SPANDIDOS PUBLICATIONS

The Wisconsin breast cancer problem: Diagnosis and TTR/DFS time prognosis using probabilistic and generalised regression information classifiers

IOANNIS ANAGNOSTOPOULOS[1], CHRISTOS ANAGNOSTOPOULOS[2], DIMITRIOS VERGADOS[1], ANGELOS ROUSKAS[1] and GEORGE KORMENTZAS[1]

[1]Department of Information and Communication Systems Engineering, University of the Aegean, Karlovassi 83200, Samos; [2]Department of Cultural Technology and Communication, University of the Aegean, Mytiline 81100, Lesvos, Greece

Received September 6, 2005; Accepted September 27, 2005

Abstract. This study addresses the breast cancer diagnosis and prognosis problem by employing two neural network architectures with the Wisconsin diagnostic and prognostic breast cancer (WDBC/WPBC) datasets. A probabilistic approach is dedicated to solve the diagnosis problem, detecting malignancy among cases (instances) as derived from fine needle aspirate (FNA) tests, while the second architecture estimates the time interval that possibly contains the right endpoint of disease-free survival (DFS) of the patient. The accuracy of the neural classifiers reaches nearly 98% for the diagnosis and 93% for the prognosis problem, while the prognostic recurrence predictions were evaluated using survival analysis through the Kaplan-Meier approximation method. Both architectures were compared with other similar approaches. The robustness and real-time response of the proposed classifiers were further tested over the web as a potential integrated web-based decision support system.

Introduction

According to the American National Cancer Institute, it is estimated that 13.4% of women born today will be diagnosed with breast cancer at some time in their lives (1,2). For the diagnosis of breast cancer cases and prognosis of disease, many techniques have been proposed. Surgical biopsy can confirm malignancy with high-level sensitivity, but is considered a costly operation and has a negative impact on the psychology of the patient. Towards these considerations, machine learning techniques aim to provide the same level of accuracy, without the negative aspects of surgical biopsy (3-11).

This study addresses the breast cancer diagnosis and prognosis problem using the Wisconsin diagnostic breast cancer (WDBC) and Wisconsin prognostic breast cancer (WPBC) datasets, which are publicly available via an anonymous ftp (12). These datasets involve measurements taken with the fine needle aspirate (FNA) test. The role of diagnosis is to provide a distinction between malignant and benign breast masses. If a patient is diagnosed with breast cancer, the malignant mass must be excised. After the operation, the expected course of disease must be determined. However, prognostic prediction does not belong to the classic learning paradigms of function approximation or classification. This is because a patient can be classified as a 'recurrence' case (instance) if the disease is observed, but there is no threshold point at which the patient can be considered a 'non-recurrence' case. The data are therefore censored since the time-to-recurrence for only a selected subset of patients is known. For other patients, the length of time after treatment during which malignant masses are not found is known. This time interval is the disease-free survival (DFS) time, which can be reported for an individual patient or a study population. In particular, the right endpoints of the recurrence time intervals are censored, as some patients will inevitably change hospital or doctors or die of causes unrelated to the cancer. Therefore, the training dataset for the learning phase is not well-defined. Several groups have approached prognosis as a separate problem using different architectures, such as back-propagation artificial neural networks (13), entropy maximization networks (14,15) decision trees (16) and fuzzy-based measurements (17).

Materials and methods

The WDBC and WPBC datasets are the result of efforts made at the University of Wisconsin Hospital for the diagnosis and prognosis of breast tumours solely based on the FNA test. This test involves fluid extraction from a breast mass using a small-gauge needle, then a visual inspection of the fluid under a microscope.

Correspondence to: Dr Ioannis Anagnostopoulos, Department of Information and Communication Systems Engineering, University of the Aegean, Karlovassi 83200, Samos, Greece
E-mail: janag@aegean.gr

Key words: Wisconsin breast cancer problem, diagnosis, prognosis, information classifiers

Table I. WDBC/WPBC cell nuclei characteristics/attributes.

