University Centre

Abd Elhafid Boussouf Mila

Institute of science and technology          Department of Mathematics and Computer Science

# Thesis Prepared For The Master's Degree

## In : Computer Science

### Speciality : Information and Communication Science and Technology (ICST)

# Realization of a system of prediction and patients guidance within a clinic.

**Presented by :** Khitas Mahdi

Allam Khalil

**Board of examiners :**

| | |
|---|---|
| Supervisor : Mme Boufaghes Hamida | MAA U.C.Abd Elhafid Boussouf |
| Chairman : Mme Bouchekouf Asma | MAA U.C.Abd Elhafid Boussouf |
| Examiner : Mme Boumessata Meriem | MAA U.C..Abd Elhafid Boussouf |

**Academic Year: 2021/2022**

# Acknowledgements

**Dedication:**

In the name of Allah, Most Gracious, Most Merciful ♡ ♡ ♡ I would gratefully dedicate this work to My guardian angel, my first and only love, the candle that lighten my path, my lovely mother ♡ ♡ ♡ ♡ **Zakia** ♡ ♡ ♡ ♡ . to my role model, my superhero, my right arm, my constant source of inspiration, my great father ♡ ♡ ♡ ♡ ♡ **Hassan** ♡ ♡ ♡ ♡ . To the most precious and best brothers ever: ♡ **Ikbal** and **Kenzi** ♡ . To the sweetest and kindest sisters: ♡ ♡ **Tesnim** and **Imen** ♡ ♡ . Special thanks to my brother from another mother and my partner in this work ♡ ♡ ♡ **Khalil** ♡ ♡ ♡ for his patience and collaboration. To my coach, my second father and mentor ♡ ♡ **Ali Mellah** ♡ ♡ . To my all demented but kind friends whom they ought to be grateful for such a ray of sunshine they blessed with **Bilel**, **Yahia**, **Djaber**, **Aymen**, **Charaf**, the two **Yassers** and the one and only **Snake**. To every friend and brother, I've proudly won in my 25 years. To the teachers that left a huge impact in my life i wish all the best ♡ ♡ Messrs. **Merabet**, **Laalou**, **Djaboub**. ♡ ♡ and ♡ ♡ Mses. **Zekiouk**, **Boufaghes**, **Hedjaz** ♡ ♡ . ♡ ♡ To the one, the special one and the only one. ♡ ♡ as a wise man once said: "i wanna thank me, I wanna thank me for believing in me, I wanna thank me for doing all this hard work, I wanna thank me for never quitting, I wanna thank me for doing all this hard work, I wanna thank me for tryna do more right than wrong, I wanna thank me for just being me at all times" ♡ ↬ "**Noufel**" ↫ ♡ ♡ ♡ ♡ ♡ ♡ ♡ ♡

**Khitas Mahdi**

**Dedication:**

I thank God Almighty first and foremost for the great grace that He has bestowed upon me, then I thank those who favored them. My beloved mom **Djahida Nezzar** do not cease to me for all their efforts from the moment of my birth to these blessed moments, my second mom and my friend **Fofo**. my brother and my partner **Noufel**, and brother **Bilel**. For everyone who advised me, guided me, contributed, or directed me preparing this research and connecting me to the required references and sources at any of the stages it went through, and I especially thank the distinguished professor: "**Boufeghse Hamida**", for helping me Supporting me and guiding me with advice, education, correction, and all that she did with me. I am also pleased to thank the esteemed college administration: only my partner, **Yahia** and Flose **Issam**, **Borhane**, **Lamin**, and the department of specialization, without forgetting my brother **Yasser**, **Yasser** Ben Nasser and . ♡ ♡ ♡ ♡ **Khalil** ♡ ♡ ♡ ♡ ♡ ♡ ♡ ♡ ♡ ♡ ♡ ♡ ♡ ♡ ♡

***Allam Khalil***

# Résumé

À l'ère du numérique, il faut que le monde fournisse un excellent système de santé pour que les citoyens et la société soient en vie. Cette étude propose une application Web qui fonctionne avec des algorithmes de prospection de données pour prédire la maladie et qui peut finalement former un système de prévision de la santé qui sert les patients et les médecins.

Les principales caractéristiques de ce système sont de donner le diagnostic immédiat de l'utilisateur sur la base des symptômes dont il souffre et d'obtenir une liste de médecins spécialisés dans cette maladie ainsi que de nombreux services fournis par cette adhésion comme la réservation de rendez-vous, numérisation des ordonnances, analyses médicales et suivi précis de l'état du patient.

Pour réaliser cette aspiration, l'UML2 a été choisi comme méthode, 2TUP comme approche de conception, Java EE comme plate-forme opérant en Java, hybrnet pour le traitement des bases de données et de l'informatique en nuage sarver lié à la construction des systèmes intelligents et des réseaux de Microsoft en tant qu'environnement de développement ainsi qu'Apach tomcat en tant que serveur et mysql en tant que base de donée

---

**Mots clés :** Apprentissage automatique, Java, prédiction de maladies, orientation patient, application Web.

---

# Abstract

The age of digital technology requires the world to provide an excellent health system, in order to ensure that citizens and society are alive and well. This study proposes a web application that operates with data prospecting algorithms to predict disease and that can ultimately form a health forecasting system that serves patients and doctors.

The main features of this system are to give the immediate diagnosis of the user based on the symptoms he suffers from and to obtain a list of doctors specializing in this disease as well as many services provided by this accession such as reserving appointments, digitizing prescriptions, medical analyses and accurately tracking the patient's condition.

In order to achieve this aspiration, the UML2 has been selected as a method, 2TUP as a design approach, Java EE as a platform operating in Java, hybrnet for dealing with databases and sarver cloud computing related to building Microsoft's smart systems and netbeans as a development environment as well as apach tomcat as a server and mysql as a database

**Keywords :** Machine learning, Java, disease prediction, Patient orientation, Web application.

# ملخص

في العصر الرقمي، يجب على العالم توفير نظام رعاية صحية ممتاز حتى يكون المواطنون والمجتمع على قيد الحياة.

تقترح هذه الدراسة تطبيقًا قائمًا على الويب يعمل مع خوارزميات التنقيب عن البيانات للتنبؤ بالمرض ويمكن أن يشكل في النهاية نظام تنبؤ صحي يخدم المرضى والأطباء.

تتمثل السمات الرئيسية لهذا النظام في إعطاء التشخيص الفوري للمستخدم على أساس الأعراض التي يعاني منها والحصول على قائمة بالأطباء المتخصصين في هذا المرض بالإضافة إلى العديد من الخدمات التي تقدمها هذه العضوية على النحو التالي: حجز المواعيد، ومسح الوصفات الطبية، والاختبارات الطبية، والتتبع الدقيق لحالة المريض.

ولتحقيق هذا التطلع، تم اختيار UML2 كطريقة، 2TUP مع اقتراب التصميم، Java EE كمنصة تعمل في hybrnet Java، لمعالجة قواعد البيانات والحوسبة السحابية المتعلقة ببناء أنظمة وشبكات Microsoft الذكية كبيئة تطوير بالإضافة إلى Apach tomcat كخادم و mysql كقاعدة بيانات.

---

**كلمات مفتاحية :** تعليم الالة, تنبؤ مرض, تطبيق, توجيه المرضى, جافا.

---

# Contents

# List of Figures

# List of Tables

# List of acronyms

**SHPS**        *Smart Health Prediction System*

**API**        *Application Programming Interface*

**AI**        *Artificial intelligence*

**ML**        *Machine Learning*

**2TUP**        *Two Track Unified process*

**UML**        *Unified modeling language*

**DL**        *Deep Learning*

**NB**        *Naive bayes*

**RF**        *Random forest*

# General Introduction

According to the Algerian institute for health metrics and evaluation, There are some scary numbers in the last statistics, 79% of deaths in the past 10 years is caused by diseases, 48% of these persons didn't knew about their disease, 27% had false diagnostic from their doctors, 56% of the Algerian people can't afford to visit the doctor, 46% of Algerian hospitals are classified as unqualified by the global health organization.

Numerous issues are faced by patients pertinent to health facilities such as being unable to provide medical services, insufficient number of qualified medical staffs, poor communication between doctors and patients, and unorganized health records and data, renewal of records at each hospitalization, loss of data and ability to trace patient history and care, data redundancy in each document that constitutes the medical record, patients should describe their history each time they visit a new health care professional.

We could keep going and cite more issues such as lack of quick and easy communication between medical actors to share a patient's medical information, absence of confidentiality while the secretary can access the patient's confidential information, but the respect of medical secrecy requires that only the doctor can consult this record, In the event of an emergency, medical staff find doctors to get the information they need in order to get some treatment, difficult search, excessive waste of time and paperwork, the management of appointments is done in a manual way which creates a risk of forgetting or overlapping appointments, false diagnosis and human error, lack of hospitals and clinics in some areas, majority of the patients don't take some symptoms seriously to not wast any time or money on doctors until it's too late, expensive healthcare system.

These issues motivated us to build a Web-based application using ML algorithms,that allows users to get instant guidance on their health issues through an intelligent health care system online, the system is fed with various symptoms and the disease/illness associated with those symptoms,it allows user to share their symptoms and issues, then processes user's symptoms to check for various illness that could be associated with it.

The main feature of the system is similar to a doctor, where the doctor asks the symptoms of the patient and predicts the disease that the patient is/maybe. The system predicts the disease based on the symptoms that the user will enter. The system analyzes the patient's symptoms and based on that searches in its huge data set for similar trends of data and predicts the best, then it will ease out the process of booking doctor's appointment. It is also useful for the doctors as they get to keep track of their patients on a system without the help of an assistant.

Our new system also allows the digitization of medical records, prescriptions and analysis. Doctors get more clients online, keeps the history of all the tasks that constitute the medical record, Paper and time savings and quick access to information, manage access rights, to allow selective access to the various menus and assign responsibilities to each user, avoid data redundancy, the automation of communication between medical staff (doctors, pharmacists, laboratory) and the patient.

## Chapters content

This work is composed of 6 chapters:

The first one **Preliminary study** In the first chapter, we will start by presenting our project including the solved issues, then will try to identify functional and Non-functional requirements of our system.

The second one **Artificial Intelligence and Machine Learning** we will target to define artificial intelligence and machine learning besides their subdivisions, algorithms and even their use cases.

the third one **Smart health prediction system** we'll talk about the intelligent model that we've created, and go throw the steps of how it works. the fourth chapter **Capturing requirements** we will present and detail its

needs by use cases which allow us to clearly specify the various tasks that must be performed by each actor in the system, and to complete the process we will also describe the technical needs. the fifth chapter **Analysis phase**

This chapter is devoted to the object analysis of our system. It consists in precisely studying the specifications to get an idea of what the system will achieve in terms of trades. We start with the diagrams of participating classes. Then we move on to the Detailed Sequence Diagrams.

the sixth chapter **Realisation** We will in this part describe the process of achieving our application. This is done by specifying the development environment and an overview of the interfaces of our application.

# Chapter 1

# Preliminary study

## 1.1 Introduction

The preliminary study is the first phase of our development process. It consists in identifying functional and operational needs, using mainly texts or very simple diagrams. It prepares the most formal activities such as capturing functional and technical needs.

At the level of this chapter we will begin to determine the functional requirements considering the system as a black box, in order to study its place.

After identifying the actors who interact with the system, we will develop a first UML model at the context level, to be able to establish precisely the functional boundaries of the system .

## 1.2 2 track unified process

2TUP is a software development process that implements the unified process. Each step of the cycle follows the previous steps. The method is also incremental: from the capture of functional needs, we define several use cases each representing an increment of the development cycle [16]. the following figure represents The Y development process:
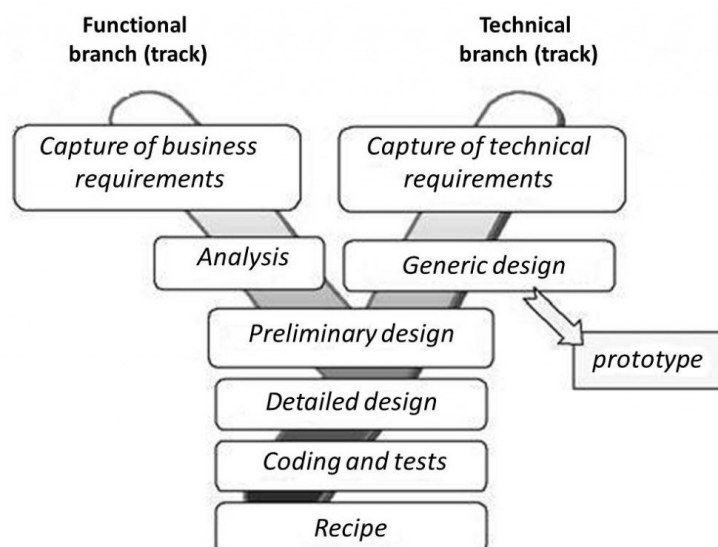


Figure 1.1: The 2TUP development process

## 1.3    Stage status in the 2TUP process

The preliminary study is the very first step of the 2TUP process, it consists in making an initial identification of functional and operational needs, using mainly text, or very simple diagrams. It prepares the most formal activities for capturing functional and technical needs [14].
the following figure represents the stage status in the 2TUP process:
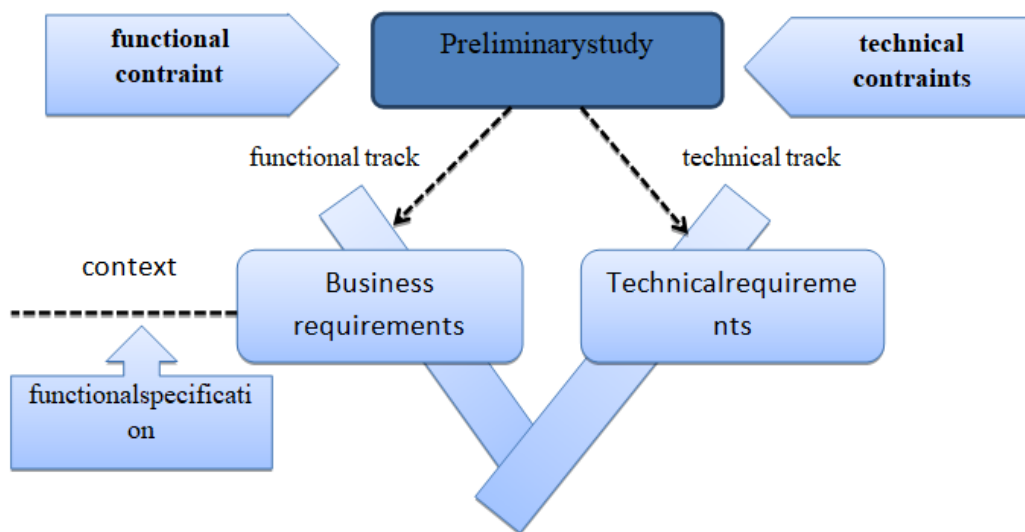


Figure 1.2: The stage status in the 2TUP process

## 1.4    Preparation of the specifications

The specifications are defined as an act, a reference document that allows a company manager to specify the conditions, rules and requirements of an assignment, an intervention, a task to be performed or a task to be performed by a management consultant, with a view to solving a specific problem or improving a given situation, while determining the expected results [27]. The specifications are a detailed service request, designed to protect the interests of the company manager and to improve the quality of the offer submitted by the consultant [18].

## 1.4.1 Project presentation

- **<u>Problematic</u>**
  During our search in some clinics and medical facilities such as "Afia medical clinic", we found several abnormalities and malfunctions in the management of patients records. We have summarized them in the following points:

  – Renewal of records at each hospitalization.

  – Loss of data and ability to trace patient history and care.

  – Data redundancy in each document that constitutes the medical record.

  – Patients should describe their history each time they visit a new health care professional.

  – Lack of quick and easy communication between medical actors to share a patient's medical information.

  – Absence of confidentiality : the secretary can access the patient's confidential information, but the respect of medical secrecy requires that only the doctor can consult this record.

  – In the event of an emergency, medical staff find doctors to get the information they need in order to get some treatment.

  – Difficult search, excessive waste of time and paperwork.

  – The management of appointments is done in a manual way which creates a risk of forgetting or overlapping appointments.

  – False diagnosis and human error.

  – lack of hospitals and clinics in some areas.

  – Majority of the patients don't take some symptoms seriously to not wast any time or money on doctors until it's too late.

  – Expensive healthcare system.

  – Incapability in managing patients medical record.

- **Objectives**

  Our mission of this project is to build a system to predict diseases, according to patient symptoms, manage patients in a medical facility. This involves defining the responsibilities of management, updating the data, organising the data collected in order to design basic files for the doctor, strengthening control and communication, to ensure better medical management of patients and consistency of information and to facilitate the work of those responsible. The main features of our application will be:

  - Patient Medical Record Management and Tracking.

  - Patient Appointment Management.

  - Patient Management.

  - Disease prediction according to symptoms.

  - Patient orientation.

### 1.4.2 Technical choices

For the modeling of our application we will use an iterative, incremental approach based on the Y process (2TUP) and the UML language, as well as JAVA programming language and servlets.
Also a MySQL database management system for data management and storage, without forgetting html, css, and javascript for front-end design.

### 1.4.3 Defining of needs

After the presentation of our project, we will outline all the necessary requirements. for the functioning of our system, both functional and non-functional.

- **Functional needs**

  Functional needs express actions that must be carried out by the system by actors each according to its role [15]. A study of the clinics employees allowed us to collect the needs expressed, which enabled us to draw up the following preliminary specifications:

  Demands an account : After the guest access the system, it allows him to create a new account by filling up all the information required, the guest could create a medical file after creating an account, than he could manage (update/delete) the his account and medical file.

  Get clinics information : The system allows the guest to consult all specialities, doctors and the clinics information including the clinics location.

  Get disease prediction : The patient gets the prediction result, the appointments and the doctors specialised in the predicted disease after entering symptoms to the prediction system .

  Making an appointment : the patient could make an appointment after getting the result from the prediction system, also he could consult, update or even cancel his appointments.

- **Non-functional needs**

  These are requirements that characterize the system, they can concern all technical constraints and implementation (programming language, DBMS, operating system...) to ensure the proper functioning of the system [15].

  So, for our application, we have drawn some needs to have a good quality functioning, we can summarize them as follows:

  - Usability: the application must be clear and easy to use.

  - Ergonomics: Application interfaces must be ergonomic.

  - Security: User data must be protected.

  - Availability: application must be available for use by any user.

  - Update: application must be open, accepts improvements to ensure the evolution.

## 1.5   Context description

Once this first collection of requirements has been completed, the description of the system context can begin [17]. It consists of three successive activities:

- actors identification.

- messages identification.

- context diagram.

### 1.5.1   actors Identification

Now, we will describe the different actors who must interact in the system, each of which is intended to achieve a set of functionalities:

Visitor (guest) : He is an unidentified person, who can access the site to consult the available specialities and doctors and even the clinics information, register to become a patient and create his medical file.

Patient : he is the person identified by the system, he is able to predict a disease and consult it, make an appointment and manage it, also he could consult his profile and update it.

Doctor : He is the responsible of consulting both medical files and appointments list and manage theme, he is able to insert prescriptions and analyses.

pharmacist : He is the actor responsible of consulting patients prescriptions and adding notes on it.

Laboratory : He is the actor responsible of consulting the patients analyses and add the results

Admin : The actor responsible of managing users and medical staff, he is also the responsible of consulting and replying on feed-backs.

Prediction system : Plays the role of secondary actor,used by the patient to predict the disease.

Google maps : A secondary actor, used by the other users to get the location of the clinic.

Data base : Considered as a secondary actor, used by the admin to verify validation.

### 1.5.2   Messages Identification

After we have described all the actors of our system, in this part we must detail the various messages exchanged between the system and the actors. The system sends the following messages:

- Display list of patients, doctors, specialities.

- Display the different forms(add,update).

- Notifications.

  The system received the following messages:

- Request for subscription, consultation.

- Authentication request.

- Request for addition, modification, deletion and consultation.

### 1.5.3   Context diagram

The context diagram is a conceptual flow model that allows to have a global vision of the interactions between the system and the links with the external environment.  For our case the context diagram is given by the following figure:

## 1.6   Conclusion

At the end of this chapter, we presented the different needs necessary for the functioning of our application, whether functional or non-functional, so that the various actors interact with the system. We can say that this step is really necessary and fundamental to be advanced to the development of our project and for the preparation of the next step.
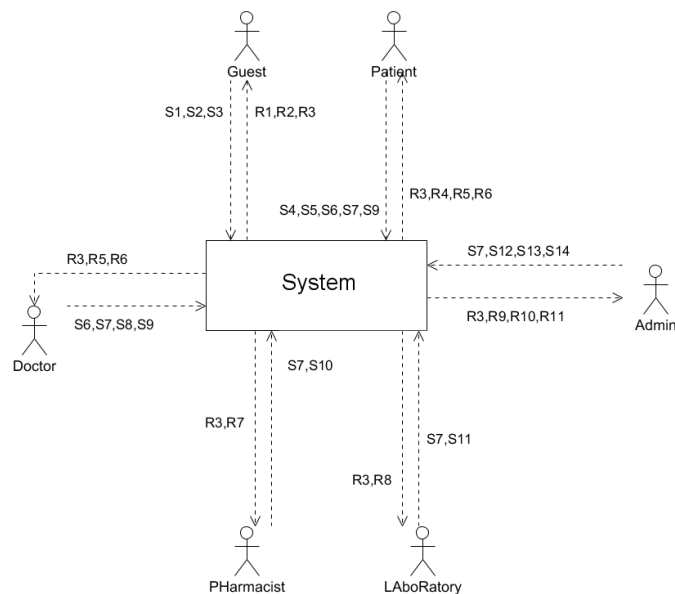
Figure 1.3: Context diagram

| Sent | Recived |
|---|---|
| S1:consult available specialties and doctors | |
| S2:consult clinics information | |
| S3:demand of subscription | R1:list of available specialties and doctors |
| S4:Predict the disease | R2:display clinics information |
| S5:consult prediction | R3:display necessary forms |
| S6:manage appointments | R4:display prediction results |
| S7:demand log-in | R5:list of appointments |
| S8:manage medical file | R6:display medical file |
| S9:consult medical file | R7:display prescription |
| S10:consult prescription | R8:display analyses |
| S11:consult analyses | R9:list of medical staff |
| S12:mange medical staff | R10:list of users |
| S13:manage users | R11:display feed-backs |
| S14:consult feed-backs | |

Table 1.1: Group of messages between actors in context diagram

# Chapter 2

# Artificial Intelligence and Machine Learning

## 2.1 Introduction

Machine learning, or more commonly, artificial intelligence, where machine learning is a sub-domain, evokes science fiction for most people. However, machine learning is not a dream of the future but it is already part of everyday life to be honest, it has been standing for decades as anti-spam filters in the 80s or video games since the early 2000s.

## 2.2 Definitions

- **Artificial Intelligence (AI)**
  is a set of defined methods, tools and systems for resolving problems. problems that require human intelligence to solve. Intelligence is always defined as the ability to learn effectively, to react effectively, to make a good decision, to communicate in language or images in a sophisticated way, and to understand.

  AI is the science and engineering of making intelligent machines, especially intelligent computer programs, or more concrete: ". . . is the capacity of a computer program to perform tasks or reasoning processes that we usually associate to intelligence in a human being" [1].

- **Machine learning (ML)**
  is a subset of AI. ML provides software, machines, devices, and robots with the ability to learn without human intervention or assistance or static program instructions [2], Machine learning is at the heart of the data science and artificial intelligence[3].

## 2.3 machine Learning development cycle

the field of application of machine learning is varied: Prediction of financial values, intrusion in the field of IT security, the search engine can be influenced by the user's profile, theft detection The life cycle of a machine learning implementation is as follows:
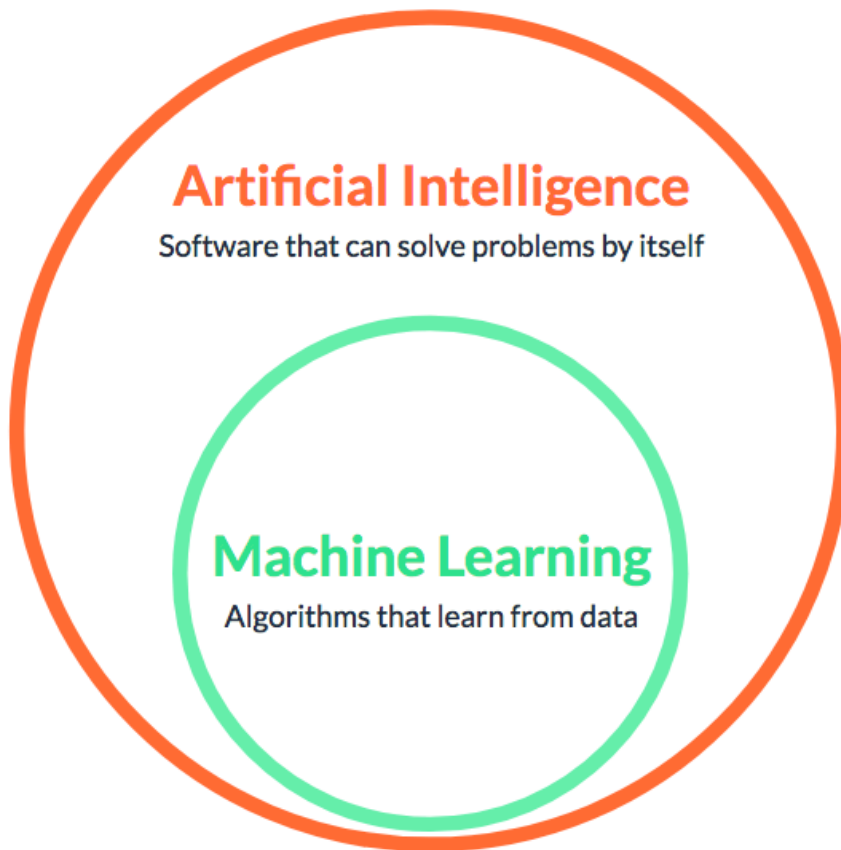
Figure 2.1: Artificial intelligence and machine learning

1. Obtaining and cleaning data.

2. Implementation of the model.

3. learning phase.

4. validation phase.

5. execution phase.

## 2.4 Machine learning types

There are several types of learning systems, and it varies depending on the type of problem you have. It is then useful to classify them into different categories.Machine learning systems can be classified according to the importance and nature of the supervision they require during the training phase. There are four main categories:

- Supervised Learning

- Unsupervised Learning

- Reinforcement Learning
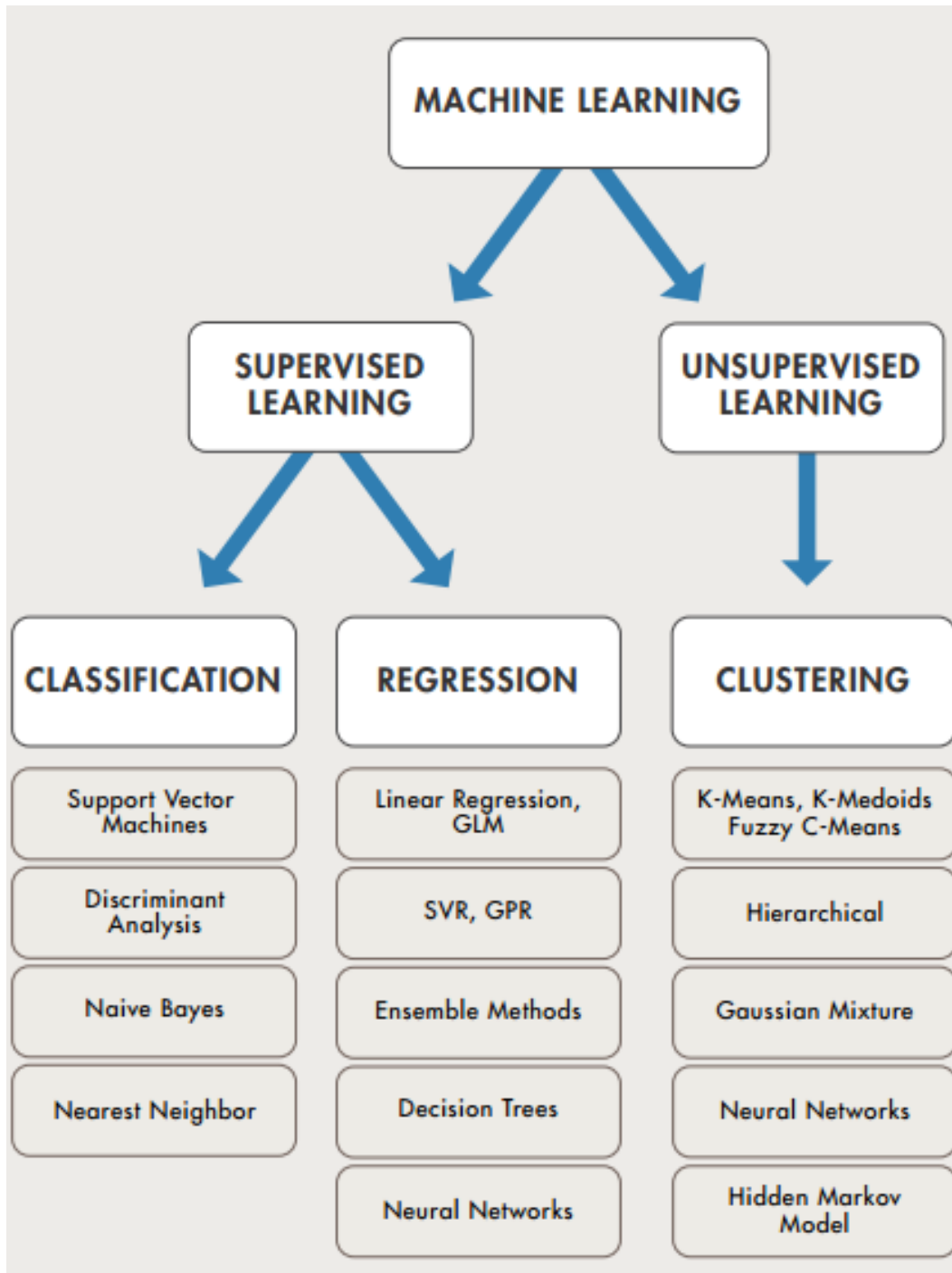


Figure 2.2: Machine learning types

## 2.5 Machine learning types

### 2.5.1 Supervised Learning

Supervised Machine Learning develops a model that makes predictions based on evidence in case of uncertainty. A supervised learning algorithm applies a known set of inputs and known responses to data (results) and results in a model that produces reasonable predictions for responses to new data. Supervised learning is used if there is known data for the outcomes to be predicted [4]. It develops predictive models using classification and regression techniques. Classification techniques provide for discrete vari-

ables [25]. For example, they determine whether an email is genuine or spam, or whether a tumour is cancerous or benign. Classification models classify input data into categories. Medical imaging, voice recognition and credit reporting are standard applications. Classification is used if the data can be marked, categorized or divided into specific groups or classes. For

example, handwriting recognition applications use classification to recognize letters and numbers. In computer image processing and vision, unsupervised pattern recognition techniques are used for object detection and image segmentation. Commonly used algorithms for classification include support vector machines (SVM), boosting/bagging of decision trees, the nearest neighbouring k method, naive Bayesian classification, discriminant analysis, logistic regression and neural networks. Regression techniques

predict continuous variables, such as temperature variations or fluctuations in energy demand. Power load forecasting and algorithmic trading are classic applications. Regression techniques are preferred if a data range is to be used or if the answers to be predicted are real numbers, such as temperature or time to next equipment failure. Common regression algo-

rithms include linear motion, non-linear model, regularization, step-by-step regression, decision tree boosting/bagging, neural networks, and adaptive Nero-blur learning.
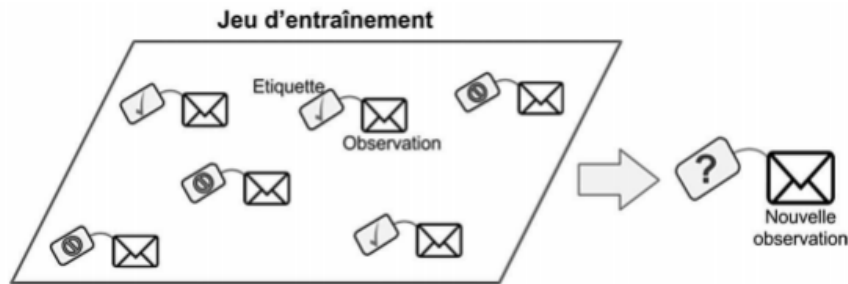
Figure 2.3: Supervised Learning

## 2.5.2  Unsupervised Learning

Unsupervised learning identifies hidden patterns or intrinsic structures in the data[5]. It is used to draw conclusions from data sets composed of input data without labeled responses . Clustering is the most common unsupervised learning technique. It is used to conduct exploratory analysis of the data to find hidden models or groupings in the data . Examples of

cluster analysis applications are gnomic sequence analysis, market research and object recognition. [6] For example, if a mobile phone company wants to optimize the construction locations of its cell towers, it can estimate the number of user clusters using their antennas using Machine Learning. A phone can only communicate with one antenna at a time. Consequently, the team uses clustering algorithms to determine the best location of the relay tower, in order to optimize the quality of signal reception for user groups (or clusters).

Commonly used algorithms for clustering include k-means and k-medoids, hierarchical clustering, Gaussian mixture models, hidden Markov models, self-organizing maps, clustering c-means blurring and subtractive clustering .
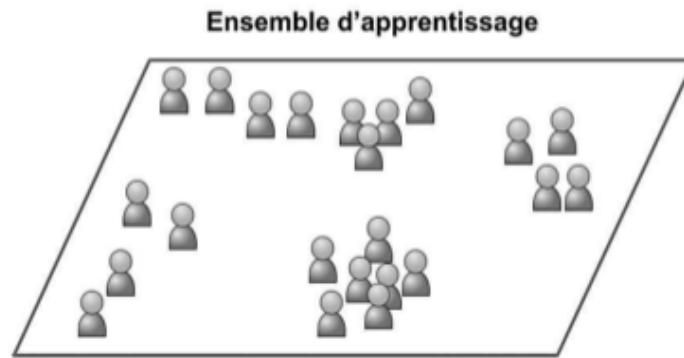
Figure 2.4: User group



Figure 2.5: clustering patterns

### 2.5.3 Reinforcement Learning

Stress learning is an approach to artificial intelligence that focuses on learning the system through its interactions with the environment[5]. With reinforcement learning, the system adapts its parameters according to the reactions received from the environment, which then provides feedback on the decisions taken. For example, a system that models a chess player who uses the results of previous steps to improve his performance, is a system that learns with reinforcement. Current research on enhanced learning is highly interdisciplinary and includes researchers specialising in genetic algorithms, neural networks, psychology and control techniques.

## 2.6 Machine learning algorithms
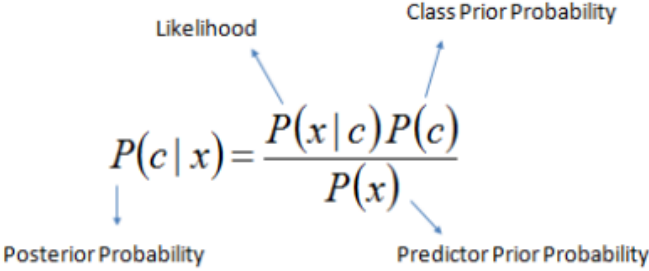
### 2.6.1 Naive Bayes classifier

One of the simplest and one of the most effective method for classification is the Naive Bayes Classifiers. The main idea on which this method works is on the Bayesian Networks which is a probabilistic graphical model signifying a set of random variables and their conditional Independence's. In the Bayesian Networks, there are several efficient algorithms that perform inference and learning. In basic requirement for this method to be applied is that the features of the data set should be different. On considering data set features different, the classification is done based on the Naïve bayes method. Naïve Bayes classifier is a simple probabilistic classifier based on applying Bayes' theorem with strong independence assumption [7]. The algorithm for Naïve Bayes classification is as follows:

- Step 1: The data set has to be first converted into a frequency table

- Step 2: Find the probabilities in order to obtain the likelihood table.

- Step 3: Then use the Naive Bayesian equation to calculate the posterior probability for each class. The class with the highest posterior probability is the outcome of prediction.

A Naive Bayesian model is easy to build, with no complicated iterative parameter estimation which makes it particularly useful in the field of medical science for diagnosing heart patients. Despite its simplicity, the Naive Bayesian classifier often does surprisingly well and is widely used because it often outperforms more sophisticated classification methods. Bayes theorem provides a way of calculating the posterior probability, P(c|x), from P(c), P(x), and P(x|c). Naive Bayes classifier assumes that the effect of the value of a predictor (x) on a given class (c) is independent of the values of other predictors. This assumption is called class conditional independence.

- P(c|x) is the posterior probability of class (target) given predictor (attribute).

- P(c) is the prior probability of class.

- P(x|c) is the likelihood which is the probability of predictor given class.

- P(x) is the prior probability of predictor

Likelihood

Class Prior Probability

$$P(c \mid x) = \frac{P(x \mid c) P(c)}{P(x)}$$

Posterior Probability

Predictor Prior Probability

$$P(c \mid X) = P(x_1 \mid c) \times P(x_2 \mid c) \times \cdots \times P(x_n \mid c) \times P(c)$$

Figure 2.6: Naive Bayes classifier

Where C and X are two events (e.g. the probability that the train will arrive on time given that the weather is rainy). Such Naïve Bayes classifiers use the probability theory to find the most likely classification of an unseen (unclassified) instance . The algorithm performs positively with categorical data but poorly if we have numerical data in the training set.

## 2.6.2 Advantages and limitations of the Naive Bayes Classifier

**Advantages:**

- the Naive Bayes Classifier is very fast for classification: indeed the probability calculations are not very expensive.

- Classification is possible even with a small data set.

