SenSE: A Toolkit for Semantic Change Exploration via Word Embedding Alignment

Maurício Gruppi, [1] **Sibel Adalı,** [1] **Pin-Yu Chen** [2]

[1]Rensselaer Polytechnic Institute
[2]IBM Research
{gouvem, adalis}@rpi.edu, pin-yu.chen@ibm.com

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 demonstration, 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.

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, a phenomenon named Lexical Semantic Change (LSC) (Schmidt 1963). For example, the English word *awful* was used to indicate something impressive before the year 1800 while, in contemporary English, it is used to describe something objectionable. Language is also subject to variation across domains, such as in different cultures or communities (Schlechtweg et al. 2019).

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, Goldberger, and Goldberg 2019; Gruppi, Chen, and Adali 2021; Gruppi, Adali, and Chen 2020). To explore this further, our system allows for the real time comparison of the following alignment methods: Global alignment (Hamilton, Leskovec, and Jurafsky 2016a), Noise-Aware alignment

(Yehezkel Lubin, Goldberger, and Goldberg 2019) and the S4 alignment (Gruppi, Chen, and Adali 2021).

In this demonstration, we present the **Se**mantic **Sh**ift **E**xploration Toolkit (SenSE), a system to display analyses of semantic change in an interactive environment where the user is able to select which words to inspect to analyze and explore the semantic change across several datasets, using multiple visualization methods. The system is available at http://sense.mgruppi.me which also contains a video demonstration.

Experience

During the demonstration, the system exhibits the differences in LSC detection for the different embedding alignment strategies. Specifically, we will develop a novel system that focuses on showing how the choice of anchor words affects the outcome of the models. To that end, we will show the results for a given input pair of sources, listing the most semantically changed words between these sources when using global alignment (Hamilton, Leskovec, and Jurafsky 2016a), noise-aware anchor selection (Yehezkel Lubin, Goldberger, and Goldberg 2019), and using a self-supervised method for choosing landmarks (Gruppi, Chen, and Adali 2021). The audience may interact with the results, highlighting the output for each alignment method through a web based graphical interface. Users are also able to adjust hyper-parameters and immediately observe the change in the results. The system will display explanations for the semantic change based on the nearest neighbors of a selected target word in each of the two corpora. A interactive plot allows the user to explore the embedding space around the target word. As an additional part of the explanation, the system displays example sentences where the word is used in different senses across the corpora. The demo will run entirely as a web application, where the audience can access the system via a web browser and try the system on several provided datasets.

Demonstration overview The flow of the demonstration 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, Sachi-

dananda, and Prabhakar 2018), and British vs. American English (Gruppi, Chen, and Adali 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 1).

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.
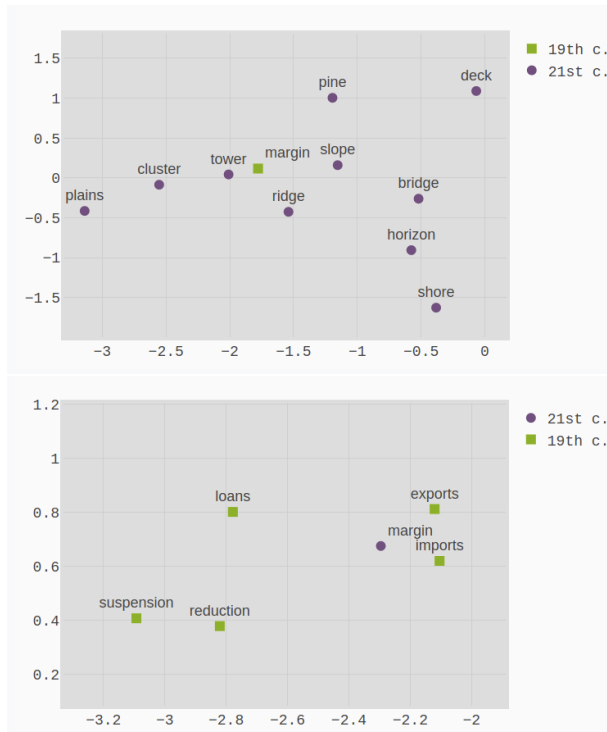

Figure 1: Snapshot of the system showing the nearest neighbors of the target word *margin*. In (a), the 19th century version of the word is shown in the 21st century embedding after alignment, which shows the neighbors with the 21st vocabulary that indicate a sense related to geographical and spatial elements such as horizon, slope and shore, 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 import, exports, loans, giving it a sense related to finances. Note that this does not mean *margin* has lost its original sense, but that these are different senses associated with the word in the two corpora.

