Anemic Status Prediction using Multilayer Perceptron Neural Network Model

C.H. Yu, M. Bhatnagar, R. Hogen, D. Mao, A. Farzindar, K. Dhanireddy

University of Southern California (USC), Los Angeles, USA

`{chinghay,dilinmao,farzinda}@usc.edu,`
`{Manas.Bhatnagar,Rachel.Hogen,Kiran.Dhanireddy}@med.usc.edu`

Abstract

Artificial Neural Network (ANN) has been well recognized as an effective tool in medical science. Medical data is complex, and currently collected data does not flow in a standardized way. Data complexity makes the outcomes associated with intra-operative blood management difficult to assess. Reliable evidence is needed to selectively define areas that can be improved and establish standard protocols across healthcare service lines to substantiate best practice in blood product utilization. The ANN is able to provide this evidence using automatic learning techniques to mine the hidden information under the medical data and come to conclusions.

Blood transfusions can be lifesaving and are used commonly in complex surgical cases. Blood transfusions come with associated risks and are costly. Anemia and clinical symptoms are currently used to determine whether a packed red blood cell transfusion is necessary. In this paper, we worked with unique datasets of intra-operative blood management collected from the electronic medical record of the Keck Medical Center of USC. We apply Multilayer Perceptron Neural Network to estimate missing values and predict the degree of post-operative anemia. Successful predictions of postoperative anemia may help inform medical practitioners whether there is a need for a further packed red blood cell transfusion.

Keywords: Machine learning, Artificial Neural Network, Prediction, Blood Transfusion, Multilayer Perceptron (MLP).

1 Introduction

Blood transfusions can be lifesaving and are commonly used in complex surgical cases. The type of blood transfusion needed depends on the situation: autologous transfusion, the collection and re-infusion of the patient's own blood or allogenic transfusion, which involves someone collecting and infusing blood of a compatible donor into him/herself. However, allogeneic blood transfusions come with associated risks such as allergic reactions, immunologic reactions, transmission of infections,

and increased susceptibility to post-operative infections (Sarode). In addition, the storage and distribution of allogeneic blood products is costly and time consuming for the hospital staff.

Reducing unnecessary blood transfusions has been a long standing goal of the medical community [Carson et al. Transfusion thresholds and other strategies for guiding allogenic red blood cell transfusion. Cochrane Database Syst. Rev. 2016]. Reducing the use of unnecessary allogeneic blood transfusions would decrease the complications associated with blood transfusion and benefit patients. In addition, it would reduce the costs of providing medical care. (Fisahn et al. Association Between Allogeneic Blood Transfusion and Post-Operative Infection in Major Spine Surgery. Clinical Spine Surgery, 2017; Li et al. Association between Allogeneic or Autologous Blood Transfusion and Survival in Patients after Radical Prostatectomy: A Systematic Review and Meta-Analysis. Plos One. 2017; Saag et al. Reducing Liberal Red Cell Transfusions at an Academic Medical Center. Transfusion. 2017; ect.]

Currently, hematologic laboratory values including anemic status in conjunction with clinical symptoms and history are used to determine whether transfusion is necessary. Machine learning could be used as an auxiliary tool to better inform clinical decision making and reduce unnecessary blood transfusions.

Artificial neural networks (ANN), one of the powerful machine learning approaches, have been increasingly used in medical science. Neural Networks have assisted in building reliable diagnostic systems to help detect cancer and heart problems. Neural Networks have also made successful contributions to biochemical analysis. The proposed method of Multilayer Perceptron (MLP) by Rosenblatt, 1957 has a very wide application in classification and regression which applied in different fields such as healthcare. Collobert and Benjio, 2004, studied the link between the classification algorithms by MLP and Support Vector Machines (SVM) (Vapnik, 1995). They showed it can be computationally expensive in time to train an SVM and can get a better performance in generalization and speed the models in MLP.

The objective of our study is to apply Multilayer Perceptron (MLP) ANN to accurately predict anemic status after surgery. The prediction of anemia status will help medical practitioners prepare the exact number of packed red blood cells for transfusion.

2 Data

This study was approved by the University of Southern California (USC) Institutional Review Board: #HS-16-00504. This was a single center, retrospective study using data on all surgical cases from the USC, Keck Medical Center from August 2014 to October 2016. Data was collected from the electronic medical record of the Keck Medical Center of USC. Intra-operative data was documented by the intra-operative circulating nurse.

The data included seventeen columns including: Surgery ID, Surgery Procedure, Start Date and Time, Stop Data and Time, Surgery Specialty, Patient Type, Blood B/M, BM Red Blood Cell, BM Fresh Frozen Plasma, BM Platelets, Results Before Surgery, Results After Surgery, Cell Salvage, Cell Salvage Recovered, Cell Salvage Returned, ANH, ANH Removed.

