Integrating Data Mining and Agent Based Modeling and Simulation

Omar Baqueiro[1], Yanbo J. Wang[2], Peter McBurney[3] and Frans Coenen[3]

[1] Department of Structural Development of Farms and Rural Areas,
Institute Of Agricultural Development in Central and Eastern Europe (IAMO),
Theodor-Lieser-Str.2, Halle (Saale), D-06120, Germany
baqueiro@iamo.de
[2] Information Management Center, China Minsheng Banking Corp. Ltd.,
Room 606, Building No. 8, 1 Zhongguancun Nandajie, Beijing, 100873, China
wangyanbo@cmbc.com.cn
[3] Department of Computer Science, University of Liverpool,
Ashton Building, Ashton Street, Liverpool, L69 3BX, UK
{Mcburney, Coenen}@liverpool.ac.uk

Abstract. In this paper, we introduce an integration study which combines Data Mining (DM) and Agent Based Modeling and Simulation (ABMS). This study, as a new paradigm for DM/ABMS, is concerned with two approaches: (i) applying DM techniques in ABMS investigation, and inversely (ii) utilizing ABMS results in DM research. Detailed description of each approach is presented in this paper. A conclusion and the future work of this (integration) study are given at the end.

Keywords: Agents, Agent Based Modeling and Simulation, Data Mining, KDD (Knowledge Discovery in Databases) Process.

1 Introduction

Two promising fields of current studies in computer science are Data Mining (DM) and Agent Based Modeling and Simulation (ABMS). DM represents the process of identifying hidden and interesting knowledge in large amounts of data. It is "*a multidisciplinary field, drawing work from areas including database technology, machine learning, statistics, pattern recognition, information retrieval, neural networks, knowledge-based system, artificial intelligence, high-performance computing, and data visualization*" [22]. In the past decade, DM techniques have been widely applied in bioinformatics [43], e-commerce [35], financial studies [25], geography [30], marketing and sales studies [9, 38], etc.

ABMS is "*a new modeling paradigm and is one of the most exciting practical developments in modeling since the invention of relational databases*" [32]. It has "*connections to many other fields including complexity science, systems science, systems dynamics, computer science, management science, the social sciences in general, and traditional modeling and simulation*" [28]. In the past decade, ABMS has been used to model real-life systems in a diversity of domains such as biology [14], manufacturing [39], computing [13] and economics [8] among others. This variety of applications demonstrates the acceptance of ABMS as a useful system modeling and simulation approach to gain knowledge regarding complex systems in such domains. The work presented in [6] cites three reasons for the importance of ABMS for social sciences: first, that other approaches have been proved not suitable for the modeling of these (social sciences related) systems; second, that the agent based approach is a natural representation of many social systems; and third, that the emergence property in agent based models is not easily achieved with other approaches.

In this paper, we introduce the integration study of DM and ABMS (at a *conceptual* level). It is concerned with two approaches/directions:

1. **Applying DM in ABMS** aiming to provide solutions to the *open problem* (further described in section 2.2) in ABMS investigation, based on DM techniques; and
2. **Applying ABMS in DM** with the objective to surpass the *data limitation* (further described in section 2.4) of DM research, based on the simulation results of ABMS.

It is also one objective of this work to make a call for a closer interaction of DM community and ABMS community. We believe that such interaction will benefit both fields. In the case of ABMS it can be used to provide more generalized methods for the validation of agent-based models.

The rest of this paper is organized as follows. In section 2 we describe both ABMS and DM in detail as the background relevant to this study. Section 3 presents the approach of "Applying DM in ABMS". In section 4, we propose the idea of "Applying ABMS in DM". Finally our conclusions and the open issues for further research are given in section 5.

2 Background

2.1 Agent Based Modeling and Simulation

Computer based modeling and simulation of (real-life) complex systems has been one of the driving forces in the development of computer systems. A general definition of a simulation is the imitation of the operation of a process or a real world system through time [7]. A computational model is the representation of a real-life system through a computer program, expressed by a set of algorithms and mathematical formulas implemented as code in a programming language.

