Journal of Theoretical and Applied Information Technology
20th May 2014. Vol. 63 No.2

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

APPLYING DECISION TREE ALGORITHM AND NEURAL NETWORKS TO PREDICT FOREST FIRES IN LEBANON

[1,4] ALI KAROUNI, [1,2]BASSAM DAYA, [3,4]PIERRE CHAUVET

[1] Lebanese University, Ecole Doctorale des Sciences et de Technologie, Lebanon

[2]Umm Al-Qura University, Faculty of Social Sciences, Information Science Department, Makkah – KSA

[3] LARIS EA, L'UNAM Université, Université Catholique de l'Ouest, 3 place André-Leroy BP 10808, 49008 Angers – France

[4]LARIS EA, L'UNAM Université, Université d'Angers, Angers - France

E-mail: [1]ali.karouni@etud.univ-angers.fr , [2]b_daya@ul.edu.lb, [3]pierre.chauvet@uco.fr

ABSTRACT

Fires have been threatening green forestry all over the world. In Lebanon, green areas declined dramatically during the last decades, what imposes an urgent intervention with strict governmental policies and support of non-governmental organizations. The orientation is towards techniques that predict high fire risks, allowing for precautions to preclude fire occurrences or at least limit their consequences. Two data mining techniques are used for the purpose of prediction and decision-making: Decision trees and back propagation forward neural networks. Four meteorological attributes are utilized: temperature, relative humidity, wind speed and daily precipitation. The obtained tree drawn from applying the first algorithm could classify these attributes from the most significant to the least significant and better foretell fire incidences. Adopting neural networks with different training algorithms shows that networks with 2 inputs only (temperature and relative humidity) retrieve better results than 4-inputs networks with less mean squared error. Feed forward and Cascade forward networks are under scope, with the use of different training algorithms.

Keywords: *Forest Fires Prediction, Decision Tree, Neural Networks*

1. INTRODUCTION

Data mining is the process of identifying correlations in large data sets involving methods of artificial intelligence, machine learning, statistics, and visualization. Its aim is to find knowledge from databases that support the decision-making process [1]. The two primary functions of data mining are: prediction, which involves extracting unknown relationships/patterns from known values; and description, which detects concepts and gives interpretations for large database. A data mining process includes four steps: Data acquisition, data preprocessing, data exploration and model building, interpretation and evaluation [2]. The most commonly used techniques in data mining applications are Neural Networks and Decision Trees.

1.1. Artificial Neural Networks

Artificial neural networks (ANN) are modeling techniques capable of modeling extremely complex functions. A neural network is a group of interconnected artificial neurons with linear or nonlinear transfer functions. ANN learns the relation between inputs and outputs of the system through an iterative process called training. It uses the obtained relation to make claims about something that will happen. Neural networks have been commonly used in many applications: character recognition, image compression, stock market prediction, medicine, security and many other miscellaneous applications.

Maes et al. (2002) applied ANNs to automated detection of credit card fraud: the fraud detection process is found to be faster with ANN than with Bayesian belief networks, which is very essential in financial transactions [3]. Simard et al. (2003) showed that neural networks are a powerful technology for classification of

Journal of Theoretical and Applied Information Technology
20th May 2014. Vol. 63 No.2

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

visual inputs arising from documents and achieve the best performance on a handwriting recognition task [4]. Vintan et al. (2004) proposed neural prediction techniques to anticipate a person's next movement focusing on multi-layer perceptron with back-propagation learning. The simulation results showed accuracy in next location prediction reaching up to 92% [5]. Mostafa et al. (2005) created and trained a two-layer feed forward neural network with training set of faces and non-faces. The new system was designed to detect upright frontal faces in color images with simple or complex background. The system has acceptable results regarding the detection rate, false positives and average time needed to detect a face [6]. Knox et al. (2007) developed an efficient method to recognize laughter segments, ultimately for the purpose of speaker recognition. They found neural networks to be a particularly good fit for this problem and achieved an equal error rate of 7.9% for laughter recognition [7]. Khashman et al. (2008) verified that neural networks can be trained to establish the non-linear relationship between the image intensity and its compression ratios in search for an optimum ratio and thus providing substantial improvements in picture quality at higher compression ratios. Their neural networks yielded 96.67% correct recognition rate of optimum compression ratios [8]. Mazloumi et al. (2012) developed an ANN to predict bus travel time on the basis of a range of input variables including traffic flow data and yielded improvements in prediction accuracy [9]. Akoum et al. (2010) utilized ANN to verify and to classify the different types of the vehicles, and a ratio of identification of about 97% was obtained. [10].

