KNOWLEDGE GRAPH-BASED METAPHOR REPRESENTATION

B. Tech. Project Report

Submitted by

Tushar Godbole 111903115

Under the guidance of

Prof. Dr. Y. V. Haribhakta

College of Engineering, Pune

DEPARTMENT OF COMPUTER ENGINEERING AND INFORMATION TECHNOLOGY, COLLEGE OF ENGINEERING, PUNE-5

May 2023

DEPARTMENT OF COMPUTER ENGINEERING

AND

INFORMATION TECHNOLOGY,

COLLEGE OF ENGINEERING, PUNE

CERTIFICATE

Certified that this project, titled "KNOWLEDGE GRAPH-BASED METAPHOR REPRESENTATION" has been successfully completed by

Tushar Godbole 111903115

and is approved for the partial fulfillment of the requirements for the degree of "B.Tech. Computer Engineering".

Dr. Y. V. Haribhakta
Project Guide
Department of Computer Engineering
and Information Technology,
College of Engineering Pune,
Shivajinagar, Pune - 5.

Dr. P. K. Deshmukh
Head
Department of Computer Engineering
and Information Technology,
College of Engineering Pune,
Shivajinagar, Pune - 5.

Turnitin Originality Report

Processed on: 15-Apr-2023 9:02 AM IST
ID: 2062498760
Word Count: 4365
Submitted: 3

Knowledge Graph-Based Metaphor Representation By Tushar Godbole

Similarity Index	Similarity by Source
19%	Internet Sources: 16% Publications: 13% Student Papers: 0%

2% match (Internet from 14-Sep-2020)
https://aniketbhadane.github.io/AB-dissertation-en.pdf

1% match (Internet from 13-Sep-2009)
http://www.aclweb.org/anthology-new/W/W07/W07-01.pdf

1% match (Internet from 27-Sep-2021)
https://readthedocs.org/projects/ampligraph/downloads/pdf/1.1.0/

1% match (Internet from 19-Sep-2022)
https://aclanthology.org/W07-0103.pdf

1% match (Internet from 05-Dec-2022)
https://academic.oup.com/dsh/article/36/3/698/6098012

1% match (Internet from 11-Jan-2023)
https://www.politesi.polimi.it/bitstream/10589/166574/1/thesis.pdf

1% match (Tong Shen, Fu Zhang, Jingwei Cheng. "A comprehensive overview of knowledge graph completion", Knowledge-Based Systems, 2022)
Tong Shen, Fu Zhang, Jingwei Cheng. "A comprehensive overview of knowledge graph completion", Knowledge-Based Systems, 2022

< 1% match (Internet from 03-Apr-2017)
http://www.aclweb.org/anthology/E/E17/E17-2.pdf

< 1% match (Internet from 22-Sep-2022)
https://aclanthology.org/2020.figlang-1.pdf

< 1% match (Internet from 27-Sep-2022)
https://direct.mit.edu/tacl/article/doi/10.1162/tacl_a_00471/110550/TopiOCQA-Open-domain-Conversational-Question

< 1% match (Internet from 26-Apr-2021)
https://direct.mit.edu/coli/article/39/2/301/1426/Statistical-Metaphor-Processing

< 1% match (Internet from 07-Apr-2023)
https://www.arxiv-vanity.com/papers/1711.05435/

< 1% match (Internet from 07-Jan-2022)
http://www.arxiv-vanity.com/papers/1902.10197/

< 1% match ("Web Information Systems Engineering – WISE 2019", Springer Science and Business Media LLC, 2019)
"Web Information Systems Engineering – WISE 2019", Springer Science and Business Media LLC, 2019

< 1% match (Lei Chen, Dandan Wang, Lin Shi, Qing Wang. "A Self-enhanced Automatic Traceability Link Recovery via Structure Knowledge Mining for Small-scale Labeled Data", 2021 IEEE 45th Annual Computers, Software, and Applications Conference (COMPSAC), 2021)
Lei Chen, Dandan Wang, Lin Shi, Qing Wang. "A Self-enhanced Automatic Traceability Link Recovery via Structure Knowledge Mining for Small-scale Labeled Data", 2021 IEEE 45th Annual Computers, Software, and Applications Conference (COMPSAC), 2021

15/4/23

< 1% match (Internet from 17-Jan-2022)
https://buildintuit.com/post-page-2/

< 1% match (Wei Song, Jingjin Guo, Ruiji Fu, T. Liu, Lizhen Liu. "A Knowledge Graph Embedding Approach for Metaphor Processing", IEEE/ACM Transactions on Audio, Speech, and Language Processing, 2020)
Wei Song, Jingjin Guo, Ruiji Fu, T. Liu, Lizhen Liu. "A Knowledge Graph Embedding Approach for Metaphor Processing", IEEE/ACM Transactions on Audio, Speech, and Language Processing, 2020

< 1% match (Internet from 17-Sep-2022)
https://www.grin.com/document/1138163

Abstract

A metaphor is a figure of speech which is very commonly used in natural language. Metaphor processing involves identifying and interpreting them and presents challenges to automatic natural language processing. This project is a study of the knowledge graph embedding approach to noun-pair metaphor representation. A knowledge graph is a directed graph containing entities and the relations between them. Knowledge graph embedding involves representing the entities and relations of a knowledge graph in the form of low-dimensional vectors to simplify processing while preserving the structure and meaning of the knowledge graph.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

Detection and interpretation of figurative language is one of the major challenges in Natural Language Processing (NLP). A metaphor is a figure of speech in which the properties of one entity are described implicitly by referring to another [3]. Since metaphors are frequently used, their interpretation is important for better results in NLP [4]. To simplify interpretation and processing of data, Peng et al. [5] have used knowledge graphs to represent metaphors. This project closely follows their work.

1.1 Knowledge Graph

A knowledge graph is a directed graph that models entities and the relations between them. Every entity in is represented by a node in the knowledge graph and the relation between two entities is represented by the edge directed from the node corresponding to the head entity to the node corresponding to the tail entity [5]. A triple that consists of $(head, relation, tail)$ can be used to model a pair of entities and the relation between them. For example, in the Figure 1.1, the fact that Mike is a person is shown by the triple (*Mike, isA, person*).

