Adaptive Business Intelligence (ABI)

Technology Curricular Unit (UTC) proposal for the MAP-I PhD Program

A – Programmatic Component

1. Motivation

Nowadays, business organizations are increasingly moving towards decision-making processes that are based on information. **Business Intelligence (BI)** is an umbrella term that includes methodologies, architectures, tools, applications and technologies to enhance managerial decision making [1]. The goal of BI is to: access data from multiple sources, transform these data into information and then into knowledge.

Very recently, a new trend emerged in the marketplace called **Adaptive Business Intelligence (ABI)** [2]. Besides transforming data into knowledge, ABI also includes the decision-making process. BI systems often include elements of databases, data warehouses and data Mining [1], while ABI systems also encompass forecasting [3] and optimization [4], in order to enhance adaptability. In effect, adaptability is a vital component of any intelligent system and this issue is expected to gain popularity in the next years. The final ABI goal is to use computer systems that can adapt to changes in the environment, solving complex real-world problems with multiple objectives, in order to aid business managers to make better decisions, increasing efficiency, productivity and competitiveness.

Although being a recent field, the topics covered by ABI (i.e. data mining, forecasting, modern optimization and adaptive systems) have a large research community, with several prestigious international scientific journals (e.g. Data Mining Knowledge Discovery, Decision Support Systems, Machine Learning, IEEE Trans. Neural Networks, International Journal of Forecasting, IEEE Trans. Evolutionary Computation, Journal of Heuristics, Applied Soft Computing) and conferences (e.g. ACM KDD, ACM CIKM, ACM ICIS, IEEE ICDM, IEEE IJCNN, IEEE CEC) available.

There are also several international examples of Computer Science PhD programs that include ABI topics, such as:

- Carnegie Mellon University (CMU), USA:
  - **Ph.D. Program in Computer Science** (machine learning, optimization);
  - **Ph.D. Program in Computation, Organizations and Society (**machine learning and advanced artificial intelligence concepts);
  - **Ph.D. Program in Machine Learning** (data mining, machine Learning, database management systems, optimization);
- Standford University, USA:
  - **Ph.D. in Computer Science** (databases, data mining, machine learning, modern optimization)
- Berkeley University of California, USA:
  - **Ph.D. in Computer Science**, specialization ins Communication, Computation and Statistics (database management systems, data mining, decision support);
- University of Texas at Austin, USA:
  - **Ph.D. in Computer Science** (database systems, data mining, machine learning, neural networks, reinforcement learning);
- Auburn University, USA:
  - **Ph.D. in Computer Science and Software Engineering** (database Systems,

artificial intelligence, machine learning, evolutionary computation and modern optimization, neural networks)

The proposed ABI unit was also part of the **2008/09 course edition of the MAP-I** doctoral program (http://www.map.edu.pt/i/2008/courses).

2 Objectives and Learning Outcomes

To learn about the basic ABI concepts, including: characteristics of complex business problems, BI and ABI, data mining, prediction, modern optimization and adaptability;

To master the state of the art methods and models (e.g. Decision Trees, Neural Networks, Support Vector Machines, Learning Classifier Systems, Evolutionary Algorithms, Relational Learning Algorithms), methodologies (e.g. CRISP-DM, SEMMA), and tools (e.g. R, WEKA, SAS, Evolution Machine, SCS-C);

To perform a review essay over an advanced research ABI topic (e.g. mining complex data; forecasting using sophisticated machine learning methods; multi-objective evolutionary optimization; adaptive decision-making);

To apply ABI in real-world applications (e.g. Finance, Economy, Marketing).

3 Detailed Program

1 - Introductory ABI concepts: characteristics of complex business problems, BI and ABI, ABI case studies, data mining, prediction, optimization and adaptability.
2 - Modern Learning and Optimization methods for ABI: supervised learning (e.g. decision trees, neural networks, support vector machine, learning classifier systems, multi-relational learning algorithms), clustering (e.g. hierarchical and relational clustering), inductive logic programming, heuristic search (e.g. hill-climbing, tabu-search, evolutionary computation).
3 - Data mining and Forecasting for ABI: CRISP-DM, knowledge understanding, data selection, data preprocessing; data mining, data interpretation and knowledge maintenance, univariate and multivariate forecasting, time series, conventional forecasting methods.
5 - Exploration of ABI tools (e.g. R, WEKA, RapidMiner, SAS, Evolution Machine, SCS-C) when applied to real-world problems (e.g. Finance, Economy, Marketing).

