Business Analytics as a Framework for an Evolving Multi-Agent System

Carol Anne Hargreaves, Liya Ding
Institute of Systems Science
National University of Singapore
Singapore
E-mail: carol.hargreaves@nus.edu.sg; ding.liya@nus.edu.sg

Abstract— **This article first reviews business analytics (BA) process with its key tasks, and then proposes a framework of an evolving multiagent system from the perspective of an automated business analytics. Under this view, typical BA tasks, such as prediction, evaluation, and decision are carried out by individual intelligent agents in a multiagent system (MAS) with an evolving environment to cope with the target business application.**

Keywords-business analytics; multiagent system, intelligent system, knowledge-based system.

I. Introduction

Business Analytics can be applied in any business where data is recorded and stored. It can be used for a variety of purposes, ranging from improving service or performance to detecting fraud or waste, to analyzing scientific and research information. The key process of Business Analytics is the turning of data into insights and insights into knowledge. Current emerging trends that are changing traditional ways of doing business are business analytics, real-time analysis and knowledge-based systems. Many organizations are exploiting business analytics to enable proactive decision making and are moving from reacting to situations to anticipating them. The main focus of this paper is a review of the business analytics process, and actions, and to conceptualize the business analytics process, by deriving insights from the model and then updating and formalizing these insights as an evolving Multi-agent System. The Business Analytics process as an evolving Multi-Agent System is important because it will speed up the decision making process of organizations and allow them gain competitive advantage over their competitors.

The rest of the paper is organized as follows. Section 2 gives a review of the key tasks in business analytics process; Section 3 proposes a framework of an evolving multiagent system from the perspective of an automated business analytics; an illustrated application on costumer analytics is discussed in Section 4 as an automated evolving multiagent system for specific application; and Section 5 provides the conclusion.

II. The Business Analytics Process

As a consequence of global competition, organizations need to act and react more quickly. Many organizations have difficulty in providing the right data insights to the right person at the right time, to assist them in making the right decisions. In dynamic markets with high competition, “gut feelings” and “trial & error” are ineffective for managing an enterprise. Organizations need business analytics and business intelligence to make accurate and timely decisions in a systematic approach [1]. Knowledge is the most important competitive weapon an organization can have, so it is important for organizations to develop a reliable, systematic process to integrate knowledge from various sources to generate superior competitive advantage. The objective of a reliable system is to allow the organization to be more proactive and responsive while operating in its dynamic market environment. In this paper, we investigate Business Analytics as a system where data is gathered in a centralized multidimensional database. Because many individuals will be running tasks throughout the organization from the same centralized database, it is a challenge for the organization to ensure that the Business Analytics system uses quality data, and the insights derived are reliable as everyone has the same centralized data view.

This section outlines the stages necessary to broaden the effectiveness of analytics in the business process in order to achieve full impact and value by making the business analytic solutions actionable and measurable [2]. The idea of building a Business Analytics framework is to identify relevant business analytic metrics that can align to the organizations’ strategic goals in a systematic process [3]. ‘Data Preparation’ is the first stage in the Business Analytics Process. This stage includes, cleaning the data, making computations for missing data, removing outliers, and transforming combinations of variables to form new variables.

Once the data has been cleaned, the analyst will try to make better sense of the data through -‘Data Understanding’. At this stage, the analyst will plot the data, visually check all possible slices of data and summarize the data using appropriate visualization and descriptive statistics. Here the task of the analyst is to identify interesting patterns in the data and to determine what actionable insights can be derived to achieve the organizational goal. The use of scatter plots (to identify possible correlation or non-linearity) or time series graphs can indicate any patterns or unusual observations. The standard descriptive statistics, e.g., mean, standard deviation, range, mode, median, are useful on providing a basic understanding of

DOI: 10.5176/2010-3043_2.3.200

the data. It is here that the analyst is already observing and looking for general patterns and insights. The analyst will also compute the correlation of the target variable and all other variables to determine the relationship between different variables.