In contrast with pure mathematical models, the objective of computational models is not usually to obtain analytical solutions to specific questions. Instead, computational models allow the design of experiments to test the developed models under different scenarios (with different parameter configurations). These experiments are carried out with the objective of testing the behavior of the modeled systems under a certain set of assumptions [7]. This experimentation allows the designer to obtain insight of certain aspects of a complex system which would not be possible to detect using mathematical analysis, or for problems for which there is no tractable mathematical representation.

In ABMS, a system is modeled as a set of autonomous entities, namely *agents*. Each of these agents is positioned in an environment (either virtual or real) from which the agent obtains information by the use of sensors and makes decisions based on its perceived state of the environment and its objectives. These decisions are then reflected as actions performed to modify the state of the environment (i.e. direct actions to the environment, communication with other agents, further reasoning).

An agent can have different behaviors according to the system it populates [46]. Agents also have three basic properties: (a) *reactivity* – the ability to respond to events in the environment; (b) *pro-activity* – the ability to demonstrate some behavior determined by its particular objectives, taking the initiative to satisfy its necessities; and (c) *sociability* – the ability to interact with other agents and humans to fulfill its objectives [47]. These properties give agent based systems a great versatility in comparison with typical object based systems by providing a new type of abstraction for the representation of problem domains.

2.2 The Open Problem in ABMS

Within the agent-based modeling community there have been continuous efforts to create standard processes for the development of agent-based models, such as the standardization of model description (described in [51] and [52]) via the ODD[1] protocol, or the use of software engineering techniques (like UML diagrams [53]). However, one of the main issues inhibiting researches from fields outside computer science to accept ABMS as a tool for the modeling and simulation of real-life systems is the lack of standard verification and validation methodologies. In fact, verification and validation of multi-agent simulations is a concept which has been investigated only in conjunction with the development of specific models. It has only been in recent times that researchers have engaged in independent development of techniques for verification and validation (see [29, 45, 48]).

Verification and validation are two independent actions that need to be performed in order to achieve the accreditation of a simulation [5]. Verification aims to test whether the implementation of the model is an accurate representation of the abstract model. Hence, the *accuracy* of transforming a created model into the computer program is tested in the verification phase. Model validation is used to check that the implemented model can achieve the proposed objectives of the simulation experiments. That is, to ensure that the built model is an accurate representation of the modeled phenomena to simulate.

In contrast with pure mathematical models which use analytical equations, there is still no consensus among the scientific community on the appropriate methods for verifying and validating an agent based simulation [29]. However, part of the reason for the lack of formalisms which validate agent based simulations is the inherent

[1] Short for Overview, Design Concepts and Details, the three main groups conforming the protocol.

complexity that these systems try to represent. Some verification methods (discussed in [29]) are source code analysis, automatic theoretic verification and finite state verification. However, there is still some debate in the agent based modeling community on whether formal proofs of systems are useful [20]. Similarly, there is some debate on whether the verification of complex models with many parameters is possible [40].

2.3 Data Mining and Knowledge Discovery in Databases

Data Mining (DM) is an active research field in computer science, which is attracting more and more attention from a wide range of different groups of people. It aims to extract various types of hidden, interesting, previously unknown and potentially useful knowledge (i.e. rules, patterns, regularities, customs, trends, etc.) from sets of data, where the size of a collected dataset can be measured in gigabytes. In DM common types of mined knowledge include: association rules [1], classification rules [34], prediction rules [21], clustering rules [31], sequential patterns [42], emerging patterns [18], etc.

