and multifaceted data, and it can be particularly well-suited for exploring subjective experiences, emotions, and perceptions. Its weaknesses include potential subjectivity in the coding and interpretation process, as well as the reliance on the researcher's analytical skills and expertise.

Nevertheless, thematic analysis was chosen for processing the interviews and answering the research questions due to its compatibility with the exploratory nature of the study, its ability to capture the nuances of participants' perspectives, and its flexibility in adapting to the unique characteristics of the data collected.

5.3.2 Results

The First Interview

Application Development: The speakers discuss the development of an application that is connected to their thesis. They talk about the parameters and models used in the application, including how it can be accessed via DANBIO.

Trust in AI: The speakers discuss the issue of trust in AI models, particularly in the context of clinicians using the application. They acknowledge that more complex AI models can act as "black boxes" where the inputs and outputs are not easily understandable, and discuss how they are using explainability methods to make the model more transparent. They also discuss the importance of including as many relevant parameters in the model to create trust among clinicians.

Parameter Considerations: The speakers discuss various parameters that could be included in the application, including the duration of symptoms, type of Prednisolone prescription (tablet or injection), and history of biologic treatments. They discuss the potential impact of these parameters on the model's output and the importance of considering them in the application to make it more relevant and useful for clinicians.

Gamification Implementation: The speakers discuss the gamification features in the application, including a probability slider and a textarea for justifying guesses. They discuss the usefulness of these features and potential improvements, such as making the textarea more meaningful by including it in the medical note and changing the placeholder text. They also discuss the relevance of the probability slider after removing the textarea.

SHAP Plot: The speakers discuss the SHAP plot, which is used to explain how the model arrived at its output. They discuss the surprising effect of a specific variable (VAS Global) on the plot and how the arrows in the plot indicate the direction of effect on the probability. They also discuss the usefulness of the SHAP plot for validating the effects of specific drugs.

Overall, the themes in the interview revolve around the development, trust, and improvement of an application that is connected to a thesis and involves AI models for clinicians. The speakers discuss various parameters, gamification features, and the SHAP plot, and consider ways to make the application more relevant, transparent, and user-friendly for clinicians.

The second interview

Introduction and setting: The conversation begins with introductions and setting the context for the discussion, including the purpose of the application developed by Mihkel, which is to predict DAS-28-CRP remission in rheumatoid arthritis patients after six months of Adalimumab treatment.

Feedback on the application: The participants discuss the input parameters of the application, such as clinical test results and patient information, and how changing these parameters affects the final probability of remission. They also emphasize the need for integration of contra-indications of drugs to ensure patient safety, and the importance of inflammation as a parameter in the application.

Interpretation of results: The participants analyze the impact of different input parameters on the final probability of remission and interpret the findings. They discuss the positive and negative effects of different parameters on the results, and express interest in additional predictors, such as contra-indications, in the app.

Technical issues: The conversation briefly touches on technical issues, such as the need to replace codenames with proper names in the application and the limitation of input fields to integers only.

Positive feedback: Despite the technical issues discussed, the participants express positive feedback about the application and its potential for future integration with DANBIO or other databases to streamline the process of accessing and inputting patient information. They emphasize the value of AI in decision-making and the need for collaboration between clinicians and app developers to ensure the app meets the needs of healthcare practitioners.

Decision-making process: The rheumatologist follows a guideline from MedicinNord for prescribing drugs based on efficacy, safety, and cost. They mention that cheaper conventional synthetic drugs can be prescribed independently, but more expensive drugs require consultation with specialists. The rheumatologist emphasizes patient safety as a priority in the decision-making process.

Collaboration and improvement: The rheumatologist suggests having a follow-up meeting with the app developers to discuss improvements in the model and further evaluate its usefulness. They express a willingness to recommend the app to colleagues after improvements are made, and highlight the importance of collaboration between clinicians and app developers in improving the app.

Emphasis on data-driven decision-making: The rheumatologist appreciates the potential of the app to provide numerical outputs based on available data, indicating a preference for data-driven decision-making in clinical practice.

Value of AI in decision-making: The rheumatologist acknowledges the potential of AI-based tools, including the app, in aiding decision-making in clinical practice, but also emphasizes the need for the app's output to make sense and follow clinical intuition.

Interest in additional predictors: The rheumatologist expresses interest in additional predictors in the app, specifically the integration of contra-indications as a predictor to ensure patient safety.

Need for data integration: The rheumatologist discusses the possibility of integrating data from patient journals or databases, such as DANBIO, into the app to access relevant patient information, including contra-indications.

Importance of collaboration with clinicians: The conversation highlights the importance of collaboration between clinicians and app developers to ensure that the app meets the needs of healthcare practitioners and provides meaningful and useful outputs for clinical decision-making.

Overall, the main themes identified in the combined thematic analysis revolve around feedback on the application, interpretation of results, technical issues, positive feedback, the decision-making process, collaboration and improvement, emphasis on data-driven decision-making, value of AI in decision-making, interest in additional predictors, need for data integration, and the importance of collaboration with clinicians. The conversation emphasizes patient safety, the need for integration of contra-indications, and the potential of the app to aid data-driven decision-making in clinical practice.

The Thrid Interview

Introduction: The transcript begins with introductions among the participants, where Mihkel, an IT student, introduces himself and his project related to AI in healthcare. The theme of introducing oneself and setting the context for the discussion is established.

Project Description: The conversation then focuses on the ImagineMore project, specifically the AI tool developed by Mihkel's team to predict remission after Adalimumab treatment in patients with DANBIO data. The theme of discussing the details of the AI tool, its purpose, and its potential impact on clinical decision-making emerges.

Demonstration of the AI Tool: Mihkel proceeds to demonstrate the AI tool, explaining its parameters, gamification feature, and the use of SHAP explanations to interpret the model's predictions. The theme of explaining the functionalities and features of the AI tool, including its interpretability, is prominent in this part of the conversation.

Integration with DANBIO: The discussion shifts to the integration of the AI tool with DANBIO, a national database, and the potential benefits of automatically transferring data from DANBIO to the tool. The theme of data integration and ease of use for clinicians is discussed in relation to DANBIO.

Potential Use Cases: The participants discuss potential use cases of the AI tool, such as assisting clinicians in decision-making about drug administration based on the model's output, taking into account national guidelines and regulations. The theme of discussing potential scenarios where the AI tool could be useful in clinical practice and its alignment with existing guidelines and regulations is evident.

