Language Shift Odyssey and Dynamic Word Embedding

Xiaolei Huang, Michael J. Paul
Department of Information Science, University of Colorado Boulder

Language Shift Overtime

Why? Three perspectives of language shifts

Dynamic Word Embedding (DWE)

References

Data Overview

Table 1: Top 5 closest words to “apple” across time on NYTimes.

Impacts on ML Models

Classifiers perform best when applied to the same interval they were trained. Performance diminishes when applied to other time intervals.

Does every word shift same speed?

Frequentist words vs. Polysemous words

Understanding Temporality by DWE

Figure 1. Language shift examples.

Figure 2. Document classification performance when training and testing on different years.

Figure 3. Dynamic semantic similarity between iphone and apple (blue) & cake and apple (orange) on the NYTimes dataset.

Acknowledgement

The presenter would thank for poster payment from Computational Linguistics, Analytics, Search and Informatics MS Program. This work was supported in part by the National Science Foundation under award number 1557338.