

Usage investigation of non-linear discriminate functions for classification purposes

Egmonts Treiguts
Institute of Information Technology
Faculty of Computer Science and Information Technology
Riga Technical University
1 Kalku Street, Riga LV-1658, Latvia
E-mail: egmonts@mail.turiba.lv

The effectiveness of non-linear discriminate functions using for classification purposes has been investigated. Many classification problems are non-linear in origin, which are difficult to solve with conventional classification methods - Quinlan ID3 decision trees induction method [1, 2] or linear discriminate machine [3]. In many papers devoted to the classification problems certain attention is focused for solving non-linear data domains. For that purpose usually polynomial functions are used. With such functions new attributes and new combinations of attributes are constructed which can be used in decision tree induction method [4]. Those functions may be used as discriminate functions in non-linear machines too [3]. It is essential to determine the required complication of the chosen function. Usually the criteria of choose the complexity of function absents and the second order polynomial function is chosen.

In present investigation the Group Method of Data Handling (GMDH) is used for choosing a set of non-linear functions for given data domain. The GMDH method allows to determine the influence of attributes and combinations of attributes to the value of class of data domain. Using external selection criteria the optimal set of non-linear functions can be found. Such selected functions are optimal even by complexity and effectiveness.

Using the set of selected non-linear functions and real world data domain the classification was performed. The learning of non-linear machine with different numbers of records of learning data set was performed. The error of classification with the cross - validation method is determined.

Bibliography

1. J.R.Quinlan. (1986) Induction of decision trees. Machine Learning 1, P. 81 - 106.
2. J.R.Quinlan. (1993) C4.5 Programs for machine learning. San Mateo, CA, Morgan Kaufmann.
3. Н. Нильсон. () Непараметрические методы обучения Ф-машин.
4. A.Ittner, M.Schlosser. (1996) Non-Linear Decision Trees - NDT. Proceedings of the 13th International Conference on Machine Learning (ICML'96)

5. А.Г.Ивахненко, Ю.П.Заиченко, В.Д.Димитров (1976) Принятие решений на основе самоорганизации. Москва, Сов. радио. 116 – 204.
6. A.G.Ivakhnenko, G.A.Ivakhnenko. (1995) The Review of Problems Solvable by Algorithms of the Group Method of Data Handling (GMDH), Pattern Reecognition and Image Analysis, Vol. 5, No. 4, 527 – 535.