Ontology-Based Clinical Decision Support System for Predicting High-Risk Pregnant Woman

Umar Manzoor

Faculty of Computing and Information Technology
King Abdulaziz University
Jeddah, KSA

Mohammed A. Balubaid

Industrial Engineering Department, Engineering Faculty
King Abdulaziz University
Jeddah, KSA

Muhammad Usman

Department of Computer Science
National University of Computer and Emerging Sciences
Islamabad, Pakistan

Ahmed Mueen

Faculty of Computing and Information Technology
King Abdulaziz University
Jeddah, KSA

Abstract—**According to Pakistan Medical and Dental Council (PMDC), Pakistan is facing a shortage of approximately 182,000 medical doctors. Due to the shortage of doctors; a large number of lives are in danger especially pregnant woman. A large number of pregnant women die every year due to pregnancy complications, and usually the reason behind their death is that the complications are not timely handled. In this paper, we proposed ontology-based clinical decision support system that diagnoses high-risk pregnant women and refer them to the qualified medical doctors for timely treatment. The Ontology of the proposed system is built automatically and enhanced afterward using doctor's feedback. The proposed framework has been tested on a large number of test cases; experimental results are satisfactory and support the implementation of the solution.**

Keywords—High-risk patient; Pregnant woman; Ontology-based CDSS; Clinical Decision Support System

I. INTRODUCTION

The world has a shortage of professional medical doctors; even in the most developed countries have inadequate position regarding the availability of medical doctors. According to Association of American Medical Colleges, U.S is facing a shortage of approximately 20,000 medical doctors [1]. In Pakistan, the situation is even much worse. According to Pakistan Medical and Dental Councils (PMDC), Pakistan is facing a shortage of approximately 182,000 medical doctors [2]. Because of this shortage, pregnant women are also affected due to lack of proper and timely treatment, which increased the mortality rate of pregnant women over the years. According to world health organization (WHO), almost 500,000 women die every year from pregnancy-related complications [3]. Especially focusing on Maternal Mortality Rate, it was observed that most of the deaths occur because of few basic complications. Moreover, these complications can be easily treated once the reasons are diagnosed; therefore, the major problem is the unavailability of proper diagnosis because of the shortage of medical doctors.

The four high risk pregnancy complications handled in this work are Hypertension, Obstructed Labor, Septicemia and Hemorrhage. Increase in blood pressure during pregnancy indicates Hypertension, Obstructed Labor is an anomaly that may arise during the process of labor, Septicemia pollutes the patient's blood and may occur due to infections caused by bacteria and Hemorrhage occurs due to excessive loss of blood from the patient's body.

To overcome this problem, we have proposed ontology-based clinical decision support system which can partially work in place of doctors to diagnose high-risk pregnant woman and refer them to the qualified medical doctors for treatment. This way, the high-risk patients will get proper treatment well in time, and many lives can be saved. The main focus of this system is to build a diagnostic procedure which can work independently of qualified doctors and identify high risk patients; Once these patients are identified, they can be treated by medical doctors. So our system will help in reducing the workload of doctors as well as providing basic health care to more and more patients. The framework is composed of three components: 1) Automatic Ontology Construction, 2) Feedback System and 3) Ontology Enhancement Component.

The rest of the paper is organized as follows. In section 2, existing work in ontology based clinical decision support system and automatic construction of ontology is discussed. In section 3, the proposed approach is discussed. Experimental results are presented in section 4. Finally, the conclusion is drawn in section 5.

