Personalized Medical Reading Recommendation: Deep Semantic Approach

Tatiana Erekhinskaya[✉], Mithun Balakrishna,
Marta Tatu, and Dan Moldovan

Lymba Corporation, 901 Waterfall Way, Bldg 5, Richardson, TX 75080, USA
{tatiana,mithun,mtatu,moldovan}@lymba.com
http://www.lymba.com

Abstract. Therapists are faced with the overwhelming task of identifying, reading, and incorporating new information from a vast and fast growing volume of publications into their daily clinical decisions. In this paper, we propose a system that will semantically analyze patient records and medical articles, perform medical domain specific inference to extract knowledge profiles, and finally recommend publications that best match with a patient's health profile. We present specific knowledge extraction and matching details, examples, and results from the mental health domain.

Keywords: Deep semantic extraction · Medical information retrieval · Diagnostic inference

1 Introduction

With new scientific findings and studies being reported every day, a therapist is faced with the overwhelming task of identifying, reading, and incorporating new information from a vast volume of publications into their daily clinical decisions. Arming the therapist with the most current literature would help the therapist make the best clinical decisions for their patients throughout the course of diagnosis and treatment.

This paper addresses the task of recommending relevant professional reading for doctors based on their current patient cases. In comparison to standard Information Retrieval task, this task has several complications that make keyword-based search inefficient. First, the query is not a short set of keywords, but a set of relatively large text files, which requires keyword importance evaluation and high performance. Second, the language of patient records is different from the language of papers, which makes keyword matching insufficient. Finally, some publications are more research oriented and do not address therapist needs directly, for example discussing experiments on rats, statistic analysis on population, etc. - the knowledge that does not have immediate clinical implications.

This paper presents a novel NLP-based approach to compute relevance of the candidate papers to the set of cases a therapist has on hand based on deep semantic processing of publications and electronic health records (EHR).

Both EHR and publications are converted into semantic profile. The relevance is computed based on the profiles matching. In addition to relevance, the system computes novelty score to measure how much new knowledge is provided by a candidate publication.

2   Related Work

2.1   Concept Extraction and Expansion

The problem of long queries in medical domain brings the task of extraction important concepts and assigning corresponding importance weight in a ranking formula. MedSearch system [10] was designed to assist ordinary Internet users to search for medical information by accepting queries of extended length. The system rewrites long queries by selectively dropping unimportant terms based on tf-idf scores.

Zheng and Yu [15] also targeted patients as end users. They trained LDA topic models to identify prominent topics. Queries are generated from n-grams, taking the top 5 phrases as queries from the topics that has a combined probability of over 80 %. The authors also employed Conditional Random Fields (CRF) model to identify key concepts, which are most in need of explanation by external education materials. The authors have shown that using full EHR notes is ineffective at retrieving relevant education materials.

Query expansion is a well-known technique in traditional Information Retrieval [13]. Liu and Chu proposed a knowledge-based query expansion technique to support scenario-specific retrieval [9], when query contains general terms like *treatment* that need to be matched to specific terms like *chemotherapy* in the document. The method utilized co-occurrence thesaurus, UMLS and vector space model.

2.2   Usage of Dependencies

The key concepts in the query and in the documents are forming structures that are important for relevance scoring. Choi et al. [7] uses implicit dependencies with the standardized medical concepts to favor the documents that preserve those implicit dependencies to improve ranking performance. The implicit dependence features were harvested from the original query using MetaMap [2]. These semantic concept-based dependence features were incorporated into a semantic concept-enriched dependence model (SCDM).

2.3   Negative Findings

Negative findings in patient records are expressed by means of negation or by using terms which contain negative qualifiers. From IR point of view, negative findings should be recognized and treated in a special way. Namely, EHR and relevant publications should agree on whether the finding is negative, or the negative finding in EHR might be not mentioned in the publication.
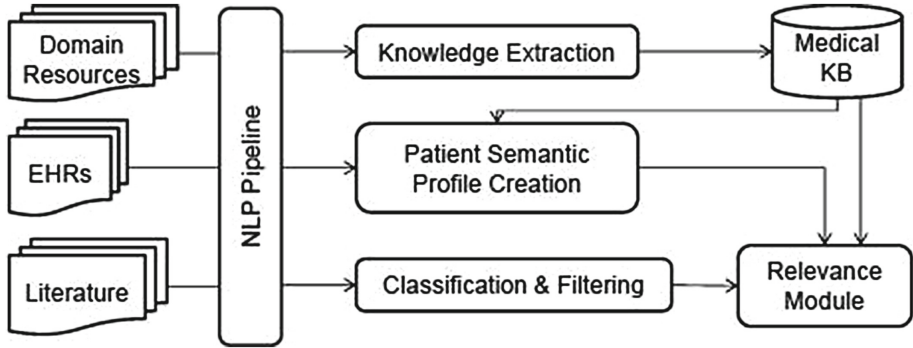
Fig. 1. Dataflow for medical literature recommendation approach.

