



Ain Shams University
Faculty of Engineering
Computer and Systems Engineering Department

Machine Learning in e-Learning Environments

A Thesis

Submitted in Partial Fulfillment of the Requirements of the
Degree of Doctor of Philosophy in Electrical Engineering
Computer and Systems Engineering Department

Submitted by

Sally Sameh Abd El Ghaffar Attia

M. Sc., Electrical Engineering
(Computer and Systems Engineering)
Ain Shams University, 2004

Supervised by

Prof. Dr. Hani M. K. Mahdi

Prof. Dr. Hoda K. Mohamed

Cairo, Egypt

March, 2011



Faculty of Engineering

Examiners Committee

Name : Sally Sameh Abd El Ghaffar Attia
Thesis : Machine Learning in e-Learning Environments
Degree : Doctor of Philosophy in Electrical Engineering

Name, Title, and Affiliation	Signature
1. Prof. Dr. Hani Hagrass Director of the Computational Intelligence Centre Head of the Fuzzy Systems Group School of Computer Science and Electronic Engineering University of Essex, United Kingdom	
2. Prof. Dr. Abdel-Moneim A. Wahdan Computer and Systems Engineering Department Faculty of Engineering, Ain Shams University, Cairo, Egypt
3. Prof. Dr. Hani. M. Kamal Mahdi Computer and Systems Engineering Department Faculty of Engineering, Ain Shams University, Cairo, Egypt (Supervisor)
4. Prof. Dr. Hoda Korashi Mohamed Computer and Systems Engineering Department Faculty of Engineering, Ain Shams University, Cairo, Egypt (Supervisor)

Date: 24 / 3 / 2011

Abstract

Sally Sameh Abdel Ghaffar Attia

Machine Learning in e-Learning Environments

Doctor of Philosophy dissertation

Ain Shams University, 2011

E-learning environments have evolved over the past decades due to the availability of high speed networking and Internet. E-learning has become a popular teaching method in recent years. One of the modes of e-learning is the blended learning where learners can view teaching materials asynchronously from a teaching website and collaborate with their peers at their own pace, while providing for necessary face-to-face explanation, discussion, and physical operation in the classroom. The learners consequently require support while using web based tools, but it is very difficult and time consuming for instructors to track and assess all the activities performed by all learners on the e-learning tools.

At the same time, the intelligent agent paradigm is growing to be a continuously evolving and expanding area. It is a perfect technology to be used in the web-based learning field since it fits the dynamic and distributed nature of e-learning where all of the teachers and students are independent entities, and they need to collaborate with each other to help the teachers teach better and students learn better.

Thus the main objective of this research is the development of a Multi-Agent System for Collaborative E-learning (MASCE) as a “proof of concept” which will use machine learning techniques for data mining of the huge amount of data collected for the reasoning of agents.

Three different machine learning techniques are being used for data mining of data collected to find best candidate helpers. These are Rough sets, Decision trees and finally an integrated approach using rough sets and entropy. The results obtained were in the form of decision rules which were used by the multi-agent system to recommend best potential helpers for a student who needs help in a particular topic. The results are compared to find which one is the most suitable for this particular application as it is known that there is no one universal machine learning solution for all the domains of applications. It was found that the integrated rough sets and entropy approach gave simpler, shorter rules with higher coverage and better accuracy at the same time.

In conclusion, supported by machine learning techniques, the researcher was able to enhance the application of multi-agent system to collect, manage, share, and analyze data and extract useful information in a real-time, dynamic collaborative e-learning environment. The results show that this is quite a promising approach that successfully combines machine learning techniques with agent technology in e-learning systems in order to provide higher quality services towards the end users of e-learning systems (both students and instructors).

Keywords:

E-learning, Collaborative Learning, Multi-Agent Systems, Machine Learning, Data Mining, Rough Sets, Decision Trees

Acknowledgements

I owe a special acknowledgment to Prof. Dr. Hani Mahdi for giving me a lot of his time during the years of preparing this thesis. I could never had done it without his support, technical advice and suggestions, thorough reading of all my work.

I am deeply grateful to Prof. Dr. Hoda Korashi for supporting me during the research and giving me her valuable comments and advice and helping me to finish the thesis in time.

I will never be able to thank my parents enough for supporting me during all my life. They were the only reason that gave me hope to finish this work. I tried to accomplish it to make them proud of me. I would like to dedicate this work to them and to my sister also.

I am very grateful to my husband for helping me whenever I needed his help and for bearing that I was always busy during the last period. Finally, this work is dedicated to my lovely two kids, Youssef and Nour, who used to pray for Allah that I finish my thesis. I want to compensate them all for being busy and nervous especially during the last period.

Statement

This dissertation is submitted to Ain Shams University for the degree of Doctor of Philosophy in Computer and Systems Engineering Department.

The work included in this thesis was carried out by the author at Computer and Systems Engineering Department, Ain Shams University.

No part of this thesis has been submitted for a degree or qualification at other university or institution.

