Towards an Efficient and Secure Blood Bank Management System

P.A.J.Sandaruwan
Department of Information Technology, Sri Lanka Institute of Information Technology, Malabe, Sri Lanka
janithsandaruwan29@gmail.com

U.D.L.Dolapihilla
Department of Information Technology, Sri Lanka Institute of Information Technology, Malabe, Sri Lanka
lakkanad@gmail.com

D.W.N.R.Karunathilaka
Department of Information Technology, Sri Lanka Institute of Information Technology, Malabe, Sri Lanka
naveendew96@gmail.com

W.A.D.T.L.Wijayaweera
Department of Computer Systems Engineering, Sri Lanka Institute of Information Technology, Malabe, Sri Lanka
dhanusha.wij@gmail.com

W.H.Rankothge
Department of Computer Systems Engineering, Sri Lanka Institute of Information Technology, Malabe, Sri Lanka
windhya.r@sliit.lk

N.D.U.Gamage
Department of Computer Systems Engineering, Sri Lanka Institute of Information Technology, Malabe, Sri Lanka
narmadha.g@sliit.lk

2020 IEEE 8th R10 Humanitarian Technology Conference (R10-HTC) | 978-1-7281-1110-0/20/$31.00 ©2020 IEEE | DOI: 10.1109/R10-HTC49770.2020.9356980

***Abstract*— a blood bank plays an important role in a hospital as well as in a country, ensuring safe and timely blood transfusions. However, there are several challenges faced by blood banks around the world, specifically when securing the blood supply chain. Reducing the supply-demand imbalance, protecting the data privacy of donors as well as receivers, are some of them. Therefore, there is a timely requirement for an effective and secure management system for the blood bank.**

We have proposed a management platform for the blood bank operations with the following modules: (1) forecast blood demand, (2) suggest blood donation campaign locations and (3) secure blood supply chain. The proposed platform has been implemented using techniques such as Long Short-Term Memory (LSTM), k-means clustering, Geographic Information Systems (GIS), and blockchain. Our results show that using our proposed modules, we can minimize the imbalance between supply and demand of blood, find the most suitable donor in an emergency, and enhance the privacy of data.

Keywords- *Machine Learning, LSTM, GIS, K-means Clustering, Blockchain*

I. Introduction

Human blood is a vitally important fluid in the body that circulates through our body and delivers essential substances like oxygen and nutrients to the body's cells. Hence, it is one of the most in-demand medical resources in the world [1]. Blood transfusion is mostly required in surgeries, organ transplants, and childbirths and for patients who are receiving treatments for diseases such as cancers, dengue, and anemia. Blood banking is the process that takes place in the lab to make sure that donated blood, or blood products, are safely stored before they are used in blood transfusions and other medical procedures.

According to [1], about 36,000 units of blood are needed every day, and about 6.8 million volunteer donor blood each year that results in 13.6 million blood units' donations per year. Therefore, a blood bank plays an important role in a hospital as well as in a country, ensuring safe and timely blood transfusions.

The first aspect of blood bank management is to accurately forecast future blood unit requirements and reduce blood demand and supply imbalance in the health sector. It will result in reducing the blood unit shortage and also blood wastage (i.e., decrease the overstocking of blood). Therefore, having a mechanism to forecast blood demand, specifically for each blood type, is an important aspect of conserving the valuable blood resource.

The second aspect of blood bank management is to plan blood donation campaigns effectively and timely so that a high number of blood donors can gather for the donation process. In the current context, the organizers of the blood donation campaign find it difficult to select the most effective place to have the donation campaign. There are campaigns with fewer donors, and campaigns with higher donors and campaigns that have fewer donors are not very effective. Therefore, suggesting the most effective area where it is most likely to be able to gather many donors who are willing to donate blood, results in conducting an effective blood donation campaign.

The third aspect of blood bank management is to ensure a secure blood supply chain. Every blood transfusion process faces two security challenges. First, preserving the confidentiality of donors' and receivers' data privacy, so that sensitive data is not disclosed to unauthorized personal. Second, ensuring the integrity of blood transfusion requests, so that fake blood transfusion requests are avoided. Therefore, maintaining confidentiality and integrity during the blood transfusion process will ensure a secure blood supply chain.

