Biomedical Event External Reference Entity Linking Using Neural Networks

Ian Gross **Shruthi Chari**
Rensselaer Polytechnic Institute, Troy, NY, USA
{grossi2, charis}@rpi.edu

1 Introduction

Within [1] the semantic web field, there is a strong desire to represent knowledge in a linked data format. Although large portions of knowledge for the biomedical domain already exist, we encounter challenges in relation to incorporation of new biomedical literature utilizing domain-relevant ontologies. Solutions to this problem exist in the form of Information Extraction, which is the formal process of interpreting events and entities from unstructured or semi-structured documents (Jurafsky and Martin, 2008). Entity linking is a sub-task within Information Extraction that allows for potential grounding of terms to external references.

Although entity linking for biomedical literature has been accomplished previously, the approach taken lacks specificity in its ontology hierarchy. To remedy this issue, we develop an architecture to automatically link ontology references to biomedical entities from event extraction output. This approach utilizes the semantics of Abstract Meaning Representation (AMR) and biomedical ontology terms. By representing this data in vector representations, we can maximize similarity between entity to ontology term pairs using deep learning models.

2 Motivation

Information and smart knowledge representation is considered to be a valuable asset in the pursuit of a wide variety of data applications. The focus of these resources is to effectively manage, analyze, and preserve domain-specific information. The semantic web attempts to solve some of these issues by providing standard definitions, hierarchical structure of knowledge, and effective methods to classify relationships between data. The idea behind incorporating semantics into pre-existing knowledge is to enable applications to utilize the information effectively and provide pathways to future studies on this knowledge.

For this study, we focus primarily on the biomedical domain, due to the availability of its literature via publicly available on-line resources and the wealth of biomedical concepts grounded within biomedical ontologies. The primary task of this work is to be able to automatically extract information from biomedical literature and produce logical representations of this information known as knowledge bases, adhering to common semantic web principles. In order to accomplish this ideology, we must incorporate natural language processing techniques to derive meaning behind biomedical text. The specific field of natural language processing that refers to the interpretation of information is known as Information Extraction.

From a high-level perspective, we are able to extract knowledge from biomedical literature in the form of events and entities. Events serve as the "predicate" nodes with the knowledge graph construction, and entities help to describe events in the context of some action being performed on some object(s). The entity-linking task provides the ability to ground entities from extracted text to concepts or individuals within biomedical ontologies. Although this problems has been successfully attempted in previous studies, these attempts lacked specificity in terms of full representation of the employed ontology's hierarchy and term representation. Additionally, we discovered that traditional Machine Learning approaches try to minimize the error between the predicted output and actual output, which is insufficient in cases where the output label is rich in context. Instead we could maximize the similarity between the mapping rep-

[1] Code available at : https://github.com/IanGross/BioEventXrefEntityLinker/tree/master/ELTask

resentation of input (word) and the output representation (Ontology). As a result of these conclusions, we propose a novel framework for entity linking for biomedicial entities to ontologies terms utilizing the rich context provided by the dependency semantics sentences of its sentence and the properties describing ontology concepts.

3 Background

3.1 Linked Data

The semantic web is designed such that information can be both machine and human readable via the World Wide Web. Data is typically expressed in the Resource Description Framework (RDF) language, which is a web standard method for data representation and retrieval. The formatting for this information is in a subject, predicate, object form known as RDF triples. To demonstrate, a term may have a label, an ontology it belongs, a comment, a superclass, subclasses, and a Uniform Resource Identifier (URI). Ontologies are simply defined as a logical hierarchy of data to describe a particular subset of knowledge. Knowledge graphs build upon ontologies by utilizing several ontologies to provide even further context to a set of knowledge.

Although it is not commonly expressed within the current of data proliferation, it is necessary to consider the importance of introducing semantics into datasets. URIs provide the ability to express knowledge as pre-defined concepts. For instance, if the term “cell” is mentioned in a database, how can we interpret the meaning of the term such that this meaning is expressed uniformly for each usage of the word? By providing a URI, we can provide a standard definition for our knowledge. Additionally, the hierarchical structure of triples introduce the basis for a wide variety of knowledge discovery mechanisms.

