

# To Improve Blood Donation Process using Data Mining Techniques

Hrushikesh Panchabudhhe<sup>1</sup>, Komal Palkar<sup>2</sup>, Aditee Padgilwar<sup>3</sup>, Krishna Rathod<sup>4</sup>, Prof. Shital Tatale<sup>5</sup>,

<sup>1,2,3,4</sup>B.E Student of Computer Engineering, JCOET, Yavatmal, Maharashtra, India

<sup>5</sup>Assistant Professor At JCOET, Yavatmal, Maharashtra, India, Sant Gadge Baba Amravati University, Amaravati, Maharashtra, India

Email: <sup>1</sup>hrushikeshpanchabuddhe@gmail.com, <sup>2</sup>palkarkomal7@gmail.com,  
<sup>3</sup>aditeepadgilwar@gmail.com, <sup>4</sup>krishnavrathod21@gmail.com,  
<sup>5</sup>tataleshital12@gmil.com

**Abstract:** *Blood saver of all existing lives just in case of emergency wants. throughout the insertion method, the acceptor receiving blood ought to be thought about before donating the blood. The donor data ought to be checked before displaying their details on the website. during this paper, we tend to propose an internet application to timely update the knowledge relating to the donors, acceptors, and patients wherever the administrator accesses the complete data regarding the bank distribution system. It additionally maintains the number of accessible blood teams, if the stock of a specific person is under the desired quantity then the projected technique notifies the donor to present blood. The web-based application is quickly climbable, economical, and filmable to satisfy the complicated want of the bank UN agency is an essential facilitator for the care sector. Thence the life at threat is often saved by this optimization technique.*

**Keywords:** *Blood bank, Donors, Clustering Techniques, Data Mining, Classification.*

## 1. INTRODUCTION

Human blood is very important in hospitals to improve the health of people. Hospitals always face a shortage of blood during emergency necessities. Blood bank sectors face encounters to meet the demand for blood across the country. The demand and supply of blood requirements can be proficient with the increase in voluntary donors. The blood and its benefactors are notably important because it can't be made and just originate from generous givers. Voluntary and non-remunerated donors have been considered employing the safest donors in blood bank sectors. Hence it is important to explore the information of donor data and convert them to appreciated information to recollect the voluntary donors for blood donation. the large volume of data offered in the health sectors, the important data is hidden in the database, and the knowledge of mining the information is necessary. As the volume of the data in the electronic health record rises the analysis of data and converting them into meaningful information is a major necessity in the health sector. To achieve the correct and exact result, we are using data mining techniques here because Data mining provides automatic pattern appreciation and attempts to uncover patterns in data that are difficult to detect with ancient statistical methods. A Data Warehouse is a semantically reliable data store that serves as a physical implementation of a decision support data model and stores the information on

which an initiative needs to make strategic decisions. A data warehouse is also often viewed as planning constructed by integrating data from multiple heterogeneous sources to support structured and/or ad-hoc queries, analytical reporting, and decision-making. Data mining can be used to expose patterns in data but is often carried out only on samples. The mining process will be unproductive if the samples are not a good representation of the larger body of data. Data mining cannot show up patterns that may be presented in the larger body of data if those patterns are not present in the sample being "mined". A crucial part of the method is the verification and validation of patterns on different samples of information.

Data mining commonly involves four classes of tasks:

**Classification** - Arranges the data into predefined groups.

For example, an email program might attempt to classify an email as valid or spam. Common algorithms include Decision Tree Learning, Nearest Neighbor, Bayesian classification, and Neural Network.

**Clustering** - This is like classification but the groups are not predefined, so the algorithm will try to group similar items.

**Regression** - Attempts to find a function that models the data with the least error.

**Association rule learning** - Searches for relationships between variables.

For example, a supermarket might gather data on customer purchasing habits. Using association rule learning, the supermarket can determine which products are frequently brought together and use this information for marketing purposes.

### **Need and Importance of Research Methodology**

Currently date, the world has experienced a dramatic rise in the amount of data stored online. With the extensive use of medical information systems including databases, there is a volatile growth in their sizes; Physicians and Surgeons are faced with the problem of making use of stored data. Massive healthcare data needs to be converted into information and knowledge, which can help control cost and continues the high quality of patient care. Without data mining, it is troublesome to realize the full potential of data collected at intervals in healthcare organizations as data under analysis is massive, highly dimensional, distributed, and uncertain.

### **Existing System**

The recruitment of blood donors, when associated with other countries, is very less in the overall blood donating percentage annually. Besides this employment, the screening of donors and the management system is not well continued. The details of the information of donors are given for the procedure of the users for contacting them when in need of blood in case of any emergency. The problem which currently happens in the medical field is that blood is needed immediately for an injured person or any major operation, it is not easily existing even though blood banks are present. There are some websites present for donating blood where the phone numbers of the donors are present which are not dependable since they don't get often updated. At present there are no suitable websites there is no proper care for people who donates blood to patients. That is the medical history of the donor is not

offered on the website. If a donor has had any medical problem and comes onward to donate blood to a patient then it may lead to threat.

Medical history like:

- A person who has anemia or is malnourished for height from their height should not donate blood.
- People who have diseases that are infectious via blood are disqualified from donating.
- Donors who have had their ear, tongue, or other body part penetrated are allowed to donate blood as long as the needle used in the penetrating was sterile. If it was not or if this is unknown, the potential donor must wait 12 months from the time of the penetrating.
- Being positive for the AIDS or hepatitis viruses ruled out as a blood donor.
- Pregnant women and recent childbirth ruled out as a blood donors.

