Modern Education
and Computer Science
PRESS

Predicting the Behavior of Blood Donors in National Blood Bank of Ethiopia Using Data Mining Techniques

Teklay Birhane, Brhanu Hailu
Department of Information Science, Mekelle University, Tigray, Ethiopia
Email: {teklaybirhane12, brhanuhaylu}@gmail.com

Abstract: A modern technology used for extracting knowledge from a huge amount of data using different models and tasks such as prediction and description is called data mining. The data mining approach has a great contribution on solving a different problem for data miners. This paper focuses on the application of data mining in health centers using different models. The model development process helps to identify or predict the behavior of blood donors whether they are eligible or ineligible to donate blood by their right status way and protects any blood bank health center from the collection of unsafe blood. Classification techniques are used for the analysis of Blood bank datasets in this study. For continuous blood donors, it will help to enable to donate voluntary individuals and organizations systematically. J48 decision tree, neural network as well as naïve Bays algorithms have been implemented in Weka to analyze the dataset of blood donors. The study is used to classify the blood donor's eligibility or ineligibility status based on their genders, deferral time, weight, age, body priced, tattoos, HIV AIDS, blood pressure, donation frequency, hepatitis, illegal drug use attributes. From the 11 attributes, gender does not affect the result. We have used 1502 datasets for the train set and 100 datasets for testing the model using cross-fold validation. Cross-fold data, partition was used in this study. The efficiency and effectiveness of the algorisms are measured automatically by the system. The obtained result showed that the J48 classifier outperforms the best result as well as both neural network and navies, Bayes, in terms of matrix evolution, with its 97.5% overall model accuracy has offered interesting rules.

Index Terms: Data Mining, Blood Donation, Classification, Decision Tree, Neural Network, and Navies Bayes.

1. Introduction

The process of finding information from large repositories through the manual system is difficult nowadays since data size has increased at alarming times. A computer-based method has been applied to find necessary information from data warehouses, repositories in which data is available such as business organizations whether internet-based or offline. Based on that data mining have conceptualized, in the 1990s as a means of analyzing a vast amount of data that is available in different organizational repositories such as healthcare, education system, market places, and governmental offices[1]. Data mining is an automatic knowledge discovery process from immense databases in a given application and its main-goal is to find out hidden and interesting information, patterns, and useful knowledge [2]. As blood banks has massive volume of data, the use of data mining techniques is necessary to analyze and convert them into useful and interesting knowledge. Therefore, the researchers used data-mining techniques for predicting continuous behavior of healthy blood donors in the Ethiopian national blood bank in the case of the Harari branch. The main aim of predicting blood donor's behaviors is to know and gather relevant information about the donor's health status and group under continuous giving blood for a blood bank in the proper way.

Currently in Ethiopia, an urgent need for blood is growing at an alarming rate as the number of medical surgeries, accidents and health diseases is growing on increasing rate. There is too few number of blood donors in Ethiopia. Because of that, the Ethiopian National Blood Bank falls to crucial issues of blood shortage and unsafe blood collection. Those problems leads to death and serious bad health consequences. Since the behavior of blood donors is unknown, it is becoming too difficult to extract information using conventional database techniques to solve such kinds of a serious problem. In fact, that the researchers stand to conduct the study using data mining techniques to predict the continuous behavior of blood donors in the case of the Harar branch.

The main objective of this study is to predict the behavior of blood donors in the national blood bank of Ethiopia as eligible or ineligible to donate using data-mining techniques to extract useful information, generate new patterns and knowledge that helps to collect safe blood and increase the number of blood donors in Ethiopia.

2.   Review of Related Literatures

A.   National Blood Bank of Ethiopia

Health and community care have become a cornerstone of humanitarian assistance, Therefore the Ethiopian Ministry of Health establishes the Ethiopian Red Cross society National Blood Bank Services (ERCS-NBBS) in 1969 that aims to reduce the vulnerability of disease and solve public health crises in Ethiopian communities[7].

ERCS is the sole organization providing Blood bank services across the county, with its central blood bank located at Addis Ababa and eleven regional blood banks found in Adama, Harar, Diredawa, Jijiga, Yirgalem, Arbaminch, Jimma, Bahir Dar, Gondar, Dessie, and Mekelle [7].