**Disadvantages:**

- the Naive Bayes Classifier algorithm assumes the independence of the variables: This is a strong hypothesis and is violated in most real cases [7] however, it is very sensitive to their correlation[8].

### 2.6.3 Decision tree classifier

Decision trees are one of the powerful methods commonly used in various fields, such as machine learning, image processing, and identification of patterns [20].

This technique divides the classification problem into sub-problems. It creates a decision tree that is used to develop a model used for classification purposes. Classifying is relatively simple. Simply answer the questions by going down the decision tree [21].

Decision trees are non-parametric models and find rules in general quite powerful. They can process very large data sets and they can also use mixed preachers (categorical and numbers). Redundant variables are eliminated because the tree does not even take them into account [22]. At each node, we must therefore find two things: which feature variable to use and what is the point of separation of the two zones, so as to minimize the quadratic error for the regression and the impurity for the classification. We grow the tree as big as possible21.
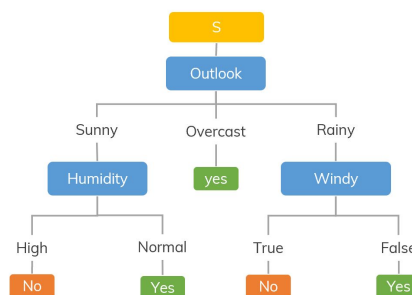


Figure 2.7: Decision tree classifier

### 2.6.4 Advantages and limitations of decision Classifier

**Advantages:**

- They are easy to understand.

- They allow you to select the most appropriate option among several.

- Reasonable turnaround time.

- They are easy to associate with other decision-making tools.

**Disadvantages:**

- The learning of the optimal decision tree is NP- complete concerning several aspects of optimality.

- Decision tree learning can lead to very complex decision trees that poorly generalize the learning package

### 2.6.5 Random forest classifier

Random Forest Classifier is a supervised learning algorithm composed of several decision trees. By averaging the impact of several decision trees, random forests tend to improve prediction [23]. It works in four steps:

- Select random samples from a given data set.

- Build a decision tree for each sample and get a result of prediction of each decision tree.

- Vote on each outcome.

- Select the prediction result with the most votes as the final prediction.
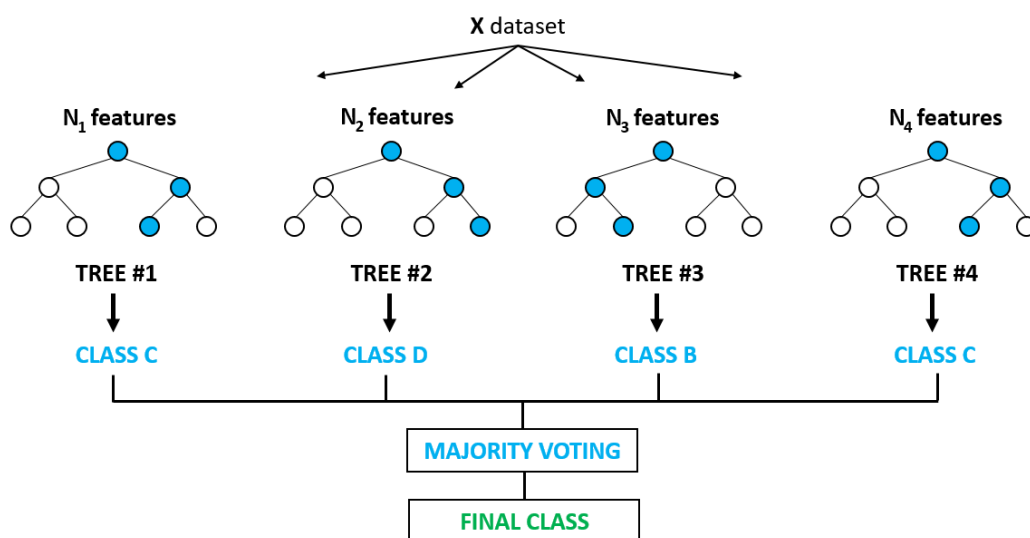


Figure 2.8: Random forest classifier

### 2.6.6 Advantages and limitations of Random forest Classifier

**Advantages:**

- a very accurate and robust method due to the number of decision trees involved in the process.

- The algorithm can be used in classification and regression problems.

**Disadvantages:**

- The model is difficult to interpret in relation to a decision tree, where one can easily make a decision by following the path in the tree [22].

### 2.6.7 Neural network

An artificial neural network (ANN) is a machine learning algorithm inspired by biological neural networks. Each ANN contains nodes (analogous to cell bodies) that communicate with other nodes via connections (analogous to axons and dendrites)[25]. in the way synapses between neurons are strengthened when their neurons have correlated outputs in a biological neural network (the Hebbian theory postulates that "nerves that fire together, wire together"), connections between nodes in an ANN are weighted based upon their ability to provide a desired outcome.[26]



Figure 2.9: Neural network

## 2.6.8 Advantages and limitations of Neural networks

**Advantages:**

- Ability to represent any function, linear or not, simple or complex

- Resistance to noise or unreliable data

**Disadvantages:**
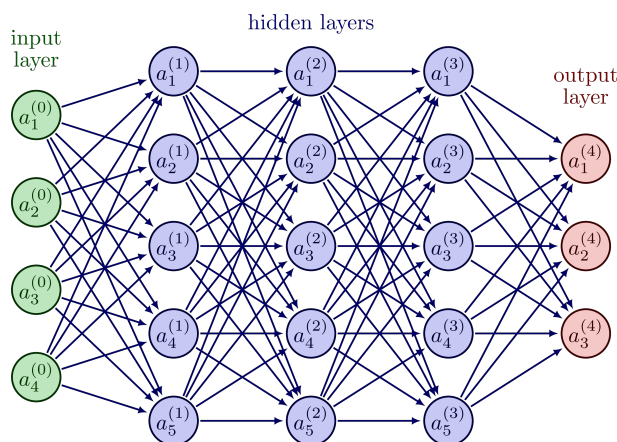
- Lack of a systematic method to define the best network topology and the number of neurons to be placed in the hidden layer(s).

- The choice of initial values of the network weights and the adjustment of the learning step play an important role in the speed of convergence [22].

## 2.6.9 Logistic regression

Logistic regression is a classification algorithm where the goal is to find a relationship between features and the probability of a particular outcome.

Rather than using the straight line produced by linear regression to estimate class probability, logistic regression uses a sigmoidal curve to estimate class probability (Fig:2.10) [26].

This curve is determined by the sigmoid function, y=1/1+e−x, which produces an S-shaped curve that converts discrete or continuous numeric features (x) into a single numerical value (y) between 0 and 1 [26].

The major advantage of this method is that probabilities are bounded between 0 and 1 (i.e., probabilities cannot be negative or greater than 1). It can be either binomial, where there are only two possible outcomes, or multinomial, where there can be three or more possible outcomes [25]
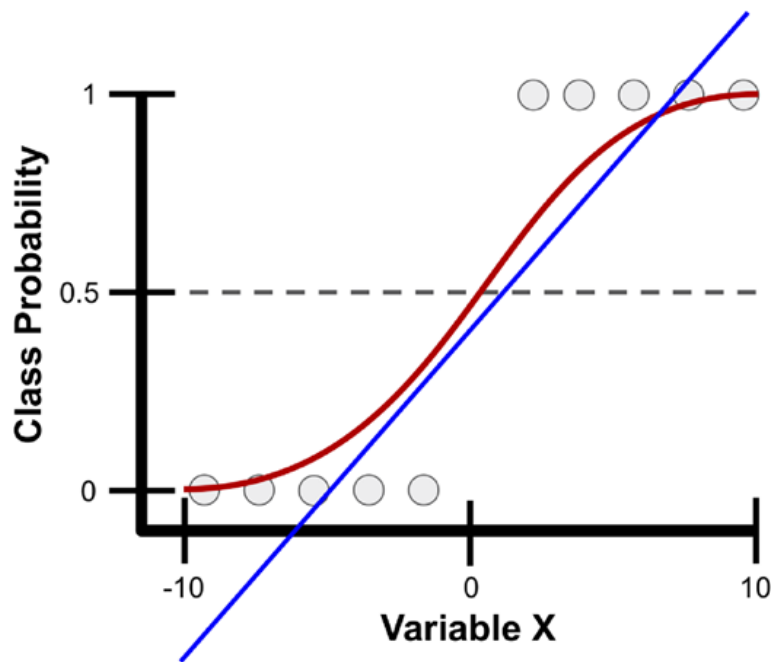
Figure 2.10: Neural network

### 2.6.10 Advantages and limitations of Logistic regression

**Advantages:**

- Logistic regression is easier to implement, interpret, and very efficient to train.

- It makes no assumptions about distributions of classes in feature space.

**Disadvantages:**

- If the number of observations is lesser than the number of features, Logistic Regression should not be used, otherwise, it may lead to over fitting.

- The major limitation of Logistic Regression is the assumption of linearity between the dependent variable and the independent variables. [24].

## 2.7   Some examples of AI uses

- Making a diagnosis: For example medical diagnosis (subject of our brief). AI can make a medical diagnosis more accurately than a human.

- Facial recognition: Allows to recognize the features of an individual in the street, in institutions, anywhere with a camera or on any photo.

- Smartphones: Many of these apps are based on AI. Integrated smart phones like Siri, Alexa and Google Assistant are the most obvious examples of AI that most of us know and use.

- Social Networks: Moving from apps on your smartphones to the presence of AI in various social media applications. Whether it's your Facebook, Twitter, Instagram or other platforms, what you see and do on these apps is to a large extent influenced by machine learning. AI has a lot of control over the flows you see when you browse these platforms or the notifications you receive.

- Smart Cars: Smart cars are another area where artificial intelligence increases its presence in our daily lives.

## 2.8   Conclusion

This chapter showed us the definitions we need to know in the AI, also the different algorithms that could be used in our model, finally the uses of this algorithms in the real life.

# Chapter 3

# Smart health prediction system

## 3.1 Introduction

As we already mentioned in Chapter two that machine learning is included almost everywhere, in this chapter we followed the steps of a machine learning project to predict diseases according to a data set of symptoms and got fine results.

## 3.2 Design of our solutions

As our solution about the prediction of diseases based on machine learning, we must first go through the machine learning stages in order to get our final results.

We used Microsoft azure ML studio which allows us to author new models and store our compute targets, models, deployments, metrics, and run histories in the cloud. operation of the proposed health prediction system is illustrated in a schema shown in the following figure:
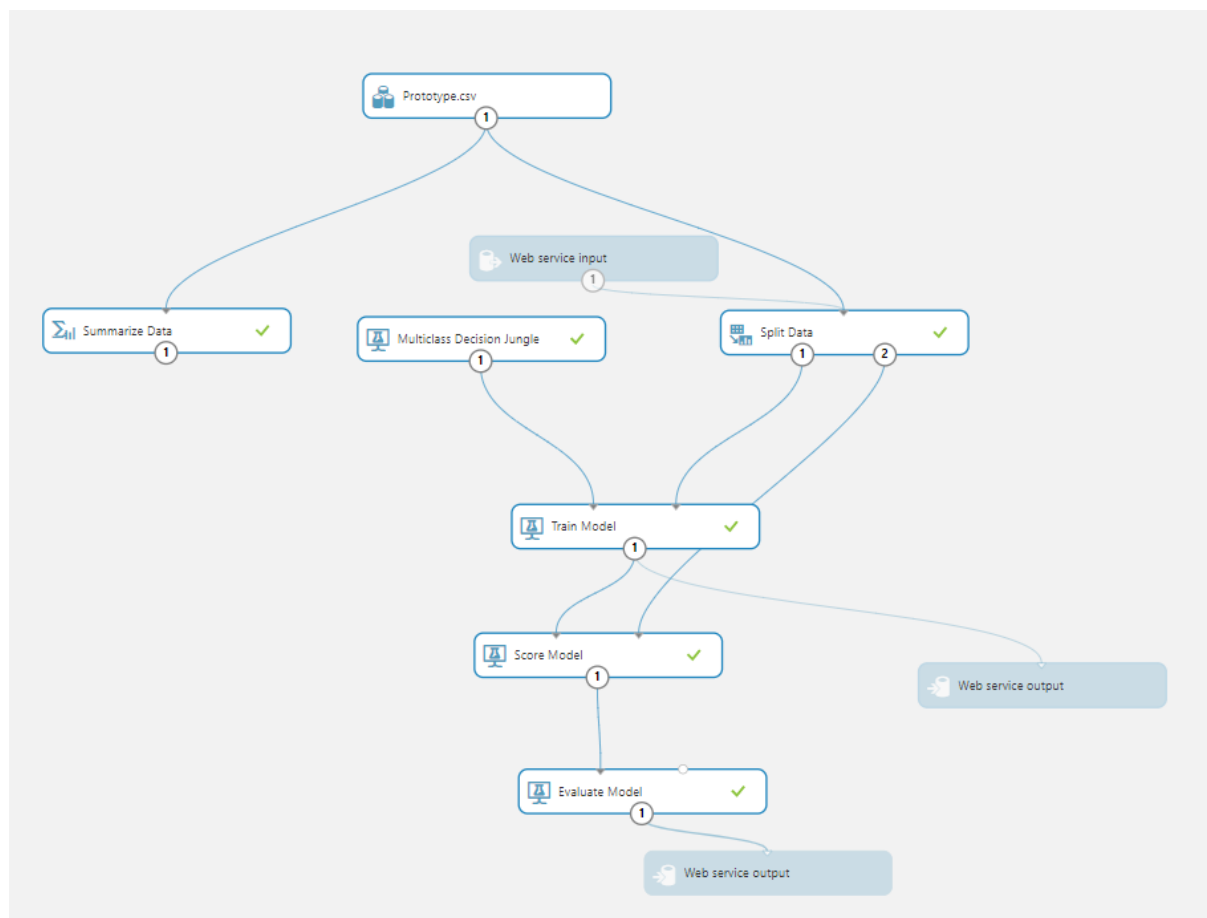


Figure 3.1: The general model of the proposed health prediction system

## 3.3 Data set choice

In this study, the data set used was obtained from Kaggle. Kaggle is an online community of data scientists and machine learners acquired by the Google in March of 2017. This data set contains 132 different symptoms and 492 different diseases, and it is a unique data set because this kind of project is rare, generally developers use small data sets (heart disease, covid 19) but we wanted to be inclusive, so we hardly found this one, and here are some statistics about the data set we used taken from the Kaggle site showing its quality:
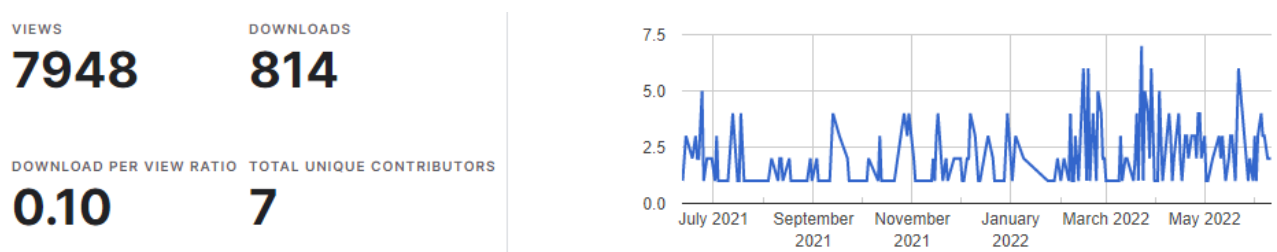


Figure 3.2: Statistics about the uses of the data set

## 3.4 visualization and analysing the data set

In this step we have done a statistical study on our data set to provide insights both to others and to ourselves, calculating the number of columns, rows and some statistics such us unique values, missing values and features type, and finally a visualization using a histogram.
The metrics provided are the ROC curve, precision-recall diagram, and lift curve. Look at these metrics to see that the first model performed better than the second one. To look at what the first model predicted, right-click on the Score Model component and click Visualize Scored dataset to see the predicted results.
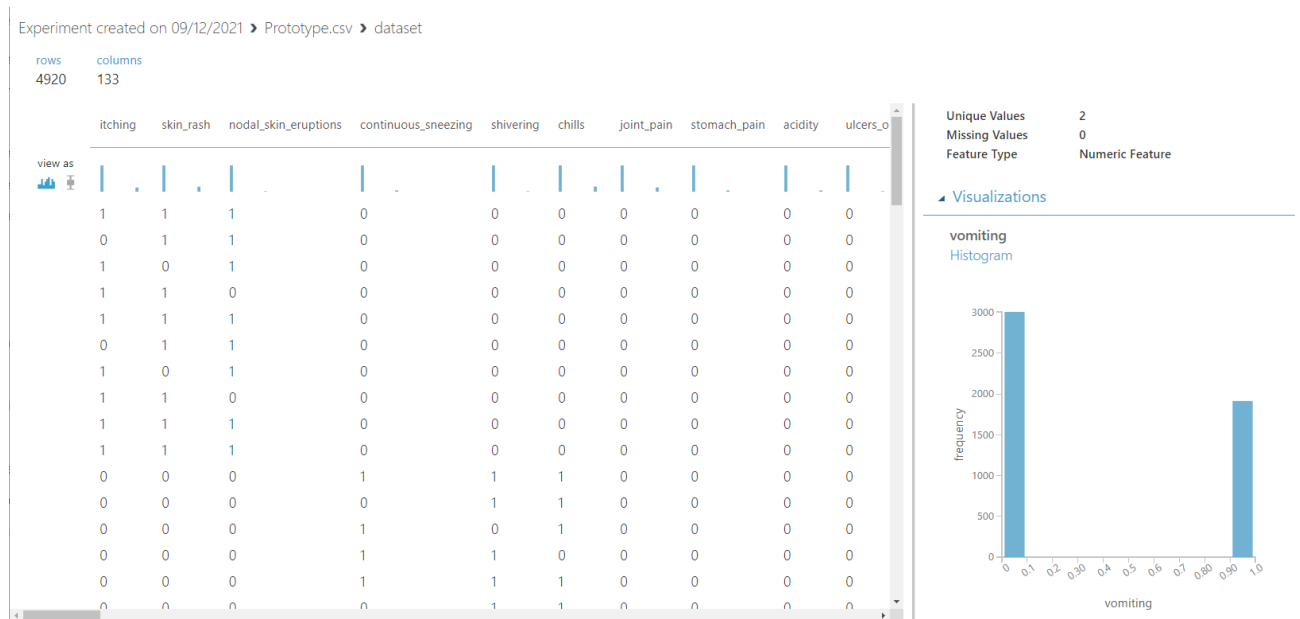
Figure 3.3: Data visualization

## Statistics

| | |
|---|---|
| Mean | 0.389 |
| Median | 0 |
| Min | 0 |
| Max | 1 |
| Standard Deviation | 0.4876 |
| Unique Values | 2 |
| Missing Values | 0 |
| Feature Type | Numeric Feature |

Figure 3.4: Data visualization statistics
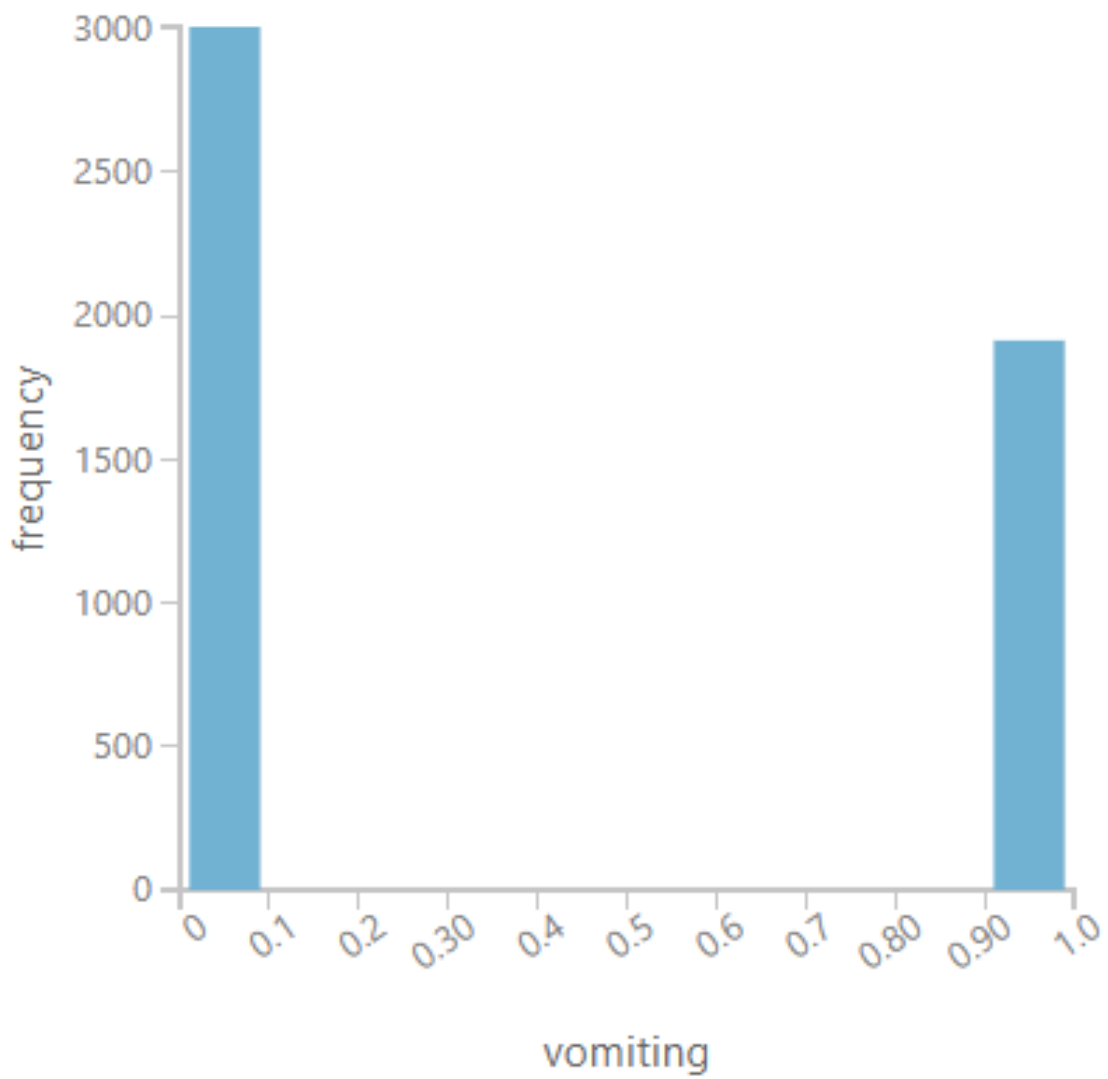
**vomiting**
Histogram



Figure 3.5: Visualization histogram

## 3.5   Summarize Data:

In this second step of our model which considered one with its previous, because it's a more detailed statistic review of our data set.
Summarize Data component is used to create a set of standard statistical measures that describe each column in the input table.
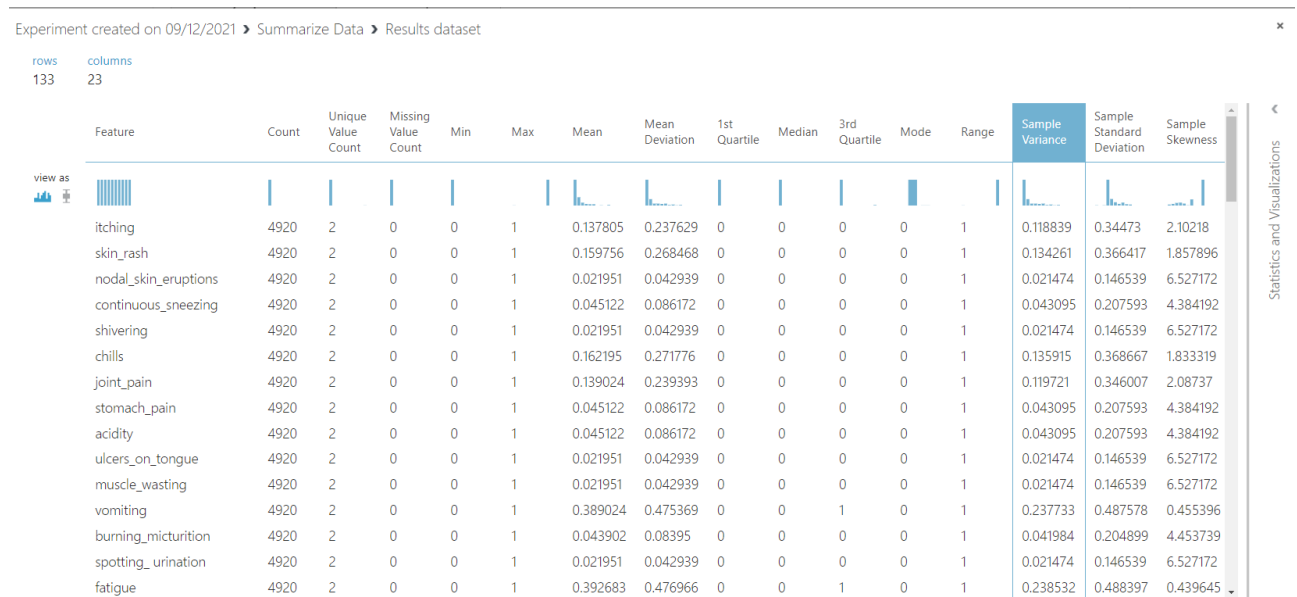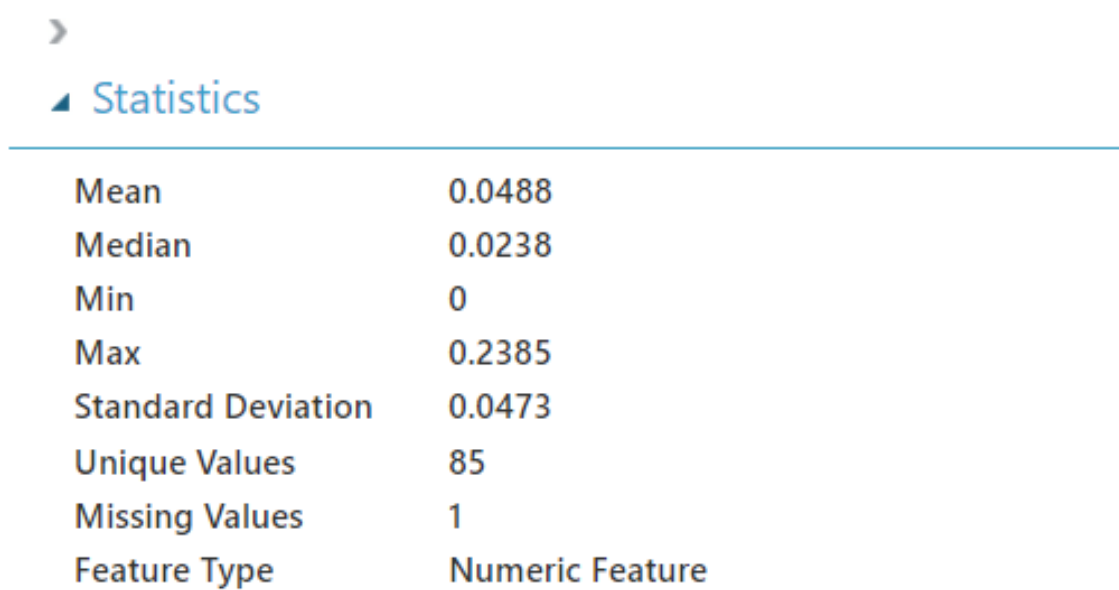
Experiment created on 09/12/2021 ❯ Summarize Data ❯ Results dataset

rows 133    columns 23

| Feature | Count | Unique Value Count | Missing Value Count | Min | Max | Mean | Mean Deviation | 1st Quartile | Median | 3rd Quartile | Mode | Range | Sample Variance | Sample Standard Deviation | Sample Skewness |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| itching | 4920 | 2 | 0 | 0 | 1 | 0.137805 | 0.237629 | 0 | 0 | 0 | 0 | 1 | 0.118839 | 0.34473 | 2.10218 |
| skin_rash | 4920 | 2 | 0 | 0 | 1 | 0.159756 | 0.268468 | 0 | 0 | 0 | 0 | 1 | 0.134261 | 0.366417 | 1.857896 |
| nodal_skin_eruptions | 4920 | 2 | 0 | 0 | 1 | 0.021951 | 0.042939 | 0 | 0 | 0 | 0 | 1 | 0.021474 | 0.146539 | 6.527172 |
| continuous_sneezing | 4920 | 2 | 0 | 0 | 1 | 0.045122 | 0.086172 | 0 | 0 | 0 | 0 | 1 | 0.043095 | 0.207593 | 4.384192 |
| shivering | 4920 | 2 | 0 | 0 | 1 | 0.021951 | 0.042939 | 0 | 0 | 0 | 0 | 1 | 0.021474 | 0.146539 | 6.527172 |
| chills | 4920 | 2 | 0 | 0 | 1 | 0.162195 | 0.271776 | 0 | 0 | 0 | 0 | 1 | 0.135915 | 0.368667 | 1.833319 |
| joint_pain | 4920 | 2 | 0 | 0 | 1 | 0.139024 | 0.239393 | 0 | 0 | 0 | 0 | 1 | 0.119721 | 0.346007 | 2.08737 |
| stomach_pain | 4920 | 2 | 0 | 0 | 1 | 0.045122 | 0.086172 | 0 | 0 | 0 | 0 | 1 | 0.043095 | 0.207593 | 4.384192 |
| acidity | 4920 | 2 | 0 | 0 | 1 | 0.045122 | 0.086172 | 0 | 0 | 0 | 0 | 1 | 0.043095 | 0.207593 | 4.384192 |
| ulcers_on_tongue | 4920 | 2 | 0 | 0 | 1 | 0.021951 | 0.042939 | 0 | 0 | 0 | 0 | 1 | 0.021474 | 0.146539 | 6.527172 |
| muscle_wasting | 4920 | 2 | 0 | 0 | 1 | 0.021951 | 0.042939 | 0 | 0 | 0 | 0 | 1 | 0.021474 | 0.146539 | 6.527172 |
| vomiting | 4920 | 2 | 0 | 0 | 1 | 0.389024 | 0.475369 | 0 | 0 | 1 | 0 | 1 | 0.237733 | 0.487578 | 0.455396 |
| burning_micturition | 4920 | 2 | 0 | 0 | 1 | 0.043902 | 0.08395 | 0 | 0 | 0 | 0 | 1 | 0.041984 | 0.204899 | 4.453739 |
| spotting_ urination | 4920 | 2 | 0 | 0 | 1 | 0.021951 | 0.042939 | 0 | 0 | 0 | 0 | 1 | 0.021474 | 0.146539 | 6.527172 |
| fatigue | 4920 | 2 | 0 | 0 | 1 | 0.392683 | 0.476966 | 0 | 0 | 1 | 0 | 1 | 0.238532 | 0.488397 | 0.439645 |

Figure 3.6: Summarize data

## Statistics

| | |
|---|---|
| Mean | 0.0488 |
| Median | 0.0238 |
| Min | 0 |
| Max | 0.2385 |
| Standard Deviation | 0.0473 |
| Unique Values | 85 |
| Missing Values | 1 |
| Feature Type | Numeric Feature |

Figure 3.7: Summarize data statistics

Missing Values          0
Feature Type            Numeric Feature

◢ Visualizations
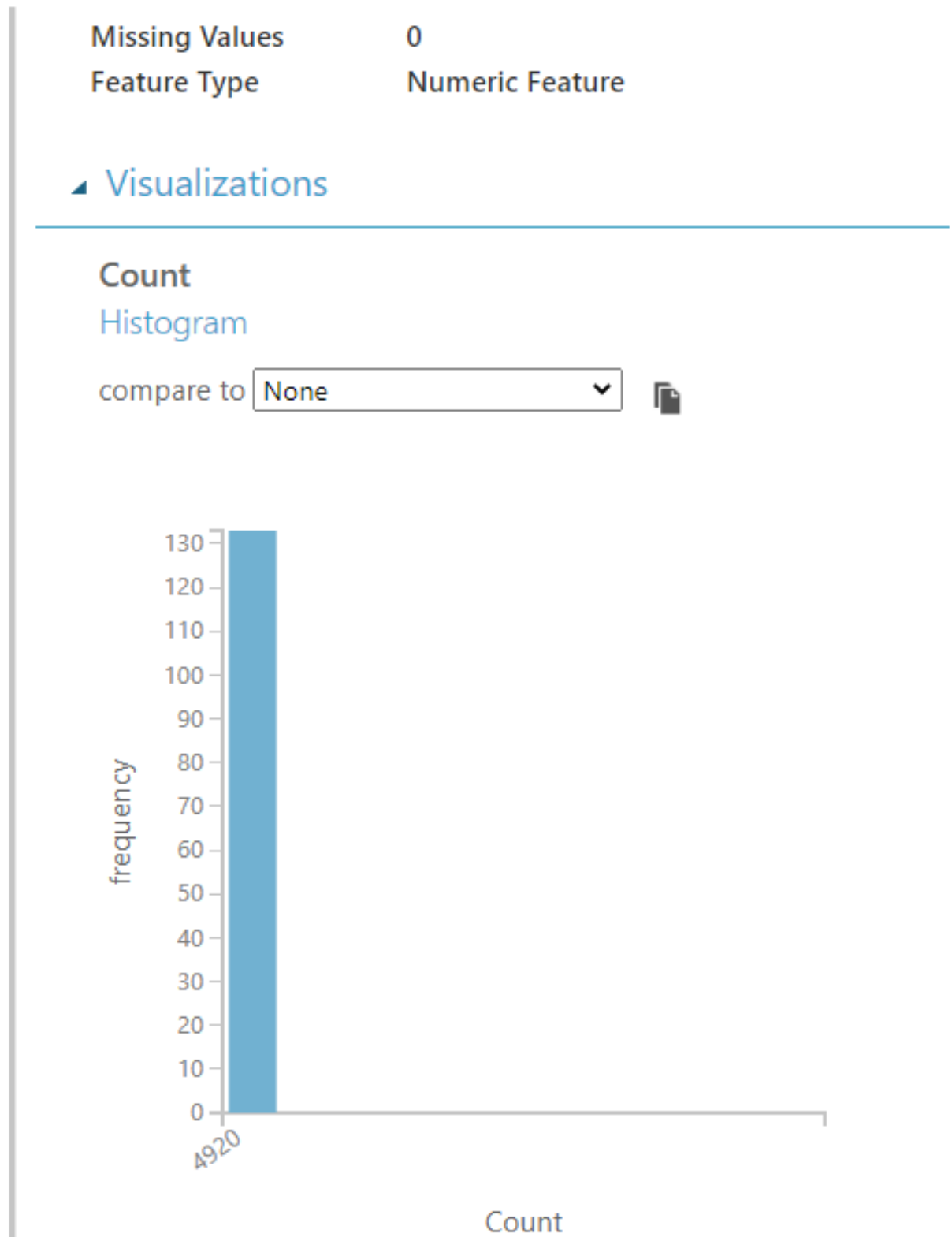
Count
Histogram

compare to  None  ⌄  ▣



Figure 3.8: Summarize histogram

## 3.6 Split data

In the following step of our model we'll be splitting the data into 70% training and 30% testing, will need both of these fractions in further steps, we'll define some characteristics such as the split mode, the class of classification of the model which is called "prognoses".

we can also customize the way that data is divided. Some options support randomization of data. Others are tailored for a certain data type or model type..

Splitting mode: we need to Choose one of the following modes, depending on the type of data we have and how we want to divide it. Each splitting mode has different options.

- **Split Rows:**Use this option if you just want to divide the data into two parts. You can specify the percentage of data to put in each split. By default, the data is divided 50/50.
  we can also randomize the selection of rows in each group, and use stratified sampling. In stratified sampling, we must select a single column of data for which we want values to be apportioned equally among the two result datasets.

- **Regular Expression Split:**we Choose this option when we want to divide our dataset by testing a single column for a value.
  For example, if we're analyzing sentiment, we can check for the presence of a particular product name in a text field. we can then divide the dataset into rows with the target product name and rows without the target product name.

- **Relative Expression Split:** we Use this option whenever we want to apply a condition to a number column. The number can be a date/time field, a column that contains age or dollar amounts, or even a percentage. For example, we might want to divide our dataset based on the cost of the items, group people by age ranges, or separate data by a calendar date.

Figure 3.9: Split characteristics and information

## 3.7  Train Model:

In this step we use this component to train a classification or regression model. Training takes place after we have defined a model and set its parameters, and requires tagged data. we can also use Train Model to retrain an existing model with new data.
In our model we'll train the Model with Multi class Decision algorithm and the 70% result of the split phase, It consists of the sample output data and the corresponding sets of input data that have an influence on the output. The training model is used to run the input data through the algorithm to correlate the processed output against the sample output.

## 3.8  Evaluate model:

the process of using different evaluation metrics to understand the model's performance, as well as its strengths and weaknesses.
We use this component to measure the accuracy of a trained model. we provide a dataset containing scores generated from a model, and the Evaluate Model component computes a set of industry-standard evaluation metrics.

