Applied Mathematics and Computation 221 (2013) 484–490

Contents lists available at SciVerse ScienceDirect

Applied Mathematics and Computation

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

A neural network algorithm to pattern recognition in inverse problems

Ali Sever

Computer Information Department, Pfeiffer University, Misenheimer, NC 28109, USA

ARTICLE INFO

Keywords:
Artificial neural networks
Inverse problems
Machine learning
Pattern recognition
Regularization

ABSTRACT

Considerable attention is currently being devoted to new possibilities of artificial neural networks (ANN) using in view of their increasing importance for solving the problem of automated reconstruction of the inner structure of an object. Accompanying algorithms that effectively quantify uncertainties, deal with ill-posedness, and fully take the nonlinear model into account are needed. A new ANN-based regularization model is generated and applied to the task of reconstruction of an inhomogeneous object.

Pattern recognition may be viewed as an ill-posed inverse problem to which the method of regularization can be applied. In this study, applications of methods from the theory of inverse problems to pattern recognition are studied. A new learning algorithm derived from a well-known regularization model is generated and applied to the task of reconstruction of an inhomogeneous object as pattern recognition. Particularly, it is demonstrated that pattern recognition can be reformulated in terms of inverse problems defined by a Riesz-type kernel. This reformulation can be employed to design a learning algorithm based on a numerical solution of a system of linear equations. Finally, numerical experiments have been carried out with synthetic experimental data considering noise level up to 5%. Reasonable good recoveries have been achieved with this methodology, and the results of simulations of this are compatible with the existing methods. This method can be used in practice of pattern recognition technology and its development and deployment for applications in industry.

1. Introduction

Inverse problems frequently arise in experimental modeling situations when one is interested in the description of the internal structure of a system and is given indirect, noisy data. Estimating the response of a system given a complete specification of the internal structure, on the other hand, is the forward problem.

The modeling problem originates when one is given noisy data, observed over irregular intervals of space and time, and is asked to develop a reasonable model to fit that observed data.

With the advent of high-speed computers and artificial intelligence techniques, this modeling problem underwent a metamorphosis and emerged as a machine learning problem [2,3]. Tikhonov and Lanweber regularized learning algorithms have recently received an increasing interest due to both theoretical and computational motivations [1,4,11]. Great deal of attention is currently being devoted to new possibilities of using artificial neural networks due to their increasing importance for solving the problem of automated reconstruction of the inner structure of an object. Accompanying algorithms that effectively quantify uncertainties, deal with ill-posedness, and fully take the nonlinear model into account are needed There-

E-mail address: ali.sever@pfeiffer.edu

0096-3003/$ - see front matter © 2013 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.amc.2013.06.094

fore, it is necessary to both look for possible ways to improve the classical learning algorithms already existent in literature, and to identify new methods which can compete with the traditional ones in speed, robustness, and quality of results.

Inverse problems are often formulated by assuming that the underlying phenomenon is a dynamic system characterized by mathematical equations, although no such assumption is always essential. Often the goal is to build an algorithmic model of the underlying phenomena. In some contexts a model is only a means to an end. The ultimate goal in such cases is to test the validity of a hypothesis. In these cases, the model is used as a classifier (e.g., neural nets and decision trees), and it matters little whether the model is parametric or non-parametric; the classification accuracy becomes more important. From this point of view the entire field of Machine Learning can be treated as inverse problems [2,7,8]. By their very nature, inverse problems are difficult to solve. Sometimes they are ill-posed. A *well-posed* mathematical problem must satisfy the following requirements: existence, uniqueness and stability. The existence problem is really a non-issue in many realistic situations because the physical reality must be a solution. However, due to noisy and/or insufficient measurement data, an accurate solution may not exist. More often, the major difficulty is to find a unique solution; this is especially when solving a parameter identification problem. Different combinations of parameter values may lead to similar observations. A useful strategy to handle the non-uniqueness issue is to utilize a priori information as additional constraints. These constraints generally involve the imposition of requirements such as smoothness on the unknown solution or its derivatives, or positivity, or maximum entropy or some other very general mathematical property. A more aggressive approach would be the use of regularization. Given an observed data set, genetic algorithms and programming can be used to search a hypothesis space.

No patterns can be derived solely from empirical data [12]. Some hypotheses about patterns have to be chosen and, from amongst patterns satisfying these hypotheses, a pattern with a good fit to the data must be sought.

In neurocomputing framework, searching for parameters of their input/output functions is called learning, and samples of data training sets and a capability to satisfactorily process new data that have not been used for learning is called generalization.

The capability of generalization depends upon the choice of a hypothesis set of input/output functions, in which one searches for a pattern (a functional relationship) that matches the empirical data. So a restriction of the hypothesis set to only physically meaningful functions can improve generalization.

In this paper, starting from a reformulation of the pattern recognition as an inverse problem, we introduce an alternative learning algorithm derived by a well-known regularization method. We use a Riesz-type kernel to solve classification tasks by transforming the geometry of input space by embedding them into higher dimensional, inner product spaces, and introducing a regularization method which adds to the derived integral equation a new term, called stabilizer, which penalizes undesired input/output functions. We split the problem into a simpler, ill-posed problem (an integral equation with a Riesz-type kernel) and a well-posed problem. In this way, we isolate and better control the propagation of errors due to the ill-posedness [8]. Then we show that this reformulation can be employed to design a learning algorithm based upon a numerical solution of a system of linear equations.

This method can be used in practice of pattern recognition technology and its development and deployment for applications in industry such as automated machine recognition of objects, signals or images and automated decision-making based on a given set of parameters.

The rest of the paper is organized as follows: Section 2 describes our model and justifies its use. In Section 3, we formulate the proposed regularized learning algorithm. Section 4 presents main simulation results. Finally, we conclude the paper with a summary of the work in Section 5.

2. Generalization model as regularization

Let us formulate the generalized problem as regularization in the following way:

Find a function $\sigma_1 \in L_\infty(\Omega)$, $\Omega \in \Re^n$, given the function $B(x_k) = w(x_k)$, $x_k \in \Omega_k$, $\Omega_k \in \Re^n$.Therefore, we have the following integral equation of the first kind

$$A\sigma_1(x) = B(x), \quad x\,\bar{C}\,\Omega_k \tag{2.1}$$

where

$$A\sigma_1(x) = \int_\Omega k(x,y)\sigma_1(y)dy \text{ and } k(x,y) = (1/2\pi)^2|x-y|^{-2}$$

and A is considered as an operator from $L_\infty(\Omega)$ into $L_\infty(\Omega_k)$. This integral equation is a Fredholm integral equation of the first kind with a Riesz-type kernel.

First we need to show that Eq. (2.1) represents a severely ill-posed problem. Then we have to prove that a solution $\sigma_1\,(\alpha)$ to the Eq. (2.2) exists and is unique.

Theorem 1. Let assume that Ω and Ω_k are nonintersecting domains in $\Re^3$. Then the integral Eq. (2.1) with the Riesz-type kernel represents an ill-posed problem.