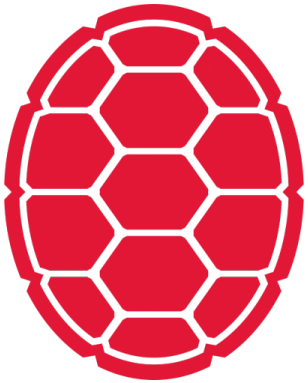


Distributional Semantics

Computational Linguistics: Jordan Boyd-Graber
University of Maryland
SLIDES ADAPTED FROM YOAV GOLDBERG AND OMER LEVY

Word Representation

- Before, we saw how valuable hidden layers were for representation (much more language today)
- How can we use it for words?

Word Representation

- Before, we saw how valuable hidden layers were for representation (much more language today)
- How can we use it for words?
- How similar is “pasta” to “pizza”
- Computers often use one-hot representations
- Or fragile knowledge bases

Word Representation

- Before, we saw how valuable hidden layers were for representation (much more language today)
- How can we use it for words?
- How similar is “pasta” to “pizza”
- Computers often use one-hot representations
- Or fragile knowledge bases
- Distributional Hypothesis (Harris, 1954; Firth, 1957)
- Know the word by the company it keeps

Intuition (from Boroni)

Marco saw a furry little wampimuk hiding in the tree

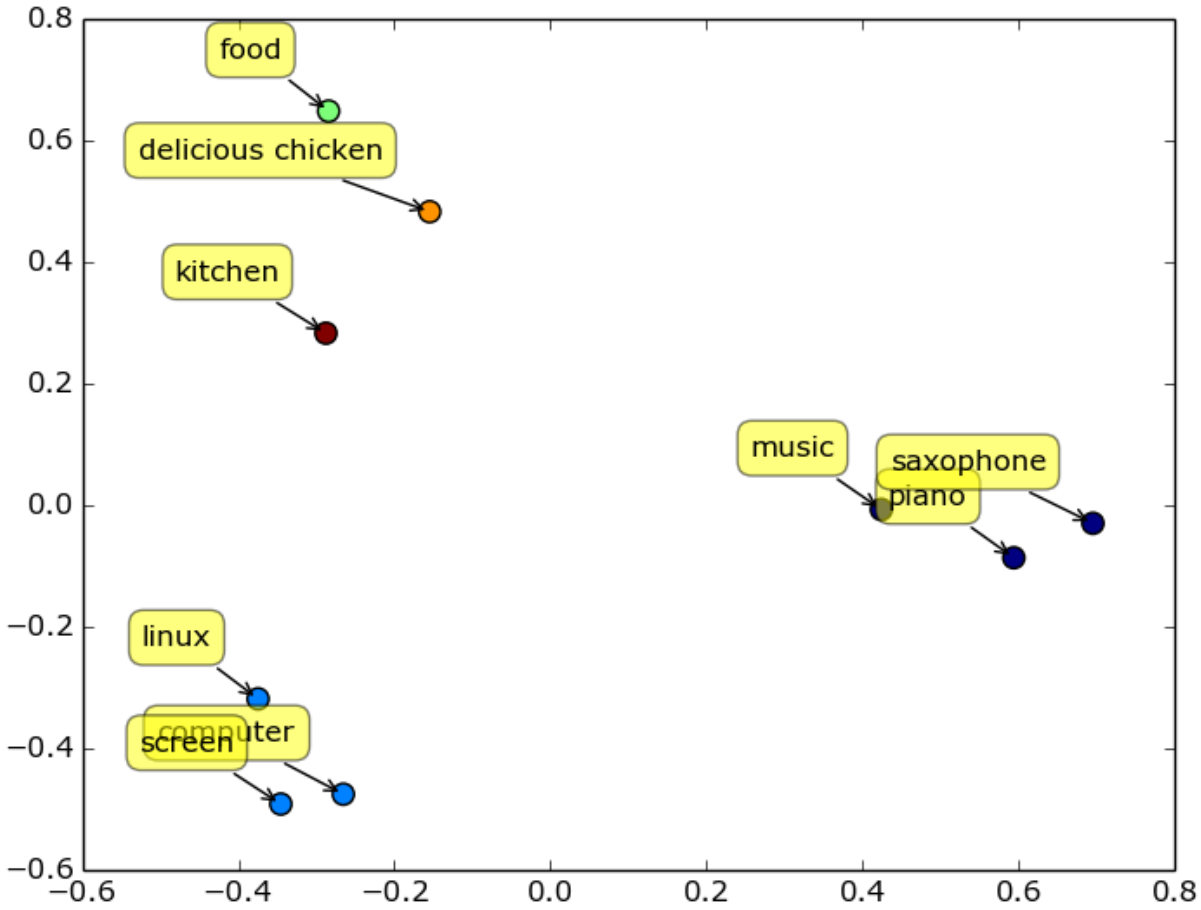
Intuition (from Boroni)

Marco saw a furry little wampimuk hiding in the tree

Intuition (from Boroni)

Marco saw a furry little wampimuk hiding in the tree

Representation



Usefulness

- Multimodal
- Multilingual
- Useful downstream feature