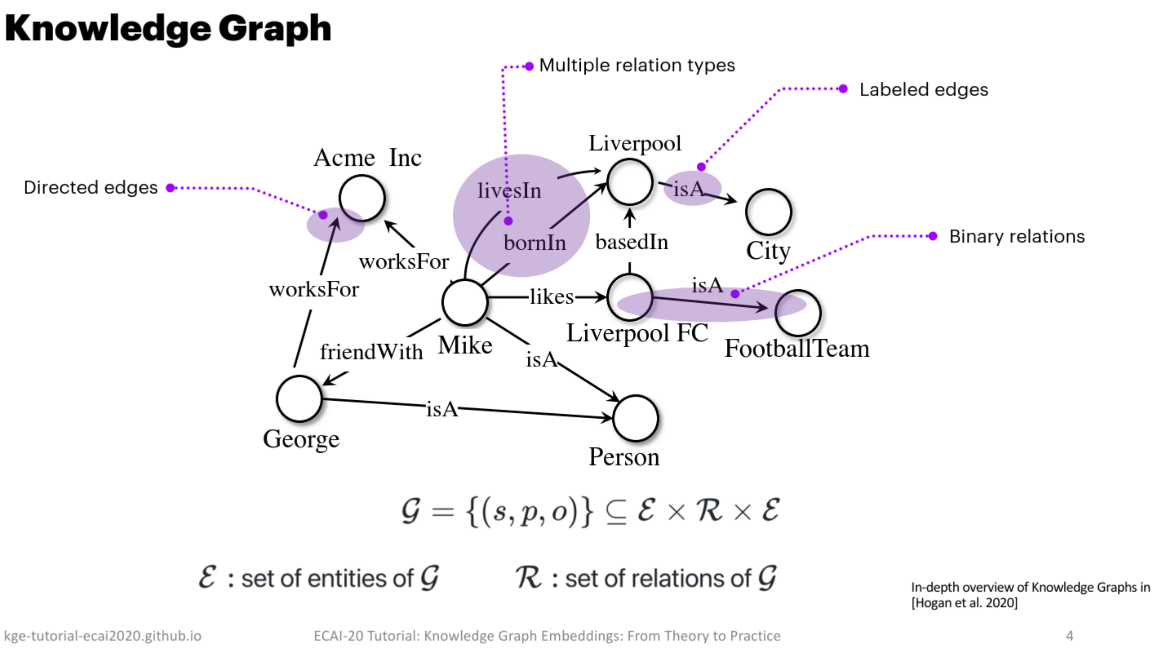

Figure 1.1: Example of a knowledge graph [1]

1.2 Knowledge graph embedding

Using knowledge graph embedding methods for metaphor graphs allows generating representations of metaphor relations and can make metaphor identification and interpretation more efficient and easier [5].

Knowledge graph embedding involves automatic, supervised learning of embeddings, which are representations of the entities and relations of a knowledge graph, using continuous, low-dimensional vectors. This preserves the original structure and meaning of the graph while making it easier to manipulate. [3]. Knowledge graph embedding involves automatic, supervised learning of embeddings [1]. Machine learning tasks like knowledge graph completion or the evaluation of the truth of triples make use of knowledge graph embedding. Real-world applications include recommender systems and drug repurposing [2]. Some examples of knowledge graph embedding models are TransE [6], TransH [7], TransR [8], TransD [9], TransA [10], DistMult

[11], ComplEx [12], HolE [13], SimplE [14] and RotatE [15].

These models use similar methods to learn the meanings of facts[2] but differ in their scoring function and loss function used. As mentioned in the Ampligraph documentation [16], "The training phase always consists of minimizing a loss function. The goal of the optimization procedure is learning optimal embeddings, such that the scoring function is able to assign high scores to positive statements and low scores to statements unlikely to be true." In this project, the TransE model is used, which was one of the first embedding models and is still in common use[1].

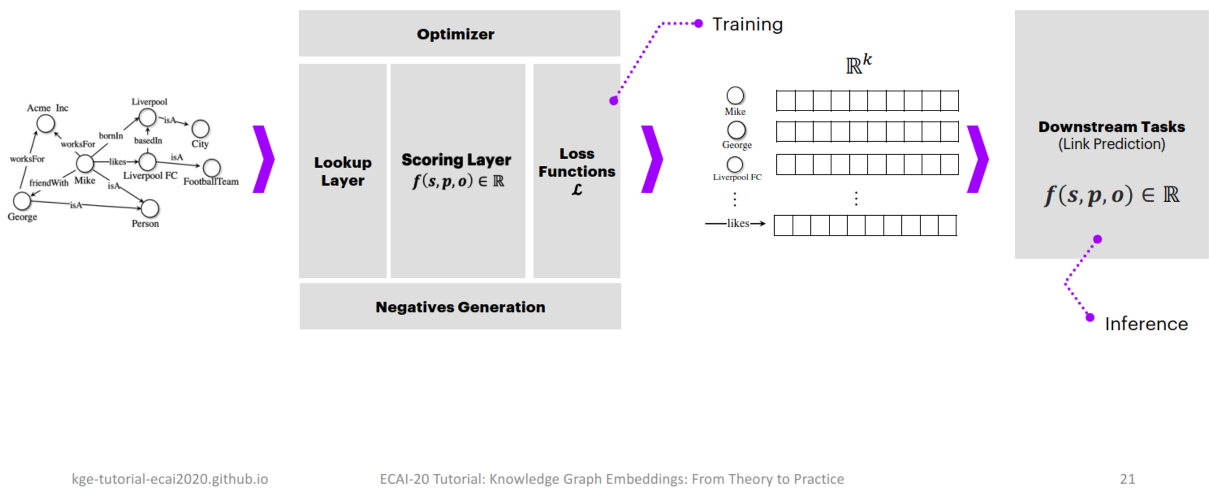

Figure 1.2: Embedding process [1]

1.3 TransE

TransE is a model which attempts to force the embeddings of each triple to fit the vector sum equation:

$$\mathbf{h} + \mathbf{r} \simeq \mathbf{t}$$

for a triple $(h, r, t) \in \mathcal{K}$, where h and t are the head and tail entities, respectively, and r is the relation directed from the head to the tail. $\mathcal{K}$ is the knowledge graph. $\mathbf{h}, \mathbf{r}$ and $\mathbf{t}$ are the embeddings of h, r and t, respectively. In other words, the distance between the embedding of the head entity plus the embedding of the relation and the embedding of the tail entity is minimised [1, 3].