1.2. Decision Trees

Decision trees are a form of multiple variable analysis. They allow predicting, explaining, describing, and classifying an outcome. Decision trees classify instances by sorting them down the tree from the root to some leaf node. The final result is a decision tree in which each branch represents a possible scenario of decision and its outcome. Decision tree learning algorithm has been successfully used in expert systems in capturing knowledge for the purpose of decision making. Hey are also used to model behaviors in multi-agent systems (behavior trees).

Suontausta et al. (2000) studied a decision tree based text-to-phoneme mapping. The phoneme accuracy of the most effective mapping was 99% on the training set and 91% on the test set of the pronunciation dictionary [11]. Kim et al. (2001) proposed a marketing rule-extraction technique for personalized recommendation on Internet storefronts using decision tree induction techniques. They found that decision trees can generate marketing rules that match customer demographics to product categories [12]. Dey et al. (2002) stated that risk management using a combined analytic hierarchy process and decision tree approach could provide an effective means for managing a complex project efficiently, and for fighting against time, cost, and quality non-achievement [13]. Pavlopoulos et al. (2004) used a fully expanded decision tree classifier which resulted in a classification accuracy (total corrects/total tested) of 90% (45 correct/50 total records) [14]. Singh et al. helped drug discoverers to predict the functions of proteins which are responsible for various diseases in human body using decision tree technique. They found that the tree with greater depth ensures more number of tests before functional class assignment and thus results in more accurate predictions than the existing prediction technique. The percentage accuracy of the new Human Protein Function predictor is 72% and that of the existing prediction technique is 44% [15]. Markey et al. (2011) used decision tree analysis to support an intrusion detection team with the many challenges of defending a network. They found that due to the ever-increasing volume of data, decision trees have the potential to save time for security experts and assist in the analysis of malicious data [16]. Kabra et al. (2011) applied decision tree algorithm on engineering students to predict their performance. It enabled to identify the students in advance who were likely to fail, with a true positive rate of 0.907 [17]. Hailemariam et al (2011) studied the ability to accurately determine localized building occupancy in real time, including intelligent control of building systems to minimize energy use and real-time building visualization. Decision Trees were used to perform the classification and to explore the relationship between different types of sensors, features derived from sensor data, and occupancy. When used with a simple threshold, this individual feature detected occupancy with 97.9% accuracy. Combining multiple motion sensor features with a decision tree, the accuracy improved to 98.4% [18].

Journal of Theoretical and Applied Information Technology
20th May 2014. Vol. 63 No.2

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

1.3. The Problem

During the last 40 years, more than 35 % of the existing forest cover in Lebanon has deteriorated. Forest fires, among other natural and human threats, have a major cause of this decline [19]. To date, they continue to be one of the most dangerous threats endangering Lebanon's forest and causing their decline. According to AFDC, the Association for Forest Development and Conservation, forest fires that occurred between 1993 and 2005 amount to 70.600 fires in different parts of the country. The number of yearly burned areas has tremendously increased in the year 2006-2007 due to the July war 2006 and to the October 2007 fires, which burned huge forested areas in only a few days. In this paper, forest fire occurrence prediction is studied using the techniques of decision trees and artificial neural networks. We used weather data of the year 2012 collected from North Lebanon, Kfarchakhna station, provided by Lebanese Agricultural Research Institute (LARI). 361 sets were under study. Each set is related to a day of the year containing four measured attributes: noon temperature T, relative humidity RH, maximum speed of the average wind WS and 24h-precipitation P. Fire incidents for the year 2012 scoping North Lebanon were recorded from various sources: meteorological sites and newspapers. Our purpose is to examine the performance of decision tree algorithm as well as artificial neural networks in natural phenomena like forest fires. The latter will point toward relying on data mining techniques which are not expensive and easy to use to predict and then take the necessary and prompt actions to prevent a fire occurrence or at least minimize its disorders. In this way, we will be dispensing the costly use of the certified and complicated widely used weather indices that claim to predict fire incidents which is not feasible in developing countries like Lebanon.

2. APPLYING DECISION TREE ALGORITHM

A decision tree is a tree in which each branch node represents a choice between a number of alternatives. Decision tree starts with a root node on which it is for users to take actions. From this node, users split each node recursively according to decision tree learning algorithm. Leaf nodes represent the decisions [20]. There are various algorithms in this area like ID3, C4.5, ASSISTANT etc. ID3 algorithm is selected in this study because it builds tree based on the information obtained from the training instances and then uses the same to classify the test data.