The third stage is 'Modelling'. The analyst will model the data using techniques that include Decision Trees, Neural Networks, Logistic Regression, etc. These techniques aim to discover insights and highlight relationships and 'hidden evidences' of the most influential variables.

The fourth stage is the 'Evaluation of the model accuracy and outcomes'. It is in this step that the analyst will compare the predictive values with the actual values and compute the predictive errors. It is here that the analyst will also run 'what-if' scenarios, using targets set by managers. The problem of finding and confirming targets becomes an optimization problem. The analyst will select the optimal solution and model based on the lowest error, management targets and the intuitive understanding of the model coefficients that most align to the organization's strategic goal.

In the fifth stage, the analyst will 'Make Decisions and Take Action' based on the derived insights from the model and the organizational goal.

Stage six takes place after an appropriate period of time that the action was taken, where the 'Action is Measured' and in stage 7, the 'Results of the Action and the new insights that were derived from the model is Recorded and Updated' in the database.

Figure 1 gives the top level processing flow of a typical Business Analytics process.

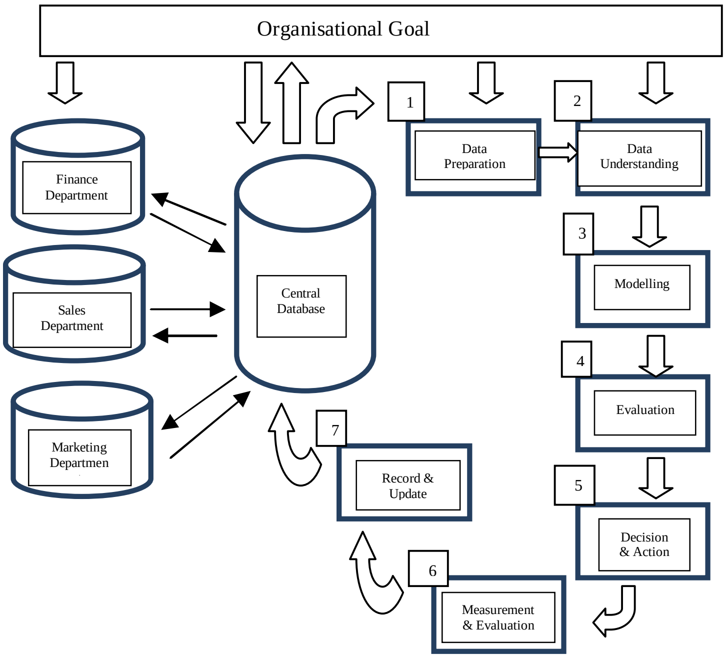

Figure 1: Business Analytics Process

III. An Evolving Multiagent System

A. *Multiagent Systems*

Multiagent Systems (MAS) is considered the subfield of Artificial Intelligence whose main aim is to provide both principles for construction of complex systems involving distributed multiple agents and mechanisms for coordination of independent agents' behaviors [4]. Though there is no generally accepted definition of "agent" in AI, in our discussion, we understand an agent to be an entity, with goals, actions, and domain knowledge, situated in an environment. The ways it acts is considered "agent's behavior" and is characterized by several important characteristics [5]:

- Autonomy: the agents are at least partially autonomous;
- Local views: no agent has a full global view of the system;
- Decentralization: there is no designed controlling agent.

Through a composed agent's behaviors, an MAS is capable of providing different design benefits such as parallelism, robustness, scalability, geographic distribution and cost effectiveness [6]. In particular, the latest development of multicore processors allows MAS to organize distributed computing in a hierarchical way where agents are conceptually arranged in a treelike structure where each node represents a processor core [4].

These design benefits provided by MAS will be able to help BA process in many ways considering the current BA practice as playing with "big-data" which is often in distributed sources and with frequent updating, and making decision in an evolving environment with changing business target.

