Rule based Medical Content Classification for Secure Remote Health Monitoring

J. Balachander
Department of Information Technology
Thiagarajar college of Engineering
Madurai, India

E. Ramanujam
Department of Information Technology
Thiagarajar college of Engineering
Madurai, India

ABSTRACT
Now days, Medical data records are computerized and it is essential to predict the disease / symptoms of a patient in near future. The medical data records are referred during multiple disease prediction and to implement patient health monitoring system. The clinical data is essential to create clinically applicable information in healthcare system for treatment analysis. Ontology based implementation helps to classify clinical data and provide better results which is helpful to identify relevant symptoms and causes for a disease identification in patient health monitoring from the past clinical records. The proposed system implements rule based ontology classification method such as pattern matching for patient's diseases and compare with other methods such as Resource Description Framework implementation and SPARQL query language method. The results from this analysis can be helpful to developers and hospitals that can use the results to reduce the incorrect medical decisions facilitated by these systems.

Keywords
Ontology, Telemedicine, Quality of data, Clinical Decision Support Systems, Clinical Practice Guidelines.

1. INTRODUCTION
Medical healthcare system uses advanced computing techniques and its capability to produce the quality and effectiveness in patient health care monitoring system. The Medical healthcare system maintains patient health care information of different hospitals in different forms and stored in a database. These databases are used in a patient monitoring system for the prediction of diseases. The implementations are effective to establish a National Healthcare Information Infrastructure (NHII) system that must provide the health information to the patient in the absence of medical experts. In any medical emergency situation, there is a requirement of authoritative emergency information to the patient. The major problem in providing secure authoritative information to remote and rural area is the difficult task.

Data Engineering helps in providing remote and secure medical healthcare system using the concept of telemedicine system[1]. Telemedicine system uses Information and Communication Technology (ICT) with statistical, expert techniques, to obtain clinical data from remotely monitored ambulatory patients and it employs clinical Decision Support Systems (DSS). Telemedicine has been helpful in rural areas, where the shortage of healthcare providers had lead to lack of access ability to both basic healthcare and specialty maintenance. It too serves to eradicate distance obstacles and can improve access to medical services that would sometimes not be reliably available in remote rural communities. Through telemedicine system, the patients can get access to health care in faster time.

In advanced technology, a new way of accessing and processing clinical data are emerging for effective patient care. Pervasive healthcare system uses ICT tools to obtain clinical data from monitored ambulatory patients and it often employs automated clinical Decision Support Systems to process the data at the point of care. This way the patient is more closely preserved, in the absence of medical experts and human resources. In the modern Clinical Decision Support Systems the data are heterogeneous in nature so, the ontology based is used to collect the relation among technological context, quality of clinical data, and outpatient treatment. The ontology gives a formal way to represent the knowledge to specify the effect of technological context changes in the clinical data quality and the impact of the clinical data quality on an outpatient's treatment.

To handle huge amount of heterogeneous medical data from various database, to make decisions for emergency situation this paper proposes rule based ontology classification. The Rule based ontology classification predicts the disease analysis from the past history of clinical records for different symptoms at a time. To provide efficiency of the proposed, the results of the proposed technique are compared with ontology classification techniques such as SPARQL [2] query language and Resource Description Framework method (RDF) for predicting disease in medical healthcare systems.

2. PROBLEM STATEMENT
Medical data analysis shares disease history across the world in various medical applications and uses different computation, control and classification techniques to predict the diseases in future. For example, if a doctor initiates a clinical test for a patient, the report of the test is stored as quantitative data in a medical data repository as mentioned in [3] and used to predict the disease in near future. In this scenario, the data conversion from raw medical record to computerized data and storage of data is a critical one. The human errors, typo errors, machine errors play a major impact in data conversion. Also, the data storage requires heterogeneous format to store data in the database. In the last decade, more research has carried out the conversion, storage and data analysis technique and achieves effective results. However, there need more concentration in the storage and analysis for effective results. This paper proposes a data analysis of patient's medical record to predict the future disease using ontology based storing and classification methods.