II. LITERATURE REVIEW

A. Ontology Based CDSS

There are many mistakes made on regular basis by humans in clinical environments. Hazmy Iman Abas et al in [16] have identified the three common mistakes made by clinicians that are; they failed to meet guidelines, they are not educated on regular basis and they are not aware of their responsibilities. According to the authors, these mistakes can easily be avoided by the use of ontology based clinical decision support system. The early detection of Alzheimer Disease is a challenging task in medical domain. Eider Sanchez et al in [17] proposed an ontological CDSS approach to detect Alzheimer in early stages. In this system multidisciplinary knowledge is used (i.e.

the system uses three ontologies that are SWAN, SNOMED CT and MIND). SWAN is used for the diagnosis of Alzheimer Disease, SNOMED CT's purpose is standardization and MIND ontology is used to carry out patient tests. Farahidayah Bt. Mahmud et al in [15] designed a CDSS that finds the right time of weaning a patient from ventilator. According to the authors, the proposed Ontology based CDSS is very helpful because of the ontology's ability of presenting complex concepts, reusability and specification of shared conceptualization. Similarly, Adnan et al in [18] proposed ontology based clinical decision support system to assist electronic discharge summary (EDS) while prescribing the patient's medications. Matt et al in [20] developed an ontology based CDSS for preoperative risk assessment. The proposed system takes the patient's data as input, stores it in database and also passes it to Ontology Modeler for conversion into OWL format (i.e. the same record is stored in two different formats). Afterwards, inference is performed on both the database and ontology. Rule engine is used on the database to calculate the numeric scores (i.e. cardiac scores) whereas classification algorithm is run on the ontology to assign category to the patient, the results of both are then combined to calculate the patient preoperative risk assessment.

B. Automatic Construction of Ontology

Abd-Elrahman Elsayed et al in [19] used data mining technique (c4.5 decision tree) on structured data to construct ontology automatically. In this approach, the authors proposed decision tree to ontology mapping where the tree decision nodes are mapped to ontology classes and leaf nodes are mapped to individuals. The authors tested the proposed approach on the soybean disease dataset and showed the efficiency of the same. Seongwook Youn et al in [20] proposed architecture for the classification of emails as spam or legitimate using ontology based approach. The authors created a dataset D based on the features of the email (i.e. spam email) and used WEKA (J48 decision tree) to generate decision tree which afterwards is passed to JENA for conversion into RDF ontology format. Authors divided the original dataset into two parts (i.e. training and testing dataset) and tested the proposed approach on testing dataset; according to the authors the results are satisfactory. Amit Bhagat et al [21] used association rule mining to construct ontology from large transaction databases. In the proposed approach, multiple level association rule mining is used to extract more specific and relevant knowledge as compared to single level rule mining. Patrick Clerkin et al in [22] proposed automatic construction of ontology using the COBWEB algorithm. COBWEB is a clustering algorithm which creates different clusters of the data in a hierarchical manner. In the proposed technique, the hierarchy of the clusters given by COBWEB is mapped into ontology classes in such a way that parent cluster(s) is mapped as parent or super class(es) in the ontology whereas the sub or child cluster(s) is mapped as sub-class(es) in the ontology. Henrihs Gorskis et al in [23] reviewed 1) the work done in the field of ontology building using data mining techniques and 2) the potential of different techniques in the construction of ontology. According to the authors, the ontologies created with data mining technique(s) may be inferior to those constructed manually.

III. PROPOSED TECHNIQUE

In this paper, we have proposed a framework for predicting high risk woman using ontology based CDSS. The framework is composed of three main modules: 1) Automatic Ontology Construction 2) Feedback System 3) Ontology Enhancing Process.

A. Automatic Ontology Construction

This module automatically constructs the ontology of high risk pregnant woman using pregnant woman dataset and is composed of two sub-components namely Rules extractor and Rules to Ontology Mapper. Rules extractor extracts the rules from the pregnant woman dataset using WEKA and is developed in Java. WEKA API is used to call WEKA functions from the program. The dataset is given as input to this component which uses WEKA API to extract rules from the dataset. The dataset is given in the ARFF (Attribute-Relation File Format) format which is compatible with WEKA. ARFF format has two sections, the first section is Header which contains the name of relation, the attributes of the relation and the attributes' data types whereas the second section has Data which contains the real instances of the relations. Example of partial IRIS dataset in ARFF format is given below.