Ceusters et al. [6] classified these phenomena in terms of the various top-level categories and relations defined in Basic Formal Ontology [8] and taking into account the role of negation in the corresponding descriptions. The authors introduced the *lacks*-relation that allowed them to represent nearly all negative findings that occur in patient charts.

3   Proposed Approach Overview

3.1   Problem Formulation

Given a set of patient cases $\{P_1, P_2, ..., P_k\}$ and past knowledge of the therapist K, the literature recommendation module will return a ranked list $R = [r_1, r_2, ..., r_n]$ of publications with the links between suggested publications and original patient cases $r_i \rightarrow p_j$. Past knowledge K consists of medical profiles of past cases and previously read papers.

The relevance should be computed taking into account the following therapist information needs: (1) diagnosis methods; (2) new, more efficient treatments for known diseases; (3) adverse effects of prescribed treatment; (4) potential risk factors for new health problems.

As the therapist updates a patient's file and adds case notes, the semantic model for the patient will continue to update such that relevant reference articles are presented that may justify the current diagnosis.

The literature recommendation to the clinician can be presented directly at the point of care, as they type in session notes during an ongoing clinical interview as well as in an offline, proactive manner.

3.2   Dataflow Overview

Figure 1 shows the dataflow of the proposed approach. First, the patient records are processed via the NLP Pipeline. The key task is to extract medical concepts: symptoms, diseases, administered treatment, medication, life events, etc.

Then symptoms are normalized, for example, *eating without control* would be matched to *binge eating*. This information about the patient is put into Semantic Patient Profile. Then, the inference module suggests possible diagnosis with some confidence score. This diagnosis can be used as a suggestion for doctors in the beginning of patient care process, as an alternative consideration for doctor-provided diagnosis, and as additional strong keyword for retrieval in case no diagnosis was provided by a therapist. The diagnosis and standardized symptoms are taken from Medical Knowledge Base, that was created based on existing resource like Mesh [1] and SnoMED [14] and extracted from textbooks and manuals. The publications are processed with NLP tools and semantically indexed. In addition, the publications are classified according to the therapist needs. There is a boolean Naive Bayes classifier for each need. The publications that do not match any of the needs are filtered out.

4   NLP Pipeline

The first step of deep semantic processing of medical text is the NLP Processing that spans the lexical, syntactic, and semantic layers of knowledge extraction from text.

Our concept detection methods range from the detection of simple nominal and verbal concepts to more complex named entity and phrasal concepts. This hybrid approach to *concept extraction* makes use of machine learning classifiers, cascade of finite-state automatons, and lexicons to label more than 80 types of concept classes. The concept categories with examples are shown in Table 1. Note, that the categories can be expressed not only with nouns which are easy to extract from ontologies, but with other part of speech words as well, also a concept can have nested concepts in it, as the ones in the bottom of the table.

Table 1. Partial list of recognized medical concept types.

NE type	Examples
Symptom	headache, depressive mood, behavior change, gained weight
Sign	blood pressure, temperature
Diseases & disorders	Aarskog Ose Pande syndrome, obese
Temporal qualities	acute, chronic, episodic, history, x 2 weeks
Body part	leg, mid-calf, cardiac, dental
Life event/condition	toxic exposure, stress, traveling abroad
Addictions	smoker, 1 glass of wine every day
Medical examination	ophthalmoscopy, lung exam, screening
Medical procedure	UPPP, reconstructive surgery, organ transplant
Medication	NSAID, angiotensin-converting enzyme inhibitor
Allergy and sensitivity	hay fever, cryesthesia, anaphylaxis
Medical problem	abnormal lung exam, abdominal pain with palpation

The extracted concepts are normalized using standard formulations in existing knowledge bases via semantic matching. For example, *lost 5 pounds* in EHR is matched to *weight loss* in Medical Subject Headings.

