Middle-East Journal of Scientific Research 24 (8): 2567-2576, 2016
ISSN 1990-9233

DOI: 10.5829/idosi.mejsr.2016.24.08.23829

Investigation on the Performance of Classifiers in Prediction of Erythemato-Squamous Disease: An Automated Ontology Learning (AOL) Methodology

S Sivasankari and Shomona Gracia Jacob

Anna University, SSN College of Engineering,
Department of Computer Science Engineering, Kalavakam, Chennai, India

Abstract: Erythemato-squamous disease (ESD) is a persistent skin disease that affects the lives of individuals all over the world. Diagnosis of this ailment remains a challenge since most skin disorders have similar characteristics and this ailment predominantly affects people in tropical regions who are exposed to unhygienic conditions of living. The authors in this research attempt to computationally investigate the use of data mining techniques in deriving meaningful patterns of ESD through Ontology construction and learning for knowledge reuse and inference. Since ontology construction is a time-intensive task, automating the process of ontology construction for ESD was the rationale for this research. This work proposes the formulation of an automated ontology structure through classification of available medical data related to ESD. This will enable disease diagnosis using the developed knowledge base and rule system. Classification of data is carried out by the decision tree algorithm. SWRL (Semantic Web Rule Language) and JESS (Java Expert System Shell) were combined to analyze the outcome of various classifiers (J48, Random Forest, LD Tree, Rep Tree and Random Tree). The best knowledge structure with respect to classification accuracy was the basis for automatically constructing the ontology through OWL. SWRL was employed for framing the rule base while the inference was made by JESS. Among the various classifiers investigated, J48 (Decision tree) yielded the best performance with an accuracy of 95% with the optimal set of discriminative features. This led to the utilization of rules and attributes yielded by the Decision Tree algorithm for automated ontology construction and rule formation.

Key words: Ontology · Data mining · Classification · Dermatology · Reasoning · SWRL · JESS

INTRODUCTION

Dermatology is a field of medicine dealing with hair, nail and skin related diseases. Skin diseases is a complex problem which is difficult to diagnose, ultimately leading to severe conditions like skin cancer. ESD is one such complicated problem in dermatology [1]. Most of the skin diseases share the clinical features of erythema and scaling, with very minimal differences. There are six group of diseases: Psoriasis, seboreic dermatitis, lichen planus,pityriasis rosea, cronic dermatitis and pityriasis rubra pilaris.

Erythemato-squamous diseases have certain morphological features in common, which may at times cause diagnostic difficulties. Hence diagnosis totally depends on visual task done by expert doctors based on their expertise. Thereby, Ontology construction for ESD diagnosis can help serve this situation.

Ontology is a combination of artificial intelligence and machine language [2-11] to help share and reuse knowledge which contains natural language processing and knowledge representation, etc. Ontology can be used as channels of communication between the human being and system. Ontology is, therefore, defined as a formal explicit specification of a conceptualization of a domain and its relationships [12]. Ontologies finds its application in almost all fields some of which include knowledge engineering, artificial intelligence, bioinformatics, medicine, e-commerce, chemistry, education and knowledge management, [13, 14] etc. Though ontology

Corresponding Author: S. Sivasankari, Anna University, SSN college of Engineering,
Department of Computer Science Engineering, Kalavakam, Chennai, India.

has many-fold applications [15] construction of the ontology structure still remains a complex task. Human intervention to manually construct an ontology makes the entire process time, labor and resource intensive. In view of this, it is believed that automatic ontology construction of medical data sets would be a great enhancement in this sphere of research. Hence it would be highly useful if the process of Ontology building could be automated and reasoners infer the constructed ontologies. In this paper, SWRL and JESS were combined to analyze the symptoms of the ESD and on the basis of this, the most appropriate class is recommended. In this proposed method, automatic ontology construction is attempted through data mining techniques [16-18] and subsequently diagnosing ESD using knowledge base and rule base system. The contribution is this work is depicted from two perspectives: (i) The ontology to represent the terms associated with the diagnosis of ESD is automatically constructed through classification techniques. (ii) The constructed ontology is utilized to diagnose the presence/absence of ESD in the unsupervised test data through SWRL. This paper is organized as follows: Section 2 presents the survey of related work in this area of research. Section 3 details on the proposed methodology for automatic ontology constructions using classification techniques and the computational diagnosis on an unknown dataset. Section 4 details the results and discussion. Section 5 concludes the paper and also gives direction on the possible extensions to this work.