Header Section:
```
@RELATION iris
    @ATTRIBUTE sepallength NUMERIC
    @ATTRIBUTE sepalwidth NUMERIC
    @ATTRIBUTE petallength  NUMERIC
    @ATTRIBUTE petalwidth   NUMERIC
    @ATTRIBUTE class        {Iris-setosa,
Iris-versicolor, Iris-virginica}
```
Data Section:
```
@DATA
    5.1, 3.5, 1.4, 0.2, Iris-setosa
    4.9, 3.0, 1.4, 0.2, Iris-setosa
    4.7, 3.2, 1.3, 0.2, Iris-setosa
    4.6, 3.1, 1.5, 0.2, Iris-setosa
    5.0, 3.6, 1.4, 0.2, Iris-setosa
```

Once the dataset is loaded, attribute selection algorithm is called using WEKA API. We have used the genetic search for attribute selection and once the most relevant features are selected, JRIP algorithm is called to extract rules which later are used for the construction of ontology as summarized in figure 1.

Rules to Ontology Mapper is responsible to construct ontology using the rules extracted in the first component as shown in figure 2. It first creates ontology classes for each dataset (i.e. hypertension, Obstructed labor, Hamorhage, Septicemia) then the properties are created based on the attributes found in the rules. Furthermore, the range and domain of each property is set based on the information provided by each dataset. Afterwards, for each class, its definition is created that reflects the classification criteria for that class.
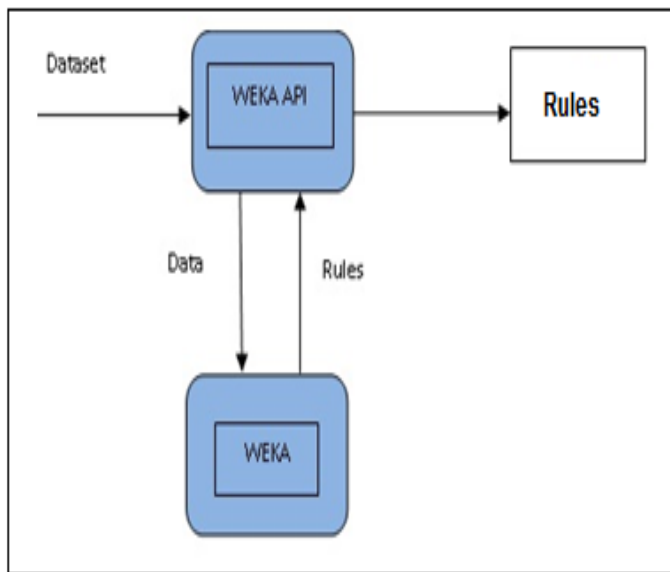
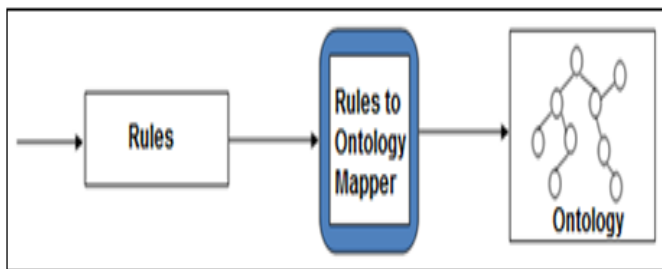Fig. 1.   Rules Extractor

