APPLICATIONS OF RADIAL BASIS NEURAL NETWORKS FOR AREA FOREST

Angel Castellanos, Ana Martinez Blanco, Valentin Palencia

Abstract: This paper proposes a new method using radial basis neural networks in order to find the classification and the recognition of trees species for forest inventories. This method computes the wood volume using a set of data easily obtained. The results that are obtained improve the used classic and statistical models.

Keywords: Neural Networks, clustering, Radial Basis Functions, Forest Inventory.

ACM Classification Keywords: I.5. Pattern Recognition – I.5.1. Neural Nets; I.5.3. Clustering

Introduction

The research community has developed several different neural network models, such as backpropagation, radial basis function, growing cell structures [Fritzke 1994] and self-organizing feature maps [Kohonen 1989]. A common characteristic of the aforementioned models is that they distinguish between learning and a performance phase. Neural networks with radial basis functions have proven to be an excellent tool in approximation with few patterns. Most relevant research in theory, design and applications of radial basis function neural networks is due to Moody and Darken [Moody and Darken, 1989], Renals [Renals, 1989] and to Poggio and Girossi [Poggio and Girosi, 1990].

Radial basis function (RBF) neural networks (Figure 1) provide a powerful alternative to multilayer perceptron (MLP) neural networks to approximate or to classify a pattern set. RBFs differ from MLPs in that the overall input-output map is constructed from local contributions of Gaussian axons, require fewer training samples and train faster than MLP.

In this paper, we propose a method using radial basis neural networks in order to find the classification and the recognition of trees species for forest inventories. We use an unsupervised technique called the k-nearest neighbors rule to estimate centers and widths of the functions of radial base. The centers of the clusters give the centers of the RBFs and the distance between the clusters provides the width of the Gaussian function. Computation of the centers, used in the kernels function of the RBF neural network, is being the main focus to study in order to achieve more efficient algorithms in the learning process of the pattern set.

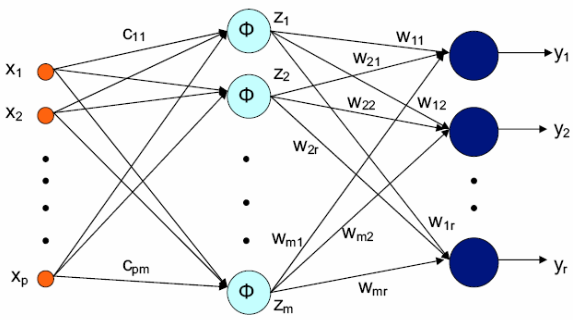


Figure 1.- Radial Basis Function Neural Network.

Problem Description

This paper seeks to estimate the wood volume for area forest inventory and find a classification of trees species, using a set of data that can be easily obtained such as: diameter, thickness bark, grow of diameter and height.

Volume parameter is one of the most important parameters in forest research when dealing with some forest inventories [Schreuder, H.T., Gregoire, T.G. and Word, G.B. 1993]. Usually, some trees are periodically cut in order to obtain such parameters using cubical proofs for each tree and for a given environment. This way, a repository is constructed to be able to compute the volume of wood for a given area and for a given tree specie in different environments. Stem volume formula is function of a tree's height, basal area, shape, etc. Volume is one of the most difficult parameters to measure, because an error in the measure or assumptions for any one of the above factors will propagate to the volume estimate. Volume is often measured for specific purposes, and interpretation of the volume estimate will depend on the units of measurement standards of use, and others specifications. Calculations of merchantable volume may also be based on true cubic volume. Direct and indirect methods for estimating volume are available [Hamilton, F. and Brack, C.L. 1999].

The method more usual to estimate volume in forest is the tree volume tables or tree volume equations. Huber's volume equations are a very common equation used to estimating volume:

$$V = h\pi\left(\frac{d}{2}\right)^2$$ V denotes volume, h denotes length, d denotes diameter.