1. Radius (mean of distances from centre to points on the perimeter)
2. Texture (standard deviation of grey-scale values)
3. Perimeter
4. Area
5. Smoothness (local variation in radius lengths)
6. Compactness [(perimeter)^2/area) - 1]
7. Concavity (severity of concave portions of the contour)
8. Concave points (number of concave portions of the contour)
9. Symmetry
10. Fractal dimension ('coastline approximation' - 1)

This paper proposes two neural network architectures to address the breast cancer detection/prognosis problems. The first is a probabilistic classifier that can detect malignancy, while the second architecture consists of a probabilistic neural network that employs a generalised regression algorithm, estimating the recurrence time (TTR; time-to-recurrence) and the period in which the patient exceeds her disease-free survival (DFS) time. The prognosis of the specific time interval is considered difficult since the training data are right censored (13,14,18,19).

The Wisconsin breast cancer datasets. The Wisconsin diagnostic breast cancer (WDBC) dataset consists of 569 instances (357 benign and 212 malignant), where each represents FNA test measurements for one diagnosis instance. For this dataset, each instance has 32 attributes, with the first 2 attributes corresponding to a unique identification number and diagnosis status (benign/malignant). The remaining 30 features are computations for 10 real-valued features, along with their mean, standard error and the mean of the three largest values ('worst' value) for each cell nucleus, respectively. These 10 real values, which are depicted in Table I, are determined from a digitised image of a fine needle aspirate (FNA) from the breast tumour, describe characteristics of the cell nuclei present in the image and are recorded with four significant digits.

The Wisconsin prognostic breast cancer (WPBC) dataset consists of 198 instances (151 non-recurrences and 47 recurrences), where each represents follow-up data for one breast cancer case. These were consecutive in-patients at the University of Wisconsin Hospital from the period of 1984 to 1995 and include only those cases exhibiting invasive breast cancer and no evidence of distant metastases at the time of diagnosis. Each instance has 35 attributes, with the first 3 attributes corresponding to a unique identification number, prognosis status (recurrence/non-recurrence) and DFS time or time-to-recurrence (TTR) in months, respectively. They follow the above-mentioned 30 features, and the last 2 attributes are the diameter of the excised tumour (in cm) and number of positive axillary lymph nodes observed at the time of surgery.

Both WDBC and WPBC datasets were used in several published studies in medical literature (10,16,20-23). In addition, due to their consistency and robust creation, these datasets have also been used to measure the classification or predict the performance of information systems in other scientific areas (24,25).

The proposed neural network architectures. For the diagnosis problem, the proposed neural network belongs to the probabilistic type (probabilistic neural network; PNN), since this kind of network presents a high-generalisation ability and does not require a large amount of training data (26,27). The PNN decides whether the input corresponds to a benign or malignant case.

Topology of the diagnosis neural network (DiagNN) was 31-568-2. The input layer consists of 31 nodes, which correspond to diagnosis status, followed by the 30 calculated values (mean, standard error and 'worst' value) from the digitised image of each instance. The second layer is the middle/pattern layer, which organises the training set so that an individual processing element represents each normalised input vector. Therefore, it consists of 568 nodes, which correspond to the total amount of WDBC instances except for one that is used to test each training epoch according to the jacknife test method. Finally, the network has an output layer consisting of 2 nodes, representing the decision of malignancy or not.

For the prognosis problem, the neural network belongs to the generalised regression type (Generalised Regression Neural Network, GRNN). These neural networks have the special ability to deal with sparse and non-stationary data where non-linear relationships exist among inputs and outputs. The role of the neural network is to calculate a time interval that corresponds to a possible right endpoint of patient DFS time.

Topology of the prognosis neural network (ProgNN) was 14-193-z-z. The input layer consists of 14 nodes, where each node corresponds to the prognosis status, TTR or DFS time, 10 cell nuclei characteristics/attributes of Table I, diameter of the excised tumour and number of positive axillary lymph nodes observed at the time of surgery. Due to the small amount of WPBC instances and to avoid the 'curse of dimensionality' problem, the standard error and 'worst' values from the 10 real-valued features were removed during the training phase. Four instances were not included in the training/testing set since their lymph node values were missing. Thus, the second layer consists of 193 nodes, which correspond to the total amount of patterns for the training epoch according to the leave-one-out method. Finally, the summation/division layer consists of z nodes that feed the same amount of processing elements in the output layer, and represent the classified time intervals that correspond to a possible right endpoint of DFS time.