B.   Blood Donation Process

Donating blood is too simple and safe. The process takes approximately 30 minutes starting from registration to refreshments. The overall procedures and steps followed in the course of blood donation processes are clearly stated as follows:

Step 1: **Registration**: The donors general information (such as name, age, sex…..etc.) is recorded in the reception area to maintain accurate information.

Step 2: **Medical interview**: every donor meets privately with blood bank experts to review their recorded medical history and this information is kept confidential.

Step 3: **Physical status check**: The health status of the donor like weight, Hepatitis, HIV test, body priced, tattoo, and blood pressure are checked to assure if they fit the minimum requirement to donate during blood donation process.

Step 4: Every donor is taken to the actual donation area, where a phlebotomist sterilizes the area of the donors' arm from which the blood is drawn.

Step 5: **After donation**, the donor is allowed to the canteen area to take enough rest for approximately 10-15 minutes and is suggested afterward to stop any activity.

Step 6: Finally taking a very intensive screening for every unit of collected blood for measuring safety and the transfusion-transmissible infections.

C.   Definition of Data Mining

Data Mining [3] is the process of discovering interesting patterns and hidden knowledge from a large amount of data sources. The source of data for data mining can be data warehouses, repositories, webs, and organizations. It is also an iterative sequence of data for the knowledge discovery process [3].

According to Fawzy and Moussa[4], Data mining is explained as the process of analyzing datasets in novel ways to find unsuspected relationships. The unknown and useful information is gained from data through data mining models or patterns that extract implicitly. This is needed for future prediction trends and behaviors to make decisions in low time consumption to answers business questions. To process and analyze several types of data patterns different data mining techniques are studied. The most popular data mining tasks are classification, summarization, association rules, and clustering.

D.   Data Mining and KDD Process

Data mining is a technology that is used for analyzing a huge amount of data and screen out the relevant ones. It refers to extracting or discovering knowledge from large amounts of data [1]. However, the process of Knowledge Discovery from Database (KDD) in data mining is very expensive, especially in the stages of business objectives elicitation, data mining objectives elicitation, and data preparation. This is especially the case each time data mining is applying to a blood bank. It can also define as fetching relevant information i.e. Knowledge from the large repositories of data. That is the reason why many people are also called Knowledge Mining.

The following figure (Fig.1.) shows data mining as simply an essential step in the process of KDD.
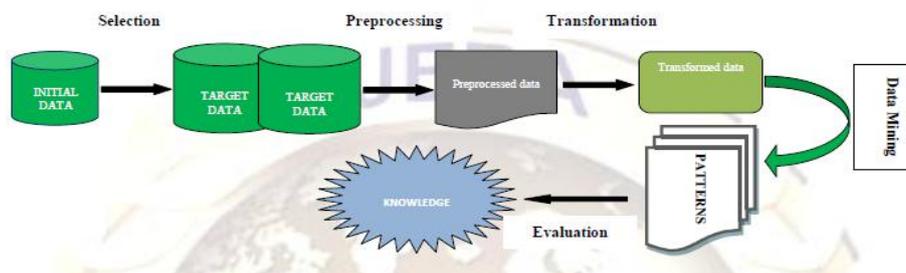

Fig.1. Data mining is the core of the KDD Process[1].

