Industrial Engineering Journal
ISSN: 0970-2555
Volume : 53, Issue 5, No.4, May : 2024

A MACHINE LEARNING-BASED BLOOD DONOR RECOMMENDATION SYSTEM TO ENHANCE BLOOD DONATION EFFICIENCY

Dr.Srinivasa Rao Akella, Chamanthi Thatiparthi, Karimunnisa Binth Hassen Mohammad Lakshmi Deepika Thammaneni Computer Science and Engineering, Andhra Loyola Institute of Engineering and Technology, Vijayawada, Andhra Pradesh ,India
akella.srinivas08@gmail.com chamanthithatiparthi@gmail.com karimunnisa612@gmail.com deepikathammaneni@gmail.com

Abstract—
Blood is an essential part of receiving medical care, yet maintaining a sufficient supply of blood is still difficult. This research provides a machine learning (ML) based recommendation system for blood donor availability. Given the growing need for particular blood types and the urgency of donations, this kind of technology is crucial for effectively matching donors and recipients. We employ the Random Forest algorithm to build the recommendation system. Through the use of a dataset containing attributes such as age, location, and blood type, the model predicts the likelihood of available donors. The preferred approach is Random Forest since it can handle complex data and generate accurate predictions. The system can determine and order the top ten potential blood donors according to relevant criteria and the required blood type. This information improves blood donation operations by streamlining the process of rapidly locating eligible donors. By putting this strategy in place, healthcare facilities can better handle emergencies and blood drives. Subsequent investigations could concentrate on including real-time data streams and broadening the model to encompass supplementary elements like donor eligibility standards and contribution chronicles, thereby enhancing the precision of recommendations and system usefulness.

Index Terms—Blood donor availability, Machine learning recommendation system, Random Forest algorithm, Healthcare logistics optimization, Donor-recipient matching,Donation history analysis.

INTRODUCTION

Medical analytics is crucial to the advancement of humanity. Many manually performed operations must be automated in order to avoid delays in medical care. Treatments administered on time will always help to save lives. Blood is one of these, and its prompt availability is essential for treating any emergency cases[1]. Accidents and other life-threatening situations happen rather frequently in this fast-paced society. The hospitals are required to treat each of these individuals as soon as possible.

One admirable idea to aid in this process is to deal with the problem of arranging blood at a time. People are voluntarily offering their assistance to one another these days. However, one of the most difficult tasks is connecting these donors with patients in need[16]. These procedures are always started as soon as an emergency occurs. However, automating this would facilitate the process overall and eliminate the need to wait for the circumstance to arise before responding. This project would, in general, be more proactive than reactive.

Because machine learning algorithms can forecast the future, they are assisting in the resolution of numerous realworld problems. It is capable of handling data that is labeled or unlabeled. There are several supervised methods for managing the labeled data[18], and classification is one of the duties it must carry out. Many of these problems would be resolved if the observed input data were properly classified. Conversely, unsupervised methods such as clustering will tackle the problem of unlabeled data. There are numerous well-liked supervised methods for solving classification issues. Applications in medicine stand to gain the most from these methods.

The frequency of a donor's donations and information regarding their most recent donation are two important pieces of donor data that should be taken into account. It will be crucial to maintain track of these two records in order to resolve this need-based donation[15]. In the modern era of social media, it is simple to contact a donor; nevertheless, the matters of eligibility and availability must be verified.

The development of a system that would automate this procedure will be much appreciated by humanity to aid in an emergency, yet the traditional blood bank system will coexist and continue to provide its services as it has in the past [10]. because, regardless of the urgency of the case, the traditional system is always laborious and takes its sweet time to reply to any inquiry.

To a certain extent, many issues could be resolved by keeping records properly. The execution would be delayed if there was any human intervention in this. This is the primary reason that total system automation would, to a certain extent, resolve these problems. This technique of forecasting the future based on the present records would be supported by machine learning algorithms. Accurate record availability would guarantee that this process proceeds as planned.

This study is structured as follows: an introduction, relevant literature, a suggested system, and an analysis of the results. We will be able to determine which strategy is most likely to perform more accurately and provide us with the flexibility to act through its categorization abilities thanks to this analysis. quick decision-making process.

LITERATURE SURVEY

A clinical decision support system for matching patients with suitable donors and managing situations more skillfully by raising standards was suggested [2] as a decision support system-based blood banking system. It was predicted that occasionally 50 percent of potential donors would not be available, which is another significant issue that needs to be properly addressed[3]. For managing such unforeseen situations, individual-level estimations based on a machine learning algorithm would be useful.