Methods The methods utilized in each step of the demonstration are described below:

1. Each dataset consists of two corpora from distinct domains (being from different time periods, or different communities).

2. Two separate Word2Vec (Mikolov et al. 2013) models are trained using each corpus, producing embedding matrices U and V.

3. In order to make the embeddings comparable, the embeddings are aligned using one of the three methods: Global (Hamilton, Leskovec, and Jurafsky 2016b), Noise-Aware (Yehezkel Lubin, Goldberger, and Goldberg 2019), S4 (Gruppi, Chen, and Adali 2021).

4. The semantic shift is computed for every word in the common vocabulary as the cosine distance between its word vectors in each corpus. That is, for a word w, the semantic shift is given by $s_w = 1 - cos(v_w, u_w)$.

5. Each word in U is mapped to V'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.

Technology

We will demonstrate the use of embedding alignment for Lexical Semantic Change, which allows for the tracking and detection of semantic differences across time and domain. Typically, the alignment of embeddings for semantic change detection is done via Orthogonal Procrustes (OP) (Schönemann 1966). One of the greatest challenges of this problem is the selection of anchor words for alignment. OP aligns two embeddings by minimizing the euclidean distance between pairs of vectors (anchors). When all words are used as anchors, we call it global alignment. However, recent studies have shown that a more principled selection of anchor words may lead to better performance in semantic change related tasks (Yehezkel Lubin, Goldberger, and Goldberg 2019; Gruppi, Chen, and Adali 2021). This demonstration aims at exposing the quantitative and qualitative differences between alignment of word embeddings using different anchor selection strategies, and at providing explanations for semantic changes across different corpora.

Acknowledgements

References

Gruppi, M.; Adali, S.; and Chen, P.-Y. 2020. SChME at SemEval-2020 Task 1: A Model Ensemble for Detecting Lexical Semantic Change. In *Proceedings of the Fourteenth Workshop on Semantic Evaluation*, 105–111.

Gruppi, M.; Chen, P.-Y.; and Adali, S. 2021. 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, 12893–12901.

Hamilton, W. L.; Leskovec, J.; and Jurafsky, D. 2016a. 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, 2116. NIH Public Access.

Hamilton, W. L.; Leskovec, J.; and Jurafsky, D. 2016b. 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)*, 1489–1501.

Mikolov, T.; Sutskever, I.; Chen, K.; Corrado, G. S.; and Dean, J. 2013. Distributed representations of words and phrases and their compositionality. In *Advances in neural information processing systems*, 3111–3119.

Schlechtweg, D.; Hätty, A.; Del Tredici, M.; and Schulte im Walde, S. 2019. 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*, 732–746. Florence, Italy: Association for Computational Linguistics.

Schlechtweg, D.; McGillivray, B.; Hengchen, S.; Dubossarsky, H.; and Tahmasebi, N. 2020. SemEval-2020 Task 1: Unsupervised Lexical Semantic Change Detection. In *To appear in Proceedings of the 14th International Workshop on Semantic Evaluation*. Barcelona, Spain: Association for Computational Linguistics.

Schmidt, K. H. 1963. Ullmann Stephen. Semantics. An Introduction to the Science of Meaning (Book Review). *Indogermanische Forschungen*, 68: 183.

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

Yehezkel Lubin, N.; Goldberger, J.; and Goldberg, Y. 2019. 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)*, 460–465. Minneapolis, Minnesota: Association for Computational Linguistics.

Yin, Z.; Sachidananda, V.; and Prabhakar, B. 2018. The global anchor method for quantifying linguistic shifts and domain adaptation. In *Advances in neural information processing systems*, 9412–9423.