Neural Networks Based System for Cancer Diagnosis Support

Witold Jacak and Karin Pröll

Dept. of Software Engineering
Upper Austrian University of Applied Sciences
Softwarepark 11, A 4232 Hagenberg, Austria

Extended Abstract

Neural networks are proven tools for prediction tasks on medical data[1]. The paper presents the analysis of two different approaches for a system to support cancer diagnosis based on tumor marker values and blood parameters. Both systems use several heterogeneous artificial neural networks, which in parallel compute values for the estimation of tumor markers and additionally the risc of different cancer occurrence. The typical cancer prediction system is based on data coming from a vector $C = (c_1, , c_m)$ of tumor marker values. We need thousands of datasets for training and evaluation of neural networks. Missing values of tumor markers values in patient blood probes occur frequently and cause problems in neural network training. To overcome this problem we also make use of a blood parameter vector $P = (p_1, , p_n)$ of each patient containing values usually measured in standard blood counts to support training. As this vector might be incomplete too, the system must also work on partially incomplete blood parameters as input values for estimation of missing tumor marker values and their probable classification. We use two independent neural networks based systems: The first one is based on complete or incomplete tumor marker datasets C, the second one makes also use of a corresponding blood parameters vector P for computation of cancer risc.

Cancer diagnosis support system based on tumor marker values: We use parallel working systems ($Cancer_k$) with the same structure of heterogeneous neural networks, each of them trained on different types of cancer. The input of each $Cancer_k$ system is the complete or incomplete vector C of tumor marker specific for the chosen type of cancer, and the output represents possibility values between 0 and 1 of a real cancer disease. Output values greater than 0,5 are treated as high risk of cancer occurrence. Each system $Cancer_k$ consists of different groups of neural networks (Fig.1): Group of neural networks (C‗net) for individual marker C_i, FF neural network (CGroupFFnet) and pattern recognition neural network (CGroupPRnet) for a vector of a group of markers C and an aggregation method for final calculation of cancer risc.

Group of individual neural networks for specific marker: We compare three methods of aggregation for the individually calculated output values of C‗ nets. Maximum value or average value of all individual network outputs and a separately trained perceptron network. The diagnosis prediction based on aggregation of
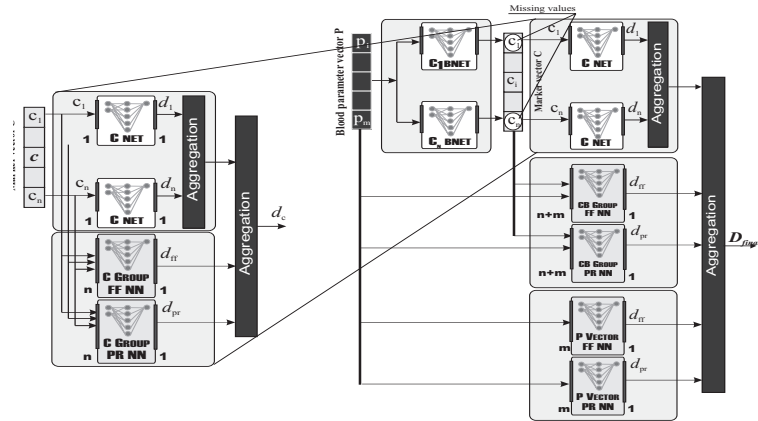
Fig. 1. Structure of Cancer prediction system based on tumor marker values

separately cancer predictions of individual marker networks C_ net is not sufficient for generally predicting cancer. We use two additional neural networks based on cumulative marker groups trained only for a specific cancer type.

Feed forward and pattern recognition neural networks for a group of tumor markers: When the tumor marker value in vector C is not available, then this value is set to -1. Based on such an assumption we can generate training sets for a specified cancer type $Cancer_k$ and train two neural networks: Feed forward neural network with 16-20 hidden neurons and tansig/linear activation functions (CgroupFFnet) and pattern recognition network with 16-20 hidden neurons (CgroupPRnet).

Tumor markers and blood parameters based cancer diagnosis support system: Partially available cancer marker values can be amended by values of standard blood parameters. We use the vector of blood parameters values as input for estimating missing tumor markers to provide input values for the neural networks for individual and group of markers[2]. The vector of blood parameters is also used as a third neural network system to support final computation of risc of cancer.

Experiments were taken on breast cancer using values of tumor markers C153, C125, CEA and C199 and groups of them with which we obtained reasonable results.

References

1. Astion, M.L., Wilding P., 1992, Application of neural networks to the interpretation of laboratory data in cancer diagnosis. Clinical Chemistry, Vol 38, 34-38.
2. Jacak, W., Proell, K., 2010, Neural Network Based Tumor Marker Prediction, Proceedings EMSS 2010, Marocco