Feedback and Suggestions: The conversation concludes with feedback and suggestions from the participants, with a focus on the interesting aspects of the AI tool and its potential impact on clinical decision-making. The theme of feedback, suggestions, and reflections on the AI tool's usefulness and potential limitations is prominent in this part of the conversation.

Overall, the thematic analysis reveals themes related to the introduction of participants, project description, demonstration of the AI tool, integration with DANBIO, potential use cases, and feedback and suggestions. The conversation revolves around the features, functionalities, and potential impact of the AI tool on clinical decision-making, as well as its alignment with existing guidelines and regulations.

6. Summary

In conclusion, the proof of concept application has been successfully developed and valuable feedback has been gathered from key professionals, providing evidence of its positive outcome.

The findings from the thematic analysis support our hypothesis that integrating an XAI app within the EMR-system can offer transparent and understandable explanations for AI-based predictions and recommendations in rheumatology. This can ultimately result in improved clinical decision-making and better patient outcomes. However, it is important to acknowledge that there may be limitations to the use of XAI, including increased cognitive load for healthcare practitioners and challenges in explaining complex AI models to patients.

For future work it is important to further validate these hypotheses and address the limitations identified. It is thus advisable to conduct further research and evaluation of the XAI app in a real-world scenario, such as a randomized controlled trial investigating the impact of augmenting the work of healthcare providers. Additionally, ongoing collaboration between healthcare practitioners and app developers, to incorporate feedback from clinicians will be fundamental for building a user-centric XAI app in the context of clinical decision-making in Rheumatology.

The interviews conducted have revealed some of the underlying challenges that hinder medical professionals from effectively utilizing AI in their practice. One prominent concern is related to privacy, particularly in handling sensitive patient data. One early proposal was to incorporate SMART on FHIR in the application, allowing doctors to import patients' baseline characteristics directly from the EMR-system. However, due to emergent privacy concerns, this feature will be implemented in a later stage. This application can serve as a foundation for the development of a robust data agreement, addressing privacy concerns comprehensively.

As new machine learning models are planned in the pipeline, the development and improvement of the XAI app will continue based on user feedback, with a specific focus on enhancing its effectiveness in supporting clinical decision-making. Furthermore, prioritizing the implementation of error-handling mechanisms, adding a confusion matrix to the FAQ to further explain the accuracy of the model, and improving the medical note based on feedback, as suggested for UI improvements, will be important steps in the ongoing refinement of the XAI app.

References

[1] Sara Momtazmanesh, Ali Nowroozi, and Nima Rezaei. "Artificial intelligence in rheumatoid arthritis: Current status and future perspectives: A state-of-the-art review". en. In: *Rheumatol. Ther.* 9.5 (Oct. 2022), pp. 1249–1304.

[2] Farah Magrabi et al. "Artificial intelligence in clinical decision support: Challenges for evaluating AI and practical implications". en. In: *Yearb. Med. Inform.* 28.1 (Aug. 2019), pp. 128–134.

[3] Divya Gopinath. *A taxonomy of ML explainability techniques to understand your model.* Accessed on April 11, 2023. URL: https://towardsdatascience.com/picking-an-explainability-technique-48e807d687b9.

[4] Scott Lundberg and Su-In Lee. *A Unified Approach to Interpreting Model Predictions*. 2017. arXiv: 1705.07874 [cs.AI].

[5] Marco Tulio Ribeiro, Sameer Singh, and Carlos Guestrin. "" Why should i trust you?" Explaining the predictions of any classifier". In: *Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining*. 2016, pp. 1135–1144.

[6] Anupam Datta, Shayak Sen, and Yair Zick. "Algorithmic transparency via quantitative input influence: Theory and experiments with learning systems". In: *2016 IEEE symposium on security and privacy (SP)*. IEEE. 2016, pp. 598–617.

[7] Jacob R Greenmyer et al. "DAS28-CRP cutoffs for high disease activity and remission are lower than DAS28-ESR in rheumatoid arthritis". In: *ACR Open Rheumatology* 2.9 (2020), pp. 507–511.

[8] Johan Askling. *Development of clinical support tools for personalised medicine implementation*. Unpublished Project Proposal. 2021.

[9] *39th Scandinavian Congress of Rheumatology*. accessed on April 11, 2023. URL: https://scr2022.dk/abstract-submission/.

[10] *React Documentation*. Accessed on April 10, 2023. URL: https://react.dev/.

[11] *Google Trends for JavaScript frameworks*. Accessed on April 10, 2023. URL: https://trends.google.com/trends/explore?cat=31&q=React%20jobs,Angular%20jobs,Vue%20jobs.

[12] *Vite Guide: Why Vite?* Accessed on April 10, 2023. URL: https://vitejs.dev/guide/why.html.

[13] *The TypeScript Handbook*. Accessed on April 10, 2023. URL: https://www.typescriptlang.org/docs/handbook/intro.html.

[14] *SASS Documentation*. accessed on April 10, 2023. URL: https://sass-lang.com/documentation/.

[15] *Mantine documentation*. accessed on April 10, 2023. URL: https://mantine.dev/.

[16] *Emotion documentation*. accessed on April 10, 2023. URL: https://emotion.sh/docs/introduction.

[17] *Axios documentation*. accessed on April 10, 2023. URL: https://axios-http.com/docs/intro.

[18] *Zustand Github*. accessed on April 10, 2023. URL: https://github.com/pmndrs/zustand.

[19] *Wouter Github*. accessed on April 10, 2023. URL: https://github.com/molefrog/wouter.

[20] *FastAPI documentation*. accessed on April 10, 2023. URL: https://fastapi.tiangolo.com/.

[21] *Tidymodels documentation*. accessed on April 10, 2023. URL: https://www.tidymodels.org/learn/.

[22] *SHAPviz package github*. accessed on April 10, 2023. URL: https://github.com/ModelOriented/shapviz.

[23] Tianqi Chen and Carlos Guestrin. "Xgboost: A scalable tree boosting system". In: *Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining*. 2016, pp. 785–794.

[24] Chadia Abras, Diane Maloney-Krichmar, Jenny Preece, et al. "User-centered design". In: *Bainbridge, W. Encyclopedia of Human-Computer Interaction. Thousand Oaks: Sage Publications* 37.4 (2004), pp. 445–456.

[25] *DANBIO IT System*. Accessed on April 13, 2023. URL: https://danbio-online.dk/om-danbio/it-system-indrapportering-og-tilbagemelding.

[26] *Thematic Analysis: What it is and How to Do It*. Accessed on April 13, 2023. URL: https://www.questionpro.com/blog/thematic-analysis/#:~:text=Thematic%5C%20analysis%5C%20is%5C%20a%5C%20method,making%5C%20sense%5C%20of%5C%20the%5C%20data..