Allogeneic blood transfusions included in the data set were packed red blood cells, fresh frozen plasma, and platelets. Allogeneic blood transfusions are transfusions derived from another person. Autologous blood transfusions included in the data set were ANH and cellsaver. Autologous blood transfusions are those derived from one-self. Acute normovolemic hemodilution consists of the removal of autologous whole blood prior to the start of the surgical case, maintenance of normovolemia with crystalloids and colloids, and re-infusion of that whole blood when required during surgery or at the completion of surgery. Cellsaver consists of the collection of blood from the

surgical field, the washing and resuspension of red cells in isotonic solution, and the re-infusion of the red cells when required during surgery.

3 Artificial Neural Networks (ANN)

An ANN is an interconnected group of nodes that simulate the vast network of neurons in a brain. In the medical field, ANNs have been used since the late 1980s, and more recently, widely used in medical research. An artificial neuron is an information-processing unit conceived as a model of biological neurons. Artificial neurons receive signals from input links in an ANN. The transfer function (activation function) defines the output of that artificial neuron given an input or set of inputs. The behavior of the ANN depends on both the weights and the transfer function that is specified for the neuron. The cost function (objective function or loss function) is to return a number representing how well the neural network performed to map training examples to correct output. The algorithm's goal is to find a function that has the smallest possible cost.

The backpropagation is a common method used to train ANNs. The algorithm repeats a two phase cycle, propagation and weight update, until the algorithm gets the optimum solution. The algorithm first runs a "forward pass" to go throughout the network. By using a cost function, an error value is calculated for each of the neurons in the output layer. The error values are then propagated backwards from the output, layer by layer, until each neuron has an associated error value. The error value and learning rate are used to update weights.

The MLP that use backpropagation training is most commonly used. Backpropagation networks are networks where signals travel in one direction from an input neuron to an output neuron without returning to their source. Multilayer Perceptron consist of at least three layers of units: an input layer, at least one hidden layer, and an output layer. The output from the input layer is connected as an input into the hidden layer. Similarly, the output from the hidden layer is connected as an input into the output layer to produce the final output of the ANN (Hamdan & Garibaldi). MLP model can be illustrated in Figure 1: **MLP Model.**
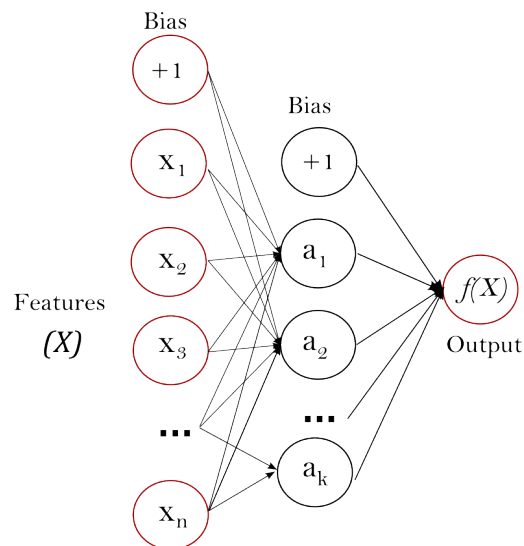

Figure 1: MLP Model (1.17. Neural network models (supervised))

4 Data Preprocessing

In this study, we first applied the Data Preprocessing techniques to convert the raw data from the Blood transfusion dataset into a clean dataset. In Data Preprocessing, we used the following approaches: data cleaning, data integration, transformation, data reduction, and data discretization.

4.1 Data Reduction

Data reduction is reduction in volume with production of the same or similar analytical results (Data Preprocessing Techniques for Data Mining). The following columns were removed from the data-set because they were determined to be unnecessary to the analysis: Surgery Start Date and Time, Stop Date and Time, Blood BM, Cell Salvage and ANH have been removed. Cases with missing pre-operative anemic status or surgical procedure were also removed from the dataset.

4.2 Missing Data Imputation

Gelman at el. (Gelman & Hill, 2006) in *Data Analysis Using Regression and Mutilevel/Hierachical Models* described ways to impute missing data values. These included analysis of randomness of missing values. The methods of imputing missing data values described included simply discarding the data itself, using mean imputation, using the last value carried over, using information from related observations, imputation based on logical rules. Our study used mean imputation, data discard and imputation based on logical rules to impute missing data and to clean the data in order to train the neural network model.

4.3 Outlier Detection and Removal

Permissive clinically accepted ranges for blood transfusion parameters as well as anemic status values were used to detect outliers. Outliers were removed from the data frame.

