Applied Computing and Informatics xxx (2017) xxx–xxx

Contents lists available at ScienceDirect

Applied Computing and Informatics

journal homepage: www.sciencedirect.com

Original Article

Clinical decision support system for venous thromboembolism risk classification

Zelal Qatawneh [a], Mohammad Alshraideh [a], Nada Almasri [b,*], Luay Tahat [b], Abdullah Awidi [c]

[a] Computer Science Department, University of Jordan, Amman, Jordan
[b] MIS Department, Gulf University for Science and Technology, West Mishref, Kuwait
[c] Faculty of Medicine, University of Jordan, Amman, Jordan

ARTICLE INFO

Article history:
Received 10 March 2017
Revised 7 September 2017
Accepted 7 September 2017
Available online xxxx

Keywords:
Clinical decision support system
Venous thromboembolism
Artificial neural network
Multilayer perceptron
Resilient backpropagation

ABSTRACT

This paper presents a clinical decision support system using Artificial Neural Networks (ANN). The system uses Multilayer Perceptron (MLP) feed forward neural network to predict the risk of developing Venous Thromboembolism (VTE) in hospitalized patients. The developed system classifies the risk of VTE into five risk levels ranging from low to high. The input layer of the system consists of 35 input variables grouped into six categories representing the risk factors of VTE according to Caprini model. The output layer consists of one node indicating a value representing the level of VTE risk. The number of hidden nodes and layers is determined through an iterative process. The system is trained using Resilient Backpropagation algorithm (Rprop). The dataset used for training and testing the system consists of 150 medical records obtained from Jordan University Hospital (JUH). Stratified ten-fold cross validation scheme is applied to assess the generalization of the proposed system. The results of the experiment show that the accuracy of the system is 81%.

1. Introduction

Venous thromboembolism (VTE) is a serious and potentially fatal disorder, and it often complicates the course of hospitalized medical patients. Although it can be fatal, this disorder can be prevented by screening patients based on the potential risks of developing this condition. Many risk factors for VTE have been well recognized [1], and some basic and clinically relevant risk assessment models are available to facilitate VTE risk assessment in hospitalized medical patients [2,3]. However, as reported in our previous research [4], many researchers have found that a large number of high risk VTE patients go undetected and consequently do not get the proper treatment [5–7].

In this paper we present a clinical decision support system to predict the risk of developing VTE for hospitalized patients. Based on the patient's medical record, the system can classify the risk level of developing VTE into five levels ranging from low to high. The proposed system uses Artificial Neural Networks which are perfect fit for Medicine, the science of uncertainty. Neural Networks are especially effective when the nature of the problem is fuzzy and requires complex decision making process. Indeed, VTE is multifactorial with more than 35 different identified cumulative factors. This makes the decision making process for identifying patient's risk levels of developing VTE very complex [35].

The proposed system uses Multilayer Perceptron (MLP) feed forward neural network using Resilient Backpropagation algorithm (Rprop) to train the system. The system receives as inputs all of the 35 risk factors identified by Caprini model [8]. The MLP network is trained using 150 medical records collected from Jordan University Hospital (JUH) with the consent of the patients.

To implement the system, an experimental study was performed to choose the best structure of the neural network used in the proposed decision support system. The network was then trained, tested, and validated. Stratified ten-fold cross validation scheme was applied to assess the generalization of the proposed system. The results of the experiment showed that the accuracy of the system was 81%.

The remaining of this paper is organized as follows: Section 2 introduces the background and related work. Section 3 presents the methodology followed to design the system and it introduces

* Corresponding author.
E-mail address: almasri.n@gust.edu.kw (N. Almasri).

Peer review under responsibility of King Saud University.

Production and hosting by Elsevier

http://dx.doi.org/10.1016/j.aci.2017.09.003
2210-8327/

the specification and architecture of the system. Section 4 presents the experiment study which was conducted to train different neural network designs in order to choose the best performing model. In addition, it presents and discusses the results obtained when applying ten-fold cross validation on the selected network design. Finally, Section 5 provides the conclusion and future work.

2. Background and related work

2.1. Venous thromboembolism (VTE)

Venous Thromboembolism (VTE) is a major cause of morbidity and mortality and it is a serious health issue which often complicates the course of hospitalized medical patients. It results when a thrombus forms in a vessel and obstructs the blood flow. Many risk factors for VTE have been well recognized [1] and a review of 1231 consecutive patients treated for VTE, showed that 96% of the patients had at least one recognized risk factor [9,10].

There is convincing evidence that VTE risk increases in proportion to the number of predisposing factors [1,35]. These risk factors are generally cumulative [11] and may include old age, cancer, surgery, prolonged immobilization, hip and knee surgery, multitrauma, puerperium, spinal cord injuries, paralysis, heart failure, chronic lung conditions, acute myocardial infarction, sepsis, obesity, use of oral contraceptives, the antiphospholipid antibody syndrome, and inherited thrombophilic conditions [1,8,12].

A previous study performed by our research team [4] addresses risk factors for VTE assessment and the use of heparin in Jordan University Hospital (JUH) in the Middle East as a prevention/treatment measure. It concludes with the importance of implementing strategies such as educational sessions and standardized methods for identifying high-risk patients to ensure appropriate treatment.