Results

The training set of the DiagNN consists of the WDBC dataset, and its role is to classify an instance as benign or malignant. The quality of prediction was examined using the jacknife test in which each instance was singled out as a test instance, with the remaining instances being used to train the neural network.

lean time needed for the completion of one training a Pentium V at 3 GHz with 1024 MB RAM was 14.4 sec, while for a T1 connection at 1.544 MBps with the same computational power, the respective training time needed from a remote client was 23.1 sec. Equations 1 and 2, outline the Akaike's information criterion (AIC) as well as the Rissanen's minimum description length (MDL) during the training period. Values $|d_{ij} - y_{ij}|$ correspond to the distances among the desired and actual network output for the i^{th} testing instance (exemplar) residing at the j^{th} processing element of the middle layer.

$$AIC(k) = N * ln(MSE) + 2 * K \quad (1)$$

$$MDL(k) = N * ln(MSE) + 0.5 * K * ln(N) \quad (2)$$

where $$MSE = \frac{\sum_{j=1}^{P}\sum_{i=1}^{N}(d_{ij} - y_{ij})^2}{N \cdot P}$$

In the above criteria, P is equal to the number of output processing elements (malignant/positive-benign/negative; P=2), while N and K define the amount of exemplars in the training set and number of network weights, respectively (N=568 and K=18744). AIC measures the trade-off between training performance and network size, while MDL combines the error of the model with the number of degrees of freedom for determining the level of generalisation. The aforementioned indicators were used to fine-tune the mean and variance biases of the respective local approximators and produce a confusion matrix with the best possible values in the diagonal cells. The condition that must be fulfilled is the stabilisation of AIC and MDL over the β coefficient, where $\beta=1/2\sigma^2$ and σ represents the standard deviation of the Parzen estimator expressed by:

$$\frac{1}{n\sigma}\sum_{i=0}^{n-1} W\left(\frac{x - x_i}{\sigma}\right)$$

This estimator is used as a Bayesian classifier at the middle layer, where n is the total sample size, χ and χ_i are the input and sample points, and W is the weighting function.

Table II. Testing phase confusion matrix (Diag NN, 2 classes).

DiagNN		Tested	
		Positive	**Negative**
Actual	**Malignant**	209 (TP)	3 (FN)
	Benign	7 (FP)	348 (TN)
Sensitivity $(100 * \frac{TP}{TP + FN})\% = 98.58\%$		**Specificity** $(100 * \frac{TN}{TN + FP})\% = 98.03\%$	
Predictive value (positive) $(100 * \frac{TP}{TP + FP})\% = 97.76\%$		**Predictive value (negative)** $(100 * \frac{TN}{TN + FN})\% = 99.14\%$	
Efficiency $(100 * \frac{TN + TP}{TP + TN + FP + FN})\% = 97.89\%$			

It was evaluated that the neural network was not properly trained when $0.1<\beta<1$, due to large mean square error (MSE) value and therefore large AIC and MDL values. Conversely, MSE was significantly decreased when $\beta=1$, revealing acceptable percentage values in the diagonal cells of the training confusion matrix. However, further investigation of the influence of β in the learning ability of the classifier exposed that MSE and AIC/MDL values were increased when $1<\beta<3$, then further reduced when $3<\beta<100$. Finally, for $\beta\geq100$, the AIC/MDL criteria were stabilised (AIC =19615 and MDL =41621), satisfying in parallel the optimisation condition $\delta MDL/\delta\beta = 0$. Therefore, during the training period, the 'beta' coefficient for all local approximators of the middle layer was set equal to 100 (β=100), while the mean and variance values of the randomised biases were equal to 0 and 0.5. These fine-tuning measurements are depicted in Fig. 1.