Another form of previous equation is:

$$V = \eta h\pi\left(\frac{d}{2}\right)^2$$ η = factor for the merchantable volume

We present a study of the potential wood forest amount, that is, the maximum amount of wood that can be obtained. All data are taken from an inventory of the M-1019 area at "Elenco" in Madrid (Spain), at "Atazar" village. Most of the trees belongs to the Pinus Pinaster family and a small amount to the Pinus Sylvestris family. All this area is focused on the wood production. The area is divided into two different sub areas with a surface of 55.6 Ha and 46.7 Ha respectively.

The main aim is to be able to forecast the wood volume and detect relationships between all the variables that are in our study. Variables taken into account are: normalized diameter, total height, thickness bark, and radial growth in the last ten years. Normalized diameter has been measured in the whole feet of the two sub areas that made up the samples, provided they are larger than 12.5 cm till the last cluster of 60 cm.

In this work we compare solutions obtained with statistical regression analysis and classical methods as Huber's, with the results using radial basis function neural network.

Classifiers for the prediction in forest products

A radial basis function neural network has been implemented with four input neurons: diameter, thickness bark, grow of diameter and height, in order to estimate the volume of wood that can be used.

The net uses a competitive rule with full conscience in the hidden layer and one output layer with the than function, all the learning process has been performed with the momentum algorithm. Unsupervised learning stage is based on 100 epochs and the supervised learning control uses as maximum epoch 1000, threshold 0.01. We have performed an initial study using 260 patterns in training set; after a 90 patterns in training set and finally with only 50 patterns in training set, and the error MSE, are similar in three cases.

Problem under study is prediction of volume of wood, and it is compared to other methods such as the Huber's formula and the statistical regression analysis in order to estimate the amount of wood using typical tree variables: diameter, thickness and diameter growth. Neural networks had approximated in a good manner tested examples, getting a small mean squared error, see table below. Radial basis function neural network learns with only a few patterns, that is the way results using only 50 patterns are really excellent. For each of the tree species tested, the RBF gives less MSE estimated than the standard formulas Huber and Multivariate Analysis Regression.

	Error-Huber	Error-RBF	Error-Regression Multivariate
MSE	0.05	0.007	0.01

Next step consists on forecasting the input variable importance (sensitive analysis) in the learning process. Our neural network is a mapping $f(x_1, x_2, x_3, x_4) : \Re^4 \to \Re$ where $x_1 = diameter(cm)$, $x_2 = thickness\,bark(cm)$, $x_3 = growth\ of\ \ diameter(cm)$, $x_4 = height(cm)$, in order to forecast variable $x_5 = volume(dm^3)$. All centers are stable in two points that are those who signal the two main clusters, and that the net has been able to detect the two tree species.

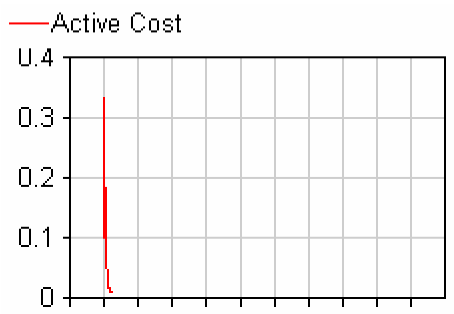
Several matrixes have been computed; where columns are input variables to forecast and rows are hidden neurons. These matrixes show the center values. Variable $x_3 = diameter\,growth$ takes the same value in both centers what it means that the study can be done without such variable obtaining similar values of MSE. Main centers of RBF approximate real clusters in the two forest areas, following table shows the real clustering.

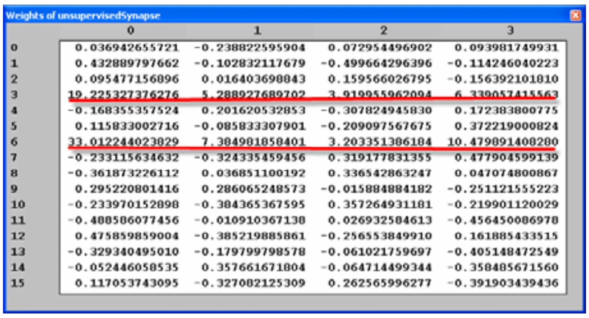
Zone - species	x_1	x_2	x_3	x_4
1	19,49	5,28	3,19	6,45
2	33,71	7,38	3,91	10,66

Previous table shows the matrix where the columns represent the input variable and the rows represent the hidden neurons. The hyperspace is divided into different regions or clusters starting from 16. Later, the number of clusters has been decreased till the minimum number of possible clusters is reached in order to solve the problem minimizing the mean squared error. The number of hidden neurons must be greater than the number of input variables to perform a correct learning.