Appendix 1 – Non-Exclusive License for Reproduction and Publication of a Graduation Thesis[1]

I Mihkel Kaalep

1. Grant Tallinn University of Technology free licence (non-exclusive licence) for my thesis "Explainable Artificial Intelligence Application for Clinical Treatment Decision-Making in Rheumatology", supervised by Gert Kanter and Niels Steen Krogh
   1.1. to be reproduced for the purposes of preservation and electronic publication of the graduation thesis, incl. to be entered in the digital collection of the library of Tallinn University of Technology until expiry of the term of copyright;
   1.2. to be published via the web of Tallinn University of Technology, incl. to be entered in the digital collection of the library of Tallinn University of Technology until expiry of the term of copyright.
2. I am aware that the author also retains the rights specified in clause 1 of the non-exclusive licence.
3. I confirm that granting the non-exclusive licence does not infringe other persons' intellectual property rights, the rights arising from the Personal Data Protection Act or rights arising from other legislation.

22.05.2023

[1]The non-exclusive licence is not valid during the validity of access restriction indicated in the student's application for restriction on access to the graduation thesis that has been signed by the school's dean, except in case of the university's right to reproduce the thesis for preservation purposes only. If a graduation thesis is based on the joint creative activity of two or more persons and the co-author(s) has/have not granted, by the set deadline, the student defending his/her graduation thesis consent to reproduce and publish the graduation thesis in compliance with clauses 1.1 and 1.2 of the non-exclusive licence, the non-exclusive license shall not be valid for the period.

Appendix 2 – Project Links

Item	**Link**
prognosis-ai front source code	https://gitlab.cs.ttu.ee/mikaal/prognosis-ai
prognosis-ai API source code	https://gitlab.zitelab.eu/zitelab-common/zl.aiexperiments
Main issues	https://gitlab.cs.ttu.ee/mikaal/prognosis-ai/-/issues
Deployed frontend	https://prognosisai.zitelab.eu/
Deployed API endpoint	https://prognosis-ai-api.zitelab.eu/

Table 3. Project links.

