International Journal of Information, Control and Computer Sciences
ISSN: 2517-9942
Vol:10, No:3, 2016

Identifying a Drug Addict Person Using Artificial Neural Networks

Mustafa Al Sukar, Azzam Sleit, Abdullatif Abu-Dalhoum, Bassam Al-Kasasbeh

***Abstract*—**Use and abuse of drugs by teens is very common and can have dangerous consequences. The drugs contribute to physical and sexual aggression such as assault or rape. Some teenagers regularly use drugs to compensate for depression, anxiety or a lack of positive social skills. Teen resort to smoking should not be minimized because it can be "gateway drugs" for other drugs (marijuana, cocaine, hallucinogens, inhalants, and heroin). The combination of teenagers' curiosity, risk taking behavior, and social pressure make it very difficult to say no. This leads most teenagers to the questions: "Will it hurt to try once?" Nowadays, technological advances are changing our lives very rapidly and adding a lot of technologies that help us to track the risk of drug abuse such as smart phones, Wireless Sensor Networks (WSNs), Internet of Things (IoT), etc. This technique may help us to early discovery of drug abuse in order to prevent an aggravation of the influence of drugs on the abuser. In this paper, we have developed a Decision Support System (DSS) for detecting the drug abuse using Artificial Neural Network (ANN); we used a Multilayer Perceptron (MLP) feed-forward neural network in developing the system. The input layer includes 50 variables while the output layer contains one neuron which indicates whether the person is a drug addict. An iterative process is used to determine the number of hidden layers and the number of neurons in each one. We used multiple experiment models that have been completed with Log-Sigmoid transfer function. Particularly, 10-fold cross validation schemes are used to access the generalization of the proposed system. The experiment results have obtained 98.42% classification accuracy for correct diagnosis in our system. The data had been taken from 184 cases in Jordan according to a set of questions compiled from Specialists, and data have been obtained through the families of drug abusers.

***Keywords*—**Artificial Neural Network, Decision Support System, drug abuse, drug addiction, Multilayer Perceptron.

I. Introduction

ONE of the main problems that face investigators and families is to identify whether the patient is really an addict or not; this can be achieved through some observations on the patient's body or reactions to certain questions and/or the change in his/her social behavior; however, all such measures may diagnose other diseases, as well. Thus, it is not easy identified whether the person is an addict or not which requires more effort from the investigators and families in this regard.

Risk of drug abuse increases greatly during times of transition. For an adult, a divorce or loss of a job may lead to drug abuse; for a teenager, risky times include moving or changing schools. Accordingly, there is an urgent need to prevent its expansion, over control on borders and smugglers, and enact strict legislations to apply advanced and updated means and measures to deal with smugglers' operations [1]. Further, the spread of the drugs is attributed to the enormous development in the means of communications and transport, and to the length of country borders, in addition to the fact that some drugs are produced locally [2], [3].

Drugs are all natural or unnatural materials that contain inhibitory or doping substances; if used in non-medical purposes, they cause an imbalance in the brain and lead to a state of habituation or addiction and cause damage of human health physically, psychologically and socially [4].

Accordingly, there is a strong need for the decisive decision which supports creating a new system to speed up identifying addict's cases early treatment and decrease the addicts' cases as much as possible.

Machine learning techniques have been one of the important fields of the researchers in computer science during the last years; the essential abilities associated with intelligence caused it in that way of the importance [5]. Machine learning has a lot of literature concerning learning theory; it improves the speed and the accuracy of the learning mechanism in different concepts [4].

The technology of machine learning is now convenient for analyzing the medical data. Many types of research have been done in medical diagnosis. Some of the specialized hospitals save data of the correct diagnosis in the form of medical records. All that to be done is to input the patient records with known correct diagnosis into a computer program to run a learning algorithm [6]. In principle, the medical diagnostic knowledge can be automatically derived from the description of cases solved in the past. The derived classifier can then be used either to assist the physician when diagnosing new patients in order to improve the diagnostic speed, accuracy, and/or reliability, or to train the non-specialist students or physicians to diagnose the patients in some special diagnostic problem [7].