The metrics returned by Evaluate Model depend on the type of model that you are evaluating:

- Classification Models

- Regression Models

- Clustering Models

we've got an excellent result shown in the following figures:
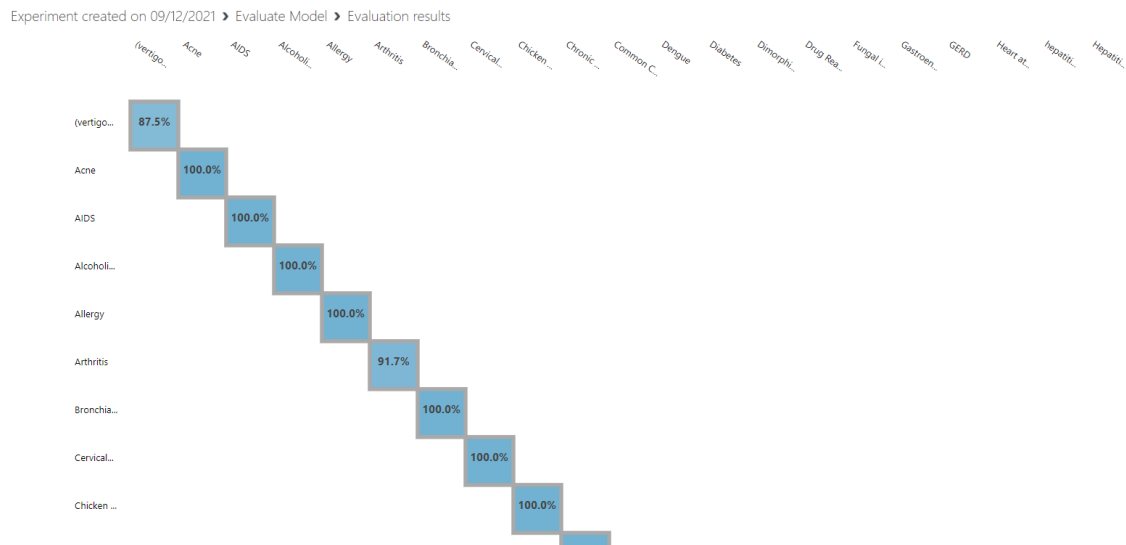
Figure 3.10: Evaluation matrix

Experiment created on 09/12/2021 ❯ Evaluate Model ❯ Evaluation results

▲ Metrics

| | |
|---|---|
| Overall accuracy | 0.97832 |
| Average accuracy | 0.998942 |
| Micro-averaged precision | 0.97832 |
| Macro-averaged precision | 0.979822 |
| Micro-averaged recall | 0.97832 |
| Macro-averaged recall | 0.97832 |

▶ Confusion Matrix

Figure 3.11: Model's evaluation

## 3.9   Conclusion

In this chapter we manged to explain the model that we've created, the deployment method and all the statistics and the results of our model.

# Chapter 4

# Capturing requirements

## 4.1 Introduction

After defining the functional requirements necessary for the operation of our application, in this chapter we will present and detail its needs by use cases which allow us to clearly specify the various tasks that must be performed by each actor in the system, and to complete the process we will also describe the technical needs .

## 4.2 Stage status in the 2TUP process

The second step of the process is Capturing Needs, the preliminary studies goes through two branches on the one hand a left branch presents Capturing functional needs and on the other hand a right branch presents Capturing technical needs
the following figure represents the stage status in the 2TUP process:
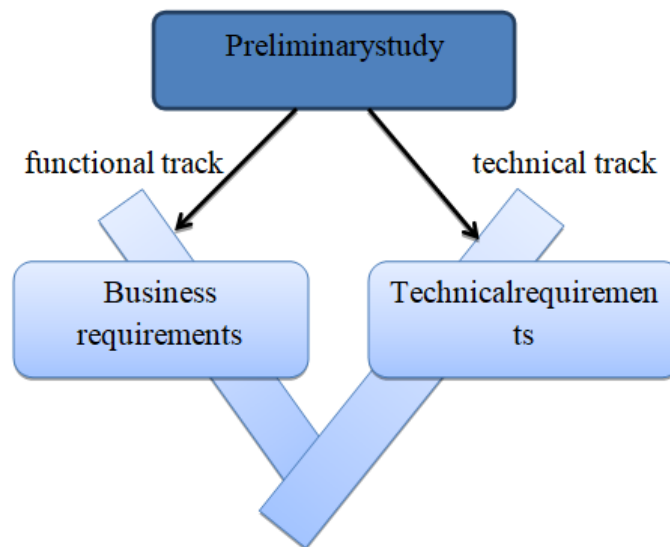


Figure 4.1: The stage status in the 2TUP process

## 4.3 Capturing functional requirements

### 4.3.1 Introduction

A functional point of view of the system architecture. Through the use cases, we will be in constant contact with the actors of the system in order to define the limits of the system, and thus avoid getting too far from the real needs of the end user. [14].

### 4.3.2 Identification of use cases

The Use Case Diagram is an UML diagram used to give a global view of the functional Specification of a system. it consists of a set of tasks known by the name of use cases, each of which represents an activity carried out by actor [9].This diagram shows the functional interactions between the actors and the system under study. It is one of the most structuring diagrams in the analysis of a system [10].

### 4.3.3 use case diagram
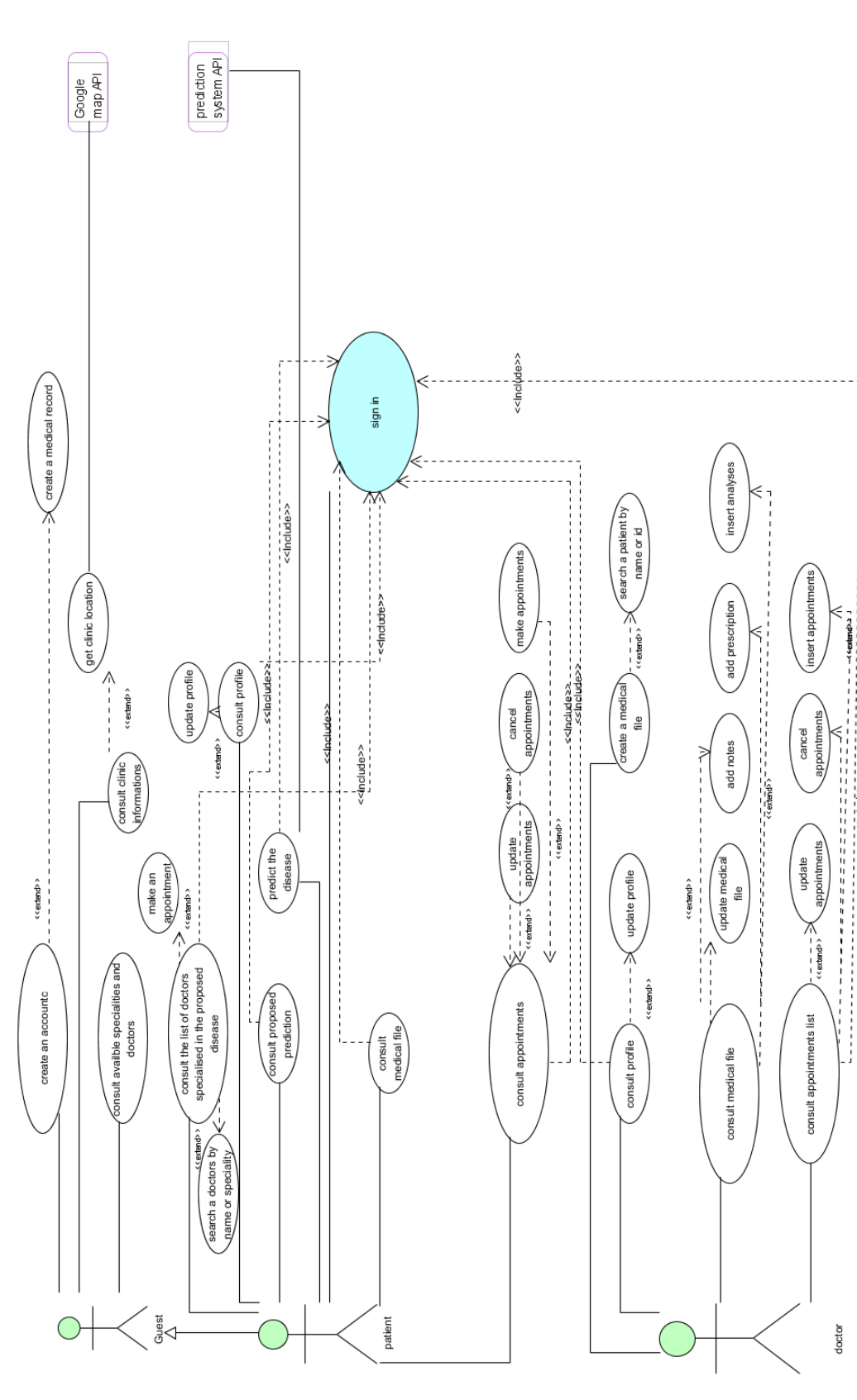


Figure 4.2: Use case diagram part 1
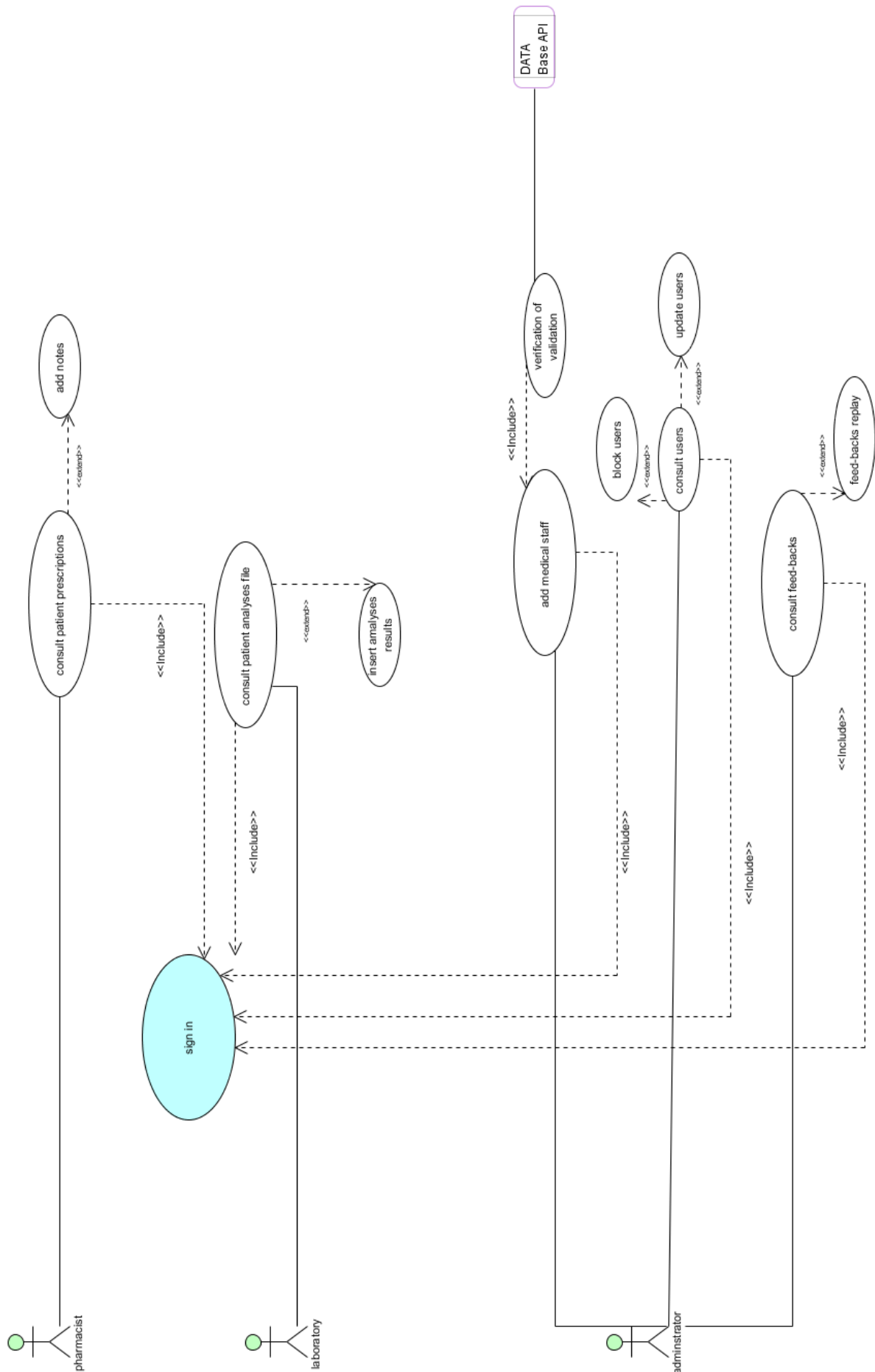
Figure 4.3: Use case diagram part 2

### 4.3.4   Use cases description

**Textual description**

Each use case must be associated with a textual description of the interactions between the actor and the system and the actions that the system must perform in order to produce the results expected by the actors.

**Graphical description**

For a better presentation of use cases, the sequence diagram is the best choice. A sequence diagram is a schematic illustration of the scenarios specified in the text description [13].

Guests use cases description: The tables and figures coming represents a textual description for all use cases of the actor "Guest".

| Use case | Create an account |
|---|---|
| **primary actor** | Guest |
| **secondary actor** | / |
| **Description** | Allows the guest to create a new account and create or add a medical record |
| **Precondition** | / |
| **Post-condition** | account creation is successfully completed (new profile created) |
| **primary scenario** | 1- The guest demands the creation of a new account<br>2- The system returns the sign in form<br>3- The guest fills the form and validate<br>4- The system verify and saves all data entered and displays a message "your account has been created"<br>5- The guest demands the creation of his medical record<br>6- The system returns the form of creating a new medical record<br>7- The guest fills the form and validate<br>8- The system verify and saves all data entered and displays a message "your medical record has been created" |
| **secondary scenario** | 4-1- The entered data is incorrect or incomplete, the system displays a message "check your information" and returns to step 2 |

Table 4.1: textual description: create account.
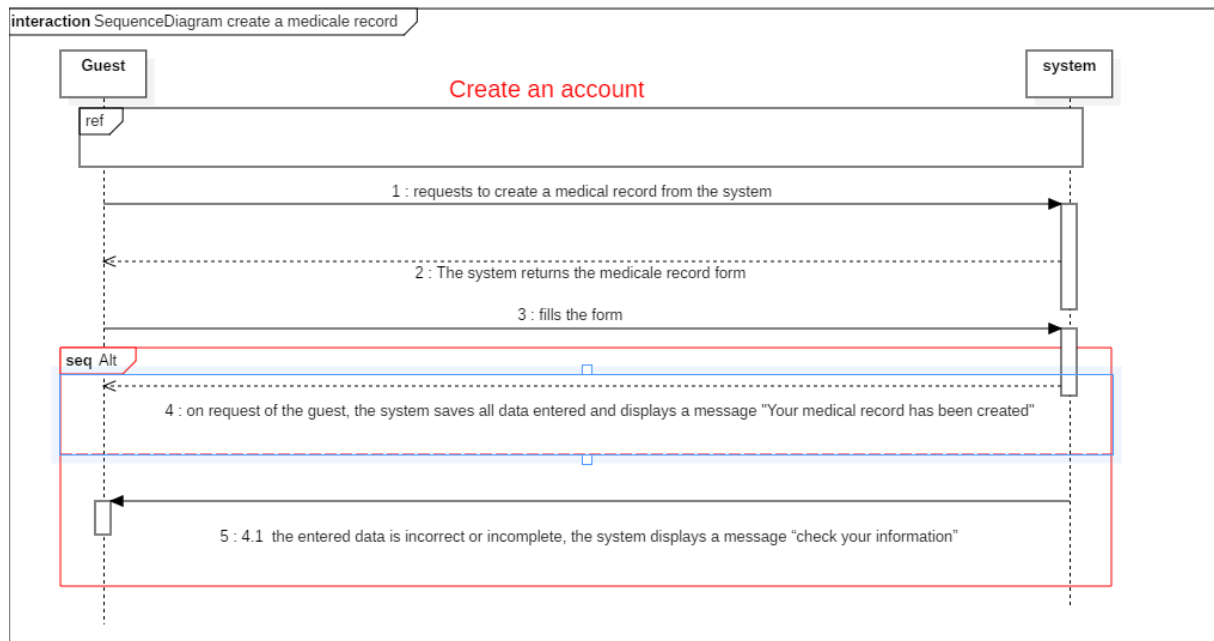
Figure 4.4: Create a medical record by guest

| Use case | consult available specialties and doctors |
|---|---|
| primary actor | Guest |
| secondary actor | / |
| Description | Allows the guest to see all the available specialities and doctors in the clinic |
| Precondition | / |
| Post-condition | available specialties and doctors list is displayed |
| primary scenario | 1- The guest requests available specialties and doctors information from the system<br>2- The system returns available specialties and doctors list |
| secondary scenario | / |

Table 4.2: textual description: consult available specialties and doctors

| Use case | consult clinics information |
|---|---|
| **primary actor** | Guest |
| **secondary actor** | Google maps |
| **Description** | Allows the guest to see details about the clinic |
| **Precondition** | / |
| **Post-condition** | Clinics information consulted / clinic location is displayed |
| **primary scenario** | 1. The guest requests clinics information from the system<br>2. The system returns the clinics information page |
| **secondary scenario** | 3.1.The guest requests the clinics location from the system<br>3.2.The system use Google Maps api to display the location |

Table 4.3: textual description: consult clinics information

Patients use cases description: The tables and figures coming represents a textual and graphical description for all use cases of the actor "Patient".

| Use case | consult profile |
|---|---|
| **primary actor** | Patient |
| **secondary actor** | / |
| **Description** | Allows the patient to consult and update his profile information |
| **Precondition** | The patient must be authenticated |
| **Post-condition** | The patient profile is displayed and/or updated |
| **primary scenario** | 1- The patient requests to access his profile<br>2- The system returns the patients profile |
| **secondary scenario** | 3-1- The patient requests profile update<br>3-2- The system returns the users information form<br>3-3- The patient fills the form and validate<br>3-4- The system saves the changes |

Table 4.4: textual description: consult profile

interaction SequenceDiagram sign in

actors                                                                System

1 : asks to sign in

2 : System returns login  interface

3 : fills the username and password

seq Alt

4 : System displays the sign in form

5 : the system  displays  message  " incorect password or username "
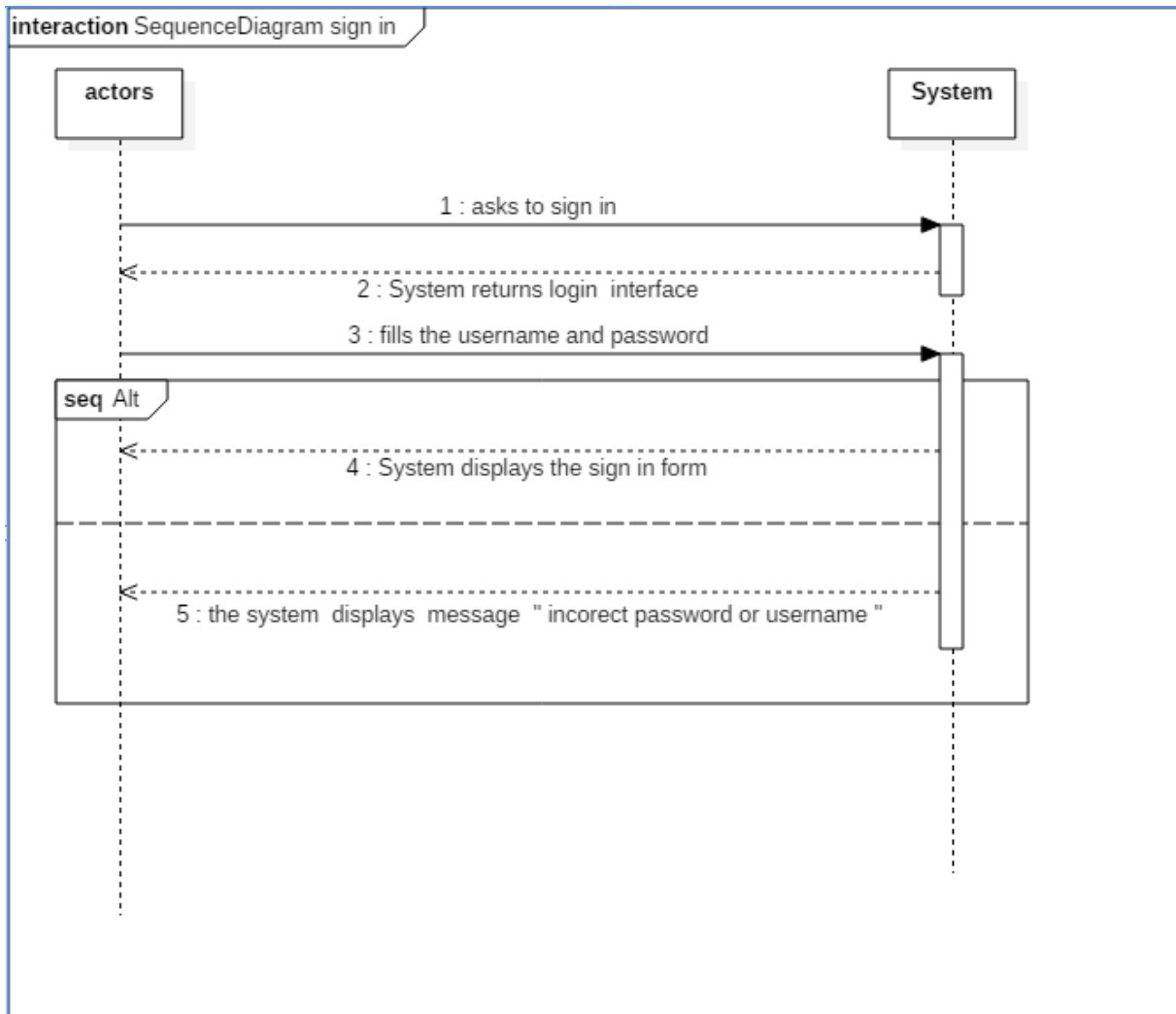
Figure 4.5: Sign in
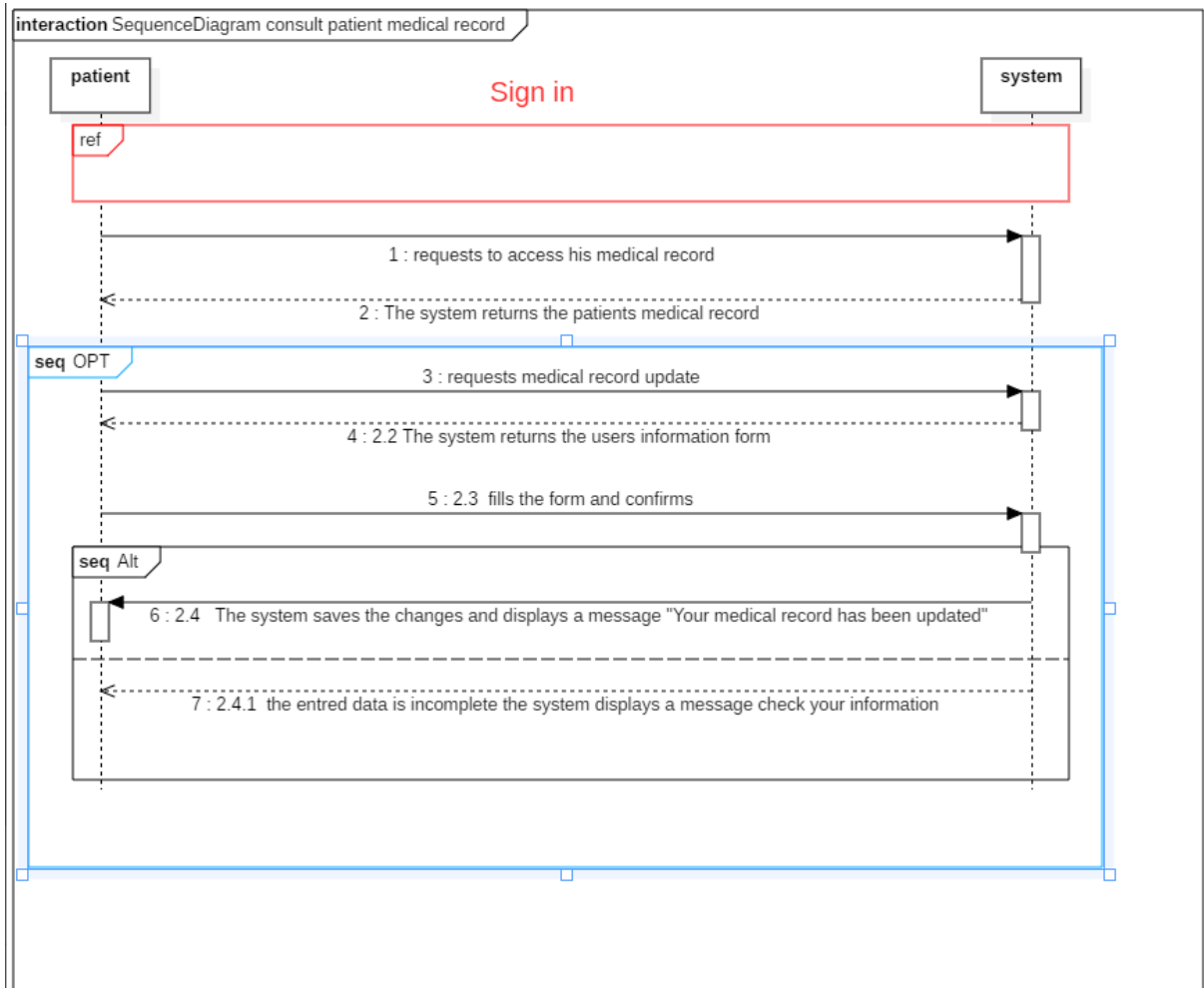
Figure 4.6: Consult medical record by patient

| Use case | Predict the disease |
|---|---|
| **primary actor** | Patient |
| **secondary actor** | Prediction system |
| **Description** | Allows the patient to predict the disease that he has |
| **Precondition** | The patient must be authenticated |
| **Post-condition** | the proposed prediction consulted |
| **primary scenario** | 1- The patient demands to predict his disease<br>2- The system returns the symptoms formthe prediction system<br>3- The patient fill the form and validate<br>4- The system prediction verify and display the predicted disease |
| **secondary scenario** | 4-1- The entered data is incorrect or incomplete , the system displays a message "verify your information" And returns to step 2 |

Table 4.5: textual description: Predict the disease

| Use case | Consult appointments |
|---|---|
| primary actor | Patient |
| secondary actor | / |
| Description | Allows the patient to consult or update appointments |
| Precondition | The patient must be authenticated |
| Post-condition | Appointments consulted or updated |
| primary scenario | 1- The patient requests the list of his appointments<br>2- The system returns the list of patients appointments |
| secondary scenario | 3.1 The patient selects an appointment to update<br>3.2 The system returns appointment update form<br>3.3 The patient fill the form and validate |

Table 4.6: textual description: Consult appointments

Doctors use cases description: The tables and figures coming represents a graphical description for all use cases of the actor "Doctor".
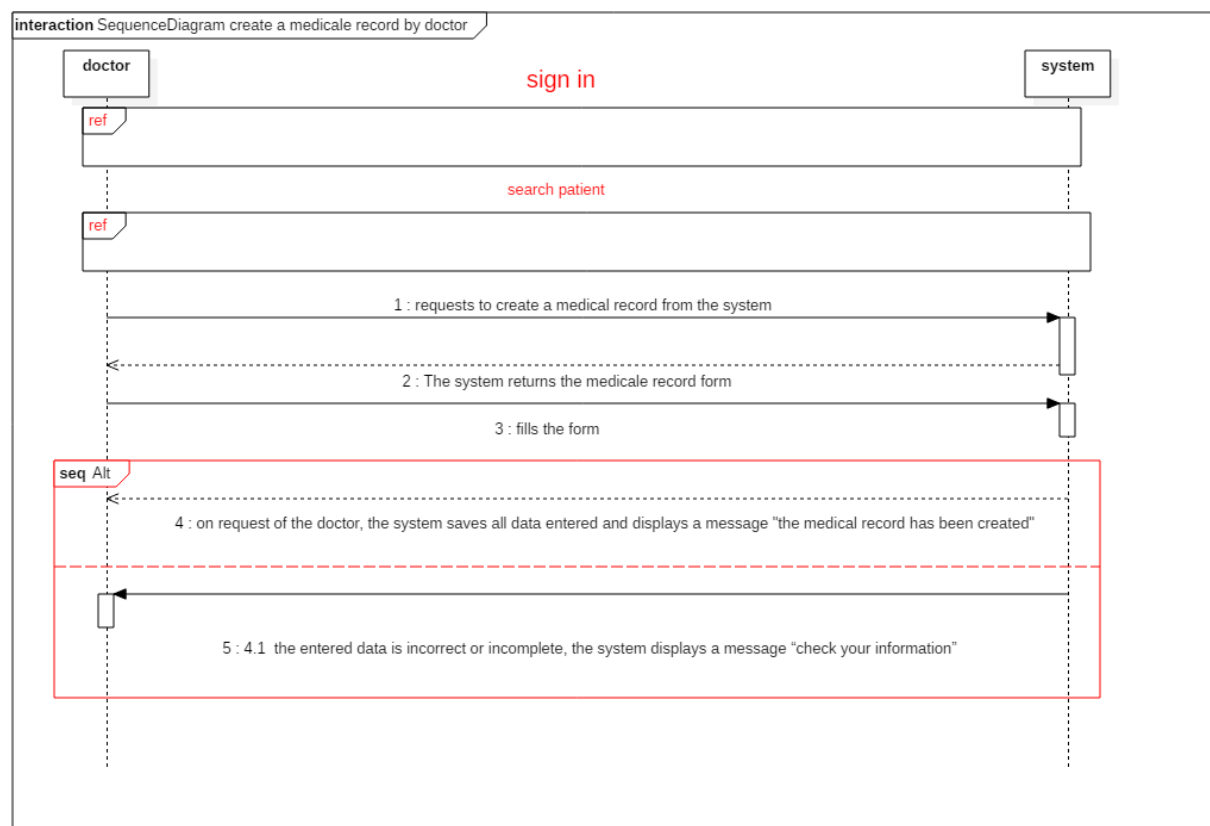


Figure 4.7: Create a medical record by doctor
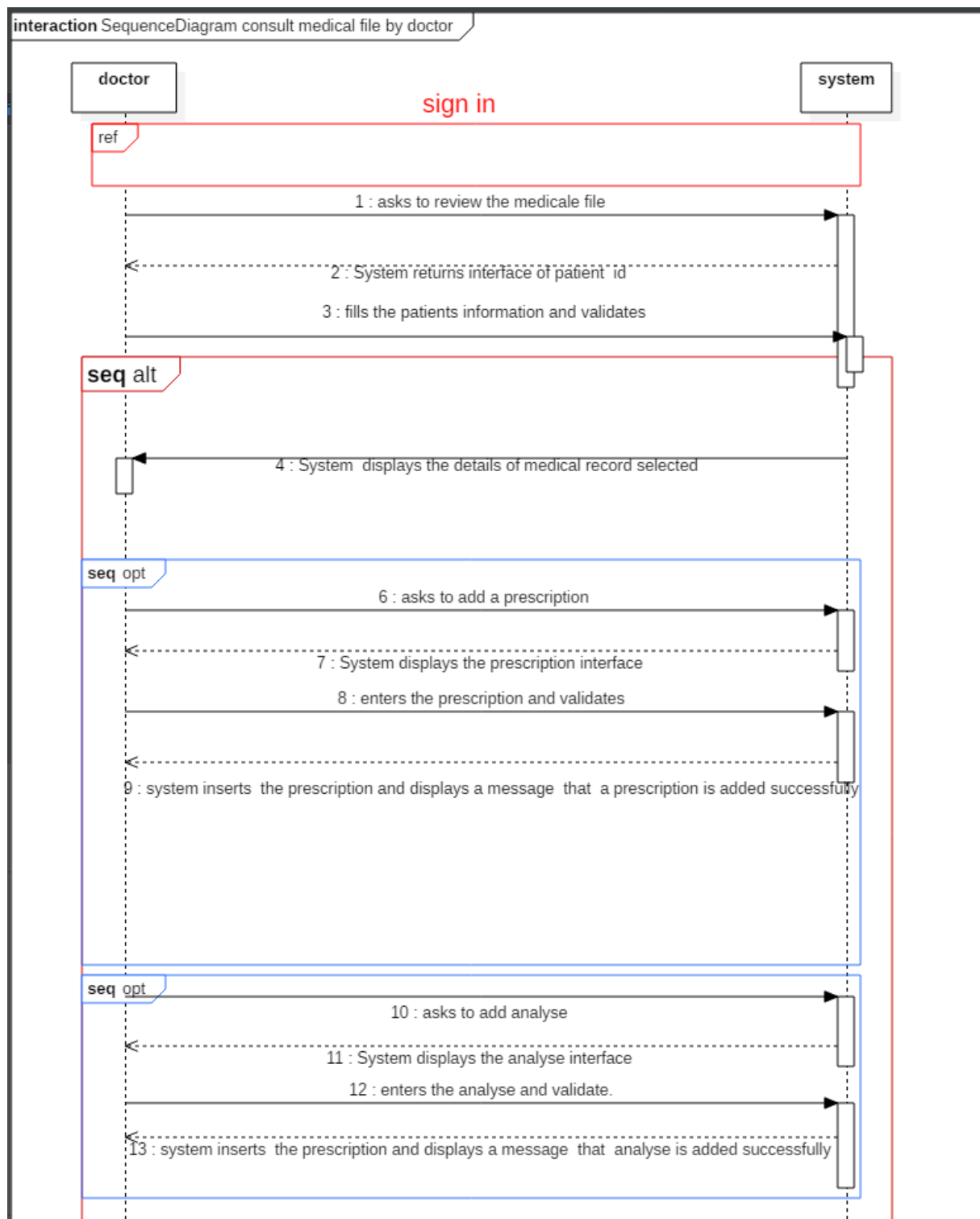
Figure 4.8: Consult medical file by doctor - part 1
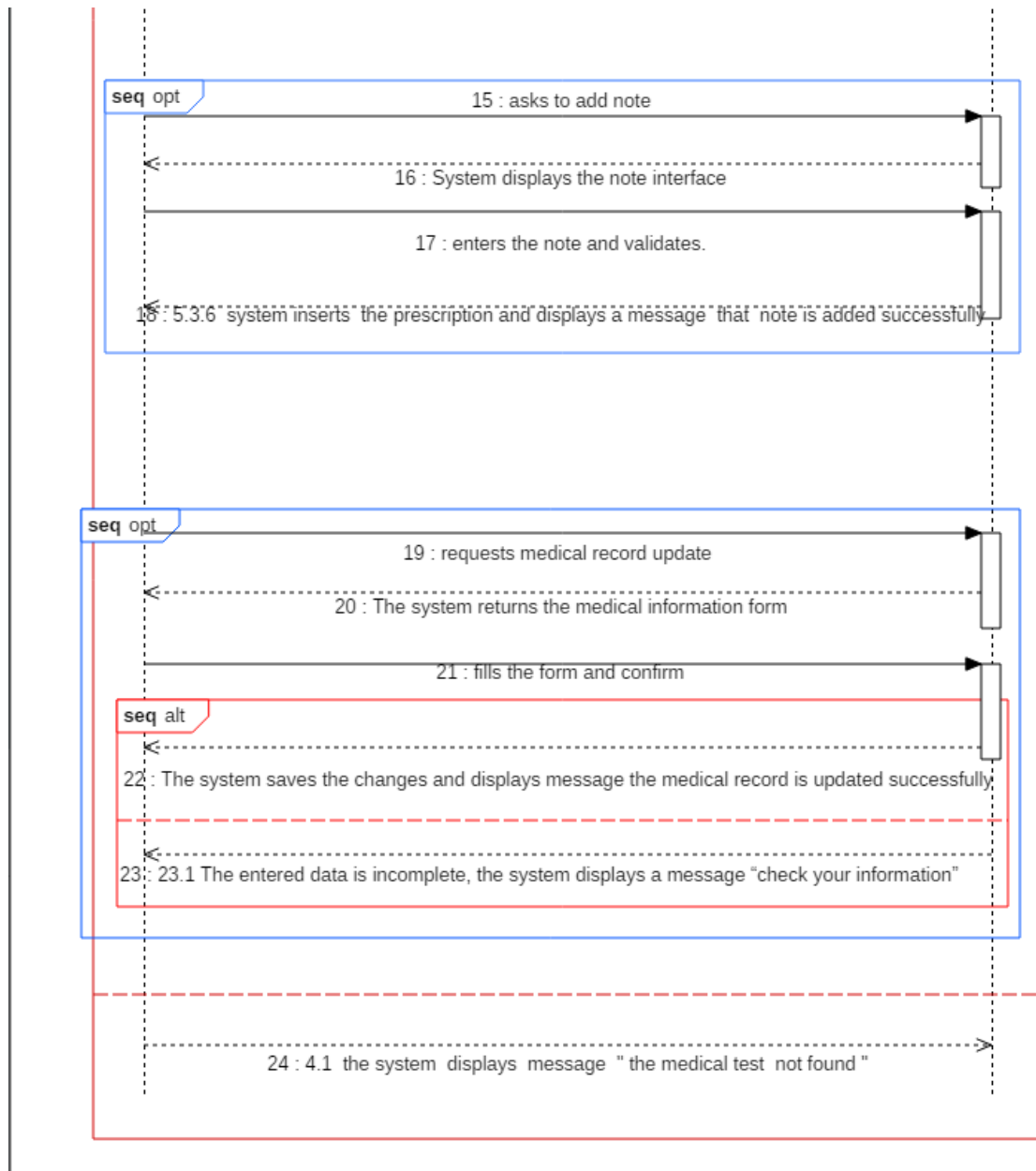
Figure 4.9: Consult medical file by doctor - part 2

Pharmacists use cases description: The tables and figures coming represents a graphical description for all use cases of the actor "Pharmacist".
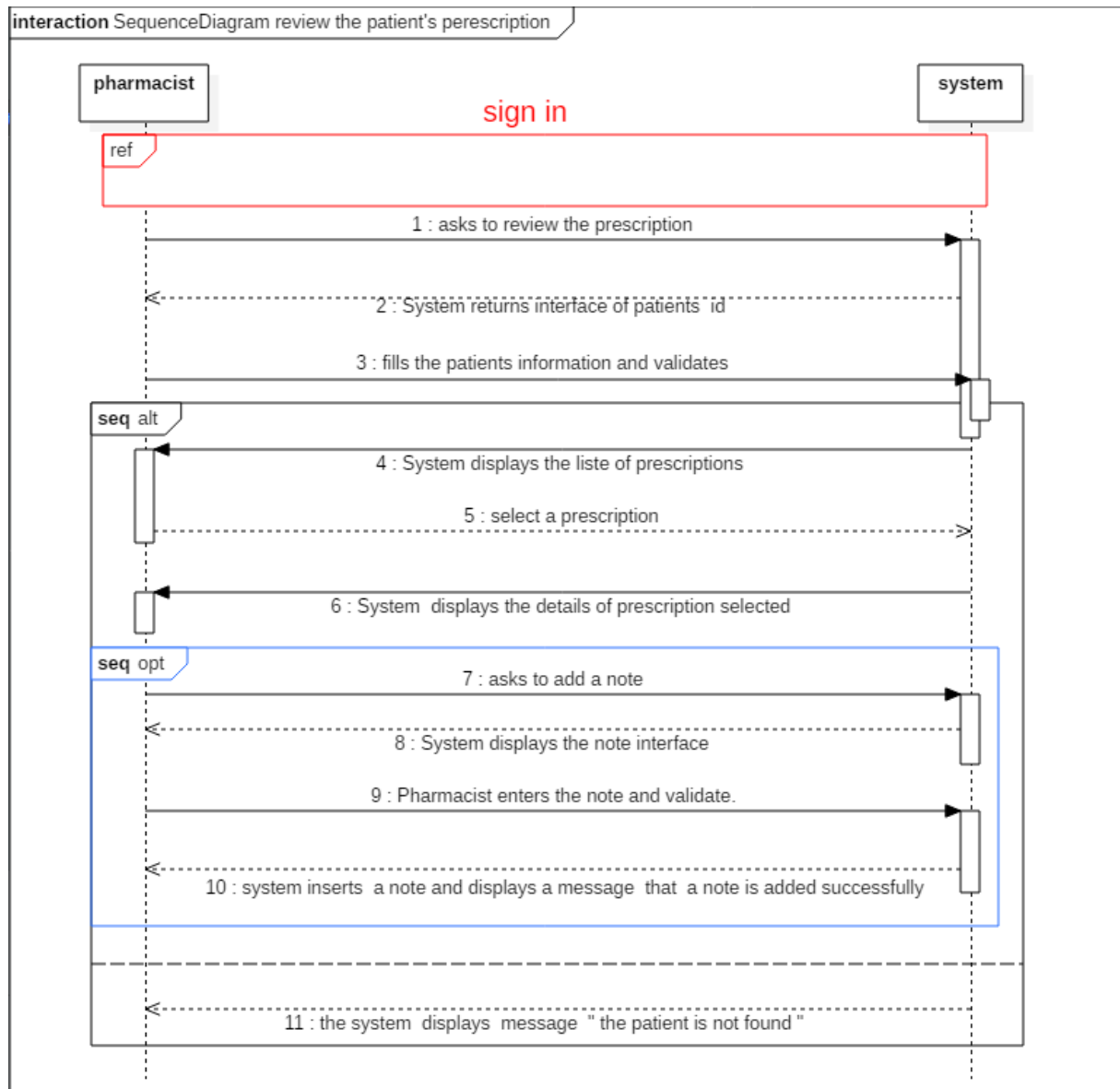


Figure 4.10: consult prescription file

Laboratory's use cases description: The tables and figures coming represents a graphical description for all use cases of the actor "Laboratory".