ID3 is a simple decision tree learning algorithm developed by Ross Quinlan (1983). The basic idea of ID3 algorithm is to construct the decision tree by employing a top-down flowchart testing each attribute at every node using Information Gain property which separates the training examples according to their target classification. To get this gain, entropy is to be computed [21].

Entropy E is a measure characterizing the impurity of an arbitrary set of examples S. Its complete formula is as follows:

$$E(S) = -\sum_{i=1}^{c} P(c_i) log_2 P(c_i)$$

Where c is the set of desired classes: *fire/no fire*
Gain is the information gained by selecting attribute A_i to branch the data; it is given by the difference of prior entropy and the entropy of selected branch. The attribute with the highest gain is chosen to split the tree. Then, the entropy of Ai is needed; it can be calculated according to the following: If we have attribute A_i, with z values, this will partition S into z subsets $S_1, S_2, \ldots, S_z$. The expected entropy if A_i is used as the current root;

$$E_{A_i}(S) = \sum_{i=1}^{z} \frac{S_i}{S} E(S_i)$$

Thus

$$Gain(S, A_i) = E(S) - E_{A_i}(S)$$

In reference to our case study, meteorological data of the year 2012 (361 cases) are used where we have four attributes (noon temperature T, relative humidity RH, maximum speed of the average wind WS and 24h-precipitation P). Each attribute has 3 categories: low, moderate and high according to Table 1. The decision will be then "F" denoting extreme fire potential and "NF" denoting unpredictable fire occurrence.

Table 1: Categories of Attributes

	T(˚C)	RH(%)	WS(m/s)	P(mm)
High	T≥30	RH≥65	WS≥3	P≥5
Moderate	25≤T<30	50≤RH<65	1≤WS<3	1≤P<5
Low	T<25	RH<50	WS<1	P<1

A sample of categorization is found in Table 2.

Journal of Theoretical and Applied Information Technology
20th May 2014. Vol. 63 No.2

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

To draw the decision tree, we have to find the root attribute to start splitting down. Upon reviewing all the cases, we can conclude Table 3.

Note that in 2012, 26 fire occurrences were detected out of 361 days, then

$$E(S) = -\left(\left(\frac{26}{361}\right) log_2\left(\frac{26}{361}\right) + \left(\frac{335}{361}\right) log_2\left(\frac{335}{361}\right)\right) = 0{,}373425$$

Now to get the gain of temperature, we have to compute the entropies $E_{T_l}(S)$, $E_{T_m}(S)$ and $E_{T_h}(S)$:

$$E_{T_l}(S) = -\left(\left(\frac{3}{271}\right) log_2\left(\frac{3}{271}\right) + \left(\frac{268}{271}\right) log_2\left(\frac{268}{271}\right)\right) = 0{,}0878066$$

$$E_{T_m}(S) = -\left(\left(\frac{21}{88}\right) log_2\left(\frac{21}{88}\right) + \left(\frac{67}{88}\right) log_2\left(\frac{67}{88}\right)\right) = 0{,}7927652$$

$$E_{T_h}(S) = -\left(\left(\frac{2}{2}\right) log_2\left(\frac{2}{2}\right) + \left(\frac{0}{2}\right) log_2\left(\frac{0}{2}\right)\right) = 0$$

Then $E_T(S) = \left(\left(\frac{271}{361}\right) E_{T_l}(S) + \left(\frac{88}{361}\right) E_{T_m}(S) + \left(\frac{2}{361}\right) E_{T_h}(S)\right) = 0{,}259166$

Hence $Gain(S,T) = E(S) - E_T(S) = 0{,}114259$

Similarly,

$$Gain(S,RH) = E(S) - E_{RH}(S) = 0{,}0498435$$
$$Gain(S,WS) = E(S) - E_{WS}(S) = 0{,}0376338$$
$$Gain(S,P) = E(S) - E_P(S) = 0{,}0134699$$

It's clearly shown that the highest gain is that of temperature and it's the root. In the 2nd phase, we considered the cases S_1 where we have a low temperature as a 1st step. Now we have 3 fire cases out of 271. The entropy is $E(S_1)$= 0,373425**.** Upon calculating the gains**,** we obtained that relative humidity has the highest gain:

$$Gain(S_1,RH) = E(S_1) - E_{RH}(S_1) = 0{,}2934059$$
$$Gain(S_1,WS) = E(S_1) - E_{WS}(S_1) = 0{,}2922296$$
$$Gain(S_1,P) = E(S_1) - E_P(S_1) = 0{,}2905827$$

Then the cases S_2 and S_3 were taken where we have moderate temperatures and high temperatures respectively. Also this time, relative humidity is found to have the highest gain.