The score function for TransE is:

$$f_{TransE} = -||\mathbf{h} + \mathbf{r} - \mathbf{t}||_n$$

where n is 1 or 2 ($L1$ or $L2$ norm, respectively).

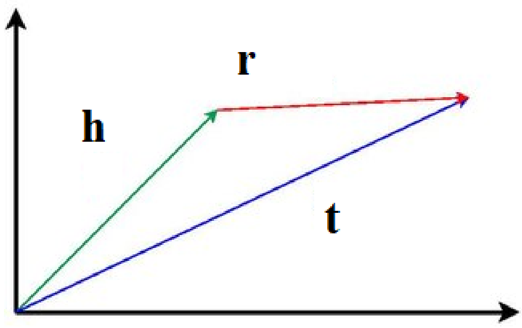

Figure 1.3: TransE embedding model [2]

The loss function used in this project is self-adversarial loss:

$$\mathcal{L} = -\log \sigma(\gamma + f(t^+;\Theta)) - \sum_{t\in\mathcal{G}}^{N} p(t^-;\Theta) \log \sigma(-f(t^-;\Theta) - \gamma)$$

where σ is the sigmoid function, γ is a fixed margin, $f(t^-;\Theta)$ is the score assigned to a synthetic negative, $f(t^+;\Theta)$ is the score assigned to a true triple, $\mathcal{G}$ is the knowledge graph and $p(t^-;\Theta)$ is the weight for the negative sample t^-. [1, 15]

Chapter 2

Literature Review

- **"Similarity-based Word Sense Disambiguation (Karov and Edelman, 1998)"** [17]

  This paper describes a technique to automatically disambiguate the meaning of a word by using a text corpus and a dictionary that can be read by a machine. The approach is based on word similarity and context similarity. The system learns representative usages for each of the meanings of the polysemantic word from the machine-readable dictionary. A new appearance of such a word is assigned the meaning related to the representative usage that is most similar to its context.

- **"Active Learning for the Identification of Nonliteral Language (Birke and Sarkar, 2007)"** [18]

  This paper describes a largely automated active learning approach for creating an annotated corpus of literal and nonliteral usages of verbs. Recognition of nonliteral language is reduced to a word-sense disambiguation problem, for which the algorithm proposed by Karov and Edelman [17] is used to calculate the similarities between the target sentence and the feedback sets. The target sentences are attracted to the feedback set containing the most similar sentence. Active learning is used to improve the system's accuracy by combining the system's predictions and the annotations of a human expert.

- **"Models of Metaphor in NLP (Shutova, 2010)"** [19]
  This paper is a review of the existing computational models of metaphor. Metaphor always involves two concepts or conceptual domains: the target and the source. This paper reviews the works of Wilks (1978), Faas (1991), Goatly (1997), Krishnakumaran and Zhu (2007) and the Pragglejaz procedure, among others.

- **"A Survey of Figurative Language and its Computational Detection in Online Social Networks (Abulaish et al., 2020)"** [20]
  This paper deals with similes, sarcasm, irony, satire, humour, metaphor and hyperbole. It describes a brief history of figurative language studies. It gives an overview of feature extraction techniques and detection approaches for each of the mentioned kinds of figurative language.

- **"Automatic Metaphor Interpretation as a Paraphrasing Task (Shutova, 2010)"** [4]
  This paper presents a method to interpret metaphors such that metaphorical expressions are paraphrased literally. A corpus study was conducted where metaphorical expressions were manually annotated and a high frequency of the use of metaphors was found. Automatic processing of metaphors has two subtasks: metaphor recognition and metaphor interpretation. The system uses automatically induced selectional preferences and has a high accuracy of paraphrasing. The system produces a list of all possible paraphrases for a metaphorical expression, ranks them according to their likelihood, uses selectional preference violation

to output the literal paraphrases, and uses WordNet to disambiguate their meanings.

- **"Hunting Elusive Metaphors Using Lexical Resources (KrishnaKumaran and Zhu, 2007)"** [21]

  This paper describes an algorithm for classifying sentences into metaphoric and normal usages. The algorithm uses WordNet and bigram counts and doesn't require training. Novel metaphors 'die' and become automatic and idiomatic with repeated use. This paper is concerned with identifying live metaphors only. It focuses on a subset of metaphoric usages involving nouns in a sentence. It describes the method to identify three types of metaphor. The paper uses the hyponymy relation in WordNet and word co-occurrence information for detecting metaphoric uses in subject-object, verb-noun and adjective-noun relationships.

- **"Knowledge graph-based metaphor representation for literature understanding (Peng et al., 2021)"** [5]

  Here, a method is presented in which metaphor embedding based on knowledge graph is used to build metaphorical relations networks of nouns. The metaphorical meanings of two nouns are interpreted by the metaphor relation between them, which differs from their general relation. It is assumed that nouns having similar concepts can be described by the same adjectives. Finding common adjectives is the basis of interpreting noun-form metaphors. The common adjectives of two nouns are the metaphor relations between them. A metaphor triple is formed from the two nouns and their metaphor relation. Metaphorical triples are represented using knowledge graph. Graph embedding methods are used to project the metaphorical triples into embedding space, making it easier to understand implicit meanings of the metaphors and to predict metaphor relations between nouns.

- **"A Knowledge Graph Embedding Approach for Metaphor Processing (Song et al., 2020)"** [3]

  A metaphor is a relation between a source entity and a target entity. Any metaphor is represented as a triple of the form (*target, attribute, source*). Metaphor triples can be modeled like fact triples in a knowledge graph. Simile recognition is used to make a metaphor knowledge graph. The metaphor knowledge graph is embedded with concept-attribute collocations to learn better representations of concepts, attributes and metaphors.

Chapter 3

Research Gaps and Problem Statement

3.1 Research Gaps

Metaphor detection and processing remains an active area of research in NLP. Identifying and interpreting metaphors is a prominent area of research. There is little research on establishing metaphorical relationships between metaphorical words [5]. There are no known libraries specifically for metaphor detection and processing.