In this research, we have proposed a platform for efficient and secure blood bank management with the following modules: (1) forecasting blood demand, (2) suggesting blood donation campaign locations, and (3) securing the blood supply chain. The proposed platform has been implemented using techniques such as Long Short-Term Memory (LSTM), k-means clustering, Geographic Information Systems (GIS), and block-chain. We have consulted and visited the Blood Bank of Sri Lanka and doctors to identify the factors contributing to the above-mentioned activities and collect related data. We have conducted experiments to measure the performances of our algorithms using the Blood Bank of Sri Lanka as a case study. We have evaluated our system using real data sets, and in this paper, we are presenting the performance of our system. Our results show that using our proposed modules, we can minimize the imbalance between the supply and demand of blood, by predicting the demand patterns for the next 6 months, and by scheduling the blood donation campaigns effectively. It also enhances the security of the blood supply chain by ensuring the confidentiality and integrity of data.

The rest of the paper is organized as follows. Section II presents the related work. Section III introduces the four modules of the blood bank management platform. In section IV, the result and discussions are presented. Final Remarks and References are mentioned in Section V and VI, respectively.

II. RELATED WORK

Blood bank management is an essential component of most of the hospital management activities, but unfortunately, there are only a few research works have been conducted in this field. In this section, we will discuss the existing work on blood bank management related activities.

The first aspect of blood bank management is, forecasting the blood demand. The study in [2] shows that accurate blood demand forecasting will results in lowering costs, reduction of blood wastage, and conservation of limited resources in blood banks. They have used multiple approaches to predict blood demand, and their study concludes that the Box–Jenkins methodology gave the most accurate results. The authors of [3] conducted a forecasting analysis of blood components using POM-QM Software. The POM-QM software selected the method with the smallest number of errors of forecasting results using four methods: moving average, weighted moving average, exponential smoothing, and exponential smoothing with the trend. The work on [4] has used Artificial Neural Networks (ANNs) to predict the demand of blood components and their results show that the ANN-based model gives better results than the Auto-Regressive Integrated Moving Average (ARIMA) model. The Long Short-Term Memory (LSTM) is a special kind of Recurrent Neural Network (RNN) and has been used in several forecasting related research work because they are capable of learning long term dependencies [14-18].

The second aspect of blood bank management is to suggest more effective areas to conduct blood donation campaigns, by clustering donors' locations. The authors of [5], have used a k-means algorithm-based approach to conduct a blood donation information analysis. They propose to improve the accuracy of the k-means algorithm by increasing the accuracy in the calculation of the initial centroids. The study on [6] uses a k-means algorithm-based clustering method to describe blood donor behaviors. They have varied the value of k to have a different number of clusters and used Dunn's index for distinguishing the optimal number of clusters. The authors of [7] developed a system using clustering and classification algorithms to determine the disparities in blood donation behavior among the present donors. Further, they predict donors' intentions towards a donation to understand the problems and to increase voluntary blood donation frequency.

The third aspect of blood bank management is to ensure a secure blood supply chain. Blockchain technology is an emerging technology for securing supply chain data, and therefore, there are several on-going types of research works on using blockchain concepts to secure the blood supply chain. The work in [8] presents a framework that could be used for the implementation of blockchain technology in the healthcare sector for Electronic Health Record (EHR). Their framework aims to provide secure storage of electronic records by defining granular access rules for the users. The authors of [9] use blockchain to secure blood donor's information and to provide traceability and transparency of information in distributing blood packs. A serial number of the blood packet, which was issued by blood banks, will be stored in the blockchain, and once the blood bags are used, blood donors will receive the notification. The work on [10] has implemented a blood cold chain system based on private blockchain technology to achieve two goals: secure information visibility and reduce blood supply time. They show that the system prevents forgery and information tampering, which makes the blood management operation more transparent. Also, their system has the significance of reducing the blood supply time.

Our research has been inspired by these works, and we are targeting to build a single platform to handle blood bank related activities securely and efficiently.

III. METHODOLOGY

AIMA is a Web-based application that is intended to store, measure, recover, and investigate data worried about the managerial and stock administration inside a blood bank. We have proposed a platform for efficient and secure blood bank management with the following modules: (1) forecasting blood demand, (2) suggesting blood donation campaign locations, and (3) securing the blood supply chain.