In a downward recursive manner, splitting of the tree continued based on information gain function to conclude finally the decision tree (Fig. 1).

3. APPLYING NEURAL NETWORKS

The objective of our study is to predict a forest fire. Predicting is making expectations about something that will happen, often based on information from past. When the technique of decision trees was applied, it was shown that some attributes are more significant than others. It is clear that temperature and relative humidity are more informative than wind speed and precipitation as temperature acquired the greatest gain and the first generation of subsets was split according to the judgment of relative humidity (Refer to Fig. 1). We applied neural networks once using the four attributes and then using only two inputs: temperature and relative humidity.

Two types of networks of back propagation neural networks were tested: feed forward Function fitting neural network and cascade feed forward network. Back propagation is the optimal learning algorithm in training multilayer perceptrons (MLP), where the input signal propagates through the network in a forward direction, layer by layer, from left to right to conclude the output. In feed-forward back propagation artificial neural network model, shown in Fig.2, the propagation goes in only one direction from the input to the hidden layer(s) and towards the output layers. In our test, three different networks were used with 5, 10 and 15 neurons in the hidden layer respectively. While in cascade-forward back propagation network, shown in Fig.3, each layer of neurons related to all previous layer of neurons. Here we used two intermediate layers, the first has 3 neurons and the second has 2 neurons. In both techniques, the aim is to find for a given non-linear inputs/outputs data a transfer function that best serves to predict accordingly a good response to any posing inputs. Mathematically, a multilayer perceptron network is a function that consists of compositions of weighted sums of functions corresponding to the neurons. Initially, in the training phase, random weights are assigned to the networks. These weights are changed with

Journal of Theoretical and Applied Information Technology
20th May 2014. Vol. 63 No.2

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

every iteration in order to minimize the mean square error between the obtained output and the target output.

In our case study, the input is once a vector of four elements in case the four attributes (temperature, relative humidity, wind speed and precipitation) are introduced, and in the next time it is a vector of two inputs when only temperature and relative humidity are taken. The network has a unique output: '1' upon detection of a fire or '0' when there is no fire risk. Three training algorithms were used: *trainlm*, *traingdx* and *trainrp*.

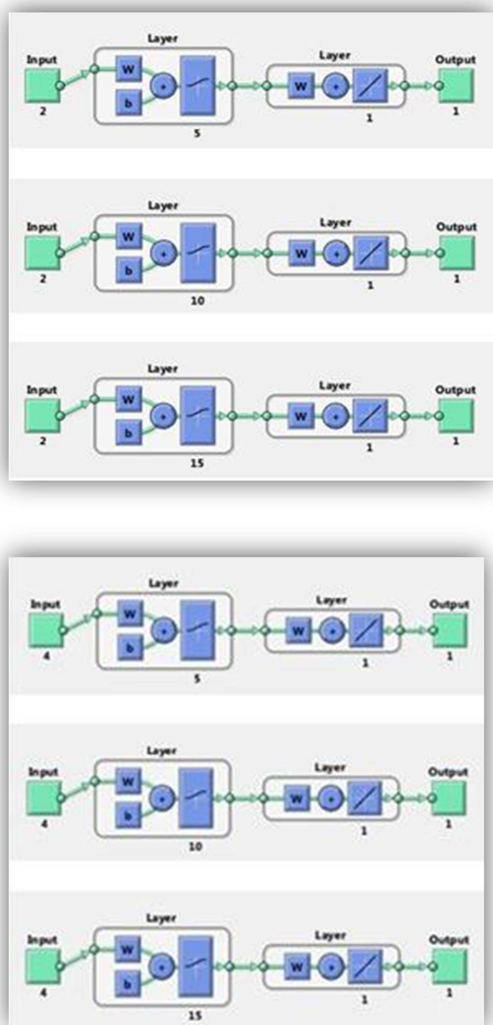

Figure 2: Feed-Forward Back Propagation Networks Used: Once with 4 Inputs and Second With 2 Inputs, and Each with 5, 10 & 15 Neurons in the Hidden Layer

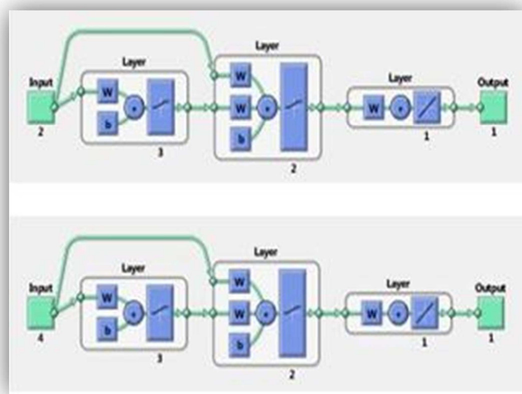

Figure 3: Cascade-Forward Back Propagation Networks Used: Once with 4 Inputs and Second with 2 Inputs

By using Matlab, we test both networks, and the results are shown in Table 4.

