A Prototype of Emergency Blood donation Instantly Using AI- An Overview

1. **RAVI VERMA**

Research Scholar, Department of Computer Science,

Institute of Engineering & Technology,Dr.R.M.L Awadh University,Ayodhya, U.P. India.

SHIKHA TIWARI

Research Scholar, Department of Computer Science,

Institute of Engineering & Technology,Dr.R.M.L Awadh University,Ayodhya,U.P. India.

ANSHITA SHUKLA

Research Scholar, Department of Computer Science,

Institute of Engineering & Technology,Dr.R.M.L Awadh University,Ayodhya,U.P. India.

Abstract— Generally, it is important to note the fact that, knowing about the rarest blood group is not enough. Availability to nearest plays a vital role. Using big data the localities where the blood group and blood donor is available can be classified. However it helps in reduction of time and increase in possibilities. Artificial intelligence plays a vital role in determining capabilities of individual for blood donation. An object of Artificial intelligence- big data is in use to for predicting, analyzing and determining. For every 2 seconds there is a requirement for blood donor. The reason is simple; the development is at the peak and also accidents relating to them. Every second 100's of patients are taking admission in hospitals. Alone development is non justifiable for such incidents, along joins stabbing, road accidents, half murders and several cases. All things considered, there are different types of blood group and some are still unrecognized to be truthful. In such cases it is very important to store data of which blood group is rare and readily available. It is very important to understand that a person needs to be charged with the same blood group he/she is born with. Acquiring wrong blood group definitely leads to unnecessary complications otherwise to death.In this paper we will discuss how to organize blood donation in such a cost efficient way so that it reduce the death rate caused by unavailability of blood to patient at required time. The prototype system will be implemented on AI based app. unavailability of blood to patient at required time. The prototype system will be implemented on AI based appunavailability of blood to patient at required time. The prototype system will be implemented on AI based app. unavailability of blood to patient at required time. The prototype system will be implemented on AI based app.

Key words: Machine Learning and AI, Big Data, Data Extraction,AI app prototype

I. INTRODUCTION

The donation of blood is important because most often people requiring blood do not receive it on time causing loss of life. Examples include severe accidents, patients suffering from dengue or malaria, or organ transplants. Extreme health conditions such as Leukemia and bone marrow cancer, where affected individuals experience sudden high blood loss and need an urgent supply of blood and do not have it can also lead to loss of life. Sound data-driven systems for tracking and predicting donations and supply needs can improve the entire supply chain, making sure that more patients get the blood transfusions they need, which can reduce mortality risk. One of the interesting aspects about blood is that it is not a typical commodity. First, there is the perishable nature of blood. Blood has a shelf life of approximately 42 days according to the American Red Cross (Darwiche, Feuilloy et al. 2010). However, what makes this problem more challenging is the stochastic behavior of blood supply to the system. Whole blood is often split into platelets, red blood cells, and plasma, each having their own storage requirements and shelf life. For example, platelets must be stored around 22 degrees Celsius, while red blood cells 4 degree Celsius, and plasma at -25 degrees Celsius. Moreover, platelets can often be stored for at most 5 days, red blood cells up to 42 days, and plasma up to a one calendar year. Amazingly, only around 5% of the eligible donor population actually donate (according to census India,2011). This low percentage highlights the risk humans are faced today as blood and blood products are forecasted to increase year-on-year. This is likely why so many researchers continue to try to understand the social and behavioral drivers for why people donate to begin with. The primary way to satisfy demand is to have regularly occurring donations from healthy volunteers. In our study, we focus on building a data-driven system for tracking and predicting potential blood donors. We extract the data of population around any hospital then investigate the use of various binary classification techniques to estimate the probability that a person will donate blood in March 2018 or not based on his past donation behavior. There is a time lag between the demand of blood required by patients suffering extreme blood loss and the supply of blood from blood banks. We try to improve this supply-demand lag by building a predictive model that helps identify the potential donors. Based on our understanding of the problem, we follow a structured analytical process widely known in the data mining community, called the Cross-Industry Standard Process for Data Mining (CRISP-DM) . The idea behind this analysis framework is to develop and validate a model (or solution) that satisfies the requirements of problem. We used guidance in the academic literature to get ideas of how others have modeled this problem and followed a similar process. Many systems such as Facebook has partnered with blood donation centers around the world to help increase the number of donors. By encouraging blood donation as a way of life, each of us can assure that there is a readily available supply of blood.But the problem is lack of awareness among the users.And other social campaign organized but requirement still insufficient.

I. LITERATURE REVIEW

We examined the academic literature and grouped what we found into a couple different categories. First, blood banks often will survey donor volunteers to try and understand the factors that led them to donate.