2.2. Feedforward multilayer-perceptron (MLP) neural network

Artificial Neural Network (ANN) is a network of processing units designed in a way that imitates the biological neural networks of the human brain. They represent a powerful approach to process tasks that involve approximation or classification based on a large set of data, and they perform very well on these tasks [13]. Multi-Layer Perceptron (MLP) is one of the most popular networks used for predications in different scientific research [31–33] as well as in medical applications [14,29]. In this network, the processing units (perceptron; or commonly known as neurons) are arranged into multiple layers. The input layer contains the processing units that receive the input to the network. The output layer consists of the processing units that trigger the output of the network, e.g. the appropriate classification of the case entered as input. Between the input and output layers, there can be any number of intermediate layers; commonly called hidden layers. In MLP, data flows from input layer to the hidden layers and finally to the output layer in one direction only, and this is why this type of neural networks is called feedforward. The output of each neuron i in the MLP is a function of the sum of the weighted inputs in addition to the bias value, θ, as demonstrated in the following formula:

$$Y_i = f\left(\sum_{j=1}^{n} W_{ij} \cdot X_j + \theta_i\right) \quad (1)$$

where Y_i is the output of the ith neuron in the network, X_1 to X_n are the input values, and W_{ij} is the weight of the j's input at the i's neuron. This function f, through which the combined sum of the input is passed to generate the output, is called the activation function.

Activation functions are mathematical functions which convert the combined input into the expected output. Researchers have used various mathematical functions, and the most common functions are: Linear, Logistic and Tangent. In this paper, we use the Hyperbolic Tangent function; TANH.

Backpropagation is the common training algorithm used in MLPs. The concept of the algorithm it to calculate the error in the output (the difference between the output result obtained by the network and the desired output), and pass it backwards to the network so that the weights of the inputs are updated in order to minimize the error. Resilient backpropagation, Rprop, is an enhanced algorithm overcoming some of the major issues with the standard Backpropagation algorithm [15].

Finally, to estimate the performance of a prediction model, e.g. MLP, in terms of its accuracy, and in order to improve its generalization, cross-validation is commonly applied. A common practice is to divide the data set into three sets: training set, validation set and testing set [16]. The training dataset is used to train the network by adjusting the input weights for neurons in the neural network. The validation dataset is used to minimize overfitting, ensuring that any increase in accuracy over the training dataset actually yields an increase in accuracy over a dataset that the network has not seen previously. Finally, the testing dataset is used when the model design is iterated many times. In this case, after selecting the model having the best performance on the validation set, the training set is used to assess the generalization of the selected model on unseen data.

2.3. Application of clinical decision support systems in medical domain

Artificial intelligence (AI) in medicine has been an active research domain since the early 1970s. Different AI methods are used as the core concepts of clinical applications; this includes but is not limited to: diagnosis, treatment, and prediction of clinical outcome [27–30,34,36]. Many researchers proposed clinical decision support system (CDSS) to improve the ability of the physicians and the medical staff to diagnose diseases accurately. Kawamoto et al. [17] studied seventy CDSS and found that these systems significantly improved clinical practice in 68% of the systems. A large number of CDSS use artificial neural networks (ANN) as their main processing concept [18].

Hawamdeha et al. [19] suggested the use of a multilayer Perceptron (MLP) feed forward neural network to predict the rehabilitation protocol for patients with knee osteoarthritis. The result of their study showed that 87% of the system's output were accurately predicted.

Ho et al. [20] compared between three prediction models to predict disease-free survival in hepatocellular carcinoma (HCC) patients who have received hepatic resection. The three prediction models considered were artificial neural networks (ANN), logistic regression (LR), and decision tree (DT). They concluded that ANN model gave the best prediction accuracy.

Zecchin et al. [21] found that ANN improved the accuracy of short time prediction of glucose concentration. They proposed an approach for short time glucose prediction using past CGM sensor readings and information on carbohydrate intake. The approach combined a neural network (NN) model and first-order polynomial extrapolation algorithm, used in parallel to describe, respectively, the nonlinear and the linear components of glucose dynamics.

Yan et al. [14] used MLP for heart disease diagnosis. Based on 40 variables, the patient was diagnosed with one of five heart diseases. The system's accuracy was found to be more than 90%.

Durieux et al. [22] studied the effect of a knowledge based CDSS on physicians' behavior when treating patients with a potential risk of VTE. The system collects patient's information as well as the physician's prescription. The system can then show a warning message if it finds discrepancies between the current case and past cases in the knowledge base.

Please cite this article in press as: Z. Qatawneh et al., Applied Computing and Informatics (2017), http://dx.doi.org/10.1016/j.aci.2017.09.003

Kucher et al. [23] as well as Piazza et al. [24] reported that the use of a computerized alert system identifying consecutive hospitalized patients at risk of VTE has reduced the rates of VTE among hospitalized patients. The system considers only eight common VTE risk factors. Determining the risk of VTE is encoded in the system based on the expertise of the physicians who developed the study. The system thus does not take into consideration the past cases of VTE, instead it looks at each patient's medical record in isolation, and it generates an electronic alert to the treating physician in case the patient is identified with high risk of VTE.