A. Forecasting blood demand

The main goals of the blood demand forecasting module were to identify the most accurate model for blood demand prediction and then to calculate the required blood amount in the future. We have consulted the Blood Bank of Sri Lanka and doctors to identify the factors contributing to blood demand and to collect related data. The data set we used for the prediction included the date of the month and blood units count. First, the collected data were cleaned and preprocessed using Pandas and numpy libraries. Table I shows a sub-set of the data set we used for predicting the blood demand. It includes the date and number of blood units issued for each blood group. Next, the forecasting problem was modeled such that, when the issued blood count and relevant dates were given, it forecasts the blood demand for the next six months.

We used the Long Short-Term Memory (LSTM) model for the blood demand forecasting problem. The LSTMs are a special kind of Recurrent Neural Networks (RNNs) and they are capable of learning long term dependencies [11-13]. All RNNs have the form of a chain of repeating modules of neural networks. As shown in Figure 1, in standard RNNs, this repeating module will have a very simple structure, such as a single tanh layer.

TABLE I: Blood demand prediction – a subset of the data set

Date	Total No. of Units	No. of units for each blood group							
		O+	*O-*	*B+*	*B-*	*A+*	*A-*	*AB+*	*AB-*
1/1/2017	51	14	0	17	0	17	0	3	0
1/2/2017	32	11	0	4	0	13	0	4	0
1/3/2017	69	16	0	25	0	14	0	14	0
1/4/2017	49	10	0	23	0	11	0	5	0
1/5/2017	38	4	0	20	0	12	0	2	0

LSTMs also form in a chain-like structure, but the repeating module organizes differently. In an LSTM network, a module may have four neural network layers with three gates. The structure of a standard LSTM cell is shown in Figure 2.

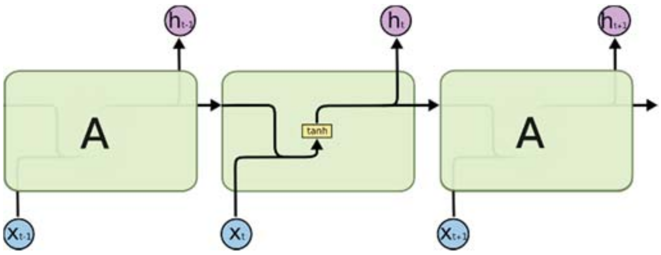

Fig 1: Standard simple RNN cell structure [12]

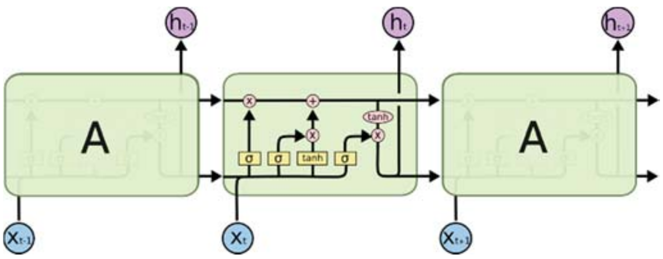

Fig 2: Standard LSTM cell structure [12]

First, the data was feature normalized as input set, X: {date, issued blood count for specific blood group} and output set, Y: {blood_demand_for_specific_blood_group}. Second, to fit an LSTM on the multivariate input data, the data was split into "Train set" (99.8%) and "Test set" (0.02%). Next, the data was fetched into Keras LSTM network with 1000 initial neurons and 1 output node.

We used the Mean Absolute Error (MAE) loss function and the efficient Adam version of stochastic gradient descent as evaluation models. The batch processing ran for 180 epochs (iterations) with a batch-size of 4. We explored both the training and test losses during training by setting the validation_data argument in the fit function.

Finally, we combined the forecast with the test dataset and invert the scaling. With forecasts and actual values in their original scale, we calculated an error score for the model. In this case, we calculated the Root Mean Squared Error (RMSE) that gives error in the same units as the variable itself.

B. *Suggesting blood donation campaigns locations*

The main goal of suggesting blood donation campaign locations module is to explore locations that have the possibility of gathering more donors and to conduct more effective donation campaigns in those locations. We have used the K-means clustering algorithm to cluster donors based on their residential addresses, and find the highest density cluster, which can be considered as the perfect location for the donation campaign.

K-means clustering is a type of unsupervised learning, which is used when you have unlabeled data (i.e., data without defined categories or groups) [19-21]. The goal of this algorithm is to find groups in the data, with the number of groups represented by the variable K. K-means, works as an iterative algorithm that tries to partition the dataset into K number of (K is pre-defined) distinct non-overlapping subgroups (clusters), where each data point belongs to only one group. It tries to make the intra-cluster data points as similar as possible while also keeping the clusters as different (far) as possible.

