An Intelligent Agent based Architecture for Visual Data Mining

Hamdi Ellouzi[1], Hela Ltifi[1,2] ,Mounir Ben Ayed[1,3]
[1]Research Groups on Intelligent Machines
National School of Engineers (ENIS), University of Sfax, BP 1173, Sfax, 3038, Tunisia
[2] University of Kairouan, Faculty of sciences and techniques of SidiBouzid, Tunisia
[3] University of Sfax, Faculty of sciences of Sfax, Tunisia

***Abstract*—the aim of this paper is to present an intelligent architecture of Decision Support System (DSS) based onvisual data mining. This architecture applies the multi-agent technology to facilitate the design and development of DSS in complex and dynamic environment.Multi-Agent Systems add a high level of abstraction. To validate the proposed architecture, it is implemented to develop a distributed visual data mining based DSSto predictnosocomial infectionsoccurrence in intensive care units. The developed prototype was evaluated to verify the architecture practicability.**

Keywords—Multi Agent System; Decision Support System; Visualization; Knowledge Discovery from Data; Nosocomial Infection

I. INTRODUCTION

Decision-making in dynamic and complex environment faces several problems as a result of the increase in temporal data size and the diversity of heterogonous data sources. Integrating Data mining technology in Decision Support System (DSS) assists decision-makers in problem solving. Data mining algorithmsprovideuseful patterns to discover data associations [1] [2].Data mining as two main goals: (1) extracting relevant information from a set of raw data according to user's request to have coherent knowledge about the system's variations (2) transforming data from textual representation into meaningful formsto specify data associations. The complexity of the previoustasks realization increases in case of temporal data. The integrationof visualization techniques[5]in data mining based DSS provides an acceptable outcome in clear graphical formto better understands temporal data and extracted patterns variation[5] [6] [7]. Thus, we are interested to visual data mining based DSS. Such systems are characterized by their complexity and dynamic character. To support this complexity, intelligent agent technology is a promising solution [3] [4].

In this context, our work aims to propose a multi-agent architecture for visual data mining based DSS,which ensureshigher level of adaptability, mobility, and intelligence. It involves a set of active goal-oriented agents to play one or more roles in the decision environment.

This paper is organized as follow: in section 2, we present our researchcontext including decision support system, visual data mining and intelligent agent technology. Then in section 3, we introduce the suggested architecture based on cognitive intelligent agents. Next, we present the architecture applicationin the medical field. In section 5, we will apply a set of evaluation utility and usability tests to validate the developed prototype.

II. RESEARCH CONTEXT

A. *Visual Intelligepnt Decision Support System*

Decision doesn't refer to a specific step clearly identified[8]. It is based on several phases defined in previous works such as[9],[10] and [11]which relyon Simon'sdecision process called ICDRbased on four phases: (1) *Intelligence* to extractrelevant information, (2) *Design* to generate a set of related models presenting different scenarios that may occur, (3)*Choice* to opt for one solution among theproposed scenarios, and (4) *Evaluation*aimingat reviewing the results found in all previous phases.

Decision processis integrated in decisional tools called decision support system (DSS)to assist domain experts tofind solutions for problems andmake decisions to improve or to adjust a current situation. The DSS developed tools suggested in previous works are applied in several domains. Among these works, we state [12] that proposed a DSS which assist user to find the best route in case of travel challenge. A decisional tool having as goalsplan and manage support in energy companies was proposed in [13]. Furthermore,a DSS relying on objective and subjective criteria to improve quality of service in digital library by generating a set of recommendations was suggested in [14] works.

The data analysis for decision-makingallows defining a set of parameters restriction to limit the search space. It provides more efficiency and clarity. Data mining algorithms gives patterns that will be analyzed by decision-makers. We are interested thus in data mining based DSS.

Several works were interested in integrating visualization methods and techniques in data mining for decision-making.In fact, it is recommended to integrate the Human in the data exploitation process, which is known as visual data mining. In this case, we ensure the integration of Human knowledge with the biggest computer capacity storage. Raw data, data mining process and generated patterns can be interactively and graphically presented to the user.

As consequence,numerous visual tools were initiatedin different context of use; e.g. CAST (Clustering And visualizing Spatio-Temporal data)[15] is a visualization tools that ensures moving entity control in order to study and evaluate them.A SIMID called tool developed by [16] provides decision-maker spatio-temporal visualization of infection. Visual temporal tool for distant monitoring was introducedMittelstädt and his colleagues [17].Based on this brief literature investigation, weset a first objective. It consists of seeking to make data-miner and decision-maker able to visualize data mining patterns, draw conclusions in real-time and interact with data in the different data miningsteps.