DM is considered to be the core stage/task of KDD (Knowledge Discovery in Databases), where KDD refers to the overall process of *knowledge discovery.* Piatetsky-Shapiro [33] clearly differentiates both terminologies of KDD and DM – "*sometimes 'knowledge discovery process' is used for describing the overall process, including all the data preparation and postprocessing while 'data mining' is used to refer to the step of applying the algorithm to the clean data*". The KDD process has been well studied and analyzed, e.g. [12, 19, 22, 41, 44]. It was pointed out in [2] that "*the problem of knowledge extraction from large databases involves many stages and a unique scheme has not yet been agreed upon*".

One possible outline of the KDD process can be presented as follows.

1. **Problem Specification:** In the first stage of the KDD process, a domain-oriented understanding of the target mining task/application is identified, which clarifies the goal of the application.

2. **Resourcing:** The second stage of the KDD process aims to create a suitable set of data on which the target application can be performed. It may be possible to find several large databases available that appear to be task-relevant for the target application, and which were originally built for other purposes and are irrelevant to the target application.

3. **Data Cleaning:** The purpose of this stage is as Han and Kamber [22] explain "*to remove noise and inconsistent data*" from a given dataset. Furthermore, the missing and distorted data [16] and data outliers are also cleansed in this stage.

4. **Data Integration:** In this stage, the cleaned data sets from different resources are combined into an integrated data set with a unified data view.

5. **Pre-processing:** The data collected may be in an unstructured format, i.e. texts, images, videos, etc. In this stage the collected data is transformed into a structured/semi-structured (e.g. XML, SGML) representation that follows the data to be further operated upon in the KDD process. For simplicity, especially

when the volume of the collected data is considered too large, this stage then selects the most significant data for the target application for further usage, and other data is discarded.

6. **Data Mining:** The purpose of this stage is to identify the most valuable information in the prepared data by utilizing "*data analysis and knowledge discovery techniques under acceptable computational efficiency limitations, and produces a particular enumeration of patterns over the data*" [50].

7. **Interpretation and Evaluation of Results:** The validity of each pattern discovered is interpreted and measured. From this the overall quality of the mining performance can be evaluated. In this stage, the discovered valuable knowledge is initially interpreted in a user-readable form (especially when the user is strongly involved in the evaluation), where the patterns, rule symbols, and/or variables are precisely and concisely expressed in human language. Suitable patterns (valuable knowledge) are then caught in this stage.

8. **Future Application:** The set of valuable knowledge mined, interpreted and measured from the stages 6 and 7 is then available to be applied for domain-oriented decision marking in the future.

It can be noted that the above stages are usually applied iteratively; with results of one stage providing feedback that allows improvement to earlier stages.

2.4 The Data Limitation of DM and KDD

With regard to the KDD process, it can be indicated that stages 2 to 5 together determine the requirements of data for DM. Although Anand *et al.* [4] indicate that "*the amount of data being stored in databases has been on the increase since the 1970's partly due to the advances made in database technology since the introduction of the relational model for data by E. F. Codd*", there is still some debate to the availability of data (for different application-domains) being stored electronically. Hence one major problem of DM and KDD is the difficulty to obtain enough data resources and collect sufficient amount of data for a mining task.

For some classical DM research tasks (i.e. classification, association rule mining, text mining, etc.), sufficient amounts of *real* data can be found (i.e. the UCI Machine Learning Repository [10], the LUCS-KDD Discretised Normalised Data Library [15], Usenet Articles [26], respectively), even though much of this data may be imperfect (being noisy, inconsistent, missing or distorted or containing data outliers). De Veaux and Hand [16] argue that "*anyone who has analyzed real data knows that the majority of their time on a data analysis project will be spent 'cleaning' the data before doing any analysis*", and "*common wisdom puts the extent of this at 60-95% of the total project effort*". Klein [24] argues that "*there is strong evidence that data stored in organizational databases have a significant number of errors*" and "*between one and ten percent of data items in critical organizational databases are estimated to be inaccurate*". Thus another problem of DM and KDD is that even there are enough data resources and sufficient amount of data, it is quite difficult to accurately detect and correct erroneous data in a given dataset.

