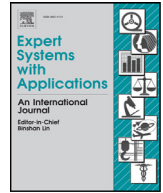




ELSEVIER

Contents lists available at ScienceDirect

Expert Systems With Applications

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

A linear model based on Kalman filter for improving neural network classification performance



Joko Siswanto^{a,b,*}, Anton Satria Prabuwo^{a,c}, Azizi Abdullah^a, Bahari Idrus^a

^a Faculty of Information Science and Technology, Universiti Kebangsaan Malaysia, 43600 UKM, Bangi, Selangor D. E., Malaysia

^b Faculty of Engineering, University of Surabaya, Jl. Kali Rungkut, Surabaya 60293, Indonesia

^c Faculty of Computing and Information Technology, King Abdulaziz University, Rabigh 21911, Saudi Arabia

ARTICLE INFO

Keywords:

Neural network
Linear model
Kalman filter
Classification performance

ABSTRACT

Neural network has been applied in several classification problems such as in medical diagnosis, handwriting recognition, and product inspection, with a good classification performance. The performance of a neural network is characterized by the neural network's structure, transfer function, and learning algorithm. However, a neural network classifier tends to be weak if it uses an inappropriate structure. The neural network's structure depends on the complexity of the relationship between the input and the output. There are no exact rules that can be used to determine the neural network's structure. Therefore, studies in improving neural network classification performance without changing the neural network's structure is a challenging issue. This paper proposes a method to improve neural network classification performance by constructing a linear model based on the Kalman filter as a post processing. The linear model transforms the predicted output of the neural network to a value close to the desired output by using the linear combination of the object features and the predicted output. This simple transformation will reduce the error of neural network and improve classification performance. The Kalman filter iteration is used to estimate the parameters of the linear model. Five datasets from various domains with various characteristics, such as attribute types, the number of attributes, the number of samples, and the number of classes, were used for empirical validation. The validation results show that the linear model based on the Kalman filter can improve the performance of the original neural network.

© 2015 Elsevier Ltd. All rights reserved.

1. Introduction

The classification problem is the problem of assigning an object into one of predefined classes based on a number of features or attributes extracted from the object (Zhang, 2000). In machine learning, classification is categorized as a supervised learning method. A classifier is constructed based on a training set with known class labels (Alpaydin, 2010). Classification problems occur in various real world problems, including problems in character recognition (Gao & Liu, 2008), face recognition (Zhifeng, Dahua, & Xiaou, 2009), speech recognition (Chandaka, Chatterjee, & Munshi, 2009), biometrics (Lyle, Miller, Pundlik, & Woodard, 2012),

medical diagnosis (Akay, 2009; Mazurowski et al., 2008; Verma & Zhang, 2007), industry (Jamil, Mohamed, & Abdullah, 2009; Kılıç, Boyacı, Köksel, & Küsmenoğlu, 2007; Nashat, Abdullah, & Abdullah, 2014; Rocha, Hauagge, Wainer, & Goldenstein, 2010), business (Chen & Huang, 2003; Huang, Chen, & Wang, 2007; Min & Lee, 2005), and science (Evet & Spiehler, 1987; Sigillito, Wing, Hut-ton, & Baker., 1989). Several classification algorithms have been proposed to solve classification problems, namely decision tree (Quinlan, 1986), linear discriminant analysis (Li & Yuan, 2005), Bayesian classifier (Domingos & Pazzani, 1997), rule-based classifier (Clark & Niblett, 1989), neural network (Lippmann, 1987), *k*-nearest neighbor (Cover & Hart, 1967), and support vector machine (Cortes & Vapnik, 1995).

Artificial neural network or simply neural network is a computational model inspired by the biological nervous system. Neural network is a nonlinear model, which is very simple in computation and has the capability to solve complex real problems including prediction and classification. Neural network has appears to be a significant classification method and an alternative to

* Corresponding author at: Faculty of Information Science and Technology, Universiti Kebangsaan Malaysia, 43600 UKM, Bangi, Selangor D. E., Malaysia. Tel.: +601121181700, +6283849409952.

E-mail addresses: joko_siswanto@ubaya.ac.id, joko.siswanto@gmail.com (J. Siswanto), antonsatria@eu4m.eu (A.S. Prabuwo), azizi@ukm.edu.my (A. Abdullah), bahari@ukm.edu.my (B. Idrus).

conventional classification methods (Zhang, 2000). It has been applied in various prediction and classification problems, such as bankruptcy prediction (Tsai & Wu, 2008), handwriting recognition (Goh, Mital, & Babri, 1997), product inspection (Kılıç et al., 2007), medical diagnosis (Mazurowski et al., 2008), and transportation (Garrido, de Oña, & de Oña, 2014).

The performance of a neural network is characterized by its structure, transfer function, and learning algorithm (Lippmann, 1987). The structure of a neural network depends on the number of hidden layers and the number of neurons in each hidden layer. However, there is no exact rule to determine the structure of a neural network. Generally, the more complex the relationship between the input data and the desired output, the more complex the structure of the neural network used in classification (Du & Sun, 2008). Therefore, a neural network classifier tends to be a weak classifier if it uses a structure that has an inappropriate number of hidden layers or an inappropriate number of neurons in its hidden layers. Although research on neural network classifiers has been widely conducted with significant results, it is still a challenging task, especially in research related to improving classification performance.

