

ML Ontology solution for Blood Donation System

Final Year Project Proposal

Session 2022-2023

A project submitted in partial fulfilment of the
COMSATS University Degree
of
BS in Computer Science / Software Engineering (CUI)



Department of Computer Science
COMSATS University Islamabad, Lahore Campus

22 December 2022

Project Registration

| | | | | | | |
|--|--------------------------------------|-----------------------|-------------------------------------|--------------------------------|-------------------------------------|------------------------|
| Project ID (for office use) | | | | | | |
| Type of project | <input type="checkbox"/> Traditional | | <input type="checkbox"/> Industrial | | <input type="checkbox"/> Continuing | |
| Nature of project | <input type="checkbox"/> Development | | <input type="checkbox"/> Research | | <input type="checkbox"/> R&D | |
| Area of specialisation | Web Development | | | | | |
| Project Group Members | | | | | | |
| Sr.# | Reg. # | Student Name | CGPA | Email ID | Phone # | Signature |
| (i) | FA19-BSE-112 | Akash Ur Rehman | 3.10 | Akashurrehman456@gmail.com | 03494611282 | <i>Akash Ur Rehman</i> |
| (ii) | FA19-BSE-033 | Muhammad Abu Hurairah | 3.02 | Abuhurairahishtiaq37@gmail.com | 03334725570 | <i>Abu hurairah</i> |
| (iii) | Fa19-bse-098 | Muhammad Salman | 2.96 | Salman.ahmed.9373@gmail.com | 03022319373 | <i>Salman Ahmed</i> |
| <p>Declaration: FYP group members have cleared all prerequisites courses For FYP-I as per their degree requirements.</p> <p><u>For BS(Computer Science)</u></p> <p>(CSC241 Object Oriented Programming, CSC291 Software Engineering Concepts, CSC371 Database Systems-I, HUM102 Report Writing Skills)</p> <p><u>For BS(Software Engineering)</u></p> <p>(CSC241 Object Oriented Programming, CSE291 Introduction to Software Engineering , CSC371 Database Systems-I , HUM102 Report Writing Skills)</p> | | | | | | |

Plagiarism Free Certificate

This is to certify that, I am Akash Ur Rehman S/o Saif Ur Rehman, group leader of FYP under registration no CIIT/FA19-BSE-112/LHR at Computer Science Department, COMSATS Institute of Information Technology, Lahore. I declare that my FYP proposal is checked by my supervisor and the similarity index is 9% that is less than 20%, an acceptable limit by HEC. Report is attached herewith as Appendix A.

Date: 28 Sep, 2022 Name of Group Leader: Akash Ur Rehman Signature: *Akash Ur Rehman*

Name of Supervisor: Miss Sana Rizwan Co-Supervisor (if any): _____

Designation: _____ Designation: _____

Signature: _____ Signature: _____

Project Abstract (mandatory)

Our project is about Blood Donation Website. A blood donation is basically a process where an individual can voluntarily donate his/her blood for future transfusions. He/she donates his/her blood to the organizations that require that blood for treatment purposes like Hospitals etc. In our project, we are majorly focused on helping people in case of emergency circumstances. We are providing the people with a platform where they can donate their blood and also, they can seek help in any case of emergency.

To solve this real-world problem, we adopt machine learning and ontology learning approaches because it make the system learn by themselves. Blood donation is a very critical process and this process can be made simple by using the Machine learning approach. In this proposal, we propose an automated blood donation system to ease the emergency situation and make the blood donation easy process. The basic aim of developing this system or web Application is that users can also make the request for blood in any case and search for blood in the Blood banks. Patients can achieve the blood by contacting the donor through the internet or a personal contact number.

Different classification algorithms of machine learning such as logistic regression, random forest, or decision tree are considered and best one will be chosen to make this system work as perfectly as possible. Chosen Algorithms will undergo multiple result metrics to evaluate their superiority over other classification algorithms using the different conditions. Different technologies have the ability to show domain interest but Ontologies have the ability to describe the interest and its relations.

