**Analyzing Blood Donation probabilities and number of possible donors**

**INTRODUCTION AND OBJECTIVES**

Blood transfusion has critical importance for human survival in risky situations that may occur. The number of possible donors and blood donation probabilities can be determined by using machine learning approaches. When the need for blood occurs in the future, medical professionals can predict potential donors for blood supply. Machine learning algorithms can support the blood transfusion process using datasets. When it comes to human health, data analysis is carried out to help prevent situations that will have critical consequences. By looking at the results of the data analysis, donors who may donate blood can be detected. In order to make this process carried out as expected, accurate and complete access to existing records must be provided. Blood transfusion has been provided for many years. The first successful transfusion was between two dogs in 1665. First medical usage of human blood in a transfusion was occured in 1818. Today, blood donation still has a vital value for saving human life. The best donor’s can be classified based on the following factors:

* R (Recency — Months since the Last Donation: this is the number of months since this donor’s most recent donation)
* F (Frequency — total number of donation: this is the total number of donations that the donor has made
* M (Monetary — total blood donated in c.c.: this is the total amount of blood that the donor has donated in cubic centimetres).
* T (Time — months since the first donation: this is the number of months since the donor’s first donation)
* Blood pressure (blood pressure at the time of last donation)
* Blood count(last time of donation)

**OBJECTIVES**

• It is aimed to create a data based system to monitor and estimate potential blood donors

• Using datasets it can analyse the last donation performance details of the donor

• Machine learning methods are used to analyse the number of donors who can donate blood

**Motivation**

Donating to the causes you care about **not only benefits the charities themselves, it can be deeply rewarding for you too**. Millions of people give to charity on a regular basis to support causes they believe in, as well as for the positive effect it has on their own lives. The blood you give is a lifeline in an emergency and for people who need long-term treatments.

**HARDWARE & SOFTWARE REQUIREMENT**

**HARDWARE REQUIREMENTS**

The selection of hardware is very important in the existence and proper working of any software. Then selection hardware, the size and capacity requirements are also important.

* Processor : Intel Pentium Core i3 and above, 64 bits
* RAM : Min 3GB RAM
* HARD DISK: 10 GB

**SOFTWARE REQUIREMENTS**

One of the most difficult task is selecting software for the system, once the system requirements is found out then we have to determine whether a particular software package fits for those system requirements. The application requirement:

* OPERATING SYSTEM: WINDOWS 10
* FRONT END: HTML, CSS, JAVASCRIPT
* BACK END: Mysql
* IDE USED: Jetbrains Pycharm, Android studio
* TECHNOLOGY USED: PYTHON JAVA
* FRAME WORK USED: Flask

**PROBLEM DEFINITION AND INITIAL REQUIREMENT**

**EXISTING SYSTEM**

In existing system when someone needs blood they have to contact other hospitals or persons,but there is only a small chance to get donor having same blood.In a risky situation we cannot take so much time for finding a perfect blood donor.

**PROPOSED SYSTEM**

In this project the performances of the two most successful classification algorithms were compared on the blood transfusion data set. By using machine learning models, it is thought to increase the connection between people who need blood and donor. In addition, an estimate was made to see whether the blood donor will donate blood in correct time. By using these findings, it may be beneficial to prevent risky situations. In this way, it may be easier to reach the right blood donor as soon as possible when blood is needed

**FEASIBILITY STUDY**

Feasibility study is defined as the practical extend to which a project can be performed successfully .When the client approaches the organization for getting the desired product developed, it comes up with rough idea about what all functions the software must perform and which all features are expected from the software.

This feasibility study is focused towards goal of the organization. This study analyses whether the software product can be practically materialized in terms of implementation, contribution of project to organization, cost constraints and as per values and objectives of the organization. It explores technical aspects of the project and product such as usability, maintainability, and productivity and integration ability.

Three key considerations involved in the feasibility analysis are:

1. Technical feasibility

2. Economic feasibility

3. Operational feasibility

**3.5.1 TECHNICAL FEASIBILITY**

Technical feasibility assess the current resources (includes the hardware and software) and technology. It centre on the existing computer system and to what extent it can support the proposed addition. Since the minimum requirements of the system like internet connection, web server in the server system is met by any average user.

**3.5.2 ECONOMICAL FEASIBILITY**

The economic analysis is to determine the benefits and savings with the current system and the proposed system that are compare with costs. The system is economically feasible, as the organization possesses the hardware and software resources required for the functioning of the system. Any additional resources, if required, can also be easily acquired. Proposed system was developed with the available resources. Since cost input for the software is almost nil the output of the software is always a profit. Hence Software is economically feasible.

**3.5.3 OPERATIONAL FEASIBILITY**

Operational feasibility assesses the extent to which the required software system performs a series of steps to solve business problems and user requirements**.** This is the longest phase in the development life cycle of a system. So, operational feasibility should be given much importance. The users of the application don’t need thorough training on the system. It has a user-friendly interface

**BASIC FUNCTIONALITIES**

**FUNCTIONAL MODULE**

• **Naive Bayes Classifier**

Naive Bayes is a kind of classifier which uses the Bayes Theorem. It predictsmembership probabilities for each class such as the probability that given record or data point belongs to a particular class. The class with the highest probability is considered as the most likely class

• **Machine learning**

Machine learning is a method of data analysis that automates analytical model building. It is a branch of artificial intelligence based on the idea that systems can learn from data, identify patterns and make decisions with minimal human intervention.