Visual data mining and DSSbecome more and more complex due to temporal data continuous progress.To face this complexity, the main problem can be divided into sub-problems; each one is assigned to a sub-system to reduce the complexity.So that, intelligence can be distributed into different parts of the system and their sub-systems.

B. *Intelligent agents technology*

The Multi-Agent technology consists in implementing a distributed intelligence in complex environment. Each task should undergo a localprocessing. An agentis a computer entity situated in an environment and capable of acting in an autonomous wayand can reach designed goals for which it was conceived [18]. We believe that each part of the whole complex and visual system can be assigned to a particular agent. Each agent belongs to one of the following categories:

1) Reactive agentdevoid from memory and environment representation, it relies on communication with agent's environment to solve problems [19]. It reacts according to its reflex without maintaining any internal state.

2) Cognitive agent: if the agent have a memory and able to realize an environmentsymbolic representation, and can take into account its past in order to reach an explicit purpose.

3) Hybrid agent: combines the twoprevious categories. The hybrid agent follows its plans. It can sometimes directly reacts to external events.

The agent technology, as we are mentioned previously, is based on autonomous andcognitive intelligent agents that havebeen applied in various domains to perform divers tasks. Agent technology is used on several DSS works to accomplish a define objective according the case. For exampleNeLH project [20] involves intelligent agents looking to find the available medical center of a given geographic area. A MAS based DSS was suggested by [21] to predict patient state according to context.

Contrary to the previous agent technology relatedworks; we will focus on proposing a set of cognitiveagents that aims to improve the effectiveness of the visual data mining based DSS. We call such system as ***v**isual **i**ntelligent DSS* (viDSS).So, our context dealswith viDSS based agent concepts that will be detailed in the next section.

III. viDSS Proposed Approach

In section2, we are studying works related to viDSS conceptsand agent technology. We are claimedthat, as far as we know,there is no work thatbrings them together. So we propose to consider then inproposing a new architecture of viDSS relying on agent technology.

A. *viDSS modules*

According to [22], theviDSS architecture is based on four principal modules (cf.Fig.1)

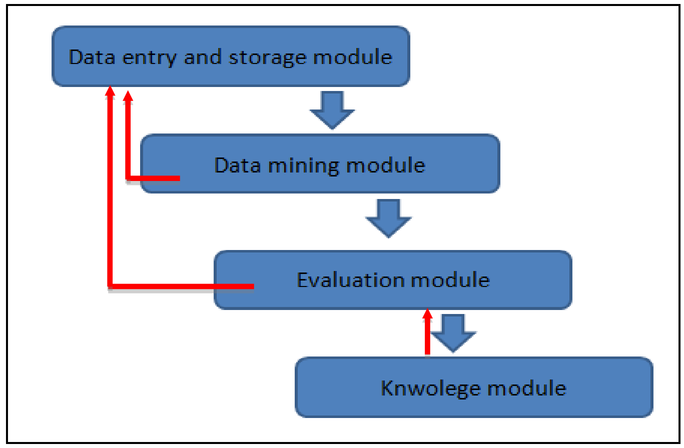

Fig. 1. viDSS principal modules

- Dataentry and storage module: this module consists ofdata selecting, pre-processing and transformation steps for data mining technique.
- Data mining module: in this module a data mining algorithm (e.g. association rules, Bayesian dynamic network, etc.) is applied totransformed data in order toextract useful patterns as output.
- Evaluation module: patterns provided by the previous module are evaluated. According to the evaluation result, it consists of moving to the next step or to make a feedback.
- Knowledge management module: after evaluation, in casepatterns are relevant and give a pertinent knowledge. This knowledge will be integrated for decision-making.

The viDSSprovides userwithvisualization of providedresults in different modules. For this reason, we added a new module to process the visualization tasks (c.f. Fig.2) called visualization module integrated in the four viDSS cited modules.

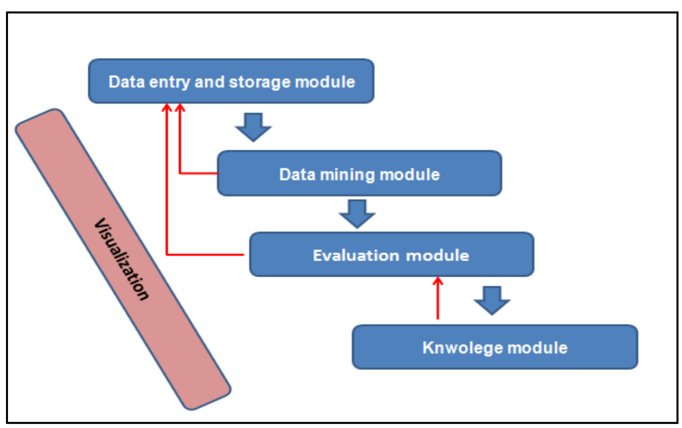

Fig. 2. Visualization module integration

B. Intelligent agent integrated on viDSS archirecture

After defining the different modules of viDSS architecture (cf. section III.A), we have to define the various intelligent agentsto be integrated,as well as the tasksassigned by each onein different modules (c.f. Fig.3).