As the above figure indicates, the KDD process has six steps, which lead from data collection to knowledge representation. Most probably, the KDD process has two phases the data preparation phase and the data-mining phase. Data preprocessing or data preparation includes data collection, data selection, data cleansing, data integration, and data transformation. Data collection is the process of gathering useful data from a different source of databases or data warehouses. Data cleansing is a fundamental step in which noise data and irrelevant data are going removed from the collection. In the Data integration step, multiple data sources are combined in a common source based on their internal relationships and common sense. At the data selection step, the decision of relevant data to the analysis process has been obtained. Data transformation is a stage in which the selected data is going transformed into appropriate forms for the mining procedure. In this stage, data is perfectly prepared for mining and the data mining process is comes after. Data mining is the core step of the KDD process in which clever techniques such as classification and clustering are applying to extract potentially useful patterns. After this step, very interesting patterns representing knowledge have been discovered. Finally, the discovered hidden information/knowledge is visually represented to the user and applied to a specified domain problem.
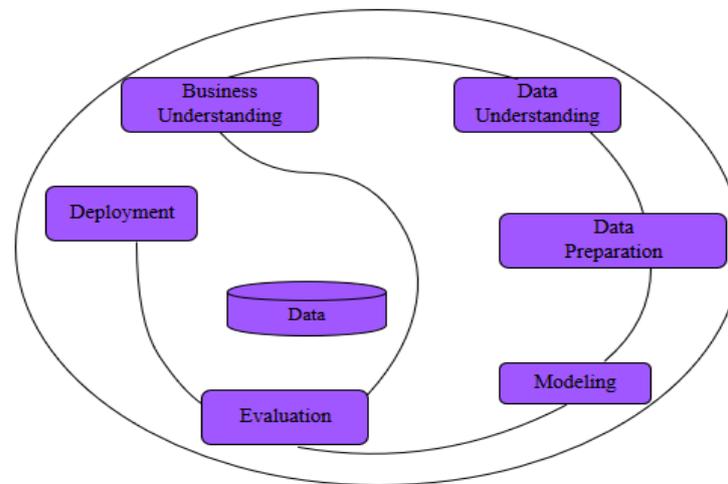

Fig.2. The CRISP-DM process model[6,7].

E.   Data Mining Models

Modeling is an iterative process and different from supervised and unsupervised machine learning [5]. It may be either description or prediction. The selection of the data mining modeling techniques is based on the problem of data mining and the type of data found in the real world. Typically, many techniques can be used for the same problem in data mining. According to Fawzy [5], Some techniques have specific requirements in the form of data. In a complex business world, it requires many coordinate efforts of different stakeholders such as experts, in organizations [5]. Many kinds of literature in data mining have tried to build a general framework to serve as a blueprint on how to organize, gathering, analyzing, and disseminating data results, implementing results, and monitoring improvements. There are different types of data mining models among them one of the most known for data mining project modeling is the CRISP-DM data mining process model.

CRISP-DM stands for Cross-Industry Standard Process for Data Mining developed by two vendors ISL (now part of SPSS) and NCR Corporation, which are the world's leading supplier of data warehouse solutions and it is a standard process model in industries which consisting of a sequence of steps that are usually involved in a data mining study [6].

The CRISP-DM process model has six life cycle processes steps. These are business understanding, data understanding, data preparation, modeling, evaluation, and deployment [6,7].

Step1: Business Understanding

The initial phase in CRISP-DM mainly focuses on understanding the objective and requirement of knowledge in business projects, based on the specific objective converting the knowledge into data mining problems definition and plan designed to achieve the objectives [6,7].

Step 2: Data Understanding

Data understanding is the second phase of CRISP-DM that starts with an initial data collection, to identify data quality problems, to find data, or detect relevant subsets of hidden information using hypotheses. [6,7].

Step 3: Data Preparation

Data preparation is the third phase of CRISP-DM and it is the process of formulating the final dataset from the initial raw data. It follows table, record, and attribute selection as well as transformation and cleaning of data for modeling tools tasks.

Step 4: Modeling

In this phase, different modeling techniques have been selected and applied and their parameters are calibrating to optimal values. For the same data, mining problem different techniques can be used. A technique depends on the specific requirement and form of data. Therefore, stepping back to the data preparation phase is necessary for problem solving [6,7].

Step 5: Evaluation

Evaluation of the model and review the steps taken to build it to be quite sure to achieve specified business objectives is very critical before proceeding to final model deployment stage. At the end of this phase, a decision should be reached on how to use the data modeling results through assessment of data mining results concerning business success criteria to approve models, review of the process, and determine next steps such as possible actions and decisions making based on patterns and relationships[6,7].

Step 6: Deployment

In general, model construction is not the end process of CRSIP-DM. Even if the purpose of the model is to increase knowledge of the data, the extracted knowledge will be organized and presented in a way that the customer can use it. The main activities in this phase are planning deployment, monitoring, maintenance, presenting, reviewing the project, and producing documentation.

F.   Related Works

Nowadays the advantage and applications of data mining are growing fast in different sectors such as the health system, education, and e-commerce, etc.