Appendix 3 - Frontend Views

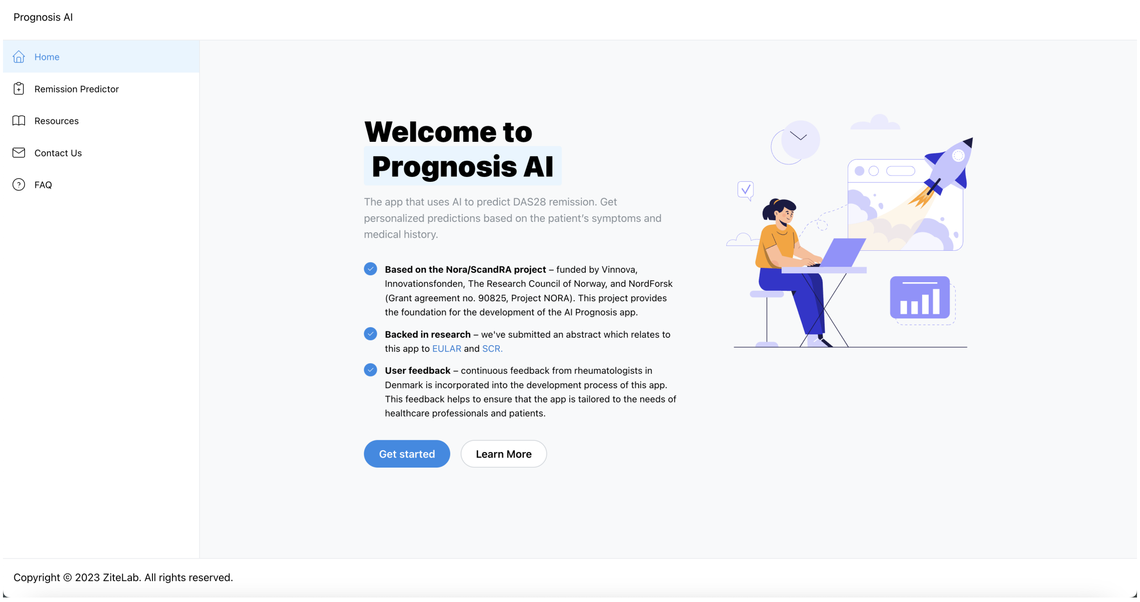

Figure 5. *Home view.*

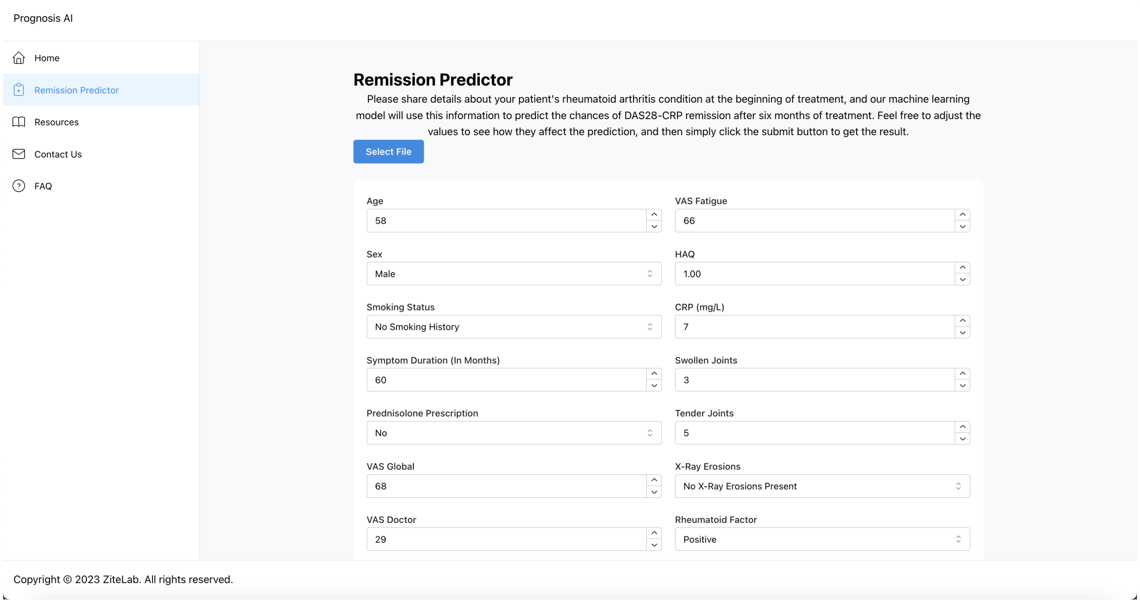

Figure 6. *Remission predictor inputs 1/2.*

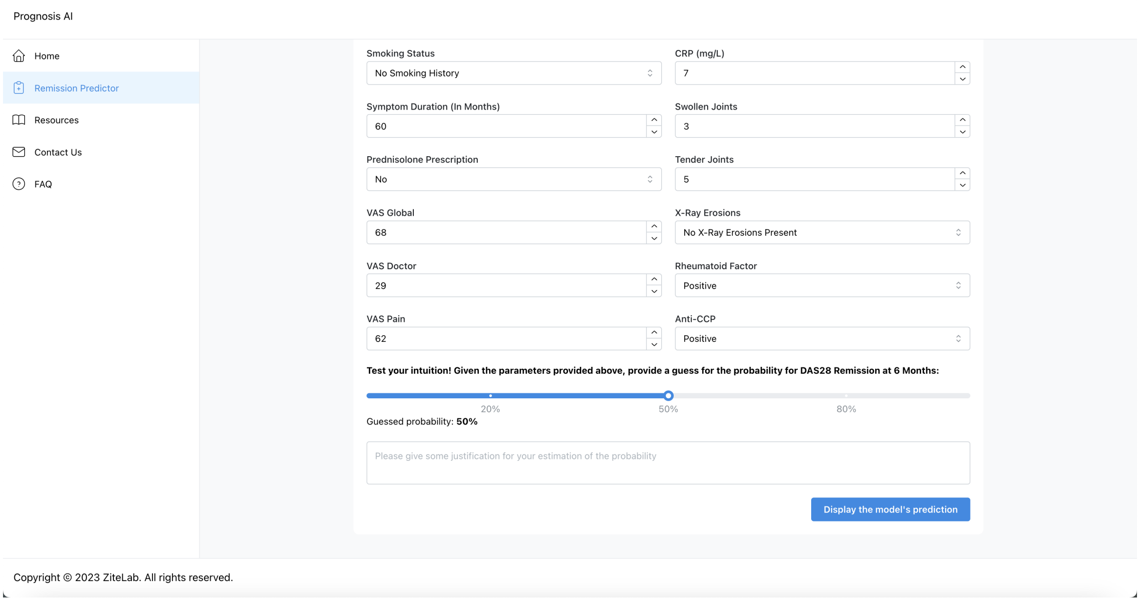

Figure 7. *Remission predictor inputs 2/2.*

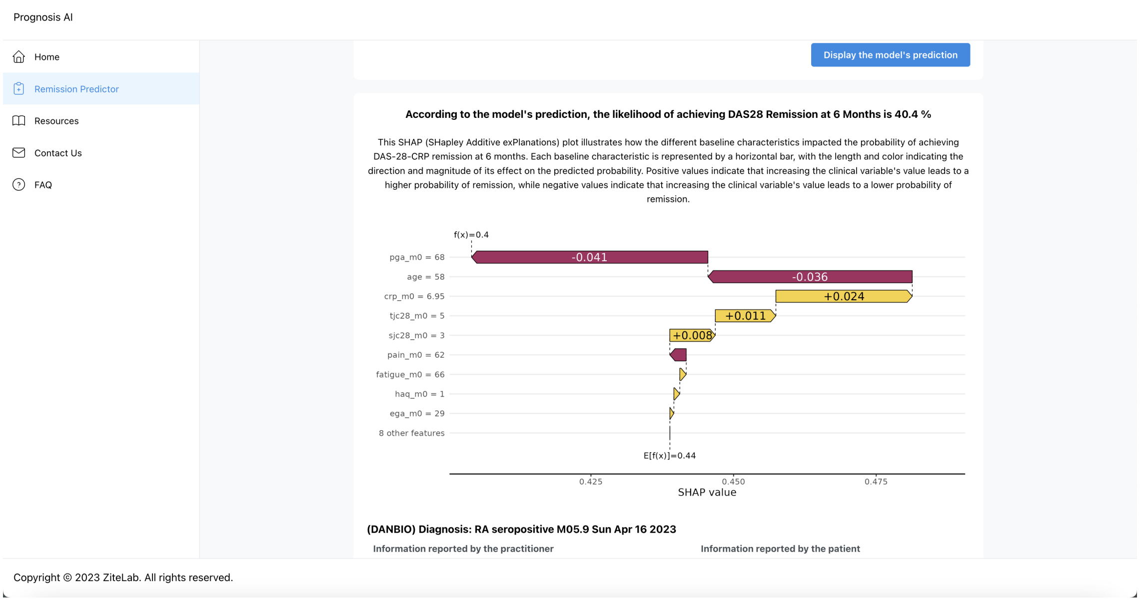