**MODULE DESCRIPTION**

• Admin

• Hospitals

• Users

• Blood bank

**ADMIN**

• Approve hospitals

• Block/Unblock hospitals

• Add and manage blood group

• View feedback

• View complaints and reply

• Predict blood donation requirements

**USERS**

• Registration

• View hospitals

• View blood groups

• Send complaint and view reply

• Send feedback

**HOSPITALS**

• Registration

• View users

• Send request for blood

• View request status

• View complaints and send reply

• View feedback

**BLOOD BANK**

• View request for blood

• Send request status

• Search blood

• Predict blood donation requirements

**LITERATURE SURVEY**

1. **Predicting Future Donor Behavior: A Machine Learning Approach**

Blood donations play a crucial role in the quality of a healthcare system. Unfortunately however, the number of recurring donors is decreasing. This has motivated blood banks to find out why some donors keep on donating and others stop, to improve their retention and recruitment strategies. This study investigates to what extend the Theory of Planned Behavior and adverse reactions can predict future donor behavior in registered blood donors in the Netherlands. Our research also looks at what other variables might influence a correct prediction, such as self-perceived health and actual health status of a blood donor. We used machine learning models to find out which variables were related to each other and influenced a correct prediction of future donor behavior; a combination that has not been used in previous studies on the subject. Overall, we found that age and number of previous donations were a strong predictor for future donor behavior. In addition, self-efficacy, intention, adverse reactions, self-perceived and actual health status, as well as a blood transfusion of a participant’s family member positively influenced the accuracy of our models. Based on these results, we conclude that the above-named variables influence whether a donor stays an active donor or if he or she stops. Moreover, we recommend Sanquin to focus on fainting prevention and enhance self-efficacy, also in returning donors, to ensure a healthy blood supply.

1. **Predicting Return Donor and Analyzing Blood Donation Time Series using Data Mining Techniques**

Since blood centers in most countries typically rely on volunteer donors to meet the hospitals' needs, donor retention is critical for blood banks. Identifying regular donors is critical for the advance planning of blood banks to guarantee a stable blood supply. In this research, donors' data was collected from a Saudi blood bank from 2017 to 2018. Machine learning algorithms such as logistic regression (LG), random forest (RF) and support vector classifier (SVC) were applied to develop and evaluate models for classifying blood donors as return and nonreturn donors. The natural imbalance of the donors' distribution required extra attention and considerations to produce classifiers with good performance. Thus, over-SMOTE sampling was tested. Experiments of different classifiers showed very similar performance results. In addition to the donors return classification, a time series analysis on the donors dataset was also considered to find any seasonal variations that could be captured and delivered to blood banks for better planning and decision making. After aggregating the donation count by month, results showed that the number of donations each year was stable except for two discovered drops in June and September, which for the two observed years coincided with two religious periods: Fasting and Performing Hajj.

# Classifying Blood Donors Using Data Mining Techniques

# Data mining refers to extracting knowledge from large amount of data. Real life data mining approaches are interesting because they often present a different set of problems for data miners. The process of designing a model helps to identify the different blood groups with available stock in Indian Red Cross Society (IRCS) Blood Bank Hospital Classification techniques for analysis of Blood bank data sets. The availability of blood groups in blood banks is a critical and important aspect in a Blood bank. Blood banks are typically based on a healthy person voluntarily donating blood and used for transfusions or made into medications. The ability to identify regular blood donors will enable blood bank and voluntary organizations to plan systematically for organizing blood donation camps in an efficient manner. The analysis had been carried out using a standard blood group donor's dataset and using the J48 decision tree algorithm implemented in Weka. The research work is used to classify the blood donors based on the sex, blood group, weight and age. This may be achieved through collecting the data utilizing the data mining technique and choosing the most suitable implementation tool for the domain.

# 4.Exploring Blood Donors’ Status Through Clustering: A Method to Improve the Quality of Services in Blood Transfusion Centers

# Urgent need for blood and the lack of an alternative for it necessitates the presence of a pattern to assist doctors in providing the proper services for the donors and also the right management of blood transfusion centers. The present study is aimed at determining blood donors’ status. Materials and Methods: Cross-sectional survey was applied in the present study through census. The population included the data extracted from Blood transfusion center of Birjand from Khordad to Shahrivar 1392 which was provided as an Excel file by the direct visit of the researcher from the blood transfusion organization. At first in the present study First, Two-Step clustering and then C50, C&R TREE, CHAID, and QUEST algorithms were executed to obtain the best ratio among different fields. Results: The obtained accuracy for executing C50, C&R Tree, CHAID, and QUEST equals 0.9998, 0.9960, 0.9930, and 0.8913, respectively. The results of indices including sensitivity, Specificity, accuracy, precision, FM, GM, FPR, FNR, and ER for C50 are indicators of better performance of this algorithm compared to the other ones. important variables in modeling are blood pressure label, blood donation status and temperature. Conclusions: The proposed model helps us in predicting faster and more precise status of blood donation and also the proper management of the blood transfusion centers and it can efficiency step for usage of blood donation and decreasing the blood maintenance costs.

# Conclusion

Now we can target the people who are interested in donating blood and which will result in getting more volunteers and we can save more people.

The demand for blood fluctuates throughout the year. As one prominent example, blood donations slow down during busy holiday seasons. An accurate forecast for the future supply of blood allows for an appropriate action to be taken ahead of time and therefore saving more lives.