E-BLOOD BANKS ANALYSIS USING LSTM MACHINE LEARNING TECHNIQUES

[1]S.Senthilvelan, [2]E.Dharshini, [3]C.Sujitha
[1]Professor, MCA Department, Paavai Engineering College, Namakkal, Tamil Nadu
[2,3]II MCA, Paavai Engineering College, Namakkal, Tamil Nadu

Abstract :As technology advances rapidly, complexity in data processing increases. To prevent and nullify these complexities, it is crucial to keep systems up to date with the latest upgrades. However, some government organizations, such as blood banks, still rely on traditional systems that lead to rising complexities and an inefficient system. To address this, a proposed solution aims to shift blood banks' traditional processes into digital technology. This approach would not only provide a highly efficient method for resolving huge complexities, but it would also enable proper analysis of vast amounts of data to achieve predictions that could help prevent future anomalies and medical emergencies caused by poor traditional data management and processing techniques. The predicted results would also help maintain a proper track record of future blood unit stock supply for the respective blood banks. Overall, the proposed system would offer an effective solution to the challenges faced by government blood banks.

Keywords : Data Analysis, Machine Learning, LSTM

I. INTRODUCTION

A blood bank functions similarly to a storage facility for a sufficient supply of blood units from various blood types. These blood banks typically store all the transaction information of donors and recipients using a straightforward data management system, and in some cases even file systems. These strategies were simple to use for transaction data records on a smaller scale. However, as complexity grew significantly as a result of the population growth boom, these systems became incredibly ineffective in processing and handling data. These techniques have serious flaws that led to serious ineptitude in medical situations. The slower data management and processing systems prevented medical aids from reaching the vital areas.

To avoid serious crisis situations, traditional techniques and technology urgently need to be upgraded. One of the most effective approaches to assist in resolving the complexities of the previous systems is the idea of E-Blood banks. The e-Blood bank system is an improved version of the previous file management system used by blood banks. All blood transaction and donation records are now properly serialised and stored in an effective digital database system that is free of any binding structures, allowing the system to process large amounts of data instantly and solving complex problems at a breakneck speed. One of the important characteristics of E-Blood bank systems is their proper analysis of data, effective segregation of that data, and capacity to handle and analyse massive amounts of data.

II. LITERATURE SURVEY

Information about blood bank systems in urban and rural areas was gathered during the survey. According to the results, hospitals in urban areas have their own blood bank facilities, but rural areas lack them because there are fewer hospitals there and maintaining them is expensive. Many businesses have their own online blood bank system, which eliminates the need for human labour and speeds up the

process of preserving and managing information. This study compares various existing systems and offers suggestions for how to make the current system better. This system's foundations are research-based conclusions from numerous blood bank analyses.

Some systems use unique techniques and materials to ensure the security of blood provided by individuals to banks, necessitating a thorough examination of the relationships between donors and banks. This project's primary goal is to assess the frequency and main causes of blood donor deferrals.The outcome provides a wide perspective on the group of blood donors whose information was taken from the DBDC database. [1]

a web application that is cloud-based. Blood banks use a cost recovery system that donors are unaware of, which could lead to awkward situations. In this proposed solution, rural areas now have access to an SMS-based service via cloud data.

because blood banks in remote locations lack the necessary infrastructure for long-term blood storage. [2]

The use of a web-based project has also sped up the delivery of blood services and improved patient care in emergency situations. This technology uses GPS position monitoring to locate the closest donors who are close enough to the patient's location so that blood can be donated promptly. Donors will receive a notification by SMS Gateway as soon as their position is tracked so they can contribute. [3]

It is required to analyse the data in each domain. Here, the term "domain" refers to a blood bank, which is in charge of managing donor blood that is collected and used as needed. The RDBMS idea is used on the System to store data. Helping public and commercial decision makers solve problems and address policy challenges is the main goal of system analysis. [4] The system shifts its primary attention to data collection methods and a recent application on data mining for blood donors. The proposed system contains extensive research on various data collection methods, which is beneficial in the study of systems that would be precise for donor and patient data mining. Data.

