A NOVEL ALGORITHM FOR FUZZY-GENETIC DISTRIBUTED DATA MINING

By

Hassan Ahmed Hassan M. Abounaser

A Thesis Submitted to the
Faculty of Engineering at Cairo University
in Partial Fulfillment of the
Requirements for the Degree of
DOCTOR OF PHILOSOPHY
in
Computer Engineering

FACULTY OF ENGINEERING, CAIRO UNIVERSITY
GIZA, EGYPT
2017

A NOVEL ALGORITHM FOR FUZZY-GENETIC DISTRIBUTED DATA MINING

By

Hassan Ahmed Hassan M. Abounaser

A Thesis Submitted to the
Faculty of Engineering at Cairo University
in Partial Fulfillment of the
Requirements for the Degree of
DOCTOR OF PHILOSOPHY
in
Computer Engineering

Under the Supervision of

Prof. Dr. Ihab El-Sayed Talkhan

Professor
Head of Computer Engineering Department
Fauculty of Engineering
Cairo University

Prof. Dr. Ahmed Fahmy Amin

Professor,
Computer Engineering Department
College of Engineering & Technology
Arab Academy for Science, Technology & Maritime Transport (AASTMT), Cairo

FACULTY OF ENGINEERING, CAIRO UNIVERSITY
GIZA, EGYPT
2017

A NOVEL ALGORITHM FOR FUZZY-GENETIC DISTRIBUTED DATA MINING

By

Hassan Ahmed Hassan M. Abounaser

A Thesis Submitted to the
Faculty of Engineering at Cairo University
in Partial Fulfillment of the
Requirements for the Degree of
DOCTOR OF PHILOSOPHY
in
Computer Engineering

Approved by the
Examining Committee

Prof. Dr. Ihab El-Sayed Talkhan , Thesis Main Advisor

Prof. Dr. Ahmed Fahmy, Member
Professor, College of Engineering & Technology, Arab Academy for Science, Technology & Maritime Transport (AASTMT), Cairo

---

Prof. Dr. Mohamed Zaki Abd El-Magid , Examiner
Professor, Faculty of Engineering, Al Azhar University

Prof. Dr. Amr Anwar Badr, External Examiner
Professor, Faculty of Computers and Information, Cairo University

FACULTY OF ENGINEERING, CAIRO UNIVERSITY
GIZA, EGYPT
2017

Engineer: Hassan Ahmed Hassan M. Abounaser
Date of Birth: 23 / 7 / 1979
Nationality: Palestinian
E-mail: hassan.abounasser@aast.edu
Phone: +201006620848
Address: 3078b, Zahraa- Nasr City, Cairo, Egypt

Registration Date: 1 / 10 / 2010
Awarding Date: 2 / 2 / 2017
Degree: Doctor of Philosophy
Department: Computer Engineering

Supervisors: Prof. Dr. Ihab El-Sayed Talkhan
Prof. Dr. Ahmed Fahmy Amin
Professor, College of Engineering & Technology- Arab Academy for Science, Technology & Maritime Transport (AASTMT), Cairo

Examiners: Prof. Dr. Amr Anwar Badr (External Examiner)
Professor, Faculty of Computers and Information, Cairo University
Prof. Dr. Mohamed Zaki Abd El-Magid (Examiner)
Professor, Faculty of Engineering, Al Azhar University
Prof. Dr. Ihab El-Sayed Talkhan (Thesis Main Advisor)
Prof. Dr. Ahmed Fahmy Amin (Member)

Title of Thesis: A Novel Algorithm For Fuzzy-Genetic Distributed Data Mining

Key Words:- Fuzzy Classification; Rule-base; Fuzzy Logic System (FLS); Genetic Algorithm (GA); Distributed Data Mining (DDM)

Summary:

A novel framework for a Parallel Fuzzy-Genetic Algorithm (PFGA) has been developed for classification and prediction over decentralized data sources as a main contribution to the scientific community. The model parameters are evolved using two nested genetic algorithms (GAs). The outer GA evolves the fuzzy sets whereas the inner GA evolves the fuzzy rules. During optimization, best rules are only distributed and exchanged among agents to construct the overall optimized model. Several experiments have been conducted and show that the developed model has good accuracy and more efficient in performance and comprehensibility of linguistic rules compared to some models implemented in KEEL software tool.

Acknowledgements

I would like gratefully to acknowledge all the following people who for various reasons were involved in contributing to this work, and for the help and time, they have given me over the work of this thesis.

First, I would like to thank my supervising committee, Prof. Ihab Talkhan and Prof. Ahmed Fahmy. This work would never have been successfully completed without their help, guidance and continuous support.

Second, special thanks go to my family for their love, patience, care and support during the period I spent working on this thesis. In particular, I would like to thank my parents for their endless encouragement, understanding and support.

Third, thanks also go to the staff of the Computer Engineering Department in AASTMT for the enlightening discussions and observations they made. In particular, I would like to thank Dr. Sherif Fadel for his valuable comments, suggestions and advice.

A final word of thanks is owed to my best friends, particularly, Dr. Mohamed Almoghalis, for his valuable support and advice.

Sincerely,

Hassan Ahmed Hassan M. Abounaser

Dedication

This thesis work is dedicated to my dear parents, who have always loved me unconditionally and whose good examples have taught me to work hard for the things that I aspire to achieve. This work is also dedicated to my lovely brothers and sisters who have been a constant source of support and encouragement during the challenges of life. I am truly thankful for having you in my life.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

ACO : Ant Colony Optimization
AdaBoost : Adaptive Boosting
AI : Artificial Intelligence
AIS : Artificial Immune System
ANN : Artificial Neural Network
ARM : Association Rules Mining
BOAT : Bootstrapped Optimistic Algorithm for Tree construction
CART : Classification And Regression Trees
CBR : Case-Based Reasoning
CI : Computational Intelligence
CoA : Center of Area
CoG : Center of Gravity
CUDA : Compute Unified Device Architecture
DDM : Distributed Data Mining
DHT : Distributed Hash Table
DM : Data Mining
EC : Evolutionary Computing
EM : Expectation-Maximization
FAM : Fuzzy Associative Memory
FGA : Fuzzy-Genetic Algorithm
FH-GBML : Fuzzy Hybrid Genetic-Based Machine Learning
FLS : Fuzzy Logic System
FS : Fuzzy Set
GA : Genetic Algorithm
GFS-RB-MF : Genetic-base Fuzzy Rule Base Construction and Membership Function tuning
GPU : Graphics Processing Unit
KB : Knowledge Base
KDD : Knowledge Discovery from Data
KEEL : Knowledge Extraction Based on Evolutionary Learning