Fig. 2.   Rules to Ontology (R2O) Mapper

Figure 3 shows the class definition (containing all attributes and their values) for Obstructed Labor in Protégé.
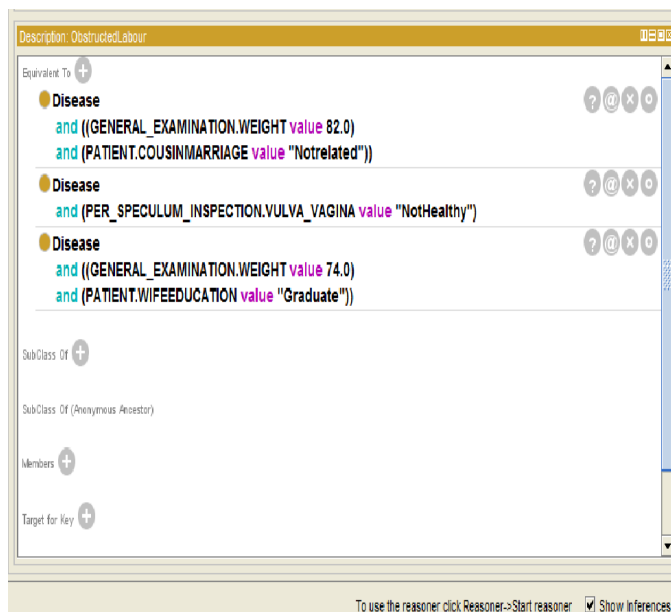

Fig. 3.   Obstructed Labor definition in Protege

B. Feedback System

Once a patient is predicted as high risk by our proposed system and recommended to the doctor for treatment, feedback is taken from the doctor whether the patient was correctly identified or not. If the doctor overrules the decision made by the system, the attributes which caused the decision are shown to the doctor. He can add / delete the attribute(s) or tune their parameters / range (i.e. change the values of an attribute to more suitable one). Afterwards, the overruled instance with tuned parameters or different attributes (if any) is added to overruled dataset. Once the number of instances in overruled dataset reaches β (configurable), this dataset is passed as input to Ontology Enhancing Process for updating the ontology. In this way, doctor's knowledge is incorporated in the system and ontology is enhanced accordingly as shown in figure 4.

C. Ontology Enhancing Process

According to Gruber [23], ontology represents knowledge as a hierarchy of concepts and their relation for a specific domain. Furthermore, ontology knowledge representation is based on the concept of conceptualization which can be defined as the objects, the concepts and the relationship that hold among them [23]. New concepts, their properties and relationship can easily be incorporated in ontology knowledge base. Therefore, once the number of overruled instances reaches β, JRIP algorithm is executed on overruled dataset to extract new knowledge, and the same is updated in the ontology either by creating new concepts or by defining new relationship between the old concepts. The complete architecture of our proposed Ontology based Clinical Decision Support System for predicting high risk pregnant woman is shown in figure 4.

IV.   EXPERIMENTAL RESULTS

A series of tests have been carried out in order to demonstrate that the proposed system is working properly. The tests have been subdivided in two classes: 1) Automatic construction of ontology, and 2) Ontology enhancement Process. Each of the experiment and its results are described in following:

A. Automatic Construction of Ontology

This experiment is designed to verify the automatic construction of ontology (i.e. how well the rules are mapped into ontology). For this purpose, the ontology and existing rule-base system are tested on the same datasets with the same training and testing ratio. In this experiment, we used four dataset (for each output class i.e. Hypertension, Obstructed Labour, Hamorhage, Septicemia) where each dataset is divided into two parts (i.e. 70% and 30%) where 70% of data is used for training and 30% is used for testing. The experimentation result showed that ontology and rule based system have same accuracy which means the rules are transformed with 100% accuracy into the ontology as shown in figure 5.

B. Ontology enhancement Process

Enhancing the ontology is the most important aspect and has been deeply validated. In this activity, fifteen doctors from five different hospitals of Pakistan participated, whenever the

patient is identified as high risk by our system, the patient is referred to one of the fifteen doctors. If the doctor overruled the decision, feedback from the doctor is taken and the instance is added to overrule dataset (in this case the β value is 25). When the instances in the overruled dataset reach β, the new