The Methodology used to build this system is Waterfall and prototyping. The technology platform for the implementation of this system uses Protégé for Making the Ontology Model, Python for training the Machine learning Model, Sparql for making the queries, MongoDB for Database, HTML, CSS, ReactJS, NodeJS, and JavaScript for Web Development.

1 Introduction (mandatory)

In the 21st century, accidents and other types of life-threatening issues quite easily happen especially in Pakistan. In many cases, blood is needed on an urgent basis. And many deaths occur due to the unavailability of blood on time. There are multiple blood donor societies in Pakistan but only a few online systems could help for arranging the blood on time and save precious lives. Manual systems are time-consuming in this fast-moving world. and By using information technology in this domain, we can decrease the workload on the Blood Banks, and NGOs, and help to identify the donors who will be eligible for donation on the basis of their medical conditions. In our project, we are majorly focused that we should provide the user the facility to donate as well as get blood in any case of emergency. That saves a lot of time and effort for the person that is already passing through some adverse circumstances. Supervised learning Models such as logistics regression, random forest or decision tree will be used for the donor or non-donor classification based on different factors. With machine learning, we will use ontology because it will improve the data quality for training the datasets. For making the ontology, we will use Protégé. Through ontology, we can enhance the machine learning algorithms' accuracy because it provides a dynamic knowledge base.

Through this website, we aim to provide blood to needy people, especially in emergency cases. Users also contact the donor if the requirements match. Modules of this system will be user and Admin. Admin can add Hospitals, Blood banks centers, and NGOs to the website. They can also view the list of hospitals, Blood bank centers, and NGOs and the list of donors which can be filtered on the basis of area and blood groups. They can also send different types of notifications or messages to the user. Admin can also view the request for blood donation. Users can register to the website after filling out the form. Any user can make the request for blood at any time. Users can also view different hospitals and blood banks with blood details. All the users can view the request for blood. They can accept or deny the request.

Goals that are to be achieved is that every user can make the request for blood. If the requested blood type is available in any blood bank or NGO notification to that user will be sent. We can add new Blood banks or Hospitals to our website through the Admin dashboard. Classification of the donor will be done through Machine learning algorithms such as a decision tree or Logistics regression that he cannot donate blood or not by checking various conditions or features. The main sub-tasks are that every user can book an appointment for donating blood to their interested Blood bank or NGO. If the user confirms the appointment, Appointment details will send to the respective blood bank or user. Every request for blood will be shown on the user's profile and at the top of the website.

In traditional methods exact steps required for solving problem and parameters values are fixed based on statistical assumptions and calculation for prediction. That's why traditional methods not produces precise accurate and correct prediction results. Different types of algorithms will be considered for the classification of the donor and the best one will be chosen that has high performance, greater accuracy, and best output. These classification algorithms are logistics regression, Random forest, and Decision Trees. Different feature variables such as age, weight, smoker or not, diabetes, any serious disease, last blood donation etc., will be considered for the classification of the donors that they are eligible for donation or not. We will use ontology and machine learning concepts to identify whether the person can donate the blood We are using the ontology concepts because it can help the Machine Learning and the cognitive Artificial Intelligence by broadening its scope. Ontology learning also helps in improving the quality of the training datasets.

2 Success Criterion

The success criteria of the project are that every interested user can book an appointment with the respective blood bank centers or NGOs for the donation of blood after checking that either they are viable for blood donation or not through Machine learning. Appointment details of every user will send to the respective user and blood bank or NGO. If Admin is able to add different types of hospitals and Blood bank centers to the website then it will be considered a success. Another success criterion is that every user can make a request for blood in the case of an emergency it will show to all the users at the top of the profile.