3.2 Problem Statement

Study the knowledge graph embedding approach to metaphor representation.

3.3 Objectives

1. To study the effectiveness of knowledge graph as a form of metaphor representation

2. To study knowledge graph embedding techniques

3. To predict the metaphor relations between noun pairs using knowledge graph embedding techniques

Chapter 4

Proposed Methodology/ Solution

The method followed in this project starting from the source datasets to the evaluation of knowledge graph embeddings is shown in figure 4.1, following which each step is explained in detail.

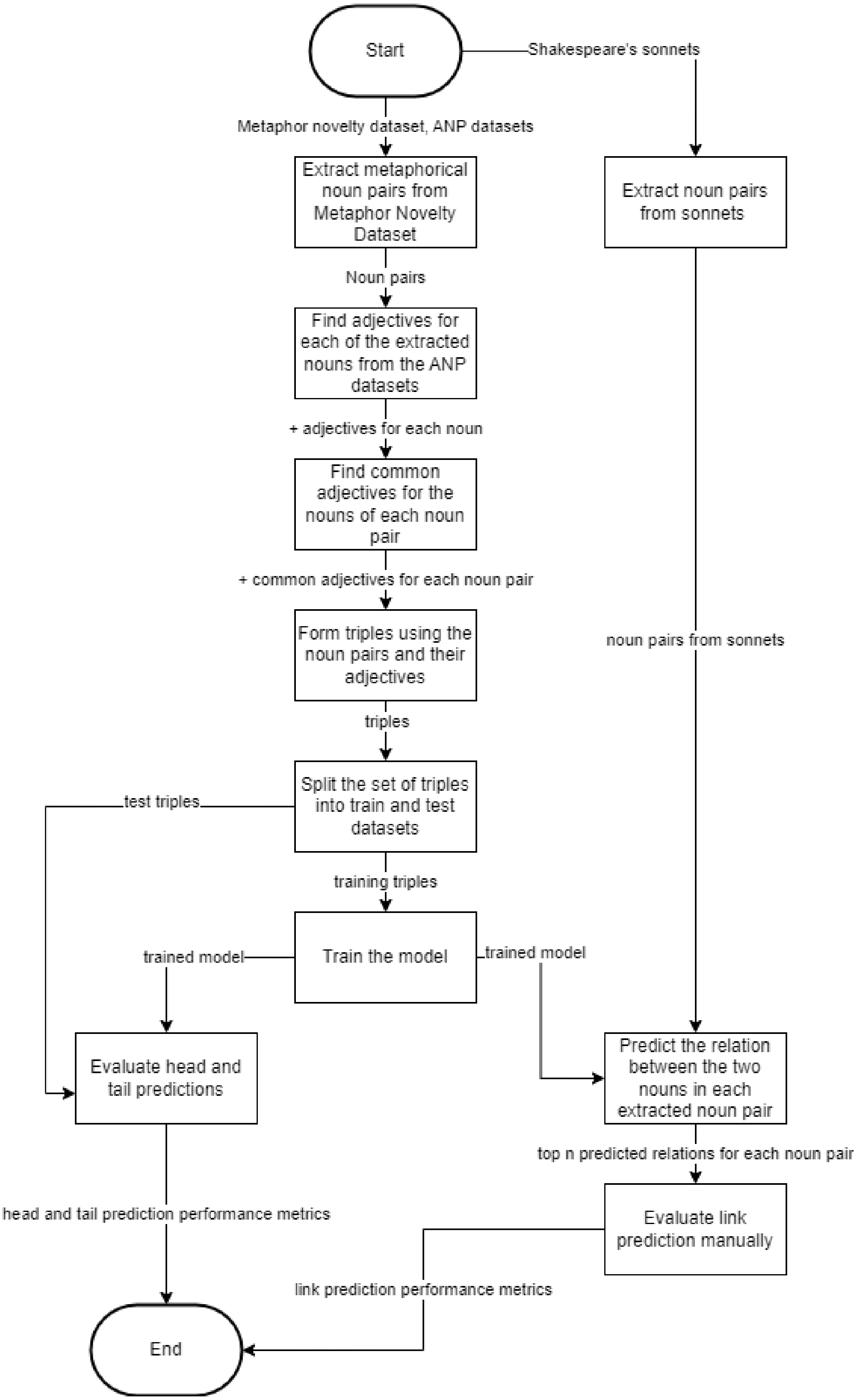

Figure 4.1: End-to-end process

1. **Get noun pairs from Metaphor Novelty Dataset**: Metaphorical noun pairs are extracted from the Metaphor Novelty Dataset [22]. The part-of-speech (POS) tags of the words as in the Brown corpus of the Python NLTK library were used for POS tagging, This was followed by lemmatizing the nouns in the noun pair and then checking that the metaphor score was greater than 0. The head and tail words are determined from each pair by assuming the word with the smaller index in the sentence to be the head and the word with the greater index to be the tail. Out of the 18,452 word pairs in the Metaphor Novelty Dataset, 6,135 are found to be metaphorical noun pairs.
   E.g., From the Metaphor Novelty Dataset, the ID column entry `7891__hair_31__flowers_29`, the extracted noun pair is `(flower, hair)`.

2. **Get adjective-noun pairs (ANPs) from ANP datasets**: Adjective-noun pairs are taken from the Visual Sentiment Ontology dataset [23], which contains 3,244 ANPs, and the dataset by Guitérrez et al., 2016 [24], which contains 8,592 ANPs. Adjectives were found for 1,466 of the extracted nouns.
   E.g., For the noun pair (`flower`, `hair`), 36 adjectives were found for `flower` and 37 adjectives were found for `hair`.

3. **Get common adjectives from the ANPs for noun pairs in the Metaphor Novelty Dataset**: Common adjectives are found from the ANP datasets for 1,063 of the noun pairs extracted from the Metaphor Novelty Dataset.
   E.g., `flower` and `hair` have 9 common adjectives: `pretty, fantastic, awesome, bright, soft, golden, dry, amazing, wild`

4. **Form triples**: A dataset of 1,762 triples of the form (*head, relation, tail*) is created, where *head* and *tail* are nouns of a noun pair extracted from the Metaphor Novelty Dataset and *relation* is their common adjective found in the ANP datasets. In case of multiple common adjectives, one triple is formed for each.