knowledge is incorporated in the ontology. We have observed that after two rounds of ontology updation the accuracy of the system increases and the false positives are decreased significantly as shown in figure 6.
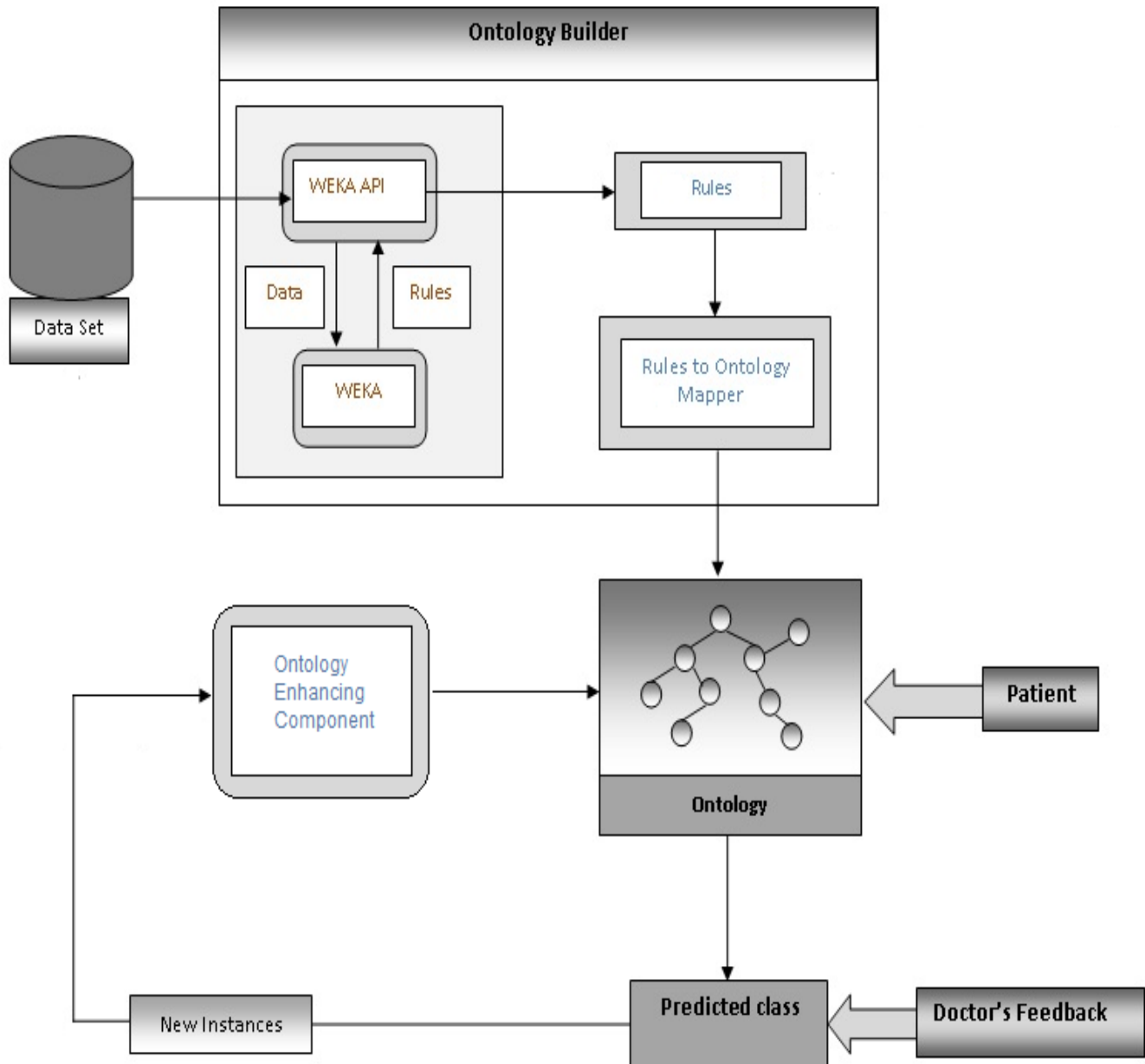