(A blood donation system is an automated system that minimizes manual work for managing donors, request makers, and blood donation centers) **SCOPE**. (The fulfillment time of the blood donation system is almost one year. During this period of time learn the usage of machine learning and ontology learning in blood donation system) **TIMELINE**. (The goal of the project is to develop a blood donation system for maintaining blood donation from donors to different blood donation centers. The main objective is to provide the best services to users and blood donation centers through machine learning and ontology learning) **GOAL AND OBJECTIVE**. (This project helps to maintain the donation history of users. By using this application users can easily contact the blood donation centers and get an appointment for donation) **DELIVERABLES**. (This project will handle queries and feedback from users. Donar can safely interact with the blood donation centers) **STAKEHOLDER SATISFACTION**.

3 Related work (mandatory)

“Haemoglobinopathies in India: estimates of blood requirements and treatment costs for the decade 2017-2026,”
[1]

“And those who saved the life of one, it shall be as if he had saved the life of all the mankind”. The best one among you is that who is beneficial to others. (Prophet Muhammad P.B.U.H). The literature shows that in developing countries collecting blood and make available for patients is very severe. Many patients die due to a lack of blood on time or transfusion with infected blood. Pakistan health Organizations want to provide facilities to the patients. Many welfare societies wrote multiple books which depict the prevention of unsafe blood.

“Profiling blood donors in Egypt: A neural network analysis, Expert system with Applications,” [2]

Through this website, we can fill the different gaps such as the Lack of communication between donor and patient. And searching the donor through their status and blood group type. WHO recommends that the collection of blood should be done from unpaid regular donors by strengthening the blood donation system. Profiling the donors is a good approach to control the emergency situation.

“A Research Paper on Blood Donation Management,” [3]

Advancement in natural science has increased the demand for blood significantly. This inspires us to grow the system that will assist the people. The limitation of the previous system was that there was no database for registered donors. The information retrieval of blood donors is a very laborious process and makes it hard for hospitals to save critical lives.

“Web-Based Information System for Blood Donation,” [4]

According to research, only 10% of people in Pakistan are willing to donate blood. Machine learning tools were implemented in New York for improving the blood donation system to save many lives.

The semantic technologies-based research are described as following for AI system's trustworthiness. A research which is literature based gives overview of use of semantic technologies related to machine learning method describes explanation and embedding of Neural Network used in domain of health-care and entertainment. Data must be in such a format that allows different relationships to be discovered. Ontology makes all these relationships visible. The research description framework and the web ontology language are the two standards that govern the construction of ontologies.

The basis of these ontologies is made of two main components including classes and relationships.

These are the following types of ontologies:

- 1) Top Level ontologies that represent very common ideas
- 2) Domain ontologies explain ideas related to a specific domain.
- 3) Task ontologies describe basic ideas related to the activity.
- 4) Application ontologies are only linked with specific activity or domain.

Ontology-based approach provides benefits of displaying, positioning and organising data in a better way.

Making machine learning system accountable

The 3 phases are described:

1) Predictive Model Development and deployment phase in which development of a model is recognized to check the standard and quantity of data and after this it checks the performance, assessing the performance and making it better. Then making sure the availability in production as their deployment.

2) Semantic annotation phase:

In this phase information is collected and data for training is collected on the basis of its quality and amount for predicting its standards and amount for predictive model. The reuse process of Ontologies resources is used and the basic standards of Ontologies is specification, to know the motive of ontology, to help the user to understand the data.

It describes two areas which are:

First is Representing forecast and second is Representing predictive procedure.

Representing forecast: At the present time, different ontologies can be used to show different functions. A type of ontology EEPSA was chosen to be used again because it discovers a platform for better consistency.

An Ontology-based approach for making Machine Learning systems Accountable[5]

According to McKinsey the Artificial intelligence technologies (AI) are advanced but its application and adoption is not wide as a need of time. AI systems should be trustworthy as the users have full confidence and willingness to use it on the basis of actions and directions and decisions of the AI system. Researchers describe different factors which affect the trust of users on AI system. These factors consist of explainable artificial intelligence (XAI) which was studied from 1970s to 1990s. XAI describes the techniques for a user to understand about the trust and emerging generation of AI partners. The rise of black-box models in AI systems which do not show the main logic to users causes a lack of explanation of different issues. And due to it white box models are introduced but these are not enough to make them accountable. The accountability means that the taken decisions fulfils the standard or not and considers someone accountable. Different technologies have the ability to show domain interest But Ontologies have the ability to describe the interest and its relations.

