

FastText

Represent a word as a bag of character n-grams, e.g. for $n = 3$:

$G_{where} : \text{_wh, whe, her, ere, re_}, \text{_where_}$

Model a word vector as a sum of sub-word vectors:

SGNS:

$$\text{sim}(u, v) = \langle \phi_u, \theta_v \rangle$$

FastText:

$$\text{sim}(u, v) = \sum_{g \in G_v} \langle \phi_u, \theta_g \rangle$$

Code and pre-trained embeddings: <https://fasttext.cc/>