ISSN: 2278-4632
Vol-10 Issue-8 No. 2 August 2020

Ontology-Based Analysis Semantic Correlation Interventions in the Field of Health

D.Akila[1], S. Jeyalaksshmi[2,] , K.Rohini[4], G.Suseendran[4] , V.R.Elangovan[5]
[1,2,3,4]Department of Information Technology, Vels Institute of Science, Technology & Advanced Studies , Chennai, India
[5]Department of Computer Science, Agurchand Manmull Jain College, Chennai, India

Email id : akiindia@yahoo.com[1], pravija.lakshmi@gmail.com[2], rrohini16@gmail.com[3], suseendar_1234@yahoo.co.in[4], elangotesting@gmail.com[5]

Abstract

In recent years there is an extraordinary development in the Web Ontology Language (OWL). This makes an increase in the quantity of knowledge that is stored under ontologies. Specifically, to match the entities glossed with ontology concepts, semantic similarly has been evolved. The semantic similarity in biomedical ontologies is one of the recent technology which are applied in variety of applications. This helps in providing exact information essential for biological datasets, the biomedical literature and patient records . Also, several studies are available in recent years to define and assess the different as approaches.

The Semantic similarity is considered the best method for interpreting the effects of biome dical studies such as gene clustering, gene expression data analysis, molecular interaction prediction and validation, and prioritization of disease genes. In this paper, we reviewed the semantic similarity measures based on biomedical ontologies and their classifications. We present an outlook on biological and biomedical ontologies and focusing on the functionalities and its support in integrative research. They often concentrate on current implementations of semantine similarity measures, and i dentify examples of biomedical research applications. This will explain how biomedical re searchers can take advantage of semantine similarity measures and help them select the m ost suitable method for their studies.

Keywords : biomedical ontologies, semantic similarity, synonymous words, information retrieval

I. Introduction

The equality between two concepts can be evaluated using Semantic Similarity Measures (SSMs).These measures are particularly useful in improved understanding of textual resources.

The presence in English of semantine equality groups between lexical objects makes it hig hly attractive to use thesauri applications with synonymous or almost synonymous words.

Due to serious, comprehensive needs on such IR obligations as patient group recognition, the matter is primarily relevant in the medical domain.

The Measures can be mail dived into two clusters. The first variety of measure is Distributional based methods[1]. In this method the distributed concepts jointly with various knowledge resources are used to evaluate the similarity. The corpus Information Content (IC) and context vector methods are the two measures under this category.The Second variety of measure is Knowledge based methods. In this method, ontologies and semantic networks are used as knowledge resources. The Path-based and intrinsic IC-based measures are the two further classifications of Knowledge based methods. Semantic similarity measures have been used in wide array of applications in biomedical domain, using biomedical ontologies [2]. Some of the applications are like to rephrasing text, to recommend medicine transposition, to grouping genes based on their molecular role. Semantic similarity measures are without doubt basic parts of numerous information based frameworks. These days semantic similarity is getting more consideration because of the developing selection of both Semantic Web and Linked Data paradigms.

Fortunately, the field of biomedical has been exceptionally productive in making medical ontologies which sort out ideas in a non-questionable manner to be utilized by semanticmeasures.Many of the popular biomedical ontology models include Medical Subj ect Headings (MeSH),International Disease Classification (ICD taxonomy) and Systematiz ed Medicine Nomenclature, Clinical Terms (SNOMED-CT). [4]

2.Ontology

2.1 Biomedical ontologies

Researchers have plenty of biomedical tools available, ranging from gene and protein sequ ence databases.There has been an growing demand in recent years for the exchange and in corporation of medical data into biomedical research. To strengthen a health care system, data integration must be facilitated by promoting semantine interoperability systems and p ractices[5].

2.1.1 Unified Medical Language System

The Unified Medical Language System (UMLS) can be used as an example of jargon that includes multiple clinical words and integrates approximately 100 different vocabulary[6]. The Unified Medical Language System (UMLS) has attracted time , skills and energy from a wide variety of disciplines[7] over the past 10 years. [4] The UMLS incorpo rates over 2 million names from more than 60 families of biomedical vocabulary for some 900 000 concepts, as well as 12 million relationships amo ng these concepts.UMLS Metathesaurus vocabulary includes NCBI taxonomy, Gene Onto logy, Medical Subject Headings (MeSH), OMIM, and Virtual Anatomist Symbolic Knowl edge Base. UMLS is composed of three information sources and a collection of interactive resources that can be used to access these knowledge sources in depth. Types of informati on include the Meta Thesaurus, the Semantic Network and SPECIALIST Lexicon[4].

2.1.2. MeSH: A for Anatomy, B for Organisms and C for Diseases, et.,MeSH terms are the NLM’s controlled terminology, primarily used to organize and index information and manuscripts found in common databases such as Pub- Med [8].