Roy et al. [25] assessed the effectiveness of the use of handheld clinical decision support system to enhance the diagnostic work-up for emergency patients suspected with pulmonary embolism (PE). The software asks the physician to enter his/her own estimate of the patient's PE risk, then it suggests appropriate and inappropriate diagnostic tests, and it marks the least invasive test as "recommended". The recommendations of the system are based on the expertise of the physicians who participated in the design of the study.

3. Data encoding and system architecture

In order to design an affective medical decision support system assessing the patient's risk level of VTE, all risk factors for VTE as identified by domain experts [2,3,8] were considered. In our previous work [4], VTE risk factors were investigated, and this work is based on the findings of our previous work. We hence use Caprini model to assess risk factors of VTE.

Our proposed clinical decision support system is designed with two main components: Data entry and encoding, and risk assessment engine. The data entry and encoding component allows entering patient's data and encoding them according to the encoding scheme explained in Section 3.1. For initial system training, Patients' personal and medical data were collected from Jordan University Hospital (JUH). Experts from the domain assessed the cases according to the Caprini model [4,8], and they classified patients into five risk levels: low, lower-mild, higher-mild, moderate, and high; where the probabilities of developing VTE in these risk levels are 10%, 20%, 30%, 40%, and 50% respectively. The collected dataset was used for training, validating, and testing the neural network used in the risk assessment engine. The risk assessment engine is the main processing part of the system. It receives the patient's encoded data, and then it determines VTE risk level for the patient. The risk assessment engine uses a neural network which was trained with 150 patient's records. The optimal structure of the neural network (presented in Section 3.2) was determined experimentally as explained in Section 4.

In the next subsections we discuss the process of data collection for the dataset used to train and test the system along with the encoding scheme used to represent patient's data, and then we introduce the MLP-based architecture of the risk assessment engine used in the proposed decision support system.

3.1. Data encoding scheme

Patients' data were collected from JUH, which is a busy tertiary care medical center. The data set used to train and test the system was based on a sample of 150 complete medical records. The sample was selected so that it contains 30 records from each risk level. Only patients who are above the age of 40 were considered for this study.

The sample data set was collected from medical records for both walk-in patients and admitted patients. An informed consent form was obtained from all study subjects. The study was approved by the hospital IRB as part of quality monitoring focused study to generate hospital guidelines concerning VTE prophylaxis.

From each medical record, the values of thirty-five risk factors were extracted. These risk factors were categorized into six categories as follows: Age, Obesity, Surgery, Disease, Women related, Genetic & Drugs, where:

- Age: has only one risk factor. Its value considers 3 intervals of age for patients over 40. Each interval gets a different risk score.
- Obesity: has only one risk factor. Its values considers 3 BMI intervals with different score of risk.
- Surgery: has 7 risk factors in total with different risk scores.
- Disease: has 17 risk factors in total with different risk scores.
- For Women: has 3 risk factors in total with different risk scores.
- Genetics &Drugs: has 6 risk factors in total with different risk scores.

As demonstrated in Table 1, each risk factor is encoded as a risk score ranging from 1 to 5.

Table 1
VTE risk factors.

Category	Factor	Input score
Age	>75	3
	60–74	2
	41–70	1
Obesity	BMI > 50	3
	BMI > 40	2
	BMI > 0	1
Surgery	Major surgery (over 3 h)	5
	Major surgery (2–3 h)	3
	Laparoscopic surgery (>60 min)	2
	Major surgery (>60 min)	2
	Arthroscopic surgery (>60 min)	2
	History of prior major surgery	1
	Minor surgery planned	1
Disease	Multiple trauma (<1 month)	5
	Acute spinal cord injury (paralysis) (<1 month)	5
	Stroke (<1 month)	5
	Elective major lower extremity arthroplasty	5
	History of DVT/PE	3
	Central venous access	2
	Malignancy (present or previous)	2
	Abnormal pulmonary function (COPD)	1
	Medical patient currently at bed rest	1
	Leg plaster cast or brace	1
	Varicose veins	1
	History of inflammatory bowel disease	1
	Swollen leg (current)	1
	Acute myocardial infarction (<1 month)	1
	Congestive heart failure (<1 month)	1
	Series lung disease incl pneumonia (<1 month)	1
	Sepsis (<1 month)	1
Women related	Oral contraceptives or hormone replacement therapy	1
	Pregnancy or postpartum (<1 month)	1
	• History of unexplained stillborn infant, or • Recurrent spontaneous abortion (>3), or • Premature birth with toxemia or growth-restricted infant	1
Genetics & drugs	Positive Factor V Leiden	3
	Elevated serum homocysteine	3
	Heparin-induced thrombocytopenia (HIT)	3
	Positive prothrombin 20210A	3
	Positive Lupus anticoagulant	3
	Elevated anticardiolipin antibodies	3

Please cite this article in press as: Z. Qatawneh et al., Applied Computing and Informatics (2017), http://dx.doi.org/10.1016/j.aci.2017.09.003

3.2. System architecture

The risk assessment engine uses a feedforward MLP neural network using resilient backpropagation algorithm (Rprop).