This study aims to find a simple mechanism to help an investigator and the families to identify addicts. In this paper, The ANN has been used to identify if the person is an addict or not; this is done by studying a number of addicts' cases in Jordan.

The rest of the paper is organized as follows. Section II will describe ANN from its basic model going through its architecture, algorithms, and design. Section III illustrates the related works that tack about the method for identify addict

Mustafa Sukar (M.Sc. student), Azzam Sleit, Abdullatif Abu-Dalhoum and Bassam Al-Kasasbeh are with the Computer Science Department, University of Jordan, Amman 11942 (e-mail: mustafasukar@yahoo.com, azzam.sleit@ju.edu.jo, a.latif@ju.edu.jo, bassamkasasbeh@yahoo.com).

individuals. Section IV illustrates the system for Drug addiction while Section V illustrates the experiments evaluation and system results. Section VI suggests a possible detection technique. Finally, Section VII concludes the paper and draws possible future work.

II. Artificial Neural Networks

Artificial Neural Networks (ANNs) are a machine learning algorithm use as analytical techniques. It is insured from the neural structure of the human brain. These neurons are an attempt to bind the hugely parallel, distributed computation of biological brains for a variety of purposes. Even though a silicon chip (10-9) is much faster than a nerve cell (10-3), the human brain has great computation capabilities to perform certain operations (such as pattern recognition, perception, and motor control) more efficiently than fastest digital computers. This greater efficiency of the human brain is due the highly complex organization of more than 1010 neurons and 6× 1012 synapses connections with each other's [4], [6]. Nowadays, ANN's is one of the most common research topics and it has been used everywhere in most of the areas over the recent years. ANNs used for solving many problems and achieved good results, so it is one of the most famous models.

ANNs consist of an input layer, number of hidden layers and an output layer. The input layer consists of a number of neurons that represent the system inputs; each input neuron has a weight that adaptively changes during the training phase. The hidden layers are used when the data is not linearly separable. Finally, the output layer represents the final output. ANNs is learnt by-for example, the user needs to gather representative data sets, and then applies training algorithms to automatically learn the patterns of the data.

There are many types of ANNs models. MLP is considered the most common model of neural networks, which are consists of three or more layers (an input and an output layer with one or more hidden layers). This neural network model is known as supervised learning network because it requires a typical output in order to be learnt. The aim of this model is to create a relationship F: X→Y that maps the input X with the output Y using the historical knowledge collected so the MLP can then be used to extract the output when they are unknown [8], [9].

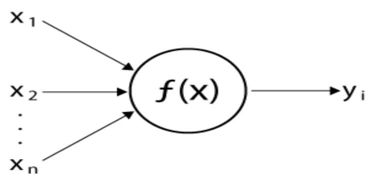

Fig. 1 MLP neural networks

Fig. 1 illustrates a simple MLP neural network where a bias with a weight value equals to one is usually added to the input layer to prevent neural networks from saturating. The input values are fed into the input layer and then get multiplied by the weights as they are passed from the input layer to the first hidden layer.

In the hidden layer, the multiplications are summed in each neuron then processed by a transfer function. The neural networks use a transfer function to move from this layer to the next layer where the transfer functions contain a mathematical function to generalize the desired output [10]. There are many types of transfer functions used in ANN and the most common types of it are: Linear transfer Function (LF); it is computed as follows: a= n where n is the sum of the weighted inputs, Log-Sigmoid Transfer Function; it has a range between 0 and 1 and the function has the following formula:

$$a = \frac{1}{1+e^{-n}} \tag{1}$$

There are a lot of algorithms that may be used in the training phase in MLP. In this paper, we will use one of the most popular neural networks called Feed Forward Back Propagation learning algorithm [11]. With this algorithm, the input repeatedly enters to the neural networks with each output the typical output and error is computed. This error is then back propagated to the neural network and is used to adjust the weight values such that the error is reduced with each iteration and the neural network gets closer to producing the typical output.