. Data analysis based on a dataset of blood donors serves as the system's foundation, which is also provided by data mining researchers. [5]

A cloud-based Android application that searches for donors nearby and periodically stores data on the cloud enables hospitals to receive faster access or service. Where accident-prone areas are present, this service is most frequently chosen since it makes it simple for patients to transport to the closest hospitals. The primary objective of blood banks is to alleviate the blood shortage. [6]

This proposed system has provided a proposal that compares the proposed and existing systems. There is no such system that connects blood banks with donors. As a result, this proposed system suggested strengthening weak points in the new system that were reflected in the previous systems. [7]

A system with Java/JSP implementation that integrates with the phone gap mobile service is also available. The primary goal of study is to create a system that can manage blood bank operations and provide a potential solution to the issues currently plaguing blood banks and blood donation facilities. [8]

A straightforward blood bank research also covers the administration of data management for donors' and patients' relevant information at blood banks. The technology will let the designated blood bank officer simply handle the records of the blood donors and the patients in need of blood by logging in using a secret password. a trustworthy system that guards against intentional threats, such as unauthorised access or data leaking. It offers automated donor code registration for each type of blood, saving data on these platforms as opposed to using large files. [9]

III. METHODOLOGY

The system aims to design the most ideal methodology for E-Blood bank donation and transaction data management and give the blood banks a path towards faster data processing and proper management technologies aiding in the resolution of complexities after conducting a thorough analysis of the various methodologies and systems being used in the blood banks. Additionally, the system will be able to

anticipate future stockpile requirements for the specific blood banks, ensuring that the blood banks are always well stocked with supplies for any serious medical emergency.

DataCollection:

Data collection from a range of sources, including NGOs, private hospitals, and organisations.The data features provided by the instances of blood donors have had an impact on some data collection methodologies.The data collection process must be precise and secure.

This is the first or initial phase of the project's implementation. Information was gathered from a number of places, including blood banks, hospitals, online resources, etc. Since the data gathered from all the sources was insufficient, some data had to be manually produced.

Donor/patient profiles and Blood Bank/Hospital profiles are among the various profiles employed in the project. Donor information with all relevant criteria is stored in donor profiles and is accessible in case of future emergencies. Blood banks' information, such as their location and the types and quantities of blood they have available, is stored in blood bank profiles

Data Manipulation:

Data Cleaning, Data Transformation, and Data Reduction are the many phases of data preprocessing. Filling in missing data and removing noisy data were the key goals of data cleaning. Data gaps were filled using either an approximation method or the most likely values. A clear and noise-free dataset will result from this. Data normalisation, discretization, and appropriate attribute selection were used to alter the dataset.