[1] For example Godin, Conner et al. (2007) found that the important factors that lead to repeated blood donation among experienced donors were intention, perceived control, anticipated regret, moral norm, age, and past donation frequency. Moreover, the factors leading to repeated blood donation among new donors were only intention and age. Others have designed studies to understand one's motives for donating blood.

[2]Sojka and Sojka (2008) surveyed over five hundred donators and found that the most commonly reported motivator among their participants was friend influence (47.2%), followed by media requests (23.5%). Lastly, they found that altruism (40.3%), social responsibility (19.7%), and friend influence (17.9%) were the primary drivers for blood donors to continue to be blood donors in the future. As stated previously, only around 5% of eligible donor population actually donate . The reasons for this are regularly reviewed by social and behavior scientists to help improve population participation (Ferguson, France et al. 2007).

[3]The first published study we found investigating machine learning classification techniques to identify donors versus non-donors was by Mostafa (2009). They show that it is possible to identify factors of blood donation behavior using machine learning techniques. They train and test two artificial neural network (ANN) variants; one using a multi-layer perceptron (MLP); the other a probabilistic neural network (PNN). They then compare these non-linear models to a linear discriminant analysis (LDA) model. They conclude that the ANN models both perform very well compared to LDA due the nonlinearities that exist in their data.

[4] Santhanam and Sundaram (2010) used the Classification and Regression Tree (CART) from the University of California – Irvine Machine Learning repository. They showed on this data set that this algorithm has the ability to classify future blood donors accurately that had donated previously (i.e. recall/sensitivity of 94%). We found a very similar study published by one of the original authors the following year with a comparison of what they call a Regular Voluntary Donor (RVD) versus a DB2K7 (Donated Blood in 2007), which led to slightly better recall and precision (Sundaram 2011). Their key contribution was that the RVD model realized better accuracy than DB2K7. Darwiche, Feuilloy et al. (2010) extend this investigation of this data set by testing ANN with a radial basis function (RBF) as well as investigate performance using Support Vector Machines (SVMs). Even though the feature space is limited they also build and evaluate these models using principal components analysis (PCA) as feature inputs instead of the raw feature inputs. The SVM (RBF) model performed best using PCA as inputs because this model achieved the highest area under the curve (AUC) on the test set (i.e. 77.5%). The ANN model achieved the best AUC of 72.5% using only the features recency and monetary value. Lastly, we found the study design of (Darwiche, Feuilloy et al. 2010) better than (Santhanam and Sundaram 2010) and (Sundaram 2011) because their models are assessed on a test (i.e. holdout) set, which provides more realistic performance on future observations. Furthermore, this design allows one to identify if a model has overfit to the data by comparing the testing set statistics to the training set statistics.

[5]Zabihi, Ramezan et al. (2011) investigate the use of fuzzy sequential pattern mining to try and predict future blood donating behavior. The features investigated in this study were (1) months since last donation, (2) total number of donations, (3) time (in months) since first donation, and (4) a binary feature indicating whether blood was donated in

March 2007 or not. These features are similar in nature to those we investigated in our study. [4]Ramachandran, Girija et al. (2011) investigated the performance of the J48 algorithm provided in Weka3 . The J48 algorithm is an implementation of the C4.5 decision tree written in Java (Wikipedia , Quinlan 1993). They found this methodology to also perform well at predicting blood donors whom had donated before having a sensitivity of 95.2%, but performed poorly at future non-donors. Sharma and Gupta (2012) also used the J48 algorithm in Weka on a different blood donation data set obtained from a blood bank in Kota, Rajasthan, India. While they were attempting to predict the "number" of donors through their age and blood group, they actually performed a classification of donors versus non-donors which raised concerns over the validity of this study. Bhardwaj, Sharma et al. (2012) provided a very limited review of data mining in blood donation and do not actually train and test any models. They propose to do this in the future research.

DATA ORGANIZATION AND DATA EXTRACTION

Data Organisation is very first stage for any big data where there is huge data and we need data mining.And for this various algorithm has been proposed.All the relevant data then extracted according to machine need.These methods are classified into four catagories as follows.

The first category is based on blood group.All the data of blood group of located population is defined.

Separate catagories of each blood group is made according to their weightage.Second category is made for data of total population whose age <16 in the defined location.That will give the exact no. of donors.And last category is made for the list of hospitals around any area for which the extracted data is used.

CLASSIFICATION AND REGRESSION TECHNIQUES

CART Classification and Regression Techniques (CART) is used to refer to decision trees used for classification and regression techniques in predicting modelling. Trees have nodes and leaves. Every node is a question. In a binary tree, it is a yes or no question. The parent node is split into two child nodes. The splitting continues until a decision is reached for the target variable. The last node where the data cannot be split further are called leaves or terminal nodes. Every node is split on certain variable which gives the maximum information gain. Hence methods such as pruning are used where the nodes are not split after reaching a predetermined maximum depth. The nodes are then replaced by a leaf.