The structure of the network; and particularly the number of hidden layers and the number of neurons in each hidden layer; is chosen based on an experimental iterative process to determine the best structure with the highest accuracy as demonstrated in Section 5. As demonstrated in Fig. 1, the adopted network structure consists of:

- 35-input neurons in the input layer.
- Three hidden layers having 19 neurons in the first layer, 10 in the second layer, and 5 in the third.
- One output neuron in the output layer.

TANH activation function was used for all hidden layers as well as the output layer. The values 0.08, 0.0000001, and 6 were used for the initial delta, the performance goal error, and the number of validation checks to avoid overfitting of the network respectively.

The details of the network architecture are further explained in the following subsections.

3.2.1. Input layer

The input layer is designed with 35 neurons for the 35 risk factors identified in Section 3.1. Each risk factor is encoded as a risk score ranging from 1 to 5 as demonstrated in Table 1.

3.2.2. Hidden layers

Since the process of determining the best number of hidden layers and neurons in each layer is very complex as it depends on: the number of input and output neurons, the number of training cases, and the complexity of the classification problem to be learned [31]; we applied an experimental iterative approach to determine the best number of hidden layers and neurons within each layer. Based on the experimental study, it was found that in order to obtain the best accuracy, the system should consist of three hidden layers, with 19, 10, and 5 neurons respectively. Consequently, we adopted this architecture for the system. The experimental methodology that was followed to find the best performing network is discussed in Section 4.

3.2.3. Output layer

The output layer consists of one neuron. It displays the output variable, with a value normalized into five possible levels of risk of VTE: low, lower-mild, higher-mild, moderate, and high.

4. Experimental study and system evaluation

An experimental study was conducted to choose the best architecture for the system. The network with the best classification accuracy was then validated with ten-fold cross validation method.

4.1. Model building

In the experimental study, we train several network models and then we compare them based on their classification accuracy. The cross validation method was used to estimate the accuracy which is determined by the overall number of correct classifications divided by the total number of instances in the dataset. A percentage of 80%, 10%, and 10% data split of the complete data samples was applied to represent the training, validation, and test subsets respectively. Consequently, we obtained a total of 120, 15, and 15, sample cases for the training, validation and testing subsets respectively.

The whole process works as follows:

(1) Set initial number of hidden layers NHL to 1
(2) Calculate the initial number of neurons (N_f) in the hidden layer according to Eq. (2)

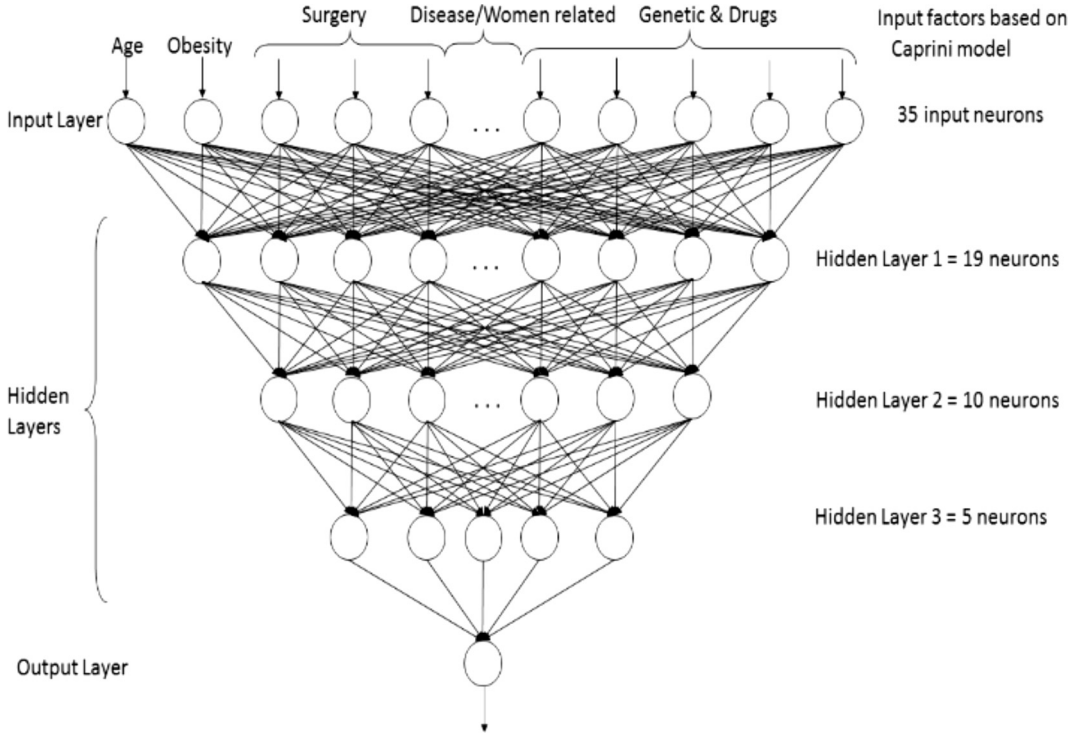

Fig. 1. MLP structure used for the risk assessment engine.

$$N_f = \left| \frac{N_{prev} + N_o}{2} \right|. \quad (2)$$

(a) Train and calculate the accuracy of the network with the hidden layer having N_f neurons
(b) Train and calculate the accuracy of the network with the hidden layer having $N_f - 1$ neurons
(c) Train and calculate the accuracy of the network with the hidden layer having $N_f + 1$ neurons
(3) Add a new hidden layer and repeat step 2 until $N_f = 2$

Where N_f is the number of neurons in the hidden layer, N_{prev} is the number of neurons in the previous layer (when considering the first hidden layer, N_{prev} is the number of neurons in the input layer), and N_o is the number of the neurons in the output layer.