3.2 Information Extraction and Entity Linking

As described previously, Information Extraction is the process of extracting implied meaning from a provided input. For this study, we focus primarily on methods for event extraction and entity linking. An event can be described by a change based on the context of its entities. Entity linking refers to the process of linking entity mentions in a document collection to entities in a reference knowledge base, otherwise known as wikification. Named-entity resolution usually handles the task of classifying terms under pre-defined categories such as organization or location. Using a pre-existing event extraction system provides a use case of the usage of our entity linker in potential applications.

One popular format that provides a logical representation for information is known as Abstract Meaning Representation (AMR). AMR is represented in hierarchical structure and is useful for tasks involving natural language understanding (Banarescu et al., 2013). Countless numbers of papers have been published using AMR as its representation format. We leverage the structure and underlying meaning conveyed using AMR as a means of understanding the context in which a term is used in its source sentence.

3.3 Deep Neural Networks

Deep Learning has proliferated all domains in which traditional machine learning techniques were applied, it automates the process of feature engineering. Deep Learning Models and Neural Networks in general, are adept at identifying intermediate representations/features that contribute to the target function $f(x)$ mapping $Input(X)->Label(Y)$. Deep Learning techniques have greatly benefited Supervised learning where there is a clear/fuzzy correspondence between the input and output. Although Deep Neural Networks were originally applied to Computer Vision tasks, more recently they have found their application in the Natural Language Processing World as well.
Several classes of DNN models like Convolutional Neural Networks, Long Short Term Memory networks have been used in tasks ranging from Named Entity Recognition(NER) to Machine Translation (Young et al., 2017). The ongoing research in word embeddings, character level embeddings, sentence embeddings, sub-word embeddings has made it feasible for Deep Neural Networks to leverage the semantics and structure of Natural Language representations. The field of word representations alone is being studied upon heavily, and there is an effort to enrich vector representations by utilising both the local and global context of occurence of words/phrases. Biomedical articles contain scientific terms whose represenations can benefit from a more granular global context, and disentangled representations (Jain et al., 2018) have been applied for the same.

Convolutional Neural Networks(CNN) (LeCun et al., 1998) are a class of Neural Network models that are used to summarise/condense inputs and also extract representative features from the input. They are often used as a feature extraction model and serve as inputs to other models like Long Short Term Memory(LSTM) models and Adverserial Neural Networks. The depth and richness of features learnt is based on the number of rounds of convolution applied and the number of filters utilised for the same. CNNs have also been used for classification and regression tasks where a softmax activation is applied on the output of the last fully connected layer to predict the outcome of the trained model. Convolutional Neural Network have been utilised to capture the semantic context necessary for the entity linking task (Francis-Landau et al., 2016), CNNs help construct topical vector representation that can be used to ground entities to the authoratative mentions. We jointly take advantage of this capability of CNNs in conjunction with the neighborhood representation captured by the AMR representation of a sentence to predict the ontology mention for an entity term.

3.4 Related Work

The primary paper that most closely resembles this study focuses on entity linking for biomedical literature (Zheng et al., 2015). Within this paper, they propose an entity linking system that utilizes context analysis graphs, candidate retrieval, noncollective entropy ranking for candidates, and collective inference between the candidates and the document graph. Additionally, the paper showcases the importance of utilizing the hierarchy structure of ontologies and the context of a term within a sentence or paragraph within the context of biomedical entity linking. The work we have completed differs immensely in terms of the methodology used our ability to incorporate more about ontology terms than their hierarchy. In order to accomplish this distinction, we utilize deep learning methods fully maximize the AMR format and the full ontology structure.
Although we haven't fully leveraged the hierarchial nature of the ontology structure, we plan to use Tree-Structured LSTMs/RNNs (Tai et al., 2015) for the same.Another important issue we plan to address is the lack of training data in the Biomedical domain, even more so the lack of annotated training data to ontologies. The Zero Shot Learning approach has been applied to the event extraction task (Huang et al., 2017) and has shown promising results, we apply this to the entity linking task wherein we are able to ground entities to unseen ontology terms.