Social media use [4] would be beneficial in disseminating any message to a large audience. An average of 3.8 hours passed until the right individual received the Twitter messages, which were examined to see if they met the blood requirement. Once more, it would only be reachable during the day, from 10 a.m. to 4 p.m. The results of testing the multivariable[5] machine learning algorithmic approach against the donor selection process for hematopoietic cell transplantation were positive.

Creating a profile of the blood donor would be a wise course of action in an emergency [7]. Given Egypt's reputation as a research hub for the myriad facets of blood donation, this study's donor profiling took cognitive and psychographic characteristics into account. Big Data techniques were used in New York [8] to encourage individuals to donate blood, which will help save many lives in emergency situations.

A transfusion system was suggested after taking into account the different aspects of two individuals: a donor and a recipient [9]. This work addressed the time-consuming procedure of manually matching requirements. Time was saved when treating patients in emergency situations thanks to the automation of this.

The presentation of integrated blood donation management [12] showed the obstacles in addressing this. Integrated methods would enable more effective handling of this. The global network [13] was deemed necessary, and its challenges—which mostly concerned blood and bone marrow transplantation—were raised.

As a result, this chapter presents some of the linked works, while the following chapter presents the planned work.

PROPOSED SYSTEM

The proposed system uses machine learning algorithms to effectively match donors and receivers, revolutionizing the process of blood donation. Healthcare facilities can expedite the donation procedure and guarantee a steady supply of blood for patients in need by precisely forecasting donor availability based on multiple parameters, including age, geography, and blood type. Because it cuts down on wait times and minimizes blood shortages, this approach has the potential to save lives and enhance healthcare results. The Random Forest algorithm, a potent machine learning method renowned for its capacity to manage complicated datasets and generate precise predictions, is

employed by the recommendation system. A large dataset comprising donor data, such as demographics, past donations, and geographic location, is used to train the model. The algorithm can determine patterns and trends to forecast the possibility of donor availability for particular blood types by examining these variables. The first two chapters restate the requirement for an intelligent system within the integrated framework. This chapter will outline the suggested framework that will allow the process to run automatically without the need for human input.

Global healthcare systems are very concerned about blood shortages. Critical shortages of blood may result from an increase in demand for the blood product during emergencies, natural catastrophes, or pandemics. The suggested system helps prevent blood shortages and guarantees a steady supply of blood, particularly in emergency situations, by precisely forecasting donor availability and effectively linking donors with recipients. Blood donation facilities frequently struggle to efficiently manage their donor resources. The contribution process may be hampered by low manpower, logistical limitations, and administrative overhead. The system increases the efficiency of blood donation operations, minimizes administrative overhead, and optimizes resource allocation by automating donor recommendation and priority.

A suitable dataset in a necessity is necessary for the entire procedure, and it must be correctly assessed. For this study, we gathered our own dataset by surveying blood donors at different blood donation locations . The dataset contains contact details and donor history in addition to demographic data like age and blood type. The considered dataset consists of various attributes such as Recency in months ,Frequency as numbers ,Monetary in cc and the donor information. By closely examining and confirming the responses, as well as by carrying out preprocessing procedures to address missing values and standardize the data format, we were able to assure the quality of the data.

A. Methodology

The methodology for the proposed system is given as follows:

- Data collection: Gathering information about blood donors, such as contact information, blood type, age, location, frequency of donation, and recentness of donation. The machine learning model is trained using this dataset as a basis.
- Data Preprocessing: Handle missing values, eliminate duplicates, and standardize data formats to clean up the dataset. To extract pertinent features and convert categorical variables into numerical representations appropriate for modeling, use feature engineering.
- Feature Selection: Identifying which characteristics are most important in forecasting the availability of donors. To choose the best subset of features for model training, apply methods including domain expertise, feature importance ranking, and correlation analysis.
- Model Selection :Select a suitable machine learning algorithm for constructing the recommendation framework. Ensemble approaches that can handle complex data and produce reliable predictions, such Random Forest, Gradient Boosting, or XGBoost, are a good fit for this project because of the characteristics of the dataset and the prediction objective.
- Model Training :To assess the performance of the model, divide the dataset into training and testing sets. Using the training data, train the chosen machine learning model. Adjust the hyperparameters to get the best results in terms of accuracy, precision, recall, or F1 score.
- Model Evaluation: Analyze the prediction power and generalization capacity of the trained model using the testing dataset. Cross-validate and analyze performance indicators to make sure the system is reliable and robust.
- System Development :Create a user-friendly application or interface and incorporate the learned machine learning model into it to develop the recommendation system. To provide smooth system interaction, implement features for user input, prediction generation, and outcome visualization.