The prime motivation for this work was supported by the evidence of involvement of ontologies in the biomedical field. Under various conditions, Web rule language has worked out to be a very beneficial tool in medical ontologies. The purpose of this, is to study various classes of units which includes substances, qualities and processes which are relevant to the biomedical domain [19].Numerous studies have shown the possible usage of ontology based information in health care. For instance, MajaHazic *et al*., [20] has proposed the development of Ontology based grid middleware for research of human diseases which helps in resolving concerns related to medical conditions and disease factors. Further, generic human disease ontology (GHDO) was designed by Maja Hadzic to represent knowledge of four main attributes of human disease, namely, types, symptoms, causes and treatments [21]. Notwithstanding this, the framework was expected to bolster the study of complex disorders which resulted due to multiple factors working at the same time. In addition, Akifumi Tokosumi *et al*., [22] put forth the future attitude towards the medical repository system after evaluating the present medical ontologies in the direction of three knowledge repositories, for instance, the confined nature, collective acquisition and usable knowledge. VnHIES, is another ontology based extraction system developed by Tran Quoc Dung *et al*., [23] for health care. This system was more accurate and new semantic elements and algorithm was used for extraction of data in semantic word while a document weighting algorithm was applied to get summary of health related information. Besides these systems, Tharam S. Dillon *et al*., [24] worked on storage and processing of information in the biomedical domain [25]. Beside these, disease specific ontology based rule system was developed by Antonio J. Jara *et al*., [26] for the detection and prediction of myocardial diseases. The system intended to predict the disease by Chronobiology algorithm, where ontology trees were developed to get the knowledge base of the disease.This concept was used by the scholarly semantic web for automatic generation of ontology. To facilitate the ontology representation, following the semantic pattern based approach, TextOntEx, construction was used to extract candidate's relations and mapping that into a meaningful representation by analyzing natural domain text [27]. All the above literature shows absence of disease specific domain ontology except for myocardial diseases.

Proposed Automated Ontology Learning Framework: The most important step in construction of ontology automatically. In this paper, automatic ontology construction through classification techniques is proposed. The proposed framework consists of domain identification, data pre-processing, Classification, building decision tree, performance evaluation, construction of ontology using SWRL and Inference engine mechanism. The proposed framework is depicted in Fig 1. The various modules in the proposed framework are presented in detail below.

Domain Identification: The dataset used for processing the proposed methodology is detailed here.

This Erythemato-squamous disease dataset is to predict whether the patients have

Psoriasis	- It is a most common chronic inflammatory skin disease, 2% to 5% of people affect around the world.
Seboreic dermatitis	- Usually occurring in younger age groups, disease often seen in the area of face, scalp, upper trunk.
Lichen planus	- A younger age group is predominantly affected.
Pityriasis rosea	- It is a common skin disorder usually occurring in children and young adult.
Chronic dermatitis	- People with chronic dermatitis often have asthma or seasonal allergies.
Pityriasis rubra Pilaris	- It is a rare chronic disorder may occur in any age group.

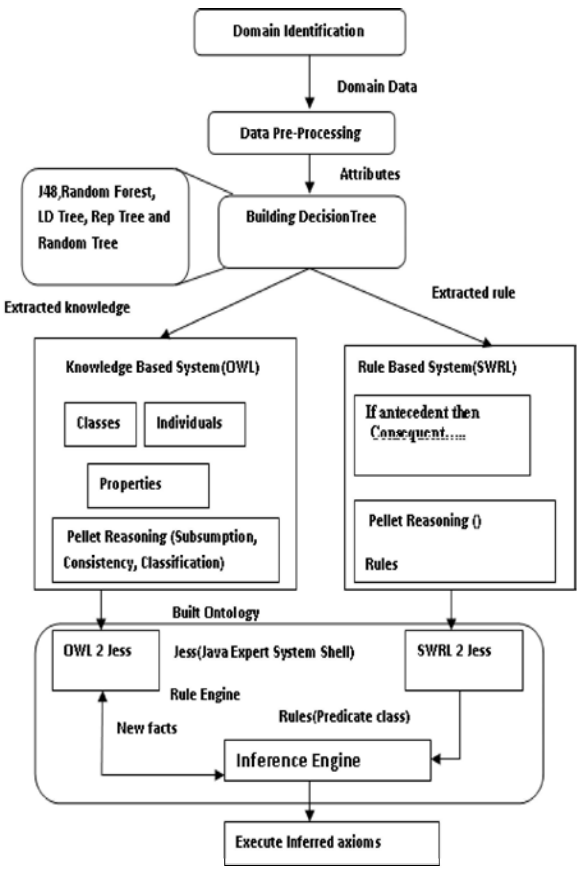

Fig. 1: Proposed Automatic ontology Learning Framework

Henceforth the patient has to undergo various diagnoses. The used data in this study was taken from UCI machine learning repository [28]. Dermatology dataset for the classification of erythemato-squamous diseases. The data contain 366 records and each record has 34 features which includes 12 of clinical and 22 of histopathological.Consequently this dataset is displayed in Table 1 and Table 2. The selected features were given degrees from ’0’ to ’3’ (’0’ representing the absence of a particular feature, ’3’ representing the biggest score possible and ’1’, ’2’ showing the relative intermediate values).