All network models trained in this experiment share the following characteristics:

- A Feedforward Backpropagation neural network was used.
- The number of neurons in the input layer was 35.
- The number of neurons in the output layer was 1.
- The training algorithm that was used for training was the Rprop.
- Each model was tested with TANH activation function.
- The following values were used 0.08, 0.0000001, and 6 for the initial delta, the performance goal error and the number of validation checks to avoid the overfitting of the network respectively.

Table 2 lists the best performing network architectures based on their classification accuracy as obtained from the experimental study.

4.2. Stratified ten-fold cross validation

To ensure higher robustness of the evaluation, stratified 10-fold cross validation was applied as suggested by [26]. Consequently, the data set was stratified into 10 folds, where each fold consisted of 15 sample cases having 3 low risk cases, 3 lower-mild risk cases, 3 higher-mild risk cases, 3 medium risk cases, and 3 high risk cases.

The validation process went through 10 different runs of the model, where each run used 9 folds for training and one fold for testing. Each run used a testing fold that was not picked in any of the previous runs. This guarantees that each fold would be considered as a testing fold in one of the runs.

In order to find the classification accuracy for the model, we calculated the classification accuracy of the test fold in each run. The overall accuracy of the model was calculated as the average accuracy of all 10 runs. The overall accuracy was calculated as follows:

- First, the classification accuracy for each fold is computed. The classification accuracy Af_i of an individual fold f_i depends on the number of samples correctly classified and is evaluated by the formula:

$$A_{f_i} = \frac{t}{n} * 100\% \quad (3)$$

Table 2
Best performing network architecture.

Number of hidden layers	Number of neurons within each layer	Best classification accuracy %
1	35–17–1	93.3
2	35–17–9–1	98
3	35–19–10–5–1	100
4	35–17–9–3–2–1	80
5	35–19–9–5–3–2–1	86.66

where t is the number of sample cases correctly classified, and n is the total number of sample cases.

- Then, the overall model classification accuracy is calculated as follows:

$$Average\ Classification\ Accuracy = \frac{1}{10}\sum_{i=1}^{10} A_{f_i} \quad (4)$$

To summarize the results of the 10 test folds, Table 3 illustrates the classification accuracy for all runs. The average accuracy classification obtained for all folds is 81%. We notice that the highest classification accuracy was obtained in run 5 and run 10 with an accuracy of 100% and the lowest accuracy was obtained in run 7 with an accuracy of 33.34%. Looking further at the testing fold of run 7, we noticed that most sample cases were border cases between two risk classes, which led to their misclassification into either the class which is directly below the actual class or the class which is directly above it. Looking at each class of risk, the overall classification accuracy is 90%, 80%, 70%, 76.7%, and 86.7% for low risk, lower-mild risk, higher-mild risk, moderate risk, and high risk classes respectively.

To further analyze the performance of the classification model, we provide the confusion matrix in Table 4.

The confusion matrix helps visualizing the performance of the model within each class of VTE risk by looking at the actual versus the predicted class of risk for each class. The rows represent the predicted classes for the sample cases from a specific class of risk. For example, the first row shows that out of 30 actual low risk cases, the model correctly predicted 27 cases within the low risk class, while it incorrectly classified the remaining three low risk cases as lower-mild risk class. The columns show the actual classification of the cases that were predicted within a specific class of risk. For example, the first column show that the model predicted a total of 31 low risk cases, while actually only 27 out of them were actually low risk cases, and the other 4 were incorrectly classified as low risk while they were actually lower-mild risk cases. Consequently, the rate 27/30 shows the measure of the procedure accuracy of the classification for the low risk class (which is also called the Recall measure), while the rate 27/31 shows the user accuracy of the classification model for low risk class (which is also called the Precision measure).

Table 5 summarizes procedure accuracy and user accuracy for each VTE risk class.

4.3. Observations and discussion

The results presented in the previous section show that the system provide very good accuracy values for both low and high risk levels of VTE, while it provides modest accuracy values for the classes in between. Since in a medical setting, what is important is to safely discharge patients with low risk of VTE while directly treating patients with high risk of VTE, we can clearly see that the system achieves this objective.

For high risk VTE, the user accuracy shows that all cases (100%) that are predicated as high risk of VTE are actually high risk of VTE. The procedure accuracy, on the hand, shows that the probability that the system will accurately classify a high risk case into its proper class is 90%. Based on the confusion matrix, we know the percentage of misclassified high risk cases (10%) are classified by the system as moderate risk; which is the adjacent class directly below the high risk class.

In general, according to the confusion matrix, we can clearly see that when a case is misclassified, it is incorrectly classified by only one adjacent risk level (directly below or directly above) the actual class. This is an indication that the range of classification errors is ±1 adjacent class.

Table 3
Classification accuracy for 10 folds.