ANNs have the ability that allows the trained network to classify new data or noisy data; that is called the generalization. The generalization is one of the main advantages of ANNs. The best generalization is reached when the dataset is split into three parts [12]:

- The training set is used to train a neural network.
- The validation set is used to determine the performance of a neural network on patterns that are not trained during learning.
- The testing set is for finally checking the overall performance of a neural network.

III. Related Works

Several researchers and governments have been conducting research about drug addiction [13]-[18]. Such studies proposed that the health care providers are using computerized systems that are capable of saving a soft copy of the medical information, including prescriptions. These studies aim to manage and reduce the exchange of the illegal recipes that save money and effort. Furthermore, system databases have the advantage of supporting relevant research and studies in order to develop a system to identify addicts' individuals.

National Institute on Drug Abuse (NIDA) encourages investigators to be interested in neuroscience research in terms of advances of relationship between the use of technology and addictive behavior for addicts' identification. It also encourages the use of programs to improving understanding of the brain mechanisms that underlie addiction to simplify preventive and therapeutic interventions [14].

In [15], the authors proposed application of mobile and wearable sensors in Cognitive Behavioral Therapy for drug addiction. This study presents one of the benefits of using mobile applications with Wearable Sensors clinically to avoid therapeutic shock that may lead to suicide in some cases.

In [16], the authors proposed motivating the use of information technology (IT) to produce better results in addicts' identification and/or raise community awareness rather than the traditional methods. They also offered solutions for drug addicts and their families through social media and simplified the procedure to stop or reduce drug use that can give other benefits including diagnosis of other diseases such as AIDS and relevant crimes that can help to build a database which contains information for researchers to use DSS to identify drug addicts.

Machine Learning is becoming increasingly popular in healthcare, if not increasingly essential, and several factors motivated the use of machine learning applications in healthcare, such as fraud and abuse detection, ability of transforming data, and benefit of healthcare providers [17].

ML applications can greatly benefit all parties involved in the healthcare industry. For example, ML can help healthcare insurers detect fraud and abuse, healthcare organizations make customer relationship management decisions, physicians identify effective treatments and best practices, and patients receive better and more affordable healthcare services [17]. The huge amount of data generated by healthcare transactions is too complex and voluminous to be processed and analyzed by traditional methods. ML provides the methodology and technology to transform these piles of data into useful information for decision making [17], [18].

In this paper, we study the causes of drug addiction in Jordan in order to obtain adequate information on the symptoms experienced by drug addicts by specialist doctors and the families of addicts. We then used this information to identify if the person is an addict or not by using ANNs in order to detect early cases of abuse to reduce the spread of addiction.

IV. DSS for Drug Addiction Using ANNs

The ANNs have been applied to many areas which are common in the medical fields. It is used to analyze different diseases and conditions to some extent. In this paper, collecting of information took a long time to design the study questionnaire with the help of specialists in drug addiction in Jordan. The specialists in drug addiction explained that each one of the symptoms in questionnaire form may point to another disease such as a psychological disorder, epilepsy, heart disease and pressure etc. Therefore, the questionnaire used symptoms of addiction for the most popular types of drugs include Captagon, Marijuana, hashish, and cannabis. The experts in this area are advised that most of the addicts were suffering symptoms of these type of drugs in Drug Treatment Centers in Jordan.

Experts stressed that the most important feature of the addicted person is lying and deception where it was found that the best result can be achieved by answering items of the questionnaire which is close to the patient, such as parents, sister, brothers, wife, husband...etc.

The basic information of an addict including (age, gender, level of education and profession) and the doctor diagnosis was not used in current system where it can benefit from such data in the future with new ideas, the name of patient is not written in the questionnaire to maintain the privacy of patient and get the correct information through answering.

The variable of "result of lab test" is used to compare the result of answers to the questionnaire with the result of lab test represented by an addict or not to use in the ANN.