Table II presents the true positive (TP), false positive (FP), true negative (TN) and false negative (FN) results. Thus, having an *a priori* known set of 357 benign and 212

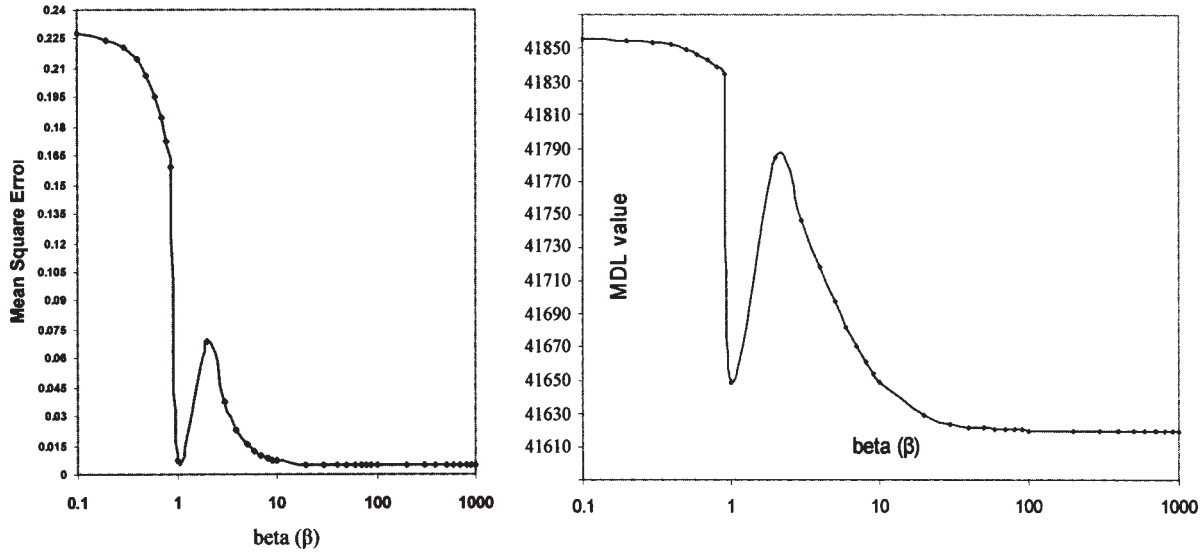

Figure 1. MSE/AIC/MDL over the beta (β) coefficient (DiagNN).

malignant instances, the neural network successfully identified 348 instances as negative and 209 instances as positive. According to these observations, Table II presents the sensitivity, specificity and efficiency of the DiagNN, as well as the predicted values of a positive/negative test result. Sensitivity, specificity and efficiency outline the percentage of patients and healthy subjects recognised by the DiagNN and the percentage of patients and healthy subjects correctly classified on the basis of test results. Efficiency also corresponds to the average precision or overall performance of the system. Finally, the predicted values of positive and negative test results define the percentage of patients among instances that were correctly classified as positive (malignant) and negative (benign) on the basis of test results. The time-to-decision by the system for each test instance was measured at 2.3 sec (mean value) with an additional delay of nearly 0.5 sec for the web-based system.

The training set of the ProgNN consists of WPBC data divided into four classes, namely C_1, C_2, C_3 and C_4, and topology of the neural network was 14-193-4-4. This categorisation was made according to the value of the third attribute, which indicates the TTR or DFS time. Thus, C_1 corresponds to the instances where DFS/TTR ≤1 year (20 recurrence/23 non-recurrence), while C_1, C_1 and C_1 correspond to time intervals where 1< DFS/TTR ≤3 years (14 recurrence/34 non-recurrence), 3< DFS/TTR ≤6 years (7 recurrence/48 non-recurrence) and DFS/TTR >6 years (5 recurrence/43 non-recurrence), respectively. The required training time for the local classifier was 53 sec, and that for the client-server architecture was 68 sec.

The total WPBC instances were presented to the network in a round-robin manner and leave-one-out method, and training ended before the average testing error on the left-out cases began to increase. The total prediction accuracy or efficiency (TP) and the prediction accuracy or precision (P) for each location were calculated to assess the prediction system according to equations 3 and 4:

$$TP = \frac{\sum_{k=1}^{4} p_k}{N} \quad (3)$$

$$P = \frac{p_k}{n_k} \quad (4)$$

In the above equations, N is the total number of sequences (NR+NN), k is the respective class, n_k is the number of instances in class k and p_k is the number of correctly predicted instances in class k.