Ramachandran *et al* [2] had studied classifying blood donors using data mining techniques. The study aims to identify voluntary, regular blood donors in blood donation camps inefficient manner. The analysis had carried out using J48 classifier algorithm applying to a standard blood group donor's dataset. J48 is selected based on the classification accuracy models. The researchers used attributes of blood donors based on sex, blood group, weight, and age. It may be succeeded by collecting necessary data using data mining techniques and selecting the best implementation tools for the domain area. Nevertheless, the researcher has used very small attributes to classify the blood donors and the class of classification type is not clear.

3.   Material and Methods

A.   Dataset Explanation

The real-life initial dataset is collected from Harar health center it is located in the capital city of Harar regional state. The data is organized in Microsoft office Excel format for further processing. The data has 11 attributes and 1502 records/instances with no missing values. The real datasets are briefly described in Table 1.

Table 1. List of initial dataset attributes used for the study.

No	Attribute Name	Data Scale	Description
1	Gender	Nominal	The donor's biological identity {male or female} and is not a criterion for donation.
2	Age	Real	The lower and upper age limit for blood donation in most countries is 18 and 60-70 years respectively.
3	Blood Pressure	Nominal	A normal blood pressure (systolic 120–129 mmHg, diastolic 80–89 mmHg) is generally regarded as an indicator of good health status.
4	Weight	Real	A donor should weigh at least 50 kg to donate.
5	Donation frequency	Nominal	The occurrence of blood donation
6	Hepatitis	Nominal	Blood donors with positive hepatitis should not accept to donate.
7	Body pierced	Nominal	A donor with any price of the body is not acceptable for donation.
8	Tattoos	Nominal	A donor with any tattoo is not acceptable for donation it should wait at least 2 years.
9	HIVAIDS	Nominal	If a blood donor has HIV positive must not donate blood.
10	Illegal Drug Use	Nominal	Using illegal drugs like cocaine and smoking cigarettes makes out of selection criteria for donation.
11	Deferral	Nominal	The gap in the donor's status.

B.   Data Pre-Processing

Data preprocessing helps to fill in some missing values, to detect some outliers that may jeopardize the result of data mining, and to detect and remove some noisy data. Additionally, it provides collecting required data from filtered data fields, indicating suitable methods to relieve symptoms of missing data and comprising redundant data, Preprocessing and transformation of the dataset are done using Waikato knowledge analysis (Java platform used for various machine learning and data mining techniques [12].

C.   Tools and Techniques

After analyzing and comparing different data mining models, classification is selected to develop this model. It is one of the leading tasks with the specific aim of strongly trying to predict the value of a single nominal variable based on the predefined groups and known values of other components. The present work intends to build a mining model based on three classification algorithms to provide a solution for the prediction of blood donor's health status behavior. The classification models used by the authors are presented as follows.

J48 Decision tree*:*

Decision tree strongly needs for the prediction and classification of medical investigation in a good manner. Multiple decision tree algorithms are resented to classify real and synthetic datasets, such as C4.5, J48, ID3, C5, CHAID, and CART [11]. In the research work done in this paper, the J48 decision tree algorithm has been chosen to build the model. It is a simple C4.5 decision tree for classification and creates a binary tree. Each input variable corresponds with edge node has possible value and easily search item through the path [8].

Naive Bayes:

A naive Bayes is a probabilistic-based classifier algorithm that assumes the value of a particular feature is unrelated to the presence or absence of any other feature, in the given class variable [9].

Neural Network:

Represents a brain metaphor for information processing. These models are biologically inspired rather than a replica of how the brain functions [10]. Neural networks have been forecasted many application systems in business classification because they can "learn" from the data, that have nonparametric nature (i.e., no rigid assumptions), and their ability to generalize [10].

D.   Proposed Model

This paper explores the aspect of Decision Tree, Naïve Bayes Classifier, and Neural Network as Data Mining techniques used to create models to predict the behavior of blood donors using classification data mining tools on real medical data. The data were divided into a training set and test set by cross-validation [11]. To access the predictive performances, 10- fold cross-validation was used on each configuration. Under such a scheme, for any provided run the dataset is randomly split into 10 subsets of similar size.