B. *Evolving Multiagent System Driven by BA*

Our aim is to automate tasks in BA process mentioned in the Section 2 by intelligent agents with necessary *local knowledge*, thus the entire BA process can be automated by a multiagent system that is characterized by several characteristics:

- There is a central database to which all the individual agents have the access;
- There is no centralized control;
- Each agent has its "influence zone" where it affects other agents' response. This influence zone can be understood as a kind of dependency relation which is not symmetric between any two dependent agents in general;
- The data in system database can be updated by any agent in the system;
- The local knowledge of an agent can be updated either by itself or by other agents that have influence to it;
- The decision model of an agent can be updated by other agents that have influence to it;

- There is a *system synchronization* based on predefined “system regulation”, and such synchronization is related to data change and model change as system evolving;
- Human involvement is allowed to access the system database for data retrieval, data update, and decision retrieval, but must be under the same spirit of system synchronization.

This is the framework of an Evolving Multiagent System driven by BA process (EMAS-BA). One of the key challenges here is the coordination of multiple agents. Most of the time, agents in EMAS-BA are given autonomy to carry out knowledge-based processing individually given their own local knowledgebase and protocols to the system database. However whenever an update from an agent is detected, other agents in its influence zone need take necessary action for the “synchronization” according to the “system regulation” for specific application. The system database here is conceptually treated as one database for easy discussion but in reality it can be in multiple distributed sources (further discussion will be handled elsewhere).

Intelligent agents in the proposed EMAS-BA framework are classified to following types with their tasks:

1) Data preparation

It looks for any outliers and removes these from the data set as they will affect the accuracy of the model if they remain in the data set. The task is carried out as a knowledge-based processing with domain specific criteria given. Typical rule-based or case-based approach can be considered.

It has its interface to the system database.

2) Data understanding

It applies the standard descriptive statistics, e.g., mean, standard deviation, range, mode, and median to obtain a basic understanding of the data and also applies statistical tools such as scatter plots, or time series graphs to indicate any patterns or unusual observations. Based on the basic understanding, it can also compute the correlation of the target variable and all other variables to determine the relationship between different variables. A set of higher level strategies (knowledge) in choosing suitable tools is given as the local knowledge to support the agent, which is usually domain specific, and can be represented and processed by rule-based or case-based approaches.

It has its interface to the system database as well as a collection of available statistical tools.

3) Model building & selection

This is the most complicated type of agent in our EMAS-BA framework. It plays the role of bridging data understanding to prediction and decision. It discovers insights and relationships of most influential variables and captures them in possible “model” for specific problem. Typical techniques that support this task include Decision Trees, Neural Networks, Logistic Regression, etc. Similar to *data understanding* agent, a set of domain specific strategies in choosing suitable techniques is necessary as the local knowledge of the agent. Having multiple possible models for a specific application discovered, this agent is also to compare the predictive values provided by different predictive models with the actual values and evaluate accuracy of each model. An optimal model will be selected in terms of the lowest error, and the best match to business target. The selected model will then be formed as an independent agent in the evolved multiagent system. In this sense, a *model building and selection* agent also acts as an “agent constructor”. The Knowware System [7] techniques for an automatic construction of knowledge-based system (KBS) with intelligent components [8] can be adopted to support this function.

It is necessary to have a set of decision strategies (e.g. rules or cases) as the local knowledge for a decision agent, when a decision or an action is required for a specific application. Such decision strategies are considered a part of initial knowledge, but necessary parameters can be decided and updated by this model building agent based on the latest insights of data discovered.

So a model building and selection agent has its interface to the system database, the collection of predictive techniques, and also the local knowledgebase of decision agent (if there is any) within its influence zone.

4) Prediction

A prediction agent can be built through two ways: (a) an agent in an initial multiagent system, when the prediction task is clearly defined and well modeled using domain knowledge from very beginning; or (b) an agent constructed by a model building & selection agent based on the predictive model selected. In whatever case, there will be a local knowledge base connected to the agent, with the corresponding knowledge representation according to the specific technique used, such as neural networks, decision tree, etc.

It has its interface to the system database.

5) Decision

A decision agent is constructed in an initial EMAS-BA with necessary local knowledge for decision making according to the business target. It can be designed as deterministic or having parameters to be set or updated by *model building & selection* agent.

It has its interface to the system database, and also to the model building & selection agent.