Fig. 4. System architecture of the proposed framework

V. CONCLUSION

According to world health organization (WHO), almost 500,000 women die every year from pregnancy-related complications. Moreover, these complications can be easily treated once diagnosed; however, the major problem is unavailability of proper diagnosis because of shortage of medical doctors. In this paper, we have proposed an ontology-based CDSS for diagnose high-risk pregnant woman and refer them to the qualified medical doctors for treatment. The

proposed framework is tested on a large number of test cases, results are satisfactory and support the implementation of the same. The work can be extended in many directions; one possible direction is including more pregnancy-related disease in the ontology. Second direction is automatic ontology enhancement process, currently based on expert (doctor) feedback the ontology is enhanced, machine learning algorithms should be incorporated in the proposed system to enhance ontology automatically.
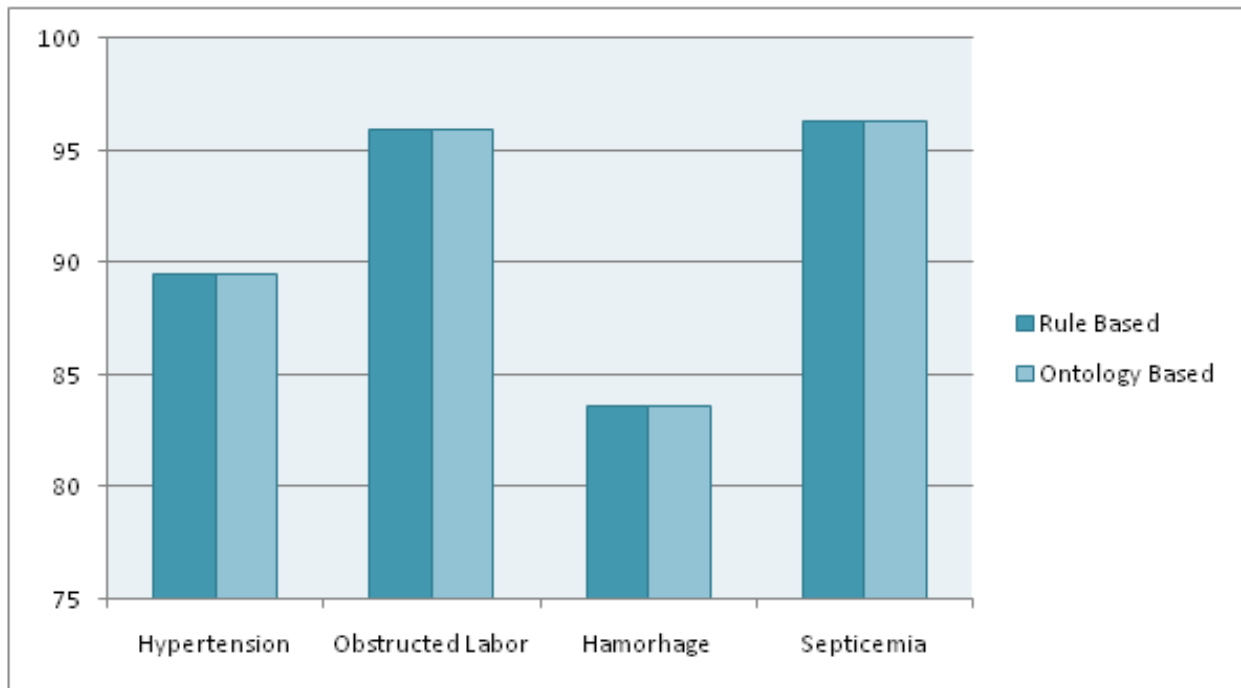
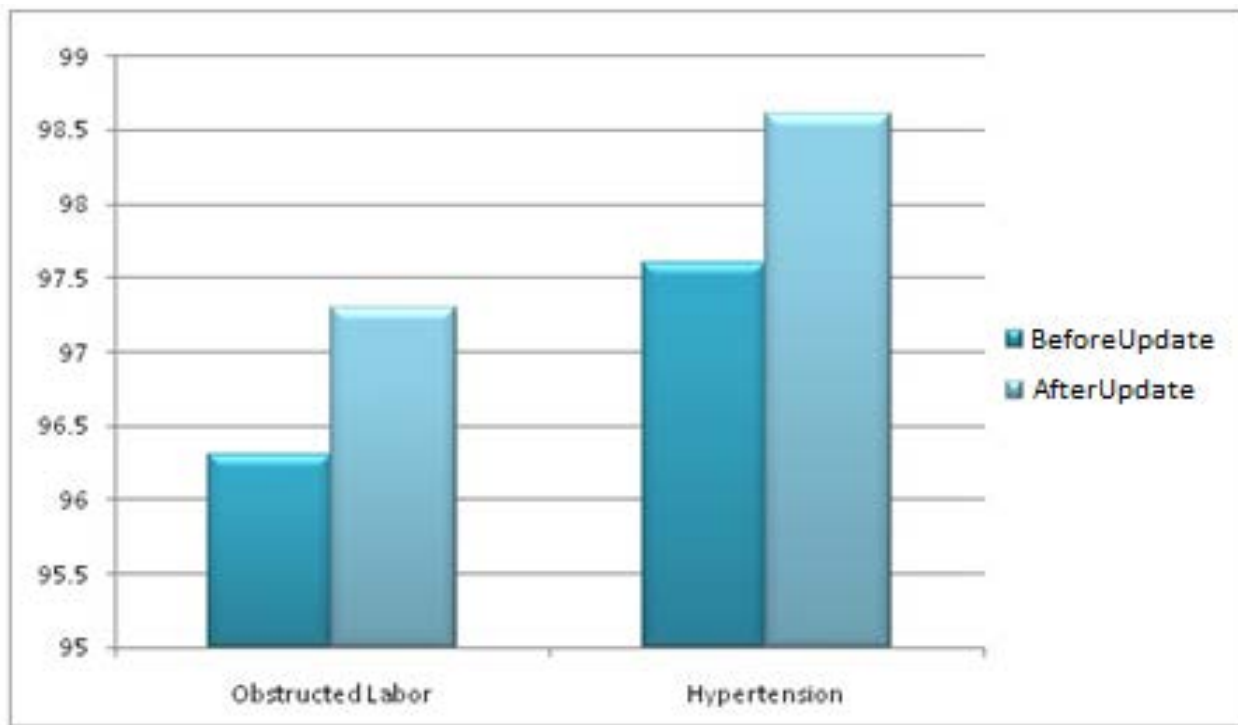Fig. 5.   Accuracy of Rule Based and Ontology Based system