Besides, the automated approaches for data integration and pre-processing may further damage the quality of the collected data. For example, during the data integration and transformation phases the significance of a data item for a particular mining task/application may be varied; and certainly the inaccurate selection of the most significant data may even cause none of any valuable pattern (knowledge) to be extracted.

3 Applying DM in ABMS

3.1 Previous Work

In [37], a method for the validation of ABMS using DM techniques is proposed. In that work, the authors present the results of the analysis of a simulation data obtained from the repeated execution of experiments while varying a single parameter. They analyze the obtained data using clustering techniques finding new patterns in the data.

As far as we know, the work by Remondino and Correndo [36, 37] is the only attempt to apply DM techniques to ABMS. In these works, the authors differentiate between *endogenous* and *exogenous* application of DM techniques. The endogenous use is concerned with providing the agents participating in the simulation with the DM techniques in order to improve their performance. For example in [3], the authors established *data mining agents* in the context of a multi-agent system, making use of DM techniques to perform distributed mining.

On the other hand, the exogenous application focuses on using DM techniques to analyze the data resulting from the simulation. Exogenous use of DM in ABMS is exemplified by the application described by Remondino and Correndo [37], where they perform a series of simulation experiments using an agent based model of a biological phenomenon. Afterwards, they proceeded to perform a DM analysis with the obtained data to detect if there was any novel pattern.

3.2 Proposed Ideas

Although we generally agree with the use of DM in ABMS discussed in [36] and [37], we believe that DM techniques can be further used to aid in the formalization of validation processes for ABMS. To achieve this, we claim that techniques such as the identification of classification rules, clustering rules, association rules and sequential patterns in the data obtained from the simulation experiments can be used. Specifically, we believe that using DM techniques would be possible to *abstract* experimental results by creating higher level description elements representing the behavior observed in the simulations, without having to focus in the specific obtained data values.

This would allow a more straightforward comparison between different experimental results, obtained from the same model or from independent models representing the same phenomena. Such comparative analysis would be possible by the comparison of the high level description elements (such as sequential patterns of agent behavior or identified clusters) obtained from the raw experimental data.

Similarly, such an approach could be used to compare the results obtained from the experiments of the simulation and data from the real-life/world system that is being simulated. This considering that, even though the values obtained from the different experiments will not be exactly the same than those from the (real-life) system, it is sensible to expect that the data obtained from the simulation experiments and the data from the system will share some general descriptive patterns. Fig. 1 depicts the process followed when applying DM techniques for as part of the verification and validation of agent based models.

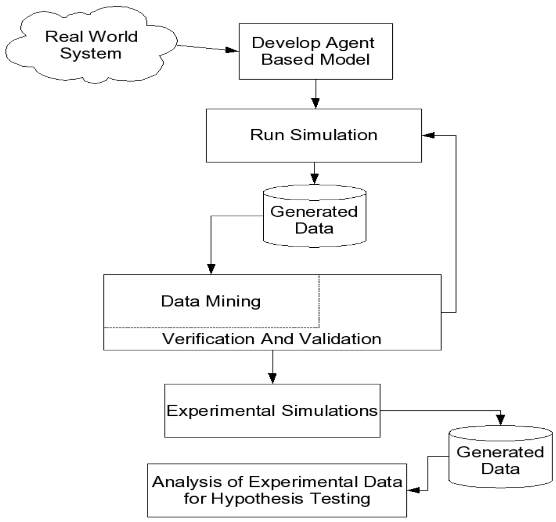

Fig. 1. Applying DM for the verification and validation of agent based models.

4 Applying ABMS in DM

Beyond the topic of "Applying DM in ABMS", in this paper we inversely, for the first time, introduce the approach of "Applying ABMS in DM" (Fig. 2). The aim of this approach is to provide sufficient amount of data with good quality for various DM mechanisms and ensure the utilization of DM techniques in different domain-applications. Specifically, the simulation results of ABMS can be used as *quasi-real* data when there is the lack of *real* data for a domain-specific DM task.