6) Measurement & evaluation

When a recommendation from a *decision* agent is accepted by user and an actual action of the user according to the recommendation is taken place, the impact should be recorded in the system database. This is the only update in the EMAS-BA framework that requires human involvement as necessity.

A measurement & evaluation agent is to measure and evaluate the impact of an action and the evaluation result will be used to guild future processing. This agent keeps “silence” most of the time and only be active when an update of impact of action is detected.

A measurement & evaluation agent has its interface to the system database.

C. EMAS-BA for Specific Application

For a specific application, an EMAS-BA can be constructed with a collection of agents of the six types. While it is not always the case that all the six types of agents involved in one application, there may also be multiple agents of the same type involved in one application.

IV. AN ILLUSTRATED APPLICATION

In this section we illustrate an evolving multiagent system for business analytics using a customer analytics example, where the organizations strategic goal is to reduce it churn.

There is a *central database* and each individual agent has access protocol to this database. There is also a *local knowledgebase* for each of the agents. In the initial system five real agents and one dummy agent are employed, and the dummy agent will be updated by the model building & selection agent after an optimal model discovered and selected.

1) Agent-1: Data Preparation

It performs knowledge-based processing to look for any outliers and remove these from the data set. Here a domain specific strategy is applied: if more than 50% of any variable's data is missing, this variable is also removed from the data set. New variables such as 'time to contract end' and 'customer lifetime value' are computed as part of the data preparation. As typical KBS techniques, rule-based inference or case-based inference can be applied.

Its influence zone includes Agent-2.

2) Agent-2: Data Understanding

Once the data has been cleaned, descriptive statistics such as the mean, mode, median, standard deviation, range correlation, will be computed so that different comparisons can be made and more understanding of the data can be made. It can also do scatterplots to identify any possible correlations between variables and with the 'churn' variable. For example, a comparison of the 'number of complaints in the last month' made by customers whose contract is about to end, compared to those who have been customers for more than 3 years.

Its influence zone includes Agent-4.

3) Agent-3: Model Building & Selection

Candidate churn models will be built based on data understanding using techniques such as decision trees, neural networks, and the logistic regression, to identify customers who are likely to churn and the factors that statistically contribute to churn. Once a set of possible churn models has been built, an evaluation of the different models accuracy and coefficients will be carried out using targets initially set by human managers. The optimal model is chosen based on the lowest error, management targets and the intuitive understanding of the model coefficients that most align to the organization's strategic goal.

Its influence zone includes Agent-4 and Agent-5. The Agent-4, i.e. prediction agent (dummy or previously updated) will be updated with the model selected by this Agent-3. The insights of the selected optimal churn model will be used to update the local knowledgebase of Agent-5. Once the dummy prediction agent is replaced by a real model and the local knowledgebase of the decision agent is updated accordingly, the model building & selection agent will keep as idle until new changes have been detected or a poor performance of the current system is recognized.

4) Agent-4: Prediction

It starts as a dummy agent and is updated every time when a new optimal model is chosen by Agent-3.

Its influence zone includes Agent-5.

5) Agent-5: Decision

Based on the derived insights of the optimal churn model, the decision rules in the local knowledge base of the decision agent are updated. Decisions will be made in terms of the marketing strategy the organization should take, in order to prevent the person who is likely to churn in the next 2 months, from churning. For example, it might decide to make those customers who are valuable but have a high probability of churning an offer –for the next 2 years their subscription fees will be reduced by 10%. Customers who are likely to churn but are in bad debt with the organization, will not be given the offer of reduced subscription fees for the next 2 years as the organization will be better off if they were not customers.

Its influence zone includes Agent-6 with a predefined delay period, e.g. 3 months.