The K-means algorithm can be described by the following four key steps [19]:

1. Initialize cluster centers (centroids) (Ex: randomly select *K* points as centers).
2. Assign observations (data points) to the closest centroid.
3. Revise centroids as a means of assigned observations.
4. Repeat from step 1 to 3, until there is no change to the centroids

We have consulted the Blood Bank of Sri Lanka and doctors to identify the locations of donors and collect related data. In our implementation, we have assumed *K* to be 50 and, we have clustered donors into 50 clusters. First, the residential addresses of donors are converted into location information: longitudes and latitude [22]. Therefore, each donor's location is represented as {latitude, longitude} [23]. Second, we select 50 centroids: 50 locations where blood campaigns are held frequently. Third, the donors are divided into 50 clusters, each donor is assigned to the closest cluster. Finally, we revise the centroids, by taking the average of all donors' locations (longitude and latitude), that belong to each cluster, as given in Equation (1) and (2). These steps are repeated until we reach for a converged set of centroids.

$$Avg_L = \frac{(\sum_{i=0}^{n} L_i)}{n} \quad (1)$$

$$Avg_T = \frac{(\sum_{i=0}^{n} T_i)}{n} \quad (2)$$

Where,

n	No. of donors of the Cluster
Li	Longitude of i[th] donor's location
Ti	Latitude of i[th] donor's location
Avg_L	Average of Longitude
Avg_T	Average of Latitude

Once we have a converged set of clusters, the centroids (suggested locations for donation campaigns) and data points (locations of donors) are displayed on a map to provide better output [24].

C. Securing the blood supply chain

The main goals of securing the blood supply chain are to preserve the confidentiality of donors' and receivers' data privacy, so that sensitive data is not disclosed to an unauthorized person, and ensure the integrity of blood transfusion requests so that fake blood transfusion requests are avoided.

In an emergency, if a patient requires a blood transfusion, the doctor submits a "blood transfusion request form" with confidential details of the patient [25]. The form includes a unique ID (patient ID), blood type, reason for blood request, if any diseases, and contact details of the patient, etc. Once the form is filled, it is submitted to the blood bank of the hospital, and then the blood bank issues a blood packet for a given patient ID.

In our proposed blood bank management platform, we have allowed hospital management to register their hospital with a set of credentials. Once the hospital is registered on our platform, an authorized user can submit the blood transfusion request form, which includes a patient's sensitive data.

We have used blockchain technology to preserve the confidentiality and integrity of patient's sensitive data. A blockchain is a distributed copy of the database or a ledger in a network, where information stored in the ledger is made up of blocks and linked together in a chain [26]. Each block is connected to all the blocks before and after it – in a series of records. The records on a blockchain are stored securely through a cryptographic hash function. Each member in the blockchain network owns a full copy of the ledger. This makes it difficult to alter a record in blockchain and it provides the confidentiality and integrity features to data stored in blocks. Therefore, in our proposed system, we have stored the blood transfusion request form details, using the blockchain concept.

Furthermore, we have used the smart contract concept of blockchain to handle blood transfusion requests. A smart contract is a piece of code that contains specific terms that are executed when triggered by specific agreed events and stored inside a blockchain. Smart contracts run on the blockchain directly, thus making themselves secure from any kind of tampering and alterations. Whenever a blood transfusion request is submitted, the details of the form are passed using a smart contract. There are two smart contracts involved here; hospital and blood bank. If the blood transfusion is approved, blood packets will be allocated as requested, and the allocated blood packets will be identified by an ID. The ID of the allocated blood packets is also passed to the user who requested the blood transfusion, using a smart contract. Other than those concepts, a delegate call is introduced here for the connection between two smart contracts.

IV. Experiments and results

In this section, we are presenting the results and observations of the following modules: (1) forecasting blood demand, (2) suggesting blood donation campaign locations, and (3) securing the blood supply chain.

A successful blood management system works through an evidence-based approach to optimize patient safety and improve management and outcomes through measurable enhancements We have consulted the Blood Bank of Sri Lanka and doctors to identify the factors contributing to the above-mentioned activities and collect related data. We have conducted experiments to measure the performances of our algorithms using the Blood Bank of Sri Lanka as a case study.