Run set	Folds	Low risk 10%	Mild Risk		Moderate risk 40%	High risk 50%	Average accuracy For each fold
			Lower-mild20%	Higher-mild30%			
1	F1	100%	66.7%	66.7%	66.7%	100%	80.02%
2	F2	100%	33.3%	0%	66.7%	66.7%	53.34%
3	F3	100%	100%	100%	100%	66.7%	93.34%
4	F4	100%	100%	100%	66.7%	100%	93.34%
5	F5	100%	100%	100%	100%	100%	100%
6	F6	100%	66.7%	100%	33.3%	100%	80%
7	F7	0%	66.7%	0%	33.3%	66.7%	33.34%
8	F8	100%	66.7%	33.3%	100%	100%	80%
9	F9	100%	66.7%	100%	100%	100%	93.34%
10	F10	100%	100%	100%	100%	100%	100%
		Average folder accuracy					(80.67)%
Average accuracy for each level of risk		90%	76.7%	70%	76.7%	90%	≈(81)%

Table 4
Confusion matrix.

Original level		Predicted low risk 10%	Predicted mild risk		Predicted moderate risk 40%	Predicted high risk 50%	Sum of actual cases
			Lower-mild 20%	Higher-mild 30%			
Actual low risk 10%		27	3	0	0	0	30
Actual mild risk	20%	4	23	3	0	0	30
	30%	0	5	21	4	0	30
Actual moderate risk 40%		0	0	7	23	0	30
Actual high risk 50%		0	0	0	3	27	30
Sum of predicted cases		31	31	31	30	27	150

Table 5
Procedure and user accuracy.

Risk level of VTE		Procedure accuracy % (recall)	User accuracy % (precision)
Low risk 10%		27/30 = 90%	27/31 = 87.1%
Mild risk	Lower-mild 20%	23/30 = 76.7%	23/31 = 74.2%
	Higher-mild 30%	21/30 = 70%	21/31 = 67.8%
Moderate risk 40%		23/30 = 76.7%	23/30 = 76.7%
High risk 50%		27/90 = 90%	27/27 = 100%
Average accuracy		80.7%	81.2%

Consequently, we recommend the use of this system for screening low/high risk levels of VTE. Additionally, we recommend that the physician manually checks cases identified as moderate risks since there is a 10% probability that it can actually be a high risk case.

5. Conclusion and future work

Venous thromboembolism (VTE) is a very common health issue with complications that carry a high morbidity and mortality rates. Although VTE is a common disease, it is a preventable cause of hospital death if it is quickly identified by the treating physician.

In this paper we propose a clinical decision system to automate and accurately predict the risk of VTE. This work represents a preliminary step in developing a methodology to predict VTE and classify it into five levels of risk based on predisposing factors which are chosen from Caprini score of VTE model as used in Jordan University Hospital.

The proposed system uses Artificial Neural Networks (ANN) in evaluating multifactorial health issue. Developing the system passed through several steps starting from determining the factors, collecting and stratifying the data set, normalizing the data input, training and choosing the most accurate system model, and finally validating the system.

The system was developed using one of the most widespread machine learning techniques; MLP feed forward neural network. It was trained using the Rprop training algorithm, and it consisted of: an input layer with 35 neurons (representing the input variables for each patient such as, age, gender, etc.), 3 hidden layers (where the number of neurons in the first, second and third hidden layer were 19, 10 and 5 respectively) and an output layer (that produced the type of the disease the patient suffered from).

Stratified ten-fold cross validation was applied, and it showed that the system had an overall user accuracy of 81.2% and an overall procedure accuracy of 80.7%. It is worth noting that the system accuracy is very high for both extreme classes of risk: low risk and high risk. For the high risk class, the user accuracy obtained from the experiment is 100% and the procedure accuracy is 90%. For the low risk class, the user accuracy is 87.5% while the procedure accuracy is 90%. When misclassification occur, the system usually misclassifies a case in either the class directly below the actual classification class or the one directly above it.

Finally, in a future research we will consider comparing the performance of our system using different classifiers such as Naïve Bayes (NB), Support Vector Machine (SVM), and Decision Tree (DT). Additionally, we will consider classifying VTE into its two major types: Deep Venous Thrombosis (DVT) and Pulmonary Embolism (PE).

References

[1] F.A. Anderson, F.A. Spencer, Risk factors for venous thromboembolism, Circulation 107(23 Suppl. 1) (2003) 1–9.
[2] W.H. Geerts, J.A. Heit, G.P. Clagett, G.F. Pineo, C.W. Colwell, F.A. Anderson, H.B. Wheeler, Prevention of venous thromboembolism, CHEST J. 119(1_suppl) (2001) 132S–175S.