Artificial Neural Network Artificial neural networks (ANNs) are learning algorithms inspired by human brains. The main architecture of ANN is the input layer, the hidden layer and the output layer. Except for the input layer, all other layers are connected to their previous layer by weights in the form of a directed graph. The nodes represent a neuron which has a linear or non-linear activation function. The learning happens in two parts, feed-forward and back-propagation. In feed forward, weights are assigned and in back-propagation, actual learning happens. The error is calculated at each node and the weights are updated. This process is repeated until the algorithm converges. We investigated the multi-layered perceptron (MLP) neural network where the activation layer is a linear function which maps the weighted inputs to the output of each node.

Support Vector Machines Support vector machines (SVMs) are supervised classification or regression techniques widely used for non-linear datasets. The kernel trick allows the user to deal with non-linear data without having to worry about its linear separability. In SVM-RBF, a radial basis function is used as the kernel. The algorithm transforms the data into higher dimension into a linearly separable space and implements quadratic programming to increase the speed.The crux of the algorithm is that the data is transformed into a linear space. The data is then separated using a hyperplane which is supported by data points. The best separating hyperplane is one with maximum support vectors and the maximum margin. However, the drawback is that a very complex or wiggly hyperplane is likely to overfit the data.

Model: Made Donation in March 2007 ~ f (Total Number of Donations,Months Since First Donation + Months SinceLast Donation)

I. PROPOSED FRAMEWORK

A . Block Diagram:

The process flow can be depicted in the figure no.2. The working architecture starts with hospital where there is urgency of blood.The hospital will use proposed AI enable app which retrieve the patient matched blood group data in that location and then send a request message to all the population using UIDAI sever data(Unique Identification Authority of India) at their registerd mobile no. and huge no. off supply could be fulfilled using these framework.

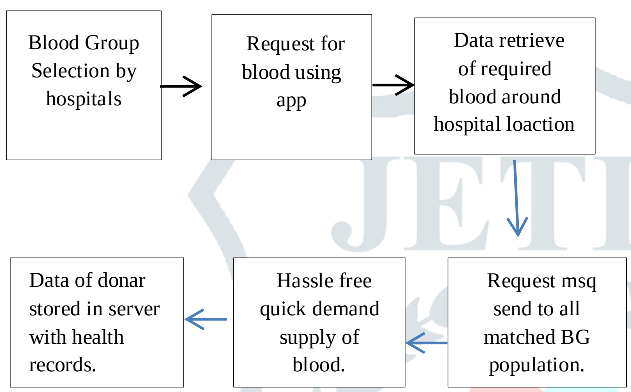

Fig. 2: Block Diagram

B. Software Part: The software part consist of a website or app that works on the data mining and AI.Whenever in case of emergency the app helps to send the alert message to the population near hospital whose blood group is same as the required.The data of population would be retrived using third party UIDAI.It would be programmed in either Python or R.The prototype of app is shown below.

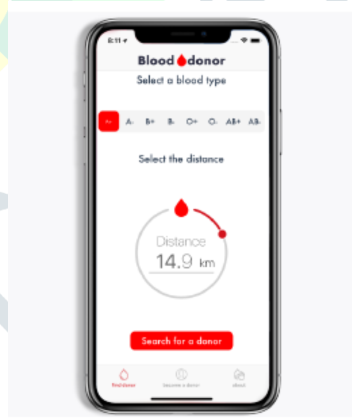

Fig 2. Prototype of Proposed System.

1

VI. CONCLUSION

The proposed system aims at developing a general framework to assist the needy person to get the blood donated wherever it needed from any location.The system will help to reduce the death rates of people who lost their lives only because of delay in blood donation.Also it would be a life saver project for everyone.

REFERENCES

[1] National Census of Indiahttp://censusindia.gov.in/Census_And_You/disa bled_ population.aspx

[2] Ashoori, M., et al. (2015). "A model to predict the sequential behavior of healthy blood donors using data mining." Ashoori, M., et al. (2017). "Exploring Blood Donors' Status Through Clustering: A Method to Improve the Quality of Services in Blood Transfusion Centers." Journal of Knowledge & Health 11(4): page: 73-82.

[3] Bednall, T. C., et al. (2013). "A systematic review and meta-analysis of antecedents of blood donation behavior and intentions." Social science & medicine 96: 86-94.

[4] Bhardwaj, A., et al. (2012). "Data mining techniques and their implementation in blood bank sector–a review." International Journal of Engineering Research and Applications (IJERA) 2(4): 1303-1309. Boonyanusith, W. and P. Jittamai (2012). Blood donor classification using neural network and decision tree techniques. Proceedings of the World Congress on Engineering and Computer Science‖