Table 1: The UCI ESDs dataset with class distribution

Class code	Class	Number of samples
1	Psoriasis	112
2	Seboreic dermatitis	61
3	Lichen planus	72
4	Pityriasis rosea	49
5	Cronic dermatitis	52
6	Pityriasis rubra pilaris	20
		366

Table 2: The UCI ESDs dataset with description of features

Clinical features	Histopathological features
(possible values: 0,1,2,3)	(possible values: 0,1,2,3)
Erythema	Melanin incontinence
Scaling	Eosinophils in the infiltrate
Definite borders	PNL infiltrate
Itching	Fibrosis of the papillary dermis
Koebner phenomenon	Exocytois
Polygonal papules	Acanthosis
Follicular papules	Hyperkeratosis
Oral mucosal involvement	Parakeratosis
Knee, elbow involvement	Clubbing of the rete ridges
Scalp involvement	Elongation of the rete ridges
Family history (0 sau 1)	Thinning of the suprapapillary epidermis
Age (Linear)	Spongiform pustule
	Munro microabcess
	Focal hypergranulosis
	Disappearance of the granular layer
	Vacuolisation and damage of basal layer
	Spongiosis
	Saw-tooth appearance of retes
	Follicular horn plug
	Perifollicular parakeratosis
	Inflammatory monoluclear infiltrate
	Band-like infiltrate

Data-Preprocessing: The erythemato-squamous domain data is given as input to the data pre-processing step. In order to make data processing interoperable between WEKA and OWL, the blank spaces in the attribute names are removed. This data set contains some missing values. The existing classifiers [29] themselves handled the missing values suitably. In the case of J48 classifier, any split on an attribute with missing value will be done with weights proportional to the frequencies of the observed non-missing values [30]. The pre-processed data are considered for further classification.

Classification Techniques and Building Decision Tree: Classification [31-37] is used to classify data into predefined class labels. Class in classification is the attribute or feature in a data set, in which users are most interested. Classification can be used to diagnose the ESD based on the clinical features and histopathological data. The Decision tree is one of the commonly used classification techniques, easy to handle huge data. As a data structure, Decision tree is used to represent the logical structure of classification rules for domain specific empirical data. In this present study, six different state-of-the-art supervised machine learning[38] algorithms, namely J48, AD Tree, BF Tree, LAD Tree, NB Tree, Rep Tree and Random Tree algorithm were analyzed. J48 implements C4.5 [39-40] decision tree learning algorithm. In this proposed method, J48 algorithm serves to be the best one with the optimal solution through the validation procedure viz,Cross Validation with 3 folds. The results obtained appeared to represent the optimal solution. Hence rules (in the form of Decision tree) generated through J48 classification algorithm was used for building the Ontology structure.

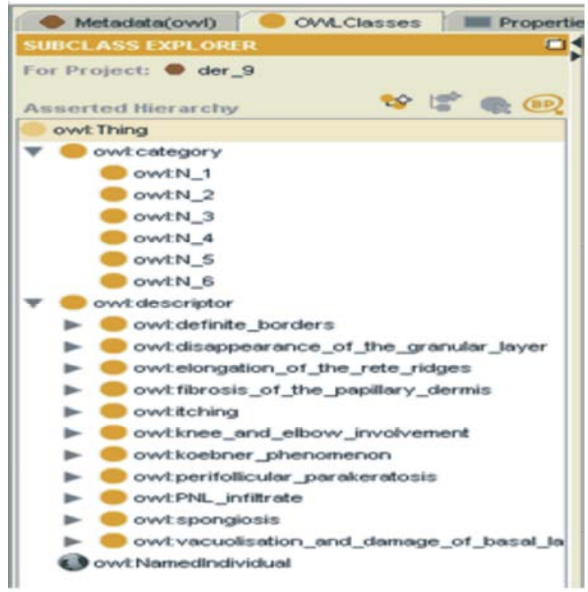

Fig. 2: Dermatology Class Hierarchy Structure

WEKA Decision Tree: A WEKA Decision tree was evolved and serialized into dot format using WEKA API to read the document and create the ontology using the OWLAPI. J48 decision tree algorithm was used to discover and extract knowledge from structured data. Then ontology was built from the generated decision tree [41]. The ontology structure consists of two systems, one is Knowledge based systems (OWL) and the other is Rule based system (SWRL).

Knowledge Based System (Web Ontology Language): Knowledge based System (OWL) can be considered as a special type of database "that holds information representing the expertise of a particular domain" [42] OWL was constructed using Java in Eclipse. The Java with OWLAPI was integrated. The implementation of this work was carried out using WEKA 3.6.10, an open source [43] data mining tool and Protégé 5, open source tool [44] for ontology framework creation.

Ontology Creation: The automatic ontology construction of hierarchical structure, consists of a set of classes organized in a structured manner to represent the domain's salient classes, a set of slots associated to classes to describe their properties and relationships and a set of instances of those classes. In OWL, classes are interpreted as sets of subclasses. The hierarchy structure of Erythemato-squamous is depicted in Fig 2.