In the literature, medical data analysis is implemented for Remote patient monitoring system and real time interaction services. These system records medical data but the prediction is critical one. Remote monitoring[4,5] also known as self regulatory or self testing which enables medical professionals to monitor a patient remotely using various technological devices such as sensors, mobile devices and other remote configurations. This method is primarily used for managing chronic diseases such as heart disease [6], diabetes mellitus or asthma and the like. These services can provide comparable health outcomes than the traditional in person patient meetings. Remote monitoring system [7] supplies greater satisfaction to outpatients and it is cost effective in purchasing the medical devices. There is some risk in the tests conducted by the patient may be inaccurate. Examples include home based nocturnal dialysis and improved joint management. Interactive service can provide immediate advice and health information to patients which also require medical attention. It provides real- time interaction between patient and internal doctor .There is several mediums utilized for this purpose which are phone, online and home visits. Many activities such as history review, physical examinations about presenting symptoms can be under taken, followed by assessment simulator to those done in traditional far-to-far appointment.

To overcome the issues of the state-of-the-art techniques and to predict the diseases from the historical patient records, the rule based ontology model is proposed. In a real time patient monitoring system, the proposed is helpful to identify the symptoms and predict the various diseases relevant to the symptom. This increases the availability of the medical information to the patients in rural areas in the absence of experts and reduces the incorrect medical decision facilitated by the other systems.

3. MEDICAL ONTOLOGY ANALYSIS

Ontology in medicine is mainly dedicated for representation, reorganization and storage of heterogeneous medical terminologies. Medical ontology is a model of the knowledge from clinical domain. It contains relevant and related information to diagnostics, treatment, clinical procedures and patient data. Ontology is designed in way that allows knowledge interpretation and reasoning. Medical ontology [1, 8] is different from terminologies which are static structures made for knowledge reference. The terminologies used for ontology are optimized for human processing is characterized by a significant amount of perfect knowledge. On the other hand, medical information systems need to communicate complex and detailed medical concepts. This paper proposes a rule based ontology classification method that connects the medical technological context of the patient's clinical records. This helps to predict future disease from patient's history of records. The medical ontology is the foundation for better data prediction and reduces more complex steps in technological data prediction.

This paper proposes a layering technique that refines a requirement elicitation method to successfully handle problems caused by technological interruptions [9]. This layering technique defines the functional relations between medical and technological dealings by using the clinical variable as an intermediary element. The technique makes certain non-functional relations between the two layers, characterized by the quality of technological context. When the quality of the technological context varies disagreeably the clinical variables like quality may degrade and affect potentially the treatment. The ontology resulting from the layering technique must define medical and technological dealings separately. The technological specific information is adapted to clinical variables and their quality of data and their clinical effects are addressed separately. Extent of this problem has been analyzed by presenting a two prolonged approach using simulation, followed by regression. In order to quantify the relative impact of weak data quality and overall computer interpretable clinical guidelines accuracy [10].