4 Methodology

4.1 Approach

The primary focus of our approach relied around the implementation of the term to biomedical entity linking architecture. In order to demonstrate its capabilities, we design a high-level overview of the formal data workflow utilizing the proposed entity linking architecture. Data from the Colorado Richly Annotated Full Text Corpus (CRAFT), a manually annotated corpus consisting of 67 full-text biomedical journal articles from the PubMed Central Open Access Subset (Bada et al., 2012), is employed for testing of our architecture. This corpus includes annotations for terms to 7 biomedical ontologies, in which our goal is to correctly identify these annotations with unsupervised deep learning methods.

4.2 High-Level Architecture

In this section, we describe the entire data workflow utilizing the CRAFT corpus and how we incorporate the developed entity liking architecture into a potential use case. This workflow visualized in Figure 1, showcasing the data transformations that we will be describing throughout the section.

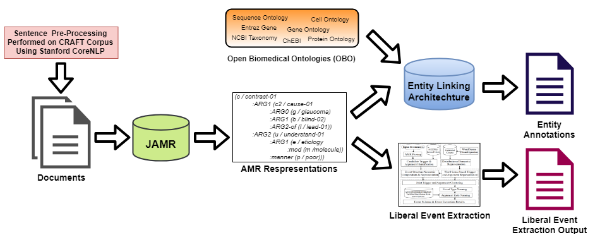

Figure 1: Diagram visualizing the high-level data workflow and potential accompanying use case for the entity linking architecture.

The CRAFT corpus provides several versions of the biomedical journal articles, of which we use the plain-text version and the annotated term files for each ontology. The initial goal of the pre-processing phase is convert each journal article into AMR, which is necessary for both the entity linking architecture and the event extraction system. The AMR format is capable of representing the literal meaning of a sentence, requiring the individual to split a journal article into sentences. To automate this process, we incorporate the tokenizer annotator of Stanford CoreNLP Natural Language Processing Toolkit to split our journal articles into sentences (required additional processing to handle edge cases). From this point, we incorporate the JAMR semantic parser for AMR (Flanigan et al., 2014) to convert our sentence files into an aligned AMR format. This parser was chosen due to its ease of use, in comparison to the other publicly available AMR parsers.

At this point, it is necessary to associate each term within the ontology annotated term files from CRAFT with the split sentences parsed through our Standford CoreNLP script and the associated AMR parsed through JAMR. An AMR reader was necessary at this step to process the AMR, which was created by (NEED INFORMATION HERE). The extracted AMR and ontology terms were then converted to vector representations using Word2Vec models (Mikolov et al., 2013) and act as input to the entity linking architecture. This mapping process was handled by a Python script written for this study. More details about the entity linking architecture are to be provided in Section 4.3.

To demonstrate the potential application of our entity linking architecture, we showcase an event extraction system alongside our entity linker. The event extraction system we utilize is the Liberal Event Extraction paradigm, developed by Lifu Huang at Rensselaer Polytechnic Institute (Huang et al., 2016). Liberal Event Extraction takes AMR as input and identifies hierarchical events based on trigger and argument joint clustering frameworks. Using the output from this event extraction system, we can directly associate entity terms within events. This has several important implications for knowledge graph application systems, providing a method of identifying events from literature and grounding entities to corresponding ontological representations.

4.3 Ontology Entity Linking Architecture

The Entity Linking framework consists of two Convolutional Neural Network Models that jointly learn the representation for an entity mention and its term mapping. The ideology behind employing a joint CNN model is that both the entity

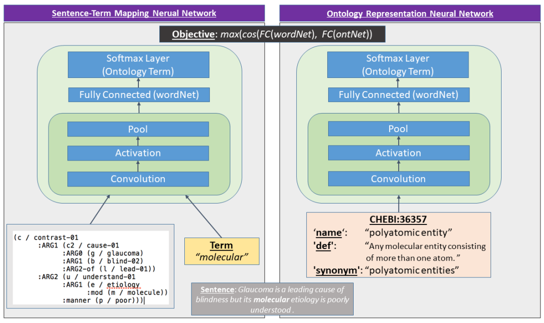

Figure 2: Diagram visualizing the ontology entity linking architecture.

mention and the ontology term mention have a rich semantic structure and neighborhood which can be leveraged. Joint Deep Learning architectures have previously been applied to Sentiment Analysis (Wang et al., 2016) and Event Extraction (Huang et al., 2017), we have explored the usability of such an architecture in the Entity Linking task. The neighborhood of an entity mention, the local context is captured by its adjacent terms in the AMR representation of the constitutent sentence, and ontology terms have representative fields like $\{def, comment, synonyms, name\}$; both of these neighborhoods described are used to train the joint CNN model(Sentence Representation Network, Ontology Representation Network) to learn the cosine similarity (Ye, 2011) between the representations of an entity mention and its mapped ontology term.