Semantic relations allow the linking of important concepts in a correct way. For example, they help connect temporal information and a medical problem, determine whether a medical problem is related to a patient or belongs to the family history, etc. Co-reference resolution module extracts co-reference chain information to help separate patient specific symptoms and features from other mentions in the patient data.

We define *semantic relations* as abstractions of underlying relations between concepts that occur within a word, between words, between phrases, and between sentences [11]. Semantic relations provide connectivity between concepts, which makes their extraction from text essential for the ultimate goal of machine text understanding. We use a fixed set of 26 relationships, which strike a good balance between too specific and too general [11]. They include the thematic roles proposed by Fillmore and others, and the semantic roles in PropBank, while also incorporating relationships outside of the verb-argument settings, representing semantic connectivity for all content words.

The important module in the pipeline is negation recognition. Negations are used to reverse polarity of a statement. In medical domain it can mean a health issue (e.g. *absent tonsil*) or absence of signs/symptoms (negative findings), which is critically important for providing diagnosis and literature recommendation. The negation module determines the scope and focus of negations and incorporate negations into semantic representation [4,12]. Negations can be expressed with auxilary words like not, without, or with content word, (e.g. denies, stop, cancel, never, absence, absent, etc.)

5   Medical Knowledge Base and Diagnostic Inference

In order to support diagnostic inference, we designed a specific knowledge extraction module that extracts diagnostic requirements such as the diagnostic criteria, diagnostic features, development and course, and the differential diagnosis for each disease described in literature. For example in Reactive Attachment Disorder, eight criteria must be evaluated, a subset is shown in Table 2.

The NLP tools read the detailed descriptions of each disorder and translate them into a graph of concepts and semantic relations. The disorder is represented as a seed node with customized semantic connections to: (1) a list of typical signs and symptoms, (2) any related medical conditions, (3) familial and culture predispositions, (4) typical faith system, (5) IQ, (6) gender, (7) age, (8) any chemical use, (9) psychosocial factors, (10) a detailed representation of the critical criteria and (11) an encoding of the differential diagnosis.

Figure 2 presents a partial view of the semantic representation that we designed to encode the diagnostic requirements such as the diagnostic criteria, diagnostic features, development and course, and the differential diagnosis. We represent the diagnostic information as structured relations with normalized

Table 2. Subset of criteria for Reactive Attachment Disorder.

| A	A consistent pattern of emotionally withdrawn behavior toward adult care giver manifested by both of the following: (1) child rarely seeks comfort when distressed, (2) child minimally responds to comfort when distressed
D	The care in Criterion C is presumed to be responsible for the disturbed behavior in Criterion A.
E	The criteria are not met for autism spectrum disorder
F	The disturbance is evident before age 5 years
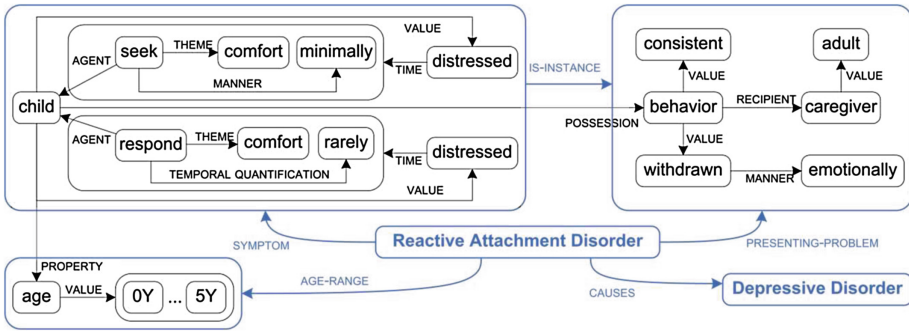
Fig. 2. Semantic representation for a reactive attachment disorder diagnostic criterion.

values for reasoning. Figure 2 also shows the inferred health-specific semantic relations (e.g. AGE-RANGE, SYMPTOM, PRESENTING-PROBLEM, etc.) that were derived using Semantic Calculus [5], a tool for combining the 26 core semantic relations into domain specific relations.