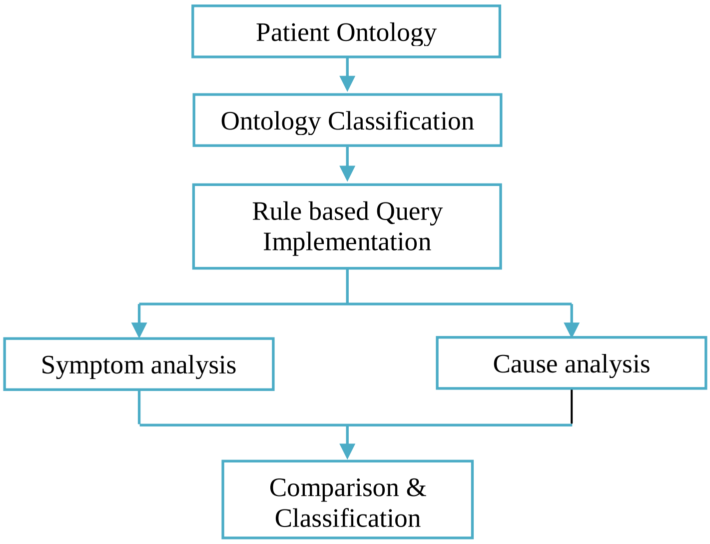

Fig.1 System Architecture

This system analyzes the effects of quality of the clinical data dishonor on outpatient treatment with medical consultants and execute these effects in the clinical decision making process during design time. Therefore the designed telemedicine system is technological connection, quality aware, preserves outpatient's safety and treatment efficiency. In this case each dimension represents the extent to which the output data's quality dimension is influenced by technological resource's of service quality dimension. For example, a technological resource's quality of service accuracy specifies the degree of correctness of resource processes data, preventing additional errors to the output data. Technique also specifies technological resources quality of service in terms of sub qualifiers as explained for quality of data and underlying Resource Qualifying Parameters. These sub qualifiers and Resource qualifying parameters are either static derived from technological resource manufacturer specified properties or dynamic derived from monitoring the technological resource's changing values properties. Resource qualifying parameters are the basic elements to compute the five quality dimensions values associated to the output data that is technological variable or at the point of decision clinical variable of a technological resource.

3.1 Ontology Classification

Ontology is classified using the Web Ontology Language and depicts the major role in Semantic Web. Ontology classification is the computation of the assumptions and sub assumptions with its hierarchies for classes and properties. Classes and properties are the reasoning services provided by all OWL reasoner's to know all details. To construct a class hierarchy, the algorithm starts with the empty hierarchy and then iteratively inserts each class from the ontology into the hierarchy. Each insertion step typically requires one or more sub assumption tests—checks whether a sub assumption relationship holds between two classes—in order to determine the proper position of a class in the hierarchy constructed so far. Significant attention has been devoted to the optimization of individual sub assumptions tests. The ontology has been executed by means of the Protege technology [11]. Class hierarchy and ontology reliability have been checked to the racer reasoned. Results demonstrates that the expected ontology based telemedicine system architecture facilitates the design of difficult medical assistance processes and the management of telemedicine messages interchange and thus should contribute to the improvement of the quality of medical services. This approach will make possible the management of the medical assistance scenarios which should be fully personalized to each geographical site. Certainly the needs of emergency services may differ consider between different countries and between urban and rural areas. Further work is however needed for rules composition optimization and to setup practicable.

The ontology specifies the relations between patient data at technological level and concepts of patient data at clinical decision making level, with clinical variables as intermediary elements. This ontology plays a significant role in the design of a technological connection and excellence aware telemedicine systems that could adjust to contextual changes to succeed the medical requirements. It plays also a role in activation medical practitioners to purify the treatment protocol.

Presents a semi automated system to evaluate the quality of medical measurements taken by patients [5]. The system realizes on data quality to evaluate various attributes aspects of measurements. The overall quality of measurement is to determine on the basis of these attributes enhanced with a troubleshooting mechanism. Particularly the troubleshooting mechanism guides healthcare experts in the investigation of the root causes of low quality values. This paper gives a solution for reviewing data reliability in home healthcare services. This approach employs a number of data qualifiers [8] to capture various quality aspects of a measurement. The overall quality of measurements is assessed using a semi automatic mechanism based on troubleshooting. This mechanism relies on workflows to determine the root causes of low quality indications for data qualifiers. Thinking over a feedback mechanism is proposed to help patients improve their future measurements.

3.2 Rule Based Query Language Method

Rule based query language method classify medical ontology data for multiple diseases. This is the classification method that classifies relevant symptoms, causes from medical ontology file and identifies the specific relationships between terms of the same or different ontologies. The major role in the rule based query language method is the sub-ontology extraction. Sub-ontology refers to the concept that identifies inferred correspondences between symptoms and causes from medical historical data. This results relevant symptoms for multiple diseases and future disease predictions.

Let us consider the Ontology as medical description logic (ML) information source $Q = < I, U >$ which is compared terms X and multiple terms X_n. Terms are defined based on medical historical records data, then,

$$ML(Q) = X(x_1, x_2, \ldots \ldots . x_n) \quad (1)$$

Where x_1 to x_n are different symptoms in medical data that relates to the causes data. Therefore

$$ML\ (Q) = I < X(x_1, x_2, \ldots \ldots . x_n) > \quad (2)$$

The relevance of Ontology is defined by an inference that provides the mapping for the following

- Disease individuals
- Disease relevance
- Symptoms individual
- Symptoms relevance

A model of ontology is such occurrences, under which all symptoms and its relevance of the ontology are satisfied. This ontology is called rule based consistent ontology and other ontology is called rule based inconsistent ontology.