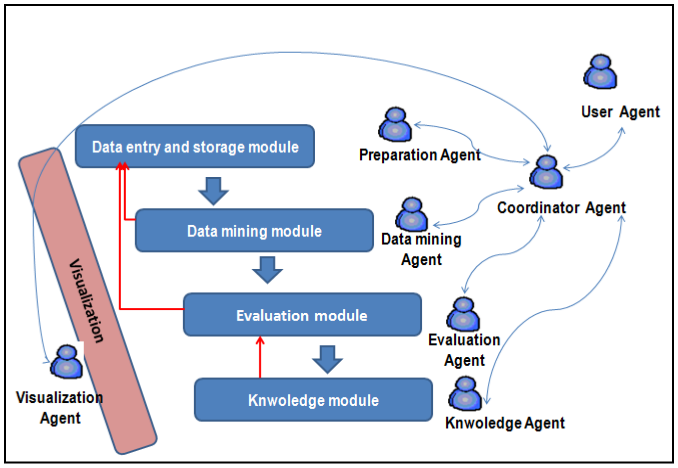

Fig. 3. viDSS integrated agents

The coordination between all the defined agents is controlled by in intelligent agent called “coordinator agent” (c.f. Fig.4).

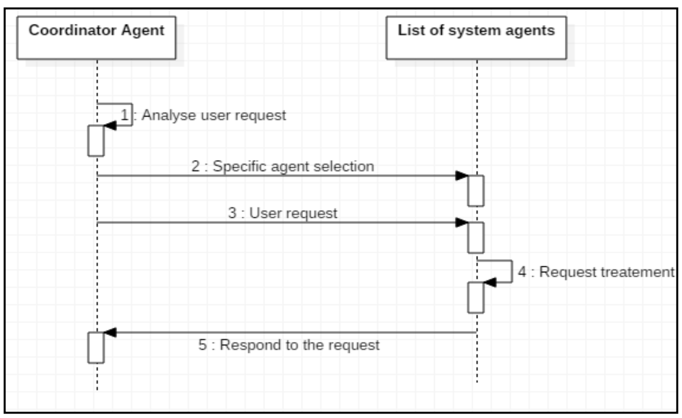

Fig. 4. Coordinator agent task

The coordinator agent is able to identify the user (Data-miner or Decision-maker), as well as the agent’s works progress and specify for each one the task to do. Following, we present the agents by module.

1) Data entry and storage module involved agent

In this first module, an agent called “data preparation agent” ensures the cleaning and the pretreatment tasks to resolve the missing valuesin the raw data (c.f. Fig.5).

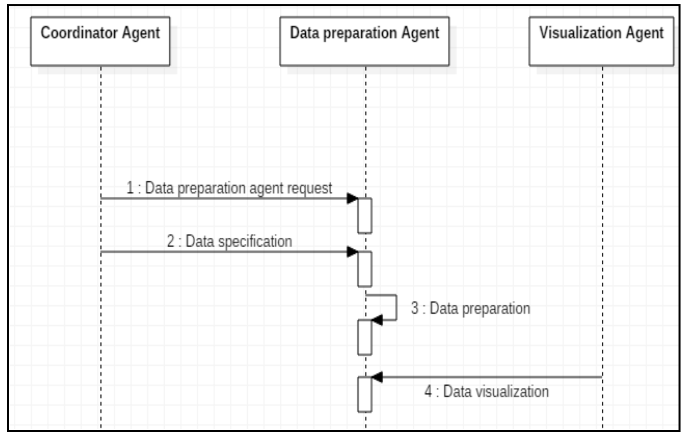

Fig. 5. Coordinator and data preparation agents interaction

The “data preparation agent” receives a request from “coordinator agent” to resolve the missing or invalid values in temporal data. The data preparation agent applies a specific algorithm to get back the missing values. The result of preparation process can be visualized by calling the “visualization agent”.Prepared data provided by this agent will be used in themodule of data mining.

2) Data mining module involved agent

After receiving prepared data from “data preparation agent”, the “data mining agent” applies a data-mining algorithm togeneratepatterns that allow the futureprediction(c.f. Fig 6).