Two main centers are found in the hyperspace, see following figures.

Four input variables and 16 clusters MSE=0.0079

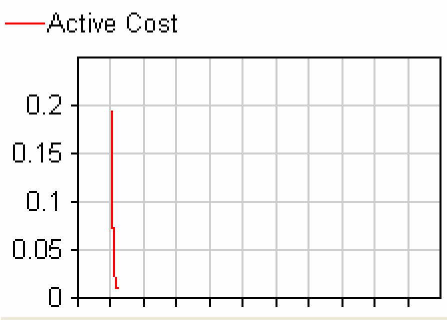


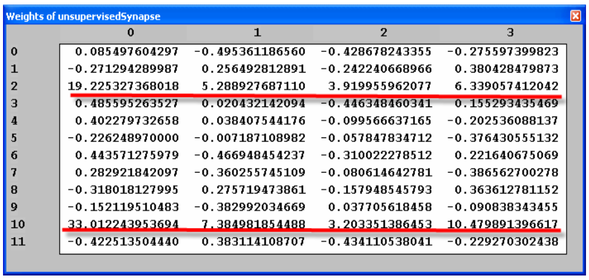


Weights of unsupervisedSynapse

	0	1	2	3
0	0.036942655721	-0.238822595904	0.072954496902	0.093981749931
1	0.432889797662	-0.102832117679	-0.499664296396	-0.114246040223
2	0.095477156896	0.016403698843	0.159566026795	-0.156392101810
3	19.225327376276	5.288927689702	3.919955962094	6.339057415563
4	-0.168355357524	0.201620532853	-0.307824945830	0.172383800775
5	0.115833002716	-0.085833307901	-0.209097567675	0.372219000824
6	33.012244023829	7.384981858401	3.203351386184	10.479891408280
7	-0.233115634632	-0.324335459456	0.319177831355	0.477904599139
8	-0.361873226112	0.036851100192	0.336542863247	0.047074800867
9	0.295220801416	0.286065248573	-0.015884884182	-0.251121555223
10	-0.233970152898	-0.384365367595	0.357264931181	-0.219901120029
11	-0.488586077456	-0.010910367138	0.026932584613	-0.456450086978
12	0.475859859004	-0.385219885861	-0.256553849910	0.161885433515
13	-0.329340495010	-0.179799798578	-0.061021759697	-0.405148472549
14	-0.052446058535	0.357661671804	-0.064714499344	-0.358485671560
15	0.117053743095	-0.327082125309	0.262565996277	-0.391903439436

Four input variables and 12 clusters MSE=0.0078

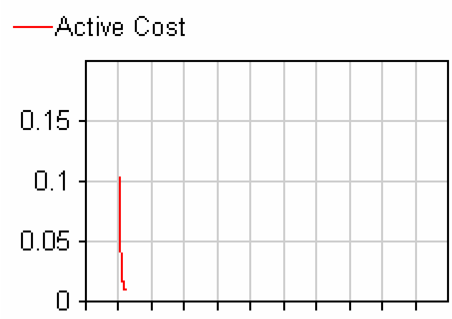




Weights of unsupervisedSynapse

	0	1	2	3
0	0.085497604297	-0.495361186560	-0.428678243355	-0.275597399823
1	-0.271294289987	0.256492812891	-0.242240668966	0.380428479873
2	19.225327368018	5.288927687110	3.919955962077	6.339057412042
3	0.485595263527	0.020432142094	-0.446348460341	0.155293435469
4	0.402279732658	0.038407544176	-0.099566637165	-0.202536088137
5	-0.226248970000	-0.007187108982	-0.057847834712	-0.376430555132
6	0.443571275979	-0.466948454237	-0.310022278512	0.221640675069
7	0.282921842097	-0.360255745109	-0.080614642781	-0.386562700278
8	-0.318018127995	0.275719473861	-0.157948545793	0.363612781152
9	-0.152119510483	-0.382992034669	0.037705618458	-0.090838343455
10	33.012243953694	7.384981854488	3.203351386453	10.479891396617
11	-0.422513504440	0.383114108707	-0.434110538041	-0.229270302438