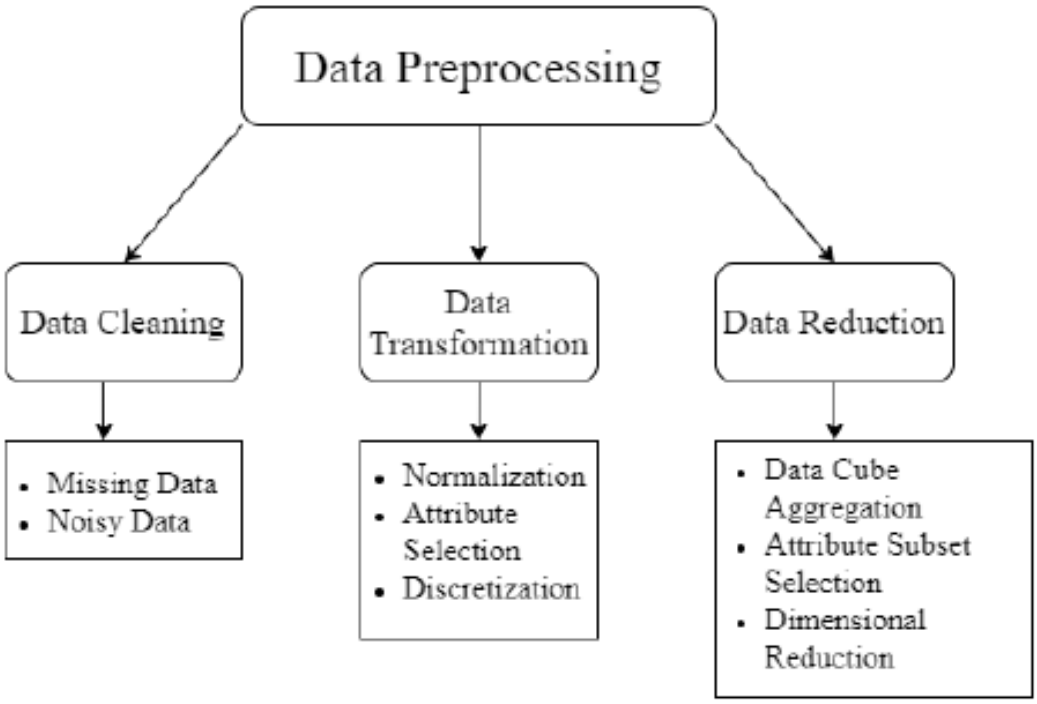

Fig 1: Data Pre-processing

The data obtained for the analysis of the needs in various modules, such as the patient and donor modules, is presented graphically in this phase.Utilise a variety of machine learning algorithms to train the module and monitor correctness.Using a training module with a larger dataset improves accuracy. Data cube aggregation, attribute subset selection, and dimensionality reduction were used to reduce the amount of data. C. Data analysis: The dataset was examined both before and after data preprocessing. Prior to preprocessing, analysis was conducted to determine which data needed to be cleaned, converted, and reduced. Analysis was then conducted to see whether all changes had been made correctly or not.The dataset must be managed correctly and in a single standard format after preparation. The dataset is therefore validated and processed into a standard format at this point.

LSTM technique

Now that the dataset is prepared, a machine-learning method should be applied to it. Automating a model or system using machine learning algorithms will produce the required output. The model is trained to accomplish this.

The stacking LSTM technique is utilised in this case since it will produce results with excellent accuracy and is simple to apply. Consequently, an LSTM algorithm is used to train the model, and predictions are made on how

The information will include the number of units of each blood group used by a certain blood bank each month. Using the LSTM technique of machine learning, this will be utilised to estimate the blood bank's future needs.

A MinMax scaler is employed with the feature range (0,1) because LSTM are sensitive to the scale of the data. The constructed dataset is initially split into training and testing data because the algorithm utilised is a supervised learning algorithm. Time series data are split into train and test sets in a certain order because each day's output depends on the preceding day's output.

A training dataset is utilised to train the model, and the results of that training are used to forecast the number of each blood type that will be needed in the test data. Following this, the amount of blood needed for the following month is predicted. The correctness of this output is then evaluated by comparison with the test dataset. The mean squared error method is also used to calculate error in addition to accuracy. A layered LSTM model is used throughout the process.

A different variable is defined for each distinct dataset. Additionally, two distinct variables are defined for both the test and train datasets. The variable designated for it holds the anticipated outcome.

A layer made of long short-term memory cells (LSTM) is an important technique employed. [16][17] Adding 'gates' to reroute the data flow through the network, LSTM cells extend feedback-based Recurrent Neural Networks. The gates enhance data durability by enabling the network to remember, forget, and choose data that is going through or being stored in it.

IV. RESULT

Blood Group	Actual Quantity	Predicted Quantity
A+	110	100
A-	125	130
B+	90	90
B-	60	65
AB+	110	105
AB-	100	90
O+	130	140
O-	150	140

The above table represents the actual and predicted quantity of blood of each blood group in a particular blood bank. Similarly, the model implements the results of each month and stores it and helps to draw the graph with that data.

The actual quantities are not calculated but they are real(randomly generated) quantities. The predicted quantities are calculated using:

$z = a1x1 + b1y1 + a2x2 + b2y2+ ... + anxn + bnyn$ z is the result

x1, y1 is the coefficient for the first feature x_n,y_n is the coefficient for the nth feature

Using timesteps values the features are calculated andthese calculated features help to predict the new output for test data. This also helps in predicting the next month's output.