Mean squared error (mse) of an estimator is one of many ways to quantify the difference between values implied by an estimator and the true values of the quantity being estimated.

If $\hat{Y}$ is a vector of n predictions, and Y is the vector of the true values, then the mse of the predictor is: $mse = \frac{1}{n}\sum_{i=1}^{n}(\hat{Y}_i - Y_i)^2$

An mse value closer to 0 indicates a fit that is more useful for prediction.

Upon analyzing 4-inputs networks, we can see that when using trainlm algorithm, feed forward generally had less mse values than cascaded networks, and within feed forward networks, the one with 15 neurons in the hidden layer got the least error with a value of 0.0033.When using traingdx algorithm, cascaded network was better than feed forward network with 5 neurons in its hidden layer, but it is not the case when using 10 and 15 neurons in the hidden layer where the testing mse obtained are 0.0073 and 0.0069 respectively. While upon using trainrp algorithm, the cascaded network got the least computed mean squared error among all feed forward networks used with a value of 2.15E-04 much closer to 0.

Upon analyzing 2-inputs networks, when using trainlm algorithm, the cascade forward network got the least mse over the three feed forward networks used with an mse of 1.98E-04. Here it is recorded that the feed forward network with 5 neurons in the hidden layer got the least mse among other feed forward networks used. The same could be observed when using trainrp algorithm. Upon using the algorithm of traingdx,

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

the cascaded network was better than all feed forward networks utilized. Entirely, the 2-inputs networks got less mean squared errors than the 4-inputs networks.

To better examine the performance of our networks and have a critical comparisons; a set of measures were used: precision, sensitivity, specificity, accuracy ACC and area under ROC curve AUC. These are statistical measures of the performance of our binary classification test (fire/no fire).

Precision is the probability of the truly predicted fire cases out of all recorded fire cases. Sensitivity relates to the test's ability to identify extreme fire risks results. Specificity relates to the test's ability to identify no fire results. ACC is the proportion of truly predicted outputs (fire and no-fire) in the population. AUC estimate can be interpreted as the probability that the classifier will assign a higher score to a randomly chosen positive example than to a randomly chosen negative example [22].
In reference to Table 5, these measures were computed for the utilized neural networks. We can clearly see that the 2-inputs networks retrieved higher rates of these measures than the 4-inputs networks which confirm the above concluded result.
Within the 4-inputs feed forward networks, the networks using trainrp algorithm showed better performance than the ones using trainlm and traingdx algorithms. The same is for the tested 4-inputs cascade forward networks.

Within the 2-inputs feed forward networks, it could be shown that the trainlm algorithm got better results than the other training algorithms. While in the 2-inputs cascaded networks, trainlm and trainrp algorithms were found to have the same performance which is better than traingdx.

4. CONCLUSION

This paper presented an overview of forest fire occurrence prediction methods: Decision Tree and Neural Networks. Data of the year 2012 for North Lebanon was taken for study. Both techniques are characterized by ease of use and low cost especially for countries like Lebanon. Decision tree was drawn beyond a certain categorization of the attributes. It can be shown that two attributes were more significant than others. Here appeared the importance of decision tree in the classification of attributes and ignorance of the least significant ones. The resulting tree seemed to be clear, adequate and could be applied to likely predict a fire occurrence. For generalizing and adopted such technique in Lebanon and Mediterranean basin, a minimum number of 10 years should be under scope for the decision to be more precise. Neural network is the second method used in this study. First, the 4 parameters of temperature, relative humidity, wind speed and precipitation were taken as inputs for our experimented networks. Then only the two most significant attributes (temperature and relative humidity) were taken as inputs. Two back propagation network architectures were used: feed-forward network and cascade-forward network. Both showed a good performance in the field of forest fire prediction, along with different training algorithms. It was obtained that 2-inputs networks retrieve better results upon computing mean squared error than 4-inputs networks. Hence, we can ignore two inputs and the decision is easier to be concluded especially for developing countries as it limits the number of measures to do, and thus reduce the problem created by potential erroneous weather forecasting and the cost of forecasting. Also, one could observe that changing the training algorithm and the number of neurons in the hidden layer for feed-forward networks had an impact on minimizing the error. When using trainlm and traingdx algorithms, feed-forward networks generally got less mse than cascaded networks while upon using trainrp algorithm, the cascaded networks were better. Another performance tests were performed by computing the measures of precision, sensitivity, specificity, ACC and AUC which all confirm the advantage of 2-inputs networks over the 4-inputs networks. They also showed that within 2-inputs networks, using trainlm gave better results. To improve neural networks by minimizing more and more the mean squared error and getting higher rates of statistical measures, the number of years taken for training and testing is also suggested to be increased.