After data pre-processing, the J48, Naïve Bayes, and NN algorithms are applied to the dataset as a reputed machine learning algorithms through the WEKA platform for data mining techniques). Data are divided into "Eligible" or "Ineligible" based on the targeted result of the model that is constructed. All models were analyzed based on correctly classified instances. Figure 3. Depicts the overall research activities to construct the model.
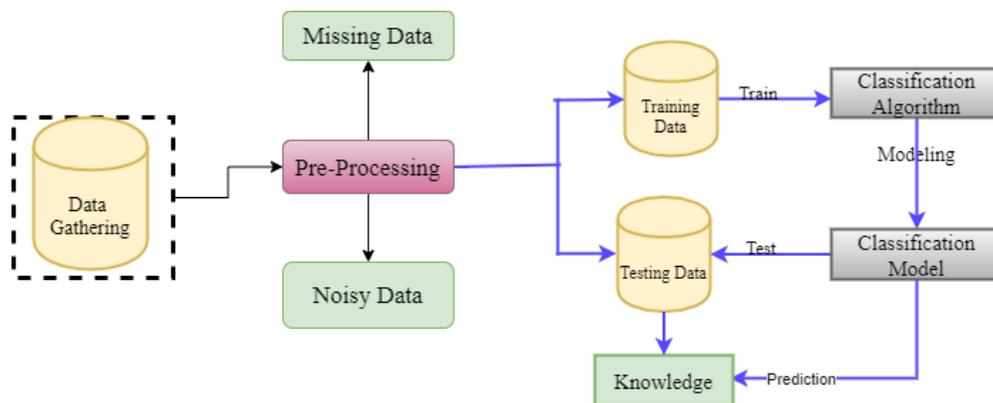

Fig. 3. Prediction Model.

E.   Evaluation Measures

To test the performance of the developed classification models; five measures were used [11, 13, 14]: accuracy, recall, precision, F-Measure, and computational time. The measures described in table 2 are interpreted based on the formula of the confusion matrix in Table 3. The confusion matrix demonstrates the correct and incorrect classification of classes. The correct classifications denoted by (TP) True Positive and True Negative (TN) whereas the incorrect classified denoted by false positive (FP), that is when an outcome is predicted as positive while it is an actual negative and false negative (FN) if an outcome is classified as negative but it is positive.

Table 2. Evaluation Measures

Measure	Formula	Meaning
Accuracy	(TP + TN) / (TP + TN + FP + FN)	The percentage of predictions those are correct.
Precision	TP / (TP + FP)	The percentage of positive predictions those are correct.
Recall	TP / (TP + FN)	The percentage of positive labeled instances that were predicted as positive.
F-Measure	2(PR)/(P+R)	A combined measure that assesses precision/recall tradeoff
Computational Time	-	Time is taken to build a model (in seconds)

Table 3.Confusion Matrix

Actual class\Predicted class	Atribute1	Atribute2
Atribute1	True Positives (TP)	False Negatives (FN)
Atribute2	False Positives (FP)	True Negatives (TN)

4.   Experimental Results and Analysis

The result gained from the discussion or experimentation of the given training dataset to predict the blood donor's behavior using classification rules with the most common algorithms such as J48, Navies Bayes and Neural Network are as follows.

A.   J48 Classifier Experimental Result Analysis

Classification of blood donors as eligible or ineligible gives the following model. The output was generated by Weka using the J48 Classifier algorithm.

