THE USAGE OF NEURAL NETWORKS FOR THE MEDICAL DIAGNOSIS

Kateryna Malyshevska

Abstract: *The problem of cancer diagnosis from multi-channel images using the neural networks is investigated. The goal of this work is to classify the different tissue types which are used to determine the cancer risk. The radial basis function networks and backpropagation neural networks are used for classification. The results of experiments are presented.*

Keywords: neural networks, backpropagation, RBF, uterine cervix, cancer, classification.

ACM Classification Keywords: I.5.1 Pattern Recognition - Neural nets

Conference: The paper is selected from XV^th International Conference "Knowledge-Dialogue-Solution" KDS 2009, Varna, Bulgaria, June-July 2009

Introduction

The possibility of uterine cervix cancer diagnosis is considered in this work. Such problem appeared from the necessity of early diagnostic of disease using the computer system, which would help the doctor to define the tissues with the high risk of the cancer transformed tissue appearance. The system is based on the assumption that the optical characteristics of healthy cells and diseased cells differ and this difference is more significant then variations in such characteristics among cells that belong to different people. At medical university of Arizona (USA) the optical system which in addition to the usual colposcopy testing provides multichannel images of uterine cervix tissue was introduced. Multichannel images of 108 patients were produced. Simultaneously, the same patients were examined by a doctor. The examination consisted of biopsy of certain tissue areas which were sent to a pathologist for an analysis. Biopsy areas were defined on the image and compared with its results. This information was used to develop the risk areas recognition algorithm based on comparison of a multichannel picture, biopsy results, and doctor's diagnosis. Based on obtained information, it is possible to create the system that allows to perform early diagnostics [Schoonmaker J. 2007].

Statement of the problem

According to the medical statistics, the uterine cervix cancer takes the fourth place among women oncological diseases, (after the stomach, skin, and breast cancer). Primary colposcopy examination of the patient defines the necessity of making biopsy and the further consultation of the oncologist [Воробьева Л. И. 2008]. The computer system considered in this work can be used by a doctor for preliminary diagnostics of a cancer by determining the presence of certain tissue types without conducting a biopsy. The European Expert Group developing the European generalized training program, has offered the following classification of UC epithelium transformation: SEA (squamous epithelial abnormalities) - benign changes of the flat epithelium, columnar without the changes, allowing to assume CIN, Squamous cell changes - changes flat epithelium without accurate signs of a tumor, CIN-I - the least risky type, represents only mild dysplasia, or abnormal cell growth, CIN-II - moderate dysplasia confined to the basal 2/3 of the epithelium, CIN-III - severe dysplasia that spans more than 2/3 of the epithelium, and may involve the full thickness (CIS - carcinoma in situ).[Koss L.G. 1989] The purpose of this work is the development of computer system which can correctly classify different tissue types (SEA, CIN-I, II, III) by multichannel images using neural networks (NN).

Methods

Backpropagation neural networks employ one of the most popular neural network learning algorithms, the Backpropagation (BP) algorithm. It has been used successfully for wide variety of applications, such as speech or voice recognition, image pattern recognition, medical diagnosis, and automatic controls.

The backpropagation algorithm trains a given feed-forward multilayer neural network for a given set of input patterns with known classifications. When each entry of the sample set is presented to the network, the network examines its output response to the sample input pattern. The output response is then compared to the known and desired output and the error value is calculated. Based on the error, the connection weights are adjusted. The backpropagation algorithm is based on Widrow-Hoff delta learning rule in which the weight adjustment is done through mean square error of the output response to the sample input.[Зайченко Ю.П. 2004]

RBF Neural Networks

Radial Basis Functions are powerful techniques for interpolation in multidimensional space. A RBF is a function which has built into a distance criterion with respect to a centre. Radial basis functions have been applied in the area of neural networks where they may be used as a replacement for the sigmoidal hidden layer transfer characteristic in Multi-Layer Perceptrons. RBF networks have two layers of processing: In the first, input is mapped onto each RBF in the 'hidden' layer. The RBF chosen is usually a Gaussian. In regression problems the output layer is then a linear combination of hidden layer values representing mean predicted output. The interpretation of this output layer value is the same as a regression model in statistics. In classification problems the output layer is typically a sigmoid function of a linear combination of hidden layer values, representing a posterior probability. Performance in both cases is often improved by shrinkage techniques, known as ridge regression in classical statistics and known to correspond to a prior belief in small parameter values (and therefore smooth output functions) in a Bayesian framework. [Domagoj Kovacevic, Sven Loncaric 1997]