Articles give us an approach based on an Ontology of the accountability of machine learning systems.

Related Works

1. The semantic technologies based researches are described following for AI system's trustworthiness. A research which is literature based gives an overview of use of semantic technologies related to machine learning method describes explanation and embedding of Neural Network used in domain of health-care and entertainment. Another research shows an approach of post-hoc explanation of decision tree algorithms. It explains about the result that Ontology supported decision trees are more easy to understand and it also creates ad-hoc of problems.

OWL Ontology - SPARQL - Java Web Project [6]

One of the other research gives 9 different explanation types to propose Ontology as **OWL** restrictions. Doctor XAI approach shows post-hoc explanation of black-box models focused on the medical domain. These researches only describes the explanation but the main purpose is making the AI system accountable and giving the method that which predictive model will make the AI system accountable.

Role of ontology [7]

Ontology derived from the Greek word ontos and logos and it was first used in the 19th century. An ontology is a model which represents knowledge as a set of concepts and its relations within a domain. It is a Form of knowledge management. People have more access to data in a day than most people have access to in a lifetime in previous decades. Data is found in many different forms which makes it difficult to understand the relationship between different data. In our environment it is totally difficult to determine the relationship of policy in a word document and business policies in models. Data must be in such a format that allows different relationships to be discovered. Ontology makes all these relationships visible.

Research description framework and the web ontology language are the two standards that govern the construction of ontologies. The basis of these ontologies is made of two main components including classes and relationships.

1.These are the following types of ontologies:

- i) Top Level ontologies that represents very common ideas
- ii) Domain ontologies explain ideas related to a specific domain.
- iii) Task ontologies describe basic ideas related to the activity.
- iv) Application ontologies are only linked with specific activity or domain.

A Machine Learning Ontology[8]

Ontology development

NeOn and Linked Open terms (LOT) are the methods for the development of ontology. But Ontology Design Pattern (ODP) is a perfect option for the development of ontology.

Reusing of previous ontologies

It helps in upgrading the quality of applications and lessen the price and upgrading the qualities of ontologies.

To reuse the Ontologies, four steps are considered:

- a) Search for suitable ontology.
- b) Evaluation of ontology
- c) Contrast or comparison of ontology.
- d) Selecting the relatable ontology.

Benefits of ontology based approach

It gives benefits of displaying, positioning and organising data in a better way.

Making machine learning system accountable

The 3 phases are described:

a) Predictive Model Development and deployment phase in which development of a model is recognized to check the standard and quantity of data and after this it checks the performance, assessing the performance and making it better. Then making sure the availability in production as their deployment.

b) Semantic annotation phase:

In this phase information is collected and data for training is collected on the basis of its quality and amount for predicting its standards and amount for predictive model. The reuse process of Ontologies resources is used and the basic standards of Ontologies is specification, to know the motive of ontology, to help the user to understand the data.

It describes **two** areas which are:

First is Representing forecast and second is Representing predictive procedure.

Representing forecast: At the present time, different ontologies can be used to show different functions. A type of ontology EEPsA was chosen to be used again because it discovers a platform for better consistency. The base of EESPA is a mixture of 3 ODPs. These three ODPs are:

i) Affected by ODP

ii) Execution Executor Procedure (EEP)

iii) Result Context (RC)

The ODP discusses two characteristics standards and objects.

EEP ODP brings the affected odo and its two classes. The main purpose of RC. ODP is the implementation of the result of EEP ODP and its background.