The accuracy of prediction using leave-one-out tests for the WPBC instances and categorized time intervals are depicted in Table III, which corresponds to the testing confusion matrix for the addressed problem. The confusion matrix is defined by labelling rows with the desired classification and columns with the predicted classifications, and displaying values where each corresponds to the percentage of effect that a particular input has on a particular output. The diagonal cells correspond to the correctly classified instances for each class respectively, while the other cells show the misclassified instances. Every row expresses the ability of the system to produce the correct classification over a tested time interval. The precision rates for the four time intervals were 90.69%, 91.67%, 92.72% and 93.75%, while the respective recall rates were 92.86%, 91.67%, 94.44% and 90.00% (overall performance = 92.3%). The terms precision and recall correspond to the sensitivity and predictive value of a tested class, respectively. Finally, the mean times needed for the decision over a tested instance was slightly higher compared with the DiagNN classifier, measured at 3.8 sec (locally) and 4.4 sec (web-based architecture).

Table III. Testing confusion matrix (Prog NN, 4 classes).

	Predicted				
ProgNN	C1	C2	C3	C4	Precision
Actual					
C1	39	2	1	1	90.69%
C2	1	44	1	2	91.67%
C3	1	1	51	2	92.72%
C4	1	1	1	45	93.75%
Recall	92.86%	91.67%	94.44%	90.00%	
Overall performance = 92.27%					

C1, DFS/TTR ≤1 year; C2, 1< DFS/TTR ≤3 years; C3, 3< DFS/TTR ≤6 years; C4, DFS/TTR>6 years.

Discussion

We compared the results derived from the proposed study with similar techniques that aim to solve the Wisconsin breast cancer problem. These techniques come from neural networks (28-30), support vector machines-decision trees (31), fuzzy logic-based approaches (32,33) and Ant Colony Optimization algorithms (34). Among them, the DiagNN presented the second best efficiency rate over the WDBC dataset (97.9% efficiency). A hybrid approach of neural network with fuzzy logic-based rules (Feature Space Mapping) presented a slightly better value (efficiency, 98.3%) (30-33). The third best performance (97.2%) was achieved by a Support Vector Machine approach, where the authors used a 5-fold cross validation over the WDBC dataset (31). Table IV summarizes the performance rates of the current and other proposed methods addressing the Wisconsin diagnosis breast cancer problem. Generally, neural network-based methods presented better efficiency values in comparison to other approaches, especially when these approaches were used as stand-alone classification systems (e.g. decision trees, fuzzy logic-based rules and Bayesian classification techniques). Ant Colony Optimization algorithms (Ant_Miner1 and Ant_Miner2) presented significantly lower efficiency rates (33), while the Quadratic Discriminant Analysis presented very poor efficiency (29).

Concerning the prognosis problem, the recurrence predictions made by the ProgNN were further examined using survival analysis. The cases were divided into four aforementioned time intervals according to the predicted DFS time. The actual recurrence probabilities of these four

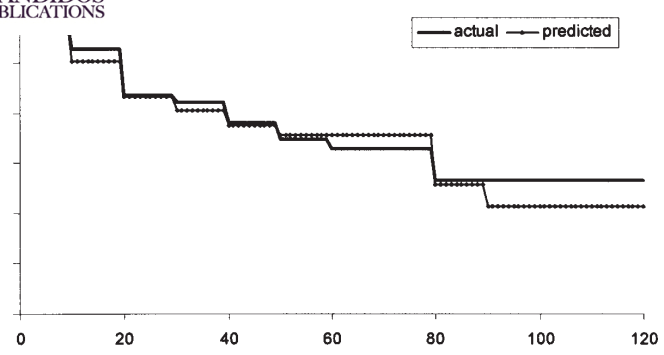

Figure 2. Survival analysis of the predicted DFS curve compared to the actual DFS curve. y-axis, DFS time probability; x-axis, time in months. Dataset: WPBC; predicted DFS ≤1 year, 1< predicted DFS ≤3 years, 3< predicted DFS ≤6 years, and predicted DFS >6 years.