The approach described in this section represents a new paradigm in data preparation of DM and KDD, which in turn create an improved KDD process. This new KDD process consists of five stages only, presented as follows.

1. **Problem Specification:** This stage contains the original problem specification contents (as previously described in section 2.3) plus the specification of an ABMS system for this application problem.
2. **Data Preparation:** Based on the determined ABMS system (from the stage 1), this stage then generates simulation results and forms these results into a corresponding dataset for the next (DM) stage. It can be pointed out that the data

generated herein shows the following properties together surpassing the data limitation of DM and KDD (see section 2.4): (a) *quasi-real* – as mentioned above, ABMS is well established and a mature technology in the analysis of (real-life) complex system, its simulation results can be very similar to the real-life/world situations (real data); (b) *suitable-sized* – for different DM techniques/ mechanisms, the amount of data may be sensitive, hence generating not too much but sufficient amount of data is considered; (c) *qualified* – the erroneous data can be controlled, so that there is neither noisy/inconsistent/missing/distorted data nor data outliner involved; and (d) *significant* – insignificant data will be avoided during the simulation process by providing correct parameter configuration(s); thus the data significance for a particular application can be ensured.

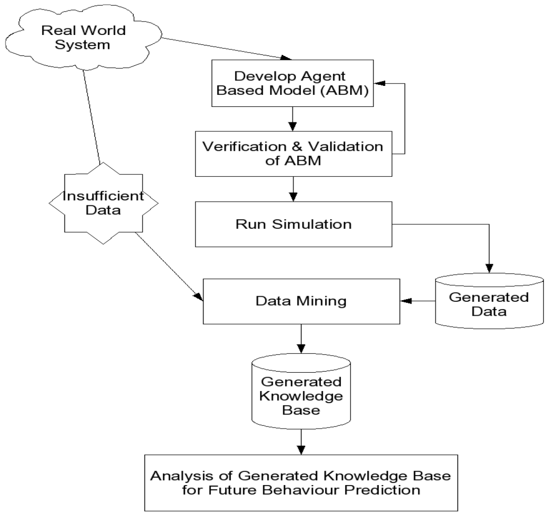

Fig. 2. Applying ABMS for the data preparation of data mining

3. **Data Mining:** keep this stage as previously described in section 2.3.
4. **Interpretation and Evaluation of Results:** Keep this stage as previously described in section 2.3.
5. **Future Application:** Keep this stage as previously described in section 2.3.

Further benefit of the proposed "Applying ABMS in DM" approach can be exemplified with the following case.

Assume that we are dealing with a classification problem – the automated categorization of an "unseen" data instance into pre-defined classes – in a financial situation, say that to classify whether an oil derivative is currently buyable, sellable or non-tradable. In this case, alternative classification mechanisms can be applied that include: decision tree induction [34], naive Bayes [17], *k*-nearest neighbor [23], neural networks [49], support vector machine [11], classification association rule mining [27], etc. In a DM context, it will be interesting to determine which (classification) mechanism is always well performed herein, so that it can be suggested in a default manner whenever a similar problem is handled. This offers the solution to avoid the problem of over-fitting – some mechanism is only good when applying it to a particular set of data, but not general applicable at all. With regard to an ABMS system available that simulates the real-life oil trading market, a set of simulation results can be generated by rationally and strategically varying the value of different parameters, where each simulation result is then formed into a dataset among which hidden rules can be mined and performance of each mechanism is then evaluated. Finally an overall averaging process can be applied to determine the best performed mechanism throughout various simulation scenarios (with different parameter configurations) in the agent based oil trading modeling and simulation.

5 Conclusions

In this paper, we proposed a general idea for the bidirectional integration of Data Mining (DM) and Agent Based Modeling and Simulation (ABMS) (i.e. "Applying DM in ABMS" and "Applying ABMS in DM"). A conceptual framework of this integration study was provided; and a broader view of the advantages that such integration can provide to both DM and ABMS was presented.

