

Feature Cheatsheet

- **Surface:** The target word itself.
- **Pos:** The word's part of speech tag.
- **Lemma:** The word's lemma.
- **Surface-backwards:** The word spelled backwards.
- **Surface-every-second:** Every second character of the word itself.
- **Is-word-longer-than-10-letters:** Either yes or no depending if the word longer than 10 letters.
- **Contains-only-vowels:** Either yes or no depending if the word consists only on vowels.
- **Is-vector-number-divisible-by-13:** Either yes or no depending if the number of vectors for the word is divisible by 13.
- **Is-flag-word:** Either yes or no depending if the word is in the list of metaphor flag words*.
- **Flag-words-in-context:** Either yes or no depending on if there is a metaphor flag word in its vicinity (5 words before and after the target word).
- **Cos-similarity-first:** Cosine similarity between the meaning in context and the most common meaning.
- **Cos-similarity-second:** Cosine similarity between the meaning in context and the second most common meaning.
- **Cos-similarity-third:** Cosine similarity between the meaning in context and the third most common meaning.
- **Context_window+1_surface:** The word right after the target word.
- **Context_window+1_pos:** The part of speech tag for the word right after the target word.
- **Context_window+1_lemma:** The lemma for the word right after the target word.
- **Context_window -1_surface:** The word right before the target word.
- ...

Take	this	sentence	for	example
Context_window -2	Context_window -1	Target word	Context_window +1	Context_window +2