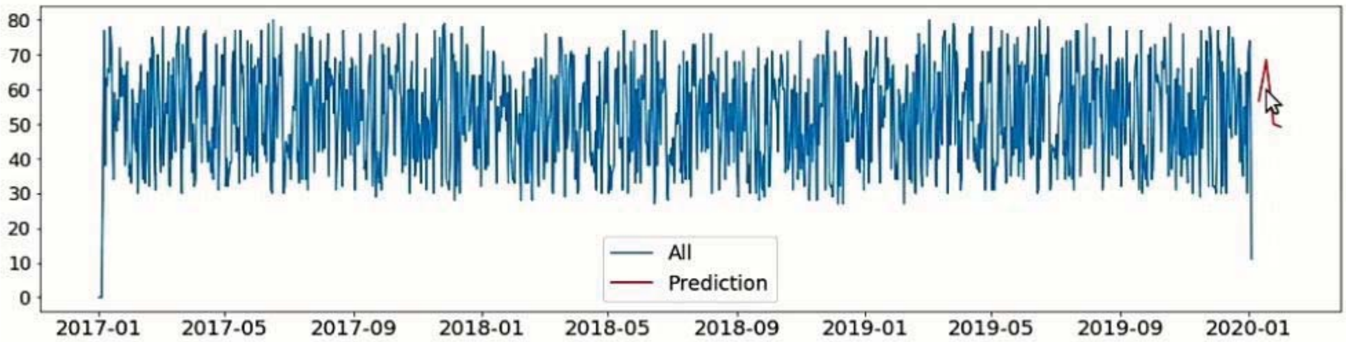

Fig 3: Blood demand prediction: Actual values VS Predicted values

A. Forecasting blood demand

For the blood demand prediction module, we ran the blood demand prediction LSTM network for 180 iterations with the blood demand prediction dataset for each blood group separately.

TABLE II: RMSE values for each blood type

Blood Group	RMSE
A+	9.75
A-	0.08
B+	6.7
B-	0.05
AB+	3.8
AB-	0.4
O+	8.1
O-	3.01

Performances of the blood demand prediction module were analyzed using the following matrices: Mean Squared Error

(MSE), Root Mean Squared Error (RMSE), training score, and test score of the models.Figure 3 shows the comparison of actual values and predicted values for test data set for all types of blood groups, as a result of the blood demand prediction. Table II shows RMSE values given by the blood demand prediction model for each blood group type when using a 98.2% train set and 0.2% test set.

B. Suggesting blood donation campaigns locations

For the suggesting blood donation campaign locations module, we have used the K-means algorithm where we assumed K to be 50. We have selected 50 initial centroids, which are the 50 locations that hold blood campaigns frequently. We have used {longitude, latitude} of each donors' location as data points for the algorithm. With the k-means clustering algorithm, we have clustered all donors into 50 clusters. The results of our clustering algorithm are visualized in Figure 4. High-density clusters represent locations that have more donors, and low-density clusters represent locations that have fewer donors. Therefore, by analyzing the cluster densities, our system can suggest locations that are likely to have more blood donors.

The geo-informatics of the above-formed clusters are shown in Figure 4 Each cluster is indexed using a cluster label. Also, the resultant {longitude, latitude} represents the centroid of each cluster, which represents the locations to be considered as blood donation campaign locations as in table III.

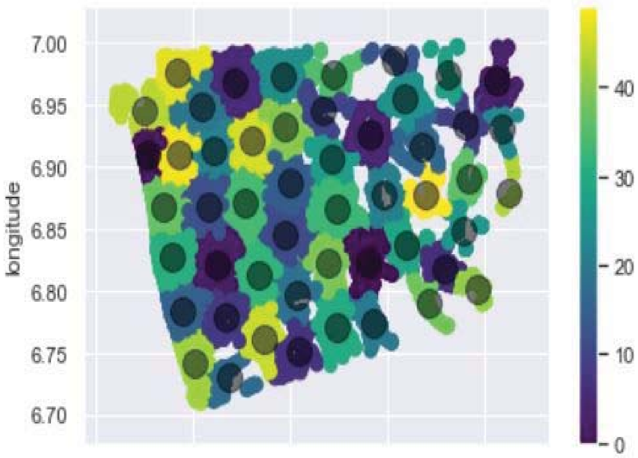

Fig 4. K-mean visualization

TABLE III. Results of the cluster

Serial No	Y17/D00001
Latitude	80.142804
Longitude	6.794448
Address	Meepe-Ingiriya Road, Wewila, Western Province, LK
ClusterLabel	15

C. Securing the blood supply chain

For the securing blood supply chain module, we have used blockchain technology and its concept of the smart contract. All the sensitive data of the blood transfusion request form is stored in blocks, and they are secured through cryptographic hash functions. Therefore, view the actual stored patient and blood packet data can be seen as in table IV, where details in Etherscan will see is a meaningless set of data as in figure 5 which contains the details of the same transaction which can be confirmed using transaction hash and contract address in table IV.

