

Chapter IV

Classification and Prediction with Neural Networks

Arnošt Veselý

Czech University of Life Sciences, Czech Republic

ABSTRACT

This chapter deals with applications of artificial neural networks in classification and regression problems. Based on theoretical analysis it demonstrates that in classification problems one should use cross-entropy error function rather than the usual sum-of-square error function. Using gradient descent method for finding the minimum of the cross entropy error function, leads to the well-known backpropagation of error scheme of gradient calculation if at the output layer of the neural network the neurons with logistic or softmax output functions are used. The author believes that understanding the underlying theory presented in this chapter will help researchers in medical informatics to choose more suitable network architectures for medical applications and that it helps them to carry out the network training more effectively.

INTRODUCTION

Medicine involves decision-making and classification or prediction is an important part of it. However, medical classification or prediction is usually a very complex and hard process at least from the following reasons:

- Much of the data that are relevant to classification or prediction, especially those received from laboratories, are complex or difficult to comprehend and can be interpreted only by experts.
- For reliable classification or prediction a large amount of data is frequently needed and some important anomalies in the data may be overlooked.

- When monitoring a patient, some of the patient's dangerous events can be too rare and therefore it may be difficult to identify them in the continuous stream of data.

Thus computer-assisted support could be of significant help. With the increasing number of clinical databases it is likely that machine-learning applications will be necessary to detect rare conditions and unexpected outcomes. The needful methods and algorithms can be found first of all in the domain of mathematical statistics and artificial intelligence. From means of artificial intelligence rule based experts systems were primarily used. Today the most important applications utilize algorithms based on neural networks, fuzzy systems, neurofuzzy systems or evolution algorithms.

Classical statistical methods require certain assumptions about the distribution of data. Neural networks can constitute a good alternative when some of these assumptions cannot be verified. From this point of view neural networks constitute a special kind of nonparametric statistical methods. Therefore the most successful neural network architectures are implemented in standard statistical software packages, as there is for example STATISTICA. Thus the eligibility of neural network algorithms for the given decision making task can be easily tested on the sample data using one of these software packages.

From abstract mathematical point of view medical classification and prediction tasks fall into the scope of either classification or regression problems. A classification or pattern recognition can be viewed as a mapping from a set of input variables to an output variable representing the class label. In classification problems the task is to assign new inputs to labels of classes or categories. In a regression problem we suppose that there exists underlying continuous mapping $y = f(x)$ and we estimate the unknown value of y using the known value of x .

Typical case of classification problem in medicine is medical diagnostics. As input data the patient's anamnesis, subjective symptoms, observed symptoms and syndromes, measured values (e.g. blood pressure, body temperature etc.) and results of laboratory tests are taken. This data is coded by vector x , the components of which are binary or real numbers. The patients are classified into categories D_1, \dots, D_m that correspond to their possible diagnoses d_1, \dots, d_m .

Many successful applications of neural networks in medical diagnostics can be found in literature (Gant, 2001). The processing and interpretation of electrocardiograms (ECG) with neural networks was intensively studied, because evaluation of long term ECG recordings is a time consuming procedure and requires automated recognition of events that occur infrequently (Silipo, 1998). In radiology neural networks have been successfully applied to X-ray analysis in the domains of chest radiography (Chen, 2002) and (Coppini, 2003), mammography (Halkiotis, 2007) and computerized tomography (Lindahl, 1997), (Gletsos, 2003) and (Suzuki, 2005). Also classification of ultrasound images was successful (Yan Sun, 2005). Neural networks have been also successfully applied to diagnose epilepsy (Walczak, 2001) or to detect seizures from EEG patterns (Alkan, 2005).

Also prediction of the patient's state can be stated as classification problem. On the basis of examined data represented with vector x patients are categorized into several categories P_1, \dots, P_m that correspond to different future states. For example five categories with the following meaning can be considered: P_1 can mean death, P_2 deterioration of the patient's state, P_3 steady state, P_4 improvement of the patient's state and P_5 recovery.

Neural networks were applied with good results for prognosis of the hepatectomized patients with hepatocellular carcinoma (Hamamoto, 1995). Also neural network predictions for patients with colorectal cancer (Bottaci, 1997) or with bladder cancer (Wei, 2003) were successful.

30 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the publisher's webpage:
www.igi-global.com/chapter/classification-prediction-neural-networks/7528

Related Content

Ooredoo Rayek: A Business Decision Support System Based on Multi-Language Sentiment Analysis of Algerian Operator Telephones

Badia Klouche, Sidi Mohamed Benslimane and Sakina Rim Bennabi (2022). *Research Anthology on Implementing Sentiment Analysis Across Multiple Disciplines* (pp. 1262-1279).

www.irma-international.org/chapter/ooredoo-rayek/308543

On the Mining of Cointegrated Econometric Models

J.L. van Velsen and R. Choenni (2010). *Dynamic and Advanced Data Mining for Progressing Technological Development: Innovations and Systemic Approaches* (pp. 122-135).

www.irma-international.org/chapter/mining-cointegrated-econometric-models/39641

Crime Hotspot Detection: A Computational Perspective

Emre Eftelioglu, Shashi Shekhar and Xun Tang (2016). *Data Mining Trends and Applications in Criminal Science and Investigations* (pp. 82-111).

www.irma-international.org/chapter/crime-hotspot-detection/157455

Semantics-Aware Advanced OLAP Visualization of Multidimensional Data Cubes

Alfredo Cuzzocrea, Domenico Sacca and Paolo Serafino (2007). *International Journal of Data Warehousing and Mining* (pp. 1-30).

www.irma-international.org/article/semantics-aware-advanced-olap-visualization/1791

Machine Learning Based Admission Data Processing for Early Forecasting Students' Learning Outcomes

Nguyen Thi Kim Son, Nguyen Van Bien, Nguyen Huu Quynh and Chu Cam Tho (2022). *International Journal of Data Warehousing and Mining* (pp. 1-15).

www.irma-international.org/article/machine-learning-based-admission-data-processing-for-early-forecasting-students-learning-outcomes/313585