The integration of DM and ABMS provides promising mechanisms for the advancement in both fields. From the ABMS point of view, DM can provide necessary mechanisms for the validation of created models and to facilitate the comparison between data obtained from different simulations. From the DM point of view, ABMS can be used to generate necessary data when the real data obtained is not enough or does not have good quality. Moreover, ABMS can be used to obtain datasets from designed simulation scenarios (for which real-life data may not be available), in order to determine the default mining approach for a group of similar applications. This offers a solution to the over-fitting problem in DM.

Further research may identify some new approaches other than the proposed verification & validation and data preparation ideas under the headings of "Applying DM in ABMS" and "Applying ABMS in DM".

Acknowledgments. The authors would like to thank Kamal Ali Albashiri, Chuntao Jiang and Maya Wardeh of the Department of Computer Science at the University of Liverpool for their support with respect to the work described here. Omar Baqueiro acknowledges the support provided by the Mexican Council of Science and Technology (CONACYT) with the grant 187564.

References

1. Agrawal, R., Imielinski, T., Swami, A.: Mining Association Rules between Sets of Items in Large Databases. In: Proceedings of the 1993 ACM SIGMOD International Conference on Management of Data, Washington, DC, USA, pp. 207--216. ACM Press (1993)
2. Ahmed, S.: Strategies for Partitioning Data in Association Rule Mining. Ph.D. Thesis, University of Liverpool, Liverpool, UK (2004)

3. Albashiri, K.A., Coenen, F., Leng, P.: Agent Based Frequent Set Meta Mining: Introducing EMADS. In: Artificial Intelligence in Theory and Practice II – Proceedings of IFIP 20th World Computer Congress, TC 12: IFIP AI 2008 Stream, Milano, Italy, pp. 23--32. Springer (2008)
4. Anand, S.S., Bell, D.A., Hughes, J.G.: Evidence based Discovery of Knowledge in Databases. IEE Colloquium on Knowledge Discovery in Databases, Digest No: 1995/021(A), 9/1--9/5, London, UK (1995)
5. Balci, O.: Verification Validation and Accreditation of Simulation Models. In: Proceedings of the 29th Conference on Winter Simulation, Atlanta, GA, USA, pp. 135--141. ACM Press (1997)
6. Bankes, S.: Agent-based Modeling: A Revolution? In: Proceedings of the National Academy of Sciences, 7199--7200 (1999)
7. Banks, J.: Handbook of Simulation. John Wiley & Sons (1998)
8. Beam, C., Segev, A.: Automated Negotiations: A Survey of the State of the Art. Wirtschaftsinformatik 39(3), 263--268 (1997)
9. Berry, M.J.A., Linoff, G.: Data Mining Techniques for Marketing, Sales, and Customer Support. John Wiley & Sons, Inc., USA (1997)
10. Blake, C.L., Merz, C.J.: UCI Repository of Machine Learning Databases. Department of Information and Computer Science, University of California, Irvine, CA, USA (1998) "http://www.ics.uci.edu/~mlearn/MLRepository. html"
11. Boser, B.E., Guyon, I.M., Vapnik, V.N.: A Training Algorithm for Optimal Margin Classifiers. In: Proceedings of the 5th ACM Annual Workshop on Computational Learning Theory, Pittsburgh, PA, USA, pp. 144--152. ACM Press (1992)
12. Brachman, R.J., Anand, T.: The Process of Knowledge Discovery in Databases: A Human Centered Approach. Advance in Knowledge Discovery and Data Mining, 37--57. AAAI/MIT Press (1996)
13. Buyya, R., Abramson, D., Giddy, J., Stockinger, H.: Economic Models for Resource Management and Scheduling in Grid Computing. Concurrency Computation Practice and Experience, 14(14), 1507--1542 (2002)
14. Castella, J., Trung, T., Boissau, S.: Participatory Simulation of Land-use Changes in the Northern Mountains of Vietnam: The Combined Use of an Agent-based Model, a Role-playing Game, and a Geographic Information System. Ecology and Society, 10(1), 27 (2005)
15. Coenen, F.: The LUCK-KDD Discrietised/Normalised ARM and CARM Data Library. Department of Computer Science, University of Liverpool, Liverpool, UK (2003) "http://www.csc.liv.ac.uk/~frans/KDD/Software?LUCS-KDD-DN"
16. De Veaux, R.D., Hand, D.J.: How to Lie with Bad Data? Statistical Science, 20(3), 231--238 (2005)
17. Domingos, P., Pazzani, M.: On the Optimality of the Simple Bayesian Classifier under Zero-one Loss. Machine Learning, 29(2/3), 103--130 (1997)
18. Dong, G., Li, J.: Efficient Mining of Emerging Patterns: Discovering Trends and Differences. In: Proceedings of the Fifth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, San Diago, CA, USA, pp. 43--52. ACM Press (1999)
19. Fayyad, U., Piatetsky-Shapiro, G., Smyth, P.: Knowledge Discovery and Data Mining: Towards a Unifying Framework. In: Proceedings of the Second International Conference on Knowledge Discovery and Data Mining, Portland, OR, USA, pp. 82--95. AAAI Press (1996)
20. Glass, R.L.: Practical Programmer: Inspections, Some Surprising Findings. Commun. ACM, 42(4), 17--19 (1999)
21. Han, J., Kamber, M.: Data Mining: Concepts and Techniques. Morgan Kaufmann Publishers, San Francisco, CA, USA (2001)
22. Han, J., Kamber, M.: Data Mining: Concepts and Techniques (Second Edition). Morgan Kaufmann Publishers, San Francisco, CA, USA (2006)