   E.g., 9 triples are formed using the words `flower`, `hair` and their common adjectives. They include (`flower, pretty, hair`), ..., (`flower, wild, hair`).

5. **Create training and test triples datasets**: The set of triples created is divided into datasets for training and testing in the ratio 7:3 such that all entities and relations present in the test data are also present in the training data. The training set had 1,234 triples and the test set had the remaining 528.

6. **Create the model with the required hyperparameters**: A model is trained with the training dataset for knowledge graph embedding by specifying the required hyperparameters. Here, the model used is the TransE model.

7. **Evaluate head and tail prediction**: The model is evaluated for head and tail prediction using the test dataset created above. The head (in case of head prediction) and tail (in case of tail prediction) is corrupted for each triple of the test data and the correct triple is ranked against the corrupted ones. The mean rank, mean reciprocal rank and hits@k are calculated for head and tail prediction.

8. **Get metaphorical noun pairs from another dataset**: A dataset containing metaphorically used nouns is chosen to test link prediction. In this case, all 154 of Shakespeare's sonnets are used. Noun pairs are

extracted from the sonnet lines. If a sonnet line contains any two nouns from the training dataset, those two nouns form a noun pair. 253 noun pairs are found.

9. **Link prediction**: The head noun and tail noun are taken from each sentence chosen for link prediction and input to a link prediction function which predicts the top n relations (adjectives) between the two entities (nouns), where n is a parameter.

10. **Evaluate link predictions manually**: The link predictions obtained in the previous step are then manually evaluated.

Note: Though this project tries to follow the work by Peng et al. [5], some of the steps could not be exactly repeated.

1. Peng et al. extracted 434 metaphorical noun pairs from the Metaphor Novelty Dataset, but 6,135 metaphorical noun pairs were extracted in this project. This is probably due to differences in the criteria for determining whether a noun pair is metaphorical.

2. The Multilingual Visual Sentiment Ontology dataset used in the paper is not available anymore. Instead, this project uses the ANP dataset by Guitérrez et al. [24].

Chapter 5

Experimental Setup

The language used for this project is Python. Google Colaboratory is used for writing and executing the Python notebook code. The NLTK library is used for POS tagging. The datasets are processed with Pandas. Ampligraph 2.0.0 is used for knowledge graph embedding. Gephi (`https://gephi.org/`) is used to visualise the knowledge graph (figure 6.1). Draw.io (`https://app.diagrams.net/`) was used to draw figure 4.1.

Tool	Use	Link
Google Colaboratory	Code development	`https://colab.research.google.com/`
Gephi	Graph visualisation	`https://gephi.org/`
Draw.io	Process flow diagram	`https://app.diagrams.net/`

Table 5.1: Tools

Library	Use	Link
Ampligraph 2.0.0	Knowledge graph embedding	`https://docs.ampligraph.org/en/latest/`
Pandas	Dataset processing	`https://pandas.pydata.org/`
NLTK	POS tagging	`https://www.nltk.org/`

Table 5.2: Libraries

Dataset	Contains
Metaphor Novelty Dataset[22][25]	Word pairs
Visual Sentiment Ontology [26][27]	ANPs
ANP dataset[24][28]	ANPs
Shakespeare's sonnets[29]	Sonnets

Table 5.3: Datasets

Chapter 6

Results and Discussion

1. **Hyperparameters**: The model is trained using the TransE scoring function with 300 dimensions and using 5 negatives during training per triple. The Adam optimizer and self-adversarial loss function are used. A batch size of 16 is used for training, with 10 epochs.

2. **Head and tail prediction**: The model is evaluated for head and tail prediction using the test set of 528 triples.

   For head prediction:

   - Mean rank = 109.6894
   - Mean reciprocal rank = 0.2142
   - Hits@1 = 0.19%
   - Hits@3 = 39.20%
   - Hits@5 = 47.16%
   - Hits@10 = 55.87%

   For tail prediction:

   - Mean rank = 107.6572
   - Mean reciprocal rank = 0.2194
   - Hits@1 = 0.19%

- Hits@3 = 39.20%
- Hits@5 = 47.16%
- Hits@10 = 58.52%

3. **Link prediction**: The top 10 predictions (of a total of 79 relations) for the link between the two nouns in every pair are found. The link prediction is manually evaluated. Only 23 of the noun pairs were found to be sufficiently metaphorical in the context of the sonnet line that they were in and have relevant adjectives in the set of all adjectives in the triples. They are shown in Figure 6.2.
   The results for the evaluation of link prediction are given below:
   - Mean rank = 3.9130
   - Mean reciprocal rank = 0.4705
   - Hits@1 = 30.43%
   - Hits@3 = 47.83%
   - Hits@5 = 73.91%

   Some of the predicted triples are shown in Figure 6.1.