REFERENCES:

[1] Soumen Chakrabarti, Martin Ester, Usama Fayyad, Johannes Gehrke, Jiawei Han, Shinichi Morishita, Gregory Piatetsky-Shapiro and Wei Wang, Data Mining Curriculum: A Proposal (Version 0.91), August 2004

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

[2] Lijia Guo, Applying Data Mining Techniques in Property/Casualty Insurance, Forums of the Casualty Actuarial Society, 2003

[3] Sam Maes, Karl Tuyls, Bram Vanschoenwinkel and Bernard Manderick, Credit Card Fraud Detection Using Bayesian and Neural Networks, Proceedings of the First International NAISO Congress on Neuro Fuzzy Technologies, 2002

[4] Patrice Y. Simard, Dave Steinkraus and John C. Platt, Best Practices for Convolutional Neural Networks Applied to Visual Document Analysis, Proceedings of the Seventh International Conference on Document Analysis and Recognition (ICDAR 2003) 0-7695-1960-1/03, IEEE Computer Society

[5] Lucian Vintan, Arpad Gellert, Jan Petzold, and Theo Ungerer, Person Movement Prediction Using Neural Networks, Technical Report, Institute of Computer Science, University of Augsburg, April 2004

[6] Lamiaa Mostafa and Sherif Abdelazeem, Face Detection Based on Skin Color Using Neural Networks, GVIP 05 Conference, 19-21 December 2005, CICC, Cairo, Egypt

[7] Mary Tai Knox and Nikki Mirghafori, Automatic laughter detection using neural networks, In INTERSPEECH-2007, 2973-2976.

[8] Adnan Khashman and Kamil Dimililer, Image Compression using Neural Networks and Haar Wavelet, WSEAS Transactions on Signal Processing, ISSN: 1790-5052 Issue 5, Volume 4, May 2008 pp. 300-309

[9] Mazloumi, Moridpour, Currie, and Rose, Exploring the Value of Traffic Flow Data in Bus Travel Time Prediction, Journal of Transportation Engineering, 2012, 138(4), 436–446

[10] Bassam Daya, Al Hussain Akoum, Pierre Chauvet, Neural Network Approach for the Identification System of the Type of Vehicle, cicn, pp.162-166, 2010 International Conference on Computational Intelligence and Communication Networks, 2010

[11] Janne Suontausta and Juha Häkkinen, Decision tree based text-to-phoneme mapping for speech recognition, INTERSPEECH 2000: 831-834

[12] Jong Woo Kim, Byung Hun Lee, Michael J. Shaw, Hsin-Lu Chang and Matthew Nelson, Application of Decision-Tree Induction Techniques to Personalized Advertisements on Internet Storefronts, International Journal of Electronic Commerce Spring2001 Vol. 5 No.3 pp 45-62

[13] Prasanta Kumar Dey, Project Risk Management: A Combined Analytic Hierarchy Process and Decision Tree Approach, Cost Engineering Vol. 44/No. 3 March 2002

[14] Sotiris A Pavlopoulos, Antonis CH Stasis and Euripides N Loukis, A decision tree – based method for the differential diagnosis of Aortic Stenosis from Mitral Regurgitation using heart sounds, BioMedical Engineering OnLine 2004, 3:21 doi:10.1186/1475-925X-3-21, June 2004

[15] Manpreet Singh, Parminder Kaur Wadhwa and Parvinder Singh Sandhu, Human Protein Function Prediction using Decision Tree Induction, IJCSNS International Journal of Computer Science and Network Security, VOL.7 No.4, April 2007

[16] Jeff Markey and Antonios Atlasis, Using Decision Tree Analysis for Intrusion Detection: A How-To Guide, GIAC (GCIA) Gold Certification, June 2011

[17] R. R. Kabra and R. S. Bichkar, Performance Prediction of Engineering Students using Decision Trees, International Journal of Computer Applications (0975 – 8887) Volume 36– No.11, December 2011

[18] Ebenezer Hailemariam, Rhys Goldstein, Ramtin Attar and Azam Khan, Real-Time Occupancy Detection using Decision Trees with Multiple Sensor Types, SimAUD 2011 Conference Proceedings: Symposium on Simulation for Architecture and Urban Design.