RESULTS AND ANALYSIS

The recommendation system will be put into place and give healthcare facilities useful information about donor availability. This information will help them make better decisions and connect with

possible donors. The technology reduces reaction times during emergencies by identifying the top donors based on predetermined parameters (such as frequency of donations or proximity to the donation site). This results in better patient care, lower medical expenses, and increased operational effectiveness for blood donation facilities. The following picture depicts the output of this system:

Figure 1 depict the result obtained sample screenshot of the top blood donors of given blood group .In the above figure 1 the blood group given as A+ .The result shows donor information such as Donor name, Email id, Location, Contact number. It is the sample only and all the results are obtained in a similar way.

```
Enter the blood group (e.g., A+, O-, AB+): A+
Top 10 donors for blood group A+
      Donor name                 Email ID       Location Contact number
       G.Jahnavi  jahnavigedda8749@gmail.com    Vijayawada    7386097503
   K.Ngalakshmi nagalakshmikalipanani@gmail.com      RL Nagar    8999547521
      M.Kamalakar     kamalakar2567@gmail.com Jangareddygudem    8989245678
      B.Satya rao        satya246@gmail.com   Machilopatnam    9951242881
       J.Sai Teja       saitejai73@gmail.com  Yanamalakuduru    6325874251
          O.Kiran   kirankumarsai98@gmail.com        Ganguru    9586211486
       Md.Haseena       haseena56@gmail.com     Rajahmundry    9505236921
  P.Amrutha vally        vally45@gmail.com     Pushpa hotel    8688673211
         N.Rajesh     nulirajesh76@gmail.com       Lingapalem    9705345671
     R.Anjaneyulu   anjaneyulu436@gmail.com           Guntur    9502772714
```

Fig. 1. Results

Accuracy	Precision	F1 Score	Recall	ROC AUC Score
0.2083	0.1810	0.1912	0.2083	0.5120

TABLE I

PERFORMANCE ANALYSIS

- Accuracy = (TP + TN) / (TP + TN + FP + FN)
- Precision = TP / (TP + FP)
- Recall = TP / (TP + FN)
- F1 Score = 2 * (Precision * Recall) / (Precision + Recall)

The result which are shown in the table 1 using above equations consists of details of True Positive, True Negative, False Positive and False Negative. With those values the various performance metrics are calculated like Accuracy, Precision etc.

In Figure 2, the confusion matrix also formed to test the results based on blood groups. The confusion matrix can be used in the context of the blood donor recommendation system project to assess how well the machine learning model performing the blood donor prediction performed. It consists of all the blood group.

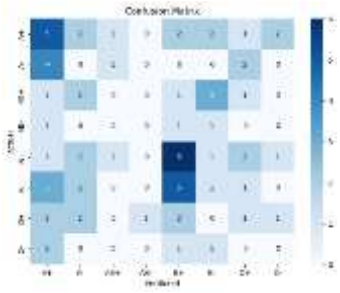

Fig. 2. Confusion Matrix

CONCLUSION AND FUTURE SCOPE

When the requirement is really high, this suggested solution is a concept that might be expanded to a big scale. A few benefits and dimensions of this effort include the narrowing of the gap between the recipient and the donor, the potential to even target rare blood types, and the potential use of prior data to learn more about the donor's background. The data from current blood donors will be linked with social media to create an application as part of the ongoing effort. In this manner, it would help a nation like India's vast populace. This innovative initiative's public-private collaboration should improve the

process of allocating proper money for the development as well as for market across the country. Sorting the donors using a machine learning algorithm would be an additional benefit in finding the ideal donor in every way. Therefore, automating has its advantages in resolving a crucial problem that will protect humanity in the event of an emergency. Additionally, this could guarantee that the primary focus and area of treatment will be healthcare.

REFERENCES

[1] P. M. D. R. Vincent, S. S. Manivannan, C. Vanmathi and S. A. M. Felicita, ”Automated Blood Donation System using Machine Learning Approach,” 2019 2nd International Conference on Intelligent Computing, Instrumentation and Control Technologies (ICICICT), Kannur, India, 2019,Volume 1,pp. 1069-1073.