Figure 8. *Remission predictor output 1/2.*

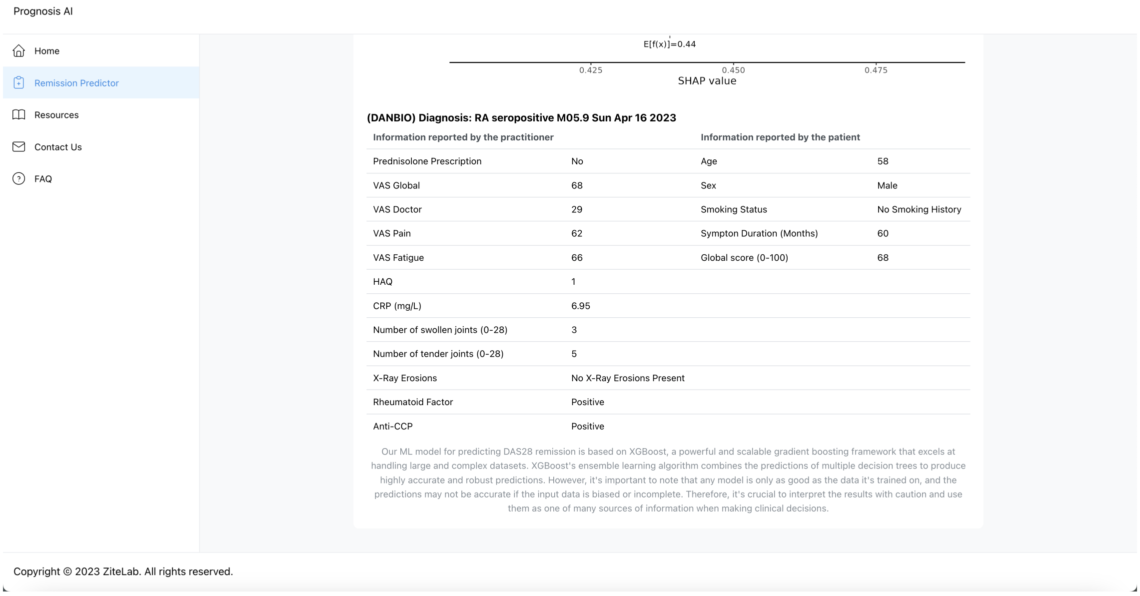

Figure 9. *Remission predictor output 2/2.*

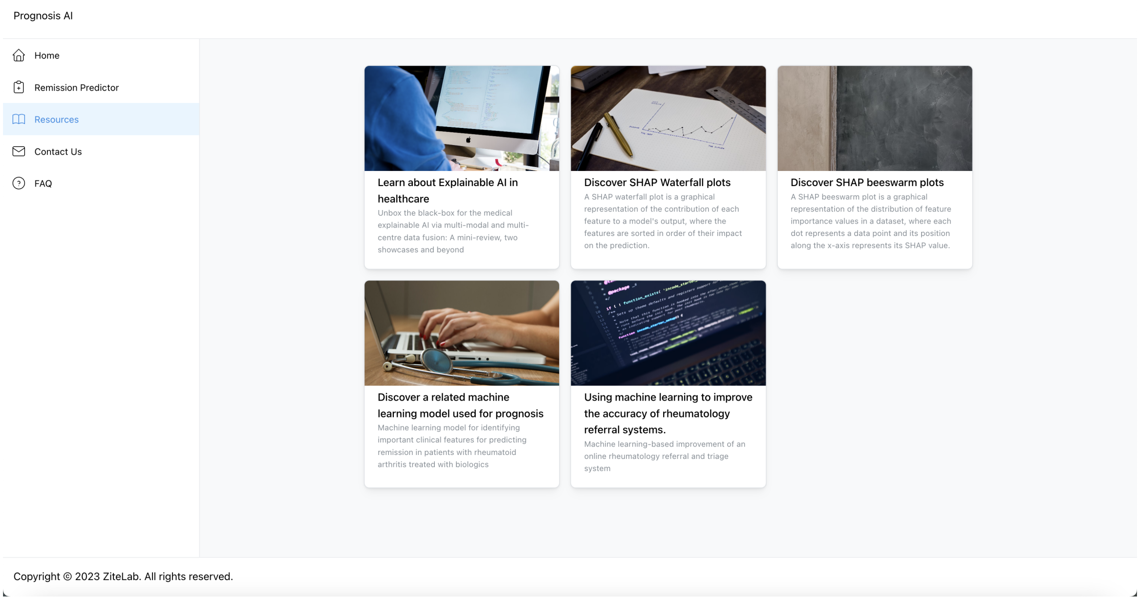

Figure 10. *Resources view.*

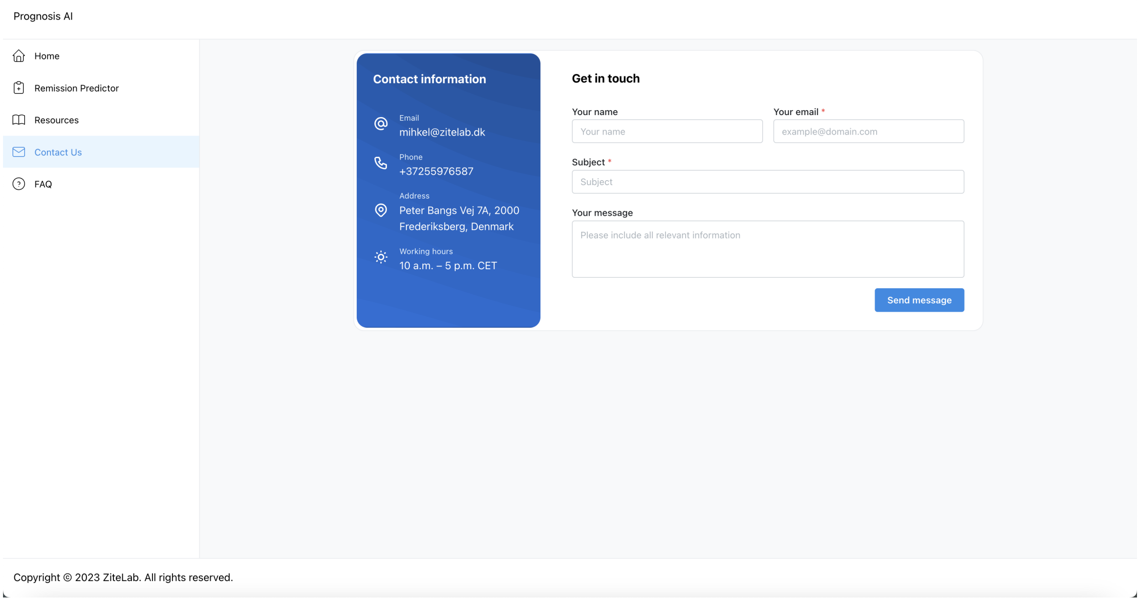

Figure 11. *Contact us view.*

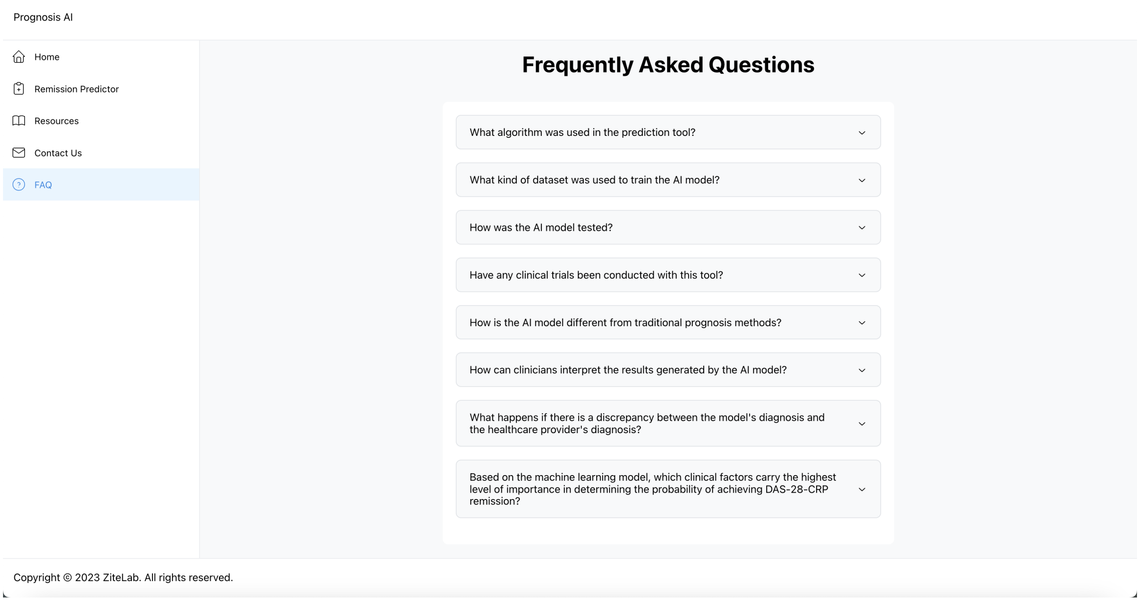

