Expert Systems with Applications 42 (2015) 6177–6183

Contents lists available at ScienceDirect

Expert Systems with Applications

journal homepage: www.elsevier.com/locate/eswa

Development of a decision support system based on neural networks and a genetic algorithm

Oleg E. Bukharov *, Dmitry P. Bogolyubov

Moscow Institute of Electronics and Mathematics, National Research University Higher School of Economics, 34 Tallinskaya Str., Moscow, Russia

ARTICLE INFO

Article history:
Available online 3 April 2015

Keywords:
Decision support system
DSS
Neural network
Genetic algorithm
GPGPU
CUDA

ABSTRACT

Given ever increasing information volume and complexity of engineering, social and economic systems, it has become more difficult to assess incoming data and manage such systems properly. Currently developed innovative decision support systems (DSS) aim to achieve optimum results while minimizing the risks of serious losses. The purpose of the DSS is to help the decision-maker facing the problem of huge amounts of data and ambiguous reactions of complicated systems depending on external factors. By means of accurate and profound analysis, DSSs are expected to provide the user with precisely forecasted indicators and optimal decisions.

In this paper we suggest a new DSS structure which could be used in a wide range of difficult to formalize tasks and achieve a high speed of calculation and decision-making.

We examine different approaches to determining the dependence of a target variable on input data and review the most common statistical forecasting methods. The advantages of using neural networks for this purpose are described. We suggest applying interval neural networks for calculations with underdetermined (interval) data, which makes it possible to use our DSS in a wide range of complicated tasks. We developed a corresponding learning algorithm for the interval neural networks. The advantages of using a genetic algorithm (GA) to select the most significant inputs are shown. We justify the use of general-purpose computing on graphics processing units (GPGPU) to achieve high-speed calculations with the decision support system in question. A functional diagram of the system is presented and described. The results and samples of the DSS application are demonstrated.

1. Introduction

Modern ideas on collecting, processing and applying knowledge are used in decision support systems (DSS), i.e. computer-based information systems designed to assist in making complicated decisions through a more profound and focused analysis of the subject area. The creation of DSS resulted from a merge of administrative information systems and database management systems.

A variety of methods are used to analyze and generate different types of decisions in DSS, e.g. search for information and knowledge in databases, situation and data analysis, precedent-based reasoning, simulation modeling, evolutionary calculations and genetic algorithms (GA), neural networks, cognitive modeling, etc.

If a DSS is based on artificial intelligence methods, it is called an intellectual DSS, or IDSS. Using a DSS, one could solve an unstructured, semistructured or even a multicriteria task.

No standard definition of the term "DSS" or its universally accepted classification are available. Researchers suggest different categorization criteria based on user-system interaction (Haettenschwiler, 1999), type of support (Power, 2007), or other approaches (Alter, 1980; Golden, Hevner, & Power, 1986; Holsapple & Whinston, 1996).

According to IDC and EMC (2011) the volume of information generated by people doubles every two years. This causes the problem of extracting the necessary information from infinite sources. Data mining is currently being actively developed as a solution to this problem. The objective of this technique is detecting the necessary information and elaborate interrelations among huge quantities of raw data.

The process of knowledge extraction at the initial stages of designing the intellectual Expert System (ES) and DSS is extremely difficult and labor-intensive, and not always successful if the databases in ill-structured subject areas contain incomplete, indistinct, polytypic or inconsistent information. The term "ill-structured" was introduced by G. Simon (1984) to designate a wide range of problems with the following characteristics: indistinct definitions,

* Corresponding author.
E-mail addresses: oleg_bukh.box@mail.ru (O.E. Bukharov), bogolub@mail.ru (D.P. Bogolyubov).

http://dx.doi.org/10.1016/j.eswa.2015.03.018
0957-4174/

changing terms, situations depending on a set of contexts, uncertainty, ambiguity, incompleteness, discrepancy, unreliability and diversity of initial data. Therefore, the intellectual data mining provides a promising approach to the solution of the problems described above. According to Fayyad, Piatetsky-Shapiro, and Smyth (1996), six types of tasks can be differentiated within data mining: classification, regression, clustering, summarization, dependency modeling, change and deviation detection. The latter types of tasks have recently attracted an increasing number of researchers, being the basis of Internet users' action analysis, identification of linguistic dependences in natural language texts, case history analysis to predict possible diseases, DNA sequencing, as well as trend prediction at financial exchanges. For example, from a researcher point of view, exchanges generate a large number of numerical sequences (corresponding to certain timepoints) called time series.

