Annals of the University of Craiova, Math. Comp. Sci. Ser.
Volume 35, 2008, Pages 112–116
ISSN: 1223-6934

Neural network-based approach to discriminate healthy people from those with Parkinson's disease

Marius Ene

Abstract. This papers deals with the application of some probabilistic neural network (PNN) variants to discriminate between healthy people and people with Parkinson's disease. Three PNN types have been used in this classification process, related to the smoothing factor search: incremental search (IS), Monte Carlo search (MCS) and hybrid search (HS). The concrete application has provided diagnosis accuracies ranging between 79% and 81% for new, undiagnosed patients.

2000 Mathematics Subject Classification. 68T05; 68Q10.
Key words and phrases. PNN algorithm, classification, diagnosis, Parkinson.

1. Introduction

Parkinson's disease (also known as Parkinson disease or PD) is a degenerative disorder of the central nervous system that often impairs the sufferer's motor skills, speech, and other functions. PD causes cognitive and mood disturbances, being in many cases related.

Typically, the diagnosis is based on medical history and neurological examination conducted by interviewing and observing the patient in person using the Unified Parkinson's Disease Rating Scale (UPDRS). The disease can be difficult to diagnose accurately, especially in its early stages. Due to symptom overlap with other diseases, only 75% of clinical diagnoses of PD are confirmed to be idiopathic PD at autopsy [9]. Thus, automatic techniques based on Artificial Intelligence are needed to increase the diagnosis accuracy and to help physicians make better decisions.

This paper deals with the application of PNN to a medical dataset concerning PD with the aim of automatically classify patients in PD or non-PD depending on their medical attributes. The results of using PNN are better than the reported accuracy obtained in human early diagnosis, encouraging this kind of methodology.

2. Probabilistic neural networks

Probabilistic neural networks are radial basis neural networks with four layers and have been introduced by Donald Specht ([3], [4]). This special type of neural network gives a solution for classification problems using an approach based on Bayes probabilistic rule. Parzen estimators (1962) [5] are used in network architecture for estimating the classes probability density function f_i (p.d.f.).

PNN's uses a supervised training set for finding p.d.f. of the pattern layer. The great advantage is that we have to pass only one time through training data. In this way is guaranteed that the p.d.f. approximation will be better as the number

Received: 20 january 2008.

of training set instances is growing. A disadvantage is that training data must be stored for a later use in classification of new data, this resulting in an increased use of memory and increased training time as the number of input vectors and training data is increasing.

The classification problems that PNN tries to solve are the ones in which we need to determine the membership of a multivariate data x to one of q possible groups Ω_i, i=*1,...,q*, based of series of measurements. Let consider a classification problem with q categories where we need to decide $d(x) = \Omega_i$, $i \in \{1, 2, ..., q\}$, where x represents an input data (vector). Knowing the multivariate p.d.f's $f_1(x)$, $f_2(x)$,..., $f_q(x)$ for Ω_1, Ω_2,... Ω_q, the *a priori* probabilities $h_i = P(\Omega_i)$ of membership to Ω_i category and *cost* parameters l_i of incorrect decisions, then, according to Bayes rule, we will classify x as being part of Ω_i category if the following is true:

$$l_i h_i f_i(x) > l_j h_j f_j(x), i \neq j$$

where f_i represents the mixt-Gaussian p.d.f:

$$f_i(x) = \frac{1}{(2\pi)^{p/2}\sigma^p m_i} \sum_{j=1}^{m_i} e^{\frac{-(x-x_{ij})^T(x-x_{ij})}{2\sigma^2}}$$

Shortly, a PNN consist of a three-layer neural network (following the input layer): pattern layer, summation layer and output layer. In some cases, a fourth layer is used (an input data normalization layer).

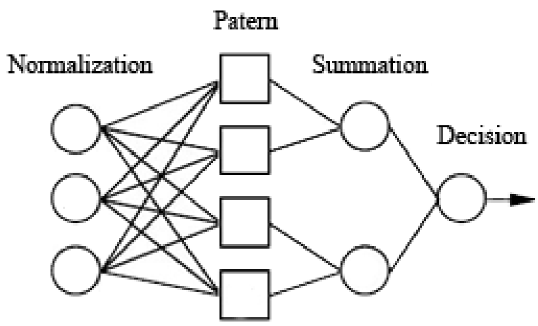
Pattern layer will have a neuron for each input vector. Each node will compute a product of input vector x (for classification) and a weight vector W_i, $Z_i = x.W_i$; then it will compute a nonlinear operation over Z_i. ($e^{\frac{(Z_i-1)}{\sigma^2}}$). Thus, the result will be given by $e^{-\frac{(W_i-x)^\tau(W_i-x)}{2\sigma^2}}$.

In summation layer each neuron will receive as input the output of preceding layer (pattern) for a specific class (so this layer will have a number of nodes equal with the number of classes/categories). The inputs are summed and we will obtain $\sum_i e^{-\frac{(W_i-x)^\tau(W_i-x)}{2\sigma^2}}$.

The output (decision-layer) neurons have each two inputs. These units will produce binary outputs related the two categories Ω_r si Ω_s, $r \neq s$, $r, s \in \{1, 2, \ldots, q\}$ using the following classification criterion:

$$\sum_i e^{-\frac{(W_i-x)^\tau(W_i-x)}{2\sigma^2}} > \sum_j e^{-\frac{(W_j-x)^\tau(W_j-x)}{2\sigma^2}}.$$

The weight of each unit depends on *cost* (loss) parameter, *a priori* known probabilities and the training instances number for each category $C = -\frac{h_s l_s}{h_r l_r} \cdot \frac{n_r}{n_s}$.