Class and Class Hierarchy: The head node denotes a descriptor. Each and every node is assigned as super-class/subclass. The hierarchical structure is framed like superclass and subclass. The tail node represents a category. Each branch in the decision tree may have a set of leaves. Each leaf in the decision tree represents a classification rule as well as the target class (category). In dermatology data, the descriptor contains 11 attributes and then six categories are shown in Fig 2. Based on the two branches, Automatic ontology construction from the extracted knowledge is represented in the decision tree.

Object Properties and Data Properties: There are two main types of properties, Object properties and Datatype properties. In the designed ontology, both types of properties exist. The total number of properties is 22 consisting of 11 Object properties and 11 Datatype properties.Each and every descriptor represents its label and the relationship between attributes as object properties. The data properties depict "has" relationship between the attributes of data literal as shown in Fig. 3.

Axioms for Classes, Attributes and instances: Axioms is a fact or rule. Each node has the relationship between classes, attributes and individuals. The descriptors have some description of the classes and its attributes. The instances are represented as its members and the value of attributes.

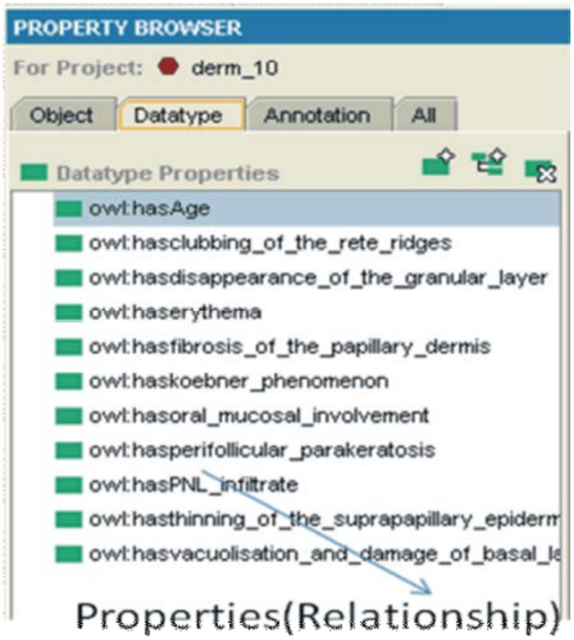

Fig. 3: Structure of Data Properties

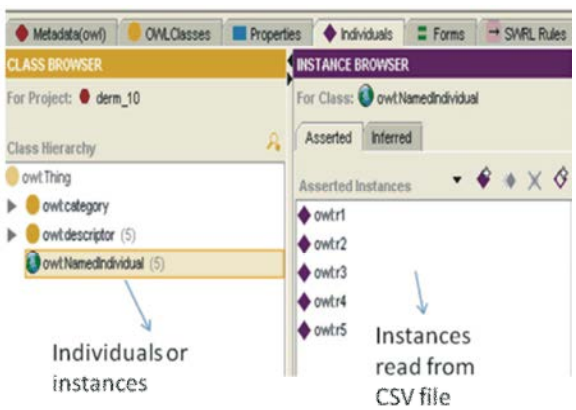

Fig. 4: Named Individual and its value

Individuals: Individual is an instances of a class. In Comma Separated Value (CSV) files, normally the test data excludes the class field. Every row is represented as r1, r2… so on. Hence the classes match with the attributes and then it would retrieve the value of the particular attribute and creates the relationship between the classes and its data type properties as shown in Fig 4.

The Rule based System: The rule based system consists of the reasoner and SWRL. To validate the SWRL rules, consistency of the relationship between classes and their properties is validated by reason. Each rule can be expressed by *If-Then* statements. The *if*-part is the antecedent (Left hand side). It consists of one or more of condition elements.The conditions represented may be categorized for integer/real intervals or combination of these for more complex problems. The *then*-part -which is called consequent (Right Hand Side) or action, consists of a number of actions.

SWRL (Automatically Constructed From Decision Tree): SWRL is a semantic Web Rule language which is grounded on expressive OWL-based rule languages. Since it is based on OWL web ontology language [45-46], it is the combination of OWL DL and OWL lite sublanguages with Unary/Binary language. For an impressive deduction of reasoning capabilities, SWRL allows users to write rules that are expressed in terms of OWL concepts that are far better than OWL alone. It is, however, centered on the same depiction rationale as OWL and provides the comparable strong formal guarantees when performing inference. In both, OWL DL and OWL lite sublanguages, SWRL [47-49] incorporates, an effective conceptual pattern for Horn-like rules. This model-theoretic semantics provides the formal intending to OWL ontologies incorporating standard written conceptual abstracts [50].

Protégé OWLAPI: Protégé is the one of the most used open source ontology editor that permits the handler to model the ontology. Protégé frames and protégé OWL are the two prime methods of ontology editing using Protégé platform. OWL, RDFs and XML schema are the most common formats for extraction of information from Protégé. In building knowledge based tool Protégé is further extended by the method of plug-in architecture and Java based Application Programming Interference (API). Protégé OWLAPI, is an open-source Java library for the OWL and RDFs. OWL is an ontology language and working with language objects API allows the tools to work at an appropriate level of abstraction and especially in case of inference it provides clear identification of functionality. API also takes care of methods and classes to load and save OWL files, that can be used to query and manipulate OWL data models. This further aids in performing reason based Description Logic engines but also optimizes implementation of user graphical interfaces.