Approximators with local basis

RBF NN looks like perceptron with one hidden layer, carrying out nonlinear reflection

$\Re^d \Rightarrow \Re^m$ $y = \sum_i h_i \phi(w_i, x)$, being a linear combination of basic functions. But unlike perceptrons where these functions depend on projections to a set of hyperplanes $\sigma(wx)$, In the RBF NN functions which depend on distances to basic centers (often Gaussian) are used: $y = \sum_i h_i \phi_i(|w_i - x|), \phi_i(z) = e^{-z^2/\sigma_i^2}$.

Both sets of basic functions provide possibility of approximation of any continuous function with any accuracy. The main distinction between them is the method of information coding on a hidden layer. If perceptrons use global variables (sets of infinite hyperplanes) RBF networks depend on the compact spheres surrounding a set of basic centers (fig. 1).

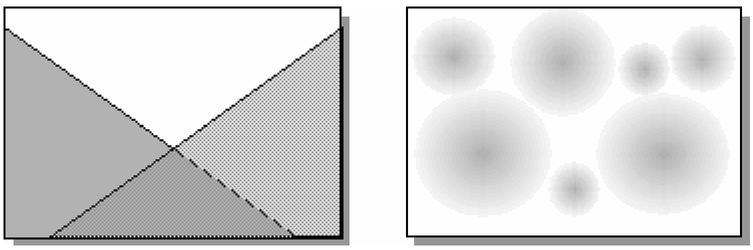

Fig. 1. Global (perceptron) and local (RBF NN) methods of approximation.

In the first case all neurons of hidden layer participate in approximation in a area of any point, in the second case - only the nearest neurons. As consequence of such inefficiency, in the latter case the quantity of the support function necessary for approximating with set accuracy exponentially increases with dimension of space. It is the main disadvantage of networks of radial basis [5].

Advantages RBF networks have the advantage of not suffering from local minima in the same way as Multi-Layer Perceptrons. This is because the only parameters that are adjusted in the learning process are the linear mapping from hidden layer to output layer. Linearity ensures that the error surface is quadratic and therefore has a single easily found minimum. In regression problems this can be found in one matrix operation. In classification problems the fixed non-linearity introduced by the sigmoid output function is most efficiently dealt with using iteratively re-weighted least squares.

Data preprocessing

Many indicators have high correlation. It can have negative impact on the neural networks work and classification of the tissues on normal/abnormal. For this purpose in order to avoid potential problems in two cases indicators have been changed. In one case, only a subset of indicators with correlation <0,95 was used for neural network training (defined experimentally). In other case, the Principal Component Analysis was used, indicators have been transformed to a set of independent non correlated indicators. [M.C. Jones and R. Sibson 1987]. These methods should improve the functioning of a neural network and increase accuracy of classification.

Experiments results

There are the diagnoses of two doctors representing presence of each of 6 possible types of tissues (Squamous, Columnar, SEA, CIN1, CIN2, CIN3). The following tables present the results of applying the neural networks. During the experiments, the Cross validation method was used. Cross-validation is a technique for assessing how the results of a statistical analysis will generalize to an independent data set. It is mainly used in settings where the goal is prediction, and one wants to estimate how accurately a predictive model will perform in practice. One round of cross-validation involves partitioning a sample of data into complementary subsets, performing the analysis on one subset (called the training set), and validating the analysis on the other subset (called the validation set or testing set). To reduce variability in the overall assessment of generalizability, multiple rounds of cross-validation are performed using different partitions, and the validation results are averaged over the rounds.

Tables 1 through 6 show the mean square error (MSE) for four types of neural networks (NN), namely Backpropagation Conjugate gradient NN, Backpropagation Quasi Newton NN, Backpropagation Scaled Conjugate gradient NN, and RBF NN each with two methods of data preprocessing (a subset of indicators with correlation <0,95 and the indicators transformed using PCA) for the following tissue types: Squamous, Columnar, SEA, CIN1, CIN2, and CIN3.