After distribution of the questionnaire on Addiction Treatment Centers and Clinics specializing in the treatment of addiction located in Jordan - Amman, 184 cases were collected of which 155 case were addicted to drugs and 29 cases were suffering from other diseases, then the data were analyzed and encoded and input to ANNs Model to get the result that will be shown in Section V. The following subsections will discuss these issues.

A. Data Preparation

During the examination of the drug abusers, the specialists must take into consideration a large number of possible relevant inputs to make an accurate diagnosis. The accurate diagnosis or treatment decision depends on observations, signs and physical examinations or lab results. The DSS system simulates the diagnostic process of transforming the patient information into mathematical variables. Then, these variables can be processed using certain reasoning mechanisms. Architecture of ANN to Identify Addict Person

There are a large number of relevant inputs must be considered during the diagnosis of addiction, families can identify addict person by answers to questions and a physician reaches an accurate diagnosis or treatment decision based upon physical examinations or lab results.

In this paper, the symptoms addiction dataset used for testing and training the system consisted of a total of 184 cases gathered from centers Treatment addiction, located in Amman, Jordan. Moreover, 57 variables essential to the diagnosis of the addiction than can be obtained from a relative of the addict by answering the questionnaire. These variables:

- Relative: This variable aims to know the relationship between the one who fill in of questionnaire and an addicted person (a member of the family like parents, brother, sisters, wife, etc.). Basic information of an addict (including the age, gender, level of education and profession).
- Attributes (50 factors of symptoms are an addiction in total) that filled in by the relative.
- The result of lab test: Result of the laboratory examination to see whether the patient is an addict or non-addict to be used it in training the neural network.
- Doctor diagnosis to know if the patient is suffering from another disease.

B. Input Variables Encoding Scheme

In this paper, 51 out of 57 variables were encoded into numerical inputs to be used in the neural network that receives numerical inputs values as Variables with two independent attributes, such as the 50 symptoms of addiction and result of lab test; they are encoded into binary values (0, 1). For instance for the answer of the attribute; 1 represents a " Yes " and 0 represents " NO ", result of lab test; 1 represents " addict " and

0 represent " Non-addict " where the basic information in this paper is left to be used with any new idea that maybe appeared in the future. After encoding all variables, the training dataset is standardized based on the information that was collected from the training dataset; the validation and test datasets are also standardized.

C. Number of Hidden Layers and Hidden Neurons

Optimizing the number of hidden layer neurons for an ANN to solve a practical problem is still one of the unsolved tasks in this research area. Therefore, a number of researchers proposed some of the general rules to determining an optimal number of hidden neurons for general applications. One of the rules is: one hidden layer is enough; to find the number of neurons in the first hidden layer; add the number of neurons in the input layer to the number of neurons in the output layer and divide the result by two. Another known rule is: if more than one hidden layer must be used to achieve a less generalization error; then the number of neurons in each hidden layer except the first one will be the number of neurons in the previous layer divided by two [8]. Moreover, [11] and [12] have shown that in most situations, there is no way to determine the best number of hidden layers without training several networks and estimating the generalization error of each, because the best number of hidden layers is 56 units that depend mainly - in a complex way on the number of input and output units, the number of training cases and the complexity of the classification problem should be learned [8].

Accordingly, we used an iterative process to determine the best number of hidden layers and neurons in each hidden layer. In the iterative process, a ten - cross validation technique and percentage split is used to access the generalization for each architecture where the whole process works as follows:

Step 1: Testing the network architecture will be started with only one hidden layer by applying the following equation to find the number of neurons in the first hidden layer:

$$\text{Number of neurons} = \frac{\text{attributes} + \text{classes}}{2} \tag{2}$$

If a number of neurons was not a fixed number, then apply ceil and floor operations, so you will get two values.

Step 2: Add another layer; the number of neurons in this layer will be the number of neurons in the previous layer divided by two.

Step 3: Repeat step two until the number of the hidden neurons in the layer is one.

D. K-Fold Cross Validation Technique

Cross-validation is one of the techniques used to evaluate prediction by models partition of the original dataset into a training set to train the system, and then test the set to evaluate it.