Jess inference Engine: Jess is a rule engine which is based on Java language, which can develop Java software with a capacity to reason from the knowledge provided in the form of rules. It comprises of rule base, fact base and execution engine. SWRLJessTab [51] is an example of Jess to be used in Protégé based tools.

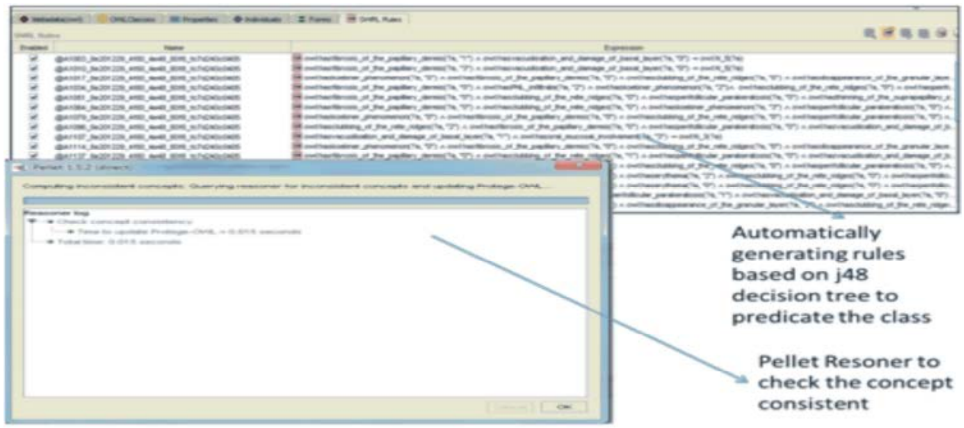

Fig. 5: Automatic rule generation and consistency check

OWL to Jess Facts: OWL to Jess is an archetypical tool for the transformation of OWL ontologies to Jess rule facts by XML- based syntax of OWL, thereby enabling extension of the OWL model by the means of rules. To implement application based reasoning services, predefined Jess rules are used to encode OWL while running Jess engine OWL [52] property information can be directly represented as Jess facts.

SWRL to Jess Rules: In order to allow interaction between an OWL knowledge base containing SWRL rule and a rule engine, SWRLJessTab plug-in is used in Protégé-OWL for the execution of SWRL rules in the Jess rule engine Fig 5. There are four important tasks that are needed to allow Jess to interoperate with SWRL: (1) representation of OWL individual knowledge as Jess facts; (2) representation of SWRL rules as Jess rules; (3) inference is performed on the basis of that rule and results are reflected in the OWL knowledge base; (4) interaction is controlled by graphical interface.

Inference Engine: In the context of the processing program, considered as an expert system, the infe-rence engine is a part which derives conclusions based on the facts and rules. It is logic-based and is considered as an interpreter of the system.

Executed Inference Axioms: During ontology creation Jess rule engine is executed to infer knowledge on the basis of SWRL rule and OWL knowledge. The knowledge inferred from Jess in the form of inferred axioms is then transferred into an OWL ontology.

RESULTS AND DISCUSSIONS

In order to estimate the accuracy of classifiers, five different algorithms, viz. J48, Random Forest, LD Tree, Rep Tree and Random Tree were analyzed. The results depicted higher accuracy when classified by the J48 classification algorithm viz., 95% with optimal number of 11 attributes Further rules were therefore generated through J48 classification and was used for building ontology structure Table 3. The output of Jess inference tab is presented in the Fig. 6-9.