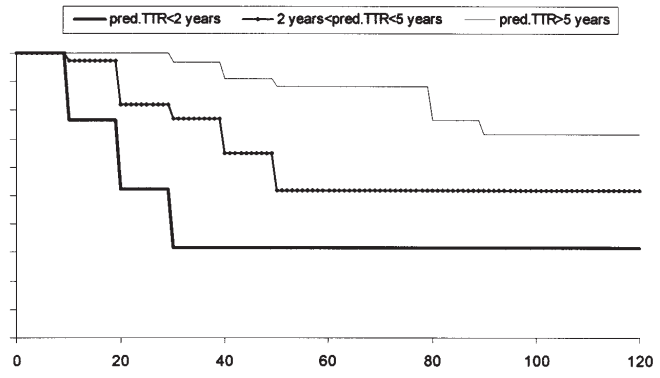

Figure 3. Kaplan-Meier DFS time probabilities based on the predicted TTR. y-axis, DFS time probability; x-axis, time in months. Dataset: WPBC; predicted TTR ≤2 years, 2 years <predicted TTR ≤5 years, and predicted TTR ≥5 years.

Table IV. Overall performance rates of several Wisconsin breast cancer problem techniques.

Diagnosis method	Efficiency (%)	Reference
Feature space mapping	98.3	(29,32)
DiagNN (minimize ß over AIC/MDL)	97.9	This study
Support vector machine (5xCV)	97.2	(30)
3-NN standard Manhattan	97.1	(27)
kNN with DVDM distance	97.1	(27)
21-NN standard Euclidean	96.9	(27)
Fisher's linear discriminant analysis	96.8	(28)
Multi-layer Perceptron/back propagation	96.7	(28)
LVQ	96.6	(28)
kNN, Euclidean/Manhattan	96.6	(28)
NB - naïve Bayes	96.4	(28)
C4.5 (decision tree)	96.0	(31)
Linear discriminant analysis	96.0	(28)
OC1 DT (5xCV)	95.9	(30)
GTO DT (5xCV)	95.7	(30)
Assistant I tree (ASI)	95.6	(28)
Rule induction over approximate classification	95.0	(31)
Assistant R tree (ASR)	94.7	(28)
Lookahead feature construction binary tree	94.4	(28)
Ant_Miner3	94.3[a]	(33)
C4.5 (5xCV)	93.4	(30)
Ant_Miner1	92.6[a]	(33)
Quadratic discriminant analysis (QDA)	34.5	(28)

[a]Mean value.

time intervals were then assembled using the Kaplan-Meier approximation. Fig. 2 presents the survival analysis of the predicted DFS curve produced by the ProgNN compared to the DFS curve of actual WPBC dataset instances. The y-axis corresponds to the probability of DFS time, and the x-axis corresponds to the time in months.

The two curves are almost identical for the time intervals of 0-11, 20-28, 40-58 and 80-88 months. The remaining predictions were similar except for the predictions made for the period of >90 months. Thus, despite the fact that the ProgNN had better precision level values for the classes that correspond to longer DFS periods (Table III), the survival analysis of the predicted results showed no significant statistical differences for shorter DFS times. This is caused by the 'unfair' division of the learning set over the categorized intervals and the fact that the range of the predicted interval elongates for longer DFS periods (classes C1, C2, C3 and C4 correspond to 12, 24, 36 and nearly 50 months, respectively).

Table V. Testing confusion matrix (Prog NN - 3 classes).

	Predicted			
ProgNN	D1	D2	D3	Precision
Actual				
D1	63	2	4	91.30%
D2	2	55	2	93.22%
D3	1	3	66	94.29%
Recall	95.45%	91.67%	91.67%	

Overall performance = 92.92%

D1: DFS/TTR ≤2 years; D2, 2< DFS/TTR ≤5 years; D3, DFS/TTR >5 years.

In a similar approach, the predicted probability of DFS values presented higher levels compared to the actual data, without important statistical differences (35). However, this is undesirable in problems with prognosis since the end of DFS time may correspond to a possible recurrence of disease. This means that the predicted DFS time curves derived from the ProgNN more adequately addresses the Wisconsin breast cancer prognosis, since the predicted probabilities for the right endpoint of patient DFS time is lower than the actual ones in most time intervals.