As mentioned earlier the joint CNN network architecture is trained to minimise the cosine distance (Wang et al., 2017) between the entity representation and the ontology term representation, as opposed to minimising the difference between the predicted ontology term mapping and the actual mapping. Cosine similarity has been used as a measure of similarity in many NLP tasks like Information Extraction and Document Summarisation, it is defined mathematically as the dot product between the vector norms:

$$cos\_sim(a, b) = \frac{a.b}{\sqrt{a^2+b^2}}$$

The input vectors fed into the Sentence Representation Network and the Ontology Representation Network were obtained by utilising a pre-trained word2Vec model (Duong et al., 2017) trained on a Biomedical corpus of $\approx 15GB$ of Pubmed articles. Out of vocabulary words not handled by the model were ignored, and randomness was introduced to encode sentences/ontology descriptions with zero words in vocabulary. The entity AMR representations and the ontology representations required a fair amount of pre-processing to be able to limit and extract meaningful neighborhood content for both inputs. A python Obonet reader (Obo) was used to extract the node content for each of the OBO foundry ontology terms mapped to in the CRAFT corpus, on close inspection of the field mappings we zeroed in on using $\{name, comment, def, synonyms\}$ as the subset of fields to represent an ontology term. The length of the Ontology term neighborhood was restricted to 50. To construct the entity representation the position of the entity term in the AMR representation of its constituent sentence was used as a pivot to extract surrounding terms($\approx$ 6, 3 ahead and 3 behind) that describe the entity. An AMR output reader (amr) was customized and leveraged for the purpose of extracting node information from the AMR representation of a sentence. For instance, for the entity mention mouse in:

Sentence: "The human and mouse genes display a high degree of synteny. Mcoln1 shows 91% amino acid and 86% nucleotide identity to MCOLN1."

AMR Representation:

```
(s / show-01
      :ARG0 (m / Mcoln1)
      :ARG1 (a / and
            :domain-of (i / identity
                  :mod (n / nucleotide))
            :op1 (a4 / and
                  :ARG1-of (d2 / display-01
                        :ARG0 (g / gene
                              :part-of (h2 / human)))
                  :op1 (m2 / mouse))
            :op2 (d / degree
                  :ARG1-of (h / high-02))
            :op3 (a2 / acid
```

Neighborhood words for "mouse"

$$\{mouse, and, display, gene, human, and, identity, nucleotide, degree\} \quad (1)$$

Ontology neighborhood for ground truth "NCBITaxon:10088"

$$\{Mus, mouse, genus, mice, mouse, Musmice, mouse, Mus\} \quad (2)$$

As seen above the process of generating representative definitions for entity mentions and mapped ontology terms was performed across mentions and annotations. These representations were then vectorised using the word2vec model described earlier, and batched and fed into the joint CNN network model. The Sentence representation network consists of three convolution layers (12 filters of 3×3, 12 filters of 3×3, 12 filters of 3×3) and a max pooling layer (2×2 with a stride of 2), and the Ontology representation network contains three convolution layers (50 filters of 7×5, 50 filters of 5×5 and 75 filters of 3×3) and 2 max pooling layers (2×2 with a stride of 2). A Softmax activation (Do and Ng, 2006) was used to predict the label mapping from the Sentence Representation network.

5 Results and Analysis

5.1 Results

The CRAFT corpus data was found to be inherently imbalanced, below is a visualization of the number of term mentions per ontology term.

