

Chapter 9

Using Particle Swarm Optimization Algorithm as an Optimization Tool Within Developed Neural Networks

Goran Klepac
Raiffeisenbank Austria, Croatia

ABSTRACT

Developed neural networks as an output could have numerous potential outputs caused by numerous combinations of input values. When we are in position to find optimal combination of input values for achieving specific output value within neural network model it is not a trivial task. This request comes from profiling purposes if, for example, neural network gives information of specific profile regarding input or recommendation system realized by neural networks, etc. Utilizing evolutionary algorithms like particle swarm optimization algorithm, which will be illustrated in this chapter, can solve these problems.

INTRODUCTION

Neural networks are not self-explainable regarding to their nature, and it is challenging to find out right methodology, which will find logical and explainable connections between inputs and outputs. These connections and their explanation are not linear and it cannot be observed in that way, but finding typical (optimal) mixture of input values in a neural network for achieving desired output is a great movement. Usage of neural networks for constructing predictive models is a demanding process. It demands precise and objective determination of sample construction, target variable construction, attribute relevance analysis, model testing and many other activities, which will guarantee that developed model is robust, stable, reliable and predictive.

DOI: 10.4018/978-1-5225-5134-8.ch009

If we are talking about predictive models with binominal output and predictive models based on logistic regression, neural networks or similar techniques, then determination of initial states of variable values for achieving specific output are relatively simple. It can be a manual process, but achievable from perspective of human effort.

Reason why someone would like to find out which values of input variables will cause best fit for specific output is that we would like to find out typical case, or profile. That means if we would like to find out typical profile of churning based on developed binominal predictive model we should find combination of input values which will result with maximum output value in zone of wanted output.

Things became more complicated when we have multinomial output from predictive models.

Main advantage of binominal output usage is ability to understand relations between target variable and potential predictors and business logic check.

From technical point of view, data mining techniques like neural networks, and logistic regression by the nature of their algorithms, prefer to operate with values between 0 and 1. Dummy variables could be interpreted as membership declaration with 0 and 1 values. If some value belongs into specific class represented as dummy variable, it is true and dummy variable has value "1" otherwise "0".

Robust and stable predictive models have few attributes incorporated into model. It could be 6-10 of most predictive attributes. As it is evident initial data sample could contain more than hundreds of potential predictors. Some of them are original variables from databases as socio demographic values assigned to each customer, and other has behavioural characteristics defined by experts and extracted from existing transactional data.

Attribute relevance analysis has two important functions:

- Recognition of most important variables which has greatest impact on target variable
- Understanding relations and logic between most important predictor and target variable, and understanding relations and logic between most important predictors from target variable perspective

Contrary to assurance that powerful hardware and sophisticated software can substitute need for attribute relevance analysis, attribute relevance analysis is important part of each kind of analysis which operates with target variable. Recognition of most important variables which has greatest impact on target variable reduces redundancy and uncertainty at model development process stage. It provides robustness of the model and model reliability. Attribute relevance analysis beside importance measuring, evaluates attribute characteristics. Attribute characteristics evaluation includes measuring attribute values impact on target variables. It helps on understanding relations and logic between most important predictor and target variable, and understanding relations and logic between most important predictors from target variable perspective. After attribute relevance analysis stage, analyst has initial picture about churning profile and behaviour. This stage often opens many additional questions related to revealed relations and sometimes induces construction of new behavioural (derived) variables, which also should pass attribute relevance analysis process.

From perspective of predictive modelling there are two basic data sample types for predictive churn model development:

- Data sample with binomial target variable
- Data sample with multinomial target variable

28 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/using-particle-swarm-optimization-algorithm-as-an-optimization-tool-within-developed-neural-networks/198928

Related Content

Biclustering of DNA Microarray Data: Theory, Evaluation, and Applications

Alain B. Tchagang, Youlian Pan, Fazel Famili, Ahmed H. Tewfik and Panayiotis V. Benos (2011). *Handbook of Research on Computational and Systems Biology: Interdisciplinary Applications* (pp. 148-186).

www.irma-international.org/chapter/biclustering-dna-microarray-data/52315

A Generalized 2-D Model for Fully Bounded Chaotic Attractors and Chaotic Seas

Zeraoulia Elhadj (2012). *International Journal of Artificial Life Research* (pp. 53-56).

www.irma-international.org/article/generalized-model-fully-bounded-chaotic/74336

A Model for a Heterogeneous Genetic Network

Ângela T.F. Gonçalves and Ernesto J.F. Costa (2010). *Handbook of Research on Computational Methodologies in Gene Regulatory Networks* (pp. 523-545).

www.irma-international.org/chapter/model-heterogeneous-genetic-network/38250

MP Modelling of Glucose-Insulin Interactions in the Intravenous Glucose Tolerance Test

Vincenzo Manca, Luca Marchetti and Roberto Pagliarini (2011). *International Journal of Natural Computing Research* (pp. 13-24).

www.irma-international.org/article/modelling-glucose-insulin-interactions-intravenous/58063

Human-Robot Interaction Design Using Smart Device Based Robot Partner

Jinseok Woo and Naoyuki Kubota (2016). *International Journal of Artificial Life Research* (pp. 23-43).

www.irma-international.org/article/human-robot-interaction-design-using-smart-device-based-robot-partner/179254