Contribution ID: **118** Type: **Poster (sponsored)**

Configuration of Genetic Programming Classification Algorithms for Financial Forecasting using Grammatical Evolution

inancial forecasting is a widely researched problem domain which is known to be quite challenging and exhibits characteristics of uncertainty. Genetic Programming (GP) [1] has been shown to be an effective tool for financial forecasting [2]. However, the manual design of GP classification algorithms for financial forecasting still remains the popular approach despite it being shown to be an error-prone, time-consuming task influenced by human bias. The volatility of financial forecasting problems require a tool that can respond to changes in a timely manner while also maintaining acceptable prediction rates. In this research, we propose the use of grammatical evolution (GE)[3] to configure GP classification algorithms for financial forecasting. Grammatical evolution and genetic programming are population-based algorithms and part of their functionality is evaluating the fitness of each individual at each generation. Inevitably this leads to high run-times furthermore, the algorithms are stochastic in nature which means a number of runs have to be performed in order to obtain a normal distribution of results. To evaluate the proposed grammatical evolution systems, we made use of the CHPC distributed architecture. Fifteen stocks were selected from the NASDAQ, NYSE, XETRA and HKSE stock exchanges. A varied selection was made because different industry stock have a varied volatility, for example stock from the technology sector is more volatile than stock from the banking sector. Each dataset comprises of data from 1500 trading days 03/01/2012 to 05/03/2018 (1000 training and 500 test).The GE algorithms are distributed over the cores and multiple runs performed on multiple cores. Certain operations of the algorithm that can be performed independently are distributed over the cores, such as population generation or fitness evaluation. For example, the GE system consists of a GE algorithm and GP algorithm running simultaneously. The population generation of the GE is distributed over a fixed number of cores. Each member of the GE population then initialises a GP algorithm which is also distributed over a specified number of cores, for a fixed number of runs. The GE system also has to perform a number of runs and these are performed on the cores. The system is coded in the Java programming language mainly using the threads API. Each individual of the GE is represented as a core therefore as a thread. The large queue on the Lengau Cluster is used, using typical values of cores ranging from 600 to 1800.

Presenter Biography

Primary authors: Mr NYATHI, Thambo (University of KwaZulu-Natal); Dr PILLAY, Nelishia (University of Pretoria)

Presenter: Mr NYATHI, Thambo (University of KwaZulu-Natal)

Session Classification: Poster session

Track Classification: HPC Techniques and Computer Science