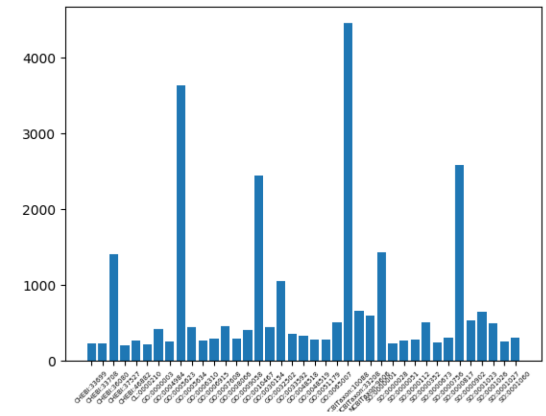

Figure 3: Bar chart representing the imbalance found in term distribution.

The accuracies and loss during the training phase of the joint CNN model were found to vary for a bit, and remain stagnant beyond a certain number of epochs. The plots for the same are depicted in figure-4 A training and testing split of 85 : 15 was used, that is $\approx$ 7500 entity mentions were held out.

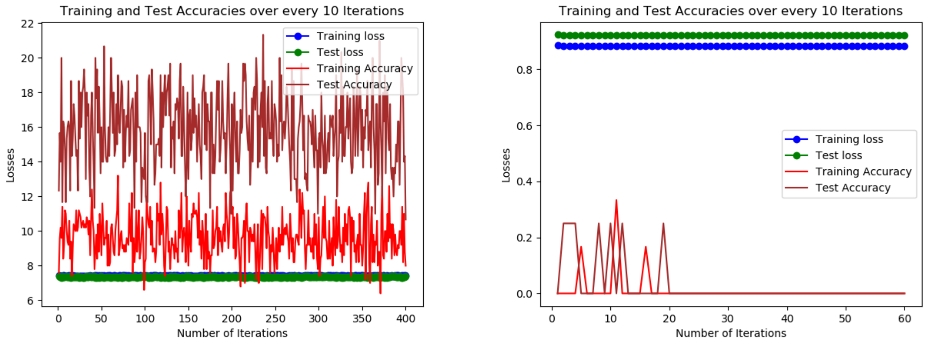

Figure 4: Plots for Conventional CNN and Joint CNN classification respectively

5.2 Evaluation

The joint CNN model was evaluated, predictions were made as shown in the table 1. All terms were predicted to the $NCIBTaxon$: 10088 term for mouse, as the data imbalance described above caused the Entity Linking architecture to be activated for a populous term mapping.

5.3 Challenges

Throughout implementation of the entity linking architecture and the associated pre-processing, we encountered numerous issues. One large problem that plagues a wide variety of NLP applications is the interoperability of its output. In the case of our study, AMR lacks a standard data format, which greatly increased the difficulty of working with AMR. The related programs programs we utilized such as JAMR and the AMR parser either severely lacked documentation, plagued with various runtime errors, or were not available for public usage.

Although a wide variety of biomedical ontologies exist, we found that there is a limited amount of ontology annotation training data available for the biomedical domain. Additionally, the CRAFT corpus we used contained some depreciated term mappings, demonstrating potential longevity issues with these ontologies. Some of terms mapped to in the CRAFT corpus were not located in the standard OBO foundry ontologies Smith et al. (2007), and these terms were ignored during the training phase.

As for the entity linking architecture, we encountered several issues during implementation of the training model, leading to inconsistent results. The data from CRAFT corpus was imbalanced in terms of the number of entity mentions per ontology term, this biased the network towards terms with multiple mappings. Additionally the CRAFT corpus linked a term to multiple ontologies, eg: cell was mapped to ontology terms both in the Cell Ontology(CL)(Bard et al., 2005) and Gene Ontology(GO) (Ashburner et al., 2000). So the above entity linking problem described in this paper should in actuality be treated as a multiset multiclass classification (Zhou et al., 2012) problem. It remains to be seen if an Multiple Instance Multi-Label(MIML) can provide better results.

5.4 Future Work

Based on the issues encountered during the construction of the entity linker, we plan on re-evaluating our results and revising the current model for the architecture. The purpose of these revisions would revolve around fixing the data imbalance issues testing new types of classification schemas. We also plan to revise the CNN architecture and settle on the best one that produces a good accuracy on a test set by using cross validation techniques. We plan to evaluate the usability of the disentangled representations (Jain et al., 2018) to enrich our vector embeddings for entity mentions. Since the entity linker portion will be exported as a semantic annotator tool for the process of automating the creation of Semantic Data Dictionaries (Rashid et al.), we also need to investigate methods to capture context for mentions in structured data.