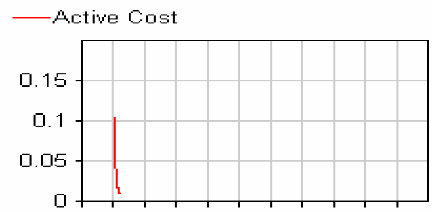
Four input variables and 8 clusters MSE=0.0075



Weights of unsupervisedSynapse

	0	1	2	3
0	0.139393292032	-0.039048432875	-0.491027558214	0.009964293344
1	0.022415845210	0.086809900204	0.144367809076	0.051072725608
2	0.098223822748	0.138782921842	-0.480559709464	0.225455488754
3	-0.001388592181	0.200430310984	-0.053514206366	0.407345805231
4	19.635513153942	5.354130101370	3.961877017807	6.443362612384
5	0.243858149968	-0.416776024659	-0.162892544328	0.434659871212
6	-0.461210974456	0.113635670034	0.494354075747	0.235343485824
7	33.531236499098	7.460325997065	3.093478779314	10.660911450110

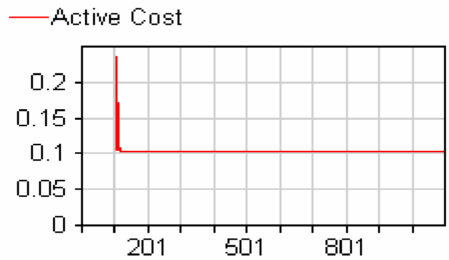
Four input variables and 5 clusters MSE= 0.0073



Weights of unsupervisedSynapse

	0	1	2	3
0	0.225424970244	0.038163396100	-0.008224738304	-0.242484817042
1	-0.326044495987	-0.496215704825	0.115955076754	-0.043595690786
2	-0.041459395123	0.219168675802	-0.375911740471	0.385494552446
3	19.635514146112	5.354130753430	3.961878389731	6.443361133609
4	33.531214807424	7.460317864873	3.093467055318	10.660923756069

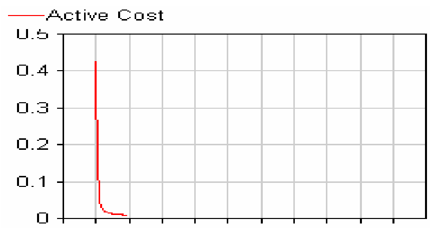
Four input variables and 4 clusters MSE=0.1



Weights of unsupervisedSynapse

	0	1	2	3
0	25.641225758815	6.267486735499	3.586643466294	8.269150221939
1	-0.469908749657	0.400753807184	-0.449949644459	0.216422009949
2	0.180684835353	-0.020706808679	0.264641254921	-0.400875881222
3	-0.177663502915	-0.422421948912	0.222281563768	-0.498290963469

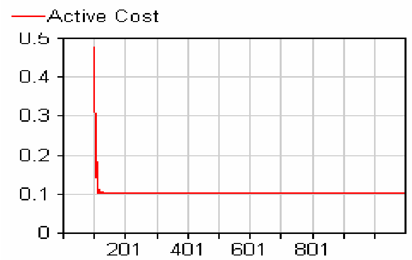
Three input variables and 4 clusters MSE=0.0078



Weights of unsupervisedSynapse

	0	1	2
0	-0.497070223090	-0.439146092105	0.293725394452
1	33.371678352723	7.454034626552	10.546546589181
2	19.492469146051	5.318156886818	6.451743076843
3	-0.314081240272	0.053331095309	0.129078035829

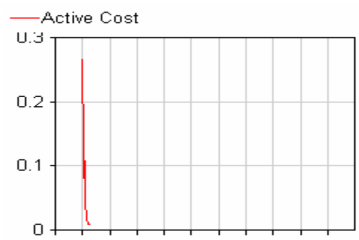
Three input variables and 3 clusters MSE=0.104