In k-fold cross-validation is used, the original dataset is randomly partitioned into k equal size sub samples. Of the k sub samples, a single sub sample which is retained as the validation data for testing the model, and the remaining k-1 sub samples are used as training data. The cross-validation process is then repeated k times (the folds), with each of the k sub samples used exactly once as the data validation. The k results from the folds can then be averaged (or otherwise combined) to produce a single estimation. The advantage of this method is that all observations are used for both training and validation, and each observation is used for validation exactly once.

For classification problems, one typically uses stratified k-fold cross-validation, in which the folds are selected so that each fold contains roughly the same proportions of class labels.

In repeated cross-validation, the cross-validation procedure is repeated n times, yielding n random partitions of the original sample. The n results are again averaged (or otherwise combined) to produce a single estimation.

V. Experiments Evaluation

In this section, we show the preliminary results obtained by the dataset described in Section III. In our experiments, we use WEKA Toolbox to evaluate our dataset. Waikato Environment for Knowledge Analysis (WEKA) is an open source data mining software suite for the Java language, developed at the University of Waikato in New Zealand. The data mining algorithms in WEKA may be applied to a dataset, and it may be called from Java code. WEKA contains a lot of algorithms for data preprocessing, clustering, classification, association rules, visualization, and regression [19], [20].

The classification was done using Intel® Core™ i5-4210U CPU @ 1.70 GHz 2.40 GHz, 8.00GB RAM with Windows 8.1 64-bit Operating System.

Because different performance metrics is appropriate in different settings, this paper utilizes seven performance metrics: True Positive Rate (TPR), False Positive Rate (FPR), Overall Accuracy (A), Precision (P), and Root Mean Square Error (RMSE). (TPR) represents the rate of cases of a case identified correctly, (FPR) represents the rate of a case of no identified incorrectly, (A) is the total rate of correct detections, (P) represents the predicted positive cases that were correctly classified, and (RMSE) provides information on the efficiency that indicates the difference between the outputs and the targets. Lower values of the RMSE indicate more accurate evaluation. Zero means no error.

$$\boldsymbol{TPR = \frac{TP}{TP + FN}} \qquad \boldsymbol{FPR = \frac{FP}{FP + TN}}$$

$$\boldsymbol{A = \frac{TP + TN}{TP + TN + FP + FN}} \qquad \boldsymbol{P = \frac{TP}{TP + FP}}$$

$$\boldsymbol{RMSE = \sqrt{\frac{\sum_{i=1}^{n}(O_i - T_i)^2}{n}}}$$

where the TP= true positive, TN= true negative, FP= false positive, FN= false negative, O_i and T_i are the output and target values, respectively, and n is the total number of data points.

All dataset which consist of 184 sample cases were trained using 10-fold cross validation method. The different architectures were compared based on the average classification accuracy for each one. These different architectures trained with 10 different seed numbers, and then the average accuracy of each one was calculated. In this paper for all models, we assume the following:

1. MLP neural network was used for building all architectures.
2. The number of neurons in the input layer was 50.
3. The number of neurons in the output layer was one.
4. Log-Sigmoid was the transfer function that was used in the layer for each architecture
5. The Learning rate (used for weight adjustment on each repeat) is 0.3.
6. The Momentum (used for weight adjustment during Back propagation, in order to speed up convergence and avoid local minima) is 0.2.

We have built the models with one; two, and three hidden layers, then we have compared the average of classification accuracy of architecture, and then find the best result in different architecture which has higher accuracy and less error ratio. Then it is used as a final architecture for the DSS.