In terms of future work for applying the higher-level architecture, the most probable future work involves improvement of the AMR parser and adaptation of the event extraction workflow. As described in Section 6.2, We encountered several

Sentence Neighborhood	Entity	Ground Truth	Prediciton
human and and display gene mouse identity nucleotide	human	NCBITaxon:9606	NCBITaxon:10088
mouse and display gene human and identity nucleotide degree	mouse	NCBITaxon:10088	NCBITaxon:10088
amino thing unique and long extraordinary amino more	amino	CHEBI:46882	NCBITaxon:10088
extraordinary carry-out thing unique and long amino amino more	amino acids	CHEBI:33708	NCBITaxon:10088
human develop human utility great birth potential investigate possible	humans	NCBITaxon:9606	NCBITaxon:10088
then blood and presence isoform red presence isoform both	red blood cells	CL:0000232	NCBITaxon:10088
protein experiment elegant express family encode they diverse gene	proteins	CHEBI:36080	NCBITaxon:10088
transcribe pair gene express apparent level thing express number equal	transcript	SO:0000673	NCBITaxon:10088
mucopolysaccharide show patient other or organomegaly resemble skeletal	excretion	GO:0007588	NCBITaxon:10088
skeletal show or resemble organomegaly mucopolysaccharide patient	skeletal changes	GO:0001501	NCBITaxon:10088
phagocytosi immune function and fundamental cell cell animal	apoptotic cells	CL:0000445	NCBITaxon:10088

Table 1: Sample entity prediction results.

issues with the JAMR parser. The most critical of these issues was related to the training of a new semantic parsing model for the biomedical domain. As a result, a default news parser was used to convert our sentences to AMR. To remedy this situation, we must continue working through the various problems associated with JAMR or attempt to utilize a new semantic AMR parser. Evaluation of the BioAMR model is necessary as part of this future work.

In order to demonstrate the capabilities of the entity linking architecture, we must modify our workflow to accommodate term input from event extraction output. This would involve taking the entities discovered in event extraction, passing the terms to our developed architecture and modifying the representation of the events to include this ontology grounding.

6 Conclusion

In this study, we have successfully implemented a data workflow to extract knowledge from biomedical literature, utilizing a novel entity linking architecture for grounding biomedical entities to corresponding ontology terms. Pre-processing of biomedical literature is necessary to convert the generated AMR and the chosen ontology terms to a vector representation as input for the entity linking architecture. Utilizing the CRAFT corpus (Bada et al., 2012), we were able to identify $\approx 2,000$ unique term mappings and $\approx 50,000$ annotations, serving as the foundation for the evaluation for our architecture. The evaluation of this entity linking architecture identifies potential for an entity to ontology term grounding system, with future iterations of the training model.

Acknowledgments

We would like to thank Dr. Heng Ji, Spencer Whitehead, Lifu Huang, and Xiaoman Pan from the BLENDER lab at Rensselaer Polytechnic Institute for their assistance and guidance in completion of this work.

References

Amr reader. https://github.com/panx27/amr-reader/tree/master/amrreader.

Obonet. https://github.com/dhimmel/obonet.

Michael Ashburner, Catherine A Ball, Judith A Blake, David Botstein, Heather Butler, J Michael Cherry, Allan P Davis, Kara Dolinski, Selina S Dwight, Janan T Eppig, et al. 2000. Gene ontology: tool for the unification of biology. *Nature genetics*, 25(1):25.

M. Bada, M. Eckert, D. Evans, K. Garcia, K. Shipley, D. Sitnikov, W. A. Baumgartner, K. B. Cohen, K. Verspoor, J. A. Blake, and L. E. Hunter. 2012. Concept annotation in the CRAFT corpus. *BMC Bioinformatics*, 13:161.

Laura Banarescu, Claire Bonial, Shu Cai, Madalina Georgescu, Kira Griffitt, Ulf Hermjakob, Kevin Knight, Philipp Koehn, Martha Palmer, and Nathan Schneider. 2013. Abstract meaning representation for sembanking. In *Proceedings of the 7th Linguistic Annotation Workshop and Interoperability with Discourse*, pages 178–186, Sofia, Bulgaria. Association for Computational Linguistics.