TABLE IV: Patients' details passing between hospital and blood bank

_id	5f5e065611462301adc3bcf3
updatedAt	2020-09-13T11:45:37.199Z
createdAt	2020-09-13T11:45:26.962Z
id	32
hospital_id	52
pnic	989234534v
pname	A.Perera
pphone	0112342674
bloodtype	B+
reason	for lung cancer
age	21
dname	DR Michael Maxvel Remance
dID	22517
date	10-10-2020
dinNo	23451
contract_address	0x084566aa525456262bbb2a40c955e9bc5ad9 3232964831f6e19fdc8b16c
tx_hash	0xad1824cc4f1a525456262bbb2a40c955e9bc 5ad93232964831f6e19fdc873ef1

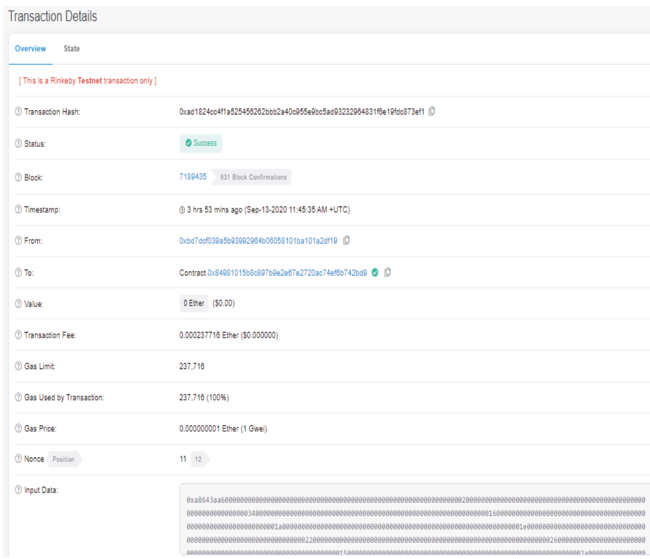

Fig 5. Patients' details passing between hospital and blood bank in Etherscan as a transaction

The hospital ID is equal to the hospital ID where the patient details are passing from. therefore the transaction is made from that particular hospital to the blood bank. Then when it requests to the blood bank, a blood packet id (dinNo) can assign to that particular patient ID.

V. CONCLUSION

We have proposed a platform for efficient and secure blood bank management with the following modules: (1) forecasting blood demand, (2) suggesting blood donation campaign locations, and (3) securing the blood supply chain. The proposed platform has been implemented using techniques such as Long Short-Term Memory (LSTM), k-means clustering, Geographic Information system (GIS), and blockchain. We have evaluated our system using real data sets, and our results show that using our proposed modules, we can minimize the imbalance between the supply and demand of blood. Long Short-Term Memory (LSTM) model was used for the blood demand forecasting problem, which is an artificial recurrent neural network architecture used in the field of deep learning to provide an accurate blood demand prediction. Other than that this helps for scheduling the blood donation campaigns effectively. From this, it can visualize a large number of data points simultaneously on a map through google map APIs.which is important to find blood campaign locations in real-time. It also enhances the security of the blood supply chain by ensuring the confidentiality and integrity of data. Furthermore, we introduce the smart contract for blockchain-based healthcare systems which is key for defining the pre-defined agreements among various involved stakeholders.

As future work, we are planning to implement a module to find the best donors for specific blood transfusion requirements, especially in emergencies. We are planning to consider factors such as donor's age, the most recent time of blood donation and distance to hospital, etc. Also further for the human organ donation.

REFERENCES

[1] Manning FJ, Sparacino L, Blood Donors and the Supply of Blood and Blood Products, Washington (DC): National Academies Press (US).

