

Removing time dependencies in semantic analysis in the New York Times corpus

Carlos Selmo¹, Mariano G. Beiró ^{2,3}, Julián F. Martínez ², J. Ignacio Álvarez-Hamelin ^{2,3},
(1) Instituto Tecnológico de Buenos Aires, Argentina;
(2) Facultad de Ingeniería, Universidad de Buenos Aires, Argentina; (3) INTECIN-CONICET, Argentina;

The meaning of words changes dynamically throughout time. We can infer social trends and language constructs related to each period by studying the words' meanings evolution, e.g., how a particular word is associated with others. Traditional techniques such as word2vec or glove do not capture these evolutions correctly. Several techniques were proposed to study the changes of these semantic representations over the time focusing on the word association changes of a group of particular terms.

In this work, we present a variation of word2vec technique to study not the evolution of a particular word, but the words that take place in a particular area of the semantic representation for each period of time; and we apply it to the New York Times corpus.