**interaction** SequenceDiagram consult analyses file

laboratory | system

sign in

ref

1 : asks to review the analyse file

2 : System returns interface of patients id

3 : fills the patients information and validates

**seq** alt

4 : System displays the list of analysis

5 : select a prescription

6 : System displays the details of analysis selected

**seq** opt

7 : asks to add the result of analysis

8 : System displays the result interface

9 : enters the result and validate.

10 : system inserts the result and displays a message that the result is added successfully

11 : 4.1 the system displays message " the patient is not found "

Figure 4.11: consult analyses file

Admins use cases description: The tables and figures coming represents a textual and graphical description for all use cases of the actor "Admin".

| Use case | Consult feed-backs |
|---|---|
| **primary actor** | Admin |
| **secondary actor** | / |
| **Description** | Allows the admin to see all the feed-backs |
| **Precondition** | The admin must be signed in |
| **Post-condition** | All feed-backs list is displayed |
| **primary scenario** | 1- The admin requests list of feed-backs from the system<br>2- The system returns available list of feed-backs |
| **secondary scenario** | 3-1- The admin selects one of the feed-backs to reply<br>3-2- The system returns the reply form<br>3-3- The admin fills in the form and validate |

Table 4.7: textual description: Consult feed-backs

| Use case | Add medical staff |
|---|---|
| **primary actor** | Admin |
| **secondary actor** | Database API |
| **Description** | Allows the admin to add a new member of the medical staff |
| **Precondition** | The admin must be signed in |
| **Post-condition** | A new member of the medical staff is added |
| **primary scenario** | 1- The admin requests to add a new member<br>2- The system returns the adding form<br>3- The admins fills the form<br>4- The system uses the Database API to verify the validity of the new member<br>5- The system demands the confirmation from the admin<br>6- The admin validates<br>7- The system displays "The new member has been added successfully" |
| **secondary scenario** | 4-1- The new member does not exist<br>4-2- the system displays "the member does not exist" |

Table 4.8: textual description: Add medical staff

| Use case | Consult Users |
|---|---|
| **primary actor** | Admin |
| **secondary actor** | / |
| **Description** | Allows the admin to consult the list of users and manage it |
| **Precondition** | The admin must be signed in/The list exists |
| **Post-condition** | the list of users is displayed and/or updated |
| **primary scenario** | 1- The admin demands to consult users<br>2- The system returns the list of users |

| secondary scenario | 3-1- The admin demand to update users information |
| | 3-2- The system returns users update form |
| | 3-3- The admin fills the form and validate |
| | 4-1- The admin demands to block a user |
| | 4-2- The system returns confirmation message |
| | 4-3- The admin confirms the process |

Table 4.9: textual description: Consult users

## 4.4 Capturing technical requirements

This section deals generally with the capture of technical requirements. This phase serves to complement the capture of functional requirements. The vector idea at this level revolves around the identification of the various constraints that are neither descriptive of the users profession, nor descriptive from an application point of view or of any business process. The model of specification is expressed in two ways [13]:

- The technical specification from a hardware perspective.

- Capture software specifications.

### 4.4.1 The technical specification from a hardware perspective:

The technical prerequisites were expressed in the preliminary study, when operational requirements and strategic development choices were expressed. These choices involve constraints related to the configuration of the hardware network and are geographic, organizational, and technical in nature. They concern data access performance, system security, volume [32].

- **Tiered architecture style** The level architecture style specifies the number of geographic and organizational levels where the system's run time environments will be located [18].
  In our system The geographical configuration requires the development of a two-tier client/server solution.

- **Architecture specification 2 levels** A two-tier architecture also
  known as client-server is an architecture where a Client Station exe-
  cutes an application on a server.

  This is done as a dialogue between the client and the server in which:
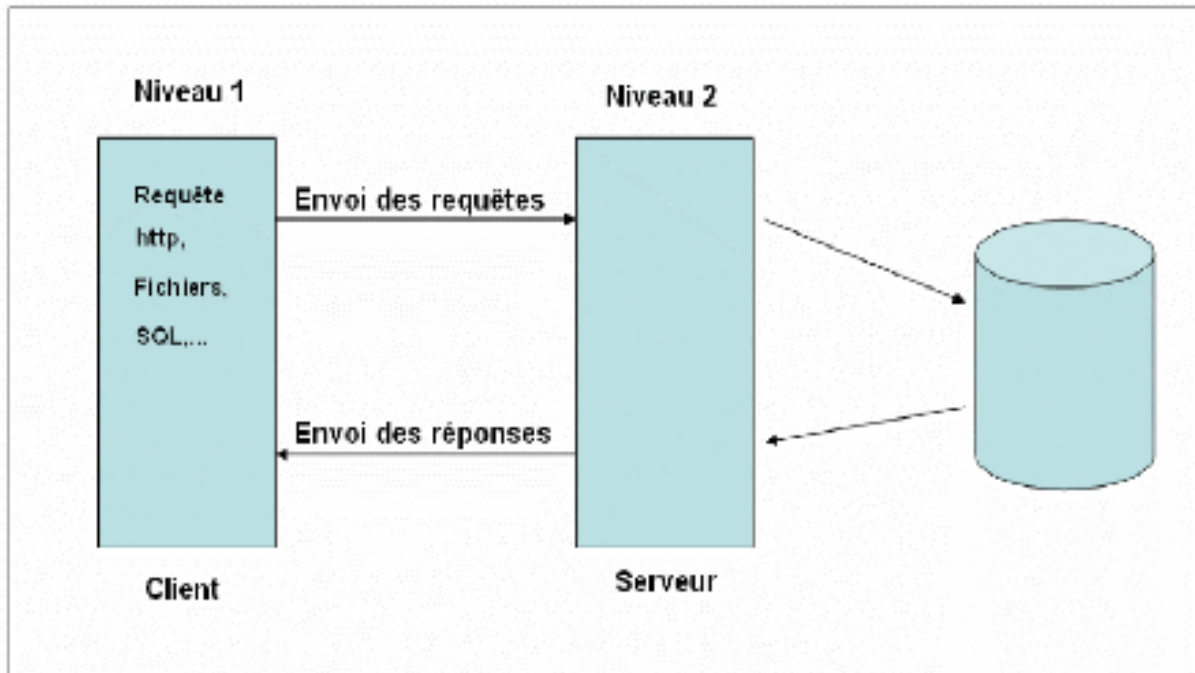


Figure 4.12: The two tier architecture

## 4.4.2 Capturing Software Specifications:

Once the technical and architectural specifications have been expressed,
the specific features of the technical system can be addressed by a software
specification.

In this case, the use cases are used differently than for the functional spec-
ification. This is why we have introduced the concept of operator and
technical use cases.

- Operator: The operator is an actor in the sense of UML, except that
  it only benefits from the technical functionalities of the system.

- Technical use case: a technical use case is intended for the operator,
  it is a sequence of action producing operational or purely technical
  added value.

Identification of system operators: The operators of our system are the
majority of the actors of the functional branch, are therefore users
in the technical dimension and who are: the patient, the doctor, the
admin, the pharmacist, the laboratory.

Identification of technical use cases: The main technical use cases of our
system are illustrated by the following figure:
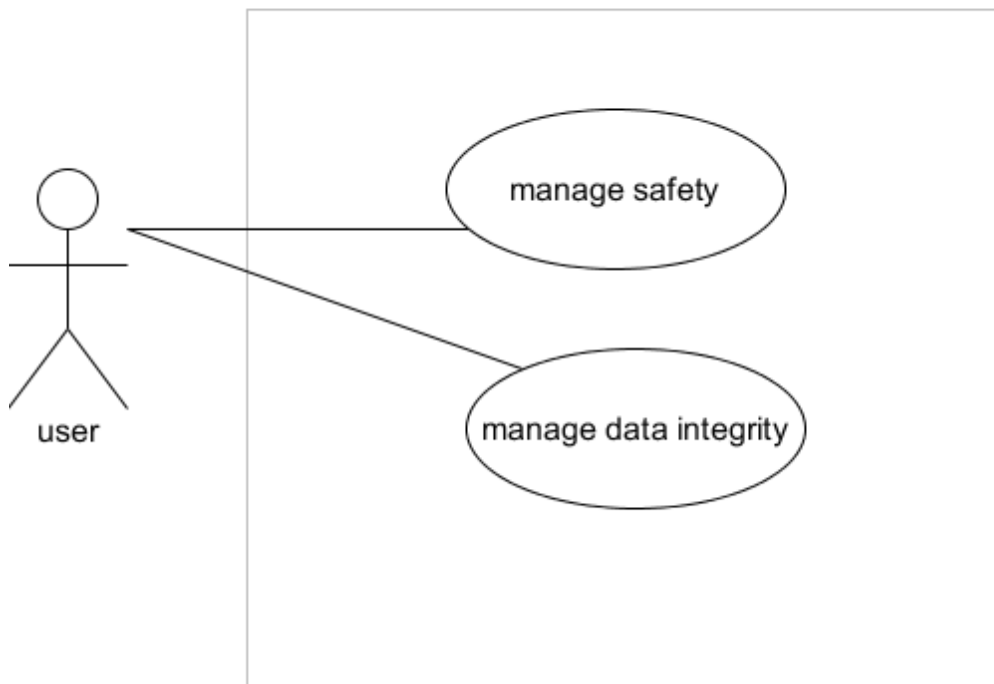


Figure 4.13: System Software Specification Template

### 4.4.3 Description of technical use cases:

**Managing Safety:**

Textual description:

| Use case | Authenticate |
|---|---|
| primary actor | User |
| Description | This case allows to verify the access authorization to application features. |
| Precondition | User has an account |
| Post condition | The user is authenticated. |
| primary scenario | 1-The user requests access to the application. 2-The system displays the authentication form. 3-The user enters his username and password and confirms authentication. 4-The system checks the authentication information. 5-The system returns the corresponding interface. |
| secondary scenario | 3-1- Incorrect ID and/or password. System displays error message. Repeat nominal scenario sequence in 2. |

Table 4.10: textual description: Authenticate

Graphical description
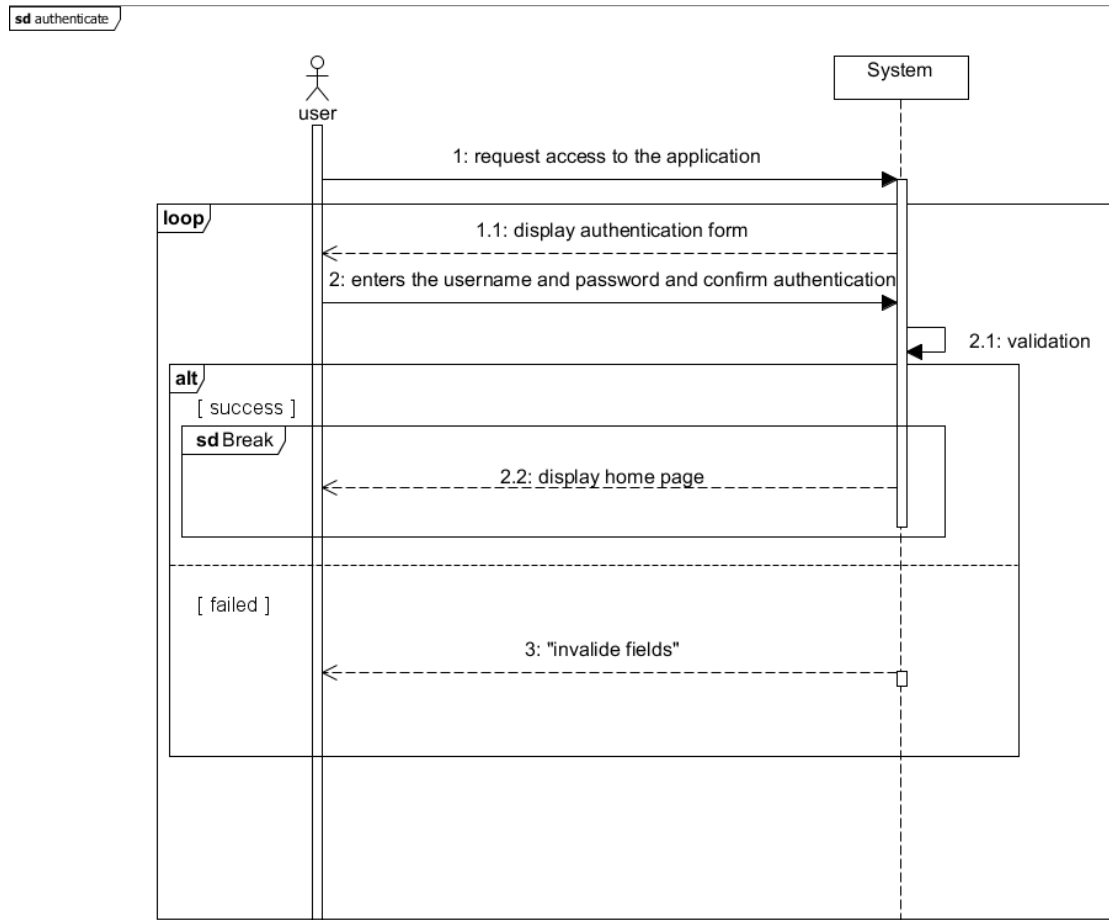


Figure 4.14: Sequence diagram of the use case: Authenticate

**Manage data integrity:**

Textual description:

| Use case | Manage data integrity |
|---|---|
| primary actor | User |
| Description | This case allows to Verify data integrity of the application features. |
| Precondition | the user authenticates |
| primary scenario | 1-The user enters the data. 2-The system checks the validity of each field. |
| secondary scenario | 2-1-The system displays an "invalid fields" notification. The scenario is repeated in point 1. |

Table 4.11: textual description: Manage data integrity

Graphical description
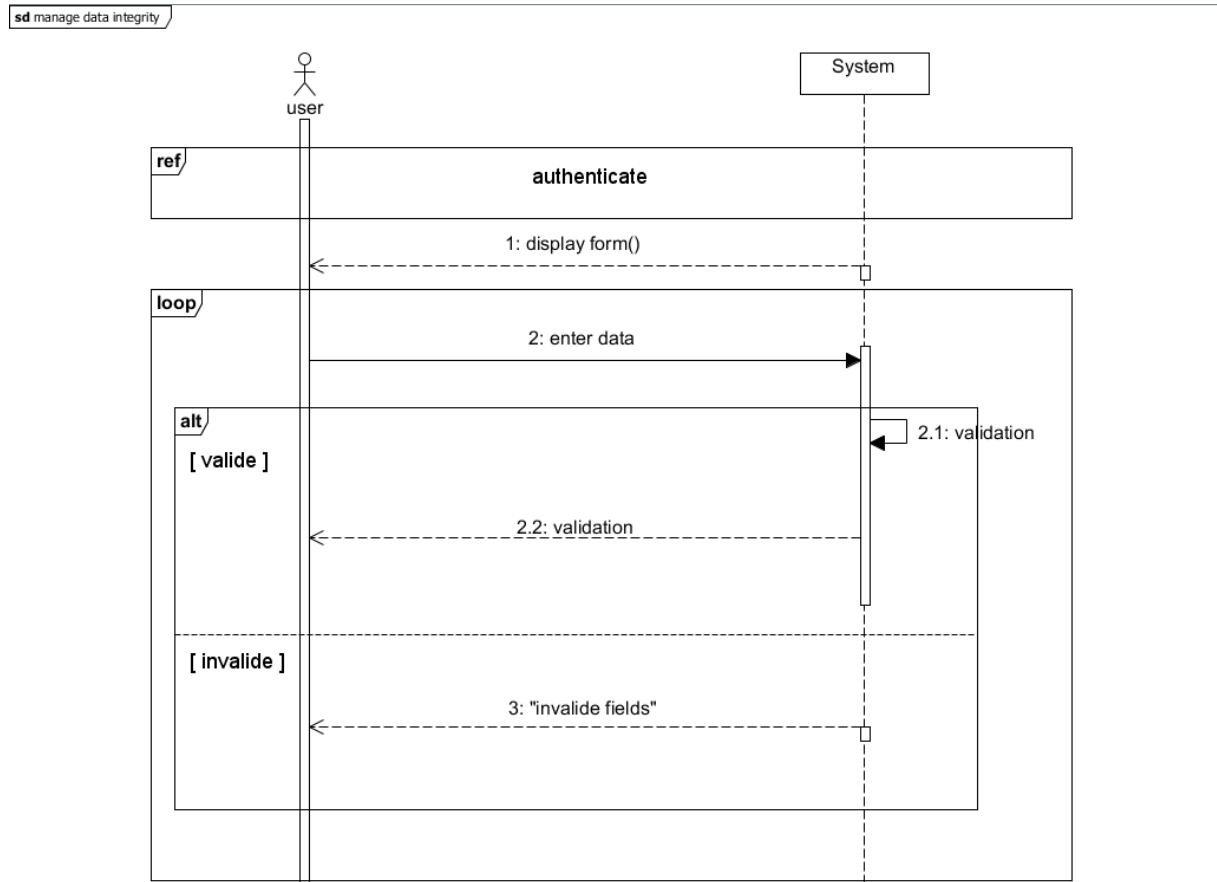


Figure 4.15: Sequence diagram of the use case: Manage data integrity

## 4.5 Conclusion

The needs capture phase is a step that works on the one hand to complete the collections of needs operated during the preliminary study phase, and on the other hand, it gives a first view for the next stage.

# Chapter 5

# analysis phase

## 5.1 Introduction

This chapter is devoted to the analysis phase. We start with the diagrams of participating classes. Then we move on to the Detailed Sequence Diagrams.

## 5.2 Stage status in the 2TUP process

The analysis phase represents the second stage of the left branch of the Y cycle.It is dedicated to the object analysis of our system just after the preliminary study and the capture of functional and technical requirements.It consists in precisely studying the specifications to get an idea of what the system will achieve in terms of trades [18].
the following figure represents the stage status in the 2TUP process:
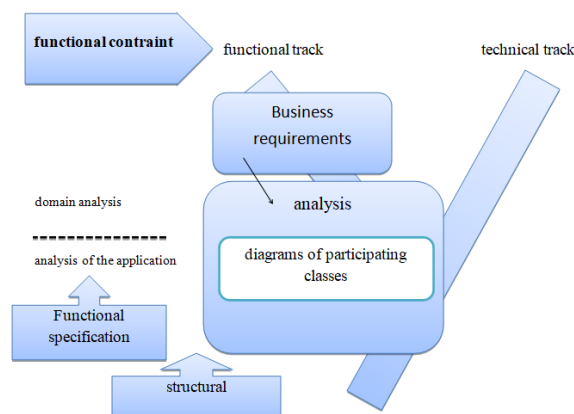


Figure 5.1: The stage status in the 2TUP process

## 5.3 Identification of participating classes diagram

This is a diagram that is used to identify the analysis classes that will participate in the implementation of use cases. Three types of analysis classes may exist in this diagram:

- "Dialogues" which represent the means of interaction with the system.

- "Controls" that contain the application logic.

- "Entities" which are the business objects handled [19].

## 5.4   diagrams of participating classes

**Guest:**

<u>Create an account</u>:
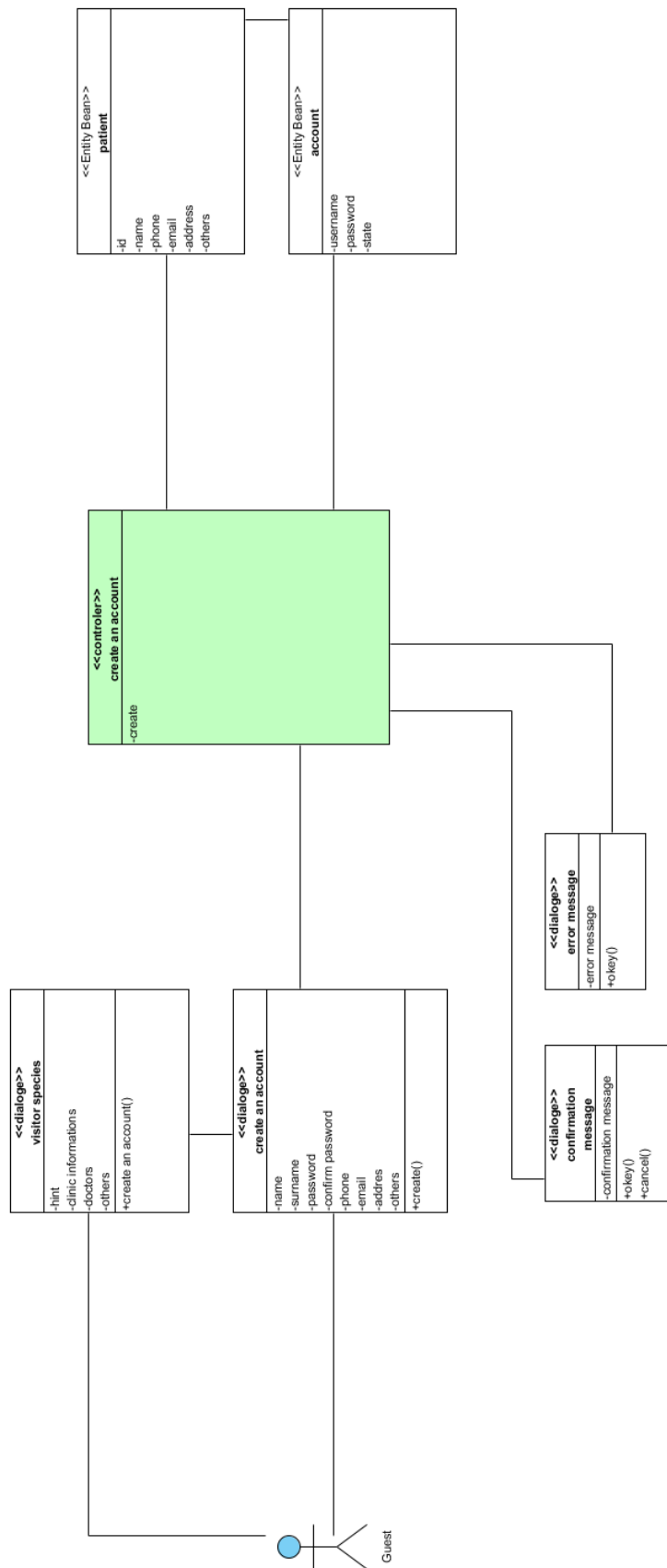
Figure 5.2: Create an account

Create medical record:
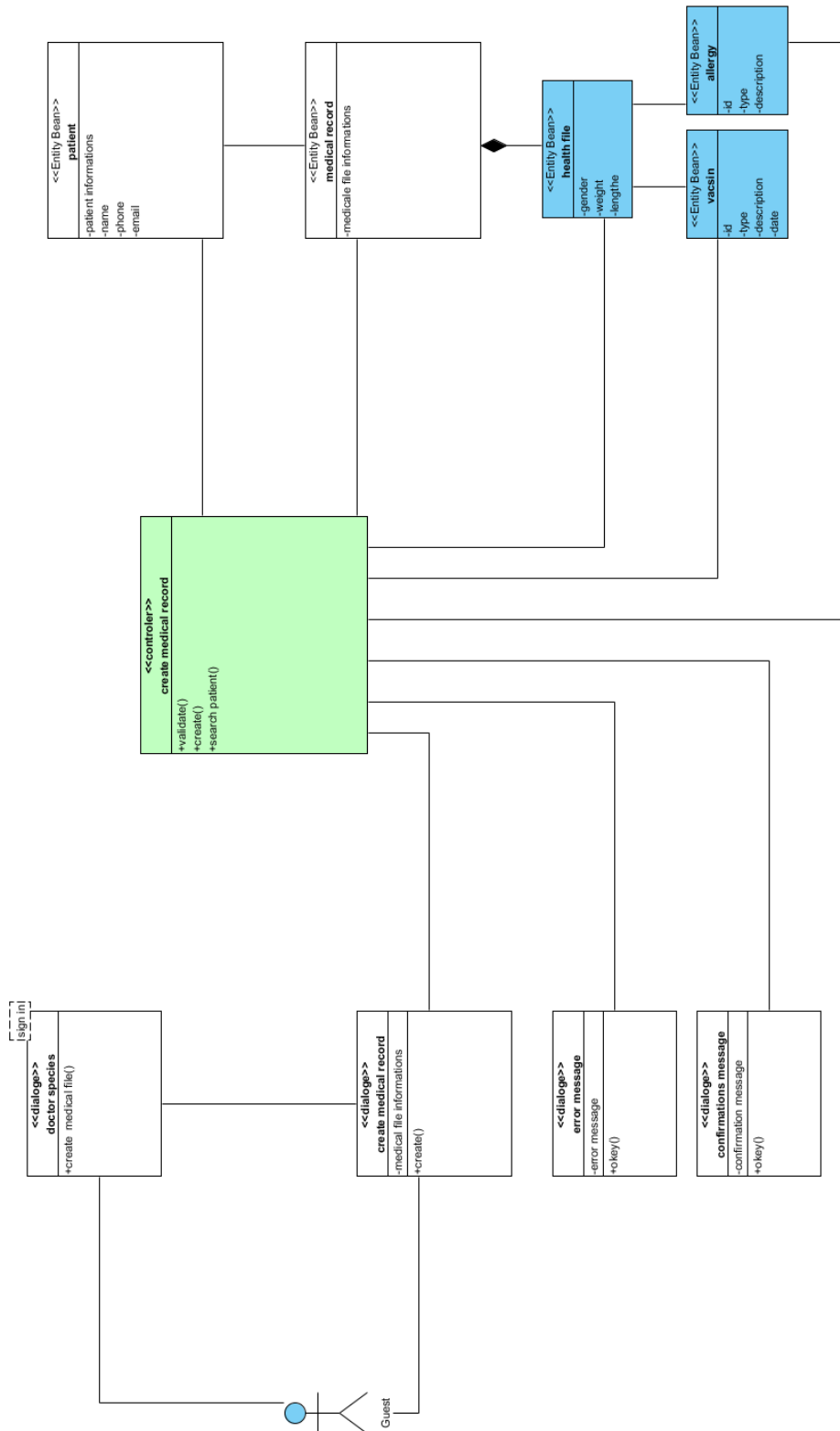


Figure 5.3: Create medical record

Consult clinics information:
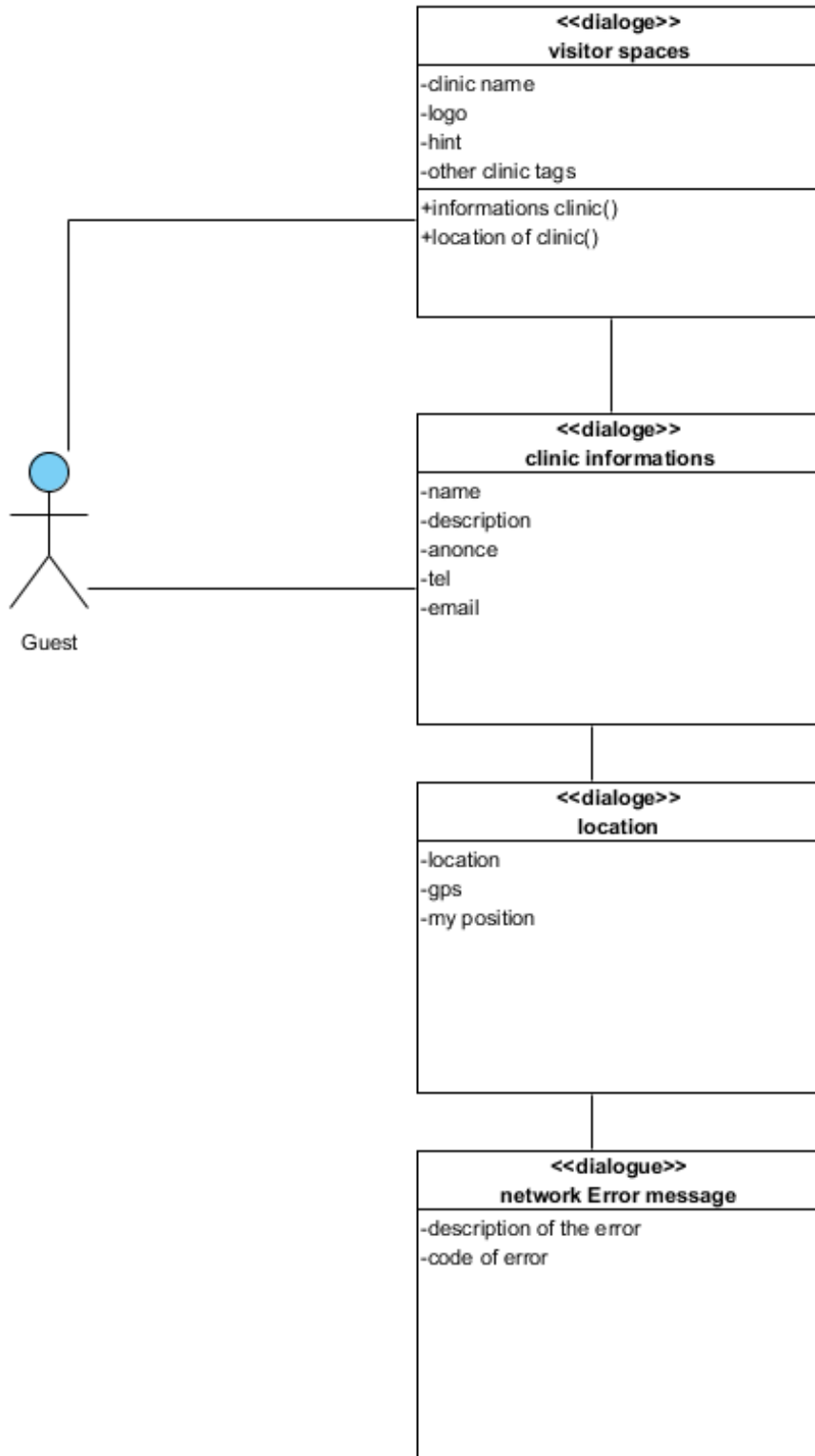


Figure 5.4: Consult clinics information

Consult available doctors and specialties:
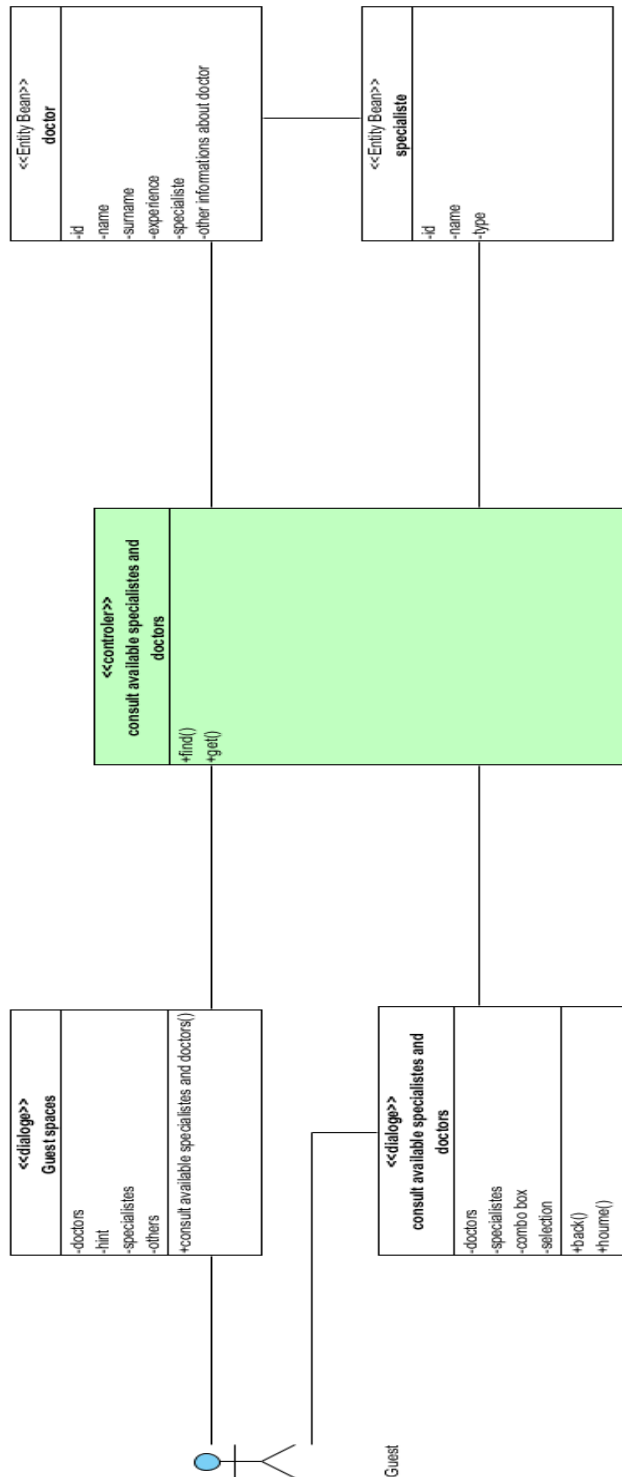