3.3 SPARQL Query Language Method

SPARQL query language is standardized in 2008 by the World Wide Web Consortium (W3C) and supported by RDF triples. SPARQL query evaluation is based on sub group matching [2,9]. This is also called simple entailment since it can equally be defined in terms of the simple entailment relationships between OWL files. The range of queries can be implemented significantly in the form of standard conjunctive queries, supported in the OWL files.

This is the query language method search the Ontology Owl file for a particular data. The data may be a class or property or attribute. SPARQL is “data-oriented” method in that it only queries the information held in the models; there is no inference in the query language itself. It returns the information from the patient Owl File and collects the relevant symptoms that are mostly occurring in file.

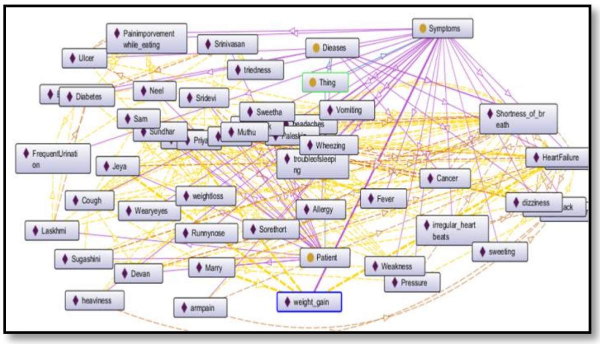

Fig.2 Ontology graph

Figure 2 describes Ontology graph for a patient’s historical records. Using this ontology, this system analyze the classes for SPARQL method, that are

- Diseases
- Symptoms
- Causes

SPARQL query language retrieves data from these classes and its attributes. Anonymous individuals in the template ontology are treated as variables whose bindings do not appear in the query’s result sequence. This is in contrast to conjunctive queries where they are treated as existential variables. Anonymous individuals in the queried ontology are treated as constants, i.e., they can be returned in a query answer. For brevity, the proposed assume here that neither the template ontology nor the queried ontology contains anonymous individuals.

3.4 Resource Description Framework (RDF) Method

The RDF is a standard for describing resources which identifies the subject, resources and object in the OWL file and also identifies its relationships to the individual patient records. All RDF standards consider properties and statements. Each one is related to the other resources. Based on the resources, patient medical diseases data set is classified and individual patient’s records are also analyzed.

This method works on the basis of atomic changes which are determined by additions or deletions of certain groups of statements to or from an RDF knowledge base in the ontology

OWL file. Such atomic changes are aggregated to more complex changes, resulting in a hierarchy of changes, thus facilitating the human reviewing process on various levels of detail (symptoms, causes). The derived compound changes may be annotated with meta-information such as the user executing the change or the time when the change occurred. The proposed technique present a simple OWL ontology capturing such information, thus enabling the distribution of change sets. Assuming that there will be no control of evolution, it must be clarified which changes are compatible with a concurrent branch of the same root ontology. Also, the proposed present a compatibility concept for applying a change to ontology on the level of statements.

Definition 1: A component is single if it may not be split into two nonempty components whose statements are disjoint.
Definition 2: A single Component is said to be a positive single component in a class C if its attributes are occurring in more times.
Using these definitions RDF methods describes the statement is a triple (S, P, O) where S is subject, P is predicate and O is a literal. These three values are mapped to diseases, causes and symptoms. Therefore

$$t(s, p, o) = c(d, cu, s) \quad (3)$$

Class hierarchy and ontology reliability have been checked and implemented results demonstrate that the expected ontology based healthcare system architecture facilitates the design of difficult medical assistance processes and the management of telemedicine messages interchange and thus should contribute to the improvement of the quality of medical services. This approach will make possible the management of the medical assistance scenarios which should be fully personalized.