The ensemble method is a well-known method to improve the classification performance of a neural network by combining a series of trained neural networks (Giacinto & Roli, 2001; Glodek, Reuter, Schels, Dietmayer, & Schwenker, 2013; Zaamout & Zhang, 2012). However, if the outputs of each neural network are biased or correlated, then there is no guarantee that ensemble can improve the classification performance of the neural network (Zhang, 2000). Feature selection is another issue in improving classification performance. Feature selection aims to find a subset of features that achieves maximum classification performance and reduces computation effort. Various feature selection methods have been developed for neural network classifiers. One such method has used a genetic algorithm to select salient features (Li, 2006; Verma & Zhang, 2007). However, employing feature selection on a neural network classifier does not always improve classification performance, as reported in T.-S. Li (2006).

Improving classification performance is a promising issue, not only for neural network classifiers but also for other classifiers. Rocha et al. (2010) proposed classifier fusion for improving fruit and vegetable classification accuracy. They employed a combination of fusion of binary classifiers and a very long feature descriptor, including global color histogram (GCH), Unser's descriptors, color coherence vectors (CCVs), Border/Interior pixel Classification (BIC), and appearance descriptors. Although high classification accuracy is achieved, it takes significant time to perform the training stage. Mastrogianis, Boutsinas, and Giannikos (2009) proposed the use of the ELECTRE methods concepts to improve the accuracy of data mining classification algorithms. Even if the proposed method can improve classification accuracy of several data mining algorithms, it can be applied to classify only categorical objects. Hacıbeyoglu, Arslan, and Kahramanli (2011) analyzed the effect of discretization on classification. This method used entropy-based discretization to transform continuous-valued features into integer-valued features. Therefore, it cannot be applied to classify objects with only categorical- or integer-valued features.

In recent years, several authors tried to combine several techniques to improve classification performance. Farid, Zhang, Rahman, Hossain, and Strachan (2014) have proposed two hybrid algorithms of decision tree (DT) and naïve Bayes (NB) classifiers for multi-class classification. The first algorithm used NB to remove misclassified instances from training dataset before used to build DT. The second algorithm used DT to find a subset of attributes that play important roles in classification. Selected attributes by DT were then used for classification using NB. Seera and Lim (2014) have used Fuzzy Min–Max (FMM) neural network, classification

and regression tree (CART), and random forest (RF) model to develop a hybrid intelligent system for medical data classification. FMM neural network was used to generate hyperbox fuzzy set. The generated hyperbox was then used to build CART. Finally, to increase classification performance an ensemble of CART was constructed using RF. Affonso, Sassi, and Barreiros (2015) have combined rough sets theory and fuzzy neural network for biological image classification. They used rough sets theory for feature selection. The selected features were used to train a multilayer perceptron neuro fuzzy network. Onan (2015) have proposed the combination of instance selection, feature selection, and fuzzy-rough nearest neighbor for automated diagnosis of breast cancer. Fuzzy-rough instance selection method was used to remove useless or erroneous instances from dataset, while consistency-based feature selection method and a re-ranking algorithm were used to select important feature. Pruengkarn, Chun Che, and Kok Wai (2015) have used clustering technique, feature selection, and ensemble of classifier to improve classification performance. Clustering technique was employed to separate dataset into misclassification dataset and clean dataset. The clean dataset was classified using a common classifier including DT, NB, ANN, and SVM. Whereas feature selection technique based on fuzzy C-means and ensemble of classifier using majority voting were used to classify misclassification dataset. Although all authors reported achieving high classification accuracy, they did not report the computing time for the proposed methods.

A neural network classifier achieves high classification accuracy when its predicted output is very close to its desired output. Therefore, to increase neural network classification accuracy, the use of a transformation that transforms the predicted output of a neural network to a value close to the desired output can be considered as a post processing. A linear model is a simple transformation that can be used to achieve such a purpose. The linear model consists of independent (input) variables, dependent (output) variables, and unknown parameters. The parameters of a linear model need to be estimated such that the error between the predicted output and the desired output is minimized. The Kalman filter (Kalman, 1960) is a method that can be used to estimate the parameters of a linear model. The Kalman filter is a recursive method for fitting a linear model to a given dataset such that the sum of square error is minimized without performing matrix inversion as in ordinary least square. Even if the model has a number of variables greater than the number of dataset elements, the Kalman filter can still calculate the estimate (Wu, Rutan, Baldovin, & Massart, 1996).

This paper proposes a method to improve neural network classification performance by constructing a linear model based on the Kalman filter. The proposed method uses the Kalman filter iteration to estimate the parameters of a linear model. The model uses the linear combination of object features and predicted outputs of a neural network as input variables to predict class labels. As in a neural network, the model can use any type of variables as input. Therefore, the model would improve neural network classification performance without considering the types of object features.

The rest of the paper is organized as follows. Sections 2 and 3 provide a brief explanation about the structure of neural network and Kalman filter, respectively. Section 4 explains the proposed method. Section 5 describes datasets and method used for validation. Section 6 presents experimental results and discussion. And finally, conclusion and future work are provided in Section 7.

2. Neural network

The neural network model consists of interconnected neurons with weights, arranged in layers. The structure of a neuron consists of inputs p_1, p_2, \dots, p_n , weights w_1, w_2, \dots, w_n , bias b ,

Download English Version:

<https://daneshyari.com/en/article/382197>

Download Persian Version:

<https://daneshyari.com/article/382197>

[Daneshyari.com](https://daneshyari.com)