Representing predictive procedures: It must be defined with specific terms. There is a difference between machine learning experiments and data minting.

c) Data exploitation:

After understanding the data from different terminologies the result should be saved at a place where we can easily find it out.

Triplestores are structured in such a way to store this type of data.

Conclusions:

1. The recently used implementation of AI systems is unable to gain the trust of users, XAI system is learned lately but different factors affect the trustworthiness of the system. This research describes the trustworthy and accountable procedure of the Machine learning system through ontologies.

1.This process is described in three phases. In the first phase the development and deployment at production conditions of predictive model is described and the second phase describes the annotation of the relatable information coming from predictive model through ontologies and third phase is about exploitation by the users.

1.This article suggests that the Semantic Technologies are helpful and can fulfil the gaps of trustworthiness on AI systems and give it the right directions for the future.

4 Project Rationale

Blood is the very most special gift that anyone can deliver or donate to another person. Blood is essential to help patients of different types survive surgeries, cancer treatment, harmful diseases, accident victims, illnesses, and traumatic injuries. Usually, people hesitate when they think of donating blood because they don't want to struggle a lot to find a suitable organization for donating their blood. And motivation is to help a person in need.

- The goal of donating blood can save a human life.
- The main purpose of doing this project is that nowadays there is no such website in our area that gives user the facility to get help in case of emergency. He/she has to struggle a lot to find the blood.
- Motivation of doing this project is that we will give them a better platform and this will help them in case of emergency and the people can also become the reason of saving lives by donating their blood.
- Through automated blood donation system, providing them a platform from where they can easily locate the organization and they can make an appointment to donate the blood.

4.1 Aims and Objectives

The goal of the project is to overcome the workload in the blood bank centers and provide ease to the people who need blood in different situations.

- Provide blood to the patients in the case of emergency by contacting with the donors.
- Through this website or system, users can search about all the blood groups in different blood group centers or NGOs.
- The main focused objective of the project is that donors can donate blood to any blood bank or NGO after taking an appointment.
- Through the system, they can also check whether they are viable for the donation of blood or not.
- Contact Information of the donors will also be available on the website with the status of whether they can donate blood or not.

4.2 Scope of the Project

The scope of this project is divided into different users:-

Admin

- Can add different hospitals and blood bank centers
- Can send a notification to users
- View a list of hospitals and blood banks with all the information

User

- Can book the appointment for the donation of blood to any blood bank
- Can request blood donation in case of emergency
- View the list of hospitals and NGOs with blood group information