Fig. 6.   Classification accuracy comparison before and after updating the ontology

REFERENCES

[1]   American Association of American Medical Colleges. (November, 2011), 2011 State Physician Workforce Data Book.

[2]   Pakistan Medical and Dental Councils, http://www.pmdc.org.pk/

[3]   U. Fund, "State of the Worlds Children 2009. Maternal and Newborn Health," New York: UNICEF, vol. 60, 2008.

[4]   Umar Manzoor, Samia Nefti, Yacine Rezgui, "Categorization of malicious behaviors using ontology-based cognitive agents", Data & Knowledge Engineering, Volume 85, May 2013, Pages 40–56.

[5]   K.L. Clark, F.G. McCabe, "Ontology schema for an agent belief store", International Journal of Human-Computer Studies, Volume 65, Issue 7, July 2007, Pages 640-658.

[6]   Protégé, http://protege.stanford.edu/

[7]   WEKA Tool, http://www.cs.waikato.ac.nz/ml/weka/

[8]   B. Orgun, J. Vu, "HL7 ontology and mobile agents for interoperability in heterogeneous medical information systems", Computers in Biology and Medicine, Volume 36, Issues 7–8, July–August 2006, Pages 817-836.

[9]   Quynh-Nhu Numi Tran, Graham Low "MOBMAS: A methodology for ontology-based multi-agent systems development", Information and Software Technology, Volume 50, Issues 7–8, June 2008, Pages 697-722.