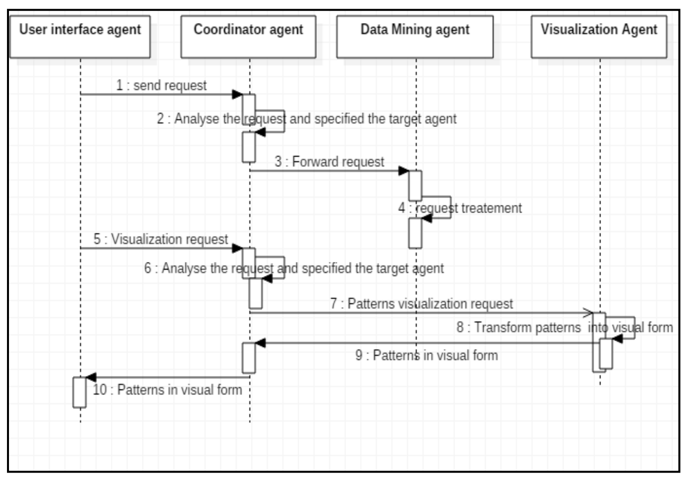

Fig. 6. Data mining agent interactions

The user (i.e. data miner / decision-maker) is able to visualize results of data treatment in different modules thanks to the “visualization agent” providing an easier interaction between user and system.

3) Evaluation module involved agent

An agent called "evaluation agent" achieves the patterns evaluation task.It ensures also thefeedback to the previous task if evaluation results are not valid by collaborating with the "coordinator agent"(c.f. Fig.7).The evaluation is realized through a set of criteria learned by the agent.

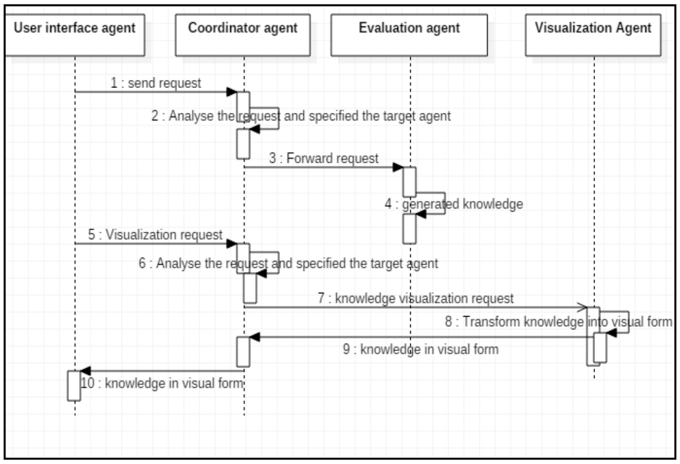

Fig. 7. Evaluation of data mining patterns

4) Knowledge management module

Knowledge management task is assigned to "knowledge agent".It has two roles; first it generates knowledge from patterns and second integrates it in knowledge base in case of it was validated by the "evaluation agent", else knowledge will be rejected and a new knowledge generation is started until we got a validate knowledge (c.f. Fig.8).

User is able to visualize the extracted knowledge, so he/she indicates the needs to visualize the extracted knowledge from models. As consequence, a communication between "Visualization agent" and "knowledge agent" occurred in order to present knowledge in a visual form.

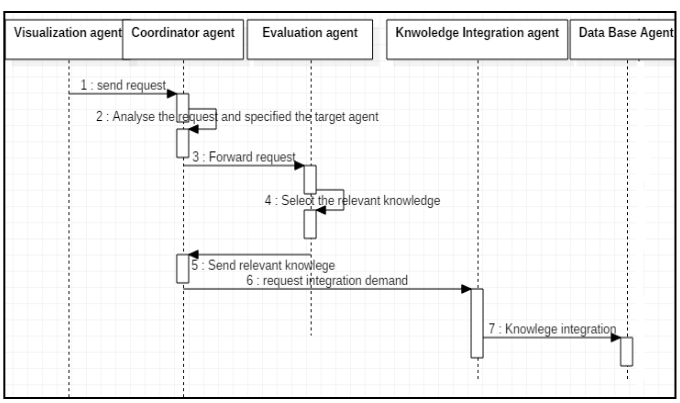

Fig. 8. Knowledge and evaluation agents interactions

C. Agents interaction modeling

The viDSS involved agents are cognitive, they can learn from their previous acts. Furthermore they have the capacityto discover environment in which they belong. Theycommunicate with each other to carry out a specific task(c.f. Fig.9).

