Proceedings of Machine Learning Research 176:283–287, 2022 NeurIPS 2021 Competition and Demonstration Track

The SenSE Toolkit: A System for Visualization and Explanation of Semantic Shift

Maurício Gruppi GOUVEMC@RPI.EDU
Sibel Adalı ADALIS@CS.RPI.EDU
Rensselaer Polytechnic Institute, USA
Pin-Yu Chen PIN-YU.CHEN@IBM.COM
IBM Research, USA

Editors: Douwe Kiela, Marco Ciccone, Barbara Caputo

Abstract

Lexical Semantic Change (LSC) detection, also known as Semantic Shift, is the process of identifying and characterizing variations in language usage across different scenarios such as time and domain. It allows us to track the evolution of word senses, as well as to understand the difference between the languages used by distinct communities. LSC detection is often done by applying a distance measure over vectors of two aligned word embedding matrices. In this paper, we present SenSE, an interactive semantic shift exploration toolkit that provides visualization and explanation of lexical semantic change for an input pair of text sources. Our system focuses on showing how the different alignment strategies may affect the output of an LSC model as well as on explaining semantic change based on the neighbors of a chosen target word, while also extracting examples of sentences where these semantic deviations appear. The system runs as a web application (available at `http://sense.mgruppi.me`), allowing the audience to interact by configuring the alignment strategies while visualizing the results in a web browser.

Keywords: Word Embeddings, Semantic Shift, Computational Linguistics

1. Introduction

Language is deeply rooted in the social, cultural and historical context that shapes it. It has been shown that word senses change over time Schmidt (1963). For instance, the English word *plane* was used to denote a surface before the invention of the aircraft and, thereafter, the word *plane* started to be used to indicate an airplane. Language is also subject to variation across domains, such as the same word being used in different cultures or communities Schlechtweg et al. (2019). For example, the word *semi* in American English refers to a type of truck and in British English is used to indicate a semi-detached house.

The detection of semantic change can be achieved by first aligning distributional word embeddings of two input corpora and then comparing the word vectors of a chosen target word through some distance measure. Alignment methods can expose different semantic changes and significantly impact the performance of downstream tasks Yehezkel Lubin et al. (2019); Gruppi et al. (2021, 2020). To explore this further, our system allows for the real time comparison of the following alignment methods: Global alignment Hamilton et al. (2016a), Noise-Aware alignment Yehezkel Lubin et al. (2019) and the S4 alignment Gruppi et al. (2021).

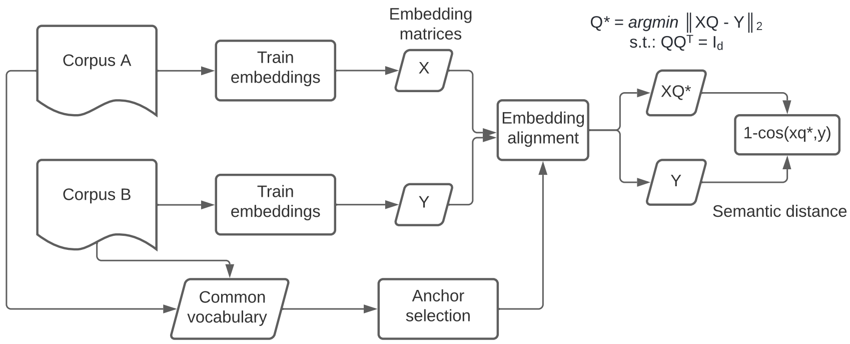


Figure 1: Overview of SenSE's pipeline from the input corpora through the final semantic distance measures. Word embeddings are trained on each input corpus. The embeddings are later aligned with Orthogonal Procrustes using a set of anchor words. This yields a transformation matrix Q^* which is used to align embeddings X to Y. After the matrices are aligned, the semantic distance between its words can be calculated with the cosine distance.

In this paper, we present the **Sem**a**n**tic **S**hift **E**xploration Toolkit (SenSE), a system to display analyses of semantic change in an interactive environment. In this system, the user is able to select words and explore the semantic change across several datasets and alignment methods, using multiple visualization methods. The system is available at http://sense.mgruppi.me.