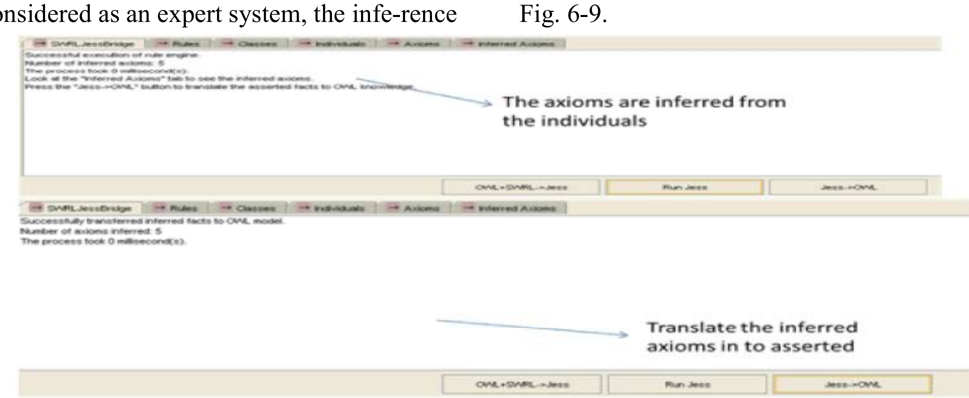

Fig. 6: SWRL output: Execution of rules in Jess engine

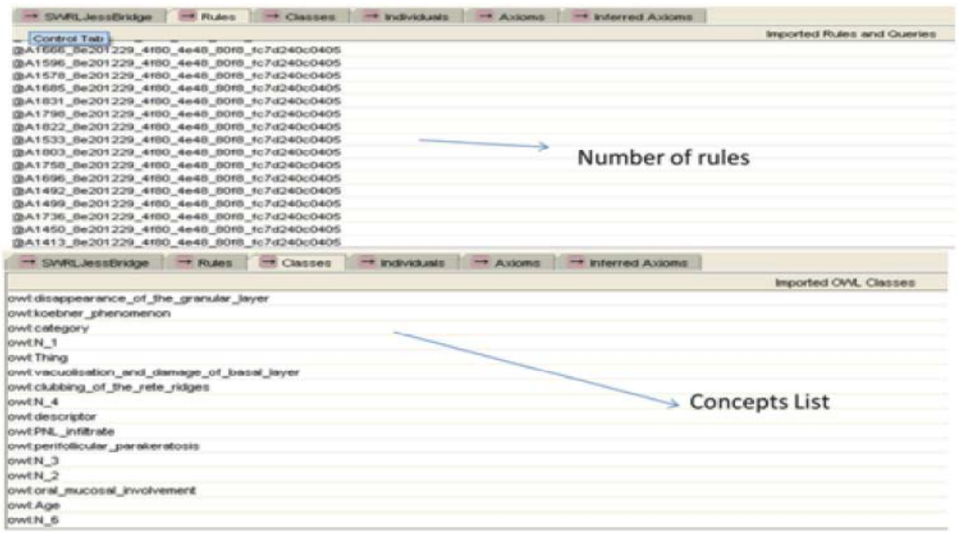

Fig. 7: SWRL output: Number of rules and concepts list

Fig. 8: SWRL output: Individuals and Axioms in Jess engine

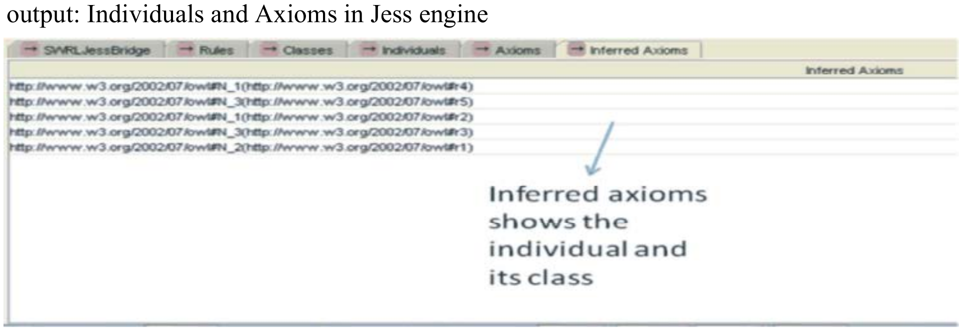

Fig. 9: SWRL output: Inferred axioms.

Table 3: Classification table

Classifiers	Accuracy
J48	95%
Random Forest	94.2%
LD Tree	94.8%
Rep Tree	92.8%
Random Tree	84.9%

Diagnosis of complicated diseases is a difficult task in medical research because of overlapping symptoms. ESD is one such complicated skin disease which is difficult to diagnose and can lead to severe conditions like skin cancer. In this paper, an attempt is made in the diagnosis of ESD by automated ontology construction and learning. This research work has utilized the classification technique for the generation of decision trees. Different classification algorithms were investigated to achieve the optimal decision tree. The J48 classifier yielded the best performance with an accuracy of 95% with optimal number of 11 attributes. The SWRL rules generated were used for the automatic construction of ontology and executed in the Jess engine for inference. The Jess results were obtained in the form of "the person with the symptoms" and stored in the ontology with important descriptors which can help in better diagnosis of disease.

CONCLUSION

This research paper has attempted to develop a system by the use of ontology as a knowledge acquisition procedure depicting the disease-symptom domain. Decision Tree with features evolved by J48 are considered for ontology construction and rule formation. Constructing ontology manually is a cumbersome task but with this research work has investigated the process of automatic ontology construction and learning. This is beleived to enhance the current state of medical diagnosis by giving a better and clearer representation of the disease-symptom domain. The proposed study is expected to aid physicians in diagnosis of ESD and hence advance the existing methods of therapy. The insight for future directions are viewed from the perspective of expanding the process of automated ontology learning to other clinical ailments. Semantic web mining of medical data and automated ontology construction from the semantic medical data is also to be explored with the proposed framework.