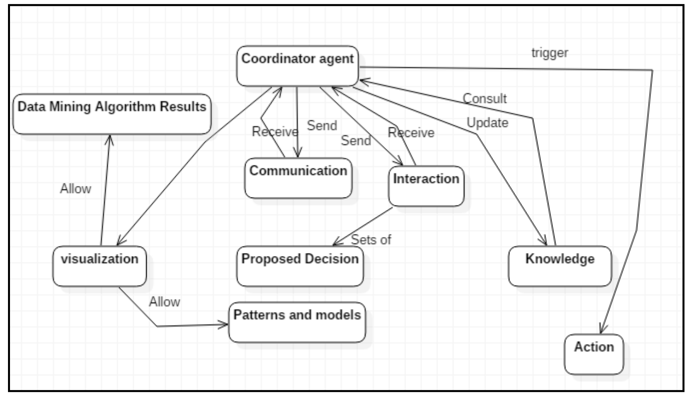

Fig. 9. Intelligent agent interaction

Each agent can send and receive message for another, it is a collaborative environment that looks for improving results quality and make the viDSS architecture center-user.

The « coordinator agent » is the principal agent of the system. It ensures the communication and the interaction between all the system agents. It receives requests and sends responses from/to agents according to the request nature. It aims to schedule the agent tasks to reach the objectives. The "coordinator agent" facilitates the interaction and collaboration between the user and the viDSS. It makes possible a quick and an easier visualization of data, patterns and generated knowledge.

IV. viDSS Developped Tool

The NI detection is acomplex process and presents a challenge because itis based on analysis and interpretation of temporal medicaldata such as performed actsand antibiotics doses. This complexity is proved by a study realized onhospitalized population that contains 280 patientsbetweenApril 17th 2002 (midnight) and April 18th2002 (midnight). This studyshown that 18 % of total number has beenaffected bythe NI [23].The NI detection is a complex task due to the big number of temporal data analyses continuously (e.g. antibiotic doses, clinical tests values, etc…). In order to supervise the NI, physicians are based on prevalence survey given by the means of several tools and devices. The divertools give different data values with different types, so that the temporal data visualization is not an easier task. Several works havepublished such as that of [24]and [25] whichlook for real solutions to decrease the number of affected persons and to detect infection in early stages. In spite of those efforts there is a lack ofdeveloped toolsbased on MAS for thepredictionofNIconcurrence.

In order to validate the intelligent agent based viDSS architecture introduced in the previous section, we have developed a prototype called "viDSS_NosDetc" for the physicians of the Intensive Care Unit (ICU) of theteaching hospitalHbibBourguibaSfax, Tunisia. This tool enables physicians toidentify infected patients inhospital intensive careunitthanks to: (1) visual data mining ofthe temporal data, providing knowledge and (2) the obtained results of temporal data interpretation.In addition, we intended toguarantee the tool capacity toimproveinteraction between user and system by meansof visualization of handled data and patterns. The applied data mining technique is the Dynamic Bayesian Networks (DNB) [26], which relates variables to each other over adjacent time phases on any day, with t is between patient entry date and leaving date(in days)(c.f. Fig.10). We present the development of our MAS based viDSS architecture by module.

A. *Data entry and storage module development*

We have first developed a “database agent” to ensure the collection and the storage of temporal ICU data in the database. The collected data can contain missing values, which need a specific treatment. Based on “data preparation agent” intelligence, the missing values are recovered.

B. *Data mining module development*

The “data mining agent” applies the DBN algorithm in order to generate model that represent the pertinent information. As mentioned above, we have developed the DBN algorithm. The causal graph of the developed algorithm is shown in the following figure (c.f. Fig 10).

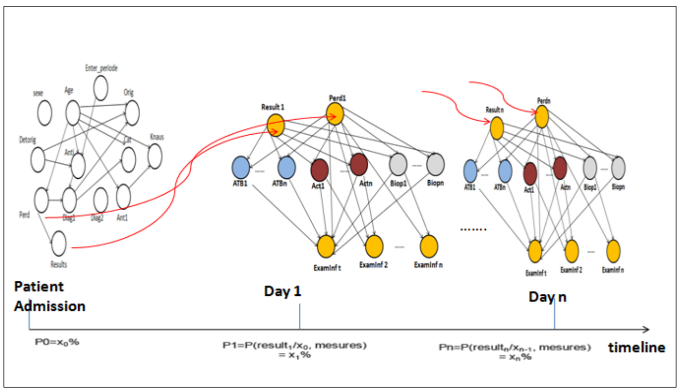

Fig. 10. Causal dependencies in a dynamic BN

The variables used as input of DBN is presented in the following table (cf. Table I).