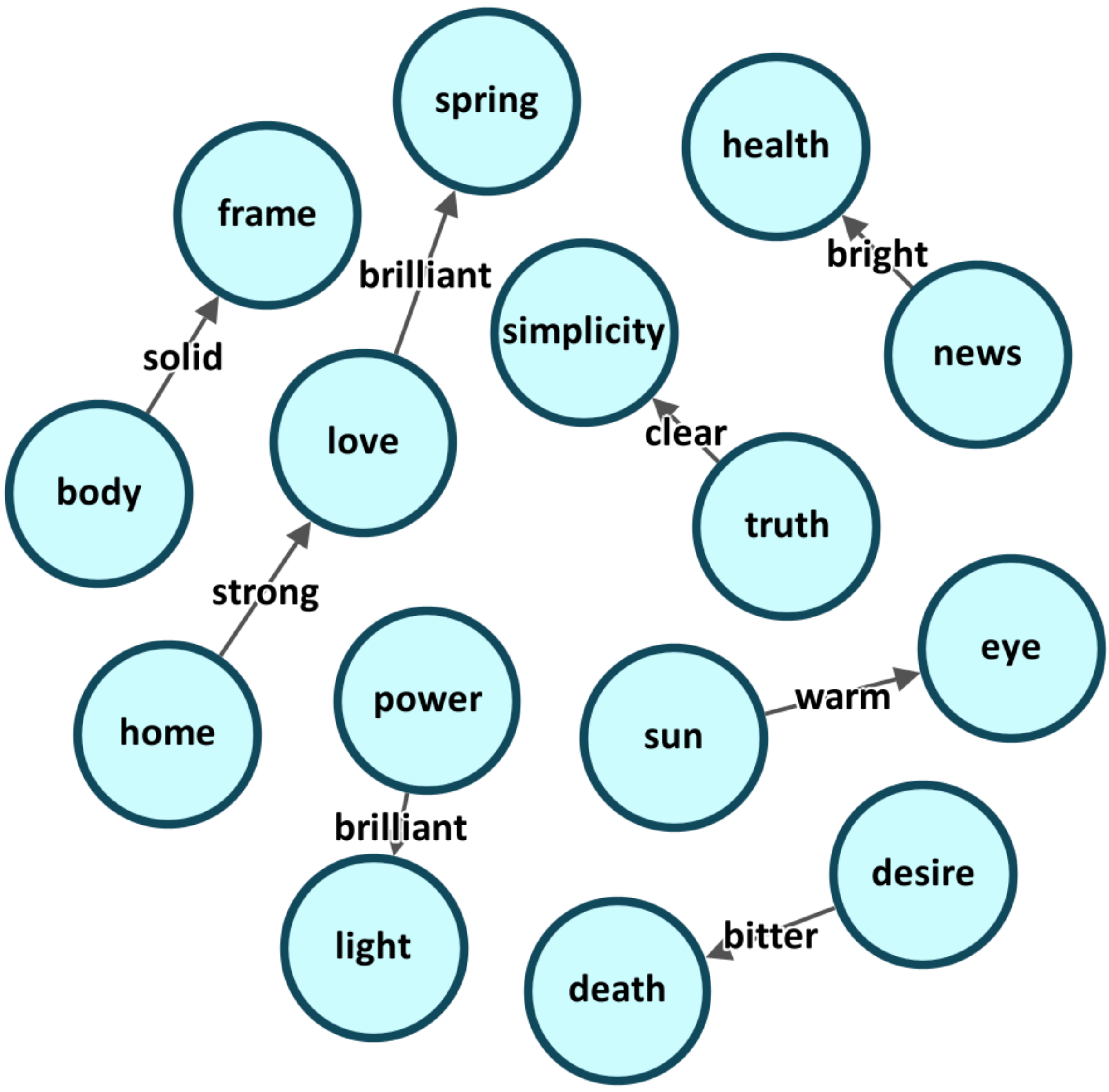

Figure 6.1: Links predicted for sample noun pairs from sonnets

LineText	Head_tail	Top 10 predicted relations	Rank
my body is the frame wherein tis held	('body', 'frame')	[**'solid',** 'soft', 'weak', 'brilliant', 'strong', 'rough', 'heavy', 'murky', 'warm', 'cold']	1
and scarcely greet me with that sun thine eye	('sun', 'eye')	['strong', 'bitter', 'solid', 'shallow', 'weak', 'sweet', 'heavy', 'soft', **'warm',** 'rough']	9
desire is death which physic did except	('desire', 'death')	['sweet', 'icy', 'murky', **'bitter',** 'warm', 'heated', 'brilliant', 'deep', 'smooth', 'rough']	4
darkening thy power to lend base subjects light	('power', 'light')	['murky', 'clear', 'rough', 'cold', **'brilliant',** 'clean', 'warm', 'bitter', 'strong', 'weak']	5
our love was new and then but in the spring	('love', 'spring')	['bitter', 'cold', 'bright', **'brilliant',** 'warm', 'heavy', 'shallow', 'weak', 'murky', 'rough']	4
that is my home of love if i have rang d	('home', 'love')	['clear', 'deep', **'strong',** 'weak', 'heavy', 'solid', 'rough', 'warm', 'clean', 'smooth']	3
no news but health from their physicians know	('news', 'health')	[**'bright',** 'rough', 'cold', 'strong', 'soft', 'clear', 'dim', 'smooth', 'brilliant', 'murky']	1
and barren rage of death s eternal cold	('death', 'cold')	['solid', 'heavy', 'clear', **'bitter',** 'soft', 'heated', 'clean', 'weak', 'shallow', 'strong']	4
but as the marigold at the sun s eye	('sun', 'eye')	[**'strong',** 'bitter', 'solid', 'shallow', 'weak', 'sweet', 'heavy', 'soft', 'warm', 'rough']	1
for precious friends hid in death s dateless night	('death', 'night')	['weak', 'solid', 'strong', 'clean', **'heavy',** 'murky', 'soft', 'clear', 'bright', 'sweet']	5
when in dead night thy fair imperfect shade	('night', 'shade')	['shallow', **'dim',** 'strong', 'rough', 'bitter', 'clear', 'cold', 'fat', 'deep', 'weak']	2
so either by thy picture or my love	('picture', 'love')	[**'sweet',** 'rough', 'strong', 'smooth', 'clear', 'murky', 'bright', 'deep', 'solid', 'heavy']	1
thy edge should blunter be than appetite	('edge', 'appetite')	['cold', 'bright', 'solid', 'weak', 'murky', 'strong', 'rough', **'shallow',** 'mad', 'warm']	8
the spirit of love with a perpetual dulness	('spirit', 'love')	[**'deep',** 'strong', 'clean', 'rough', 'smooth', 'soft', 'clear', 'weak', 'murky', 'icy']	1
with time s injurious hand crush d and o erworn	('time', 'hand')	[**'strong',** 'bright', 'solid', 'clean', 'brilliant', 'heavy', 'clear', 'icy', 'bitter', 'murky']	1
whose action is no stronger than a flower	('action', 'flower')	['bitter', 'deep', 'strong', 'murky', 'rough', 'brilliant', 'bright', 'heated', **'weak',** 'soft']	9
some in their wealth some in their body s force	('body', 'force')	['soft', 'rough', 'clear', 'shallow', 'warm', **'strong',** 'heated', 'solid', 'bitter', 'heavy']	6
the hardest knife ill us d doth lose his edge	('knife', 'edge')	['clear', 'heated', **'strong',** 'bitter', 'icy', 'rough', 'brilliant', 'shallow', 'clean', 'deep']	3
some say thy grace is youth and gentle sport	('youth', 'sport')	['smooth', 'soft', 'clean', 'rough', 'clear', **'brilliant',** 'strong', 'weak', 'bitter', 'shallow']	6
and gives thy pen both skill and argument	('skill', 'argument	['soft', **'brilliant',** 'bright', 'rough', 'weak', 'cold', 'strong', 'clean', 'solid', 'clear']	2
root pity in thy heart that when it grows	('root', 'heart')	[**'deep',** 'clean', 'rough', 'strong', 'clear', 'brilliant', 'weak', 'soft', 'murky', 'solid']	1
there is such strength and warrantise of skill	('strength', 'skill')	['sweet', 'bitter', 'rough', 'shallow', 'solid', 'heated', 'weak', 'strong', **'heavy',** 'brilliant']	8
which from love s fire took heat perpetual	('love', 'fire')	['bitter', 'sweet', 'warm', 'cold', **'brilliant',** 'bright', 'solid', 'shallow', 'weak', 'strong']	5