Date : 24/ 3 /2011

Signature :

Name : Sally Sameh Abd El Ghaffar Attia

Table of Contents

Chapter One: Introduction	1
1.1 Research Motivation	1
1.2 Research Objective	3
1.3 Thesis outline	5
Chapter Two: Machine Learning	6
2.1 Definition of Machine Learning	6
2.2 Machine Learning Categories	8
2.2.1 Supervised Learning	8
2.2.2 Unsupervised Learning	10
2.2.3 Reinforcement Learning	10
2.3 Learning Tasks	10
2.3.1 Classification	11
2.3.2 Clustering	11
2.3.3 Association	13
2.3.4 Regression, Interpolation, and Density Estimation	14
2.4 Machine Learning and Data Mining	14
2.5 Fields of Applications	16
2.6 Review on some Learning Algorithms	17
2.6.1 Decision Trees	17
2.6.2 Neural Networks	29
2.6.3 Bayesian Networks	33
2.6.4 Genetic Algorithms	34

2.6.5	Rough Sets	36
2.6.6	Fuzzy Logic	47
2.7	Challenges to Machine Learning	50
Chapter Three: E-Learning		51
3.1	Distance Education vs. Traditional Education	51
3.2	Different Terminologies	52
3.3	History of E-Learning	54
3.4	E-Learning Components	55
3.5	Transmission Modes	56
3.6	Technologies for Delivering E-Learning	60
3.7	Different Views on E-Learning	61
3.8	Spread of Distance Education	62
3.9	Current State of Art	63
3.10	Semantic Web and E-Learning	64
3.11	E-Learning 2.0	65
3.12	Challenges to E-learning	67
Chapter Four: Agent Technology		70
4.1	Definition of Agent	70
4.2	Properties of Agents	71
4.3	Multi-Agent System	72
4.4	Agent Design (Micro Perspective)	74
4.4.1	Rational Agents	74
4.4.2	PEAS	75
4.5	Agent Types	76

4.5.1	Simple Reflex Agents	76
4.5.2	Model-Based Reflex Agents	76
4.5.3	Goal-Based Agents	77
4.5.4	Utility-Based Agents	78
4.5.5	Learning Agents	79
4.6	Synthesizing Agents	79
4.7	Society Design (Macro Perspective)	81
4.7.1	Mechanism Design	81
4.7.2	Auctions	82
4.7.3	Negotiation	84
4.7.4	Argumentation	85
4.8	Agent Communication	86
4.8.1	Speech Acts	86
4.8.2	Agent Communication Languages	87
4.9	Cooperation	88
4.9.1	Types of Agents with respect to Cooperation	88
4.9.2	Coherence and Coordination	89
4.9.3	Handling Inconsistency	91
4.10	Mobile Agents	92
4.11	Agent-Oriented Analysis and Design Methodologies	93
4.12	Agent Organizations	95
4.12.1	Hierarchical Organization	95
4.12.2	Flat Organization	96
4.12.3	Subsumption Organization	96

4.12.4	Modular Organization	97
4.13	Agent Platforms	98
4.14	When is an Agent-Based Solution Appropriate?	98
4.15	Applications of Multi-Agent Systems	99
4.16	FIPA Standards and Models	100
4.16.1	FIPA Agent Reference Model	100
4.16.2	Communication Architectures	101
4.17	Agent Life Cycle according to FIPA	103
4.18	Multi-Agents Systems in E-Learning	104
4.18.1	Previous Work	104
4.18.2	Design Issues of Agents in Education	107
4.18.3	Challenges for MAS in E-learning	111
4.19	Multi-Agent Systems and Machine Learning	112
4.19.1	Machine Learning Categories in Multi-Agent Systems	113
4.19.2	Combining Learning with Recall in MAS	114
4.19.3	Types of Multi-Agent Learning	115
4.19.4	Previous Work	116
Chapter Five: Proposed Architecture for MASCE		118
5.1	Need for MASCE	118
5.2	MASCE System Requirements	119
5.3	Approaches for Building MASCE	120
5.4	System Analysis	123
5.4.1	Brief External Use Case (Step 1)	123
5.4.2	Detailed External Use Cases (Step 2)	129

5.4.3	Structuring Goals (Step 3)	129
5.5	System Design	133
5.5.1	Internal Use Cases (Step 4)	135
5.5.2	Sequence Diagram (Step 5)	137
5.5.3	Agent Belief List (Step 6)	138
5.5.4	BDI Agent Cards (Step 7)	143
5.6	Student Agency	144
5.7	Instructor Agency	145
5.8	Assistant Agency	146
5.9	Feasible Organizations for MASCE	147
5.10	Parameters for Tracking Behavior	150
Chapter Six: DM Techniques used within MASCE		153
6.1	Three Levels for Knowledge Diffusio in MAS	154
6.2	Rough Sets Mining in MASCE	155
6.3	Decision Trees Mining in MASCE	162
6.4	Integrated RS and Entropy (RSE) Mining in MASCE	163
6.5	Evaluating Machine Learning Techniques	166
6.5.1	Repeated Cross-Validation	167
6.5.2	Leave-One-Out	169
6.5.3	Bootstrapping	169
6.5.4	Cost of Evaluation	170
6.5.5	Confusion Matrix	170
6.6	Comparing Different Learning Algorithms	172
6.7	Comparison between Rough Sets and Fuzzy Logic	174