Please cite this article in press as: Z. Qatawneh et al., Applied Computing and Informatics (2017), http://dx.doi.org/10.1016/j.aci.2017.09.003

[3] A.T. Cohen, G. Agnelli, F.A. Anderson, J.I. Arcelus, D. Bergqvist, J.G. Brecht, et al., Venous thromboembolism (VTE) in Europe, Thromb. Haemost. 98 (4) (2007) 756–764.
[4] A. Awidi, N. Obeidat, A. Magablah, N. Bsoul, Risk stratification for venous thromboembolism in hospitalized patients in a developing country: a prospective study, J. Thromb. Thrombol. 28 (3) (2009) 309–313.
[5] A.T. Cohen, V.F. Tapson, J.F. Bergmann, S.Z. Goldhaber, A.K. Kakkar, B. Deslandes, et al., Venous thromboembolism risk and prophylaxis in the acute hospital care setting (ENDORSE study): a multinational cross-sectional study, The Lancet 371(9610) (2008) 387–394.
[6] A. Amin, S. Stemkowski, J. Lin, G. Yang, Thromboprophylaxis rates in US medical centers: success or failure?, J Thromb. Haemost. 5 (8) (2007) 1610–1616.
[7] S.R. Kahn, A. Panju, W. Geerts, G.F. Pineo, L. Desjardins, A.G. Turpie, et al., Multicenter evaluation of the use of venous thromboembolism prophylaxis in acutely ill medical patients in Canada, Thromb. Res. 119 (2) (2007) 145–155.
[8] J.A. Caprini, J.I. Arcelus, J. Reyna, Effective risk stratification of surgical and nonsurgical patients for venous thromboembolic disease, in: Seminars in Hematology, vol. 38, WB Saunders, 2001, April, pp. 12–19.
[9] F.A. Anderson, H.B. Wheeler, Physician practices in the management of venous thromboembolism: a community-wide survey, J. Vasc. Surg. 16 (5) (1992) 707–714.
[10] F.R. Rosendaal, Risk factors for venous thrombotic disease, Thromb. Haemost. 82 (2) (1999) 610–619.
[11] F.R. Rosendaal, Venous thrombosis: a multicausal disease, The Lancet 353 (9159) (1999) 1167–1173.
[12] P. Prandoni, Acquired risk factors for venous thromboembolism in medical patients, ASH Educ. Program Book 2005 (1) (2005) 458–461.
[13] K. Veropoulos, Machine Learning Approaches to Medical Decision Making, Doctoral Dissertation, University of Bristol, 2001.
[14] H. Yan, Y. Jiang, J. Zheng, C. Peng, Q. Li, A multilayer perceptron-based medical decision support system for heart disease diagnosis, Expert Syst. Appl. 30 (2) (2006) 272–281.
[15] M. Riedmiller, H. Braun, A direct adaptive method for faster backpropagation learning: the RPROP algorithm, in: Neural Networks, 1993, IEEE International Conference on, IEEE, 1993, pp. 586–591.
[16] Christopher M. Bishop, Pattern Recognition and Machine Learning, Springer, 2006.
[17] K. Kawamoto, C.A. Houlihan, E.A. Balas, D.F. Lobach, Improving clinical practice using clinical decision support systems: a systematic review of trials to identify features critical to success, BMJ 330 (7494) (2005) 765.
[18] K. Papik, B. Molnar, R. Schaefer, Z. Dombovari, Z. Tulassay, J. Feher, Application of neural networks in medicine – a review, Med. Sci. Monit. 4(3) (1998) MT538–MT546.
[19] Z.M. Hawamdeh, M.A. Alshraideh, J.M. Al-Ajlouni, I.K. Salah, M.B. Holm, A.H. Otom, Development of a decision support system to predict physicians' rehabilitation protocols for patients with knee osteoarthritis, Int. J. Rehabil. Res. 35 (3) (2012) 214–219.
[20] W.H. Ho, K.T. Lee, H.Y. Chen, T.W. Ho, H.C. Chiu, Disease-free survival after hepatic resection in hepatocellular carcinoma patients: a prediction approach using artificial neural network, PLoS One 7 (1) (2012) e29179.
[21] C. Zecchin, A. Facchinetti, G. Sparacino, G. De Nicolao, C. Cobelli, Neural network incorporating meal information improves accuracy of short-time prediction of glucose concentration, IEEE Trans. Biomed. Eng. 59 (6) (2012) 1550–1560.
[22] P. Durieux, R. Nizard, P. Ravaud, N. Mounier, E. Lepage, A clinical decision support system for prevention of venous thromboembolism: effect on physician behavior, JAMA 283 (21) (2000) 2816–2821, http://dx.doi.org/10.1001/jama.283.21.2816.
[23] N. Kucher, S. Koo, R. Quiroz, J.M. Cooper, M.D. Paterno, B. Soukonnikov, S.Z. Goldhaber, Electronic alerts to prevent venous thromboembolism among hospitalized patients, N. Engl. J. Med. 352 (10) (2005) 969–977.
[24] Gregory Piazza, Samuel Z. Goldhaber, Computerized decision support for the cardiovascular clinician applications for venous thromboembolism prevention and beyond, Circulation 120 (12) (2009) 1133–1137.
[25] P.M. Roy, P. Durieux, F. Gillaizeau, C. Legall, A. Armand-Perroux, L. Martino, et al., A computerized handheld decision-support system to improve pulmonary embolism diagnosis: a randomized trial, Ann. Intern. Med. 151 (10) (2009) 677–686.
[26] R. Kohavi, A study of cross-validation and bootstrap for accuracy estimation and model selection, in: Proceedings of the Fourteenth International Joint Conference on Artificial Intelligence, Morgan Kaufmann, San Francisco, CA, 1995, pp. 1137–1143.
[27] Vimla L. Patel, Edward H. Shortliffe, Mario Stefanelli, Peter Szolovits, Michael R. Berthold, Riccardo Bellazzi, Ameen Abu-Hanna, The coming of age of artificial intelligence in medicine, Artif. Intellig. Med. 46 (1) (2009) 5–17.
[28] A.N. Ramesh, C. Kambhampati, J.R.T. Monson, P.J. Drew, Artificial intelligence in medicine, Ann. R. College Surg. Engl. 86 (5) (2004) 334.
[29] Fujita U. Hamido, Rajendra Acharya, Vidya K. Sudarshan, Dhanjoo N. Ghista, S. Vinitha Sree, Lim Wei Jie Eugene, Joel E.W. Koh, Sudden cardiac death (SCD) prediction based on nonlinear heart rate variability features and SCD index, Appl. Soft Comput. 43 (2016) 510–519.
[30] K. Sudarshan Vidya et al., Computer-aided diagnosis of myocardial infarction using ultrasound images with DWT, GLCM and HOS methods: a comparative study, Comput. Biol. Med. 62 (2015) 86–93.
[31] Binh Thai Pham, Dieu Tien Bui, Indra Prakash, M.B. Dholakia, Hybrid integration of Multilayer Perceptron Neural Networks and machine learning ensembles for landslide susceptibility assessment at Himalayan area (India) using GIS, Catena 149 (2017) 52–63.
[32] Cyril Voyant, Gilles Notton, Christophe Darras, Alexis Fouilloy, Fabrice Motte, Uncertainties in global radiation time series forecasting using machine learning: the multilayer perceptron case, Energy 125 (2017) 248–257.
[33] Tien Dat Pham, Kunihiko Yoshino, Dieu Tien Bui, Biomass estimation of *Sonneratia caseolaris* (l.) Engler at a coastal area of Hai Phong city (Vietnam) using ALOS-2 PALSAR imagery and GIS-based multi-layer perceptron neural networks, GISci. Rem. Sens. 54 (3) (2017) 329–353.
[34] Kumar Ashish, Anish Dasari, Subhagata Chattopadhyay, Nirmal Baran Hui, Genetic-neuro-fuzzy system for grading depression, Appl. Comput. Inform., 2017.
[35] Lianne Parkin, Angela Balkwill, Siân Sweetland, Gillian K. Reeves, Jane Green, Valerie Beral, Million Women Study Collaborators, Antidepressants, depression, and venous thromboembolism risk: large prospective study of UK women, J. Am. Heart Assoc. 6(5) (2017) e005316.
[36] Mumini Olatunji Omisore, Oluwarotimi Williams Samuel, Edafe John Atajeromavwo, A Genetic-Neuro-Fuzzy inferential model for diagnosis of tuberculosis, Appl. Comput. Inform., 2015.