Figure 5.5: Consult available doctors and specialties

Consult profile:
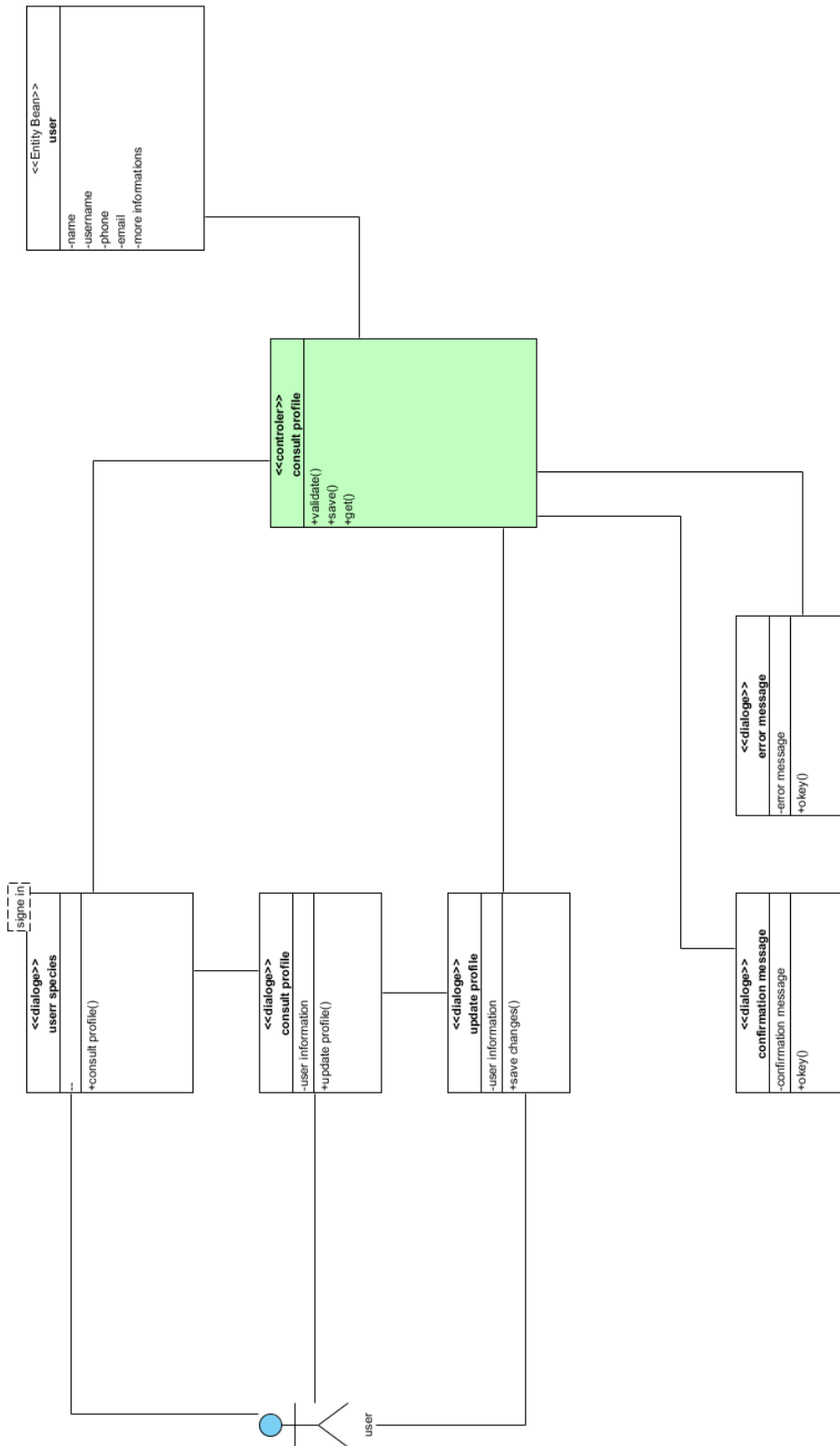


Figure 5.6: consult profile

## Sign in:
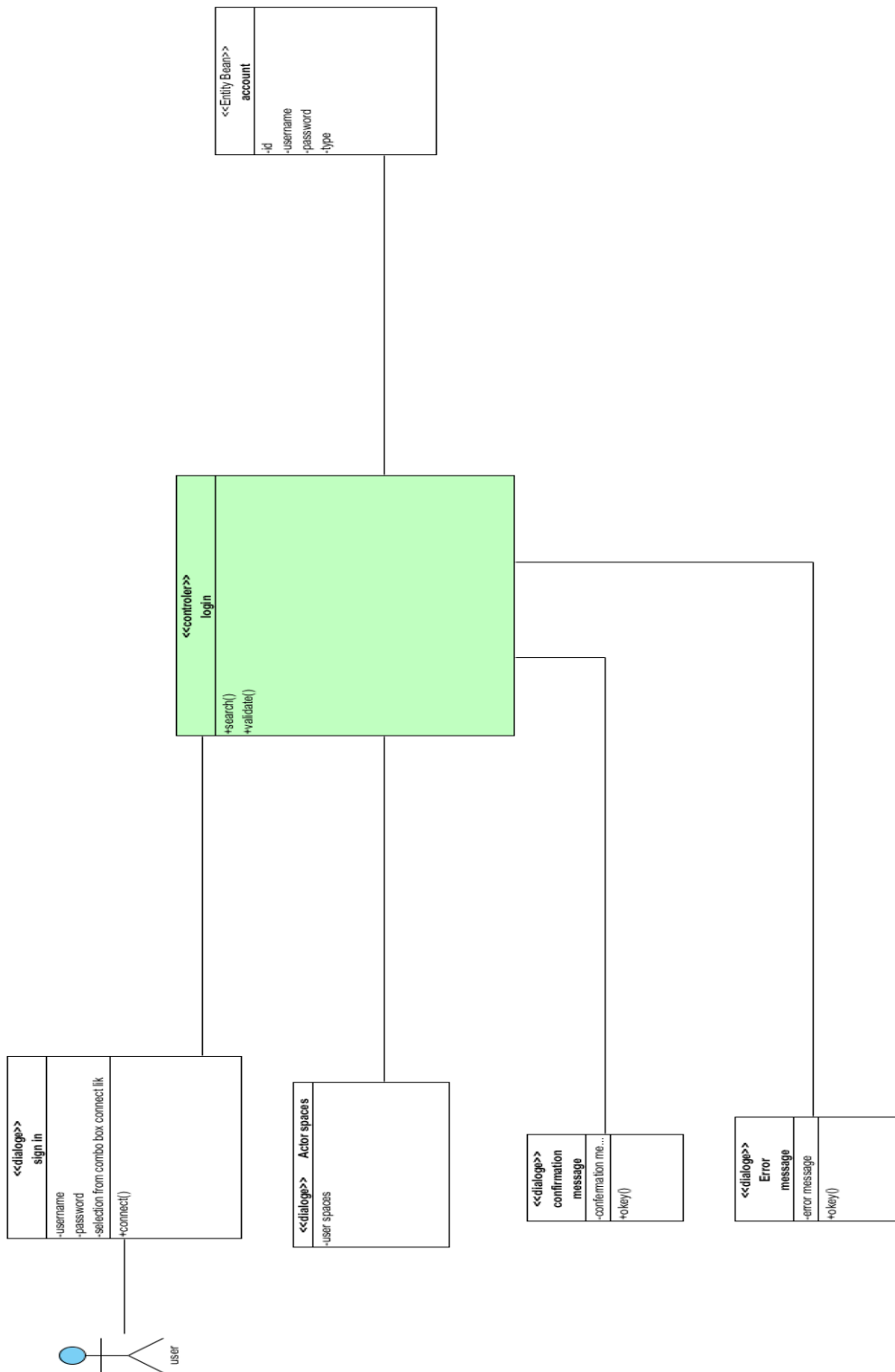


Figure 5.7: Sign in

**Patient:**

Consult available doctors and specialties:
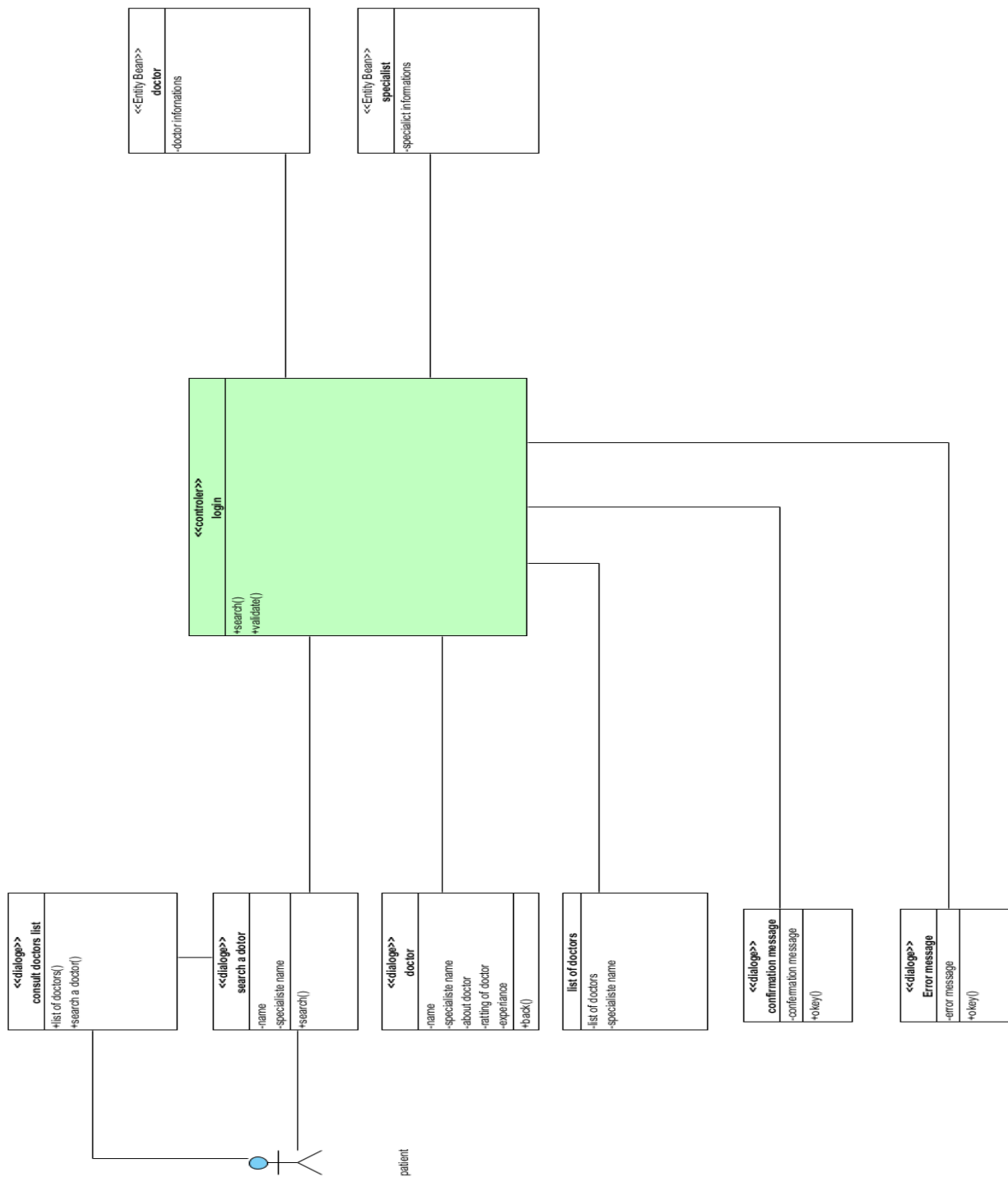


Figure 5.8: Consult available doctors and specialties

Predict the disease:
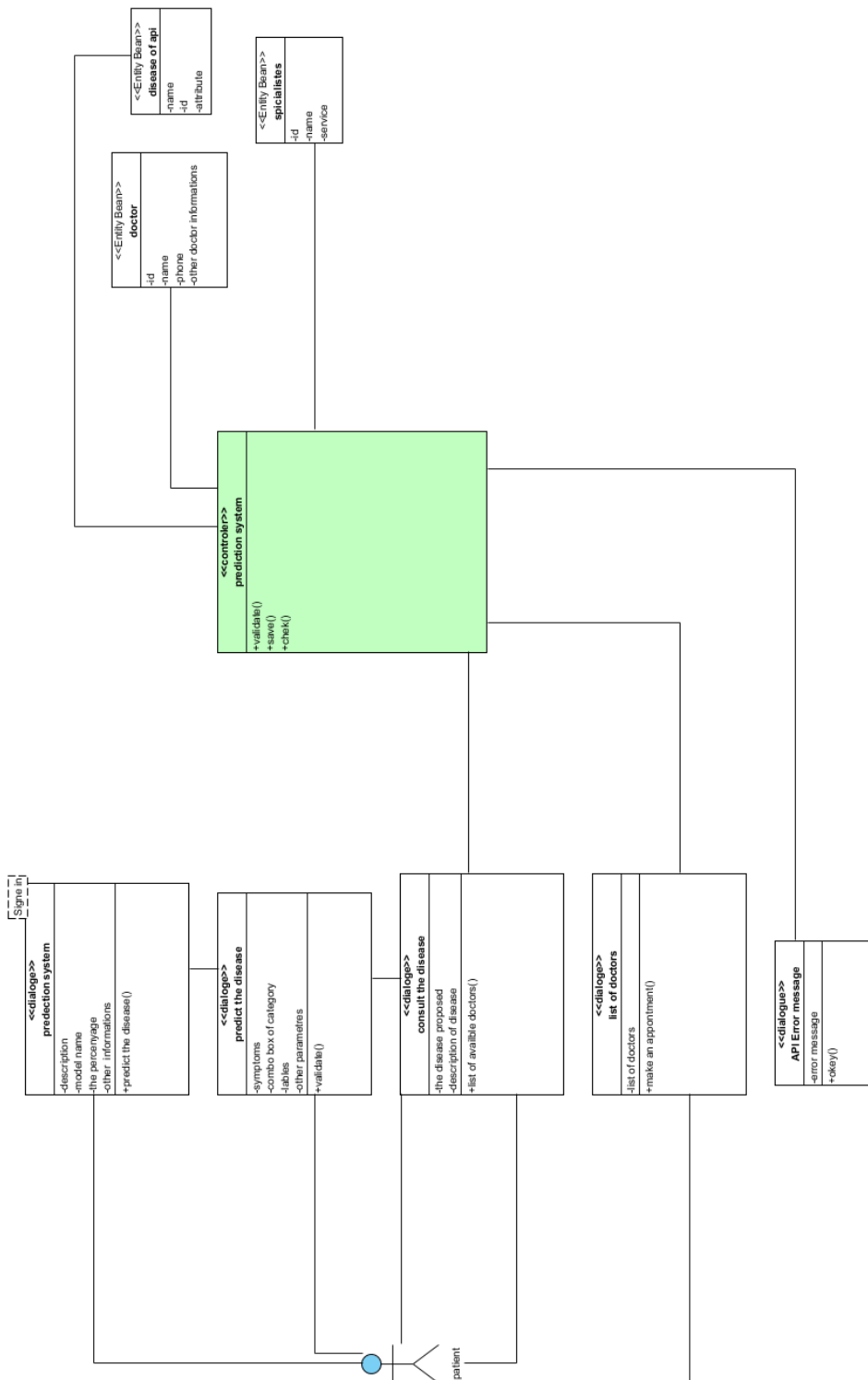


Figure 5.9: Predict the disease

**Doctor:**

Create medical file:
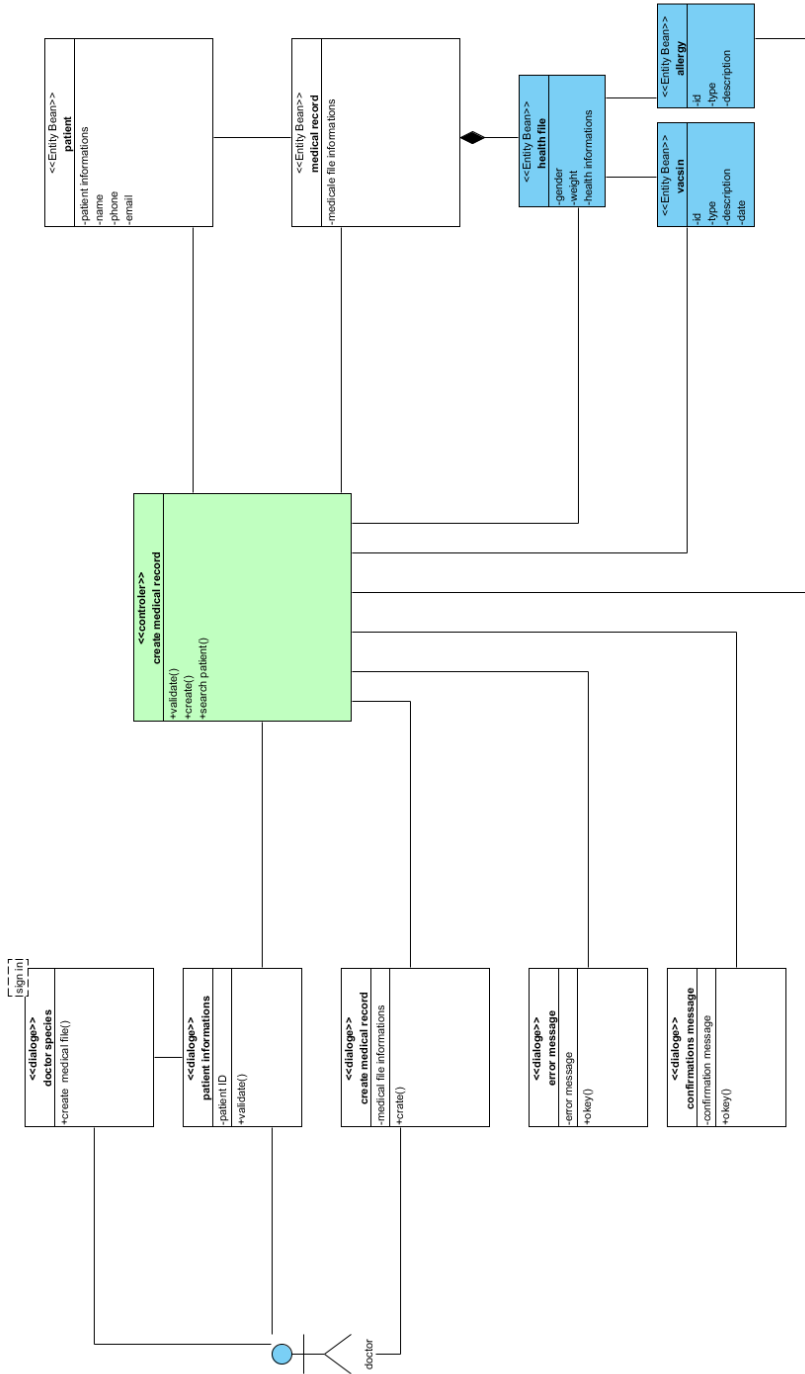


Figure 5.10: Create medical file

consult medical record:
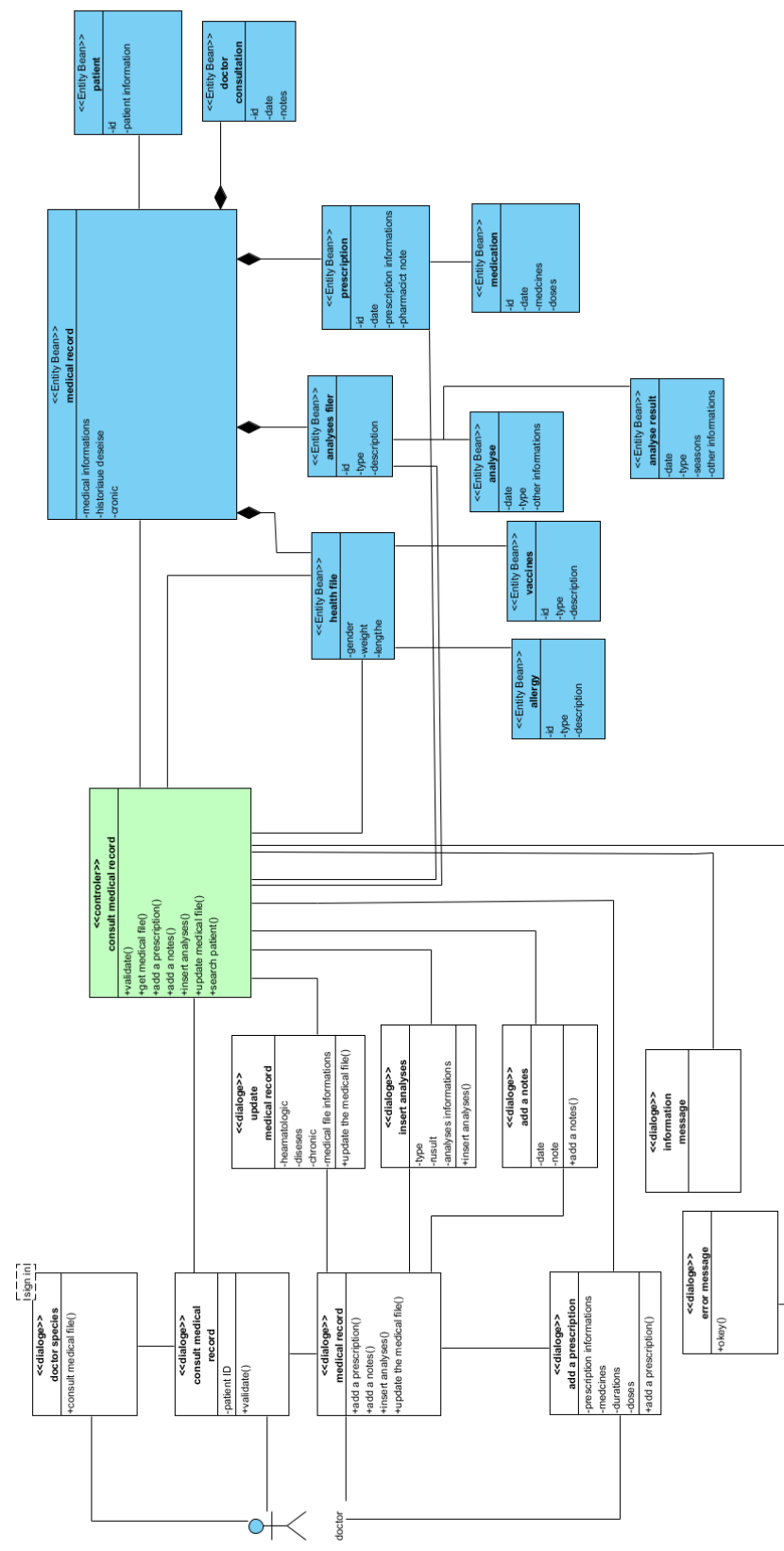


Figure 5.11: consult medical record

**Pharmacist:**

consult prescription:
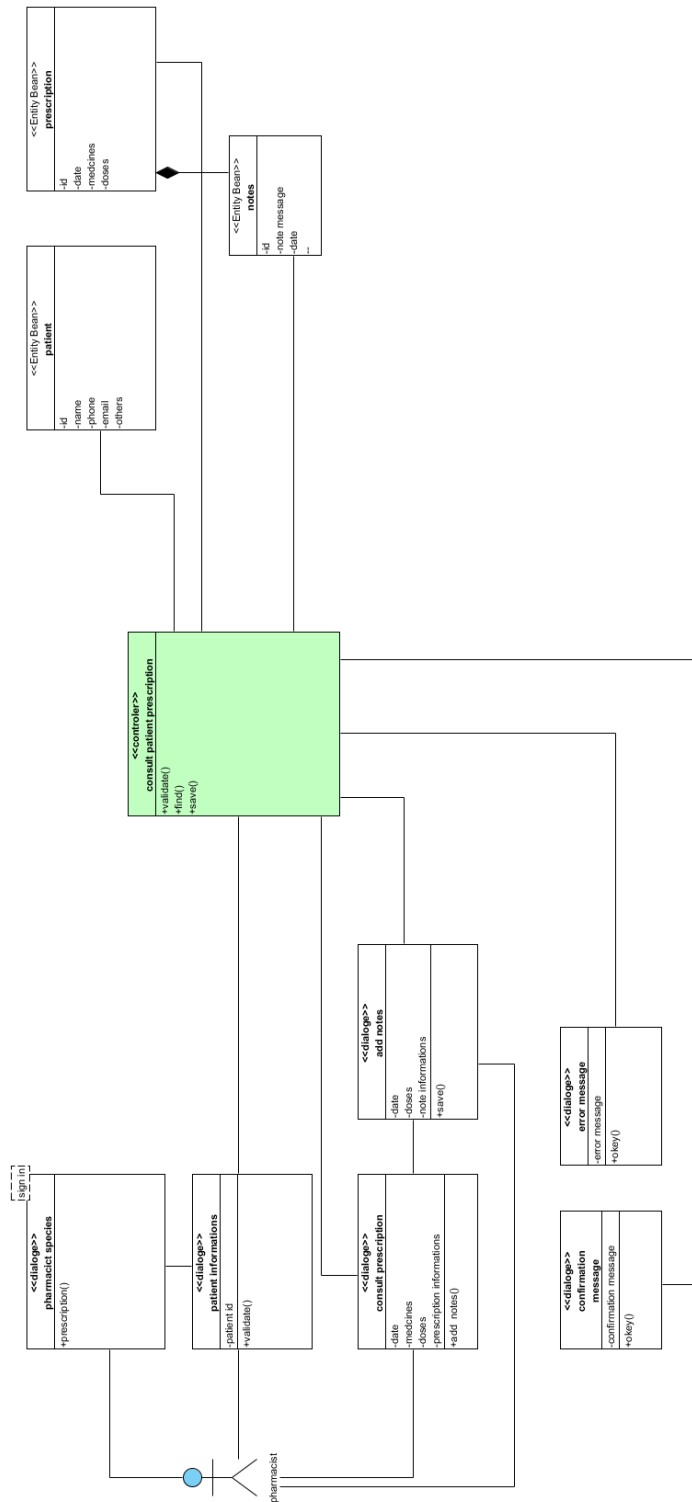


Figure 5.12: consult prescription

**Laboratory:**

consult analysis file:
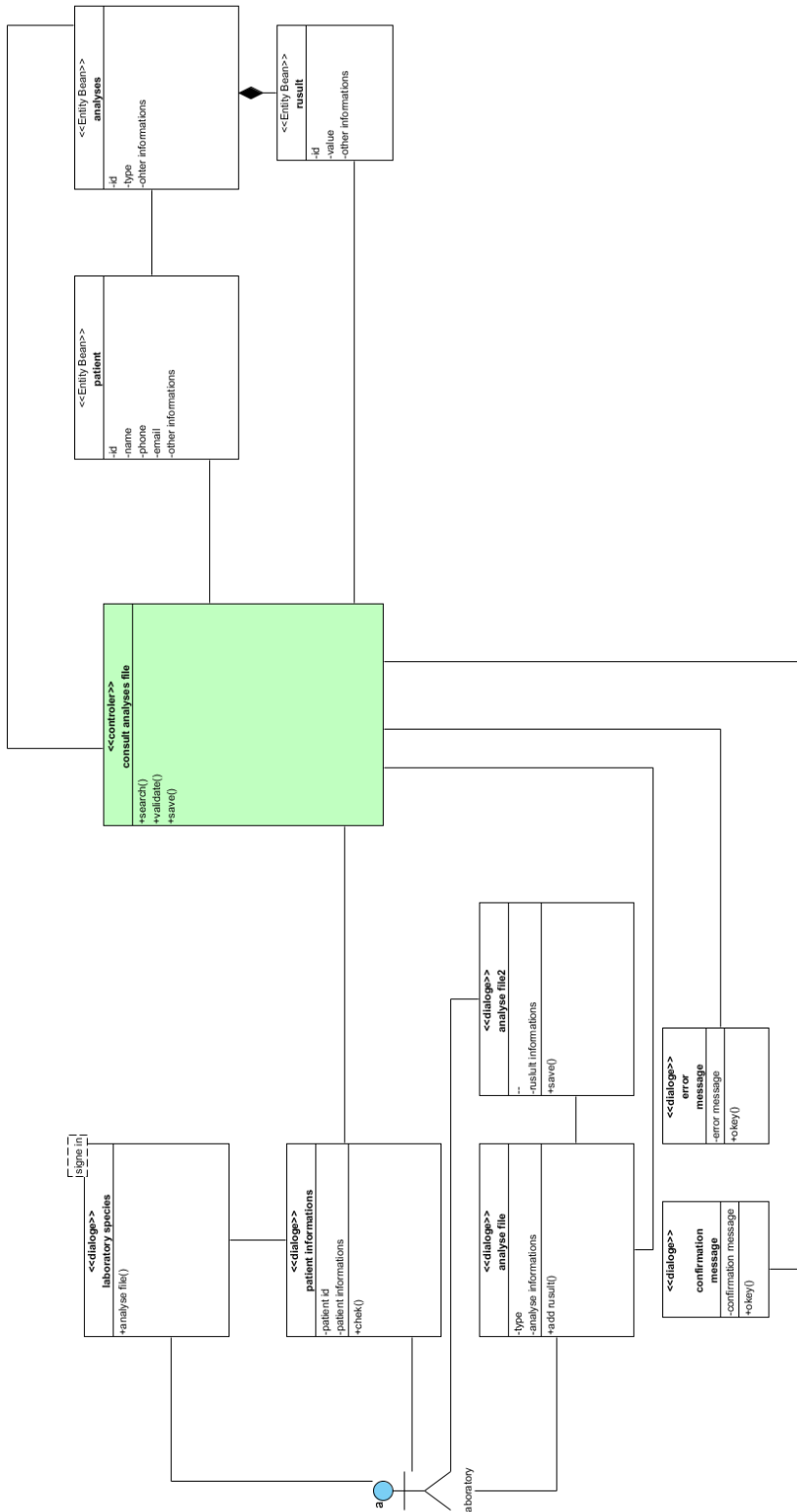


Figure 5.13: consult analysis file

**Administrator:**

# Add medical staff member:
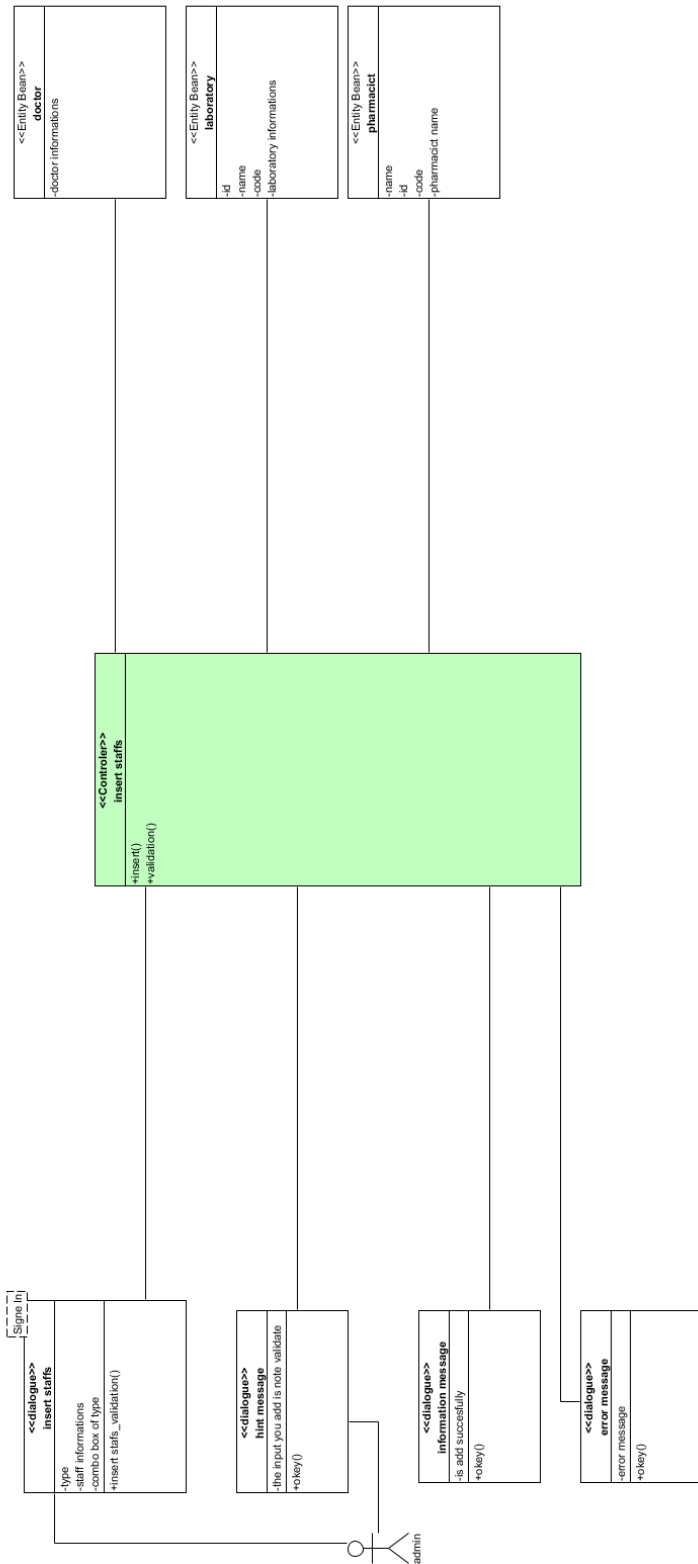


Figure 5.14: Add medical staff member
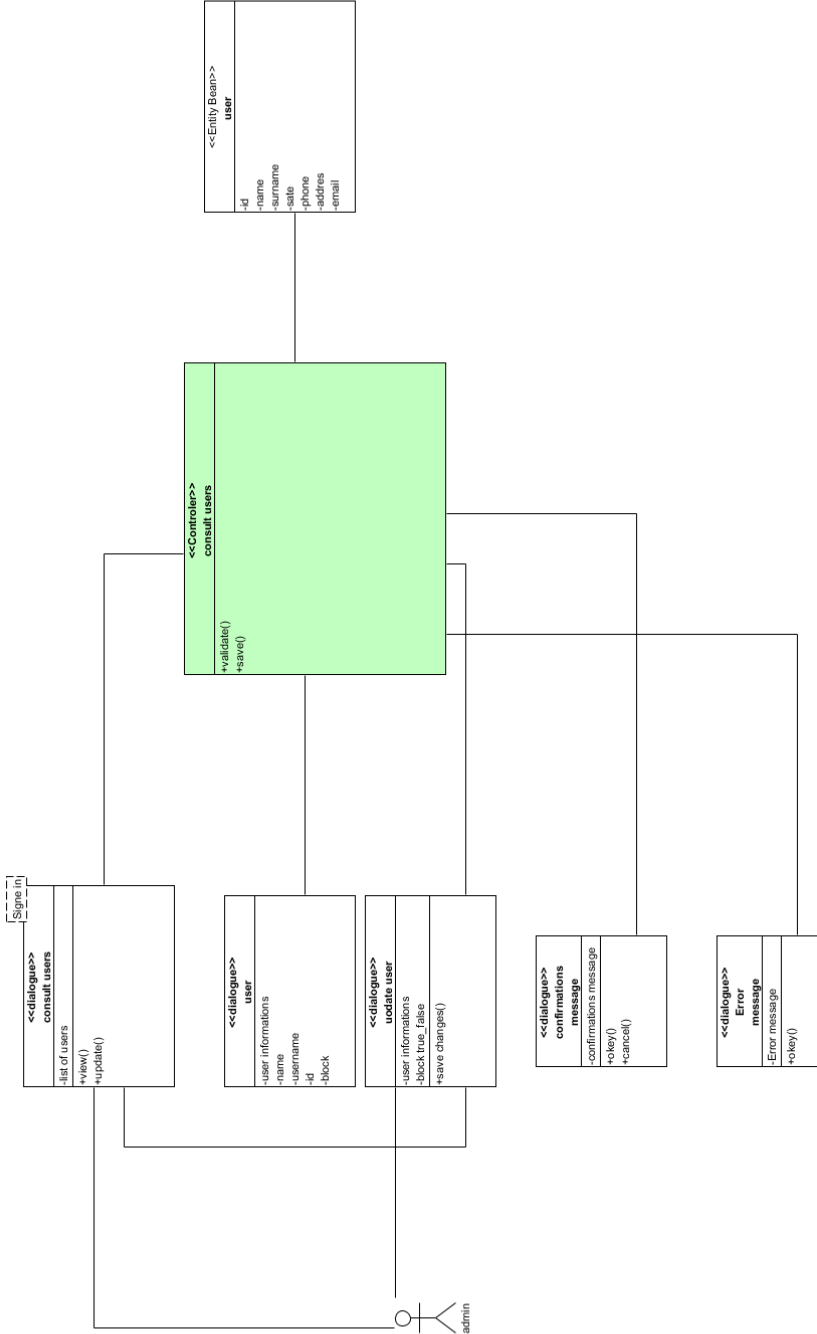
consult users:



Figure 5.15: consult users

consult feed-backs:
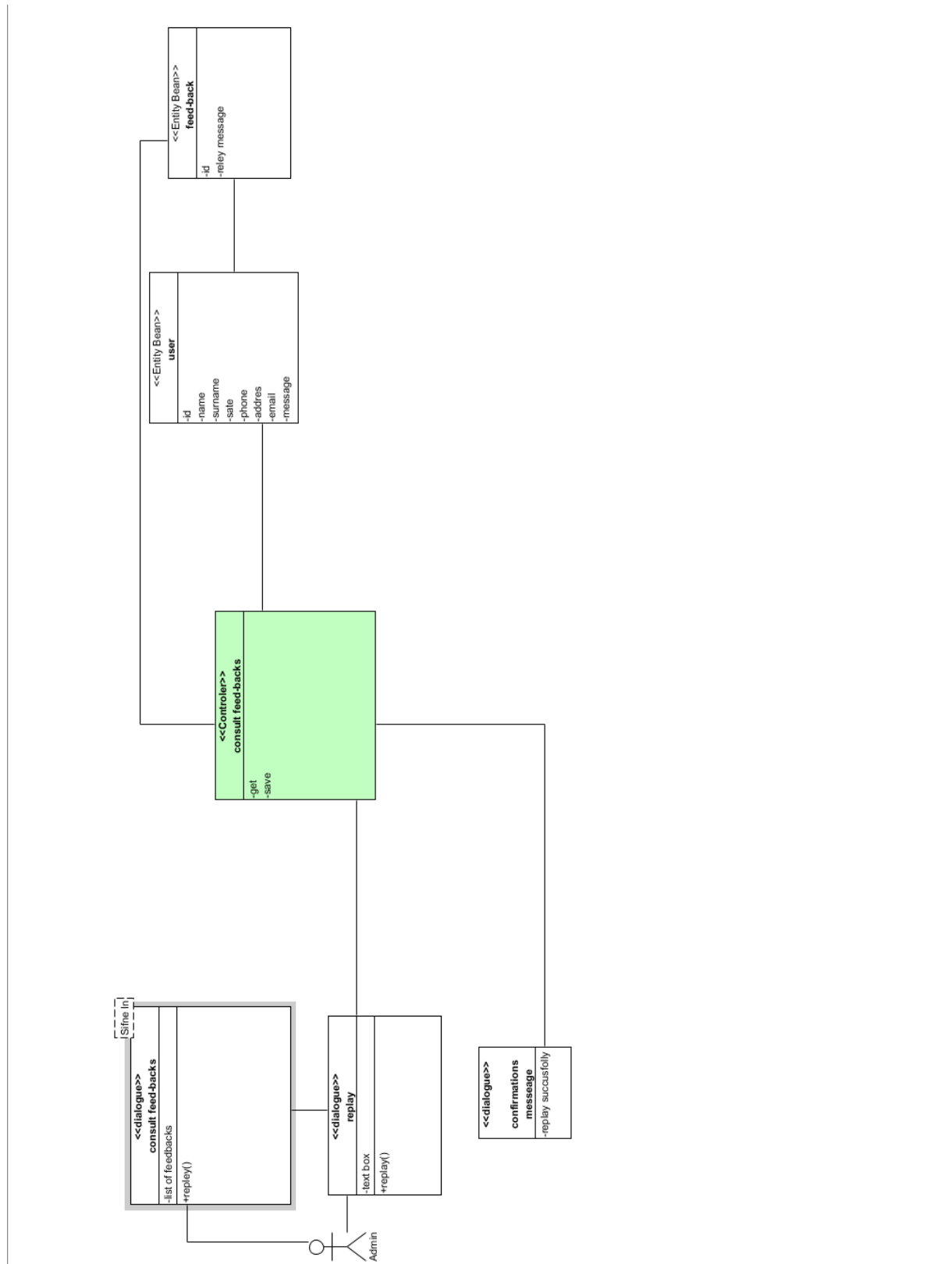


Figure 5.16: consult feed-backs

## 5.5    Identification of detailed sequence diagrams

During the requirements capture phase we saw that a use case represented by a sequence diagram describes a set of scenarios but considering the system as a black box, then to open this box the detailed sequence diagram is elaborated. this diagram allows to replace the system by a collaboration of objects in each scenario.