4 Teaching Methodology and Evaluation

Four teaching methodologies will be applied:
1 Lecture based exposition of key ABI issues.
2 Active learning (e.g. think-pair-share, in-class teams [5]).
3 Case based learning.
4 Project based learning;

Evaluation will include two elements:

A - review of key ABI research article(s) (20%);

B - an ABI project that describes the application of ABI tools to real-world datasets (80%).

5 Bibliography

Cited references:

[1] E. Turban, R. Shard, J. Aronson and D. King, Business Intelligence – A Managerial Approach, Pearson Prentice-Hall, New Jersey, USA, 2008.

[2] Z. Michalewicz, M. Schmidt, M. Michalewicz and C. Chiriac, Adaptive Business Intelligence, Springer-Verlag, Leipzig, Germany, 2007.

[3] S. Makridakis, S. Wellwright and R. Hyndman, Forecasting: Methods and Applications, John Wiley & Sons, New York, USA, 1998.

[4] S. Ruhul, M. Masoud and X. Yao (Eds.), Evolutionary Optimization, Springer-Verlag, 2002.

[5] D. Johnson, R. Johnson and K. Smith, Active Learning: Cooperation in the College Classroom, 2nd edition, Edina, Interaction Book Company, 1998.

Additional references:

[6] Michalewicz, Z. and Fogel, D. B. (2000). How to Solve It: Modern Heuristics. 2nd edition, Springer.

[7] Michalewicz, Z., Schmidt, M., Michalewicz, M. and Chiriac, C. (2005). Case study: an intelligent decision support system. In IEEE Intelligent Systems, 20(4):44-49.

[8] Michalewicz, Z., Schmidt, M., Michalewicz, M. and Chiriac, C. (2007). Adaptive Business Intelligence: Three Case Studies, In Evolutionary Computation in Dynamic and Uncertain Environments, pp. 179-196, Springer.

[9] Turban, E., Sharda, E., Delen, D. (2011). Decision Support and Business Intelligent Systems. Prentice Hall (BGUM).

B Lecture Team

1. Summary

The lecture team includes the three members of the previous ABI unit that was lectured in the MAP-I 2008/09 edition. **Manuel Filipe Santos (MFS)** and **Paulo Cortez (PC)** belong to the Business Intelligence Group (BIG) of the Algoritmi R&D Centre, University of Minho. Both research on Adaptive Business Intelligence, Decision Support Systems and Data Mining. MFS performed his PhD in Distributed Learning Classifier Systems, while PC performed his PhD in Forecasting, Neural Networks and Evolutionary Optimization. **Rui Camacho (RC)** is from the **Laboratory of Artificial Intelligence and Decision Support (LIAAD) R&D centre, University of Porto. RC researches in** Inductive Logic Programming and Relational Data Mining. For more details, see the CVs in section B.3.

The team is willing to write didactic texts related to this unit. If necessary and funds are available, one of the team members can visit CMU for the curricular unit accreditation.

2. Coordinator

Manuel Filipe Santos (MFS)

3. CVs

3.1 Manuel Filipe Santos

Biography: Manuel Filipe Santos received his Ph.D. in Computer Science (Artificial Intelligence) from the University of Minho (UMinho), Portugal, in 2000. He is auxiliary professor at the Department of Information Systems, UMinho, teaching undergraduate and graduate classes of Business Intelligence and Intelligent Data Analysis. He his also researcher at the Business Intelligence Group (big.dsi.uminho.pt) of the R&D Algoritmi Centre, with the current research interests:

- Business Intelligence and Decision Support Systems;
- Data Mining and Machine Learning (Learning Classifier Systems);
- Grid Data Mining.

Relevant publications in the last 5 years:

He is co-author of more than sixty publications in international conferences and journals. His most significant publications for the field in the last 5 years are:

[1] Á. Silva, P. Cortez, M.F. Santos, L. Gomes and J. Neves. Rating organ failure via adverse events using data mining in the intensive care unit. In Artificial Intelligence in Medicine, Elsevier, In Press (ISI impact factor 1.882).

[2] Gago, P. Silva, A., Santos, M., Adaptive Decision Support for Intensive Care, Lecture Notes in Artificial Intelligence, Progress in Artificial Intelligence, 13 th Portuguese Conference on Artificial Intelligence, EPIA 2007, Guimarães, Portugal, December 2007, Neves, J., Santos, M.F., Machado, J. (Eds.), LNAI 4874, Springer, Germany, 2007, ISBN 978-3-540-77000-8 (ISI proceedings).