[2] Sima M. Fortsch, "Reducing uncertainty in demand for blood", Operations Research for Health Care, 2016

[3] Fitra Lestari, Ulfah Anwar, "Forecasting Demand in Blood Supply Chain (Case Study on Blood Transfusion Unit)", World Congress on Engineering 2017

[4] Rohaifa Khaldi, Abdellatif El Afia, "Artificial Neural Network Based Approach for Blood Demand Forecasting: Fez Transfusion Blood Center Case Study", International Conference on Big Data, Cloud and Applications 2017

[5] Bondu Venkateswarlu, "Mine Blood Donors Information through Improved K Means Clustering", International Journal of Computational Science and Information Technology 2013

[6] Maryam Ashoori, "Using Clustering Methods for Identifying Blood Donors Behavior", Iranian Conference on Electrical and Electronics Engineering 2013

[7] W. C. Lee, "An intelligent system for improving performance of blood donation", Journal of Quality, 2011

[8] Ayesha Shahnaz, "Using Blockchain for Electronic Health Records", IEEE Access 2019

[9] TITH Dara, "Blockchain-Based Blood Bank Ecosystem for Improving Public Health and Encouraging Voluntary Blood Donors", ASEAN IVO Forum 2019

[10] Seungeun Kim, "Implementation of a Blood Cold Chain System Using Blockchain Technology", Journal of Applied Sciences, May 2020

[11] S. Hochreiter and J. Schmidhuber, "Long Short-Term Memory", Neural Computation, vol. 9, no. 8, pp. 1735-1780, 1997.

[12] "Understanding LSTM Networks -- colah's blog", Colah.github.io, 2018. [Online]. Available: http://colah.github.io/posts/2015-08-Understanding-LSTMs/. [Accessed: 15- Aug- 2018].

[13] "An Introduction to Recurrent Neural Networks", Towards data science, 2018. [Online]. Available: https://towardsdatascience.com/an-introduction-to-recurrent-neural-networks-72c97bf0912. [Accessed: 14- Aug- 2018].

[14] M. R. S. Muthusinghe, Palliyaguru S. T., W. A. N. D. Weerakkody, A. M. Hashini Saranga, W. H. Rankothge, "Towards Smart Farming: Accurate Prediction of Paddy Harvest and Rice Demand", R10-HTC 2018.

[15] Dhanuka Perera, Chamod Rathnayaka, Sachin Dilan, Lasitha Siriweera, W. H. Rankothge, "Sustainable Tourism: Application of Optimization Algorithms to Schedule Tour Plans", R10-HTC 2018.

[16] M.C. Jayasuriya, K.T. Galappaththi, D. Sampath, N. Nipunika, W Rankothge, "Experimental Study on an Efficient Dengue Disease Management System: Planning and Optimizing Hospital Staff Allocation", International Journal of Advanced Computer Science and Applications 2018

[17] C Thiranjaya, R Rushan, P Udayanga, "Towards a Smart City: Application of Optimization for a Smart Transportation Management System" ICIAfS 2019

[18] S Perera, S Dissanayake, D Fernando, "Supply and Demand Planning of Electricity Power: A Comprehensive Solution", CICT 2019

[19] SPSS Inc. Introduction to Clementine®. 2003.

[20] SPSS Inc. Clementine® 8.0 User's Guide. 2003.

[21] Ray, S., Turi, R.H. Determination of Number of Clusters in K-Means Clustering and Application in Color Image Segmentation. 4th International Conference on Advances in Pattern Recognition and Digital Techniques (ICAPRDT'99), 1999, PP: 137-143.

[22] X. Ge, "Address geocoding," New York,US, 2003.

[23] R. C. K. N. Michael Chau, "Uncertain Data Mining: An Example in Clustering Location Data," in Springer, Berlin, Heidelberg, Pokfulam, Hong Kong, 2006.

[24] Yifan Hu,Emden R. Gansner,Stephen Kobourov , "Visualizing Graphs and Clusters as Maps," in IEEE, 2010.

[25] "studylib.net. (2020). Blood transfusion special requirement request form. [online]," https://studylib.net/doc/7562869/blood-transfusion-special-requirement-request-form, [13 Feb. 2020].

[26] Dylan Yaga , "Blockchain Technology Overview", National Institute of Standards and Technology Internal Report 8202, 2018

[27] Simanta Shekhar Sarmah, "Understanding Blockchain Technology", Computer Science and Engineering 2018