A main challenge is the way of finding σ parameter (Gaussian standard deviation of p.d.f.) using the training process. Usually σ will be chosen heuristically. A too large value or a too small value will lead to a larger value of the error classification rate. A smaller deviation will produce a too selective" approximation while a bigger deviation will make us see the input data as being in more than one category.

Because the only factor that can be selected for training is the standard deviation σ of Gaussian functions, the neural network will be initialized using as weight vector W_i (for a neuron in first layer) each x vector from training set, then connecting pattern layer with summation layer.

3. Parkinson dataset

This dataset is composed of a range of biomedical voice measurements from 31 people, 23 with Parkinson's disease (PD). Each column in the table is a particular voice measure and each row corresponds one of 195 voice recording from these individuals. The main aim of processing the data is to discriminate healthy people from those with PD, according to the "status" attribute which is set to non-PD for healthy and PD for people with Parkinson's disease, which is a two-decision classification problem [1], [2].

The attributes information consists of the following technical details:

- name - ASCII subject name and recording number
- MDVP:Fo(Hz) - Average vocal fundamental frequency
- MDVP:Fhi(Hz) - Maximum vocal fundamental frequency
- MDVP:Flo(Hz) - Minimum vocal fundamental frequency
- MDVP:Jitter(%),
- MDVP:Jitter(Abs),
- MDVP:RAP,
- MDVP:PPQ,
- Jitter:DDP - Several
- MDVP:Shimmer,
- MDVP:Shimmer(dB),
- Shimmer:APQ3,
- Shimmer:APQ5,
- MDVP:APQ,
- Shimmer:DDA - Several measures of variation in amplitude
- NHR,
- HNR - Two measures of ratio of noise to tonal components in the voice status
- RPDE,D2 - Two nonlinear dynamical complexity measures
- DFA - Signal fractal scaling exponent
- spread1
- spread2
- D2
- PPE

4. Results

The PNN algorithm has been implemented in Java programming language which gives an easy and cost free modality to implement the algorithm and portability to other operating systems. Data for training and test have been read from a Microsoft

Excel database (because of its extended use) but also can be used text files, Microsoft Access databases or other types of data storage systems that can be retrieved using registered Data Source Name.

The following table presents the accuracy obtained by training the probabilistic neural network using 70% of data for training and the rest of 30% for testing. The neural network was used 10 times for each method (IS, MCS or HS), the results below being averaged over the 10 computer run.

Method	Training accuracy (%)	Testing accuracy (%)
IS	81.73	79.78
MCS	81.48	80.92
HS	81.74	81.28

5. Conclusions

The aim of this paper was to verify the effectiveness of the application of PNN to a medical dataset, related to Parkinson's diseases. Such a dataset involves many different attributes, with various range of values, which make difficult the use of many machine learning techniques.

The results obtained applying PNN show the robustness of this methodology, even the data is very varied. We observe that there is no major difference between the three techniques of searching the smoothing parameterσ, although the hybrid technique seems to better perform.

Future work will try to refine the search, focusing on other heuristic techniques and to compare PNN performance with other neural networks methods, such as MLP, RBF etc.

Acknowledgement

The dataset was created by Max Little of the University of Oxford in collaboration with the National Centre for Voice and Speech, Denver, Colorado.

Thanks to Dr. Kenneth Revett, University of Westminster, London for providing the dataset and useful knowledge.

References

[1] Little M. A., McSharry P.E., Hunter E.J., Ramig L.O. (2008), *Suitability of dysphonia measurements for telemonitoring of Parkinson's disease*, IEEE Transactions on Biomedical Engineering (*to appear*).

[2] Little M. A., McSharry P.E., Roberts S.J., Costello D.A.E., Moroz I.M. (2007), *Exploiting Nonlinear Recurrence and Fractal Scaling Properties for Voice Disorder Detection*, BioMedical Engineering OnLine, 6:23 (26 June 2007).

[3] Specht D. F. (1988), *Probabilistic neural networks for classification mapping or associative memory*, Proceedings IEEE International Conference on Neural Networks, 1, 525-532.

[4] Specht D.F. (1990), *Probabilistic Neural Networks*, Neural Networks. 3 ,1, 109-118.

[5] Parzen E. (1962), *On estimation of a probability density function and mode*, Ann. Math. Stat., 33, 1065-1076.

[6] M. Gorunescu, F. Gorunescu, M. Ene, E. El-Darzi (2006), *A probabilistic neural network-based method for hepatic diseases diagnosis*, Buletinul Stiintific, Universitatea din Pitesti, Seria Matematica si Informatica, Nr. 11, pp. 59-66

[7] F. Gorunescu (2006), *Data Mining: concept, modele si tehnici*, Ed. Cartea Albastra - Microinformatica, Cluj-Napoca

[8] *http://www.statsoft.com/textbook/stneunet.html*

[9] *Wikipedia - http://en.wikipedia.org/wiki/Parkinson's_disease*

(Marius Ene) Department of Mathematics, Biostatistics and Informatics, University of Medicine and Pharmacy of Craiova
str. Petru Rares, 2-4, zip 200349 - Craiova, Dolj, Romania, Tel.& Fax: 40-251523929
E-mail address: enem@umfcv.ro