23.James, M.: Classification Algorithms. Wiley-Interscience, New York, NY, USA (1985)
24.Klein, B.D.: Data Quality in the Practice of Consumer Product Management: Evidence from the Field. Data Quality Journal, 4(1) (1998)
25.Kovalerchun, B., Vityaev, E.: Data Mining in Finance: Advances in Relational and Hybrid Methods. Kluwer Academic Publisher (2000)
26.Lang, K.: NewsWeeder: Learning to Filter Netnews. In: Machine Learning – Proceedings of the Twelfth International Conference on Machine Learning, Tahoe City, CA, USA, pp. 331--339. Morgan Kaufmann Publishers (1995)
27.Liu, B., Hsu, W., Ma, Y.: Integrating Classification and Association Rule Mining. In: Proceedings of the Fourth International Conference on Knowledge Discovery and Data Mining, New York, NY, USA, pp. 80--86. AAAI Press (1998)
28.Macal, C.M., North, M.J.: Agent-Based Modeling and Simulation: Desktop ABMS. In: Proceedings of the 2007 Winter Simulation Conference, Washington, DC, USA, pp. 95--106. IEEE Computer Society (2007)
29.Midgley, D., Marks, R., Kunchamwar, D.: Building and Assurance of Agent-based Models: An Example and Challenge to the Field. Journal of Business Research, 60(8), 884--893 (2007)
30.Miller, H.J., Han, J.: Geographic Data Mining and Knowledge Discovery. Taylor & Francis (2001)
31.Mirkin, B., Mirkin, B.G.: Clustering for Data Mining: A Data Recovery Approach. Chapman & Hall/CRC (2005)
32.North, M.J., Macal, C.M.: Managing Business Complexity: Discovering Strategic Solutions With Agent-Based Modeling And Simulation. Oxford University Press, Oxford, UK (2007)
33.Piatetsky-Shapiro, G.: Knowledge Discovery in Databases: 10 Years After. ACM SIGKDD Explorations, 1(2), 59--61 (2000)
34.Quinlan, J.R.: C4.5: Programs for Machine Learning. Morgan Kaufmann Publishers, San Francisco, CA, USA (1993)
35.Raghavan, S.N.R.: Data Mining in E-commerce: A Survey. Sadhana, 30(2&3), 275--289 (2005)
36.Remondino, M., Correndo, G.: Data Mining Applied to Agent Based Simulation. In: Proceedings of the 19th European Conference on Modelling and Simulation, Riga, Latvia (2005)
37.Remondino, M., Correndo, G.: MABS Validation through Repeated Executing and Data Mining Analysis. International Journal of Simulation Systems, Science & Technology, 7(6), 10--21 (2006)
38.Rypielski, C., Wang, J.-C., Yen, D.C.: Data Mining techniques for customer Relationship Management. Technology in Society, 24, 483--502 (2002)
39.Shen, W., Norrie, D.H.: Agent-based Systems for Intelligent Manufacturing: A State-of-the-art Survey. Knowledge and Information Systems, 1, 129--156 (1999)
40.Shervais, S., Wakeland, W., Raffo, D.: Evolutionary Verification and Validation of Software Process Simulation Models. The 5th International Workshop on Software Process Simulation and Modelling, Edinburgh, Scotland, UK (2004)
41.Smith, G.D.: Meta-heuristics in the KDD Process. In: Proceedings of the First UK Knowledge Discovery in Data Symposium, Liverpool, UK (2005)
42.Wang, W., Yang, J.: Mining Sequential Patterns from Large Data Sets. Springer-Verlag New York Inc. (2005)
43.Wang, J.T.L., Zaki, M.J., Toivonen, H.T.T., Shasha, D.E.: Data Mining in Bioinformatics. Springer-Verlag London Limited (2005)
44.Weiss, S.M., Indurkha, N.: Predictive Data Mining: A Practical Guide. Morgan Kaufmann Publishers, San Francisco, CA, USA (1997)

