

# Linguistic Regularities in Continuous Space Word Representations [Mikolov et al., 2013]

Advanced Natural Language Processing / Engineering Seminar 2

February 2, 2024

1. What are the main findings of this paper? Are you convinced?
2. What do the authors claim is the main advantage of using distributed representations of words (aka. embeddings) over classical n-gram language models?
3. What is meant by a syntactic analogy? Give some examples of your own. Use some examples to explain why word2vec embeddings might be good at capturing syntactic regularities. Do you think the same would apply to other distributional word representations?
4. What is meant by a semantic analogy? Give some examples of your own. Use some examples to explain why word2vec embeddings might be good at capturing semantic regularities. Do you think the same would apply to other distributional word representations?
5. Explain the vector offset method used to answer an analogy question. In particular, what happens when no word exists at the predicted position?
6. What do you think of the results for identifying syntactic regularities? Is answering more than 1 in 3 questions correctly a good result? Are there any obvious trends in the results?
7. Do you think the comparisons with other methods are fair? Why do the authors use a different test set when comparing with Collobert & Weston (2008) and Mnih & Hinton (2009)? Does this test set appear to be easier or harder than the original one?
8. What do you think of the results for identifying semantic regularities? Why are results given for Spearman's Rank and MaxDiff Accuracy rather than Accuracy (as before)?

## References

- [Mikolov et al., 2013] Mikolov, T., Yih, W., and Zweig, G. (2013). Linguistic regularities in continuous space word representations. *NAACL*.