4. ALGORITHM IMPLEMENTATION

This system implements an algorithm that computes the effective symptoms predictions from RDF method output. RDF classifies the patient's health care records OWL files and the algorithm implements multiple relevant diseases for the particular class of symptoms and causes.

```
Algorithm : disease predictions
Input   : patient history OWL file
Output : multiple relevant diseases and causes
Algorithm
        Read OWL File
        Create Ontology Model For that OWL file
        Read Data in Ontology Model
                For each RDF Data
                        Get Properties for Symptoms and Causes
                        For each class
                                Search value for properties
                                If found the
                                        Search the value
                                  For matched property
                        Next Class
                        For each Individual
                                Search value for Properties
                                If found then
                                        Search the value
                                  For matched property
                        Next Individual
For each search value
                        Identify relevant data
Next value
                Next Data
```

Developed a computer interpretable clinical guide lines (CIGs) for use during clinical actions. Clinical guide lines use task network model for demonstration but differ in their approaches to addressing particular modeling challenges. The paper studied the resemblances and differences between clinical guide lines in order to identify issues that must be solved before a consensus on a set of common components can be developed [4]. The components enable modelers to encode guidelines as plans that manage decision and action tasks in networks. They also enable the encoded guidelines to be linked with patient data a key requirement for permissive patient specific decision support. The proposed found consensus on many components, including plan society expression language conceptual medical record model medical concept model, and data abstractions. Variances were most apparent in underlying decision models, goal representation, use of scenario and structured medical actions.

5. RESULTS AND DISCUSSION

The proposed model is implemented on a PC with a Pentium Dual-Core processor running Microsoft Windows 7. All the algorithms are implemented in Java. OWL file is classified using Protégé Tool [11]. Using the classified patient's medical history OWL file, Rule Based Query Method, SPARQL and RDF methods are implemented. Various symptoms are feed and resultant data is recorded. Figure 3 shows the basic GUI which gets the input of symptoms for the RDF and SPARQL model. The GUI displays the diseases with the relevant symptoms for the prediction of disease in near future. The chart displays the resultant for single symptom with the comparison of other symptoms. The basic procedure of RDF and SPARQL can be processed only for a single symptom. Figure 4, 5, 6 and 7 shows the basic GUI for Rule based ontology model which gets the symptoms as input as same as the RDF and SPARQL model. The GUI shows the relevant symptoms with the diseases which affect the number of male and female patients. Also the analytical graph shows the comparison of one symptom with the other symptoms in terms of given input.

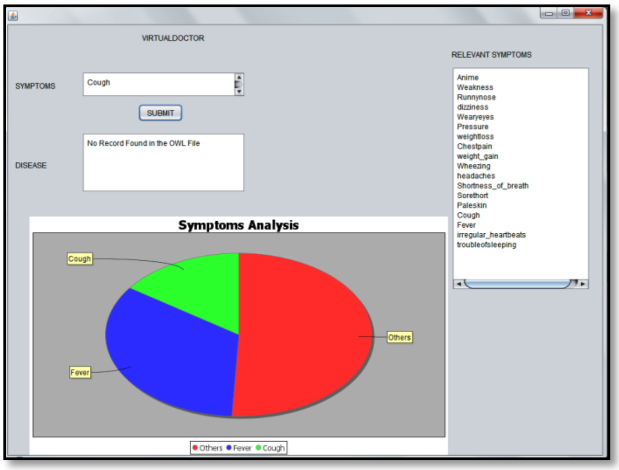

Fig. 4. GUI shows RDF and SPARQL implementation.

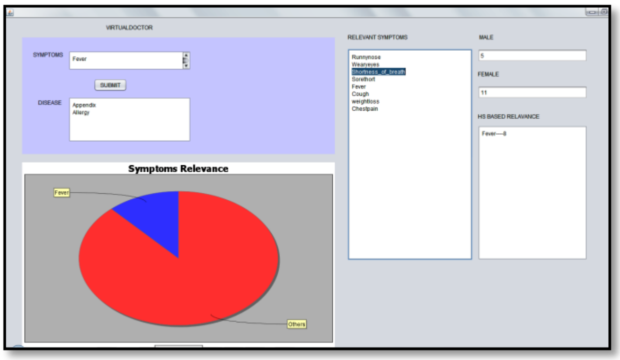

Fig. 5. GUI shows the input of single symptom with its relevant symptoms and diseases for the Rule based QL.