[10] C. Su and C. Yang, "Feature selection for the SVM: An application to hypertension diagnosis," Expert Systems with Applications, vol. 34, no. 1, pp. 754–763, 2008.

[11] Y. Chae, S. Ho, K. Cho, D. Lee, and S. Ji, "Data mining approach to policy analysis in a health insurance domain" International journal of medical informatics, vol. 62, no. 2-3, pp. 103–111, 2001.

[12] A. Tanwani, J. Afridi, M. Shafiq, and M. Farooq, "Guidelines to Select Machine Learning Scheme for Classification of Biomedical Datasets," in Proceedings of the 7th European Conference on Evolutionary Computation, Machine Learning and Data Mining in Bioinformatics. Springer, 2009, pp. 128–139.

[13] Umar Manzoor, Samia Nefti "iDetect: Content Based Monitoring of Complex Networks using Mobile Agents", Applied Soft Computing 12 (5), 1607-1619.

[14] Farahidayah Bt. Mahmud, Maryati Mohd Yusof, Shahrul Azman Naoh. "Ontological Based ClinicalDecision Support System." Electrical Engineering and Informatics (ICEEI), 2011 International Conference on. Bandung, 17-19 July 2011.

[15] Hazmy Iman Abas, Maryati Mohd. Yusof, Shahrul Azman Mohd Noah. "The Application of Ontology in a Clinical Decision." Semantic Technology and Information Retrieval (STAIR), 2011 International Conference on. Putrajaya, 28-29 June 2011.

[16] Eider Sanchez, Carlos Toro, Eduardo Carrasco, Gloria Bueno, Patricia Bonachela, Carlos Parra, Frank Guijarro. "A Knowledge-based Clinical Decision Support System for the diagnosis of Alzheimer Disease." e-Health Networking Applications and Services (Healthcom), 2011 13th IEEE International Conference on. Columbia, MO, 13-15 June 2011. 351 - 357.

[17] Mehnaz Adnan, Jim Warren, Martin Orr. "Ontology Based Semantic Recommendations for Discharge Summary Medication Information for Patients." Computer-Based Medical Systems (CBMS), 2010 IEEE 23rd International Symposium on. Perth, WA, 12-15 Oct. 2010. 456 - 461.

[18] Matt-Mouley Bouamrane, Alan Rector,Martin Hurrell. "Development of an Ontology for a Preoperative Risk Assessment." Computer-Based Medical Systems, 2009. CBMS 2009. 22nd IEEE International Symposium on. Albuquerque, NM, 2-5 Aug. 2009.

[19] Abd-Elrahman Elsayed, Samhaa R. El-Beltagy, Mahmoud Rafea, Osman Hegazy. "Applying data mining for ontology building." The 42nd Annual Conference On Statistics, Computer Science, and Operations Research. Cairo, 2007.

[20] Seongwook Youn, Dennis McLeod. "Efficient Spam Email Filtering using Adaptive Ontology." Information Technology, 2007. ITNG '07. Fourth International Conference on. Las Vegas, 2-4 April 2007. 249 – 254

[21] S Nefti, U Manzoor, S Manzoor "Cognitive agent based intelligent warning system to monitor patients suffering from dementia using ambient assisted living", International Conference on Information Society (i-Society),Pages 92-97, 2010.

[22] F Rea, S Nefti-Meziani, U Manzoor, S Davis "Ontology enhancing process for a situated and curiosity-driven robot", Robotics and Autonomous Systems 62 (12), 1837-1847

[23] Amit Bhagat, Sanjay Sharma, K.R.Pardasani. "Ontological Frequent Patterns Mining by potential use of Neural Network." International Journal of Computer Applications, Vol 36, Issue 10, 2011.

[24] MA Balubaid, U Manzoor, B Zafar, A Qureshi, N Ghani "Ontology Based SMS Controller for Smart Phones", International Journal of Advanced Computer Science and Applications 6 (1), 133-139.