Table 1. RMSE for the 1-st tissue (Squamous)

	BP Conjugate gradient	BP Quasi Newton	BP Scaled Conju-gate gradient	RBF NN
RMSE(>0,95)	0,0607	0,0593	0,0499	0,0325
RMSE(PCA)	0,0502	0,0621	0,0507	0,0326

As seen from Table 1 RBF network gives the least error, thus both methods of data preprocessing are equivalent.

Table 2. RMSE for the 2-nd tissue (Columnar)

	BP Conjugate gradient	BP Quasi Newton	BP Scaled Conjugate gradient	RBF NN
RMSE (>0,95)	0,0459	0,0457	0,0449	0,0295
RMSE (PCA)	0,0471	0,0480	0,0408	0,0297

As seen from Table 2 RBF network gives the least error, thus both methods of data preprocessing are equivalent.

Table 3. MSE for the 3-rd tissue (SEA)

	BP Conjugate gradient	BP Quasi Newton	BP Scaled Conjugate gradient	RBF NN
RMSE (>0,95)	0,0561	0,0530	0,0492	0,0244
RMSE (PCA)	0,0580	0,0577	0,0582	0,0247

As seen from Table 3 RBF network gives the least error, thus both methods of data preprocessing are equivalent.

Table 4. MSE for the 4-th tissue (CIN1)

	BP Conjugate gradient	BP Quasi Newton	BP Scaled Conjugate gradient	RBF NN
RMSE (>0,95)	0,0462	0,0852	0,0408	0,0220
RMSE (PCA)	0,0817	0,0958	0,0458	0,0224

As seen from Table 4 RBF network gives the least error, thus both methods of data preprocessing are equivalent.

Table 5. MSE for the 4-th tissue (CIN2):

	BP Conjugate gradient	BP Quasi Newton	BP Scaled Conjugate gradient	RBF NN
RMSE (>0,95)	0,0427	0,0497	0,0345	0,0158
RMSE (PCA)	0,0392	0,0500	0,0352	0,0158

As seen from Table 5 RBF network gives the least error, thus both methods of data preprocessing are equivalent.

Table 6. MSE for the 4-th tissue (CIN3):

	BP Conjugate gradient	BP Quasi Newton	BP Scaled Conjugate gradient	RBF NN
RMSE (>0,95)	0,0734	0,0789	0,0739	0,0305
RMSE (PCA)	0,0656	0,0755	0,0877	0,0306

As seen from Table 6 RBF network gives the least error, thus both methods of data preprocessing are equivalent.

Conclusion

– The results of experiments proved that NN are applicable for solving problems described in this paper.
– Experiments showed that RBF network always gives the best results.
– Different methods of data preprocessing improves the NN performance. Both methods proved to be sufficiently effective.
– The developed approach of uterine cervix cancer diagnostics using NN can be recommended for further clinical approbation.

Acknowledgement

The paper is partially financed by the project ITHEA XXI of the Institute of Information Theories and Applications FOI ITHEA and the Consortium FOI Bulgaria (www.ithea.org, www.foibg.com).

Bibliography

[Schoonmaker J. 2007] Schoonmaker J. et al. Automatic Segmentation of Uterine Cervix for in vivo Localization and Identification of Cervical Interaepithelial Neoplasia // Apogen Technologies 2007

[Воробьева Л. И. 2008] Воробьева Л. И. Цитологический скрининг рака шейки матки // Здоровье Украины, 2008

[Koss L.G. 1989] Koss L.G. The complex test for cervical cancer detection. In: The Journal of the American Medical Association. – 1989

[Зайченко Ю.П. 2004] Зайченко Ю.П. Основи проектування інтелектуальних систем. Навчальний посібник – К.: Видавничий Дім «Слово», 2004

[Domagoj Kovacevic, Sven Loncaric 1997] Radial Basis Function-based Image Segmentation using a Receptive Field //Computer-Based Medical Systems. – 1997. Proceedings., Tenth IEEE Symposium , 1997

[M.C. Jones and R. Sibson 1987] What is projection pursuit ? // J. of the Royal Statistical Society, ser. A – 1987.

Authors' Information

Kateryna Malyshevska – Ph.d. student, National Technical University of Ukraine "Kyiv Polytechnic Institute", P.O. Box: 03056, Kyiv-56, Ukraine; e-mail: volovik_katya@yahoo.com