Blood Banks Centres

- Can response to Request
- Add blood details of different groups

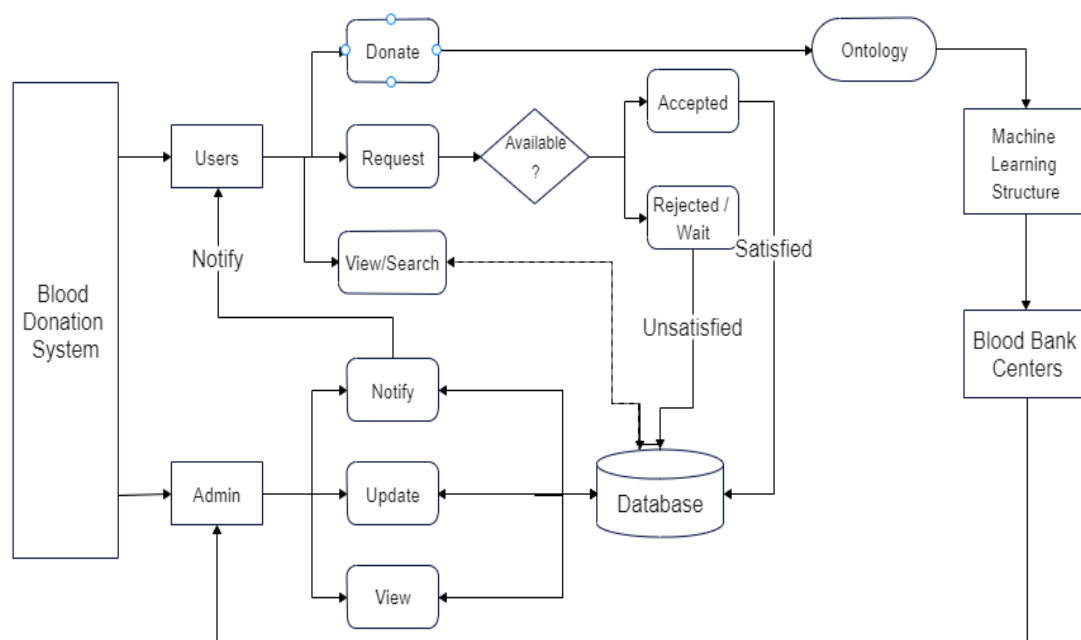
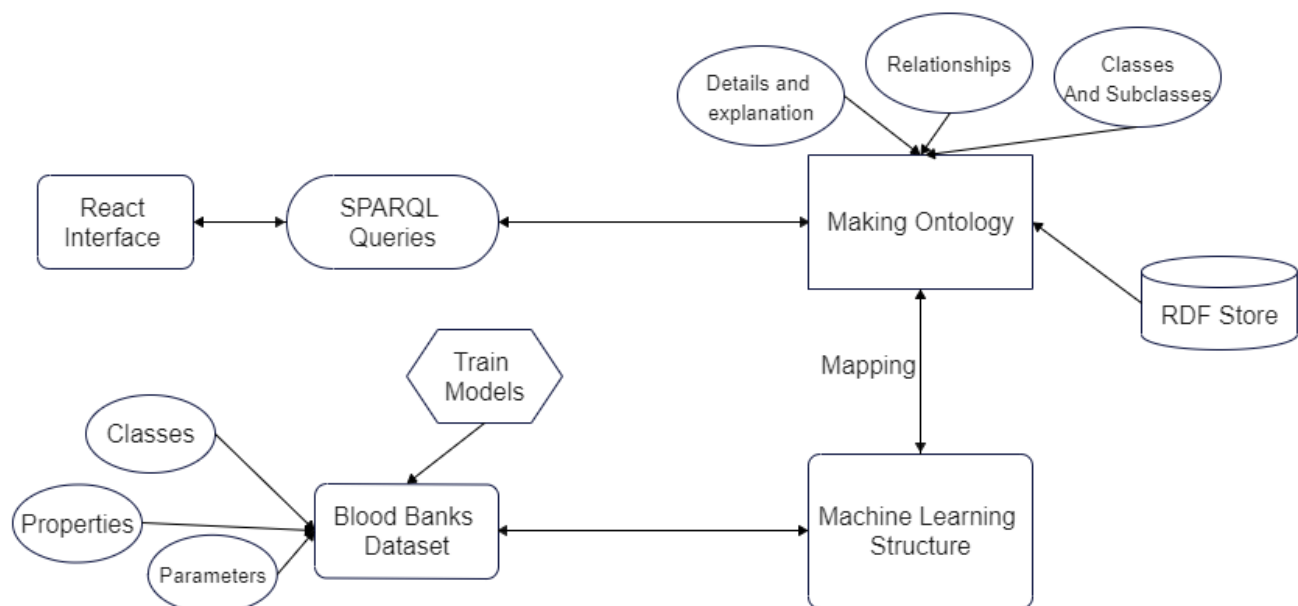
This project has to complete in one year. The documentation and design phase will complete in the first six months while the implementation and maintenance phase will complete in the next six months. Different tasks will include in this project like training the Machine learning models and making the ontology model.

5 Proposed Methodology and Architecture

The methodology we will employ in the development of the envisioned system is Waterfall. In the waterfall method firstly define function and non-functional requirements. Make the design of system from these requirements. Then implement the design using different tools and techniques. The last testing and maintenance phase will be performed for deploying our web application of blood donation system.