Zelal Qatawneh obtained her masters of science in Computer Science from the University of Jordan, Amman in 2015. She received her B.Sc. degree Computer Information Systems from Mutah University, Jordan in 2011.

Mohammad Alshraideh is a Professor of Computer Science at the University of Jordan, Jordan. He received his B.Sc. degree in Computer Science in 1988 from Mu'tah University in Jordan and a Master degree in Computer Science in 2000 from University of Jordan. He obtained his Ph.D. degree in Computer Science from University of Hull, UK, in 2007. During his graduate studies he obtained a fellowship from the University of Jordan. He was a Head Director Assistant for Computer Technology at the Hospital of the University of Jordan until June 2012. Also he was working as Human Resource Director at the University of Jordan until 2015. Prof. Alshraideh is currently working as Registrar General at the University of Jordan. His research interests include Software Testing, Artificial Intelligence, and Data Mining.

Nada Almasri is an Assistant Professor of Management Information Systems at Gulf University for Science and Technology, Kuwait, and she has held this position since 2009. Prior to that, she was a full-time lecturer at the University of Waterloo, David R. Cheriton School of Computer Science, Waterloo, Ontario, Canada. She received the M.Sc. and the Ph.D. in Computer Science from The National Institute of Applied Science of Lyon, France (INSA de Lyon), in 2000 and 2005 respectively. Dr. Almasri's research interests include Software Engineering and Component-Based Software Development and Management.

Luay Tahat is an assistant professor in the Management Information System and Computer Sceince Department at Gulf University for Sceince and technology since 2008. Prior to that, He was the lead Mobile Network Solution Architect at Alcatel-Lucent in Naperville, USA. He has a master's degree in computer science from Northeastern Illinois University in Chicago and a Ph.D. in computer science from the Illinois Institute of Technology (IIT), also in Chicago. In his time at Alcatel-Lucent, Dr. Tahat has held several positions in software development, system engineering, and system architecture and has contributed to several areas in the fields of software engineering. Dr. Tahat's research interests include software testing, software maintenance, and wireless network solutions. The results of his research were published in several Journals and conference proceedings.

Abdullah A. Awidi, M.R.C.P. (UK), FRCP, is a Professor of Medicine, Hematology and Oncology in the Faculty of Medicine, in Jordan University. He is the founder, president and director of Cell Therapy Center (CTC) in Jordan University. He received several research wards. He has well-established published research in several regional and top international journals.