## 5.6    diagrams of participating classes

**Create an account:**
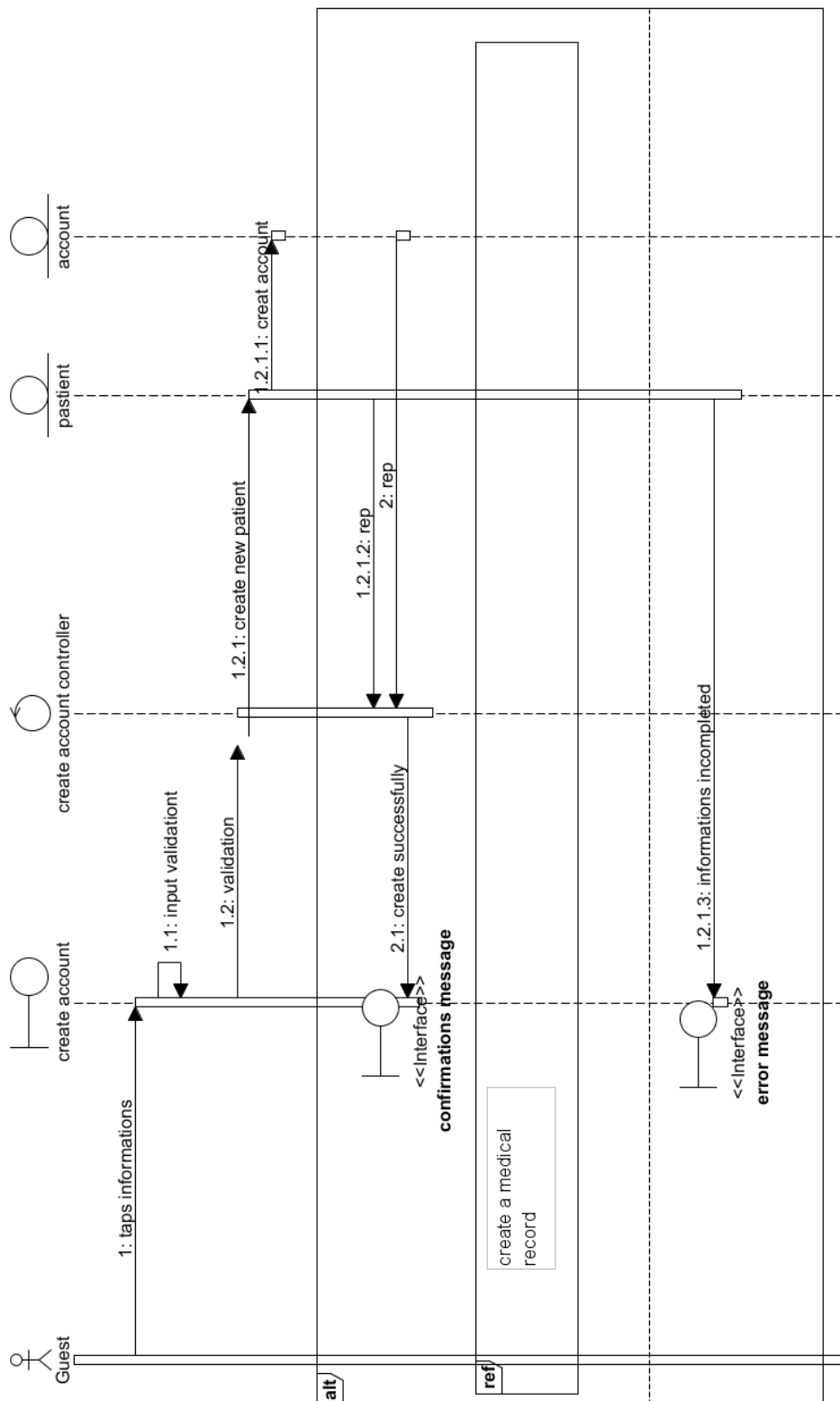


Figure 5.17: Detailed sequence diagrams of the use case:create an account

**Create medical record:**
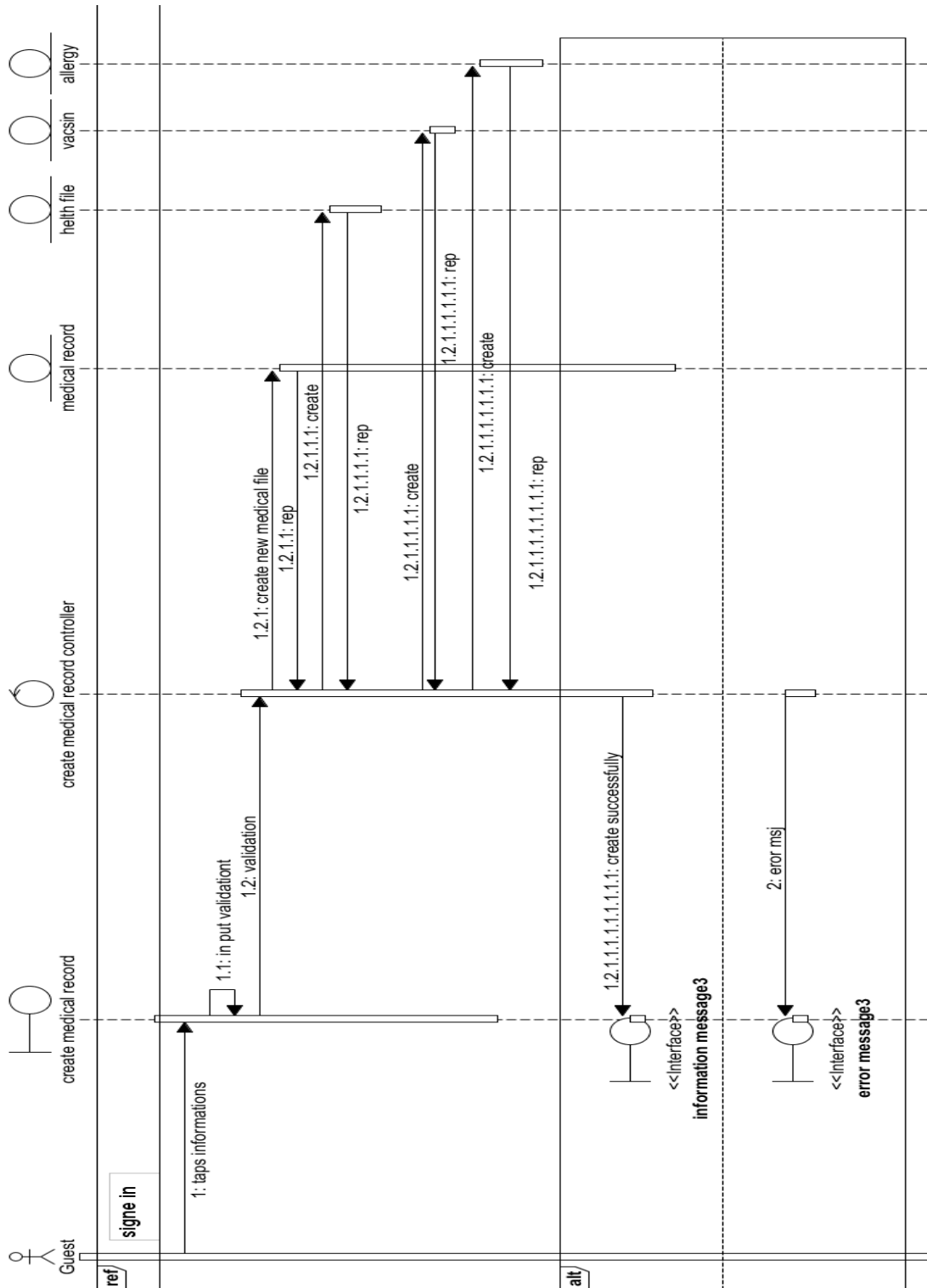


Figure 5.18: Detailed sequence diagrams of the use case: create medical record

**Consult clinics information:**
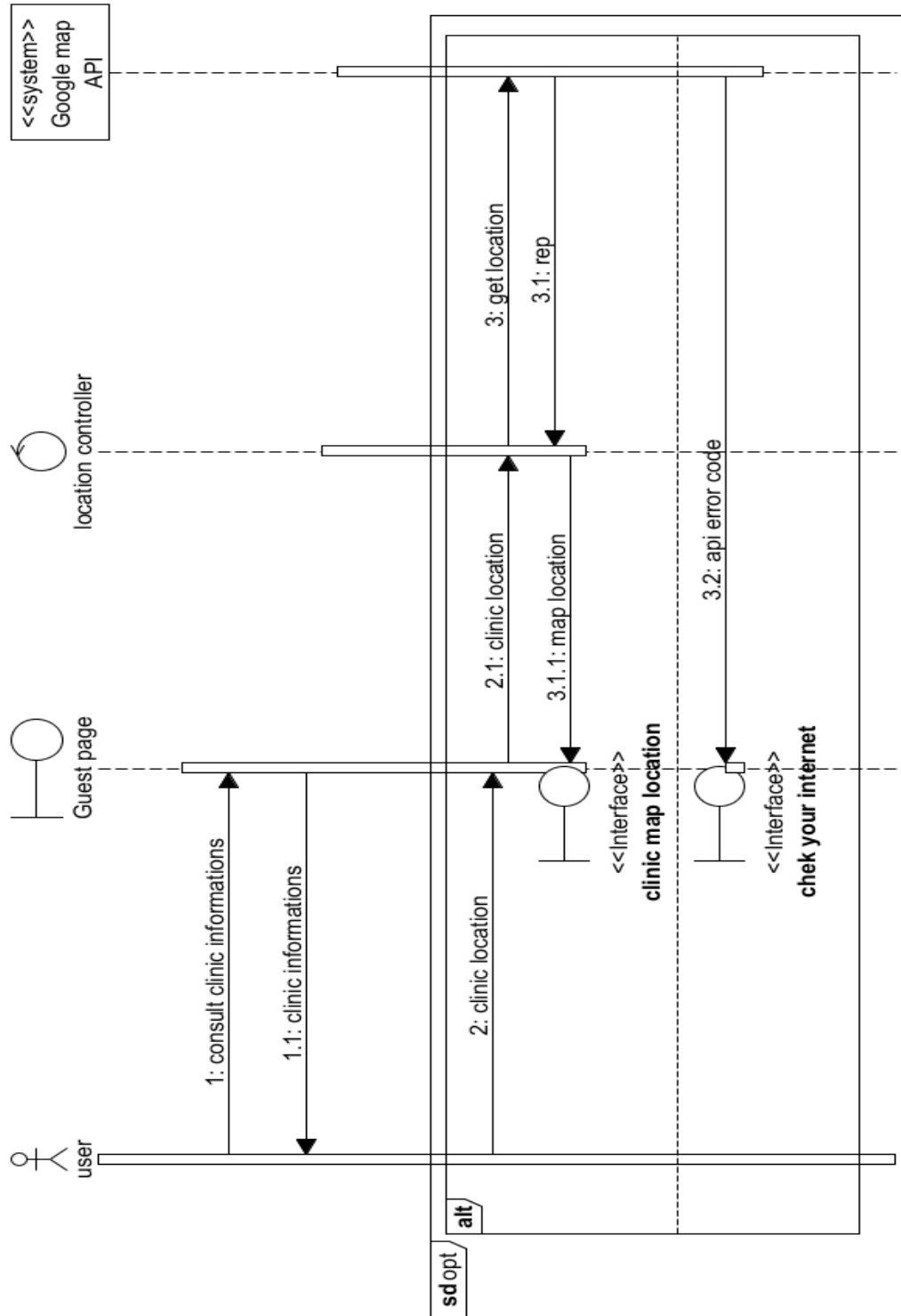


Figure 5.19: Detailed sequence diagrams of the use case: Consult clinics information

**Consult available specialties and doctors:**
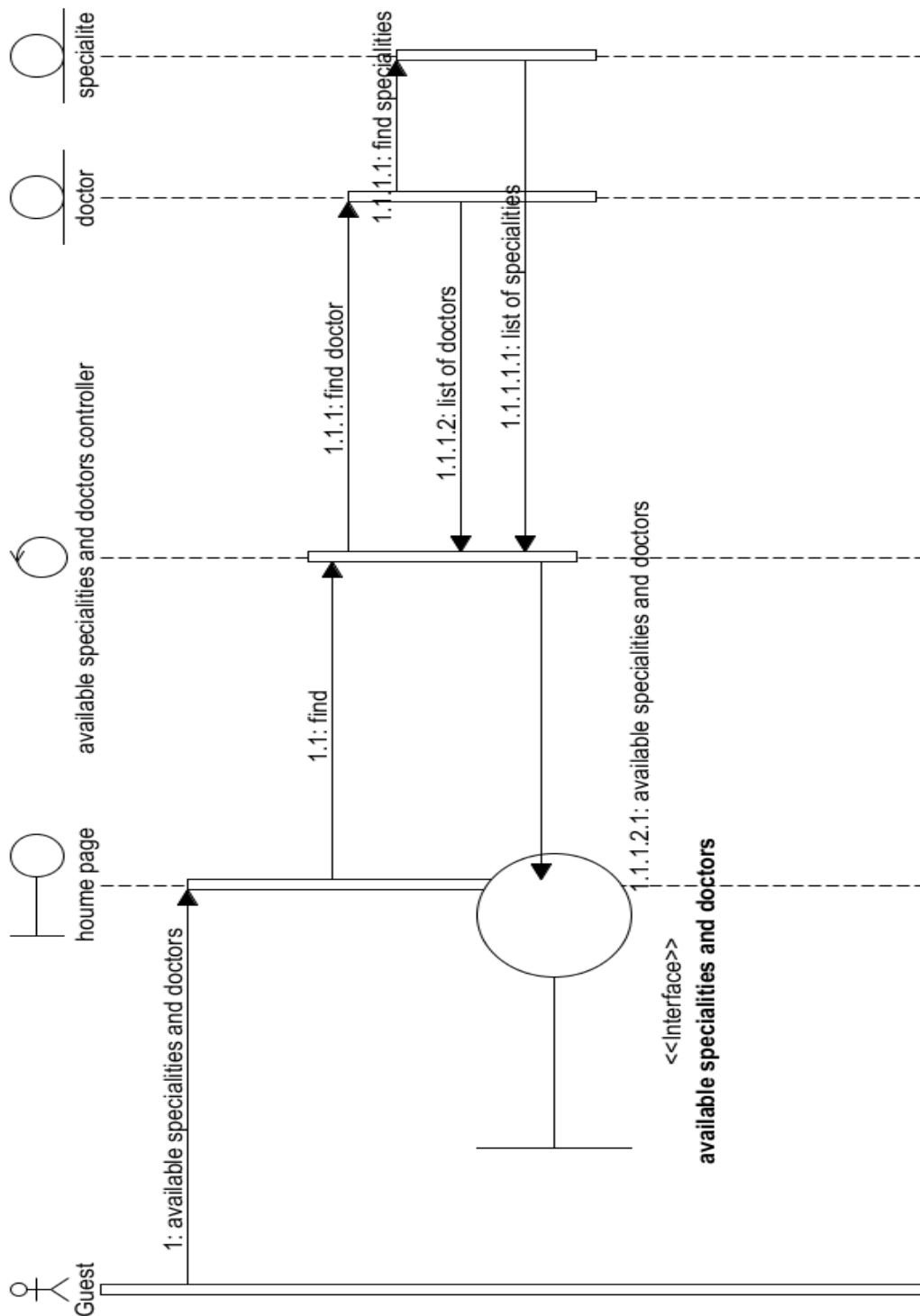


Figure 5.20: Detailed sequence diagrams of the use case: Consult available specialties and doctors

**Consult profile:**
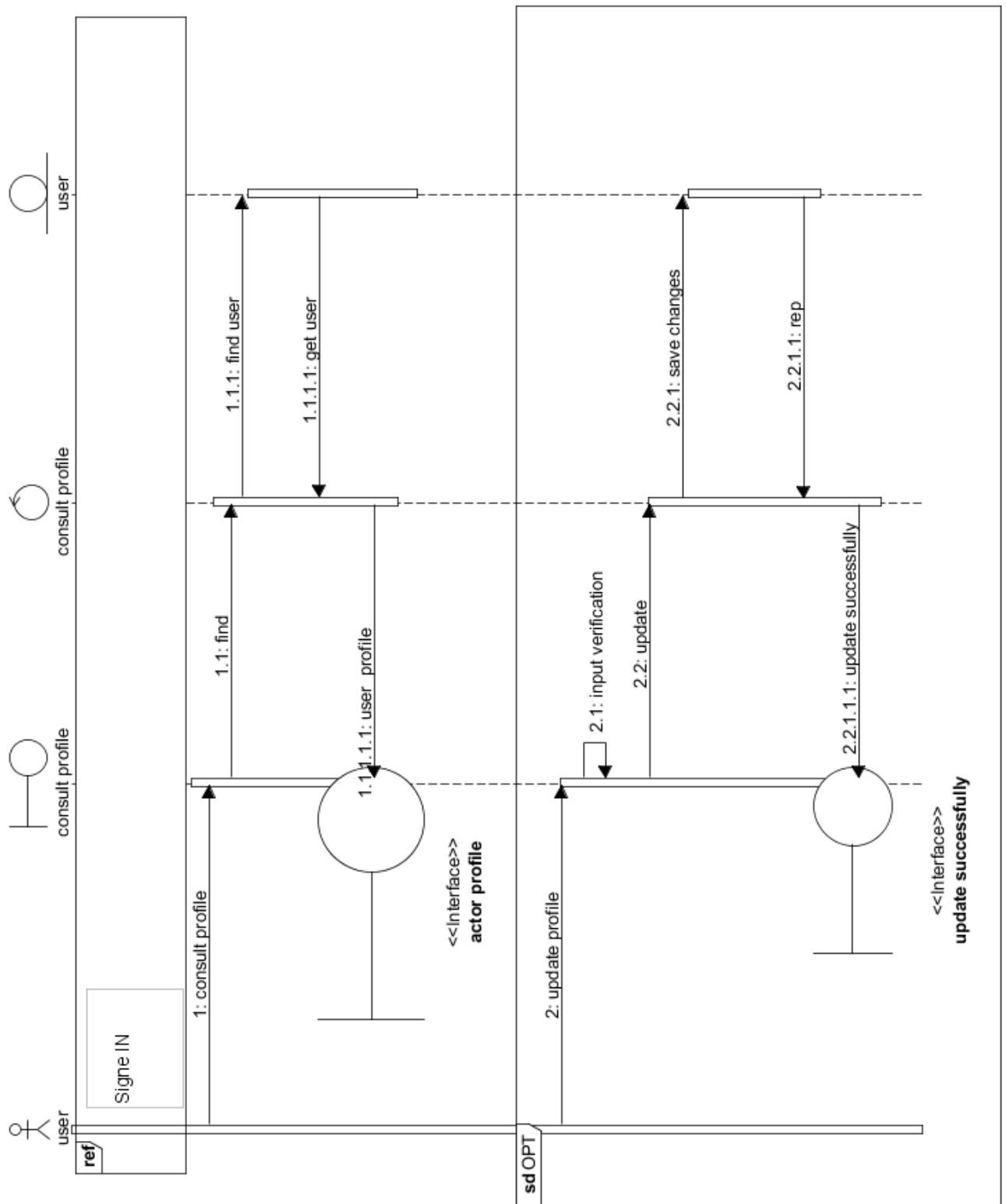


Figure 5.21: Detailed sequence diagrams of the use case:
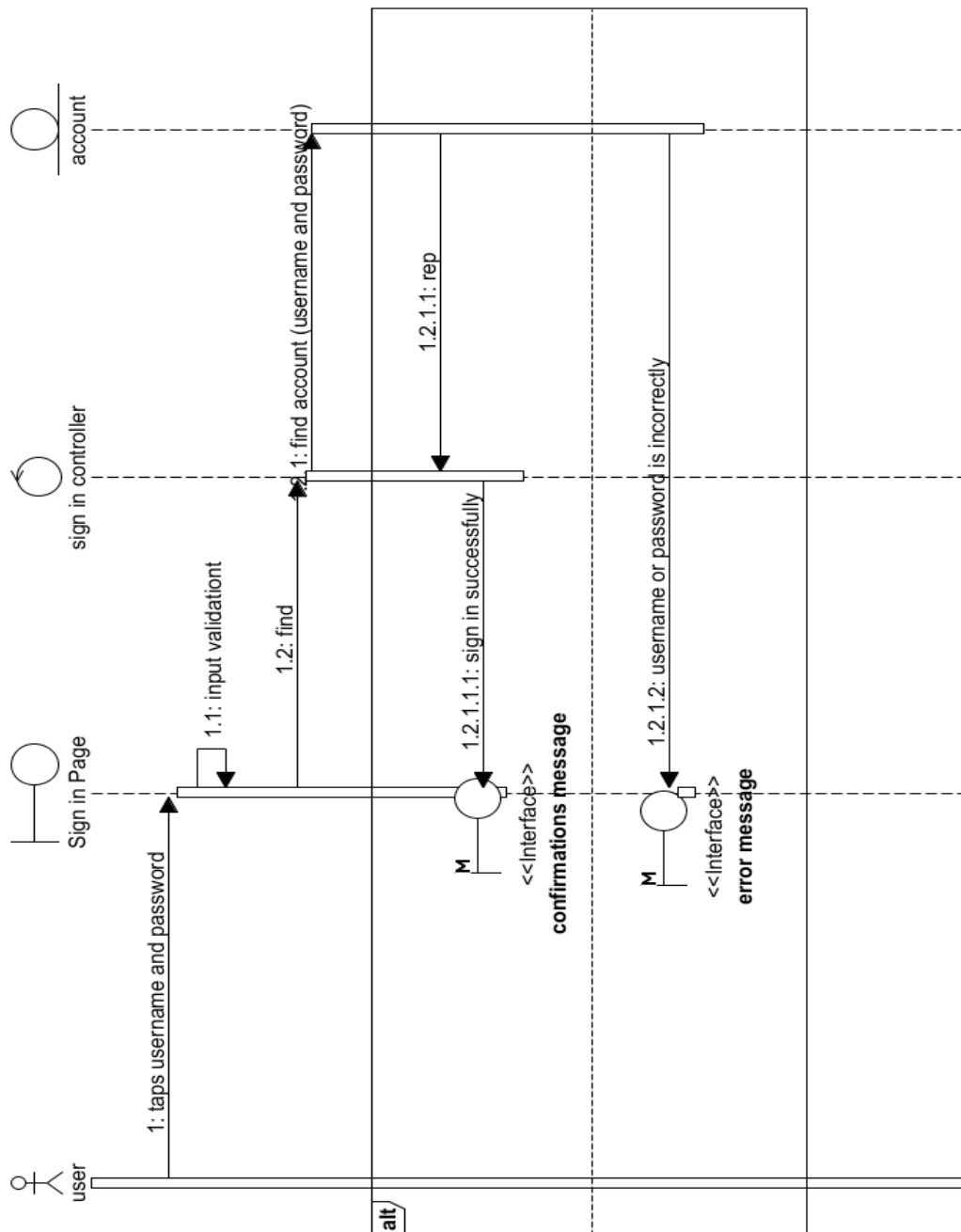
**Sign in:**

Figure 5.22: detailed sequence diagrams of the use case: sign in
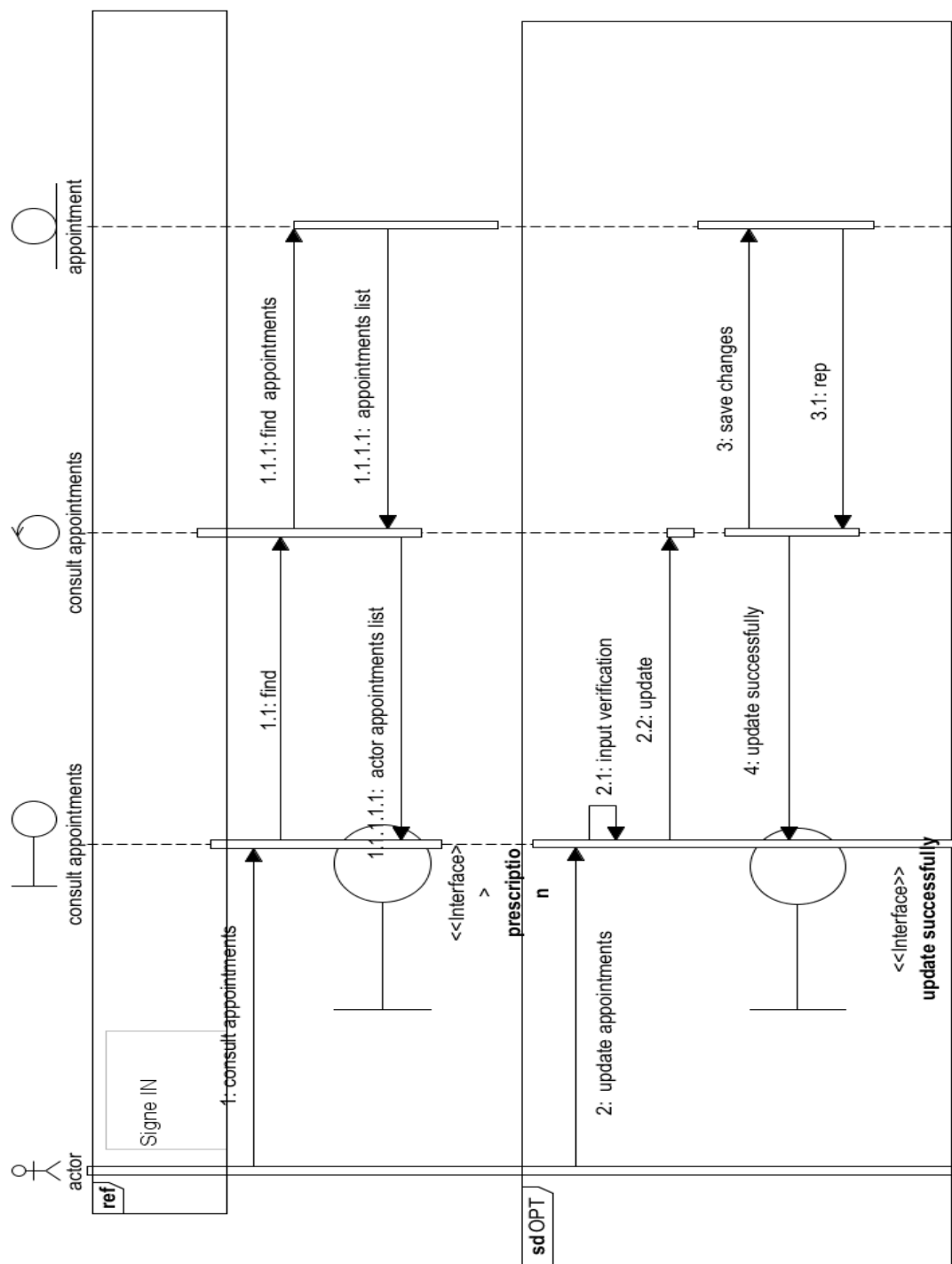
**Consult medical record:**
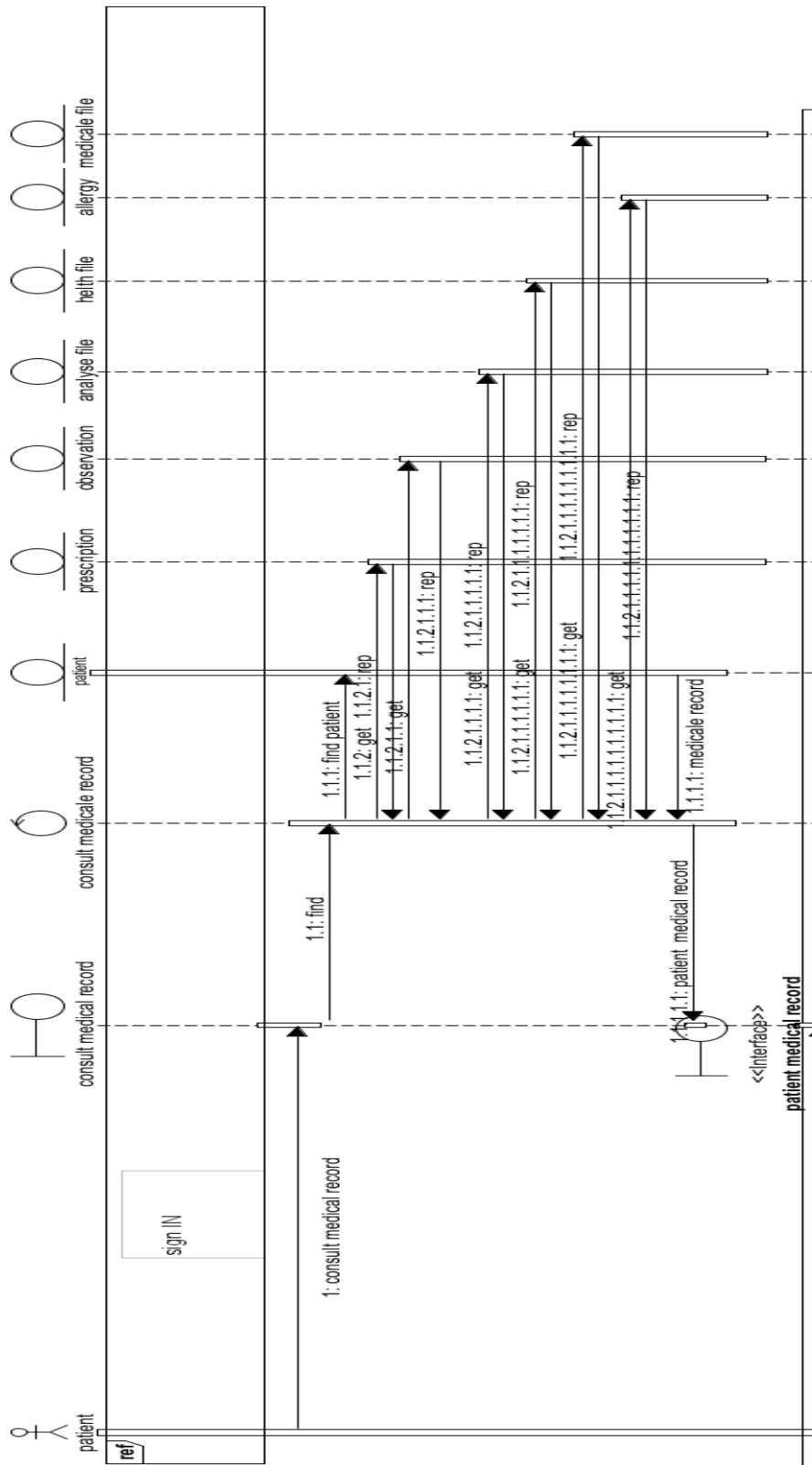


Figure 5.23: detailed sequence diagrams of the use case: consult medical record part 01
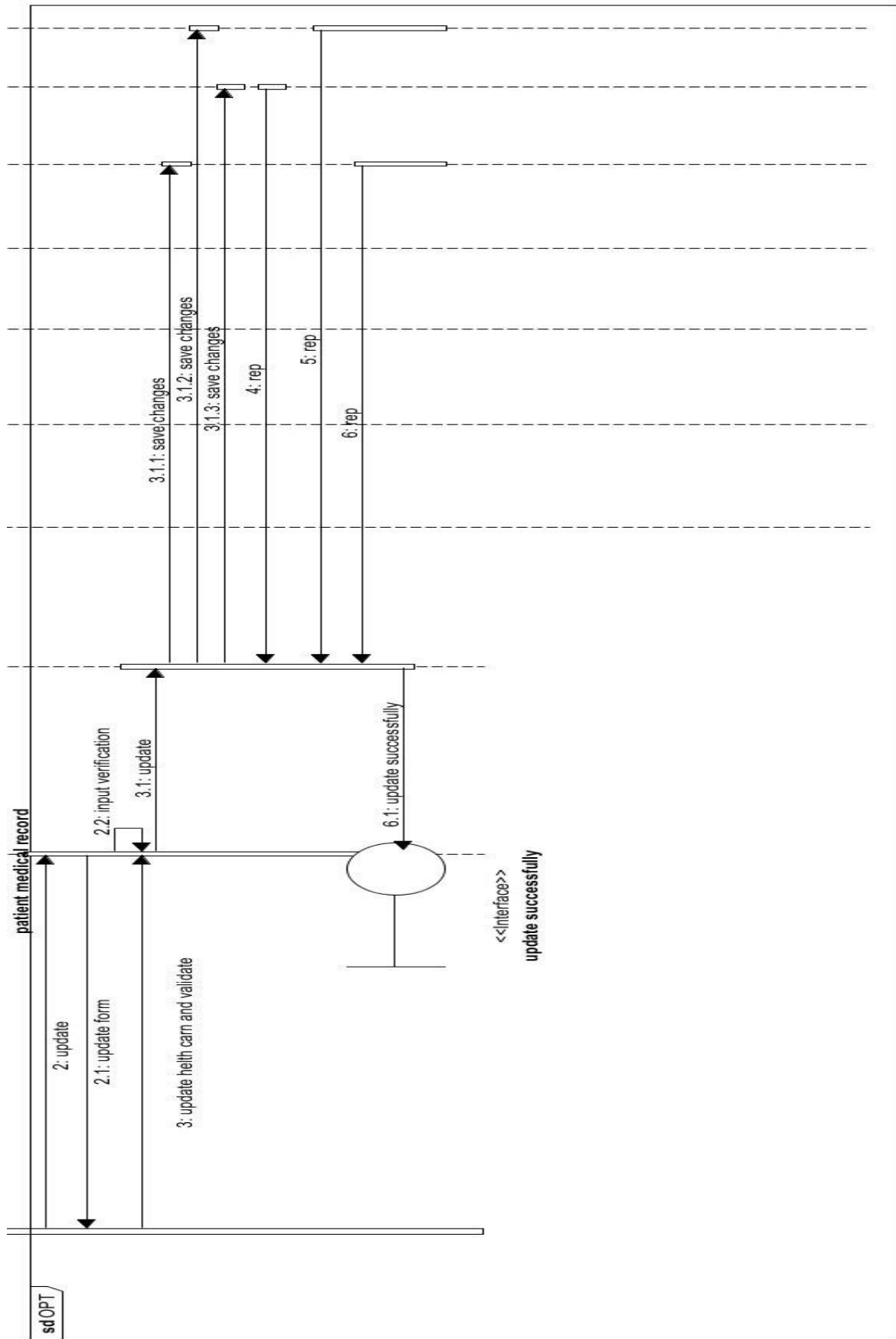
Figure 5.24: detailed sequence diagrams of the use case: consult medical record part 02