We choose this methodology because the previous phase must be complete before the next phase begins and there is no overlapping in this phase. Different Machine learning algorithms will be considered and the best one chosen on the basis of various factors.

Propose Architecture of Ontology and Machine Learning Models



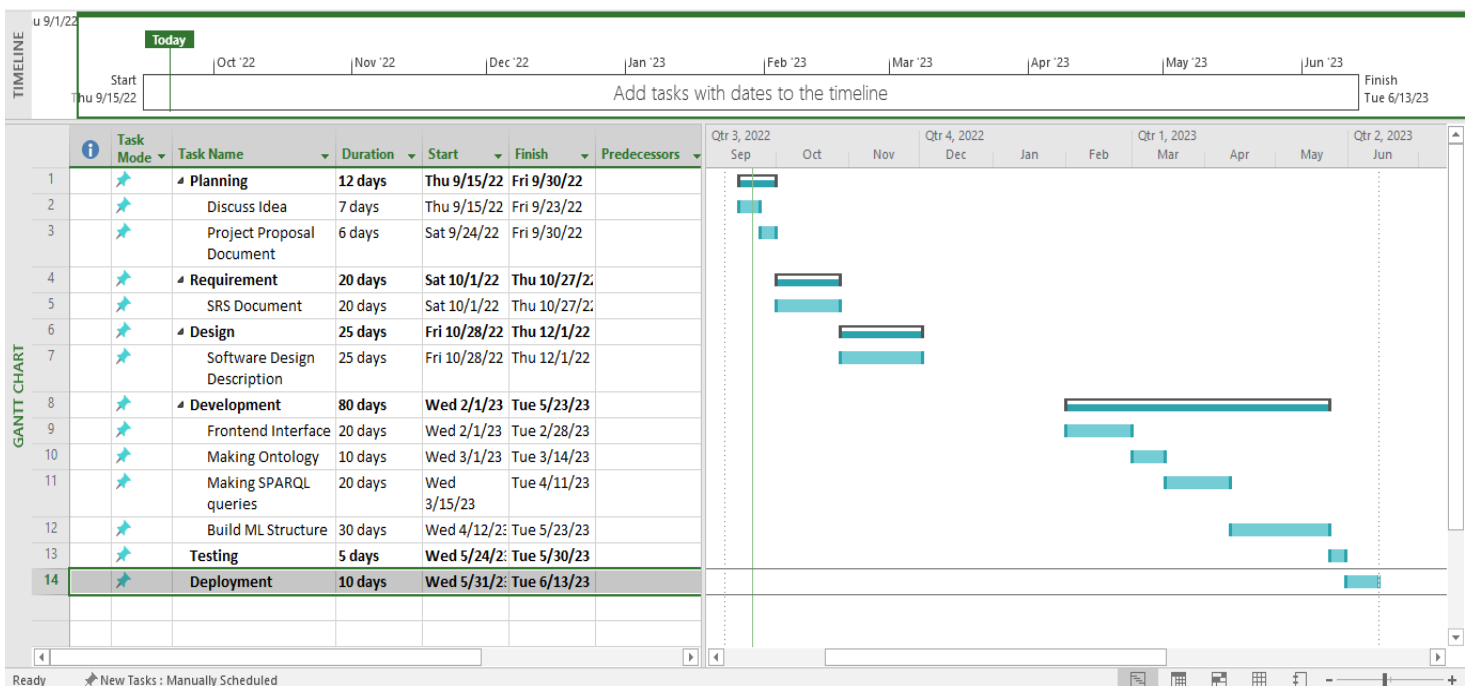
6 Individual Tasks (mandatory if group project)

Please mention individual tasks in the table indicating individual tasks and justify the one-year effort of two or three students. The expected individual task list with the tentative plan is given below.