TABLE I
SUMMARY OF RESULTS

Network Architecture	Classification Accuracy %	RMSE	FPR	P
50-25-1	98.3696	0.11002	0.003	0.985
50-26-1	**98.4239**	**0.10869**	**0.0029**	**0.9855**
50-25-12-1	96.84783	0.13385	0.0058	0.974
50-25-13-1	96.73913	0.15263	0.0059	0.9735
50-26-13-1	97.28261	0.13845	0.005	0.977
50-25-12-6-1	97.55435	0.14063	0.0045	0.979
50-25-13-6-1	97.7174	0.13853	0.0042	0.9802
50-25-13-7-1	97.33695	0.14293	0.0077	0.9772
50-26-13-6-1	97.5	0.13899	0.0046	0.9786
50-26-13-7-1	97.33695	0.14257	0.0049	0.9774

TABLE IIIII
THE CONFUSION MATRIX OF 50-26-1 ARCHITECTURE

	Non-Addict	Addict
Non-Addict	29	0
Addict	3	152

Table IVV shows the experimental results of the 50-26-1 (best architecture) test sets that have been presented as a confusion matrix.

Table VI illustrates the summary of the results obtained using different hidden layers for building the ANN that is used as the final model for the DSS. Where the best architecture was obtained by using one hidden layer; the architecture is 50-26-1 and 98.4239% average classification accuracy with 0.10869 error rate was obtained. It is noted that by using two hidden layers, good results were not achieved, where the rate of accuracy was decreased and the error ratio was increased, when another three hidden layer have been added, the best average classification accuracy was risen to 97.7174%.

We note from Table VIIVIII that there are just 3 cases classified as non-addict out of 155 Addict cases in the data set. These classification accuracies for each disease can be simply explained as the following: According to specialists, to determine if the case is addicted or not, there are some similar symptoms with each other, but they can differentiate from each other by some specific variables. Also, the practical experience of doctors in the Al-Rashid Hospital Center, Anti-Narcotics Dept in Jordan, and National Center for Addiction Treatment in Jordan allow them to diagnose their patients in high accuracy, so the samples in the training set in this system were given with high accuracy.

Consequently, the nature of the data splits at testing folds may have examples in the training sets that were very similar to the data pattern that was on the test set, so this produced high classification accuracy, so the highly accurate results obtained in this paper were realistic.

VI. CONCLUSION AND FUTURE WORKS

In this paper, we aimed to reach accurate diagnosis to identify addicts and treat civilians and preserve their lives, which are plays one of the major roles to help families, physician and investigators. Many researchers also developed DSS to monitor trafficking of drugs on borders systems which proved its effectiveness in obtaining high accuracy to discover the drugs hidden, on another hand there are many research and studies about addictions symptoms and the reasons of propagation, and drugs abuse. This paper proposed a DSS to addicted diagnoses. The proposed system uses one of the most widespread machine learning techniques; a MLP feed forward neural network. It consisted of an input layer with 50 neutrons of symptoms of addictions that represent the inputs variables, one hidden layer and an output layer that represents an Addict or Non-Addict. Moreover, 10-fold cross validation was used to access the generalization of the proposed system using 184 cases that were collected from Addiction Treatment Centers and clinics specializing in the treatment of addiction located in Jordan - Amman, of which 155 cases were addicted to drugs and 29 cases as non-addict, giving an average of 98.4239% classification accuracy. Based on these results, it can help to speed up treatment of addicts that can be used by families, physicians and investigators! Moreover, if more dataset was used for training the proposed system, that the result we got absolutely is proved with an efficient accuracy. One of the most important of major future vision to extend such work by using special sensors that can help to treat addicts by such as monitoring body temperature, the increase in respiratory rate, heart palpitations, blood pressure... etc. from symptoms of addictions. Moreover, such work can be extended by applying another machine learning types such as Neuro-Fuzzy, Artificial Immune System, and Decision Trees to compare the results with the MLP. Finally, one of the things that hopefully be conducted in the future is that the relevant questionnaires are filled in by other countries from most over the world to get different cases and more cases that will support the results obtained in this paper.

ACKNOWLEDGMENT

The authors would like to thank Dr. Refaat Al-masry head of Al Rashid Hospital Center for Psychiatry and Addiction who allow us to help us to prepare questionnaire related to these works with physicians who have good experience in treatment addiction. Also Dr. Jamal Anani, Chairman of the National Center for Rehabilitation of Addicts, Col. Firas Sulaiman Al Zu'bi, head of the Division of anti-drug in Capital of Jordanian-Amman and Maj.Fawaz Masaeed Head of Center Addiction Treatment (Anti-Narcotics Dept.) they provided us helpful information about drug abuse, guidance, encouragement and patience in helping us to collect the patient medical records and interview of families of addicts.