	M	Ma	U	Clinical valid
Results before surgery	0.	140	g	3 to 20
Results after surgery	0.	19.1	g	3 to 20
Red Blood Cells	0.	72.0	U	0 to 75
Fresh Frozen Plasma	0.	820.	U	0 to 75
Platelets	0.	490.	U	0 to 75
Cell Salvage Amt	0.	712	c	0 to 25000
Cell Salvage Amt	0.	900	c	0 to 10000
ANH Amt Removed	0.	700	c	0 to 1500

Table 1: Statistics of features and clinical valid ranges

4.4 Data Transformation

"KDnuggets." (McGinnis) explores different methodologies to transform categorical variables into numeric values by using encoding. In the data, the columns Surgery Specialty, Surgery Procedure,

and Patient Type have categorical characters. On data transformation for categorical values, it is suggested that ordinal method is applied if there are more than 500 classes. Therefore, ordinal method can be used to change types from 0 to Z values, with Z being the number of unique classes in a given field. In order to achieve a more accurate prediction and a faster convergence rate for building a better model, data scaling is the method needed to apply to the data set before it is used to train the Neural Network model.

The ranges for post-operative hemoglobin were used to transform values into categorical results that were used as Y classification results. The ranges for post-operative hemoglobin and their classification were as follows: Value < 8 is severely anemic, value from 8 to 10 is anemic, and value > 10 is not anemic. Another approach was taken in the experiment. Data was split into two classes that represent anemic and not anemic instead of splitting into three categorical classes. The ranges for post-operative hemoglobin and their classification became: Value < 10 is anemic, and value > 10 is not anemic. The values for boundaries of anemia level are indicated by Keck Medical Center.

4.5 Data Sampling

Data sampling method was the last step in data preprocessing progress. Due to the different number of anemic, severely anemic and not anemic patients, data sampling is used to ensure balance between the patients. By implementing data sampling for larger amount classes, this will increase the precision, recall and F-1 score for classes that are smaller volume.

5 Multilayer Perceptron Neural Network Model

Due to the high complexity of the dataset as well as the need for the ability to generalize any unexpected inputs/patterns, Neural Network has been specifically studied (Siganos). Han et al. (Han & May, 1997) used Neural Network and Genetic Algorithm collaboratively in order to model process parameters to achieve better quality.

According to the study, MLP Neural Network is constructed by using MLP for Classification. There are a number of hidden layers, neurons, parameters, input X and output Y used to construct the model. The X value is the whole data set without After Surgery Anemic Results and the Y value is the After Surgery Anemic Results in the data. Neural Network with one hidden layer can be demonstrated in the diagram below. The features will be X and output will be Y or f (X) in this case. The data can be separated into rows in which the Results After Surgery is a missing value or is not. The rows with a missing value are predicted and the rows with a value present are used to train/test the Neural Network model.

During the training and tuning of the predictive model of Multi-Layer Perceptron Neural Network, we have explored combinations for the activation functions as well as number of hidden layers. The different combinations consist of one hidden layer to six hidden layers, and as the number of hidden layers' increases, the accuracy of the model decreases and the time complexity increases. One hidden layer in the model was not sufficient to achieve better accuracy; therefore, two hidden layers were used for modeling. In this research, the number of hidden layer neurons was determined by trial and error. Different activation functions have also been tested, which include rectifier linear unit (ReLU), logistic and hyperbolic tangent. The ReLU function achieved the best result among the different activation functions. The optimization technique we implemented was stochastic gradient descent. This optimization was used due to its good result and time performance for the given data set.

6 Evaluation and Result

For the evaluation's purpose, we used the Cross-validation method (Stone, 1974) which is a technique to evaluate predictive models by partitioning the original sample into subsamples of training set to train the model, and a subsample of test set to evaluate it. In this paper, we used 10-fold Cross validation methods to assess overfitting and ensure accurate modeling. Using 10 folds for cross validation, the model obtained the mean value achieved performed with 67% accuracy and a 3% standard deviation. Precision and recall performance of result is shown in the table below (Table 2).

	Precisio	Recall	F1
Class 1 – Not	0.75	0.83	0.79
Class 2 – Anemic	0.54	0.58	0.56
Class 3 – Severely Anemic	0.48	0.20	0.28

Table 2: Performance scores using proposed method (three class classification)

From observing the result in the approach above, the scores for Class 3 – Severely Anemic were not acceptable. The results were unsatisfying because the class 3 data is only 5% of the whole data set. The model did not have enough knowledge to predict for class 3. Therefore, another approach was used. This approach combines class 2 and class 3 data points together to become only one class. This method can ensure a more balanced dataset, and two class classification explains the data set better. The results improved after changing the three classes classification problem into two classes with only anemic and not anemic classes for the data. The precision and recall performance of our result is shown in the table below. The performance of this model using 10 folds for cross validation and the mean value achieved performed with 77% accuracy and a 2% standard deviation. This model gets 10% improvement compared with 67% accuracy for triple classes model. This model also gets improvement in precision and recall. Thus, there was not enough training data in the Severely Anemic record to train the ANN model for it to predict with good results.