Weights of unsupervisedSynapse

	0	1	2
0	-0.451506088443	0.463896603290	0.436979277932
1	-0.466643269143	0.177205725272	0.006576738792
2	25.641225758815	6.267486735499	8.269150221939

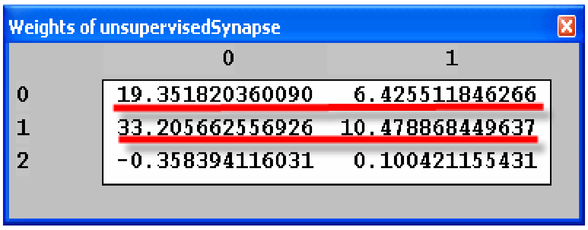
Two input variables and 4 MSE=0.0079



Weights of unsupervisedSynapse

	0	1
0	-0.219809564501	0.099536118656
1	19.351824656801	6.425514018775
2	33.205695811530	10.478871907419
3	-0.105517746513	-0.362300485244

Two input variables and 3 MSE=0.008



	0	1
0	19.351820360090	6.425511846266
1	33.205662556926	10.478868449637
2	-0.358394116031	0.100421155431

Conclusion

A radial basis function neural network has been trained with a few patterns in order to forecast the volume of wood in a given forest area. The network performs a clustering process of the trees using different input variables. A sensitive analysis can be computed observing the weight of unsupervised synapse. A previous clustering process of input data permits a better forecasting process in the output variable, in our case the amount of volume of wood in a forest area. These results improve commercial and classical forecasting methods in forest inventories, and proposed method can be applied to any tree specie or forest area without taking into account environment variables that appears in classical mathematical equations. As the number of classes that needs to be discriminated decreases, classifier accuracy increases; until obtain the real number of classes. Once the correct number of classes has been obtained using the RBF and with a supervised learning the volume of wood for a forest inventory can be estimated.

Bibliography

[Fritzke 1994] Fritzke B. Supervised learning wiht growing cell structures. Advances in neural information processing system 6, Pp. 25. Morgan Kaufmann Publishers 1994.

[Hamilton, F. and Brack, C.L. 1999] Hamilton, F. and Brack, C.L. 1999. Stand volume estimates for modelling inventory data. Australian Forestry 62(4): 360 - 367

[Kphonen 1989] Kohonen T. Self organization and associative memory. Springer-Verlag 1989

[Moody and Darken, 1989] Moody J., Darken C. Fast learning in networks of locally-tunned processing units. Neural computation, vil1, pp 281-294. 1989.

[Poggio and Girosi, 1990] Poggio T. and Girossi F. Networks for approximation and learning. IN procceding of the IEEE, volume 78, pages 1481-1497.

[Schreuder, H.T., Gregoire, T.G. and Word, G.B. 1993] Schreuder, H.T., Gregoire, T.G. and Wood, G.B. 1993. Sampling Methods for Multiresource Forest Inventory. John Wiley and Sons, Inc. New York.

Authors' Information

Angel Castellanos – Dpto. de Ciencias Basicas aplicadas, a la Ingeniería Forestal. Escuela Univ. de Ingeniería Técnica Forestal. Universidad Politécnica de Madrid, Avda. de Ramiro de Maeztu s/n 28040 Madrid, Spain. e-mail: angel.castellanos@upm.es

Ana Martinez Blanco – Dpto. de Ciencias Basicas aplicadas, a la Ingeniería Forestal. Escuela Univ. de Ingeniería Técnica Forestal. Universidad Politécnica de Madrid, Avda. de Ramiro de Maeztu s/n 28040 Madrid, Spain. e-mail: ana.martinez@upm.es

Valentín Palencia – Dpto. de Arquitectura y Tecnología de Sistemas Informáticos. Facultad de Informática. Universidad Politécnica de Madrid (Campus de Montegancedo) 28600 Boadilla del Monte, Madrid, Spain. e-mail: vpalencia@fi.upm.es