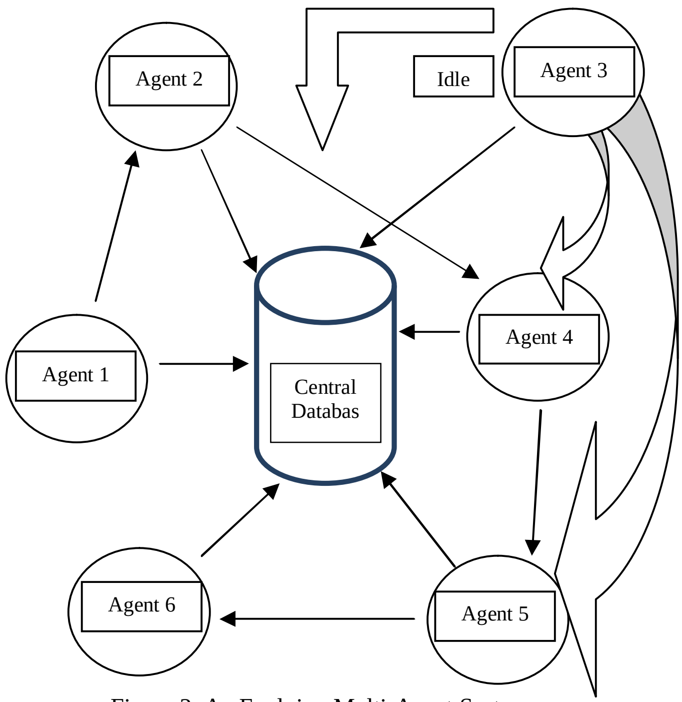

Figure 2: An Evolving Multi-Agent System

6) Agent-6: Measurement & Evaluation

When the response rate is recorded, after 3 months the agent will measure the response rate (of those who were given the offer, how many took the offer up). If the response rate is more than 70%, the marketing strategy is identified as being effective. The factors that drive the response rate are also updated into the system database. And lastly, this agent will compute the organizations new churn rate to see whether it has reduced and the objective of the company was met.

Though the guideline for where to measure, how to measure, and the evaluation criteria are predefined as local knowledge for this agent, the new churn rate will be used as new finding to update the system database thus affect the other agents' action.

Its influence zone includes Agent-3.

Figure 2 shows the influence and access relationship in the system, with the human users omitted.

V. CONCLUSION

We proposed a framework of an evolving multiagent system EMAS-BA from the Business Analytics perspective. We have shown that there is a good potential to apply existing agent technologies and other intelligent techniques to achieve an EMAS-BA. While positioning such an EMAS-BA as an automatic system, we have considered possible human involvement as real-time data provider, decision receiver, and actual impact data provider, i.e. the users (departments) mentioned in the section 2.

There are issues and challenges for our future research: (a) prediction of environment changing; (b) strategy evolving based on data change; (c) strategy evolving based on preference and criteria change; (d) system evolving driven by business target; (5) uncertainty handling at both agent level and system level; (6) management and maintenance of system database with distributed sources.

REFERENCES

[1] S. Mrdalj. "Teaching an applied business intelligence course," *Issues in Information Systems*, vol. 111, no.1, 2007.

[2] R. Kohavi, N.J. Rothleder, and E. Simoudis, *Communications of the ACM*, vol. 45, no. 8, pp. 45-48, 2002.

[3] C. Voudouris, G. Owusu, R. Dorne, and D. Lesaint, *Service Chain Management*, 205. DOI: 10.1007/978-3-540-75504-3, _c Springer 2008.

[4] G. Acampora, M. Gaeta, and V. Loia, "Exporing e-Learning Knowledge Through Ontological Memetic Agents," *IEEE Comput. Intell. Mag*., vol. 5, no. 2, pp. 66-77, 2010.

[5] M. Wooldridge, *An Introduction to MultiAgent Systems*. New York: Wiley, 2002.

[6] K.A. De Jong, "Involving intelligent agents: A 50 year quest," *IEEE Comput.Intell. Mag*., vol. 3, no. 1, pp. 12-17, 2008.

[7] L. Ding, "Knowware system for the development of intelligent systems", in *Proc. 7th Int. Conf. Intelligent Technologies, InTech 2006*, pp. 201-210.

[8] L. Ding, and S. Nadkarni, "Automatic construction of knowledge-based system using knowware system," in *Proc. the 6th Int. Conf. Machine Learning and Cybernetics, ICMLC 2007*, pp. 789-794.