	Precis	Recall	F1
Class 1 – Not	0.74	0.79	0.72
Class 2 – Anemic	0.70	0.83	0.81

Table 3: Performance scores using proposed method (binary class classification)

There are trade-offs between precision and recall scores for both classes. Due to the medical application of the proposed model, it is important to address the recall score of Class 2 Anemic prediction as the most important factor for evaluating performance. It is critical to identify and predict the number true positive (anemic) patients \and minimize false negatives since it is usually safer to assume patients are anemic than they are not.

The baseline result of the proposed model is by using instance-based learning algorithm K Nearest Neighbors (KNN) with K=1. The performance is shown in the table below. By comparing the results for recall score of class 2 anemic prediction, there is a significant rate of improvement using the proposed methodology.

	Precisi	Recall	F1
Class 1 – Not	0.81	0.81	0.81
Class 2 – Anemic	0.54	0.54	0.54

Table 4: Performance scores using baseline KNN method (binary class classification)

7 Conclusion

In this paper, we studied the data collected from the electronic medical record of the Keck Medical Center of USC related to blood transfusion and blood management during surgery and anemia perioperatively. Our goal was to predict the degree of anemia after surgery based on pre-operative laboratory values, intra-operative blood management, and characteristics of the surgery. We applied several steps of data preprocessing techniques in order to prepare the raw data. We used Neural Network to predict values in the After Surgery Anemic Status. By using N-Fold cross validation method and two-class classification, the accuracy that is obtained is 77%. The result for predicting 3 classes categorical values was not good, due to the small portion of class 3 in the data set (class 3 being only 5% in the dataset). Therefore, the resulting ANN model did not predict severe anemia after surgery. Therefore, a two-class classification was chosen. The baseline is the performance result of instance-based model KNN with k = 1. By comparing the recall score in Class 2 of ANN and baseline method KNN, it shows significant improvement. Moving forward, we hope to study the learning capability and storage capacity of two hidden layer feedforward networks and to determine the amount of neurons required to improve Multilayer Perceptron Neural Network Model performance to get a higher accuracy and precision and recall score (Huang, 2003). With better accuracy and precision recall, the model can be implemented to reduce unnecessary blood transfusion in the clinical world.

8 Acknowledgement

References

1.17. Neural network models (supervised). (n.d.) website http://scikit-learn.org/stable/modules/neural_networks_supervised.html

Collobert R. and Bengio S. (2004). *Links between Perceptrons, MLPs and SVMs.* Proc. Int'l Conf. on Machine Learning (ICML).

Data Preprocessing Techniques for Data Mining. (n.d.). Retrieved December 13, 2016, from ICAR-Indian Agricultural Statistics Research Institute: http://iasri.res.in/ebook/win_school_aa/notes/Data_Preprocessing.pdf

Gelman, A., & Hill, J. (2006, July 6). *Data Analysis Using Regression and Multilevel/Hierarchical Models.* Retrieved 2017, from Department of Statistics Columbia Univ. in the City of New York: http://www.stat.columbia.edu/~gelman/arm/missing.pdf

Hamdan H. and Garibaldi J. (2009), *Modelling Survival Prediction in Medical Data*, the 9th Annual Workshop on Computational Intelligence (UKCI2009), Nottingham, UK.

Han, S.-S., & May, G. S. (1997). Using neural network process models to perform PECVD silicon dioxide recipe synthesis via genetic algorithms. *IEEE Transactions on Semiconductor Manufacturing, 10* (2), 279-287.

Huang, G.-B. (2003). Learning capability and storage capacity of two-hidden-layer feedforward networks. *IEEE Transactions Neural Network, 14* (2), 274-281.

McGinnis, W. (n.d.). *KDnuggets Analytics Big Data Data Mining and Data Science*. Retrieved December 14, 2016, from KDnuggets Web site: http://www.kdnuggets.com/

Sarode, R. M. (n.d.). *Complications of Transfusion*. (The University of Texas Southwestern Medical Center at Dallas) Retrieved 2017, from Merck Manuals Professional Edition Web site: http://www.merckmanuals.com/professional/hematology-and-oncology/transfusion- medicine/complications-of-transfusion

Siganos, D. (n.d.). *Why Neural Networks?* (Dep. of Computing Imperial College London) Retrieved December 13, 2016, from Dep. of Computing Imperial College London Web site: https://www.doc.ic.ac.uk/~nd/surprise_96/journal/vol1/ds12/article1.html

Stone M. *Cross-validatory choice and assessment of statistical predictions.* J. Royal Stat. Soc., 36(2), 111–147, 1974.

Rosenblatt, F. (1957). *The perceptron — a perceiving and recognizing automaton* (Technical Report 85-460-1). Cornell Aeronautical Laboratory.

Vapnik, V. (1995). *The Nature of Statistical Learning Theory*. Springer. Second edition.