TABLE I. TEMPORAL VARIABLES

Measure	*Description*
Result	The NI occurrence probability
Hospital stay	The period between the patient admission and exit from the ICU
Antibiotic (ATB)	The daily antibiotic catch during the patient stay in the ICU.
Infectious examination (ExamInf)	The daily performed infectious examination
Acts	The daily carried out acts in the ICU
Biological parameters (BioP)	The daily measured biological parameters

C. *Evaluation module development*

The generated patterns resulting from the data-mining agent are evaluated by the evaluation agent. We have developed an automatic data preparation algorithm that allowsto analyze the data and identifying corrections, eliminates problematic or unnecessary fields, deriving new attributes when necessary and improves performance with intelligent scanning techniques.

D. *Knowledge managment module development*

The developed knowledge agentalgorithm integrates the evaluated pattern as new element (NI occurrence probability) in the knowledge basefor further decision-making. Based on this probability, viDSS suggests solutions. The system output helps physicians to make the best decision in real-time.

The following figure (c.f. Fig.11),is the main interface ofour viDSS prototype.

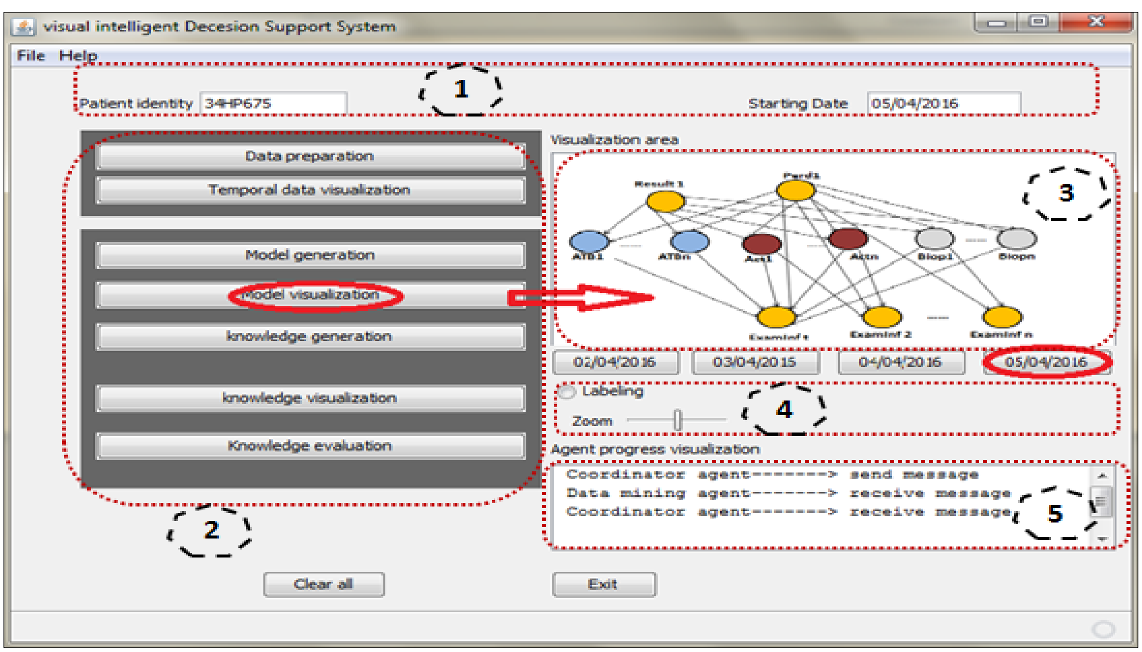

Fig. 11. viDSS tool

The user (decision-maker) specifies at the beginning the set of parameters related to the supervised patient (c.f. Fig.11(1)), then he/she chooses the task to do by clicking on the specific button (C.f. Fig.11(2)). Each task is performed by an assigned cognitive agent. The visualization task provides both temporal data (c.f. Fig.11(3)) and mining results visualization (c.f. Fig.11(4)).The user gets moredetailed abstract of temporal data variation at different time granularities by choosing one among interaction options(c.f. Fig.11(6)). After having achieved the mining tasks, interestingpatterns areextracted (c.f. Fig.11(5)) and integrated for possible decision solutions suggestion. If the physician acceptedthe suggested decision and the evaluation resultshave been acceptable, knowledge will be integrated in knowledge base to be used for future decision-making.

V. Evaluation

After developing the "viDSS_NosDetc" tool, we move to its validation that deals with utility and usability evaluation[27] [28] [29] [30] tomake sure that the user is satisfiedwith the provided services.We begin with the first evaluation dimensionaiming at measuring the system performance through confusion matrix.

A. Utility evaluation

The utility evaluation intended to verify if results provided by different agents, such as data preparation, data mining and data knowledge agents are significant by considering the observed NI probability. This comparison is realized by means ofconfusion matrix application.