A time series is a sequence of statistical data on parameter value(s) collected at different timepoints. Time series data have a natural temporal order, each value characterized by time of measurement or measurement number in order. Time series significantly differ from simple data selection because analysis considers not only statistical variety and characteristics of a selection, but also the correlation between measurements and time (Shmojlova, 2002).

Time series forecasting is an essential task for many spheres of human activities, such as:

1. Medicine (e.g. forecasting reactions to various formulations and doses in a treatment course);
2. Biology (forecasting physiological and psychological features of animals and humans);
3. Sociology (forecasting social relations and demographic indicators);
4. Economics (forecasting sales volumes, exchange rates and stock prices).

2. Forecasting

Let us consider the most widespread statistical methods of time series forecasting (Magnus, Katyshev, & Pereseckiy, 2007) (see Table 1):

In spite of a great number of existing forecasting models, very few of them are able to find a dependence between all the factors significantly influencing the predicted value. Apart from accuracy, forecasting speed is also critical for many tasks. In such tasks, it is often necessary to strike a balance between accuracy and speed focusing on the factors making a significant impact on the predicted size. Thus, the main problems of time series forecasting are as follows:

Table 1
Widespread statistical methods of time series forecasting.

Statistical method	Application scope
Extrapolation forecasting methods	A trend or long-term tendency in a time series
Spectral correlation data analysis with period and seasonality search	Changes repeated throughout a certain period or tendencies observed in a time series
Models with statistical intervention parameters	Sharp changes of a tendency resulting from an external or internal influence in a time series
Harmonic models or Autoregressive integrated moving average models	Constant trend fluctuations in a time series with the period unknown at the beginning of research

1. Lack of an efficient estimation technique to evaluate the dependence between the input parameters and the predicted value;
   a. Difficulty finding the attributes with the greatest influence on the predicted value and the past period when these attributes have a significant influence on the predicted future value;
   b. The problem of determining a dependence between the identified attributes and the predicted value;
2. Application of sophisticated statistical methods requiring a high level of user skill and knowledge.

Being a model of complicated multidimensional nonlinear regression, the neural network is more accurate than the above-mentioned methods, and has a number of other advantages (Tsaregorodtsev, 2010):

1. Possibility to work with non-informative noise input signals: the neural network can reject them as useless for the task solution;
2. Possibility to work with polytypic information (continuous and discrete, qualitative and quantitative data types), which is considered to be a difficult task for statistical methods;
3. Given several outputs the neural network can solve a number of problems simultaneously;
4. There are algorithms for inverse task solution with a neural network trained to resolve a specific task. For example, it is possible to connect the new neural network inputs with the outputs of the current neural network and train the new network to produce the previous network inputs as its own outputs;
5. A neural network has fewer requirements for the qualification of its user compared to complicated statistical models capable of obtaining similar results;
6. Having initially set synaptic weights of a neural network, it is possible to recreate and check the suggested statistical models as well as improve them by network training (Haykin, 2006).

3. Interval neural networks

When forecasting the intervals of values, it is necessary to use interval neural networks. As professor Ishibuchi visually illustrated in his paper (Ishibuchi & Tanaka, 1993), the application of two standard neural networks for this purpose may cause forecast errors when the predicted value of the upper limit of the interval is lower than that of the lower one (see Fig. 1).

An interval neural network is a system of interconnected and interacting interval neurons. Inputs and outputs of interval neurons are intervals, each of them being a continuum set of values between the two limit values (see Fig. 2).

The interval parameters for forecasting are put into each input neuron. The value at the output of the input neuron is the same value as at its input. For all the other neurons, the interval values at the output can be calculated using the following formula:

$$Y = f\left(\sum_{i=1}^{N} w_i X_i + \phi\right),$$

where Y – the output from a neuron, w_i – the connection weight of the ith input (different for each neuron), f – the activation function, N – the number of neural inputs, ϕ – bias (different for each neuron) (Holsapple & Whinston, 1996. X_i – the input to the ith input unit (interval). The value Y is calculated by the rules of interval arithmetic:

$$A * B = [a_L, a_U] * [b_L, b_U] = [\min\{a * b | a \in [a_L, a_U], b \in [b_L, b_U]\},\\ \max\{a * b | a \in [a_L, a_U], b \in [b_L, b_U]\}],$$

where $*$ – any operator.