REFERENCES

1. Guvenir, H.A., G. Demiroz and N. Ilter, 1998. Learning differential diagnosis of erythemato-squamous diseases using voting feature intervals, Artif. Intell Med., 13(3): 147-165.
2. Guvenir, H.A. and N. Emeksiz, 2000. An expert system for the differential diagnosis of erythemato-squamous diseases, Expert Syst Appl., 18: 43-49.
3. Ubeyli, E.D. and I. Guler, 2005. Automatic detection of erythemato-squamous diseases using adaptive neuro-fuzzy inference systems, Comput Biol Med., 35(5): 421-433.
4. Luukka, P. and T. Leppalampi, 2006. Similarity classifier with generalized mean applied tomedical data, Comput Biol Med, 36(9): 1026-1040.
5. Polat, K. and S. Gunes, 2006. The effect to diagnostic accuracy of decision tree classifier of fuzzy and k-NN based weighted pre-processing methods to diagnosis of erythemato-squamous diseases, Digit Signal Process., 16(6): 922-930.
6. Nanni, L., 2006. An ensemble of classifiers for the diagnosis of erythemato-squamous diseases, Neurocomputing, 69(7-9): 842-845.
7. Luukka, P., 2007. Similarity classifier using similaritymeasure derived from Yu's norms in classification of medical data sets, Comput Biol Med., 37(8): 1133-1140.
8. Ubeyli, E.D., 2008. Multiclass support vector machines for diagnosis of erythemato-squamous diseases, Expert Syst Appl., 35(4): 1733-1740.
9. Polat, K. and S. Gunes, 2009. A novel hybrid intelligent method based on C4.5 decision tree classifier and one-against-all approach formulti-class classification problems,Expert Syst Appl, 36(2): 1587-1592.
10. Ubeyli, E.D., 2009. Combined neural networks for diagnosis of erythemato-squamous diseases, Expert Syst Appl., 36(3): 5107-5112.
11. Ubeyli, E.D., 2010. Dogdu E,Automatic detection of erythemato-squamous diseases using k-Means clustering, J.Med Syst., 34(2): 179-184.
12. Gruber, T., 2009. Ontology. Encyclopedia of Database Systems. Springer-Verlag
13. Bodenreider, O. and Stevens, 2006. R, Bio-ontologies: current trends and future directions, Brief Bioinform., 7(3): 256-74.
14. Noy, N. and D McGuinness, 2001. Ontology development 101: a guide to creating your first ontology, stanford knowledge systems laboratory technical report KSL-01- 05 and Stanford medical informatics technical report SMI-2001-0880.
15. Uschold, M., 1996. Grüninger M, Principles Ontologies: Methods and applications, Knowledge Eng Rev, 11(2): 93-155.
16. Jacob Shomona Gracia and R.GeethaRamani, 2015. Prediction of Rescue Mutants to Restore Functional Activity of Tumor Protein TP53 through Data Mining Technique, J. of Scientific and Industrial Research, 4(3): 135-140.