[19] Reforest Lebanon Association, "Tremendous deforestation of Lebanon" Report

[20] A. Bahety, Extension and Evaluation of ID3 – Decision Tree Algorithm, University of Maryland, College Park

[21] J.R, Quinlan, Induction of Decision Trees, Machine Learning 1: 81-106, 1986

[22] Evaluation: From Precision, Recall And F-Measure To Roc, Informedness, Markedness & Correlation Powers, D.M.W. Journal Of Machine Learning Technologies, Issn: 2229-3981 & Issn: 2229-399x, Volume 2, Issue 1, 2011, Pp-37-63

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Table 2: A Sample from our Case Study Showing the Categorization of Attributes with the Corresponding Actual Result

Air temperature [°C]		Relative humidity [%]		Wind speed [m/s]		Precipitation [mm]		Result
25,17	Moderate	62	Moderate	0,88	Low	0	Low	**Fire**
25,5	Moderate	62	Moderate	1,07	Moderate	0	Low	**No Fire**
26,08	Moderate	57	Moderate	1,13	Moderate	0	Low	**Fire**
25,23	Moderate	60	Moderate	1,18	Moderate	0	Low	**No Fire**
25,07	Moderate	52	Moderate	0,52	Low	0	Low	**No Fire**
25,02	Moderate	63	Moderate	1,03	Moderate	0	Low	**No Fire**
30,16	High	56	Moderate	1,32	Moderate	0	Low	**Fire**
10,74	Low	77	High	3,7	High	8,3	High	**No Fire**
11,12	Low	82	High	4,1	High	18,9	High	**No Fire**
11,47	Low	78	High	2,7	Moderate	0	Low	**No Fire**

Table 3: Summary of Observations of the Year 2012-North Lebanon Upon the Categories Low, Moderate & High

	T	Fire /T	No Fire/ T	RH	Fire/ RH	No Fire/ RH	WS	Fire/ WS	No Fire/ WS	P	Fire /P	No Fire/P
Low	271	3	268	38	8	30	158	11	147	288	25	263
Moderate	88	21	67	125	15	110	112	15	97	13	0	13
High	2	2	0	198	3	195	91	0	91	60	1	59

Journal of Theoretical and Applied Information Technology
20th May 2014. Vol. 63 No.2

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

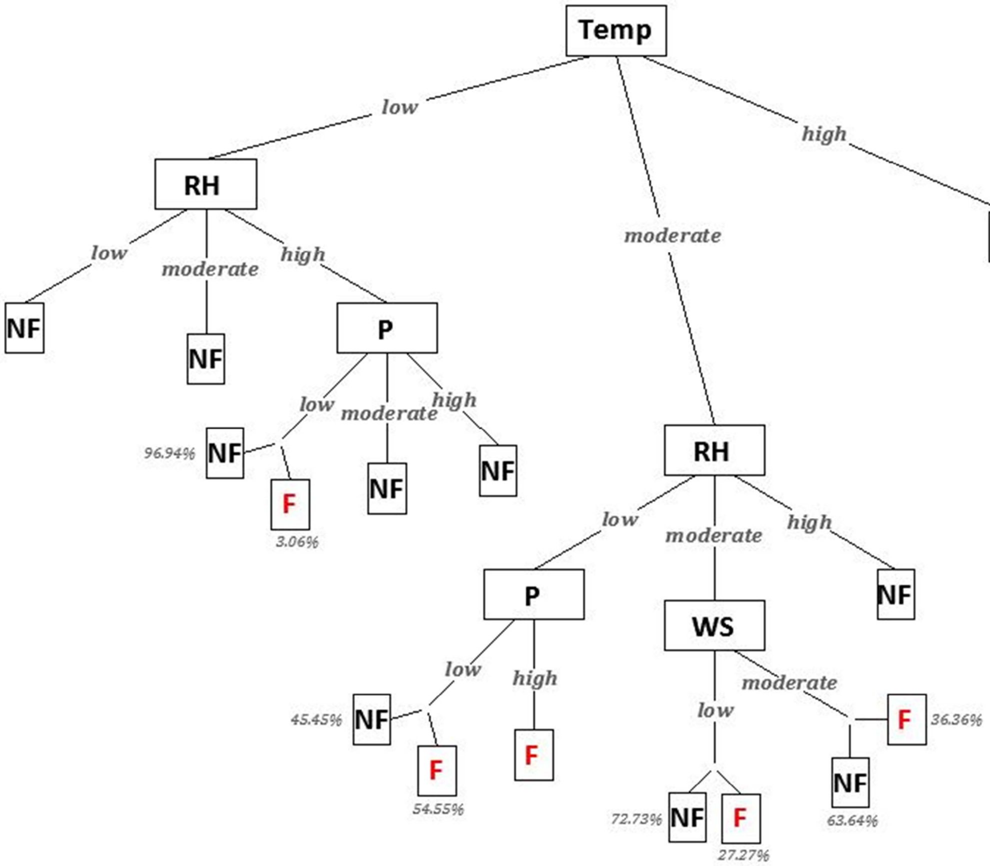

Figure 1: The Decision Tree of Forest Fire Prediction in Lebanon Based on Weather Data of the Year 2012