[2] Paolo Rebulla, Lucilla Lecchi, Laura Porretti, Francesca Poli, Ilaria Ratti, Fulvio Mozzi, Girolamo Sirchia,”Practical placental blood banking. Transfusion Medicine Reviews”, Volume 13, Issue 3, 1999, pp.205226.

[3] Ramesh S Patil, Sagar Mhetre, Madhavi Rayate, A Gafoor Karache, ”Analysis of Blood Donor Deferral causes in Solapur district”, International Journal of Biological & Medical Research, nt J Biol Med Res.2014.

[4] Neil K. Shah,”Decision Support in Transfusion Medicine and Blood Banking, Clinies in Laboratory Medicine”, 2019, In Press.

[5] Javed Akhtar Khan and M.R. Alony, ”A New Concept of Blood
Bank Management System using Cloud Computing for Rural Area (INDIA)”, International Journal of Electrical, Electronics and Computer Engineering, 2015.

[6] Adarsh Sivasankaran, Eric Williams, Mark Albrecht, Galen E. Switzer, Vladimir Cherkassky, Martin Maiers, ”Machine Learning Approach to Predicting Stem Cell Donor Availability, Biology of Blood and Marrow Transplantation ”, Volume 24, Issue 12, 2018, pp.2425-2432.

[7] Dr.J.Shanthini, Santhoshi A, Manjula T, Pavithra R,” Computer Aided Emergency Service System”, IJIRT, National Conference on Networks, Intelligence and Computing Systems, March 2017.

[8] Rabeeh Ayaz Abbası, Onaiza Maqbool, Mubashar Mushtaq, Naif R. Aljohani, Ali Daud, Jalal S. Alowibdi, Basit Shahzad, ”Saving lives using social media: Analysis of the role of twitter for personal blood donation requests and dissemination, Telematics and Informatics”, Volume 35, Issue 4, 2018, pp. 892-912.

[9] Gregory P. Prastacos, ”Systems Analysis in Regional Blood Management, Springer-Verlag Berlin Heidelberg”, 1981.

[10] Ljubomir Buturovic, Jason Shelton, Stephen R. Spellman, Tao Wang, Lyssa Friedman, David Loftus, Lyndal Hesterberg, Todd Woodring,
Katharina Fleischhauer, Katharine C. Hsu, Michael R. Verneris, Mike
Haagenson, Stephanie J. Lee,”Evaluation of a Machine Learning- Based
Prognostic Model for Unrelated Hematopoietic Cell Transplantation Donor Selection, Biology of Blood and Marrow Transplantation”, Volume 24, Issue 6, 2018, pp.1299-1306.

[11] Nor Syuhada Che Khalid, ”Review of Data Gathering Techniques on Blood Donors Data Mining”, November 2015.

[12] Chris Booth, David Highley, Crystalloids, colloids, blood, blood products and blood substitutes, Anaesthesia & Intensive Care Medicine, Volume 11, Issue 2, 2010, pp.50-55.

[13] Shubham Pande, Shweta Mate, Pradnya Mawal, Ayusha Jambulkar, Prof. N. S. More, ”E-Blood Bank Application Using Cloud Computing”, International Research Journal of Engineering and Technology (IRJET), Feb-2015.

[14] Mohamed M. Mostafa,”Profiling blood donors in Egypt: A neural network analysis, Expert Systems with Applications”, Volume 36, Issue 3, Part 1, 2009, pp.5031-5038.

[15] Tushar Pandit, Satish Niloor, A.S. Shinde, ”A Survey Proposed System on E-Blood Bank and an Idea to use on Smartphone”, International Journal of Computer Applications, March 2015.

[16] Aviva Rutkin,”Big data aims to boost New York blood donations, New Scientist”, Volume 225, Issue 3006, 2015, Page 19.

[17] Zhang. M., Zhang, G., Ynag. J. and Chen, A.C.. 2012. ”The impact of a regular blood donation on the hematology and EEG of healthy young male blood donors. Brain topography”, 25(1), pp.116-123.

[18] M. Darwiche, M. Feuilloy, G. Bousaleh and D. Schang, ”Prediction of blood transfusion donation,” 2010 Fourth International Conference on Research Challenges in Information Science (RCIS), Nice, 2010, pp. 51-56.

[19] Mohammed, S. and Essel, H.B., 2018. Motivational factors for blood donation, potential barriers, and knowledge about blood donation in firsttime and repeat blood donors. BMC hematology, 18(1), p.36.