Jonathan Bard, Seung Y Rhee, and Michael Ashburner. 2005. An ontology for cell types. *Genome biology*, 6(2):R21.

Chuong B Do and Andrew Y Ng. 2006. Transfer learning for text classification. In *Advances in Neural Information Processing Systems*, pages 299–306.

Dat Duong, Eleazar Eskin, and Jessica Li. 2017. A novel word2vec based tool to estimate semantic similarity of genes by using gene ontology terms. *bioRxiv*, page 103648.

Jeffrey Flanigan, Sam Thomson, Jaime Carbonell, Chris Dyer, and Noah A. Smith. 2014. A discriminative graph-based parser for the abstract meaning representation. In *Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 1426–1436, Baltimore, Maryland. Association for Computational Linguistics.

Matthew Francis-Landau, Greg Durrett, and Dan Klein. 2016. Capturing semantic similarity for entity linking with convolutional neural networks. *arXiv preprint arXiv:1604.00734*.

Lifu Huang, Taylor Cassidy, Xiaocheng Feng, Heng Ji, Clare R. Voss, Jiawei Han, and Avirup Sil. 2016. Liberal event extraction and event schema induction. In *ACL*.

Lifu Huang, Heng Ji, Kyunghyun Cho, and Clare R Voss. 2017. Zero-shot transfer learning for event extraction. *arXiv preprint arXiv:1707.01066*.

Sarthak Jain, Edward Banner, Jan-Willem van de Meent, Iain J Marshall, and Byron C Wallace. 2018. Learning disentangled representations of texts with application to biomedical abstracts. *arXiv preprint arXiv:1804.07212*.

Daniel Jurafsky and James H. Martin. 2008. *Speech and Language Processing (2nd Edition)*. Prentice-Hall, Inc., Upper Saddle River, NJ, USA.

Yann LeCun, Léon Bottou, Yoshua Bengio, and Patrick Haffner. 1998. Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, 86(11):2278–2324.

Tomas Mikolov, Ilya Sutskever, Kai Chen, Greg Corrado, and Jeffrey Dean. 2013. Distributed representations of words and phrases and their compositionality. *CoRR*, abs/1310.4546.

Sabbir M Rashid, Katherine Chastain, Jeanette A Stingone, Deborah L McGuinness, and James P McCusker. The semantic data dictionary approach to data annotation & integration.

Barry Smith, Michael Ashburner, Cornelius Rosse, Jonathan Bard, William Bug, Werner Ceusters, Louis J Goldberg, Karen Eilbeck, Amelia Ireland, Christopher J Mungall, et al. 2007. The obo foundry: coordinated evolution of ontologies to support biomedical data integration. *Nature biotechnology*, 25(11):1251.

Kai Sheng Tai, Richard Socher, and Christopher D Manning. 2015. Improved semantic representations from tree-structured long short-term memory networks. *arXiv preprint arXiv:1503.00075*.

Jiabao Wang, Yang Li, Zhuang Miao, Yulong Xu, and Gang Tao. 2017. Learning deep discriminative features based on cosine loss function. *Electronics Letters*, 53(14):918–920.

Xingyou Wang, Weijie Jiang, and Zhiyong Luo. 2016. Combination of convolutional and recurrent neural network for sentiment analysis of short texts. In *Proceedings of COLING 2016, the 26th International Conference on Computational Linguistics: Technical Papers*, pages 2428–2437.

Jun Ye. 2011. Cosine similarity measures for intuitionistic fuzzy sets and their applications. *Mathematical and Computer Modelling*, 53(1-2):91–97.

Tom Young, Devamanyu Hazarika, Soujanya Poria, and Erik Cambria. 2017. Recent trends in deep learning based natural language processing. *arXiv preprint arXiv:1708.02709*.

Jin Guang Zheng, Daniel Howsmon, Boliang Zhang, Juergen Hahn, Deborah McGuinness, James Hendler, and Heng Ji. 2015. Entity linking for biomedical literature.

Zhi-Hua Zhou, Min-Ling Zhang, Sheng-Jun Huang, and Yu-Feng Li. 2012. Multi-instance multi-label learning. *Artificial Intelligence*, 176(1):2291–2320.