On the other hand, if the predicted right endpoints are considered possible TTR points, then the curves depicted in Fig. 3 present the Kaplan-Meier DFS time probabilities based on the predicted TTR. However, in this case, the ProgNN classified the predicted instances into three time intervals of predicted TTR ≤2 years (29 recurrence/40 non-recurrence), 2 years < predicted TTR ≤5 years (13 recurrence/46 non-recurrence) and TTR ≥5 years (5 recurrence/65 non-recurrence), and thus its topology was 14-193-3-3. Table V shows the respective testing confusion matrix (classes D1, D2 and D3). The precision rates for the three classes were 91.30%, 93.22% and 94.29%, while the respective recall rates were 95.45%, 91.67% and 91.67%. The overall performance reached a level of 93%. All training and testing methods were similar to the methods described in Results. The selection of these intervals were made to compare the ProgNN results with a previous study where the recurrence rate was nearly 30% at 2 years and 10% recurrence up to 5 years (36). After the training phase, the respective recurrence rates outlined by the ProgNN (Fig. 3) shows a significant improvement when the prediction was <2 years (the recurrence rate reached 50%), while instances with a predicted TTR of >5 years presented similar recurrence rates compared to those of a previous study (36). For this topology, the time needed for the training phase of the ProgNN was 46 sec for the local classifier and 58 sec for the web-based architecture. The mean time required for the decision response was 3.2 and 3.7 sec, respectively.

References

1. http://seer.cancer.gov/cgi-bin/csr/1975_2001/search.pl#results, Estimated New Cancer Cases and Deaths for 2004.
2. U.S. National Institutes of Health, National Cancer Institute. http://cancernet.nci.nih.gov/
3. Wang TC and Karayiannis NB: Detection of microcalcifications in digital mammograms using wavelets. IEEE T Med Imaging 17: 498-509, 1998.
4. Huo Z, Giger M, Vyborny C, Wolverton D, Schmidt R and Doi K: Automated computerized classification of malignant and benign mass lesions on digital mammograms. Acad Radiol 5: 155-168, 1998.
5. Cheng HD, Lui YM and Freimanis RI: A novel approach to microcalcification detection using fuzzy logic technique. IEEE T Med Imaging 17: 442-450, 1998.
6. Pendharkar PC, Rodger JA, Yaverbaum GJ, Herman N and Benner M: Association, statistical, mathematical and neural approaches for mining breast cancer patterns. Expert Syst Appl 17: 223-232, 1999.
7. Setiono R: Generating concise and accurate classification rules for breast cancer diagnosis. Artif Intell Med 18: 205-219, 2000.
8. Chen D, Chang RF and Huang YL: Breast cancer diagnosis using self-organizing map for sonography. Ultrasound Med Biol 26: 405-411, 2000.
9. Giger M, Huo Z, Kupinski M and Vyborny C: Computer-aided diagnosis in mammography. In: Handbook of Medical Imaging. Sonka M and Fitzpatrick J (eds). SPIE Press, pp917-986, 2000.
10. Tourassi GD, Markey MK, Lo JY and Floyd CE Jr: A neural network approach to breast cancer diagnosis as a constraint satisfaction problem. Med Phys 28: 804-811, 2001.
11. Maglogiannis I, Pavlopoulos S and Koutsouris D: An integrated computer supported acquisition, handling and characterization system for pigmented skin lesions in dermatological images. IEEE Trans Info Tech Biomed 9: 86-98, 2005.
12. http://ftp.ics.uci.edu/pub/machine-learning-databases/breast-cancer-wisconsin/, Wisconsin Diagnostic Breast Cancer (WDBC) Dataset and Wisconsin Prognostic Breast Cancer (WPBC) Dataset.
13. Burke HB and Goodman PH: Artificial neural networks improve the accuracy of cancer survival prediction. Cancer J 79: 857-862, 1997.
14. Choong PL and deSilva CJS: Entropy maximization networks, An application to breast cancer prognosis. IEEE T Neural Network 7: 568-577, 1996.
15. Choong PL and deSilva CJS: Maximum entropy estimation vs. multivariate logistic regression: which should be used for the analysis of small binary outcome data sets? Proc 20th Ann Int Conf IEEE Eng Med Biol Soc 3: 1602-1605, 1998.
16. Wolberg WH, Street WN, Heisey DM and Mangasarian OL: Computer-derived nuclear features distinguish malignant from benign breast cytology. Hum Pathol 26: 792-796, 1995.
17. Seker H, Odetayo M, Petrovic D, Naguib RNG, Bartoli C, Alasio L, Lakshmi MS and Sherbet GV: A fuzzy measurement-based assessment of breast cancer prognostic markers. Proc 2000 IEEE EMBS International Conference on Information Technology Applications in Biomedicine. pp174-178, 2000.
18. Mangasarian OI, Street WN and Wolberg WH: Breast cancer diagnosis and prognosis via linear programming. Oper Res 43: 570-577, 1995.
19. Street WN: A neural network model for prognostic prediction. Proc 15th International Conference on Machine Learning. Madison, WI, Morgan Kaufmann, pp540-546, 1998.
20. Wolberg WH, Street WN and Mangasarian OL: Machine learning techniques to diagnose breast cancer from fine-needle aspirates. Cancer Lett 77: 163-171, 1994.
21. Wolberg WH, Street WN and Mangasarian OL: Image analysis and machine learning applied to breast cancer diagnosis and prognosis, Anal Quant Cytol 17: 77-87, 1995.
22. Jiang Y, Nishikawa R, Wolverton D, Metz C, Giger ML, Schmidt R and Doi K: Automated feature analysis and classification of malignant and benign microcalcifications. Radiology 198: 671-678, 1996.
23. Taylor P, Fox J and Todd-Pokropek A: Evaluation of a decision aid for the classification of microcalcifications. Digital Mammography. Kluwer Academic Publishers, pp237-244, 1998.
24. Hoya T and Chambers JA: Heuristic pattern correction scheme using adaptively trained generalized regression neural networks. IEEE T Neural Network 12: 91-100, 2001.
25. Kaban A and Girolami M: Initialized and guided EM-clustering of sparse binary data with application to text-based documents. Proc 15th Int Conf Pattern Recog 2: 744-747, 2000.
26. Specht DF: Probabilistic neural networks. Neural Networks 3: 109-118, 1990.