**Predict the disease:**
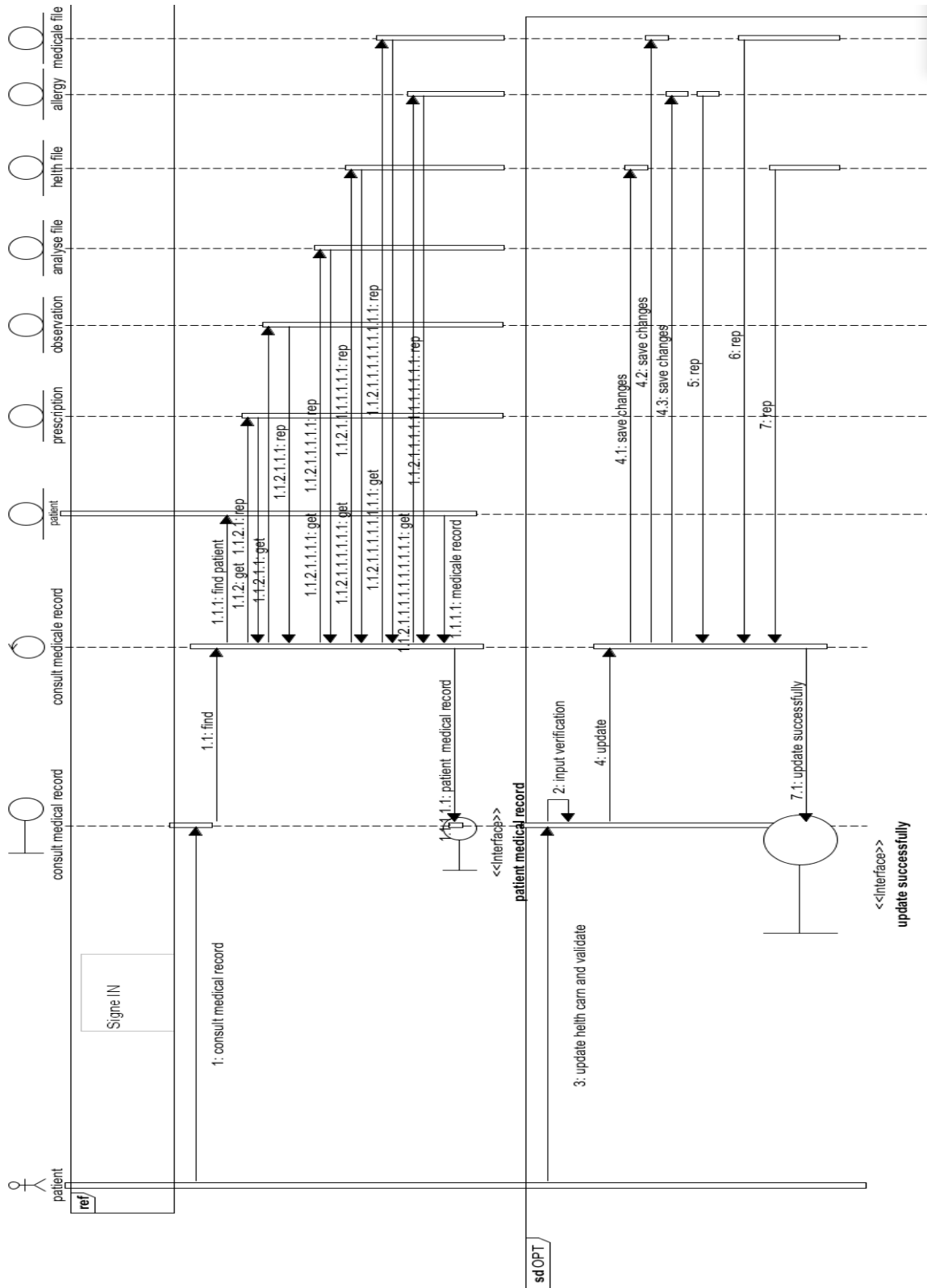


Figure 5.25: detailed sequence diagrams of the use case: predict the disease

**Consult appointment:**



Figure 5.26: detailed sequence diagrams of the use case: Consult appointment part 1
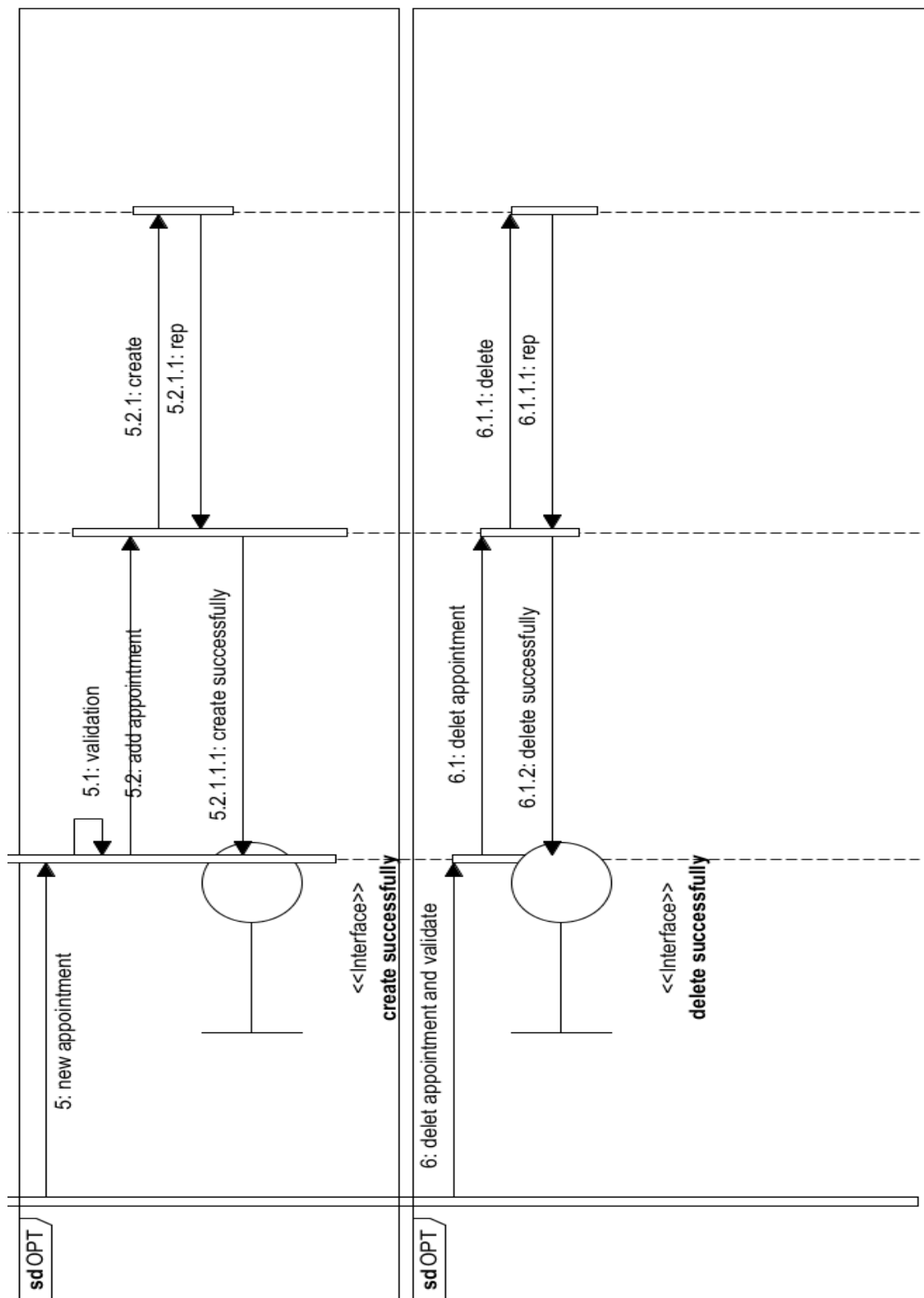
Figure 5.27: detailed sequence diagrams of the use case: Consult appointment part 2
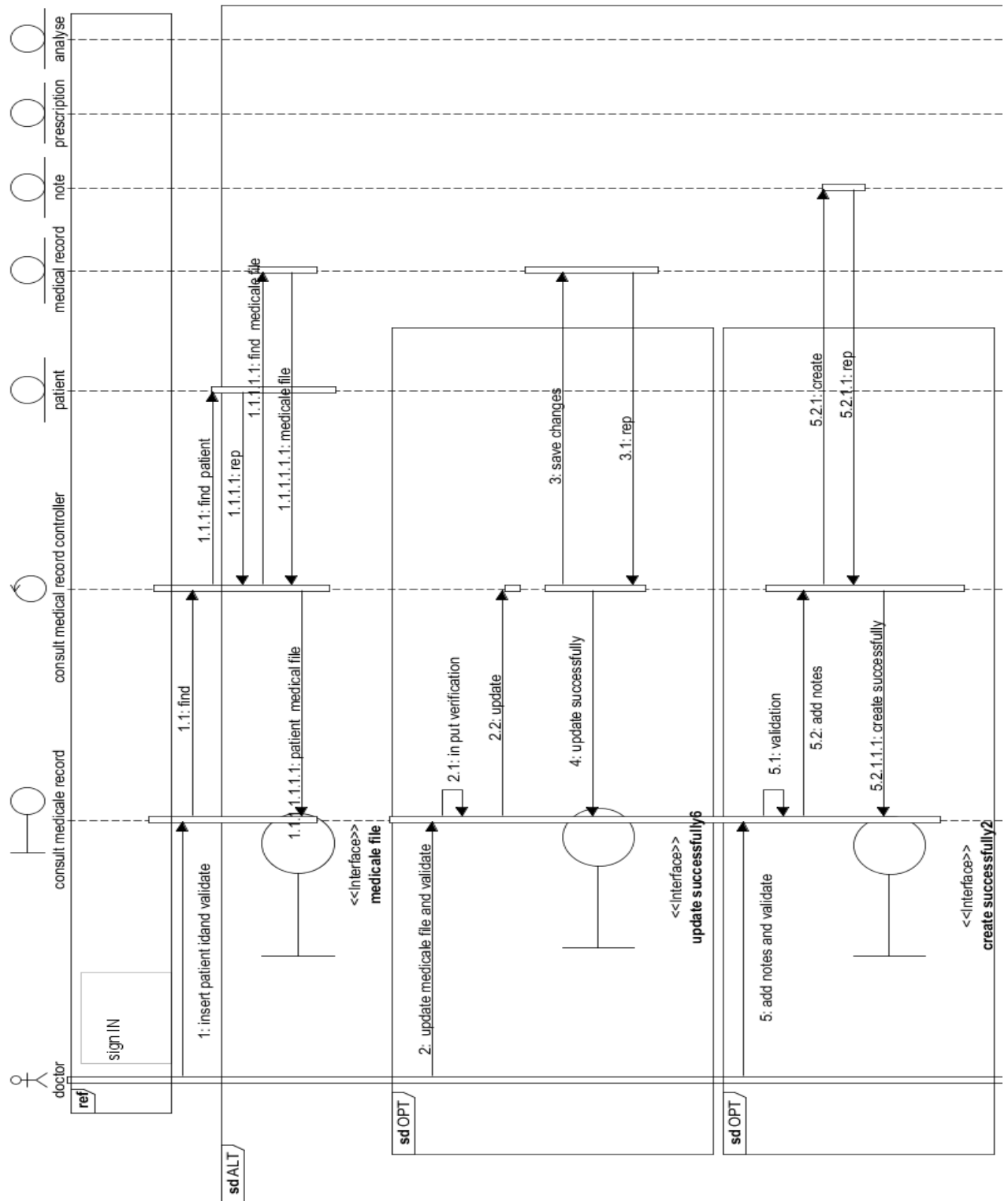
**Consult medical record by doctor**



Figure 5.28: detailed sequence diagrams of the use case: consult medical record by doctor part 1
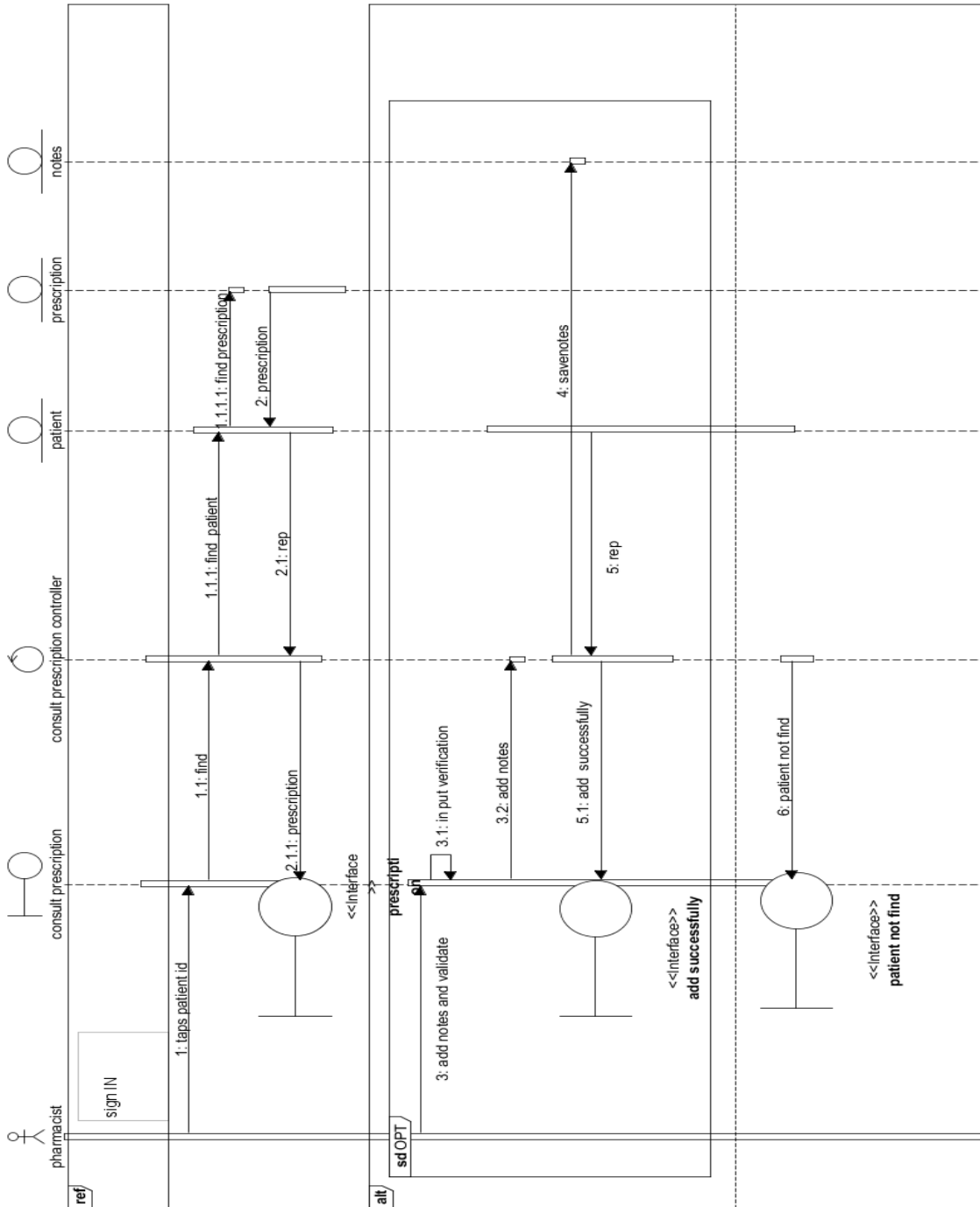
**Consult prescription:**



Figure 5.29: detailed sequence diagrams of the use case: Consult prescription
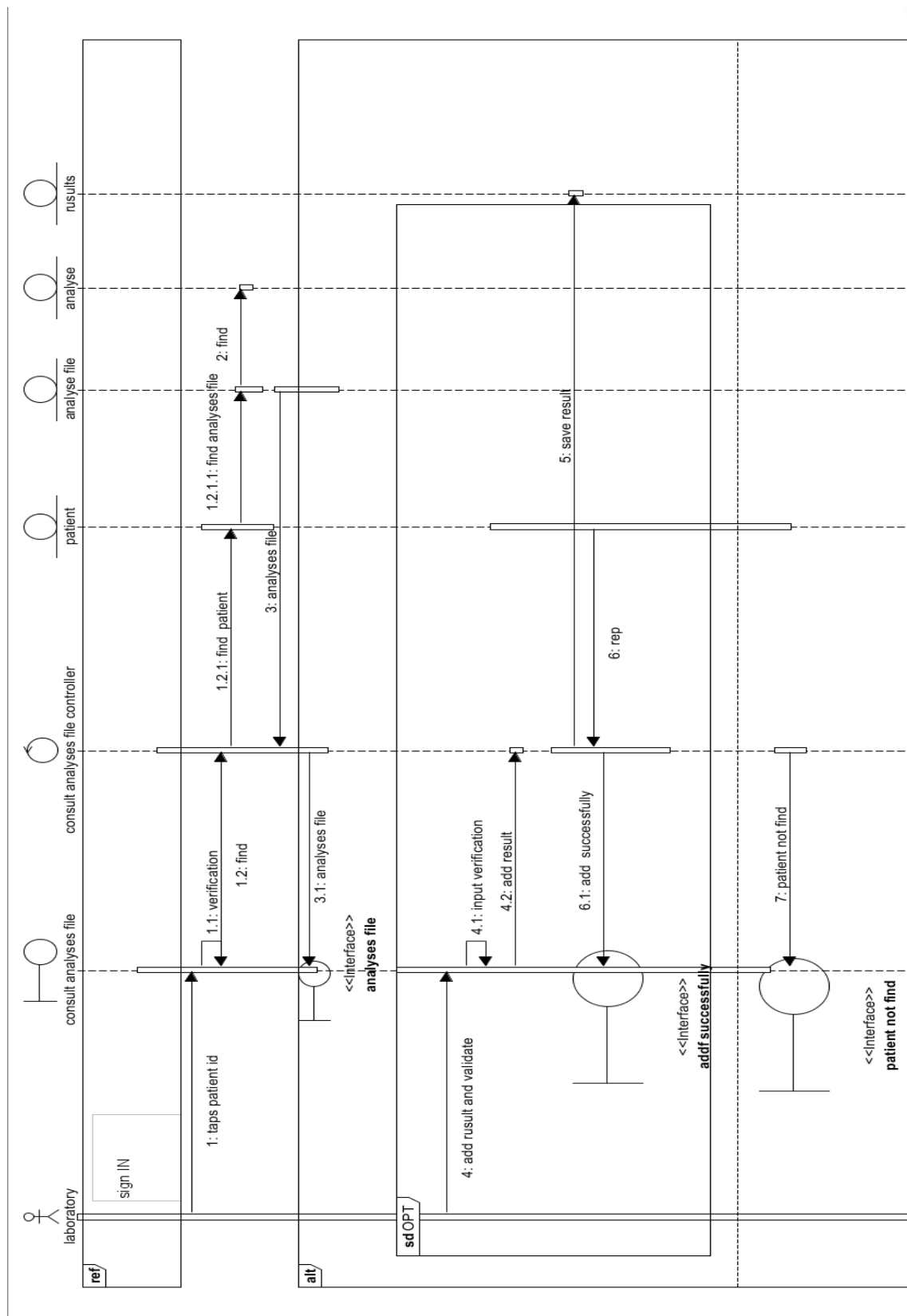
**Consult analysis:**



Figure 5.30: detailed sequence diagrams of the use case: Consult analysis
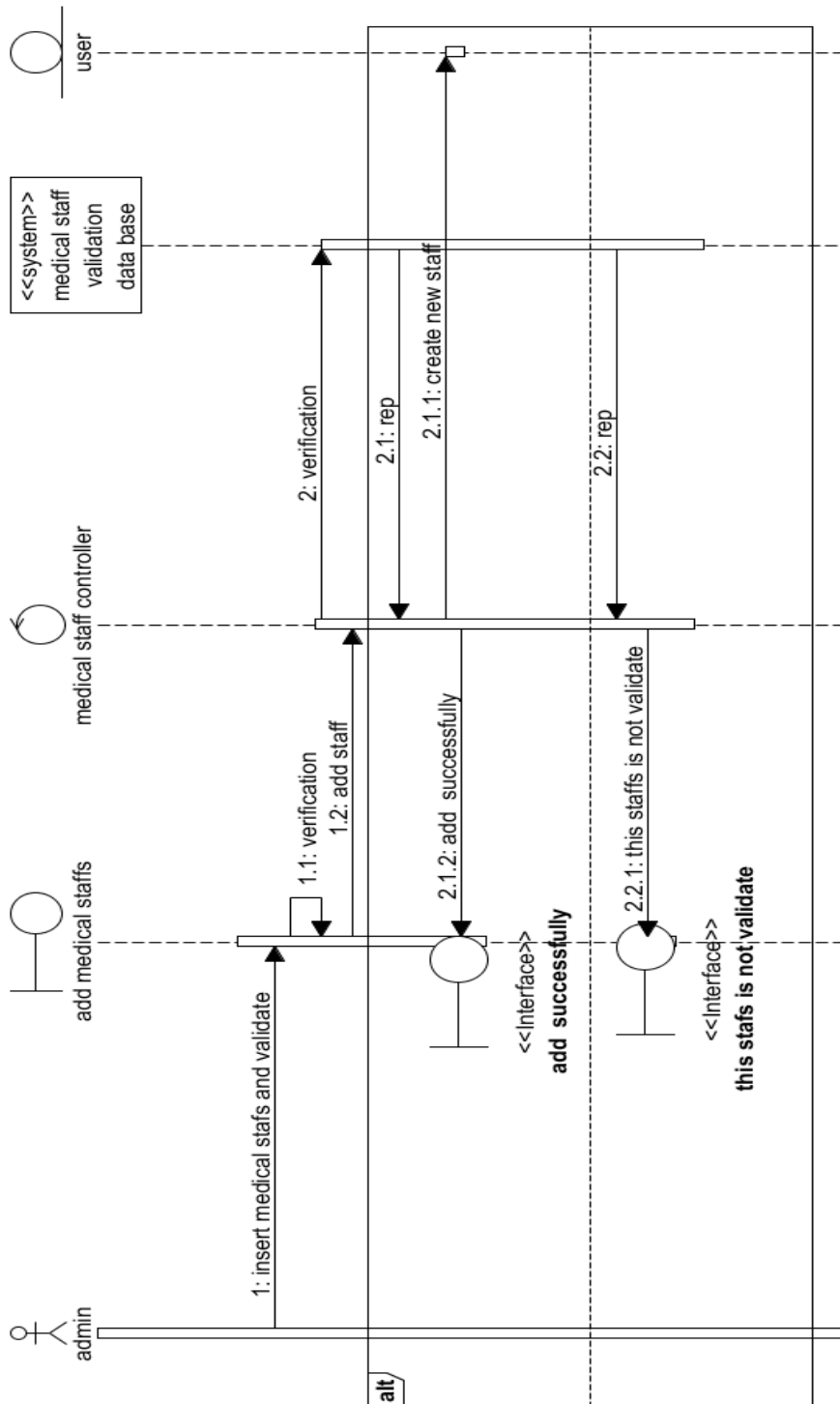
**Add medical staff member:**



Figure 5.31: detailed sequence diagrams of the use case: Add medical staff member
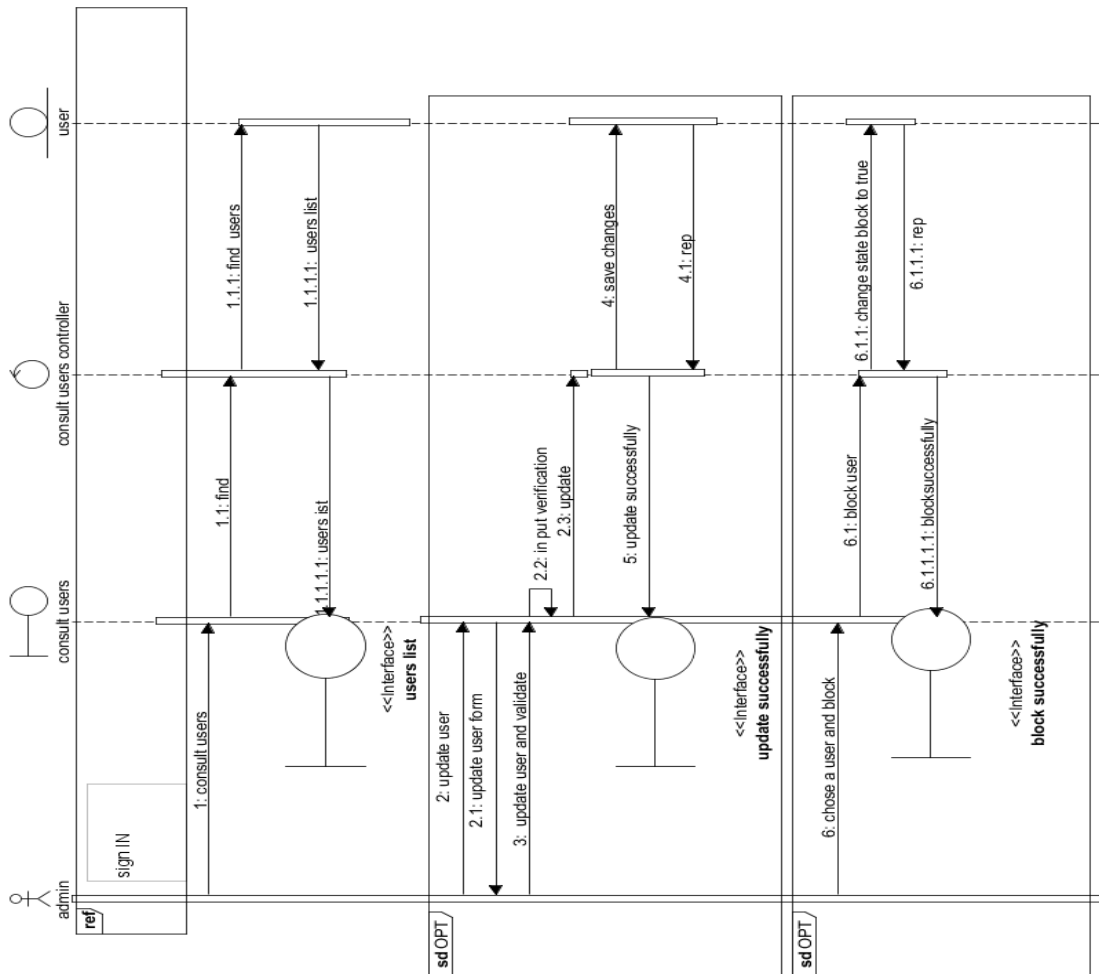
**Consult users:**

Figure 5.32: detailed sequence diagrams of the use case: Consult users
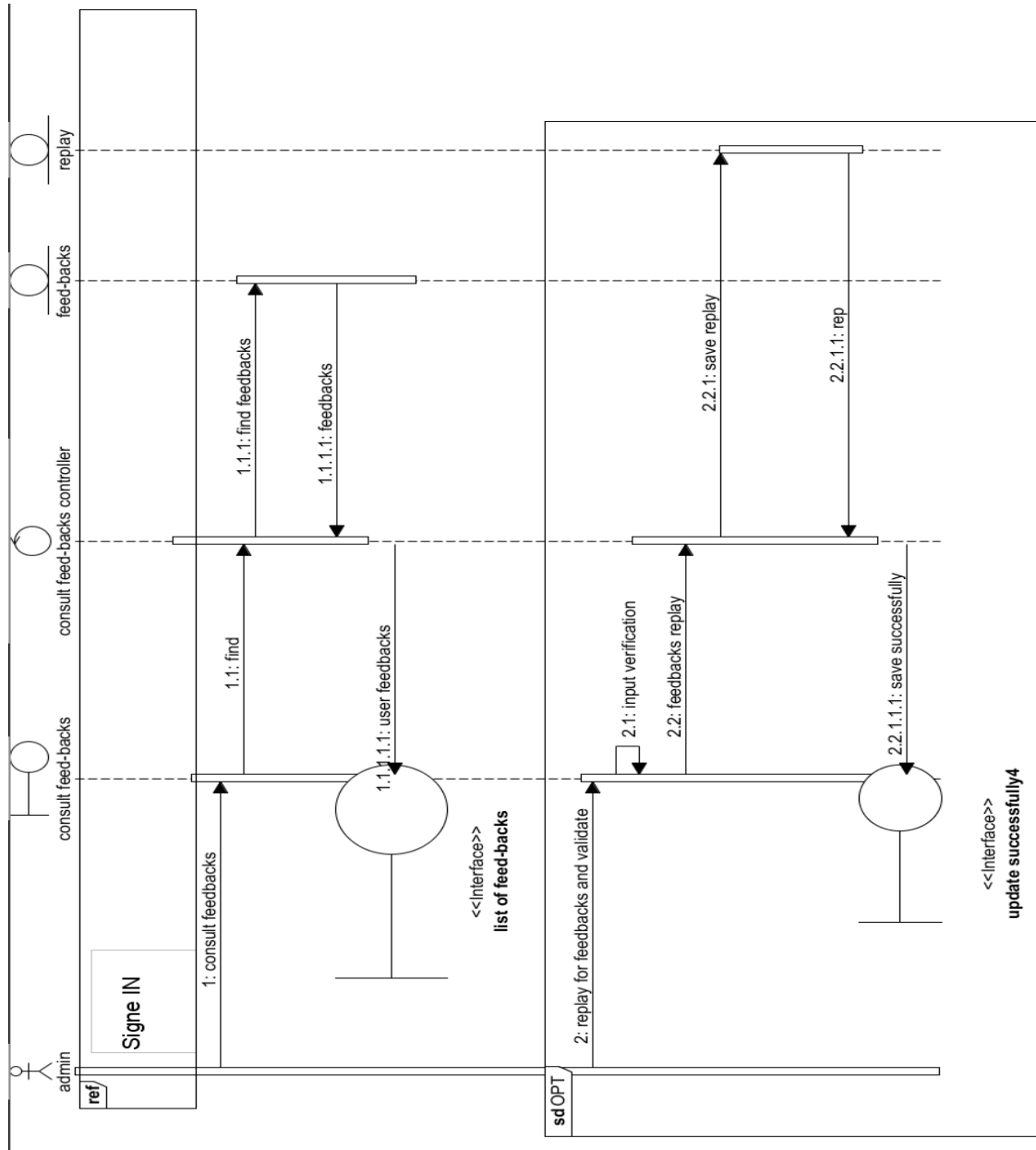
**Consult feed-backs:**



Figure 5.33: detailed sequence diagrams of the use case: Consult feed-backs

## 5.7   Detailed design

The detailed design stage is the second last step before coding. It consists of documentation classes, tables and design choices including specificities of development language or environment. [18]

### 5.7.1   Class diagram:

A class diagram in the Unified Modeling Language (UML) is a type of static structure diagram that describes the structure of a system by showing the system's classes, their attributes, operations (or methods), and the relationships among objects [36].

## 5.8   Transition to relational model

Transition to relational model To move from object model to relational model, a representation between the structure of classes and the structure of relational data is imposed while a class defines a data structure to which instances subscribe, it therefore corresponds to a table of the relational model

### 5.8.1   Relationship class diagram:

Relationships in class diagrams show the interaction between classes and classifiers. Such relationships indicate the classifiers that are associated with each other, those that are generalizations and realizations, and those that have dependencies on other classes and classifiers[36].
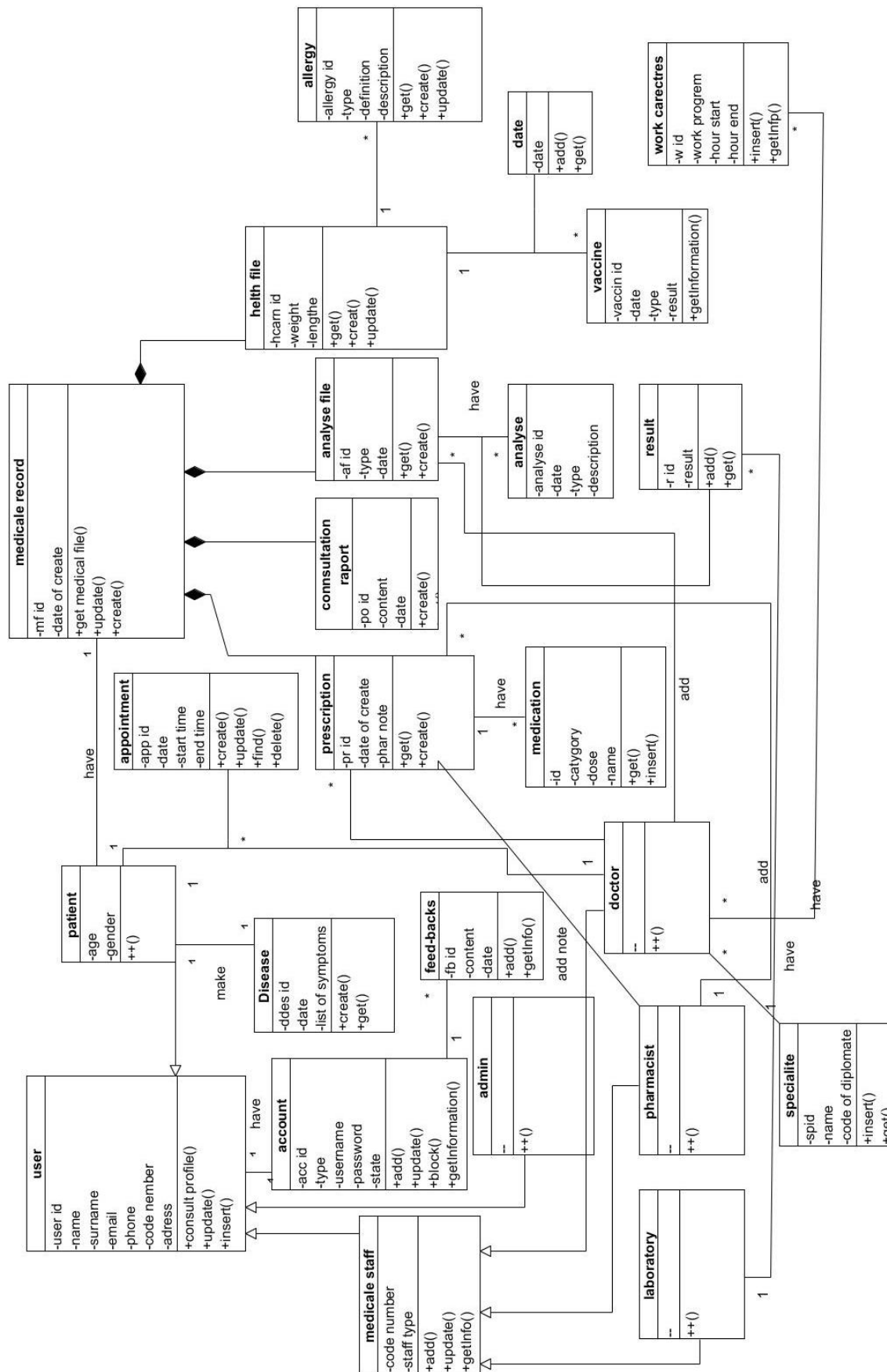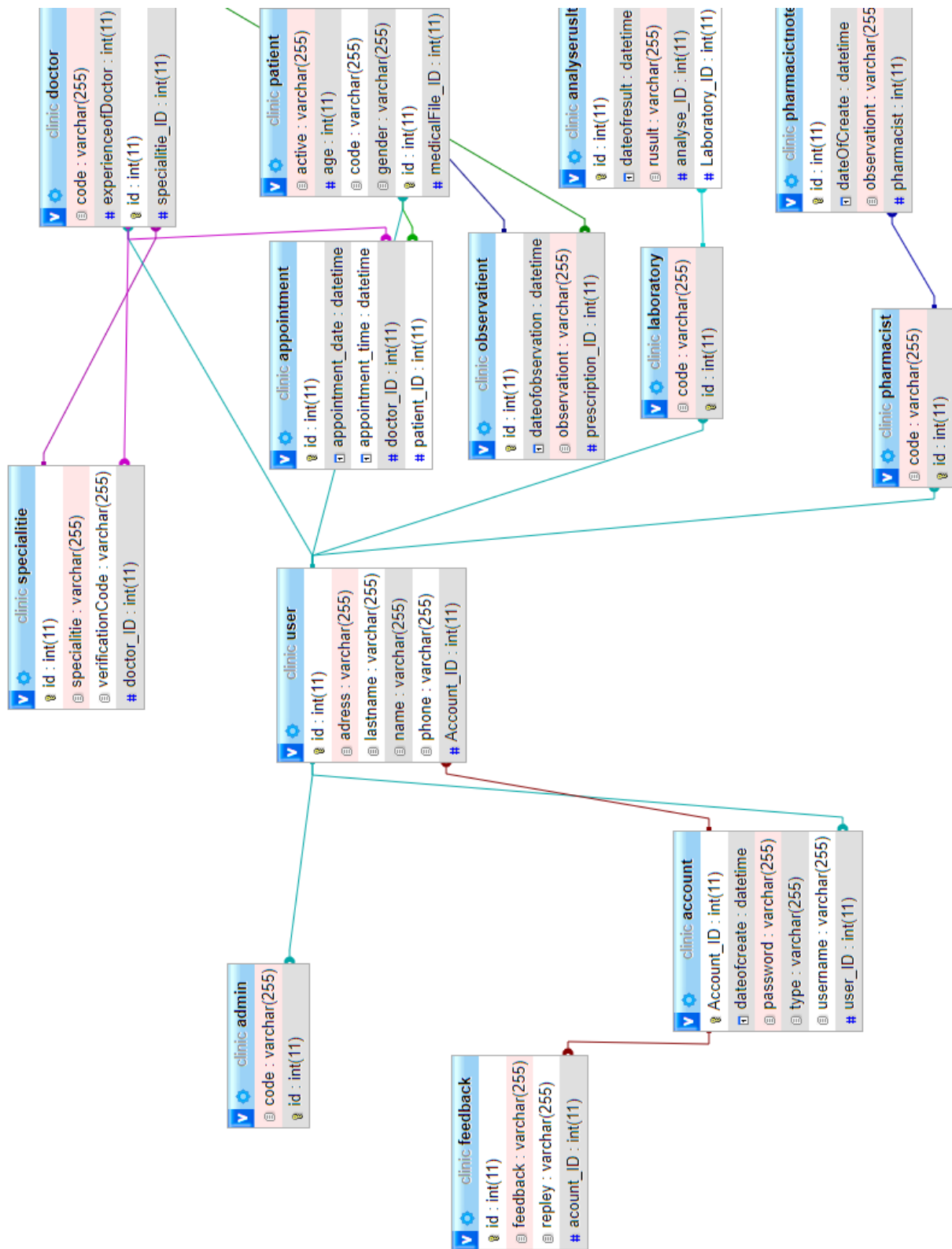
Figure 5.34: Class diagram

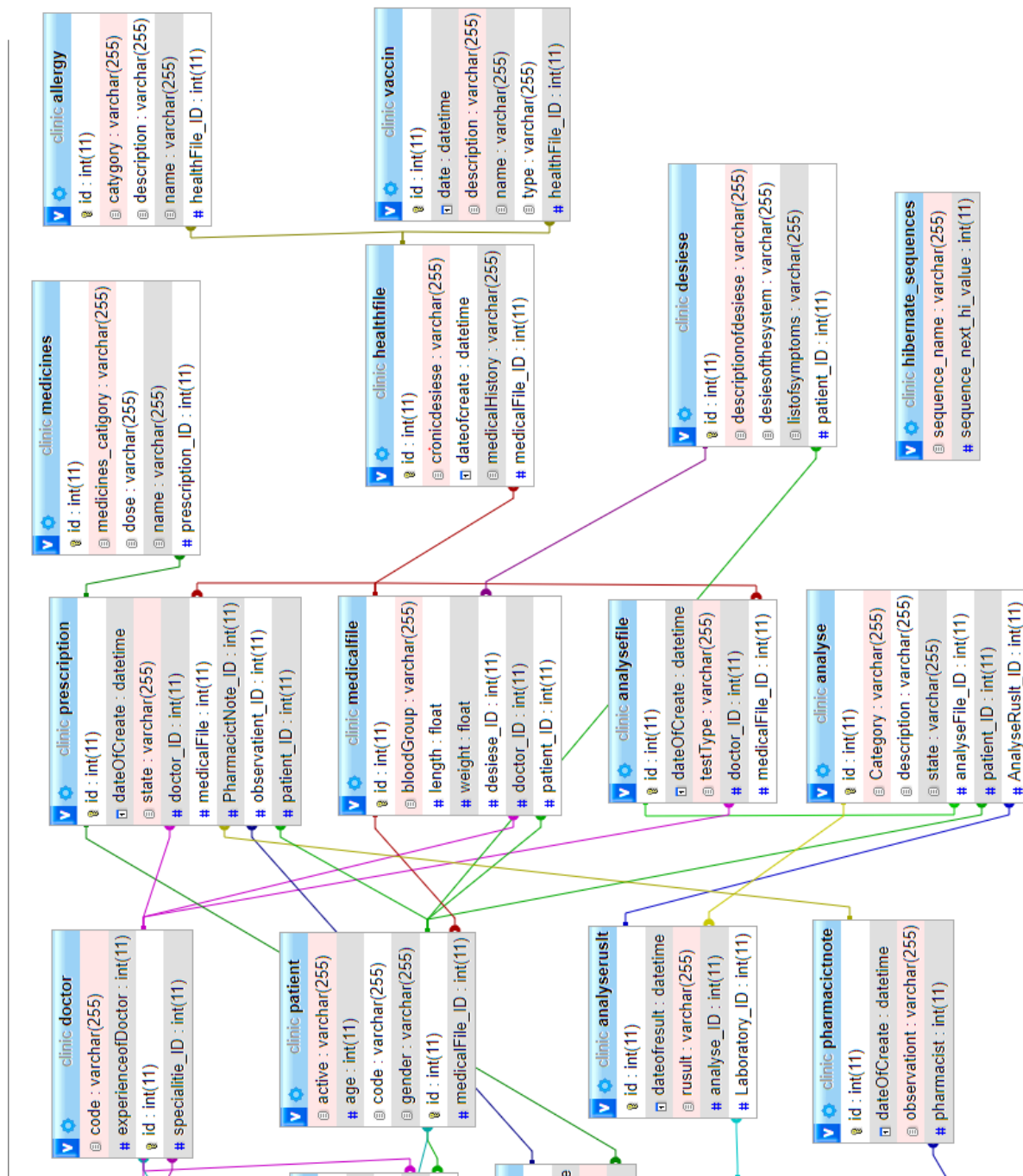Figure 5.35: Relationship class diagram Pt1

Figure 5.36: Relationship class diagram Pt2

## 5.9  Conclusion

In this chapter we tried to do a general idea for our objective and put a basic structure for the following chapters.