Figure 6.2: Predicted relations. The bold words are the correct adjectives.

Chapter 7

Conclusion

Knowledge graphs can be used as a form of metaphor representation. Knowledge graph embedding techniques can be used to predict metaphor relations between noun pairs, which helps in understanding the meaning of metaphors. This project has followed the work of Peng et al. [5] but differs in the results because of different datasets.

The datasets used were small in size, which affected the training of the model. Datasets of better quality, particularly having a larger number of metaphorical noun pairs and larger number of adjective-noun pairs with a greater number of adjectives would have probably produced better results. The predicted relations are often irrelevant because only a few examples have been used for training. The idea of common adjectives as the relation between the entities works only for metaphorical noun pairs and is not defined for other forms of metaphors like verb-noun and adjective-noun metaphors (which are referred to as 'type II' metaphors and 'type III' metaphors respectively by Krishnakumaran and Zhu) [21]. However, this idea works reasonably well with metaphor relations of the form SUBJECT IS-A OBJECT (referred to as 'type I' metaphor by Krishnakumaran and Zhu).

Appendix A

Ampligraph Functions Used

Figure A.1 shows the functions from the Ampligraph 2.0.0 library that were used in this project:

Function	Description
`load_from_csv(directory_path, file_name[, ...])`	Load a knowledge graph from a .csv file.
`compile([optimizer, loss, ...])`	Compile the model.
`fit([x, batch_size, epochs, verbose, ...])`	Fit the model on the provided data.
`evaluate([x, batch_size, verbose, ...])`	Evaluate the inputs against corruptions and return ranks.
`predict(x[, batch_size, verbose, callbacks])`	Compute scores of the input triples.
`rank_score(y_true, y_pred[, pos_lab])`	Computes the rank of a triple.
`mr_score(ranks)`	Mean Rank (MR).
`mrr_score(ranks)`	Mean Reciprocal Rank (MRR).
`hits_at_n_score(ranks, n)`	Hits@N.
`train_test_split_no_unseen(X[, test_size, ...])`	Split into train and test sets.
`query_topn(model[, top_n, head, relation, ...])`	Queries the model with two elements of a triple and returns the top n results of all possible completions ordered by score predicted by the model.
`save_model(model[, model_name_path, protocol])`	Save a trained model to disk.
`restore_model([model_name_path])`	Restore a trained model from disk.

Figure A.1: Ampligraph 2.0.0 functions used

Bibliography

[1] Luca Costabello, Sumit Pai, Chan Le Van, Rory McGrath, Nick McCarthy, and Pedro Tabacof. AmpliGraph: a Library for Representation Learning on Knowledge Graphs, March 2019.

[2] Wikipedia contributors. Knowledge graph embedding — Wikipedia, the free encyclopedia. `https://en.wikipedia.org/w/index.php?title=Knowledge_graph_embedding&oldid=1144760029`, 2023. [Online; accessed 14-April-2023].

[3] Wei Song, Jingjin Guo, Ruiji Fu, Ting Liu, and Lizhen Liu. A knowledge graph embedding approach for metaphor processing. *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, 29:406–420, 2020.

[4] Ekaterina Shutova. Automatic metaphor interpretation as a paraphrasing task. In *Human Language Technologies: The 2010 Annual Conference of the North American Chapter of the Association for Computational Linguistics*, pages 1029–1037, 2010.

[5] Ciyuan Peng, Dang Thinh Vu, and Jason J Jung. Knowledge graph-based metaphor representation for literature understanding. *Digital Scholarship in the Humanities*, 36(3):698–711, 2021.

[6] Antoine Bordes, Nicolas Usunier, Alberto Garcia-Duran, Jason Weston, and Oksana Yakhnenko. Translating embeddings for modeling multi-

relational data. *Advances in neural information processing systems*, 26, 2013.

[7] Zhen Wang, Jianwen Zhang, Jianlin Feng, and Zheng Chen. Knowledge graph embedding by translating on hyperplanes. In *Proceedings of the AAAI conference on artificial intelligence*, volume 28, 2014.

[8] Yankai Lin, Zhiyuan Liu, Maosong Sun, Yang Liu, and Xuan Zhu. Learning entity and relation embeddings for knowledge graph completion. In *Proceedings of the AAAI conference on artificial intelligence*, volume 29, 2015.

[9] Guoliang Ji, Shizhu He, Liheng Xu, Kang Liu, and Jun Zhao. Knowledge graph embedding via dynamic mapping matrix. In *Proceedings of the 53rd annual meeting of the association for computational linguistics and the 7th international joint conference on natural language processing (volume 1: Long papers)*, pages 687–696, 2015.

[10] Han Xiao, Minlie Huang, Yu Hao, and Xiaoyan Zhu. Transa: An adaptive approach for knowledge graph embedding. *arXiv preprint arXiv:1509.05490*, 2015.

