Chapter 5
GPdotNET Open Source Software for Running Genetic Programming

ABSTRACT

In this chapter, GPdotNET v5 genetic programming tool is presented from the user's perspective. GPdotNET is a computer program for running tree-based genetic programming, and its application is modelling supervised machine-learning-based problems. The chapter contains detailed information on how to use GPdotNET in order to prepare data, setup GP parameters, and to run the GP search algorithm. Since GPdotNET supports all three kinds of supervised machine learning problems, the chapter contains three use cases which demonstrate how to successfully build high quality regression, binary, and classification models. GPdotNET contains export module, where the user is able to export GP model to Excel, R language, and Wolfram Mathematica.

INTRODUCTION

GPdotNET v5 is an open source computer program for running tree based genetic programming. GPdotNET started at 2006 to be a simple command line tool for GP application in modelling. For the first time, in November 2009 it was published as an open source project, and currently is hosted at http://github.com/bhrnjica/gpdotnet. Since the beginning as the open source project, GPdNET started to be of interest of many researchers and engineers. Plenty

DOI: 10.4018/978-1-5225-6005-0.ch005

of journal articles, master thesis, PhD thesis and books, used GPdotNET. Today GPdotNET is respectable open source computer program which is used by hundreds of engineers in developing models mainly based on genetic programming.

The current software version is GPdotNET version 5 specially developed for this book edition. Since this version of GPdotNET is developed to be GP-based computer program only, so the other non-related GP modules (e.g. GA and ANN related modules have been removed and are going to be released as parts of other open source projects. History of all features which have been supported by the GPdotNET since the beginning of the project are presented in Table 1. As can be seen the GPdotNET v5 is expanded with several important features that have not been in the previous versions.

Table 1. Features development through GPdotNET versions

Feature	GPdotNET Version				
	1.0	2.0	3.0	4.0	5.0
GP Regression	+	+	+	+	+
GP Binary classification	-	-	-	+	+
GP Multiclass classification	-	-	-	+	+
GA Optimization	-	+	+	+	-
GA-LP – Traveling Salesman Problem	-	+	+	+	-
GA LP – Assignment Problems	-	+	+	+	-
GA LP – Transportation Problems	-	+	+	+	-
ANN Regression	-	-	+	+	-
ANN Binary classification	-	-	+	+	-
ANN Multiclass classification	-	-	+	+	-
Data processing	+	+	+	+	+
Data handling missing value	-	-	+	+	+
Data normalization	-	-	+	+	+
Data type (real, binary, factor)	-	-	+	+	+
Export – Excel	+	+	+	+	+
Export – Confusion matrix	-	-	-	-	+
Export - Wolfram Mathematica	-	-	-	+	+
Export – R Language Model	-	-	-	+	+
Export – R Language Data and Model	-	-	-	-	+
Multi document support	-	-	-	-	+
Handling more models per document	-	-	-	-	+