TABLE II. Confusion matrix

MAS-VIDSS			
		Observed results	
		Yes	No
Predicted results	Yes	10	4
	No	4	46
The accuracy rate: 79%			
The negative capacity of prediction: 87%			
The positive capacity of prediction: 74%			
Previous version of the system [31]			
		Observed results	
		Yes	No
Predicted results	Yes	9	8
	No	7	34
The accuracy rate: 77%			
The negative capacity of prediction: 90%			
The positive capacity of prediction: 60%			

As visible in the table 2, ourdeveloped viDSStoolprovided significant and interesting results.The tool's usability evaluation consists of assessing the visualization agent results in terms of visual representations.

B. Usability evaluation

The evaluation test checks the quality tool usability. To be sure from the user satisfaction, so we asked him/her about the degree of satisfaction (using a questionnaire). The evaluation test gives the following results visible in the figure 12:

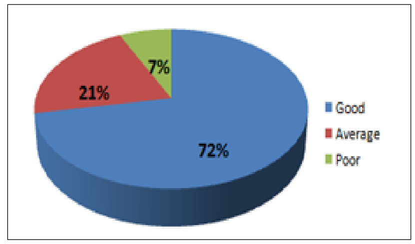

Fig. 12. Evaluation rate

This usability evaluation shows that 72% of the interviewees judged that the system was ''Good'' for decision-maker or data-miners (cf. Fig 12). We can conclude that a great number of users are satisfied with tool interfaces and the ergonomics of system.

VI. Conclusion

In the following paper, we have focused on the design and development of visual Intelligent Decision Support based on Multi-Agent architecture. The used agents are intelligent and cognitive.For each agent we have affecteda specific KDD task. The viDSS developed agents communicate using the coordinator agent. The proposed architecture make user able to easily interact with the viDSSoutputs (e.g. temporal data, patterns and knowledge) using the visualization agent.

The suggested approach is implemented for the development of amedical tool called "viDSS_NosDetc". It is used to detect the NI for the patients residing in the Intensive CareUnit of the teaching hospital HbibBourguibaSfax Tunisia. The viDSSutility and usability evaluation proves that the developed tool satisfies its users.

In our future works, we intend toextend the current architecture to support the storage and the treatment of the big data concepts. As a second contribution, we intend to develop the viDSS architecture to be applied on other distributed environments.

Acknowledgment

The authors would like to acknowledge the financial support of this research by grants from the ARUB program under the jurisdiction of the General Direction of Scientific Research (DGRST) (Tunisia).

References

[1] H. Ltifi, G. Trabelsi, M. Ben Ayed, and A.M. Alimi, "Dynamic Decision Support System Based on Bayesian Networks, application to fight against the Nosocomial Infections", International Journal of Advanced Research in Artificial Intelligence (IJARAI), Vol. 1, no. 1, pp. 22_29, 2012.

[2] K. Rajesh, A.K. Pujari and D.S. Reddy, "Clustering techniques in data mining- A survey", *IETE Journal of Research*, Vol. 47, no. 1, pp. 19_28, 2001.

[3] R. Bose, and V. Sugumaran, V. "Application Of Intelligent Agent Technology For Managerial Data Analysis And Mining", *Data Base*, Vol. 30, pp. 77_94, 1999.

[4] T. Rahwan, T. Rahwan, I. Rahwan, and R Ashri, (Eds.), Agent-Based Support For Mobile Users Using Agentspeak (L), Germany, Lncs, Springer, 2004.

[5] D. Keim, "Information Visualization and Visual Data Mining", *IEEE Transactions on Visualization and Computer Graphics*, Vol. 8, no. 1, pp. 1_8, 2002.

[6] H. Ltifi, E. Ben Mohamed and M. Ben Ayed, "Interactive visual KDD based temporal Decision Support System", Information Visualization, Vol. 14, no. 1, pp. 1_20, 2015.

[7] E. Ben Mohamed, H. Ltifi and M. Ben Ayed, "Using visualization techniques in knowledge discovery process for decision-making", The 13th International Conference on Hybrid Intelligent Systems (HIS 2013), Tunisia, pp. 94_99, 2013.

[8] A.Akharraz, Acceptabilité de la décision et risque décisionnel : Un système explicatif de fusion d'informations par l'intégrale de Choquet. Thèse de doctorat, Universitéde Savoie, 2004.

[9] P.Lévine, et J.C. Pomerol. Systèmes interactifs d'aide à la décision et systèmes experts, Hermès, 1989.

[10] E. Turban, " Decision Support and Expert Systems ". Macmillan, New York, 1993.

[11] D.J.Power,"Decision support systems: concepts and resources for managers".Westport, Conn., Quorum Books, 2002.