J48 pruned tree

```
BloodPressure = high
|  Bodypierced = no
|  |  Weight <= 49: Ineligble (157.0/8.0)
|  |  Weight > 49
|  |  |  Hepatitis = Negative
|  |  |  |  HIVAIDS  = Negative
|  |  |  |  |  Donationfrequency = abovetwomonth
|  |  |  |  |  |  IllegalDrugUse = no
|  |  |  |  |  |  |  Tattoos = no: Eligible (356.0/6.0)
|  |  |  |  |  |  |  Tattoos = belowoneyear: Ineligble (17.0)
|  |  |  |  |  |  |  Tattoos = aboveoneyear: Eligible (25.0/2.0)
|  |  |  |  |  |  IllegalDrugUse = yes: Ineligble (10.0)
|  |  |  |  |  Donationfrequency = N/A
|  |  |  |  |  |  Age <= 17: Ineligble (45.0/11.0)
|  |  |  |  |  |  Age > 17
|  |  |  |  |  |  |  IllegalDrugUse = no
|  |  |  |  |  |  |  |  Tattoos = no: Eligible (239.0/5.0)
|  |  |  |  |  |  |  |  Tattoos = belowoneyear: Ineligble (14.0)
|  |  |  |  |  |  |  |  Tattoos = aboveoneyear: Eligible (15.0)
|  |  |  |  |  |  |  IllegalDrugUse = yes: Ineligble (5.0)
|  |  |  |  |  Donationfrequency = belowtwomonth: Ineligble (110.0/1.0)
|  |  |  |  HIVAIDS  = Positive: Ineligble (55.0/1.0)
|  |  |  Hepatitis = Positive : Ineligble (5.0)
|  |  |  Hepatitis = Positive: Ineligble (55.0)
|  Bodypierced = yes: Ineligble (112.0)
BloodPressure = low: Ineligble (281.0/3.0)
Number of Leaves:    16
Size of the tree:        27
```

The following correctness are J48 algorithm result based on the given dataset.

Table 4. Performance Results from J48 Classification Algorithm

	No. of Instances	Percentage
Correctly Classified Instances	1463	97.4684 %
Incorrectly Classified Instances	38	2.5316 %

Table 5. Confusion matrix for J48 Classification Algorithm

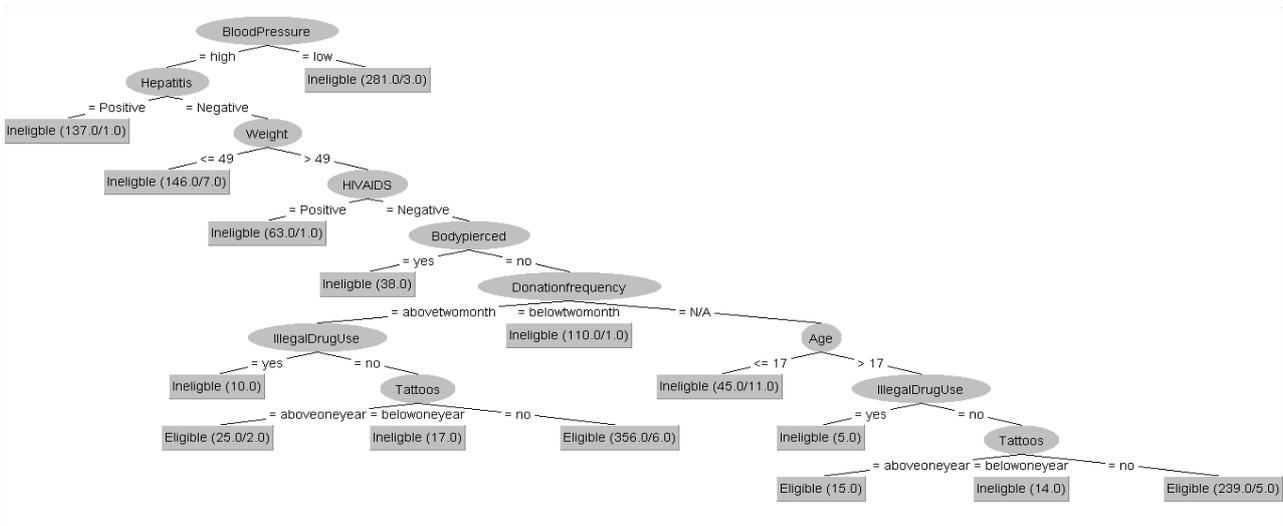| Actual class\ Predicted class	A	B	
A	622	24	A = Eligible
B	14	841	B = Ineligible

Fig. 4. Decision Tree Model using J48 Algorithm to classify blood donor's behavior.

B. Neural Network Classifier Experimental Result Analysis

The second algorithm we used to classify and predict the blood donor's behavior is the neural network classifier algorithm. The evaluation of neural network experimental result detailed accuracy by class using cross-validation generated as follows.