45. Windrum, P., Gagiolo, G., Moneta, A.: Empirical Validation of Agent-based Models: Alternatives and Prospects. Journal of Artificial Societies and Social Simulation, 10(2), 8 (2007)
46. Wooldridge, M.J.: An Introduction to MultiAgent Systems. John Wiley & Sons (2002)
47. Wooldridge, M.J., Jennings, N.R.: Intelligent Agent: Theory and Practice. Knowledge Engineering Review, 10(2), 115--152 (1995)
48. Yilmaz, L.: Validation and Verification of Social Processes within Agent-based Computational Organization Models. Computational & Mathematical Organization Theory, 12(4), 283--312 (2006)
49. Zhang, G.P.: Neural Networks for Classification: A Survey. IEEE Transactions on Systems, Man, and Cybernetics – Part C: Applications and Reviews, 30(4), 451--462 (2000)
50. Zhang, D., Zhou, L.: Discovering Golden Nuggets: Data Mining in Financial Application. IEEE Transaction on Systems, Man, and Cybernetics – Part C: Applications and Reviews, 34(4), 513--552 (2004)
51. Grimm, V., Berger, U., Bastiansen, F., Eliassen, S., Ginot, V., Giske, J., Goss-Custard, J., Grand, T., Heinz, S.K., Huse, G., Huth, A., Jepsen, J.U., Jorgensen, C., Mooij, W.M., Muller, B., Pe'er, G., Piou, C., Railsback, S.F., Robbins, A.M., Robbins, M.M., Rossmanith, E., Ruger, N., Strand, E., Souissi, S., Stillman, R. A., Vabo, R., Visser, U., Deangelis, D.L.: A Standard Protocol for Describing Individual-based and Agent-based Models. Ecological Modelling, 198(1-2), 115--126 (2006)
52. Polhill, J., Parker, D., Brown, D., Grimm, V.: Using the ODD Protocol for Describing Three Agent-based Social Simulation Models of Land-use Change. JASSS 11(2) (2008)
53. Lisec, A., Ferlan, M., Lobnik, F., Sumrada, R.: Modelling the Rural Land Transaction Procedure. Land Use Policy 25(2), 286--297 (2008)