The diagnostic inference module uses this representation to match patient profile and diagnostic criteria. The rest of the section explains the inference on the example of Reactive Attachment Disorder's criteria from Table 2. Criterion A requires that both (1) and (2) be present. For this reason, we encoded inclusion/exclusion, and minimal/maximal semantics for the critical criteria. Criterion D seeks a causation relationship between Criterion A and Criterion C. If any of the factors are true for Criterion C, the diagnostic module checks for a causation relationship with the factors in A. Criterion E introduces the complexity of negation as well as the requirement to assess autism spectrum disorder. To resolve this issue, the system navigates to autism spectrum disorder, evaluates the criteria, and then proceeds with the diagnostic assessment. Finally, Criterion F expects a temporal interval attached to the disturbance event. The system interprets the disturbance as the compilation of the signs and symptoms in order to perform temporal reasoning to decide if they occurred before age 5.

6   Relevance Computation

The relevance module matches publication profiles to semantic patients' profiles and identifies articles that bring new information to the therapist outside the body of knowledge they already have consulted.

Profile comparison algorithm computes the semantic overlap between a patient file and an article by weighed summation of matches for concepts and relations:

$$R = \sum_{i \in concepts(SPS)} w^c{}_i m^c{}_i + \sum_{i \in relations(SPS)} w^r{}_i m^r{}_i. \qquad (1)$$

In this equation, m denotes match between concept/relation from the semantic patient profile to the publication profile, range from 0 (no match) to 1 (full match) with similarity score in between. Two semantic relations are said to match if their domain and range concepts are the same. Weight w denotes importance. A concept's importance weight is based on its tf-idf score [3] and its linguistic properties. Inferred concepts (e.g. diagnosis) are scored lower than the original ones. Importance weight for relations is based on the domain/range concept importance score and its thematic properties such as its relation type and connection strength.

Figure 3 shows the concept and relation match for the patient file and the article discussing treatment for Reactive Assessment Disorder. The gray concepts show the semantic overlap used to determine relevance.

The system also measures the degree of novelty of the article with respect to past knowledge by identifying the scientific nuggets in the article that provide new information. While article relevance is derived from matching semantic profiles of the patient file and article, the novelty is derived from matching the past knowledge with the article profile. The novelty score is then computed as the semantic difference between the candidate article model and the patient file
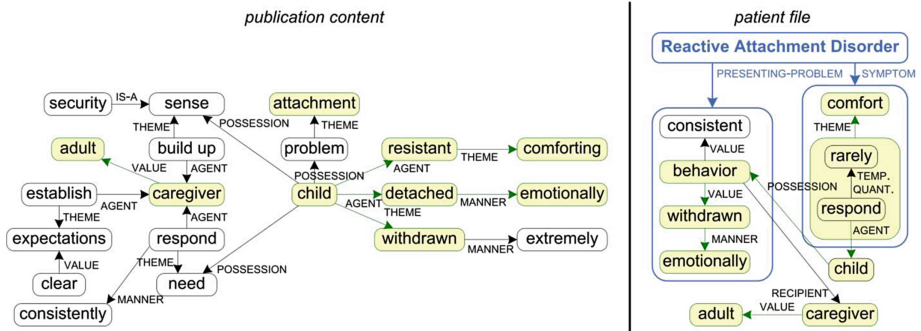

Fig. 3. Semantic representation for a reactive attachment disorder diagnostic criterion.

model augmented with models from previously suggested articles. The information conveyed by an article that could not be mapped to the knowledge stored in the patient's semantic profile is considered to be novel. The system computes the novelty score for an article using the following features: (1) weights new concepts higher than new relations that link known concepts, and (2) prefers explicitly stated knowledge to entailed knowledge from the domain ontology. The overall novelty of a scientific article is computed as the average of the novelty scores associated with each of its meaning constituents (e.g., concepts and semantic relations).

Figure 3 demonstrates the novelty computation operation for an article discussing new treatments for Reactive Attachment Disorder with the patient file from Task 1. The white concepts are the results of the semantic difference operation and indicate the novel information from the article.

7   Evaluation

The evaluation of the approach was done for mental health domain, since this domain has a comprehensive manual - DSM-5 book.

To evaluate the disorder recommendation module, we collected case studies from mental health disorder case study books or online resources. Using this data, we measured the quality of diagnosis recommended at the top-1, top-5, and top-10 levels in terms of accuracy. The disorder recommendation module obtained 62 % (top-1), 82 % (top-5), and 89 % (top-10) accuracy scores.

