Chapter 40
Hybrid Neural Genetic Architecture:
New Directions for Intelligent Recommender System Design

Emmanuel Buabin
Methodist University College Ghana, Ghana

ABSTRACT

The objective is a neural-based feature selection in intelligent recommender systems. In particular, a hybrid neural genetic architecture is modeled based on human nature, interactions, and behaviour. The main contribution of this chapter is the development of a novel genetic algorithm based on human nature, interactions, and behaviour. The novel genetic algorithm termed "Buabin Algorithm" is fully integrated with a hybrid neural classifier to form a Hybrid Neural Genetic Architecture. The research presents GA in a more attractive manner and opens up the various departments of a GA for active research. Although no scientific experiment is conducted to compare network performance with standard approaches, engaged techniques reveal drastic reductions in genetic operator operations. For illustration purposes, the UCI Molecular Biology (Splice Junction) dataset is used. Overall, "Buabin Algorithm" seeks to integrate human related interactions into genetic algorithms as imitate human genetics in recommender systems design and understand underlying datasets explicitly.

INTRODUCTION

Data explosion across business platforms has necessitated the call for efficient storage mechanisms. Research scientists estimate that close to 80% of a business's data lies in amorphous data format. This means, management decisions are likely to be optimized should unstructured data formats be considered in decision-making processes. Since its inception, the Internet has grown to become one of the largest data repositories in the world. With every computer literate as potential author on the Web, new domains have been registered and corresponding Website

DOI: 10.4018/978-1-4666-3604-0.ch040

information added or updated to increase readership/viewership. News producing giants such as CNN, BBC etc have taken great advantage of the Internet to disseminate news to a wider reading/viewing public. By transmitting news stories on the Internet, "breaking news" for example, can be accessed across the globe by the within a matter of minutes. Governments, institutions (public or private) and individuals, have also joined in the production of unstructured data by publishing pages about business, jobs, goods, and services. Text is easily to deployed, have low runtime impact on servers, and load faster. They use less hardware resources (e.g. server disk space) and much more information likely to be published—as compared to video, images, etc. The growing nature of the Internet has prompted researchers to mine the data therein to ascertain hidden trends and heuristics.

Classification based learning agents have been built to extract, mine and perform exploratory tasks on large text data in non-stationary platforms. Among approaches used, neural based methods have proved to be better candidates than their counterparts. Buabin (2011a, 2012) are typical examples of such systems. Buabin (2012) argues that, classification and recommender systems have major commonalities across their design. They both

1. Operate in search related domains,
2. Use search related mechanisms for finding optimal solutions to problems,
3. Use learning algorithms to extract knowledge from the training set into their structures and most importantly train on the exemplars.

With the aim of improving classifier performance, researchers have extracted features prior to performing classification tasks. Although implemented in many data mining systems, feature selection has taken many forms, some of which are statistical approaches (Perkins, et al., 2003; Hermes & Buhmann, 2000; Dash & Liu, 1997; Forman, 2003; Krishnapuram, et al., 2004; Piramuthu, 2004; Yang, 1999; Liu & Yu, 2005; Song,

et al., 2007; Weston, et al., 2003) and neural based approaches (Gabrilovich & Markovitch, 2004; Rakotomamonjy, 2003; Froehlich, 2002; Liu & Zheng, 2005; Taira & Haruno, 1999; Hardin, et al., 2004; Zhang, et. al., 2006; Huang & Wang, 2006; Chen, 2007; Warmuth, et al., 2003; Suna, 2004).

Biological or bio-inspired approaches have been adopted and implemented in problem domains where the solution space is extremely large, unknown, and complex. With its core principles deeply rooted in genetics, Genetic Algorithms have proven to perform better than other techniques in complex search related problems. A GA essentially comprises of a set of individual solutions or chromosomes (called the population) and some biologically inspired operators that create a new and potentially better population from an old one. It seeks to find optimal solutions for complex search problems with little knowledge of the solution space. In differentiating GAs from other techniques, GAs work by encoding the parameter set and simultaneously searching with multiple points. Whereas other approaches use deterministic rules, GAs use stochastic operators in finding optimal solutions. Feature selection is a typical search related problem and GAs tend to be good candidates. Although much work has been done in the domain of feature selection (i.e. with GAs) and are relevant, they are limited in terms of human nature, interactions and behaviour. In this book chapter, existing works in the area of feature selection and genetic based methods are investigated. In simple English language, a step by step approach to critiquing existing GA based techniques is adopted. Among others, new directions aimed at

1. Relating human nature, interactions and behaviour to search based problems,
2. Enhancing genetic operations,
3. Introducing further genetic operator restrictions, and
4. Making GA research appealing for active research.

23 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/hybrid-neural-genetic-architecture/76093

Related Content

Association Rule Mining Based HotSpot Analysis on SEER Lung Cancer Data
Ankit Agrawaland Alok Choudhary (2011). *International Journal of Knowledge Discovery in Bioinformatics (pp. 34-54).*
www.irma-international.org/article/association-rule-mining-based-hotspot/62300

Revealing the Origin and Nature of Drug Resistance of Dynamic Tumour Systems
Ricardo Santiago-Mozos, Imtiaz A. Khanand Michael G. Madden (2012). *Computational Knowledge Discovery for Bioinformatics Research (pp. 354-382).*
www.irma-international.org/chapter/revealing-origin-nature-drug-resistance/66720

PASS2: A Database of Structure-Based Sequence Alignments of Protein Structural Domain Superfamilies
Karuppiah Kanagarajadurai, Singaravelu Kalaimathy, Paramasivam Nagarajanand Ramanathan Sowdhamini (2011). *International Journal of Knowledge Discovery in Bioinformatics (pp. 53-66).*
www.irma-international.org/article/pass2-database-structure-based-sequence/73911

Human Creation and the Role of the Family: Theological Theory in Bioinformatics
Isahak Poghosyan, Tatul Manaseryanand Laura Aghajanyan (2022). *International Journal of Applied Research in Bioinformatics (pp. 1-7).*
www.irma-international.org/article/human-creation-and-the-role-of-the-family/282689

Search for Protein Sequence Homologues that Display Considerable Domain Length Variations
Eshita Mutt, Abhijit Mitraand R. Sowdhamini (2011). *International Journal of Knowledge Discovery in Bioinformatics (pp. 55-77).*
www.irma-international.org/article/search-protein-sequence-homologues-display/62301