[3] M.F. Santos, P. Cortez, J. Pereira and H. Quintela. Corporate Bankruptcy Prediction using Data Mining Techniques1. In A. Zanasi, C. Brebbia and N. Ebecken (Eds.), Data Mining VII - Data, Text and Web Mining, and their Business Applications, WIT Transactions of Information and Communication Technologies, vol. 37, pp. 349-357, 2006. WIT Press, UK, ISBN:1-84564-178-7, ISSN:1743-4463 (ISI proceedings).

[4] Á. Silva, P. Cortez, M.F. Santos, L. Gomes and J. Neves. Mortality assessment in intensive care units via adverse events using artificial neural networks. In Artificial Intelligence in Medicine, Elsevier, 36 (3): 223-234, 2006 (ISI impact factor 1.882).

Neves, J., Santos, M.F., Machado, J., (Eds), Lecture Notes in Artificial Intelligence, Progress in Artificial Intelligence, 13 th Portuguese Conference on Artificial Intelligence, EPIA 2007, Guimarães, Portugal, December 2007, Neves, J., Santos, M.F., Machado, J. (Eds.), LNAI 4874, Springer, Germany, 2007, ISBN 978-3-540-77000-8 (ISI Web of Science, Scopus)

[5] Neves, J., Santos, M.F., Machado, J., (Eds), New Trends in Artificial Intelligence, December 2007, Portugal, ISB 13978-989-95618-0-9. Supported byAPPIA, IEEE, ACM, AAAI, ECCAI.

[6] Gago, P. Santos, F., Evaluating hybrid ensembles for Intelligent Decision Support for Intensive Care, Supervised and Unsupervised Ensemble Methods and Their Applications, Springer-Verlag, Germany, 2009.

[7] Santos, M., Amado, J., Personalization of information delivery based on web mining and DNA classification, Data Mining X: Data Mining, Protection, Detection and other Security Technologies, WIT Transactions on Information and

Communication Technologies, WIT Press, Southampton, UK, 2009, ISSN: 1743-3517.

Participation in R&D projects in the last 5 years:

He participated in various R&D projects, being Principal Investigator of 2 projects, namely:

INTCARE – Intelligent Decision System for Intensive Care Medicine
FCT - PTDC/EIA/72819/2006

GRIDCLASS - Learning Classifiers Systems for Grid Data Mining
FCT - GRID/GRI/81736/2006

Supervision of Graduate Students:

Supervised 15 MSc theses and 3 PhD thesis. Currently supervises 6 PhD students, 1 MSc student and 1 post-doc student.

Other relevant topics of his CV:

- Co-organized the EPIA 2007 – 13th Portuguese Conference on Artificial Intelligence;
- Co-organized the Knowledge Discovery and Business Intelligence - KDBI 2009 thematic track of EPIA;
- Acted as external examiner of 32 MSc and 16 PhD theses;
- Participated as member of 26 Program Committees of International Conferences;
- Pereira, M.L., Neves, J., Machado, J., Santos, M., A. (Eds.), International Journal of Reasoning-based Intelligent Systems (IJRIS), ISSN (Online): 1755-0564 - ISSN (Print): 1755-0556, Inderscience Publushers.

3.2 Paulo Cortez

Biography: Paulo Cortez (http://www3.dsi.uminho.pt/pcortez) received his Ph.D. in Computer Science from the University of Minho (UMinho), Portugal, in 2002. He is lecturer at the Department of Information Systems, UMinho, teaching undergraduate and graduate classes of Computer Programming and Business Intelligence. He his also researcher at the Business Intelligence Group (http://big.algoritmi.uminho.pt/) of the R&D Algoritmi Centre, with the current research interests:

- Business Intelligence and Decision Support Systems;
- Data Mining and Machine Learning;
- Neural Networks and Evolutionary Computation;
- Forecasting.

Relevant publications in the last 5 years:

He is co-author of more than sixty publications in international conferences and journals (e.g. published by IEEE, Elsevier or Springer). His relevant publications in the last 5 years are:

[1] M. Rocha, P. Sousa, P. Cortez and M. Rio. Quality of Service Constrained

Routing Optimization using Evolutionary Computation, In Applied Soft Computing, Elsevier, In press. ISSN: 1568-4946. (ISI impact factor 1.909, Scopus)
[2] P. Cortez, Sensitivity Analysis for Time Lag Selection to Forecast Seasonal Time Series using Neural Networks and Support Vector Machines, In Proceedings of the IEEE International Joint Conference on Neural Networks IJCNN, Barcelona, Spain, 2010, in press (ISI proceedings).
[3] P. Cortez, Data Mining with Neural Networks and Support Vector Machines using the R/rminer Tool, In P. Perner (Ed.), Advances in Data Mining, Proceedings of 10th Industrial Conference on Data Mining, Berlin, Germany, Lecture Notes in Computer Science, Springer, July, 2010, in press (Scopus).
[4] P. Cortez, A. Cerdeira, F. Almeida, T. Matos and J. Reis. Modeling wine preferences by data mining from physicochemical properties. In Decision Support Systems, Elsevier, 47(4):547-553, 2009. ISSN: 0167-9236. (ISI impact factor 1.873, Scopus, ACM Portal)
[5] R. Costa, N. Cachulo and P. Cortez. An Intelligent Alarm Management System for Large-Scale Telecommunication Companies. In L. Lopes et al. (Eds.), Progress in Artificial Intelligence, 14th Portuguese Conference on Artificial Intelligence (EPIA'2009), Lecture Notes in Artificial Intelligence 5816, pp. 386-399, Aveiro, Portugal, October, 2009. Springer. ISBN-10 3-642-04685-1. (34% acceptance rate, ISI proceedings, Scopus, DBLP)
[6] P. Cortez, Clotilde Lopes, P. Sousa, M. Rocha and M. Rio. Symbiotic Data Mining for Personalized Spam Filtering. In Proceedings of the IEEE/WIC/ACM International Conference on Web Intelligence (WI-09), pp. 149-156, Milan, Italy, September, 2009. IEEE, ISBN: 978-0-7695-3801-3. (16% acceptance rate for full paper, DBLP, IEEE Xplore, ACM Portal)
[7] Á. Silva, P. Cortez, M.F. Santos, L. Gomes and J. Neves. Rating organ failure via adverse events using data mining in the intensive care unit. In Artificial Intelligence in Medicine, Elsevier, 43 (3): 179–193, 2008 (ISI impact factor 1.882).
[8] H. Quintela, M. F. Santos and P. Cortez. Real-Time Intelligent Decision Support System for Bridges Structures Behavior Prediction. In J. Neves, M. F. Santos and J. Machado (Eds.), Progress In Artificial Intelligence, 13th EPIA Portuguese Conference on Artificial Intelligence, Lecture Notes in Computer Science 4874, pp. 124-132, Guimarães, Portugal, December, 2007. Springer, ISBN 978-3-540-77000-8 (ISI proceedings).
[9] M. Rocha, P. Cortez and J. Neves. Evolution of Neural Networks for Classification and Regression. In Neurocomputing, Elsevier, 70 (16-18):2809-2816, October, 2007 (ISI impact factor 0.790).
[10] P. Cortez, M. Rio, P. Sousa and M. Rocha. Topology Aware Internet Traffic Forecasting using Neural Networks. In J. de Sá et al. (Eds.), Artificial Neural Networks - ICANN, 17th International Conference, Lecture Notes in Computer Science 4669, pp. 445-454, Porto, Portugal, September, 2007. Springer, ISBN: 978-3-540-74693-5 (ISI proceedings).
[11] D. Duque, H. Santos and P. Cortez. Prediction of Abnormal Behaviors for Intelligent Video Surveillance Systems. In Proceedings of the 2007 IEEE Symposium on Computational Intelligence and Data Mining (CIDM 2007), pp. 362-367, Honolulu, USA, April, 2007. IEEE, ISBN: 1-4244-0698-6 (ISI proceedings).
[12] P. Cortez, M. Rio, M. Rocha and P. Sousa. Internet Traffic Forecasting using Neural Networks. In Proc. of the 2006 IEEE Int. Joint Conference on Neural Networks, Vancouver, Canada, pp. 4942-4949, July, 2006, IEEE (ISI proceedings).
[13] M.F. Santos, P. Cortez, J. Pereira and H. Quintela. Corporate Bankruptcy

Prediction using Data Mining Techniques1. In A. Zanasi, C. Brebbia and N. Ebecken (Eds.), Data Mining VII - Data, Text and Web Mining, and their Business Applications, WIT Transactions of Information and Communication Technologies, vol. 37, pp. 349-357, 2006. WIT Press, UK, ISBN:1-84564-178-7, ISSN:1743-4463 (ISI proceedings).

Participation in R&D projects in the last 5 years:

He participated in 7 R&D projects, being Principal Investigator of 2 projects, namely:

- Principal Investigator (PI) of the project PTDC/EIA/64541/2006 - SPAM Telescope Miner: worldwide unsolicited email detection using data mining techniques, financed by FCT, from January 2008 to December 2010. Budget: 70000 euros.
- Principal Investigator (PI) of the Portuguese team in the project B-53/05 - Internet Congestion Control Using Neural Networks, financed by the Portuguese Association of Rectors (CRUP)/British Council, from April 2005 to April 2006. This project involved UMINHO and University College London.