[12] Y.Xie,H.Wang,J.Efstathiou,A research frame work for web-based open decision Support systems, Knowlege.-Based Syst.18, 2005, pp.309–319.

[13] H.K.Bhargava,D.J.Power,D.Sun,Progressinweb-based decision support technologies, Decis.SupportSyst.43(2007)1083–1095.

[14] F.J.Cabrerizo,J.A.Morente,Molinerab,I.JPérezc,J.LópezGijónd,E.Herrera-Viedma, "Adecisionsupportsystemtodevelopaqualitymanagementinacademicdigitallibraries", Journal of Information Sciences, 2015, pp. 48–58

[15] H. Munaga, L. Ieronutti, and L. Chittaro, "CAST: A Novel Trajectory Clustering and Visualization Tool for Spatio-Temporal Data", Proceedings of the First International Conference on Intelligent Human Computer Interaction, 2009 , pp 169-175.

[16] L. Ramírez, Y. R. Gel, M. Thompson, E. de Villa and M. McPherson, "A new surveillance and spatio-temporal visualization tool SIMID: SIMulation of Infectious Diseases using random networks and GIS", Computer Methods and Programs in Biomedicine, vol. 110, no 3, 2013, pp. 455-470

[17] S. Mittelstädt, X. Wang, T. Eaglin, D. Thom, D. Keim, W. Tolone and W. Ribarsky, "An Integrated In-Situ Approach to Impacts from Natural Disasters on Critical Infrastructures", IEEE 48th Annual Hawaii International Conference on System Sciences, HICSS-48, 2015, Los Alamitos, CA: IEEE Computer Society Press.

[18] M.Wooldridge, "An Introduction to MultiAgent Systems". London: John Wiley& Sons,2002.

[19] J. Ferber. "Multi-Agent Systems : An Introduction to Distributed ArtificialIntelligence". Addison-Wesley Longman Publishing Co., Inc., Boston, MA, USA, 1999.

[20] P. Kostkova, J. ManiSaada, G. Madle, J.Weinberg. "Applications of Software Agent Technology in the Health Care Domain", (Ch. AgentBased Uptodate Data Management in National Electronic Library for Communicable Disease, (2003) pp. 105-124.

[21] M. Tentori, J. Favela, M. Rodriguez, "Privacyaware autonomous agents for pervasive healthcare", IEEE Intelligent Systems 21 (6), pp. 55-62.,2006.

[22] H. Ltifi, G. Trabelsi, M. Ben Ayed, and A.M. Alimi, "Dynamic Decision Support System Based on Bayesian Networks, application to fight against the Nosocomial Infections", International Journal of Advanced Research in Artificial Intelligence (IJARAI), Vol. 1, no. 1, , 2012, pp. 22_29.

[23] H.Kallel, M.Bouaziz, H.Ksibi, H.Chelly, C.B.Hmida, A.Chaari, N.Rekik, et M.Bouaziz, "Prevalence of hospital-acquired infection in a Tunisian Hospital". Journal ofHospital Infection, vol. 59, 2005, pp. 343-347.

[24] S.Gafsi-Moalla,« Les infections acquises en réanimation, étude prospective réalisée dans le service de réanimation de Sfax sur une période de 3 mois ». Thèse de doctorat, Faculté de Médecine de Sfax, Tunisia, 2005.

[25] L. Hergafi, « Présentation et validation d'un nouveau système pour la surveillance de l'infection acquise en réanimation ». Thèse de doctorat, Faculté de Médecine de Sfax, Tunisia, 2006.

[26] A. Darwich, "Constant-space reasoning in dynamic Bayesian networks", International journal of approximate reasoning, Vol. 26, 2001, pp. 161-178.

[27] J. Nielsen, "Usability Engineering", Academic Press, Boston, 1993.

[28] M.Y. Ivory and M.A. Hearst. "The state of the art in automating usability evaluation of user interfaces", ACM Computing Surveys, vol. 33(4), 2001,pp. 470-516.

[29] A. Sears and J.A. Jacko, editors. "The Human Computer Interaction Handbook: Funda- mentals, Evolving Technologies and Emerging Applications". Lawrence Erlbaum Associates, Mahwah, NJ, 2nd edition, 2008.

[30] R.Shibl, M.Lawley, J.Debuse, "Factors influencing decision support system acceptance", Decision Support Systems, Vol. 54 (2),2013, pp. 953–961.

[31] D. Godoy, S. Schiaffino, and A. Amanda, "Interface Agents Personalizing Web-Based Tasks", *Cognitive Systems Research,* Vol. 5, pp. 207_222, 2004.