# Chapter 6

# Realisation

## 6.1   Introduction

After having done the proper design for our application, we will in this part describe the process of achieving our application. This is done by specifying the development environment and an overview of the interfaces of our application.

## 6.2   Used tools

In order to develop our smart prediction health system a set of tools is needed:

### 6.2.1   Latex:

LaTeX is a description language that can be used for writing and typesetting documents. It is especially useful to write mathematical notation such as equations and formulae.

It is a layout language that makes it possible to produce documents of high typographical quality and rigorously homogeneous in their presentation. This language is to be compared to (X)HTML, in that it allows to mark a text, then neatly cut into sections, subsections, paragraphs, titles or footnotes. The text with its tags is then compiled, to be formatted according to the model specifications chosen by the user. [33]



Figure 6.1: Latex logo

### 6.2.2 Overleaf:

Collaborative, cloud based environment for sharing and working with La-TeX.

In addition, it offers the possibility to write documents in a collaborative way, to offer its documents directly to different publishers (IEEE Journal, Springer, etc.) or open archive platforms (arXiv, engrxiv, etc.) for a possible publication. This platform is very compatible with various media such as tablets smartphones [35].



Figure 6.2: Overleaf logo

### 6.2.3 Star UML:

StarUML is a sophisticated software modeler aimed to support agile and concise modeling.



Figure 6.3: Star UML logo

### 6.2.4   Sublime Text:

Used to write some codes and can open various types of files. In addition, it also supports various programming languages such as C++, CSS, ASP, HTML, and so on [37].



Figure 6.4: Sublime Text logo

### 6.2.5   HTML/CSS:

HTML is the language for describing the structure of Web pages using markup.
while CSS is the language for describing the presentation of Web pages, including colors, layout, and fonts.



Figure 6.5: HTML and CSS logos

### 6.2.6   Postman:

Postman is an API platform for building and using APIs. Postman simplifies each step of the API lifecycle and streamlines collaboration so you can create better APIs—faster [34].

Figure 6.6: Postman logo

### 6.2.7 Visual Paradigm Enterprise:

Visual Paradigm Enterprise is an ArchiMate enterprise architecture tool certified by The Open Group. It supports all the vocabulary, notation, syntax, and semantics for all ArchiMate language elements and relationships [36].



Figure 6.7: Visual Paradigm Enterprise logo

### 6.2.8 Azure ML:

Azure Machine Learning is a cloud service for accelerating and managing the machine learning project lifecycle. Machine learning professionals, data scientists, and engineers can use it in their day-to-day workflows: Train and deploy models, and manage MLOps [38].



Figure 6.8: Azure ML logo

### 6.2.9 Java EE:

(Java Platform, Enterprise Edition) An application software platform from Oracle based on the Java programming language. Originally developed by Sun, which Oracle acquired in 2010, Java EE services are performed in the middle tier between the user's machine and the enterprise's databases and legacy information systems. Java EE comprises a specification, reference implementation and set of testing suites. Its core component is Enterprise JavaBeans (EJBs), followed by JavaServer Pages (JSPs) and Java servlets and a variety of interfaces for linking to the information resources in the enterprise [39].



Figure 6.9: Java EE logo

### 6.2.10 Xampp:

XAMPP is used as a stand-alone server (localhost) which consists of several programs including Apache HTTP server, MySql database, and language translator is written in PHP and Perl programming languages. This program is available GNI (General Public License), is a server that is easy to use, and can display dynamic web pages [37].

Figure 6.10: Xampp logo

## 6.3   Presentation of the interfaces of the website

**Home page**

This web page shows our project home page. All users can visit, It is the first interface presented to them.
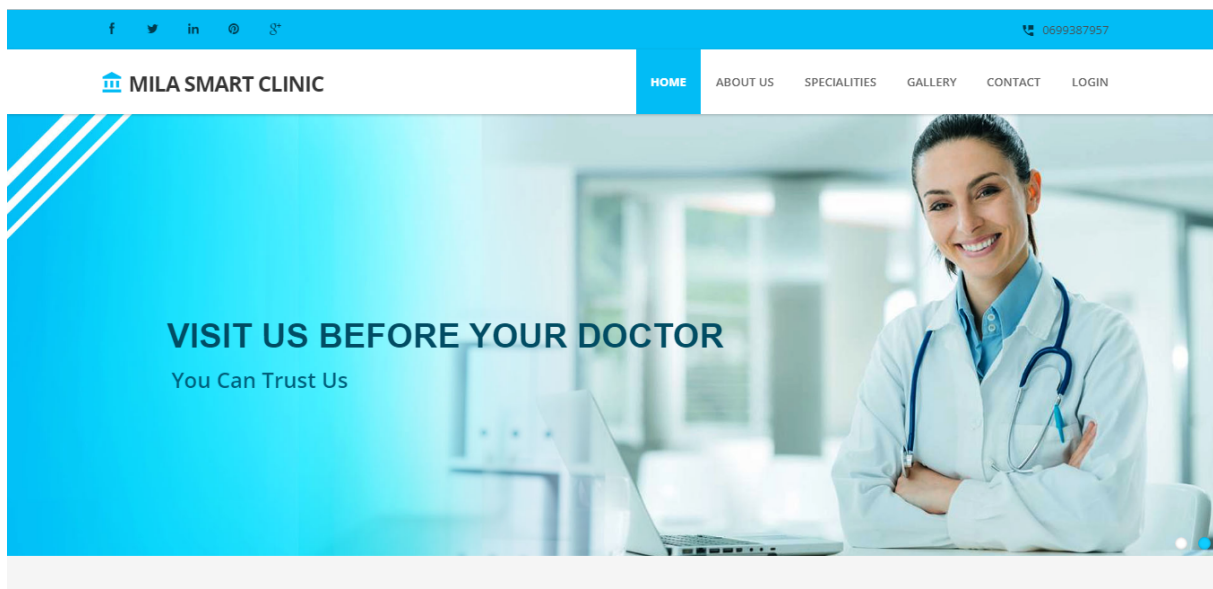

Figure 6.11: Home page

**Services page**

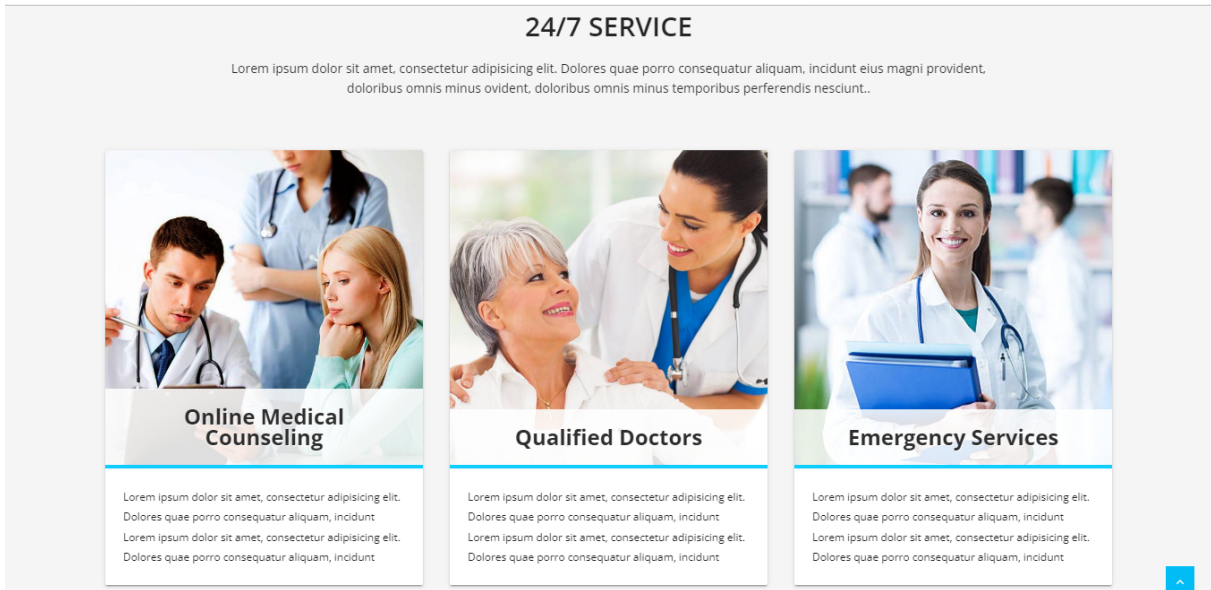Presents the different services offered by the clinic



Figure 6.12: Services page

**About us page**

shows all informations about our clinic, administration, and medical staff.
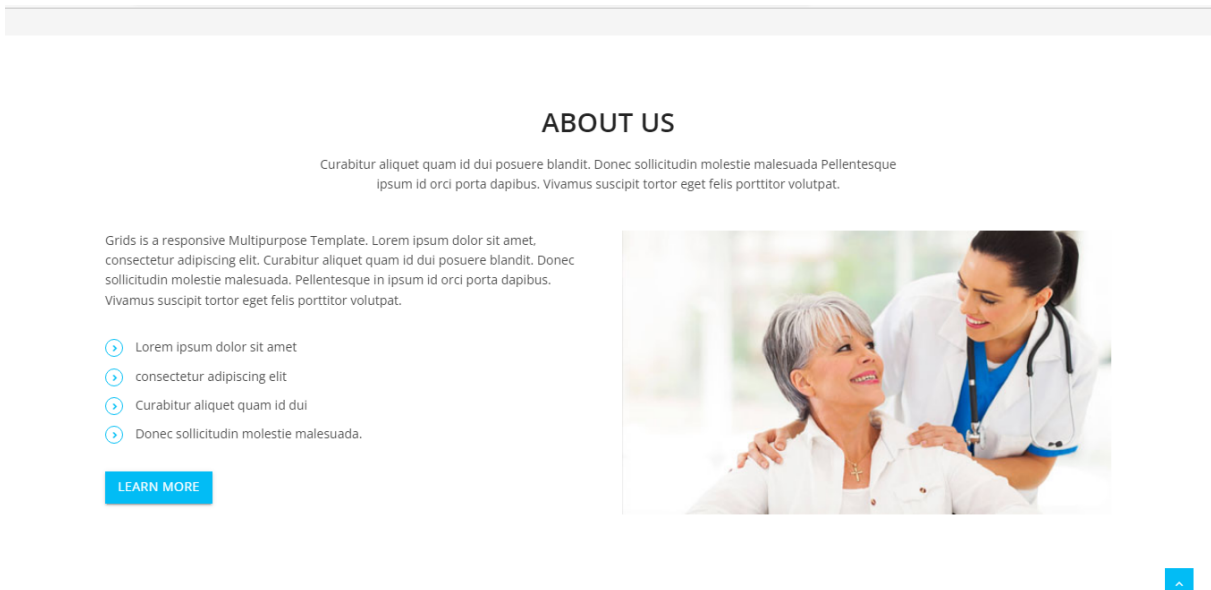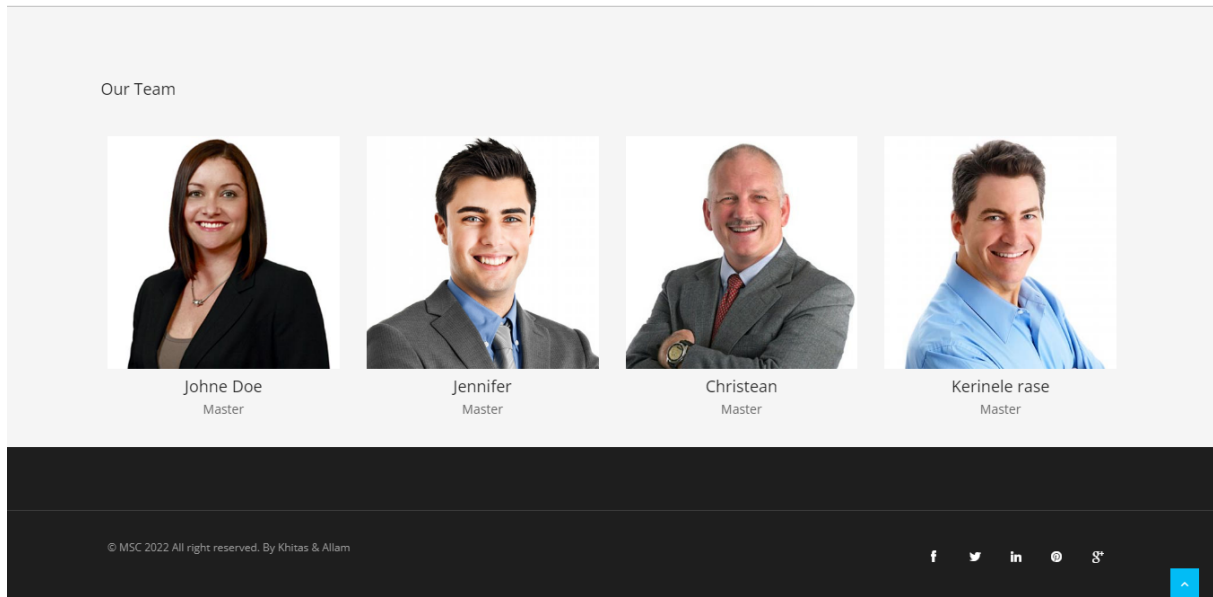


Figure 6.13: About us page Pt1

Figure 6.14: About us page Pt2

**Specialties page**

This page presents the specialties included in our clinic.
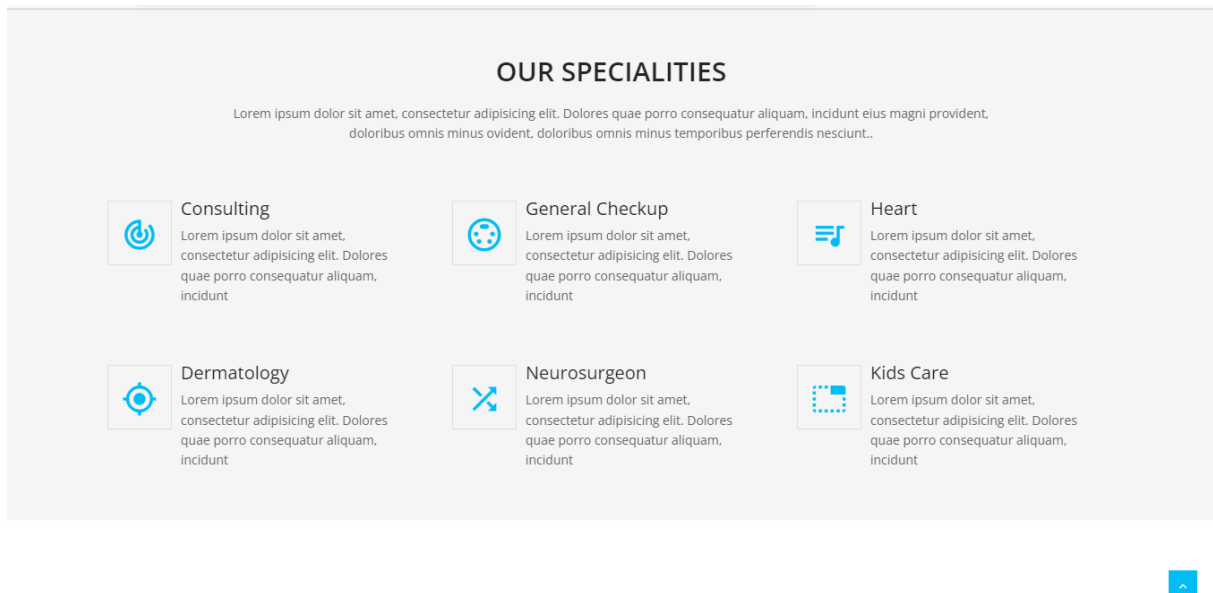


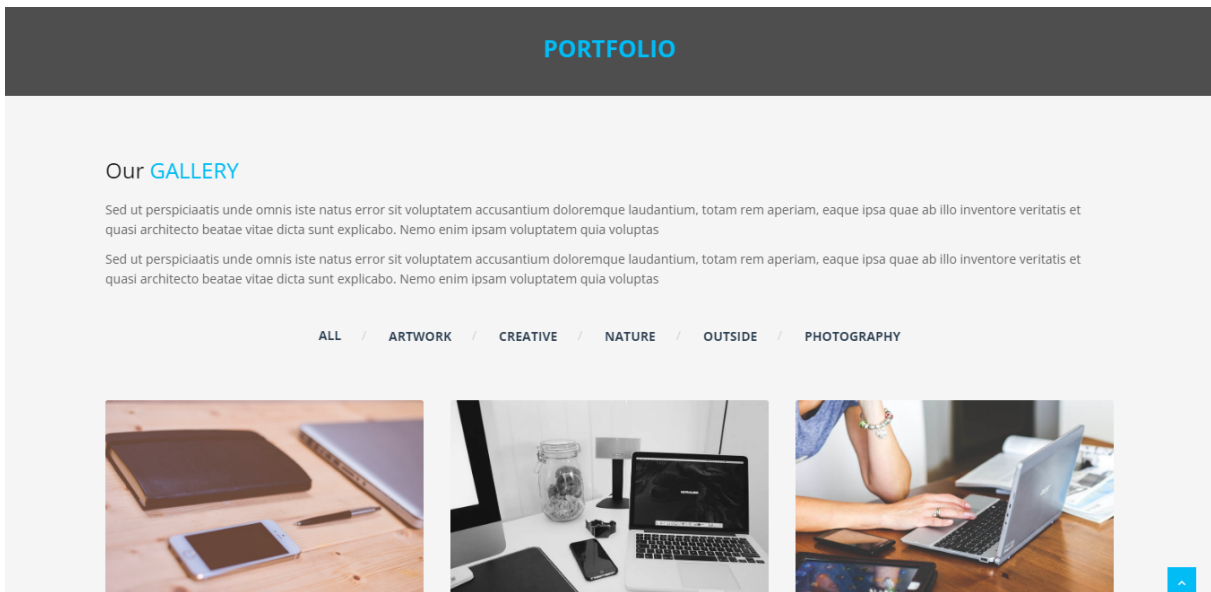Figure 6.15: Specialties page

**Gallery page**



Figure 6.16: Gallery page

**Contact us page**

From this page any user of the web application could reach us, through google maps or any way he chooses to.
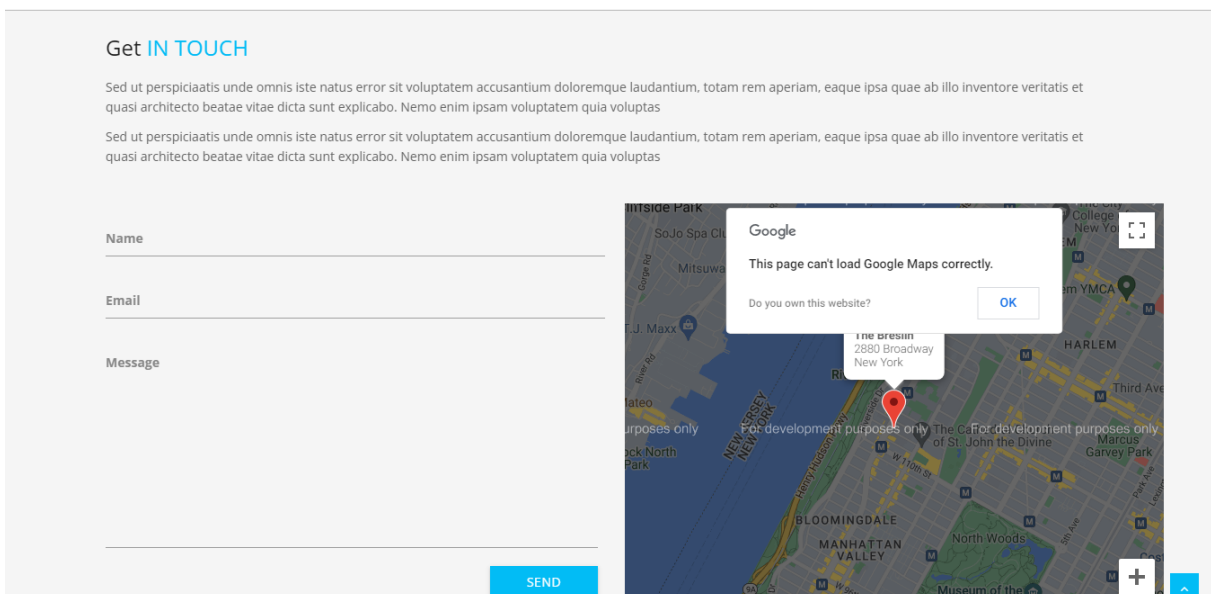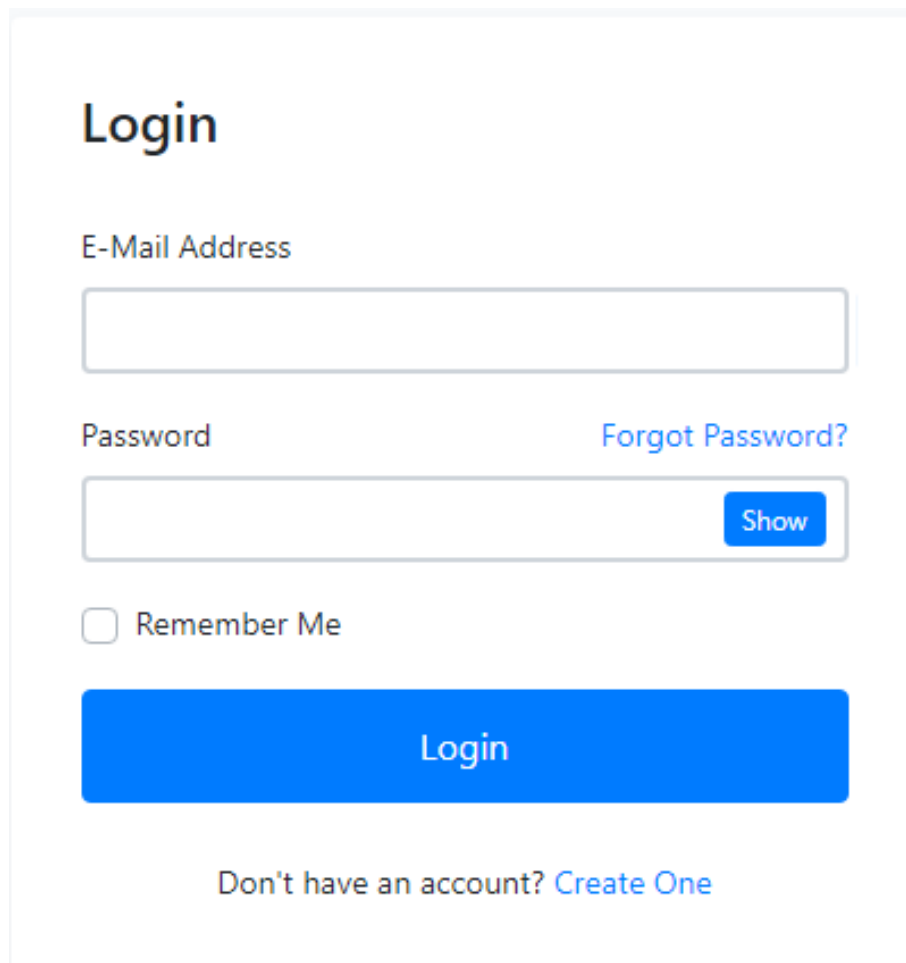


Figure 6.17: Contact us page

**Login page**
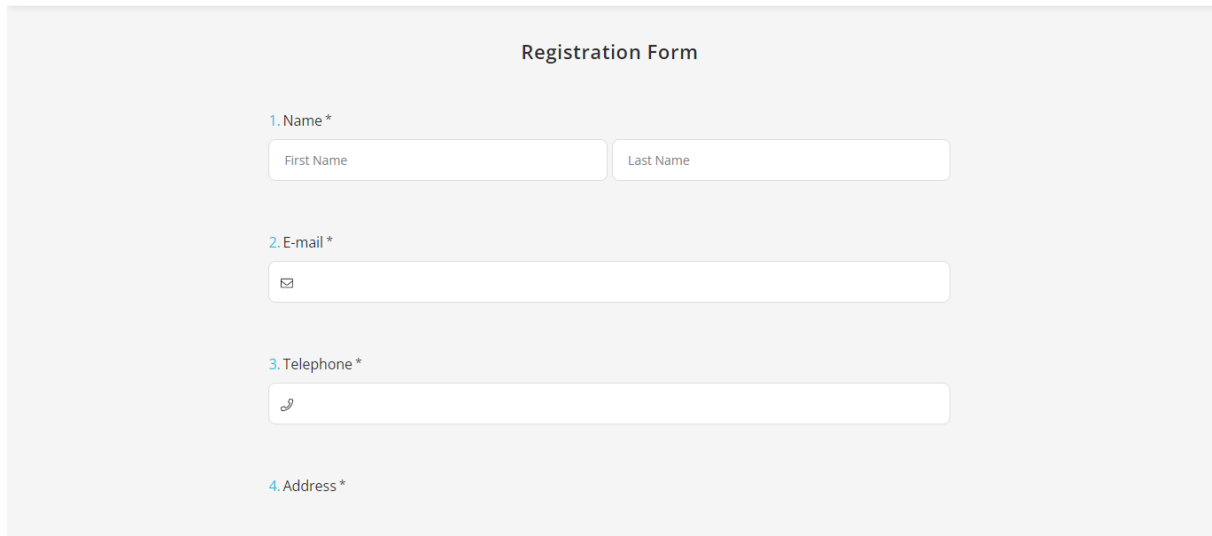


Figure 6.18: Login page

**Patient registration page**



Figure 6.19: Patient registration page

## 6.4   Conclusion

In this last chapter we put all the pieces together in order to form our final solution, and put some of the tools that we used besides the final interfaces of our system.

# Conclusion

# General Conclusion

Our work has been carried out as part of the graduation project. This work has as its main objective the prediction of diseases and patient orientation. For this we proposed to design and produce a web application to ensure an effective prediction and orientation, and to improve the communication and the exchange of information and documentation between the patients and the medical staff (the doctors, the pharmacists and the laboratory). The application should cover all stages of treatment from the opening of the session and the symptoms entering until the patient finalise his treatment.

For the design of our application, we chose a specialized approach to web development called the "2TUP" 2 track unified process and relies on the UML modeling language in these different phases. Using the 2TUP allowed us to fully understand the issue and model the functional needs of our system. He also provided us with good support during the analysis and design phases. For the realization, we used the technologies offered by the Java language for web development and MySQL as a database management system. This work was an opportunity for us to approach the web development domain and to gain practical experience with a development approach and the UML language, as well as with a range of technologies and tools such as the Java language, the JEE platform (JSPs, Servlets, Tomcat application server), and the netbeans development environment, also some web services and API's. We can say that our objectives have been achieved overall, we have achieved three main points: the realisation of the prediction model, the creation of a patient medical record, the creation of the patient orientation system. thus, as the main perspective for the current state of this project, we want to add a mobile application that do the same work as our web application.

# Bibliography

[1] Joan Cabestany Carlos Perez Albert Sama Daniel Rodriguez-Martın. "Artificial Intelligence Contribution to eHealth Application". In: (2018).https: //www.researchgate.net/publication2&_esc=publicationCoverPdf.

[2] Tom Lawry. AI in Health: A Leader's Guide to Winning in the New Age of Intelligent Health Systems (HIMSS Book Series). 2020. https: //b-ok.africa/book/5645189/bce591.

[3] F. Musumeci, C. Rottondi, A. Nag, I. Macaluso, D. Zibar, M. Ruffini, and M. Tornatore, "An overview on application of machine learning techniques in optical networks," IEEE Communications Surveys & Tutorials, vol. 21, no. 2, pp. 1383–1408, 2018

[4] T. Dietterich, C. Bishop, D. Heckerman, M. Jordan, M. Kearns, et al., "Adaptive computation and machine learning."

[5] V. Nasteski, "An overview of the supervised machine learning methods," Horizons. b, vol. 4, pp. 51–62, 2017.

[6] I. Vlahavas, A. Likas, and G. Paliouras, "Machine learning and neural networks," Progressive, no. 1, 2013.

[7] Tao, Y., Yu, S. : DDoS attack detection at local area networks using information theoretical metrics. In : 2013 12th IEEE International Conference on Trust, Security and Privacy in Computing and Communications, Melbourne, VIC, Australia, pp. 233–240. IEEE (2013).

[8] H. B. Barlow. Unsupervised Learning. Neural Computation, vol. 1, no. 3, pages 295-311,1989

[9] ATEF. Diagramme de cas d'utilisation Chapitre 02 [en ligne]. 2009.

[10] Roques, pascal. UML2. Paris : groupe Eyrolles, 2008

[11] Havertz, P. (2022). What is Unified Modeling Language (UML)?. Retrieved 13 May 2022, from https://www.visual-paradigm.com/guide/ uml-unified-modeling-language/what-is-uml/

[12] Page-Jones, M. Fundamentals of object-oriented design in UML. (1999).

[13] PASCAL ROQUES. UML en action de l'analyse des besoins à la conception en java[en ligne].édition Eyrolles, 2000.

[14] *Kazi Aouel Bassim, Rostane Zakaria. Suivie des enseignements du* LMD par application de la méthode 2TUP [en ligne]. informatique industrielle. Université Abou Bekr Belkaid de Tlemcen, 2007.

[15] WILFRED ERISCO. Conception et developpement dune application de gestion d'une bibliothèque [en ligne]. 2003.

[16] Design and Development of an e-Learning Project Management System Modelling and Prototyping Maha Khaldi , Mohamed Erradi Abdelmalek Essaâdi University, Tetouan, Morocco

[17] Roques, P., & Vallee, F. UML 2 en action. De lanalyse des besoins a la conception J2EE, 3eme edition Eyrolles. (2004).

[18] Laouar, S.Melghid, N. Conception et implémentation d'une application mobile pour la gestion du Contrôle continu. (Mémoire préparé en vue de l'obtention du diplôme de Master, Centre Universitaire Abdelhafid Boussouf Mila, Mila, Algeria). (2019).

[19] MZOUGHI, M. Développement d'une application SAAS pour les cabinets Médicaux (Doctoral dissertation, Université Virtuelle de Tunis). (2013).

[20] Stein, G., Chen, B., Wu, A. S., & Hua, K. A. Decision tree classifier for network intrusion detection with GA-based feature selection. In Proceedings of the 43rd annual Southeast regional conference-Volume 2 (pp. 136-141).(2005, March).

[21] Hastie, Tibshirani, Friedman, The elements of statistical learning: data mining, inference, and prediction. New York: Springer Verlag, 2009.

[22] Fethi, A.Litim, Y. Détection des attaques DDOS dans le Cloud Computing. (Mémoire préparé en vue de l'obtention du diplôme de Master, Centre Universitaire Abdelhafid Boussouf Mila, Mila, Algeria). (2021).

[23] Petit, J., & Risch, J. C. Identification de Relations Semantiques Supportée par des Algorithmes dApprentissage Supervise Utilisant des Mesures de Proximite Semantique. 13eme atelier sur la Fouille de Donnees Complexes (FDC) Extraction et Gestion des Connaissances (EGC 2016).

[24] Buscombe, D., & Ritchie, A. C. Landscape classification with deep neural networks. Geosciences, 8(7), 244. (2018).

[25] Hebb, D. O. Animal and Physiological Psychology. In Annual Review of Psychology (Vol. 1, Issue 1, pp. 173–188). Annual Reviews. (1950). https://doi.org/10.1146/annurev.ps.01.020150.001133

[26] Choi, R. Y., Coyner, A. S., Kalpathy-Cramer, J., Chiang, M. F.,

& Campbell, J. P. Introduction to machine learning, neural networks, and deep learning. Translational Vision Science & Technology, 9(2), 14-14. (2020).

[27] F, M. C., & A, A. (n.d.). Guide D-elaboration D-UN Cahier de charge. Fichier PDF. Retrieved June 9, 2022, from https://www.fichier-pdf.fr/2010/02/13/skvbdsy/

[28] Ben Hadid, A. Conception et développement d'un outil de facturation de voix sur IP:«BILLPRO» (Doctoral dissertation, Université Virtuelle de Tunis). (2018).

[29] Amine, M., & Afif, K. La conception dune application de supervision des equipements aeriens a distance (Doctoral dissertation, Universite Virtuelle de Tunis). (2015).

[30] T. Bailet, Architecture logiciel. edition ENI, 2012.

[31] Diop, N. Automatisation de la gestion administrative du departement dinformatique de leuniversité Assane Seck de Ziguinchor (AGADI-UASZ). (2019).

[32] Nora, R. Conception et realisation deune application mobile pour la gestion dun Cabinet davocat (Doctoral dissertation, Abdelhafid Boussouf University centre-Mila). (2019).

[33] Cantiello, M., Ginev, D., & Pepe, A. How To Write Mathematical Equations, Expressions, and Symbols with LaTeX: A cheatsheet. Authorea Preprints. (2019).

[34] What is postman? Postman API Platform. (n.d.). Retrieved June 14, 2022, from https://www.postman.com/product/what-is-postman/

[35] Hamdane Mohamed Elkamel. Overleaf a collaborative cloud-based LaTeX editor. Unpublished. (2021). https://doi.org/10.13140/RG.2.2.33217.68961

[36] Visual paradigm enterprise. Best TOGAF Software with Agile & UML - Visual Paradigm Enterprise. (n.d.). Retrieved June 14, 2022, from https://www.visual-paradigm.com/editions/enterprise/#:~:text=Visual%20Paradigm%20Enterprise%20is%20an,ArchiMate%20language%20elements%20and%20relationships.

[37] Khafidhoh, N., & Ningrum, V. F. S. Design and Build an E-Catalog Application for Typical Jombang Products. In Multidiscipline Interna-

tional Conference (Vol. 1, No. 1, pp. 79-84). (2021, December).

[38] Sdgilley. (n.d.). What is azure machine learning? - azure machine learning. Azure Machine Learning | Microsoft Docs. Retrieved June 14, 2022, from https://docs.microsoft.com/en-us/azure/machine-learning/overview-what-is-azure-machine-learning

[39] Huseini, B., & Memeti, A. Implementation Issue Analysis of Java RMI and CORBA. Journal of Natural Sciences and Mathematics of UT, 4(7-8), 85-94.(2019).

[40] Benslimane, A. Web Application for the Management and the Control of Associations and Clubs. Case Study: Sub-Directorate of Scientific, cultural and sporting activities of the University of Msila (Doctoral dissertation, UNIVERSITYS MOHAMED BOUDIAF OF MSILA Faculty of Mathematics and computer sciences-DEPARTMENT: Computer Science-BRANCH: Computer Science OPTION: SIGL). (2019).