Supervision of Graduate Students:

Supervised 1 PhD thesis and 9 MSc thesis. Currently he is supervising 3 PhD students.

Other relevant topics of his CV:

- Associate Editor of the Neural Processing Letters journal (Springer, ISI).
- Reviewer of several ISI journals (e.g. Expert Systems, Data & Knowledge Engineering, Neurocomputing) and conferences (e.g. IJCNN, DMin, ICDM, PPKDD).
- Co-organized the Knowledge Discovery and Business Intelligence - KDBI 2009 thematic track of EPIA;
- Invited lecturer in the International Summer School of Neural Networks in Classification, Regression and Data Mining (2003-06;2008; 2010).
- Acted as external examiner of 8 MSc and 3 PhD thesis.
- Author of the open source RMiner library, which facilitates the use of Data Mining applications in R (http://www3.dsi.uminho.pt/pcortez/rminer.html).

3.3 Rui Camacho

Biography: Rui Camacho received his Ph.D. in Electrical Engineering and Computers from the University of Porto (UP), Portugal, in 2000. He is Associate professor at the Informatics Engineering Department, Faculty of Engineering, UP, teaching undergraduate and graduate classes of Machine Learning and Data Mining. He his also researcher at the Laboratory of Artificial Intelligence and Decision Support (LIAAD), with the current research interests:

- Inductive Logic Programming;
- Multi-Relational Data Mining and Machine Learning;
- Applications of Bioinformatics.

Relevant publications in the last 5 years:

His most significant publications for the field in the last 5 years are:

[1] Nuno A. Fonseca, Ashwin Srinivasan, Fernando Silva, and Rui Camacho, " Parallel ILP for Distributed-Memory Architectures", Machine Learning journal, Vol. 74, Number 3, pp. 257-279, March 2009

[2] N. Fonseca, V. S. Costa, R. Rocha, R. Camacho, F. Silva, ``Improving the Efficiency of ILP Systems'', journal of Software: Practice and Experience, Vol. 39, Issue 2, pp. 189-219, Fev. 2009

[3] N. Fonseca, R. Camacho, R. Rocha, V. S. Costa, ``Compile the hypothesis space: do it once, use it often'', Fundamenta Informaticae, Special Issue on Multi-Relational Data Mining(89):45-67, 2008.

[4] editor with, João Gama, Luís Torgo, Alípio Jorge and Pavel Brazdil, of the ECML 2005 Proceedings.

[5] editor with, João Gama, Luís Torgo, Alípio Jorge and Pavel Brazdil, of the PKDD 2005 Proceedings.

[6] editor, with Ashiwin Srinivasan and Ross King, of the proceedings of the ILP 2004.

[7] Nuno A. Fonseca, Fernando Silva, Rui Camacho and Ashwin Srinivasan, Parallel ILP for Distributed-Memory Architecture, in Machine Learning journal (to appear in 2008).

[8] Ruy Ramos and Rui Camacho, Distributed Generative Data Mining, 7th Industrial Conference on Data Mining (ICDM~2007), Springer-Verlag LNAI 4597, pp 307-317, Leipzig/Germany, July 14-18, 2007.

Participation in R&D projects in the last 5 years:

He is Principal Investigator of 1 project:

- ILP-Web-Service: An Inductive Logic Programming based Web service.

Participates(ed) in the following ones:

- HORUS: Horn Representations of Uncertainty in Systems (PTDC/EIA-EIA/100897/2008);
- STAMPA - Sophisticated TAbling Mechanisms for Prolog and their Applications (PTDC/EIA/67738/2006);
- P-found (GRID/GRI/81809/2006);
- Searching for high level rules in protein folding and unfolding: from amyloid diseases to protein structure prediction

Supervision of Graduate Students:

Supervised 3 MSc theses and 1 PhD thesis. Currently supervises 6 PhD students and 2 post-doc students.

Other relevant topics of his CV:

- Co-organized the ILP 2004 – International Conference on Inductive Logic Programming;
- Co-organized the ECML/PKDD 2005 – Europena Conference on Machine Learning and the European Conference on Principles and Practice of Knowledge Discovery in Databases;
- Was guest editor of the Machine Learning journal Vol. 64, N. 1/2/3, 2006;
- Belongs to the Editorial Board of the Journal of Computational Intelligence in Bioinformatics (JCIB).