Table 4: Training and Testing Mean Square Errors (mse) For Feed Forward and Cascade Forward Networks Adopted Using Trainlm, Traingdx and Trainrp Training Algorithms

		feed forward function fitting network						**cascade network**	
		using four inputs			using two inputs			using four inputs	using two inputs
nb of neurons in hidden layer		5	10	15	5	10	15		
trainlm	nb of trials	**1000**	**1000**	**1000**	**1000**	**1000**	**1000**	**1000**	**1000**
	training mse	0,0514	0,0527	0,0503	0,0483	0,0483	0,0483	0,052	0,0483
	testing mse	0,0056	0,0093	0,0033	2,02E-04	2,07E-04	2,15E-04	0,0072	0,000198
traingdx	nb of trials	**1000**	**1000**	**210**	**203**	**1000**	**304**	**159**	**118**
	training mse	0,0571	0,0563	0,0563	0,0571	0,0563	0,0563	0,0619	0,0644
	testing mse	0,0128	0,0073	0,0069	0,0128	0,0073	0,0069	0,0111	0,0065
trainrp	nb of trials	**1000**	**1000**	**1000**	**1000**	**1000**	**1000**	**1000**	**1000**
	training mse	0,0483	0,0484	0,0484	0,0485	0,0486	0,0486	0,0477	0,048
	testing mse	0,0021	0,0024	0,0023	0,0014	0,0029	0,0015	2,15E-04	1,13E-04

Journal of Theoretical and Applied Information Technology
20th May 2014. Vol. 63 No.2

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Table 5: Measures of Precision, Specificity, Sensitivity, Accuracy and Area Under ROC Curve for the Various Neural Networks Used

			nb of trials	Precision	Specificity	Sensitivity	ACC	AUC
4-inputs feed forward	trainlm	5 neurons	1000	98,5%	69,2%	93,9%	92,8%	81,6%
		10 neurons	1000	98,1%	61,5%	93,5%	92,1%	77,5%
		15 neurons	1000	98,5%	71,4%	94,2%	93,1%	82,8%
	traingdx	5 neurons	1000	98,1%	61,5%	93,5%	92,1%	77,5%
		10 neurons	1000	98,1%	64,3%	93,9%	92,4%	79,1%
		15 neurons	210	98,1%	66,7%	94,2%	92,8%	80,4%
	trainrp	5 neurons	1000	98,9%	76,9%	94,2%	93,5%	85,6%
		10 neurons	1000	98,9%	76,9%	94,2%	93,5%	85,6%
		15 neurons	1000	98,9%	76,9%	94,2%	93,5%	85,6%
4-inputs cascade forward	trainlm		1000	98,5%	66,7%	93,5%	92,4%	80,1%
	traingdx		159	100,0%	-	91,1%	91,1%	-
	trainrp		1000	99,2%	86,7%	95,3%	94,8%	91,0%
2-inputs feed forward	trainlm	5 neurons	1000	99,2%	87,5%	95,6%	95,2%	91,6%
		10 neurons	1000	99,2%	87,5%	95,6%	95,2%	91,6%
		15 neurons	1000	99,2%	87,5%	95,6%	95,2%	91,6%
	traingdx	5 neurons	203	98,9%	70,0%	93,2%	92,4%	81,6%
		10 neurons	1000	98,1%	61,5%	93,5%	92,1%	77,5%
		15 neurons	304	98,1%	61,5%	93,5%	92,1%	77,5%
	trainrp	5 neurons	1000	98,5%	71,4%	94,2%	93,1%	82,8%
		10 neurons	1000	98,5%	71,4%	94,2%	93,1%	82,8%
		15 neurons	1000	98,5%	71,4%	94,2%	93,1%	82,8%
2-inputs cascade forward	trainlm		1000	99,2%	87,5%	95,6%	95,2%	91,6%
	traingdx		118	100,0%	-	91,1%	91,1%	-
	trainrp		1000	99,2%	87,5%	95,6%	95,2%	91,6%