2.1.3.SNOMED-CT

ISSN: 2278-4632
Vol-10 Issue-8 No. 2 August 2020

SNOMED-CT (Systematized Nomenclature of Medicine, Clinical Terminology), developed by the College of American Pathologists, was created through the integration of SNOMED RT and Clinical Terms Version 3 (formerly known as the Read Codes)[9].SNOMED CT, a clinical terminology covering the whole clinical domain is an emerging standard for the representation of semantically explicit, structured information in electronic health records by providing more than 300,000 meaning bearing representational units (concepts)[10]. SNOMED-CT as an internationally accepted standard ontology is included in the UMLS repository.[10]

2.1.4.NCI Thesaurus (National Cancer Institute Thesaurus): A ontology vocabulary that incorporates expansive inclusion of the malicious growth area, including distortion related illness, life structures, qualities and medications

2.1.5. ICD-10: remain of International Classification Diseases version 10 is a standout amongst the most essential international therapeutic expressed frameworks; it was first issued in 1893.

3. Semantic similarity measures

Semantic similarity is a fundamental and effectiveness concept to get the most relevant results[12]. Semantic similarity is related to measuring the similarity of terms that are not lexically identical. This is an important topic in research on natural language processing (NLP) and information recovery (IR) have gained significant interest in the literature[13]. The extension of queries is one of the key applications of semantic similarity. It is a method of changing an initial query to boost the efficiency of retrieval of information in retrieval operations[14]. The aim of query expansion techniques[15] is to extend the query in order to add more term that suits the original terms. Semantic similarity, semantic correlation and semantic distance are the measures used in the semantic concept[16]. Different methods were designed to calculate this notion of semantic similarity.

Here, the calculation of semantic similarity can be loosely divided into two (1) Path Based semantic similarity measure and (2) semantic similarity measure based on information quality. The First measure variety offers descriptions of the co-location of the relationships in an arrangement (i.e) classification. The second measure variety uses the specifics of the arrangement with the connections to certain definitions in addition. The two methods used to determine the calculation of semance similarity

Both intrinsic and corpus based assessments are focused on knowledge quality.Many of the semantic similarity steps were introduced in the biomedical sector by adding domain in-information from clinical data or medical ontology.

4.Results & Discussion

4.1. Datasets

There are no specific collections of terms / concepts concerning conceptual similarities and relatedness within the biomedical context, such as the M&C or general English R&G collections. The Pedersen, Pakhomov, & Patwardhan (2005) set of 30 pairs of ideas was used to compare approaches, annotated by 3 physicists and 9 experts in the theoretical index. Increasing pair had a 4-point scale annotated: "Practically identical, related, marginal, and unrelated. The average ratio between doctor and doctor is 0.68, and 0.78, respectively.

Table 1. Absolute correlation values of the five factors with respect to human decisions

Measure	Physician (rank)	Expert (rank)
Path length	0.627(4)	0.852 (3)
Leacock & Chodorow	0.672 (1)	0.856 (2)
Wu & Palmer Lin Measure	0.652 (3) 0.560 (5)	0.794 (4) 0.724 (5)
Al-Mubaid & Nyguan's measure	0.666 (2)	0.862 (1)

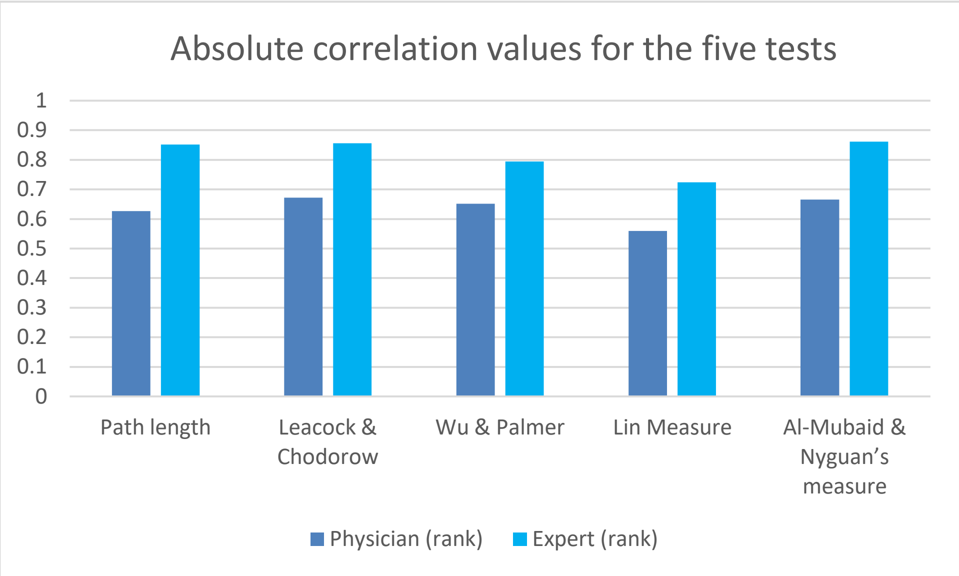

Figure 1. Absolute correlation values for the five tests

But the second marked their approach Regarding physician decisions. Since the expert scores are more correct, since the expert scores correlation (0.78) is higher There are more experts than doctors between doctors (0.68) and doctors (3 doctors & 9 experts).