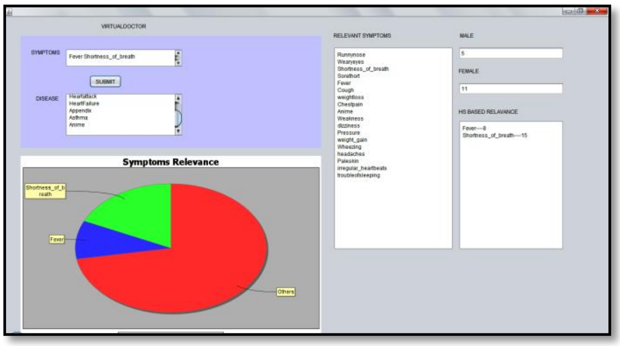

Fig. 6. GUI shows the input of two symptom with its relevant symptoms and diseases for the Rule based QL.

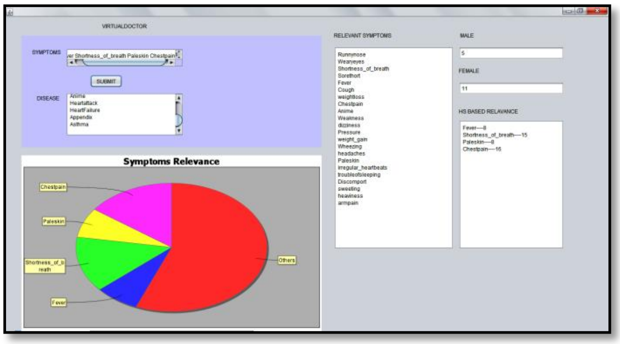

Fig. 7. GUI shows the input of Multiple symptom with its relevant symptoms and diseases for the Rule based QL.

Figure 5 shows the resultant analytical graph for single symptoms and Figure 6 shows the resultant for two symptoms and Figure 7 shows the resultant for multiple symptoms. Table 1 shows the Rule based Query Language comparison with the SPARQL and RDF based method in terms of number of individuals and the relevant symptoms it produces. Rule based query language method implements more reference symptoms based on various disease values as compared with other methods like SPARQL query language method. These methods relate patient responsive information that is then made clear by medical professionals to accomplish their diseases. The implementation of such services and hardly relies on the patient hope in a healthcare service provider in terms of privacy of the data chain and physicians trust in the reliability of information and data uncompensated by patients. Figure 9 shows the information about the number of symptoms given as input to the Rule based QL and the number of diseases predicted by the proposed technique.

TABLE 1. Relevant Symptoms Statistics

NO	Source File Length (No of Individuals)	Relevant Symptoms count		
		Rule Based Query method	SPARQL method	RDF Method
1.	10	5	3	4
2.	20	11	8	11
3.	30	15	14	17
4.	40	23	17	22
5.	50	32	26	31

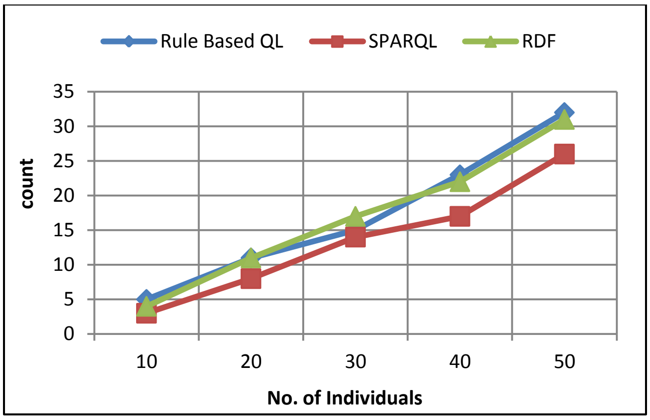

Fig 8. Comparison of relevant symptom statistics of the Rule based QL with RDF and SPARQL.