6.8	Comparative Study	177
Chapter Seven: Implementation		178
7.1	Database	178
7.2	Agent Structure Implementation	181
7.2.1	Hierarchy of Packages	181
7.2.2	The Agent class	185
7.2.3	Student Agent Class	188
7.2.4	Manager Student Agent Class	190
7.2.5	Instructor Agent Class	190
7.2.6	Manager Instructor Agent Class	191
7.2.7	Assistant Agent Class	191
7.2.8	Manager Assistant Agent Class	193
7.3	Agent Behaviors	193
7.4	Agent Communication	195
7.5	Ontology	196
7.6	Administrator Module	198
7.7	Quizzes Implementation	198
7.8	Email Service	199
7.9	Teaching Materials Service	200
7.10	Machine Learning Module	200
7.10.1	Expert Opinion	201
7.10.2	Student's Tracking System	202
7.10.3	Finding Best Helpers	202
7.10.4	Help Sessions	205

7.10.5	Peer Evaluation	206
7.10.6	Rough Sets Algorithm Implementation	208
7.10.7	Decision Tree Algorithm Implementation	209
7.10.8	Integrated RS and Entropy Algorithm Implementation	209
7.10.9	Evaluating Performance	210
7.11	MASCE Users' Guides	211
7.11.1	Administrator User's Guide	212
7.11.2	Instructor User's Guide	213
7.11.3	Student's User Guide	216
7.11.4	Assistant User's Guide	220
Chapter Eight: Results		222
8.1	Results of Rough Sets Classifier	223
8.2	Results of Decision Trees Classifier	225
8.3	Results of Integrated Rough Sets and Entropy Classifier	225
8.4	Evaluating Extracted Rules	227
8.4.1	Confusion Matrix using Rough Sets Classifier	228
8.4.2	Confusion Matrix using Decision Trees Classifier	228
8.4.3	Confusion Matrix using Integrated Rough Sets and Entropy Classifier	228
8.4.4	Comparison of Performance Measures for Three Classifiers	229
Chapter Nine: Conclusions and Future Work		231
Appendix A		235
Publications		245
References		247

List of Tables

Table 2.1 The Weather Data [9]	20
Table 3.1 The five models of e-learning [38]	59
Table 5.1 External Brief Use Case [108].....	125
Table 5.2 External Brief Use Case Divided into Sub Goals [108].....	127
Table 5.3 Help Student Collaborate Detailed External Use Case [108]	130
Table 5.4 Help Student Collaborate Internal Use Case [108]	135
Table 6.1 Confusion Matrix Parameters.....	157
Table 6.2 Different Performance Measures.....	171
Table 8.1 Snapshot of MASCE decision table filled with random data	224
Table 8.2 Extracted Decision Rules from Rough Set Classifier.....	225
Table 8.3 Extracted Decision Rules from Decision Tree Classifier.....	226
Table 8.4 Extracted Decision Rules from Integrated RSE approach	227
Table 8.5 Confusion Matrix using Rough Sets Classifier	228
Table 8.6 Confusion Matrix using Decision Trees Classifier	228
Table 8.7 Confusion Matrix using Integrated Approach.....	228
Table 8.8 Comparison of Performance Measures for three classifiers...	229

List of Figures and Illustrations

Figure 2.1 Supervised Learning [11].....	9
Figure 2.2 Basic Classification Problem	11
Figure 2.3 Various Shapes of Clusters in 2-D Space [12].....	13
Figure 2.4 Tree Stumps for the Weather Data [9]	21
Figure 2.5 Expanded Tree Stumps for Weather Data [9].....	24
Figure 2.6 Decision Tree for the Weather Data [9].....	25
Figure 2.7 Scheme of the perceptron [12]	31
Figure 2.8 Flow chart of neural networks algorithm [12]	32
Figure 2.9 Simple Bayesian Network for Weather data [9]	33
Figure 2.10 A Pictorial Representation of a Rough Set [15].....	40
Figure 2.11 Process Map and the Main Steps of RS Analysis [15].....	42
Figure 3.1 Traditional Education [36]	51
Figure 3.2 Distance Education [36].....	52
Figure 3.3 The three components of e-learning [38].....	55
Figure 3.4 Learner to Learner Connections [36]	66
Figure 3.5 Each Student can act as a teacher [36].....	67
Figure 3.6 Learning Technology Systems Architecture (LTSA) by IEEE 1484 [43]	71
Figure 4.1 An Agent in its environment [46]	73
Figure 4.2 Related research areas to Agent Technology	74