Figure 12. *Frequently asked questions view.*

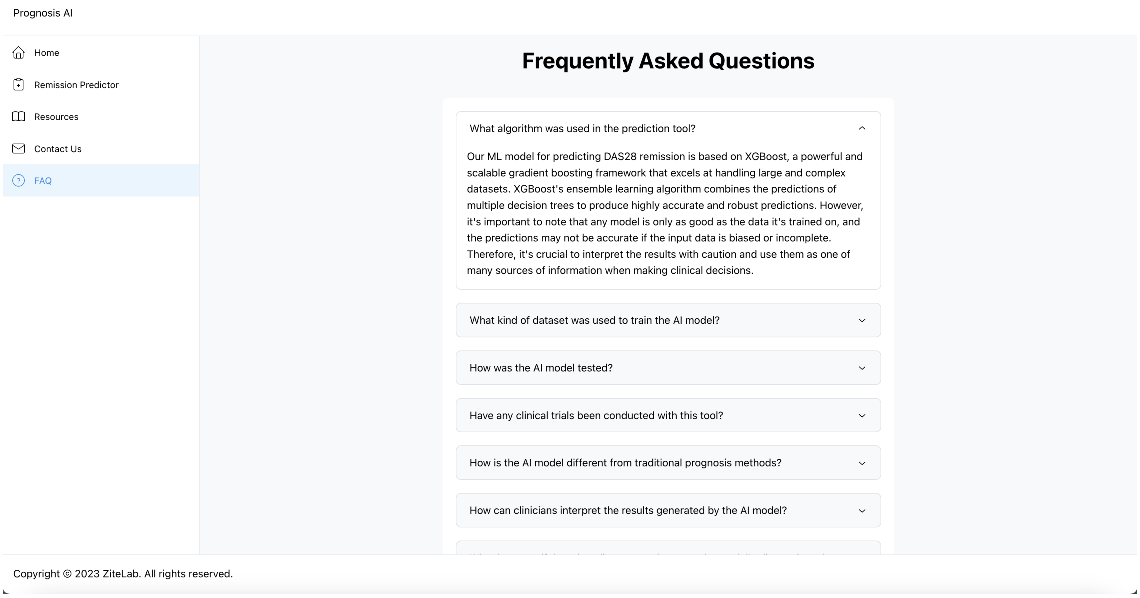

Figure 13. *Frequently asked questions - the first question's answer view.*

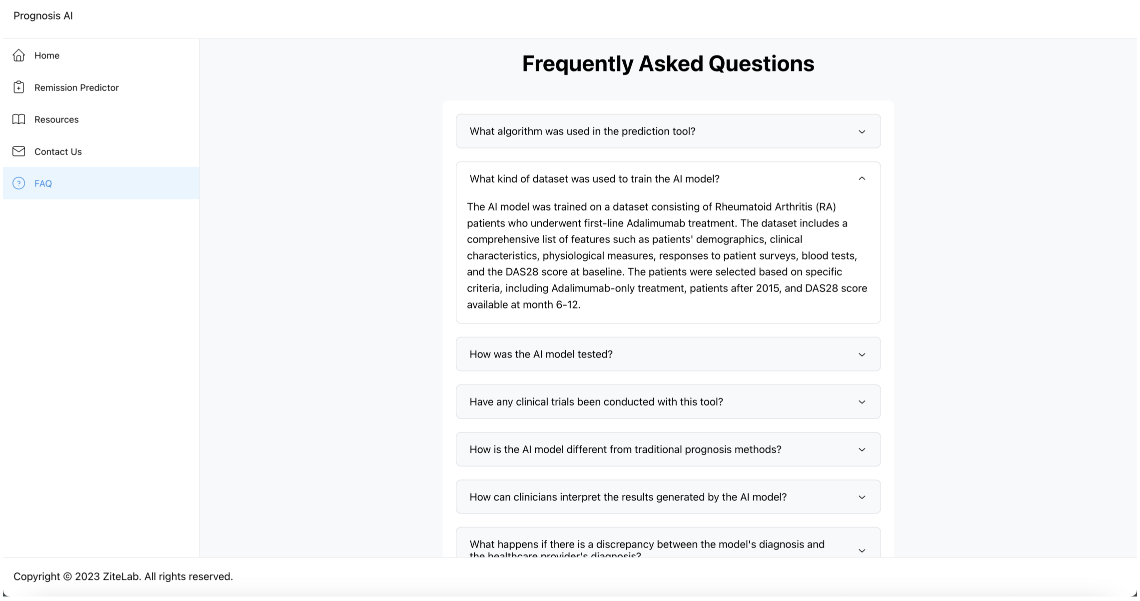

Figure 14. *Frequently asked questions - the second question's answer view.*

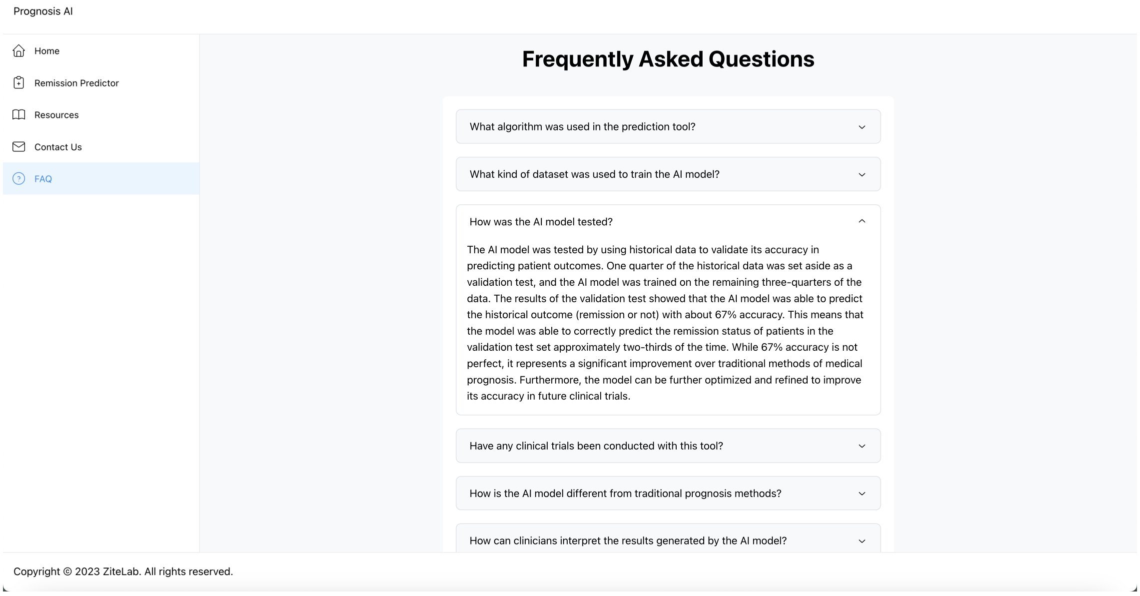