| Team Member | Activity | Tentative Date |
|-----------------|-----------------------------|-------------------|
| Akash Ur Rehman | SRS Document | October 25, 2022 |
| Muhammad Salman | Design Description Document | November 28, 2022 |

| | | |
|-----------------------|----------------------------------|-------------------|
| Muhammad Abu Hurairah | FYP-I Presentaion | December 15, 2022 |
| Muhammad Abu Hurairah | Frontend Interface | February 28, 2023 |
| Muhammad Salman | Making Ontology | March 15, 2023 |
| Akash Ur Rehman | Making SPARQL Queries | March 31, 2023 |
| Akash Ur Rehman | Build Machine Learning Structure | April 15, 2023 |
| Muhammad Abu Hurairah | Backend | June 10, 2023 |
| Muhammad Salman | Final Presentaion | June 20, 2023 |

7 Gantt Chart (Mandatory)



8 Tools and Technologies

Different types of tools and technologies will be used for the implementation of this project.

For Frontend

- HTML

- CSS
- Java
- BOOTSTRAP
- REACTJs

For Backend

- NodeJs
- ExpressJs
- Javascript

For Machine Learning

- Python
- Protégé
- Spaql

For Database

- MongoDB

9 References (Mandatory)

- [1] S. T. C. R. B. A. Sinha S, “Haemoglobinopathies in India: estimates of blood requirements and treatment costs for the decade 2017-2026,” *Journal of community genetics*, vol. 11, no. June,11, pp. 39-45, 2020.
- [2] M. M. Mostafa, “ Profiling blood donors in Egypt:A neural network anaylsis, Expert system with Applications,” vol. 36, no. 2009, pp. 5031-5038, 2009.
- [3] D. K. Srivastava, Utkarsh Tanwar, M.G.Krishna Rao and Priya Manohar, “A Research Paper on Blood Donation Management,” *International Journal of creative research thoughts*, vol. 9, no. 5 May 2021, p. 8, 2021.
- [4] M. F. Qureshi, “Web-Based Information System for Blood Donation,” *International Journal of Digital Content: Technology and its Applications*, vol. 3, no. June,2009, p. 7, 2009.
- [5] A. Rutkin, “Big data aims to boost New York blood donation,” *New Scientist*, vol. 225, no. 3006,2015, p. 19, 2015.
- [6] I. Esnaola-Gonzalez, “An Ontology-based approach for making Machine learning systems Accountable,” *IOS Press*, p. 17.
- [7] J. Braga, Joaquim L R Dias and Francisco Regateiro, “A MACHINE LEARNING ONTOLOGY,” *ResearchGate*, p. 10, November 25,2022.

- [8] Nicola Guarino, Daniel Oberle, and Steffen Staab. What is an ontology? In Handbook on ontologies, pages 1–17. Springer, 2009.
- [9] A. Gangemi and V. Presutti, Ontology Design Patterns, in: Handbook on Ontologies, S. Staab and R. Studer, eds, Springer Berlin Heidelberg, Berlin, Heidelberg, 2009, pp. 221–243. ISBN 978-3-540-92673-3. doi:10.1007/978-3-540-92673-3_10

Include here the 1st page of Turnitin Report

Blood donation system

ORIGINALITY REPORT

9%

SIMILARITY INDEX

8%

INTERNET SOURCES

5%

PUBLICATIONS

4%

STUDENT PAPERS

PRIMARY SOURCES

1

semantic-web-journal.net

Internet Source

1%

2

www.aicit.org

Internet Source

1%

3

P.M.Durai Raj Vincent, S.S. Manivannan, C. Vanmathi, S.A.M Felicita. "Automated Blood Donation System using Machine Learning Approach", 2019 2nd International Conference on Intelligent Computing, Instrumentation and Control Technologies (ICICT), 2019

Publication

1%

4

hdl.handle.net

Internet Source

1%

5

ijircce.com

Internet Source

1%

6

ieeexplore.ieee.org

Internet Source

1%

7

Shadi AlZu'bi, Darah Aqel, Ala Mughaid. "Recent intelligent Approaches for Managing

1%