17. Shomona Gracia Jacob and Dr.R. Geetha Ramani, 2012. Efficient classifier for classification of Prognostic Breast Cancer data through Data mining techniques, Proc. of the World Cong. on Engin.and Comp. Sci., 1: 493-498.
18. Jacob Shomona Gracia, Dr. R. Geetha Ramani and P. Nancy, 2012. Classification of Splice Junction DNA sequence data through Data mining techniques, International Conference on Future Communication and Computing Technology, pp: 143-148.
19. Bodenreider, O. and A Burgun, 2005. Biomedical ontologies, In Medical Informatics, Springer US, pp: 211-236.
20. Hadzic, M., E. Chang, P. Wongthongtham and R. Meersman, 2004. Disease ontology based grid middleware for human disease research study, In Industrial Electronics Society, IEEE, 1: 480-486.
21. Hadzic, M. and E. Chang, 2005. Ontology-based Multi-agent Systems Support Human Disease Study and Control., Conference Self-Organization and Autonomic Informatics (I), Glasgow, UK, 135: 129-141.
22. Tokosumi, A., N. Matsumoto and H. Murai, 2007. Medical ontologies as a knowledge repository. In Complex Medical Engineering, IEEE, 487-490.
23. Dung, T.Q. and W. Kameyama, 2007. A proposal of ontology-based health care information extraction system: Venires, In Research, Innovation and Vision for the Future, IEEE International Conference, pp: 1-7.
24. Dillon, T.S., E. Chang and M. Hadzic, 2008. Ontology support for biomedical information resources. In Computer-Based Medical Systems, 21st IEEE International Symposium, pp: 7-16.
25. Thirugnanam, M., 2013. An Ontology Based System for Predicting Disease Using SWRL Rules, International J. Comp. Science and Business Informatics, 7(1).
26. Jara, A.J., F.J. Blaya, M. Zamora and A.F. Skarmeta, 2009. An ontology and rule based intelligent information system to detect and predict myocardial diseases, In Information Technology and Applications in Biomedicine, 9th International Conference IEEE, pp: 1-6.
27. Wuermli, O., A. Wrobel, S.C. Hui and J.M. Joller, 2003. Data Mining For Ontology Building: Semantic Web Overview, Diploma Thesis-Dep. of Computer science, Nanyang Technological University.
28. UCI Machine Learning Repository Irvine, CA, University of California, School of Information and Computer Science https:// archive.ics.uci.edu/ ml/ datasets/ Dermatology
29. Ian, H., Witten; Eibe Frank; Mark A. Hall, Data Mining: Practical Machine Learning Tools and Techniques (3 Ed.). Elsevier. ISBN 978-0-12-374856-0.
30. Ramani, R.G. and S.G. Jacob, 2013. Improved Classification of Lung Cancer Tumors Based on Structural and Physicochemical Properties of Proteins Using Data Mining Model, Plos.One, 8(3)e58772.doi:10.1371/journal.pone.0058772.
31. Shomona Gracia Jacob and Dr. R. Geetha Ramani, 2012. Data mining in Clinical Data Sets: A Review, Intern. J. of Appli. Inform. Systems, 4(6): 15-26.
32. Jacob Shomona Gracia and R. Ramani Geetha, 2013. Design and Implementation of a clinical data classifier: A Supervised learning approach, Res. J. Biotech, 8(2): 16-26.
33. Ramani R. Geetha and Shomona Gracia Jacob, 2013. Benchmarking Classification Models for Cancer Prediction from Gene Expression Data: A Novel Approach and New Findings,Studies in Informatics and Control., 22(2): 133-142.
34. Jacob, Mrs. Shomona Gracia and Dr. R. Geetha Ramani, 2011. Discovery of Knowledge Patterns in Clinical Data through Data Mining Algorithms: Multi-class Categorization of Breast Tissue Data, Intern, J. of Comp. Appli., 32(7): 46-53.
35. Lakshmi Priya and S.G. Jacob, 2014. Predicting Protein-Protein Interactions through Associative Rule Mining Techniques: A Comparative Study, International Conference on Intelligent Information Technologies, pp: 198-204.
36. Prakash, G., S.G. Jacob and S. Radhameena, 2014. Mining Semantic Representation From Medical Text: A Bayesian Approach, International Conference on Recent Trends in Information Technology 978-1-4799-4989-2/14$31.00, IEEE.
37. Shomona Gracia Jacob and Dr. R. Geetha Ramani, 2012. Mining of Classification Patterns in Clinical Data through Data Mining Algorithms, ISBN:978-1-4503-1196-0,ICACCI, 997-1003.
38. Kotsiantis, S.B., 2007. Supervised Machine Learning: A Review of Classification Techniques, Informatica, pp: 31249-268.
39. Kohavi, R., 1999. Ross Quinlan:Decision Tree Discovery.

40. Wu, X., V. Kumar, J.R. Quilan, J. Ghosh, Q. Yang, H. Motoda and D. Steinberg, 2008. Top 10 Algorithms in Data Mining Knowledge and Information System, 14: 1-37.
41. Ramani, R. Geetha and S. SivaSankari, 2015. Automatic Ontology Construction Through Decision Tree Classification Techniques, Adv. in Nat. Appl. Sci., 9(6): 639-644.
42. Milton, N., 2008. Knowledge technologies Polimetrica, Monza.
43. Waikato Environment for Knowledge Analysis (WEKA) Machine Learning Tool, 2013. www.cs.waikato.ac.nz/ml/weka/downloading.html.
44. Protege Toolkit with software services. http://www. protégé.stanford.edu.
45. McGuinness, D. and F. van Harmelen, 2004. OWL web ontology language overview W3C. [cited 18.12.10].
46. Knublauch, H., R. Fergerson, N. Noy and M. Musen, 2004. The Protégé OWL plugin: an open development environment for semantic web applications, The Semantic Web- ISWC, pp: 229-43.
47. O'Connor, M.J., H. Knublauch, S.W. Tu and M.A. Musen, 2005. Writing rules for the semantic web using SWRL and Jess. In: 8th international protege conference, Madrid, Spain.
48. O'Connor, M.J., R.D. Shankar, S.W. Tu, I. Nyulas and A.K. Das, 2008. Developing a web-based application using OWL and SWRL. In: AAAI spring symposium, Stanford, CA.
49. O'Connor, M.J. and A. Das, 2009. SQWRL: a query language for OWL, OWL: experiences and directions (OWLED). In: Fifth international workshop, chantilly, VA.
50. Horrocks, I., P.F. Patel-Schneider, H. Boley, S. Tabet, B. Grosof and M. Dean, 2004. SWRL: A semantic web rule language combining OWL and RuleML. W3C Member submission, pp: 21-79.
51. Golbreich, C. and A. Imai, 2004.Combining SWRL rules and OWL ontologies with Protégé OWL Plugin, Jess and Racer, In 7th International Protégé Conference, Bethesda, MD.
52. O'connor, M., H. Knublauch, S. Tu, B. Grosof, M. Dean, W. Grosso and M. Musen, 2005. Supporting rule system interoperability on the semantic web with SWRL, In The Semant ic Web- ISWC Springer Berlin Heidelberg, pp: 974-986.