[11] Bishan Yang, Wen-tau Yih, Xiaodong He, Jianfeng Gao, and Li Deng. Embedding entities and relations for learning and inference in knowledge bases. *arXiv preprint arXiv:1412.6575*, 2014.

[12] Théo Trouillon, Johannes Welbl, Sebastian Riedel, Éric Gaussier, and Guillaume Bouchard. Complex embeddings for simple link prediction. In *International conference on machine learning*, pages 2071–2080. PMLR, 2016.

[13] Maximilian Nickel, Lorenzo Rosasco, and Tomaso Poggio. Holographic embeddings of knowledge graphs. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 30, 2016.

[14] Seyed Mehran Kazemi and David Poole. Simple embedding for link prediction in knowledge graphs. *Advances in neural information processing systems*, 31, 2018.

[15] Zhiqing Sun, Zhi-Hong Deng, Jian-Yun Nie, and Jian Tang. Rotate: Knowledge graph embedding by relational rotation in complex space. *arXiv preprint arXiv:1902.10197*, 2019.

[16] Ampligraph 2 – Ampligraph 2.0.0 documentation. `https://docs.ampligraph.org/en/latest/`.

[17] Yael Karov and Shimon Edelman. Similarity-based word sense disambiguation. *Computational linguistics*, 24(1):41–59, 1998.

[18] Julia Birke and Anoop Sarkar. Active learning for the identification of nonliteral language. In *Proceedings of the Workshop on Computational Approaches to Figurative Language*, pages 21–28, 2007.

[19] Ekaterina Shutova. Models of metaphor in NLP. In *Proceedings of the 48th annual meeting of the association for computational linguistics*, pages 688–697, 2010.

[20] Muhammad Abulaish, Ashraf Kamal, and Mohammed J Zaki. A survey of figurative language and its computational detection in online social networks. *ACM Transactions on the Web (TWEB)*, 14(1):1–52, 2020.

[21] Saisuresh Krishnakumaran and Xiaojin Zhu. Hunting elusive metaphors using lexical resources. In *Proceedings of the Workshop on Computational approaches to Figurative Language*, pages 13–20, 2007.

[22] Natalie Parde and Rodney Nielsen. A corpus of metaphor novelty scores for syntactically-related word pairs. In *Proceedings of the Eleventh International Conference on Language Resources and Evaluation (LREC 2018)*, 2018.

[23] Damian Borth, Rongrong Ji, Tao Chen, Thomas Breuel, and Shih-Fu Chang. Large-scale visual sentiment ontology and detectors using adjective noun pairs. In *Proceedings of the 21st ACM international conference on Multimedia*, pages 223–232, 2013.

[24] E Dario Gutierrez, Ekaterina Shutova, Tyler Marghetis, and Benjamin Bergen. Literal and metaphorical senses in compositional distributional semantic models. In *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 183–193, 2016.

[25] Metaphor Novelty Dataset link. `http://hilt.cse.unt.edu/resources.htm`.

[26] Visual sentiment ontology. `https://www.ee.columbia.edu/ln/dvmm/vso/download/vso.html`.

[27] VSO link. `https://www.ee.columbia.edu/ln/dvmm/vso/download/vso.html`.

[28] Gutiérrez ANP. `http://bit.ly/1TQ5czN`.

[29] Shakespeare's Sonnets by William Shakespeare. https://www.gutenberg.org/ebooks/1041.

[30] Yuri Bizzoni, Stergios Chatzikyriakidis, and Mehdi Ghanimifard. "Deep" learning: Detecting metaphoricity in adjective-noun pairs. In *Proceedings of the Workshop on Stylistic Variation*, pages 43–52, 2017.

[31] Wikipedia contributors. Literal and figurative language — Wikipedia, the free encyclopedia. https://en.wikipedia.org/w/index.php?title=Literal_and_figurative_language&oldid=1147449281, 2023. [Online; accessed 14-April-2023].

[32] Xu Liang. Summary of Translate Model for Knowledge Graph Embedding. https://towardsdatascience.com/summary-of-translate-model-for-knowledge-graph-embedding-29042be64273.

[33] What is a knowledge graph? https://www.ibm.com/topics/knowledge-graph.

[34] Balaji Kamakoti. Introduction to Knowledge Graph Embeddings. https://towardsdatascience.com/introduction-to-knowledge-graph-embedding-with-dgl-ke-77ace6fb60ef, 2020.

[35] Wikipedia contributors. Knowledge graph — Wikipedia, the free encyclopedia. https://en.wikipedia.org/w/index.php?title=Knowledge_graph&oldid=1147702633, 2023. [Online; accessed 14-April-2023].

[36] What is a Knowledge Graph? https://web.stanford.edu/~vinayc/kg/notes/What_is_a_Knowledge_Graph.html.

[37] Introduction to Knowledge Graph Embedding. https://aws-dglke.readthedocs.io/en/latest/kg.html.

[38] NLTK single-word part-of-speech tagging. https://stackoverflow.com/questions/44382254/nltk-single-word-part-of-speech-tagging.

[39] Universal POS tags. https://universaldependencies.org/u/pos/.

[40] VU Amsterdam Metaphor Corpus Online User Manual. http://www.vismet.org/metcor/manual/index.php.

[41] VU Amsterdam Metaphor Corpus Online. http://www.vismet.org/metcor/about.html.

[42] Wikipedia contributors. Norm (mathematics) — Wikipedia, the free encyclopedia. https://en.wikipedia.org/w/index.php?title=Norm_(mathematics)&oldid=1149197989, 2023. [Online; accessed 14-April-2023].

[43] What is the meaning of super script 2 subscript 2 within the context of norms? https://stats.stackexchange.com/questions/181620/what-is-the-meaning-of-super-script-2-subscript-2-within-the-context-of-norms.

[44] Github-Accenture/Ampligraph. https://github.com/Accenture/AmpliGraph/.

[45] Knowledge Graph Embeddings for NLP: From Theory to Practice. https://kge4nlp-coling22.github.io/.

[46] Knowledge Graph Embeddings Tutorial: From Theory to Practice. https://kge-tutorial-ecai2020.github.io/.

[47] Shakespeare's Sonnets. https://nosweatshakespeare.com/sonnets/.

[48] Gephi - The Open Graph Viz Platform. https://gephi.org/.