2. The System

2.1. System Overview

The system works as a web application, allowing users to select one of several pretrained datasets and to visualize the results at each step by showing: the most semantically shifted words in a dataset, the contextual mappings between the input sources, and examples of sentences with significant semantic change.

The flow of the system's experience is outlined below:

1. User chooses a dataset to explore. Examples of datasets available are: Historical documents from different periods in English, German, Latin, and Swedish Schlechtweg et al. (2020), ArXiv categories Yin et al. (2018), and British vs. American English Gruppi et al. (2021).

2. User is presented with a list of the most shifted words according to each alignment method (S4, Global, Noise-Aware).

3. User selects or searches for a target word to explore the different senses across the input corpora.

4. A 2D visualization of the embedding space is displayed to describe the semantic differences of the target word between the aligned input corpora (Figure 2).

5. User may interact by adjusting the number of neighbors shown, changing the direction of the mappings, interacting with the plot.

6. Examples of distinct senses of the target words are shown in sentences from each of the input corpora. The aligned embeddings are used to determine the candidate sentences that exhibit the most semantic difference for the target word.

2.2. Preprocessing Pipeline

To prepare the datasets to be used in the system, a preprocessing pipeline is ran to clean up the text from the input corpora, train the word embeddings and perform the alignment between them. This preprocessing is done in the backend and is completely hidden from the user. The preprocessing pipeline can be seen in Figure 1 and is described below.

1. Each dataset consists of two corpora from distinct sources.

2. Two separate Word2Vec Mikolov et al. (2013) models are trained using each input corpus A and B, producing, respectively, embedding matrices X and Y of dimensions $n_i \times d$ where n_i is the size of vocabulary of corpus i and d is the size of the embedding vectors (typically in the range [100,300]).

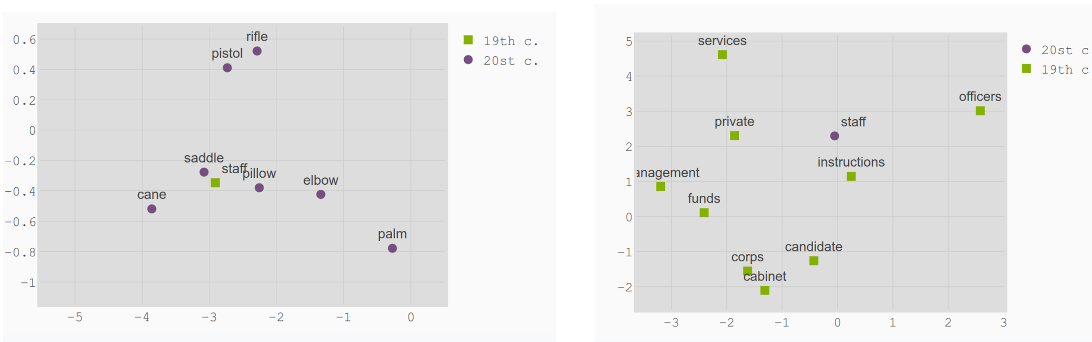
3. In order for the word embeddings to be comparable, their matrices are aligned using Orthogonal Procrustes Schönemann (1966) and the anchor words for this alignment are selected using one of the three methods: Global Hamilton et al. (2016b), Noise-Aware Yehezkel Lubin et al. (2019), S4 Gruppi et al. (2021). The goal of Orthogonal Procrustes is to learn a transformation matrix Q^* that maps X to Y according to the following objective and constraint:

$$Q^* = \arg\min_{Q} ||XQ - Y||_2 \quad s.t. : QQ^T = I_d$$

4. The semantic shift is computed for every word in the common vocabulary as the cosine distance between its word vectors in each corpus, after the alignment. That is, for a word w, the semantic shift is given by $s_w = 1 - cos(x_w Q^*, y_w)$, where xQ^* is the mapping of row vector x to the matrix Y.

5. Each word in X is mapped to Y's space and vice-versa. This creates a mapping of concepts of one domain to concepts of the other domain. This mapping is given by the nearest neighbors around the mapped position of the word.

6. Finally, we generate sentence examples that are semantically distinct based on the semantic shift of a target word. We compute the average vector representation of every sentence containing the target word w. Then, for each sentence in corpus A, we take the most distant sentences in corpus B as examples of contextually different sentences for w.