To evaluate the literature recommendation module, we selected 100 case studies from the test dataset created for the diagnosis module evaluation. Two subject matter experts searched online for articles related to the case studies and tagged two articles for each case study. They then evaluated the articles recommended by our system and scored the relevance and novelty of the articles on a scale of 1–5, with 5 being highly relevant/novel and 1 being not relevant/novel. The literature recommendation module obtained 77 % (top-1), 94 % (top-5), and 95 % (top-10) accuracy scores for relevance, and 21 % (top-1), 44 % (top-5), and 55 % (top-10) accuracy scores for novelty.

8   Conclusion

In this paper, we presented a semantic driven approach to performing literature recommendation that provides therapists with the most current, novel, and relevant literature based on their patient files. We avoided the usual pitfalls of keyword and concept driven search by semantically analyzing patient records and medical articles, performing medical domain specific inference to extract knowledge profiles, and finally recommending publications that best matches a patient's health profile. Deep semantic processing allows expansion, normalization and filtering of the publication content and the patient record. We applied our proposed system to the mental health domain and obtained promising evaluation results for the case studies specified in the DSM-5 book.

References

1. Medical subject headings (mesh). https://www.nlm.nih.gov/pubs/factsheets/mesh.html
2. Aronson, A.R.: Effective mapping of biomedical text to the umls metathesaurus: the metamap program. In: Proceedings of AMIA Symposium, pp. 17–21 (2001). http://view.ncbi.nlm.nih.gov/pubmed/11825149
3. Balakrishna, M., Moldovan, D., Tatu, M., Olteanu, M.: Semi-automatic domain ontology creation from text resources. In: Proceedings of the Seventh International Conference on Language Resources and Evaluation (LREC 2010), Valletta, Malta, May 2010
4. Blanco, E., Moldovan, D.: Semantic representation of negation using focus detection. In: Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies, pp. 581–589. Association for Computational Linguistics, Portland, Oregon, USA, June 2011. http://www.aclweb.org/anthology/P11-1059
5. Blanco, E., Moldovan, D.: Unsupervised learning of semantic relation composition. In: Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies, pp. 1456–1465. Association for Computational Linguistics, Portland, Oregon, USA, June 2011. http://www.aclweb.org/anthology/P11-1146
6. Ceusters, W., Elkin, P., Smith, B.: Negative findings in electronic health records and biomedical ontologies: a realist approach. Int. J. Med. Inform. **76**(Suppl 3), 326–333 (2007). http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2211452/?tool=pubmed
7. Choi, S., Choi, J., Yoo, S., Kim, H., Lee, Y.: Semantic concept-enriched dependence model for medical information retrieval. J. Biomed. Inf. **47**, 18–27 (2014). http://www.sciencedirect.com/science/article/pii/S153204641300141X
8. Grenon, P., Smith, B., Goldberg, L.: Biodynamic ontology: Applying bfo in the biomedical domain. Stud. Health Technol. Inform. **102**, 20–38 (2004)
9. Liu, Z., Chu, W.W.: Knowledge-based query expansion to support scenario-specific retrieval of medical free text. Technical report, Information Retrieval (2005)
10. Luo, G., Tang, C., Yang, H., Wei, X.: Medsearch: A specialized search engine for medical information retrieval. In: Proceedings of the 17th ACM Conference on Information and Knowledge Management, CIKM 2008, pp. 143–152. ACM, NY, USA, New York (2008). http://doi.acm.org/10.1145/1458082.1458104
11. Moldovan, D., Blanco, E.: Polaris: Lymba's semantic parser. In: Proceedings of LREC-2012, pp. 66–72 (2012)
12. Morante, R., Blanco, E.: *SEM 2012 Shared task: resolving the scope and focus of negation. In: Proceedings of the First Joint Conference on Lexical and Computational Semantics (*SEM 2012), pp. 265–274. Montréal, Canada, June 2012
13. Qiu, Y., Frei, H.P.: Concept based query expansion. In: Proceedings of the 16th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval (1993)
14. Spackman, K.A., Campbell, K.E., Côté, R.A.: SNOMED RT: A reference terminology for health care. In: Proceedings of the AMIA Annual Fall Symposium, pp. 640–644 (1997)
15. Zheng, J., Yu, H.: Key concept identification for medical information retrieval. In: Mrquez, L., Callison-Burch, C., Su, J., Pighin, D., Marton, Y. (eds.) EMNLP, pp. 579–584. The Association for Computational Linguistics (2015). http://dblp.uni-trier.de/db/conf/emnlp/emnlp2015.html#ZhengY15