This is because the safety of donating during and shortly after pregnancy has not been fully well-known. There may be medical risks to the mother and baby during this time. Thus the higher than following reasons don't seem to be thought-about on any websites. This kind of information is not provided by the current system and it may lead to death in person. Because the main reason is that the donor and the patient's body condition will not same all the time.

There are shortcomings in web-based applications within the blood donation system:

- Issues in handling emergencies.
- Slow internet access in some regions.
- No proper security for personal details.
- Misuse by third parties.
- No proper update about recent details.
- Want an intermediate to figure manually on information updates.
- Time intense (call or SMS to reach the donors).
- Leads to error-prone results.

## **2. PROPOSED METHODOLOGY**

The data consists of whole blood donation data together from local hospitals between 2019-2020. The data gathered consists of 8508 instances. First, the dataset is classified into permanent and temporary deferrals. According to the aim of the study, temporary deferral instances from the data are selected. The deferment sample consists of 1003 instances with 11 features. The features are donor number, name, phone number, weight, date, blood group, marital status, age, gender, donor type, and risk factors. The four features age, gender, donor type, and risk factors relevant to the study are considered. Many people come for whole blood donation at the hospital. During the donation, some of the donors are not able to donate because of health, and other reasons do not allow them for donation. As risk factors are associated with not becoming a donor, for the study the risk factor for not becoming donors and returning without blood donation are considered. In this, the risk factors that are temporarily deferred for donation are considered for analysis. K-means clustering is an unproven machine learning algorithm. Algorithm forms groups of objects which are similar. The clustering algorithm is used to cluster the input data into classes. The Elbow method used to find the value of k for k means clustering.

### **Advantages of Proposed Systems**

- 1) Less Time-consuming.

- 2) More efficient than the existing system.
- 3) The system is flexible and can work on any system.
- 4) Efficient final result.
- 5) There are many options for the recipient to look upon which will decrease the problem of unavailability.
- 6) It is comfortable to use.
- 7) No training is required.
- 8) The project is also developed in such a way that the user, new to the system will just have to install the setup and it is ready to go.

### **3. CONCLUSIONS AND FUTURE WORK**

In this research paper, we have described classification techniques for Blood Group Donors datasets and found an efficient and reliable blood donor information and management system with blood distribution based on data mining. To generate a Decision Tree we used a data mining classifier. The primary focus of this research is the development of a system that is important for the timely analysis of huge Blood Group Donors' data sets. The traditional manual data examination has become insufficient and the methods for efficient computer-assisted analysis are indispensable. This technique will be applied to the blood group transfusion database maintained in the Indian Red Cross Society (IRCS) and many Blood Bank hospitals and organizations across the country. This system will be the solution for the problems such as misuse by third parties and updating the donated blood by the donor which replaces the older systems. The proposed system is a web-based application that can help us to reduce the human mistakes which are done in the existing system. Future work can be applied to blood type classification, diagnosing diabetic symptoms, classifying blood donor type, diagnosing cancer patients based on blood cells, or predicting the cancer types on the blood groups.

### **4. REFERENCES**

- [1] Jiawei Han and Micheline Kamber "Data Mining Concepts and Techniques "2nd Edition.
- [2] T. Santhanam and Shyam Sundaram "Application of CART Algorithm in Blood Donors Classification" Journal of Computer Science 6 (5): 548-552, 2010 ISSN 1549-3636, pp.01-05 © 2010 Science Publications.
- [3] Arvind Sharma, P.C. Gupta, "Predicting the Number of Blood Donors through their Age and Blood Group by using Data Mining Tool", International Journal of Communication and Computer Technologies, Volume 01 – No.6, Issue: 02 September 2012.
- [4] Naji Moghadam, V. Ashoori, M. Alizadeh, S. and Safi, M. 2012." The Classification Algorithm for several tablet usage predictions: case study diabetes", The Sixth Iran Data Mining Conference, Tehran: Iran Data Mining.
- [5] Rashid Mehrabadi, E. and Pedram, M.M. 2010." Blood Donors Classification and Identifying Future Donors", The Fourth Iran Data Mining Conference, Sharif University of Technology, Tehran, Iran.
- [6] [6] Devchand Chaudhari & Dr. Ravindra S. Hegadi" Data Mining in Blood Platelets Transfusion Using Classification Rule." pp: 1-8.

- [7] Ankit Bhardwaj, Arvind Sharma, V.K. Shrivastava, International Journal of Engineering Research and Applications (IJERA), Vol. 2, Issue 4, July-August 2012, pp.1303-1309.
- [8] Priyanka Shinde, Pranita Taware, "Emergency Panic Button", International Journal of Scientific & Engineering Research Volume 3, Issue 3, March 2012.
- [9] Government of India, 2007. "Voluntary blood donation program" <http://www.nacoonline.org/upload/Final%20Publications/Blood%20Safety/voluntary%20blood%20donation.pdf>.
- [10] Anil Rajput, Ramesh Prasad Agarwal, Nidhi Chandel, Devendra Singh Solanki and Ritu Soni" Approaches of Classifications to Policy of Analysis of Medical Data" IJCSNS International Journal of Computer Science and Network Security, VOL.9 No.11, November 2009, pp.01-09.