(a) 19th century

(b) 21st century

Figure 2: Snapshot of the system showing the nearest neighbors of the target word *staff*. In (a), the 19th century version of the word is shown in the 21st century embedding after alignment, which shows the neighbors with the 21st vocabulary that indicate a sense related a cane, whereas in (b) the 21st century version of word is mapped to the 19th century space and shows the 21st century meaning in terms of 19th century vocabulary, related to services and personnel. This does not necessarily mean *staff* has lost its former sense, but the latter sense is present only in B.

3. Conclusions and Future Work

In this paper, we described SenSE, a system for the exploration of semantic shift between two input corpora. SenSE allows users to visualize words that undergo significant semantic shift across different sources and provides explanations to the occurrence of this shift using a mapping of embeddings across the semantic spaces. This mapping is also used to find examples of sentences for a given word that are significantly different from each other. Currently, the system is limited to a number of pre-trained datasets. In future work, we intend to extend the system, allowing the comparison between other embedding models, alignment methods, and datasets. Additionally, we intend to enable the system to work with any input corpus, allowing the user to select any pairwise combination of sources to compare.

Acknowledgments

Acknowledgements This work was supported by the Rensselaer-IBM AI Research Collaboration (http://airc.rpi.edu), part of the IBM AI Horizons Network (http://ibm.biz/AIHorizons).

References

Maurício Gruppi, Sibel Adali, and Pin-Yu Chen. Schme at semeval-2020 task 1: A model ensemble for detecting lexical semantic change. In *Proceedings of the Fourteenth Workshop on Semantic Evaluation*, pages 105–111, 2020.

Maurício Gruppi, Pin-Yu Chen, and Sibel Adali. Fake it till you make it: Self-supervised semantic shifts for monolingual word embedding tasks. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 35, pages 12893–12901, 2021.

William L Hamilton, Jure Leskovec, and Dan Jurafsky. Cultural shift or linguistic drift? comparing two computational measures of semantic change. In *Proceedings of the Conference on Empirical Methods in Natural Language Processing. Conference on Empirical Methods in Natural Language Processing*, volume 2016, page 2116. NIH Public Access, 2016a.

William L Hamilton, Jure Leskovec, and Dan Jurafsky. Diachronic word embeddings reveal statistical laws of semantic change. In *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 1489–1501, 2016b.

Tomas Mikolov, Ilya Sutskever, Kai Chen, Greg S Corrado, and Jeff Dean. Distributed representations of words and phrases and their compositionality. In *Advances in neural information processing systems*, pages 3111–3119, 2013.

Dominik Schlechtweg, Anna Hätty, Marco Del Tredici, and Sabine Schulte im Walde. A wind of change: Detecting and evaluating lexical semantic change across times and domains. In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, pages 732–746, Florence, Italy, July 2019. Association for Computational Linguistics. doi: 10.18653/v1/P19-1072. URL https://www.aclweb.org/anthology/P19-1072.

Dominik Schlechtweg, Barbara McGillivray, Simon Hengchen, Haim Dubossarsky, and Nina Tahmasebi. SemEval-2020 Task 1: Unsupervised Lexical Semantic Change Detection. In *To appear in Proceedings of the 14th International Workshop on Semantic Evaluation*, Barcelona, Spain, 2020. Association for Computational Linguistics.

Karl Horst Schmidt. Ullmann stephen. semantics. an introduction to the science of meaning (book review). *Indogermanische Forschungen*, 68:183, 1963.

Peter H Schönemann. A generalized solution of the orthogonal procrustes problem. *Psychometrika*, 31(1):1–10, 1966.

Noa Yehezkel Lubin, Jacob Goldberger, and Yoav Goldberg. Aligning vector-spaces with noisy supervised lexicon. In *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)*, pages 460–465, Minneapolis, Minnesota, June 2019. Association for Computational Linguistics. doi: 10.18653/v1/N19-1045. URL https://www.aclweb.org/anthology/N19-1045.

Zi Yin, Vin Sachidananda, and Balaji Prabhakar. The global anchor method for quantifying linguistic shifts and domain adaptation. In *Advances in neural information processing systems*, pages 9412–9423, 2018.