SPANDIDOS PUBLICATIONS
s T: Advanced Algorithms for Neural Networks. John New York, NY, 1995.
28. Wojnarski M: LTF-C: architecture, training algorithm and applications of new neural classifier. Fundamenta Informaticae 54: 89-105, 2003.
29. Ster B and Dobnikar A: Neural networks in medical diagnosis: comparison with other methods. Proceedings of the International Conference EANN '96. Bulsari A (ed). pp427-430, 1996.
30. Duch W, Adamczak R and Jankowski N: New developments in the feature space mapping model. Proc 3rd Conference on Neural Networks and Their Applications, Kule, Poland, pp65-70, 1997.
31. Bennett KP and Blue J: A Support Vector Machine Approach to Decision Trees. RPI Math Report no. 97-100. Rensselaer Polytechnic Institute, Troy, NY, 1997.
32. Hamilton HJ, Shan N and Cercone N: RIAC: a rule induction algorithm based on approximate classification. Technology Report CS 96-06, Regina University, 1996.
33. Duch W, Adamczak R and Grabczewski K: A new methodology of extraction, optimization and application of crisp and fuzzy logical rules. IEEE T Neural Network 11: 1-31, 2000.
34. Liu B, Abbass HA and McKay B: Classification rule discovery with ant colony optimization. IEEE Comp Intelligence Bull 3: 31-35, 2004.
35. Street WN: A neural network model for prognostic prediction. Proc 15th International Conference on Machine Learning. Morgan Kaufmann Publishers Inc., San Francisco, CA, pp540-546, 1998.
36. Street WN, Mangasarian OL and Wolberg WH: An inductive learning approach to prognostic prediction. Proc 12th International Conference on Machine Learning. Prieditis A and Russell S (eds). Morgan Kaufmann Publishers Inc., San Francisco, CA, pp522-530, 1995.