REFERENCES

[1] NIDA, Drug, Brains and Behavior: The Science of Addiction. National Institutes of Health Publication No. 07-5605, National Institute on Drug Abuse, 2007.

[2] Renthal, William, and Eric J. Nestler. "Epigenetic mechanisms in drug addiction." *Trends in molecular medicine* 14.8 (2008): 341-350.

[3] Schwan, Sofie, et al. "A signal for an abuse liability for pregabalin—results from the Swedish spontaneous adverse drug reaction reporting system." *European journal of clinical pharmacology* 66.9 (2010): 947-953.

[4] Pudenz, Kristen L., and Daniel A. Lidar. "Quantum adiabatic machine learning." *Quantum information processing* 12.5 (2013): 2027-2070.

[5] Bottou, Léon. "From machine learning to machine reasoning." *Machine learning* 94.2 (2014): 133-149.

[6] Fausett, Laurene. *Fundamentals of neural networks: architectures, algorithms, and applications*. Prentice-Hall, Inc., 1994.

[7] Michalski, Ryszard S., Jaime G. Carbonell, and Tom M. Mitchell, eds. *Machine learning: An artificial intelligence approach*. Springer Science & Business Media, 2013.

[8] Haykin, Simon, and Richard Lippmann. "Neural Networks, A Comprehensive Foundation." *International Journal of Neural Systems* 5.4 (1994): 363-364.

[9] Seung, Sebastian. "Multilayer perceptrons and backpropagation learning. 9.641 Lecture4. 1-6." (2002).

[10] Araghinejad, Shahab. *Data-Driven Modeling: Using MATLAB® in Water Resources and Environmental Engineering*. Vol. 67. Springer Science & Business Media, 2013.

[11] Isa, Iza Sazanita, et al. "Comparisons of MLP transfer functions for different classification classes." *Control System, Computing and Engineering (ICCSCE), 2012 IEEE International Conference on*. IEEE, 2012.

[12] Heskes, Tom. "Practical confidence and prediction." *Advances in Neural Information Processing Systems 9: Proceedings of the 1996 Conference*. Vol. 9. MIT Press, 1997.

[13] J. Frazier, 2013, Health IT helping to fight the prescription drug abuse epidemic, ONC behavioral health subject matter expert, http://www.healthit.gov/buzz-blog/health-innovation/health-helping-fight-prescription-drug-abuse-epidemic/ (accessed September 19, 2015).

[14] Goldstein, Rita Z., et al. "The neurocircuitry of impaired insight in drug addiction." *Trends in cognitive sciences* 13.9 (2009): 372-380.

[15] Fletcher, Richard Ribón, et al. "Wearable sensor platform and mobile application for use in cognitive behavioral therapy for drug addiction and PTSD." *Engineering in Medicine and Biology Society, EMBC, 2011 Annual International Conference of the IEEE*. IEEE, 2011.

[16] Bickel, Warren K., Darren R. Christensen, and Lisa A. Marsch. "A review of computer-based interventions used in the assessment, treatment, and research of drug addiction." *Substance use & misuse* 46.1 (2011): 4-9.

[17] Koh, Hian Chye, and Gerald Tan. "Data mining applications in healthcare." *Journal of healthcare information management* 19.2 (2011): 65.

[18] Viveros, Marisa S., John P. Nearhos, and Michael J. Rothman. "Applying data mining techniques to a health insurance information system." *VLDB*. 1996.

[19] Hall, M., et al. "The WEKA data mining software: an update. ACM SIGKDD Explor Newslett 2009; 11: 10–8."

[20] Bouckaert, Remco R., et al. "WEKA---Experiences with a Java Open-Source Project." *The Journal of Machine Learning Research* 11 (2010): 2533-2541.