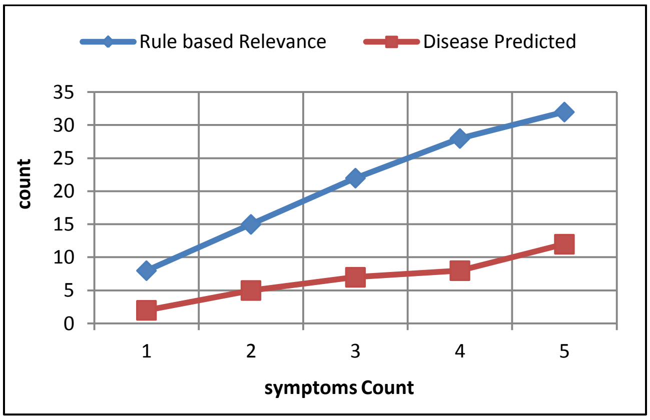

Fig 9. Prediction of diseases with the number of symptoms given as input to the Rule based QL.

6. CONCLUSION

The proposed work implements the concept of rule based ontology classification to predict relevant clinical records for multiple diseases from the history of patient's clinical records. This work implements sub-ontology classification for relevant symptoms and identification of multiple disease relevance values for predicting diseases. The proposed contribute the concept of rule based QL to overcome the issue of RDF and SPARQL in terms of multiple symptom and disease analysis. The efficiency of the proposed is compared with other methods such as SPARQL query language method and Resource Description Framework Method (RDF). Experimental results achieve optimal solution for future prediction in remote healthcare system. In future, this work will be extended to different formats of clinical records such as medical laboratory record implements ontology based classification to predict relevant clinical records.

7. REFERENCES

[1] Larburu, N., Bults, R.G., Van Sinderen, M.J. and Hermens, H.J., 2015. An ontology for telemedicine systems resiliency to technological context variations in pervasive healthcare. *IEEE journal of translational engineering in health and medicine*, *3*, pp.1-10.

[2] Larburu, N., Widya, I., Bults, R. G. A., and Hermens, H. J., 2014. Making medical treatments resilient to technological disruptions in telemedicine systems, *Proc. IEEE-EMBS Int. Conf. Biomed. Health Informat. (BHI)*, pp. 285-288.

[3] Padmavathi, S. and Ramanujam, E., 2015. Naïve Bayes Classifier for ECG Abnormalities Using Multivariate Maximal Time Series Motif. *Procedia Computer Science*, *47*, pp.222-228.

[4] Arts, D.G., De Keizer, N.F. and Scheffer, G.J., 2002. Defining and improving data quality in medical registries: a literature review, case study, and generic framework. *Journal of the American Medical Informatics Association*, vol. 9 No. 6, pp.600-611.

[5] Vavilis, S., Zannone, N. and Petković, M., 2013, Data reliability in home healthcare services. *In Computer-Based Medical Systems (CBMS), 2013 IEEE 26th International Symposium on*, pp. 377-380. IEEE.

[6] Ramanujam, E. and Padmavathi, S., 2012, December. Constraint frequent motif detection in sequence datasets. In *Advanced Computing (ICoAC), 2012 Fourth International Conference on* (pp. 1-5). IEEE.

[7] Paganelli, F., and Giuli, D., 2011. An ontology-based system for context-aware and configurable services to support home-based continuous care, *IEEE Trans. Inf. Technol. Biomed.*, Vol. 15, No. 2, pp. 324-333.

[8] Larburu, N., Bults, R.G. and Hermens, H.J., 2014. Quality of data computational models and telemedicine treatment effects., *2014 IEEE 16th International Conference on e-Health Networking, Applications and Services (Healthcom)*, pp. 364-369. IEEE.

[9] Weiskopf, N.G. and Weng, C., 2013. Methods and dimensions of electronic health record data quality assessment: enabling reuse for clinical research. *Journal of the American Medical Informatics Association*, vol. 20, No. 1, pp.144-151.

[10] Hasan, S., and Padman, R., 2006. Analyzing the effect of data quality on the accuracy of clinical decision support systems: A computer simulation approach, in *Proc. AMIA Annu. Symp.*, pp. 324-328.

[11] Rubin, D.L., Noy, N.F. and Musen, M.A., 2007. Protege: a tool for managing and using terminology in radiology applications. *Journal of digital imaging*, Vol. 20 No. 1, pp.34-46.