Figure 15. *Frequently asked questions - the third question's answer view.*

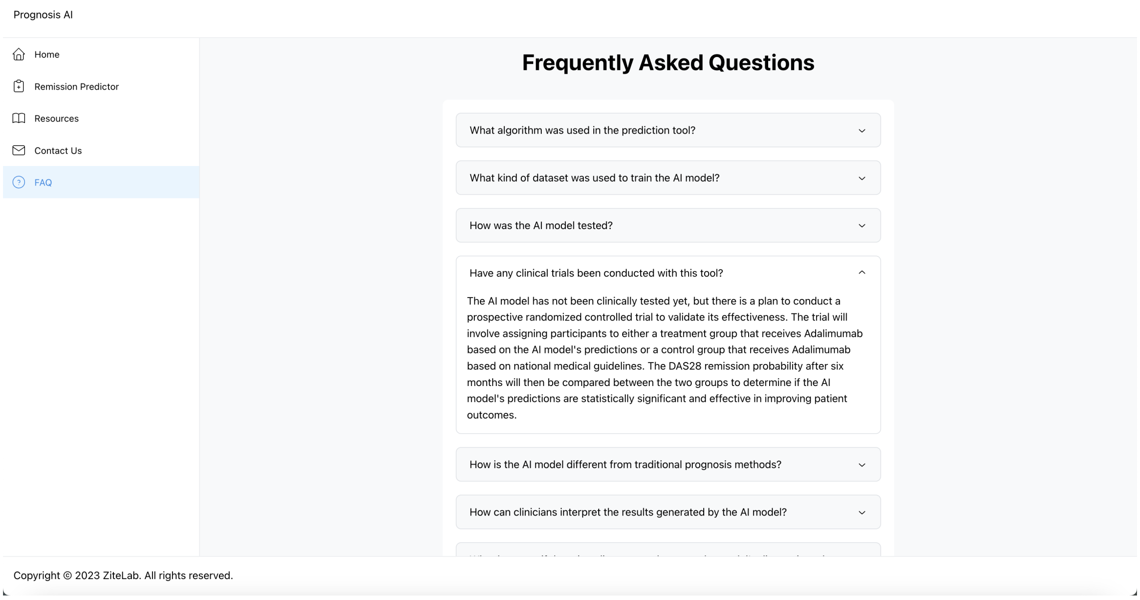

Figure 16. *Frequently asked questions - the fourth question's answer view.*

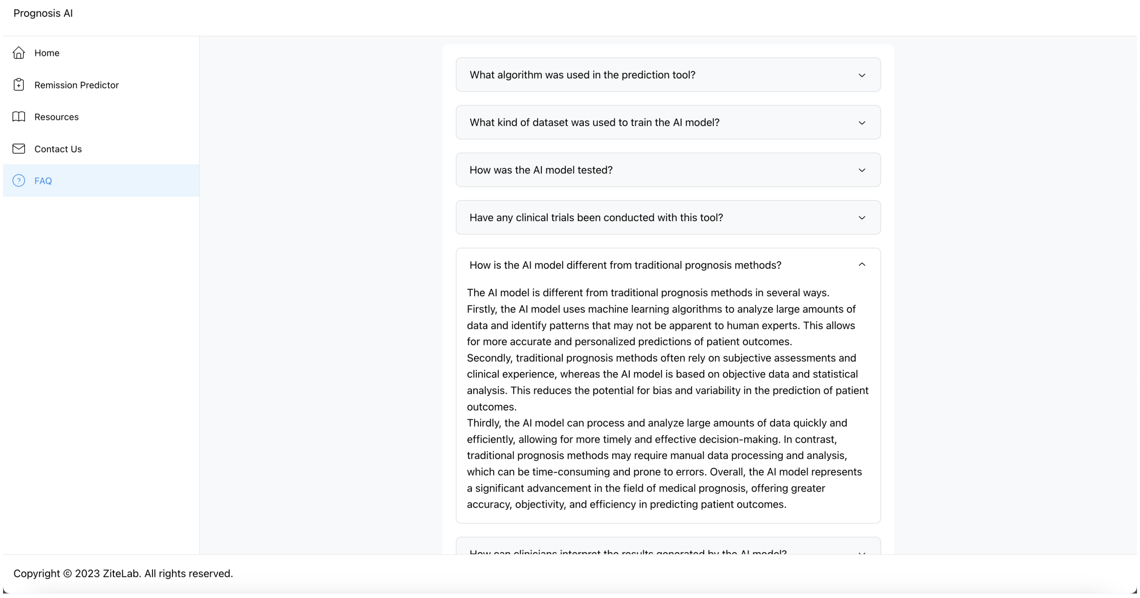

Figure 17. *Frequently asked questions - the fifth question's answer view.*

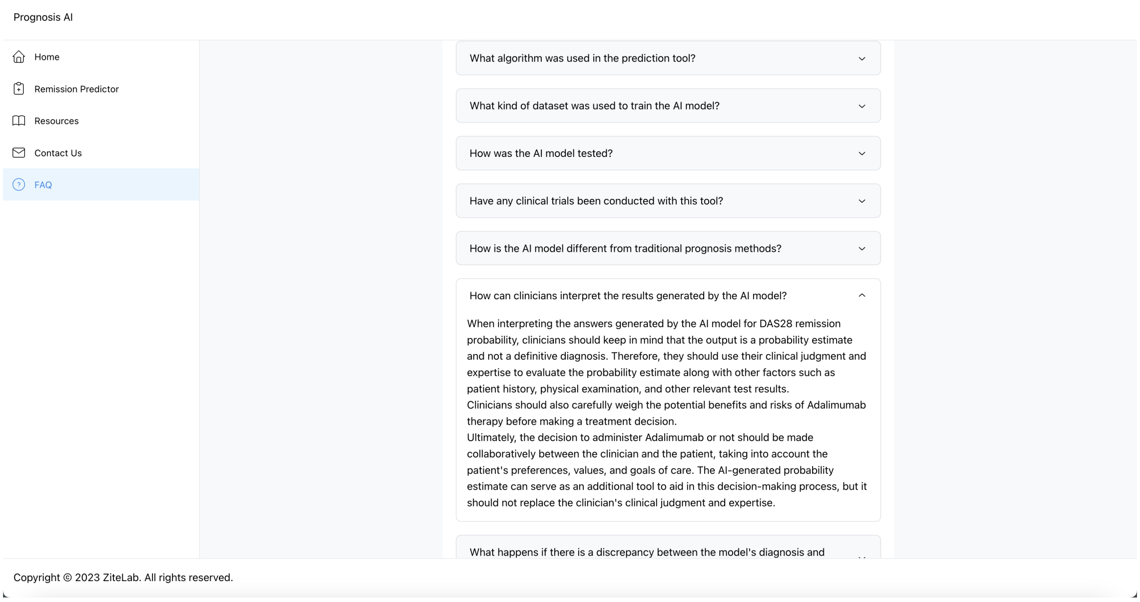