Table 6. Performance Results from Neural Network Classification Algorithm

	No. of Instances	Percentage
Correctly Classified Instances	1416	94.3371 %
Incorrectly Classified Instances	85	5.6629 %

Table 7. Confusion matrix for Neural Network Classification Algorithm

Actual class\ Predicted class	A	B	
A	613	33	A= Eligible
B	52	803	B= Ineligible

C. Naive Bayes Classifier Experimental Result Analysis

The third model experimented with for predicting the behavior of blood donors was the Naive Bayes algorithm. Some generated rules or outputs using the Naïve Bayes algorithm are here. The rules produce are unambiguous in that it does not matter in what order they are executed. The following are some of the rules extracted from algorithms even they use a different method to predict the true value but the general truth learns the algorism are the following points.

IF Age>=18 AND Blood Pressure= high AND donation frequency>=two month AND Hepatitis=no AND body pierced= no AND tattoos=no AND HIVAIDS=negative AND illegal drug use= no AND Deferral= permanent
THEN eligible ()
IF Age>=18 AND Blood Pressure= high AND donation frequency>=two month AND Hepatitis=no AND body pierced= no AND tattoos=yes AND above one year AND HIVAIDS=negative AND illegal drug use= no AND Deferral= permanent
THEN eligible ()
IF Age>=18 AND Blood Pressure= high AND donation frequency>=N/A AND Hepatitis=no AND body pierced= no AND tattoos=yes AND below one year AND HIVAIDS=negative AND illegal drug use= no AND Deferral= temporal
THEN eligible ()
Else
All in the case out of them are Ineligible.

Table 8. Performance Results from Naive Bayes Classification Algorithm

	No. of Instances	Percentage
Correctly Classified Instances	1379	91.8721 %
Incorrectly Classified Instances	122	8.1279 %

Table 9. Confusion matrix for Naive Bayes Classification Algorithm

Actual class\Predicted class	A	B	
A	631	15	A = Eligible
B	107	748	B =Ineligible

The following is the summarization of the three classification model results. A tabular comparison (table 10) shows the classification results of the J48 Decision Tree, Naïve Bayes, and Neural Network in terms of Accuracy, Precision, Recall, and F-Measure.

Table 10. Comparisons of Classification Algorithm

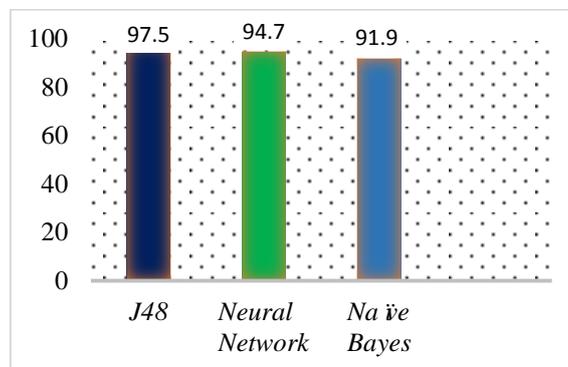| Model	Accuracy	Precision	Recall	F-measure	Computational Time
J48	0.975	0.975	0.975	0.975	0.06 sec.
Naves Bayes	0.919	0.926	0.919	0.919	0.03 sec.
Neural Network	0.947	0.947	0.947	0.947	5.54 se.

Fig. 5. The Classification Results

5. Results and Discussion

From the results obtained, the J48 decision tree has the lowest error rate; the J48 decision tree achieved high accuracy as compared to the other two algorithm implementations. Naïve Bayes has less computational time due to the omission of complex iterative estimations of the parameter. Although Neural Network model can complex, decision boundaries, the Training time of Neural Network is extremely slow. Therefore, the decision tree model can classify new blood donor's eligibility status with 97.5% accuracy. Also, the Precision = 0.975, recall =0 .975 and F-measure = 0.975 results were promising, good and acceptable.

The generated output from the sample of training test datasets for prediction for the unknown blood donor whether the current, status of his/her blood donor's behavior information is eligible or ineligible to make donation process.

Table 11. A test set of Classification Algorithm by J48.