5.Conclusion and Future Works

We contrasted a measure of semantic similarities, based on ontology. The studies discussed in this paper have shown the superiority of the approach used by Al-Mubaid & Nyguan in terms of human decisions and in contrast with other ontology-based interventions. In future work on this paper we plan to discuss studies in the biomedical domain with implementations of semantine relatedness interventions for NLP tasks such as word sense detection, information retrieval, and spelling correction. We also use the collection to compare taxonomies and calculate semantic similarities between and across two definitions in sources of UMLS terminology. Finally, we plan to develop a web-based user interface for all of these semantine similarity measures and make it freely accessible to researchers over the Web. This will be very useful for interested biomedicine researchers.

6. References

[1] T. Pedersen, S. V. S. Pakhomov, S. Patwardhan, and C. G. Chute, "Measures of semantic similarity and relatedness in the biomedical domain," *J. Biomed. Inform.*, vol. 40, no. 3, pp. 288–299, 2007.

[2] M. Zare, C. Pahl, M. Nilashi, N. Salim, and O. Ibrahim, "A Review of Semantic Similarity Measures in Biomedical Domain Using SNOMED-CT," *J. Soft Comput. Decis. Support Syst.*, vol. 2, no. 6, pp. 1–13, 2015.

[3] T. Pedersen, "Measures of Semantic Similarity and Relatedness in the Medical Domain," *Int. Classif.*, 2004.

[4] O. Bodenreider, "The Unified Medical Language System (UMLS): integrating biomedical terminology," *Nucleic Acids Res.*, vol. 32, no. 90001, p. 267D–270, 2004.

[5] F. Shen and Y. Lee, "Knowledge discovery from biomedical ontologies in cross domains," *PLoS One*, vol. 11, no. 8, pp. 1–34, 2016.

[6] A. M. B. Abdelrahman, "Evaluating Semantic Similarity between Biomedical Concepts / Classes through Single Ontology," vol. 7, no. 08, pp. 341–356, 2018.

[7] J. Brennan, R. J. Sullivan, D. J. Bryant, S. Glasel, K. Macek, and E. L. Olson, "Focus on," *Nurs. Made Incred. Easy!*, vol. 16, no. 2, pp. 44–49, 2018.

[8] T. M. Beissinger and G. Morota, "Medical Subject Heading (MeSH) annotations illuminate maize genetics and evolution," *Plant Methods*, vol. 13, no. 1, pp. 1–8, 2017.

[9] D. Rubin and N. Shah, "Biomedical ontologies: a functional perspective," *Brief. Bioinform.*, 2008.

[10] C. Martínez-Costa and S. Schulz, "Ontology-based reinterpretation of the SNOMED CT context model," *CEUR Workshop Proc.*, vol. 1060, pp. 90–95, 2013.

[11] A. M. B. Abdelrahman and A. Kayed, "A Survey on Semantic Similarity Measures between Concepts in Health Domain," *Am. J. Comput. Math.*, vol. 05, no. June, pp. 204–214, 2015.

[12] D. Akila and C. Jayakumar, "Acquiring evolving semantic relationships for WordNet to enhance information," *Int. J. Eng. Technol.*, vol. 6, no. 5, 2014.

[13] B. R. Ragot and R. Schlickeiser, "Cosmic ray acceleration by fast

ISSN: 2278-4632
Vol-10 Issue-8 No. 2 August 2020

magnetosonic waves," *Aston. Astrophys*, vol. 331, pp. 1066–1069, 1998.

[14] D. Akila, A. Vidhya, and P. Rajesh, "Optimization based information retrieval with the enhancement of annotator in wordnet application," *J. Adv. Res. Dyn. Control Syst.*, vol. 10, no. 2, 2018.

[15] D. Akila, S. Sathya, and G. Suseendran, "Survey on query expansion techniques in word net application," *J. Adv. Res. Dyn. Control Syst.*, vol. 10, no. 4, 2018.

[16] D. Akila and C. Jayakumar, "Semantic similarity- a review of approaches and metrics," *Int. J. Appl. Eng. Res.*, vol. 9, no. 24, 2014.