JOURNAL OF AGRICULTURAL INFORMATICS

AGRÁRINFORMATIKA FOLYÓIRAT

Special Issue
on
Information Systems: an Agro-Environment Perspective

Co-Editors of this Special Issue

Dr. Zacharoula Andreopoulou

Prof. Dr. Milan Mesic

Dr. Mariana Golumbeanu

EFITA

HAAI
MAGISZ
www.magisz.org

Nemzeti Fejlesztési Ügynökség
ÚMFT infovonal: 06 40 638 638
nfu@meh.hu • www.nfu.hu

New Hungary Development Plan

Content

A Genetic Programming Study on Classification of Cassava Plant

Indra Laksmana[1], Rosda Syelly[2], Nurzarah Tazar[3], Perdana Putera[4]

ABSTRACT

Cassava (*Manihot esculenta Crantz*) is an important plant that is consumed in many forms. It could be processed as vegetable, chips, fodder, or bioethanol through a fermentation process. The cyclic acid HCN of cassava varies based on the varieties. Cassava with high HCN is toxic when it is consumed directly. This research designed a system to identify the cassava varieties based on HCN content by applying a heuristic search algorithm, using genetic operations. The identification of HCN content by applying Generic programming produced a structured classification rule and represents in tree form. The experiment in this study used binary code data generated from booleanizing process. Binary code data is divided into training data and test data using 5-fold cross-validation, and then the process of genetic operation. Rules are derived from repeated experiments to get the best rule. The best rule to identify with an average accuracy of 95.26%, obtained on population parameters of 10,000, 20-30 nodes. The node consists of Function set of AND, OR, NOR and 96 terminal sets (attributes / identifiers); in addition, the best classification rules are obtained on the crossover probability of 0.9 and 0.1 mutations of 10 generations. The resulting Rule can be utilized by the community in identifying the class of HCN cassava content.

1. Introduction

The problems of classification often occur in daily life, such as choosing a vehicle, diagnosing the disease, looking for foods or drugs. It requires someone's skilled, so the mistakes in the classification of decisions could be minimized. The limitation of skilled increase the error in classifying, therefore an alternative method is needed in determining a solution to classification problem. The selection of appropriate classifier requires consideration of many factors, namely classification accuracy, algorithm and computational performance (Qurat-ul-ain et al. 2010). According to Wahyudi (2013) Classification is a collection of a record in the form of training data set, where each record contains a set of attributes and one attribute is a class.

The concept of artificial intelligence can be used to answer the classification problem. Artificial intelligence has the ability to think, guess an answer or perform the certain tasks such as human behavior that allow beyond human capabilities (Nakamura et al.2017). One of the artificial intelligence solutions that can be used in classification problems is genetic programming. Genetic programming (GP) is used to study patterns of data (Sudharmono. 2012). GP is a variant of the genetic algorithm which uses simulated evolution to discover functional programs to solve some task (Luke 2000). According to Sakprasat and Sinclair (2007), the main motivation for using genetic programming in classification rule mining is robustness and an adaptive search method making it more effective in finding patterns. Laksmana et al (2013) has applied GP programming method in identification of family of medicinal plants with an accuracy of 86.32%, resulting in a hierarchy in identifying medicinal plants.

[1] Department Agricultural Engineering, Payakumbuh State Polytechnic of Agriculture
indra.puskom@gmail.com
[2] Computer Engineering Department, Payakumbuh School of Technology
rosdasyelly@gmail.com
[3] Department of Food Technology, Payakumbuh State Polytechnic of Agriculture
perdanaputera81@gmail.com
[4] Department of Agricultural Engineering, Payakumbuh State Polytechnic of Agriculture
perdanaputera81@gmail.com