=== Predictions on user test set ===		
inst#	actual predicted error	prediction
1	1:? 1:Eligible	0.979
2	1:? 1:Eligible	0.979
3	1:? 2:Ineligble	0.952
4	1:? 2:Ineligble	0.993
5	1:? 1:Eligible	0.983
6	1:? 2:Ineligble	0.989
7	1:? 2:Ineligble	0.989
8	1:? 2:Ineligble	0.989

6. Conclusion

In this paper, we have described classification techniques for Blood Donor's behavior datasets. We have used data mining classifiers algorithms to predict the behavior of blood donors such as the J48 decision tree, Neural Network, and Naïve Bayes. The development of a model or system, which is very essential for the timely analysis of huge Blood Donors datasets in the National blood bank of Ethiopia, is our primary concern in this study. Because of large data, exploring manual data analysis is not sufficient, therefore using data mining technology (computer-based) data analysis is sufficient. This technique will be applied to the blood group donor's database maintained in the Harar branch health center. This algorithm will be adapted to find conditions under which status of blood donors are correctly eligible to donate in emergencies and help to collect safe blood.

For future work, apply data mining techniques to blood type classification, diagnosing diabetic symptoms, and diagnosing cancer patients based on blood cells or predicting the cancer types on the blood groups is required. Improve performance of the study by increase the number of attributes and classification algorithm.

References

[1] Bhardwaj, A., Sharma, A. and Shrivastava, V. K. (2012). Data Mining Techniques and Their Implementation in Blood Bank Sector , International Journal of Enginering and Reaserch Application, 2, pp. 1303–1309.

[2] Ramachandran, P., Girija, N. and Bhuvaneswari, T. (2011). Classifying Blood Donors Using Data Mining Techniques, International Journal of Computer Science & Engineering Technology, 1(1), pp. 10–13.

[3] Kalaiyarasi, P. and Kalaiselvi, A. (2018). Data Mining Techniques Using To Weather Prediction, 6(3), pp. 249–254.

[4] Fawzy, D., Moussa, S. and Badr, N. (2016). The evolution of data mining techniques to big data analytics: An extensive study with application to renewable energy data analytics, Asian Journal of. Applied Science, 4(3), pp. 756–766.

[5] Belachew, R. (2013). Application of Data Mining Techniques for Customers Segmentation and Prediction: the Case of Buusaa Gonofa Microfinance Institution.

[6] Chapman, P., (2000). Step-by-step data mining guide, SPSS inc, 78, pp. 1–78.

[7] Gebregziabher, H., Meshasha, M. and Cerna, P. (2017). Predicting the Seroprevalence of HBV, HCV, and HIV Based on National Blood of Addis Ababa Ethiopia Using Data Mining Technologies, 1(2), pp. 44–55.

[8] Santhanam, T. and Sundaram, S. (2010). Application of CART algorithm in blood donors classification," Journal of Computer Science, 6(5), pp. 548–552.

[9] Dhoke, N. and Deshmukh, P. S. S. (2015). To improve Blood Donation Process using, pp. 4834–4840.

[10] Krieger, C. (1996). Neural Networks in Data Mining ©1996, pp. 1449–1453.

[11] Han, J., Kamber, M., and Pei, J. (2011). Data mining: concepts and techniques: concepts and techniques, Elsevier.

[12] http://www.cs.waikato.ac.nz/ml/weka/, accessed 27 June 2020.

[13] Chimieski, B.F., and Fagundes, R.D.R. (2013). Association and classification data mining algorithms comparison over medical datasets, Journal of health informatics, 5, (2).

[14] Ahmed, T.M.: (2016). Developing a Predicted Model for Diabetes Type 2 Treatment Plans By Using Data Mining, Journal of Theoretical and Applied Information Technology, 90,(2), pp. 181.

Authors' Profiles

Teklay Birhane received his BSc. Degree in Information Science from Mekelle University in 2017. MSc. in Information Science from Haramaya University in 2021. Currently he is a lecturer in College of Natural and Computational Science at Mekelle University. His research interests include Natural Language Processing, Machine Learning and Knowledge Discovery Process.

Brhanu Hailu is earned in 2017 his BSc. Degree in Information Science from Mekelle University. His research interests are in the areas of artificial intelligence, information retrieval system, knowledge management system, archival and document management system.