V. CONCLUSION

The standard methods used by blood banks to store records of blood transactions are examined and assessed in this proposed system. On the basis of this, a more effective method of E-Blood banks is created, one that does away with the complexities and shortcomings of the current technologies and uses a high processing system with systematic storing of enormous amounts of data. With the use of machine learning algorithms, transaction data is properly managed and stored, and this analysed data set is then used to anticipate the availability of blood units based on the graphical output from the algorithms.

This suggested approach will assist in comprehending effective data management techniques for blood banks that are used for further analysis and, based on that information, the forecasting of future blood unit needs in the

specific blood banks. Additionally, any serious medical emergencies can be avoided by timely blood supply replenishment.

REFERENCES

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

[2] 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

[3] Dr.J.Shanthini,SanthoshiA,Manjula T, Pavithra R," *Computer Aided Emergency Service System", IJIRT, National Conference on Networks, Intelligence and Computing Systems*,March 2017.

[4] Gregory P. Prastacos, "*Systems Analysis in Regional Blood Management, Springer-Verlag Berlin Heidelberg"*,1981

[5] Nor Syuhada Che Khalid,"*Review of Data Gathering Techniques on Blood Donors Data Mining"*, November 2015.

[6] Shubham Pande, Shweta Mate, PradnyaMawal, AyushaJambulkar, Prof. N. S. More,"*E-Blood Bank Application Using Cloud Computing", International Research Journal of Engineering and Technology (IRJET)*,Feb-2015

[7] 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.

[8] Ibrahim FawzeAkar, Tukur Anas Mohammad, Mohamed Ismail Z ,"*CBBR Centralized Blood Bank Repository", International Journal of Information System and Engineering Vol. 3*, April, 2015.

[9] Clemen Teena, K. Sankar and S. Kannan," *A Study on Blood Bank Management ", Middle-East Journal of Scientific Research*, 2014.

[10] Shen Rong ,Zhang Bao-wen, "*The research of regression models in the machine learning field", MATEC Web Conf. Volume 176, 2018 6th International Forum on Industrial Design (IFID 2018)* , MATEC Web of Conferences 176,01033 (2018).

[11] AfrahAlmamri,"*Applied Research Project Report Enhancing Blood Transfusion Safety Through the Use of Online Blood Bank Management System*", October 2018.

[12] BharathwajMuralidaran , Akshay Raut , Yogesh Salve , ShivshankarDange ,LikheshKolhe, *"Smart Blood Bank as a Service on Cloud"-IOSR Journal of Computer Engineering (IOSR-JCE),* (Mar-Apr. 2016).

[13] J. Arul Valan,Dr. E. Babu Raj," *Machine Learning and Big Data Analytics in IoT based Blood Bank Supply Chain Management System*", Vol-4, Issue-12, Dec-2018.

[14] Dr. E. Babu Raj," *Machine Learning and Big Data Analytics in IoT based Blood Bank Supply Chain Management System*", Vol-4, Issue-12, Dec-2018.

[15] Khushbu Kumari, Suniti Yadav," *Linear Regression Analysis Study", Journal of the Practice of Cardiovascular Sciences 4(1)*,33 January 2018

[16] Chao-Ying Joanne Peng, Kuk Lida Lee, Gary M. Ingersoll, "*An Introduction to Logistic Regression Analysis and Reporting", The Journal of Educational Research 96(1):3-14*, September 2002.

[17] Anita Yadava, C K Jhaa , Aditi haranb (2019), "*Optimizing LSTM for time series prediction in Indian stock market", International Conference on Computational Intelligence and Data Science* (ICCIDS 2019)

[18] SimaSiami-Namini, Neda Tavakoli, Akbar SiamiNamin (2018), *"A Comparison of ARIMA and LSTM in Forecasting Time Series", 17th IEEE International Conference on Machine Learning and Applications (ICMLA),* December 2018

[19] Xun Sheng Du, Huaqing Zhang (2017), "*Stacked LSTM Deep Learning Model for Traffic Prediction in Vehicle-to-Vehicle Communication" , IEEE 86th Vehicular Technology Conference* (VTC-Fall). September 2017.

[20] *https://www.researchgate.net/profile/Ayesha-Sahar/publication/328819708/figure/fig4/AS:845833 8377046027@1578674992640/Simple-LSTM-VsStacked-LSTM.png*