Figure 18. *Frequently asked questions - the sixth question's answer view.*

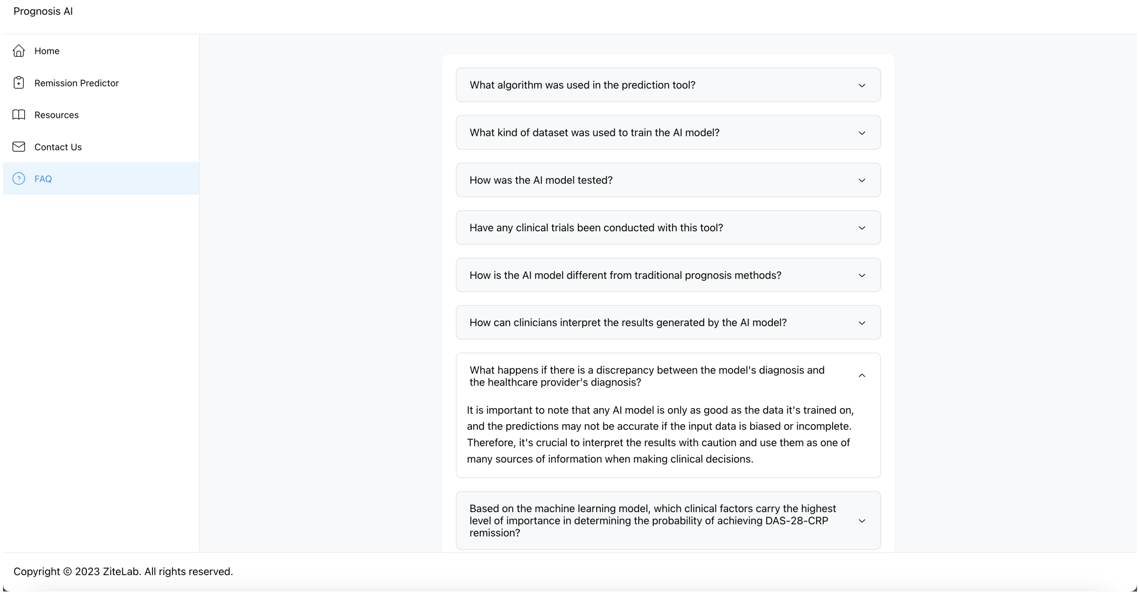

Figure 19. *Frequently asked questions - the seventh question's answer view.*

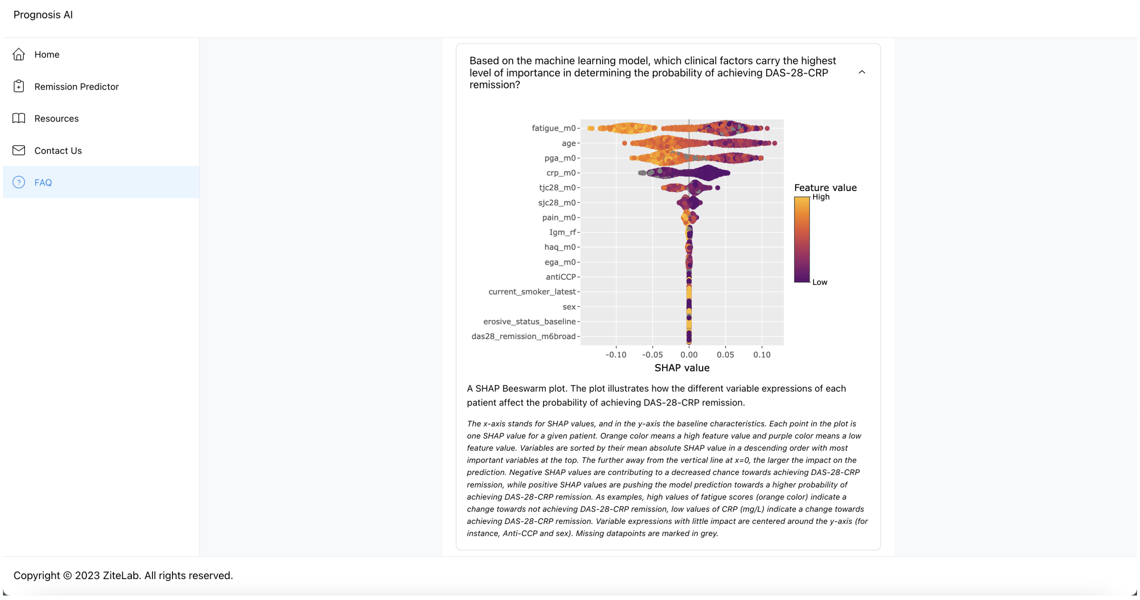

Figure 20. *Frequently asked questions - the eighth question's answer view.*