

Z Wikipedie, otevřené encyklopedie

Word2Vec je způsob [zpracování přirozeného jazyka](#), kdy jsou slova reprezentována vektory.^[1] Bezkontextový model publikovali v roce 2013 společně výzkumníci [Tomáš Mikolov](#), Kai Chen, Greg Corrado, [Ilya Sutskever](#) a [Jeff Dean](#), kteří pracovali pro Google. Algoritmus word2vec využívá model [neuronové sítě](#) k učení slovních asociací z velkého [korpusu](#) textu. Po natrénování může takový model odhalit [synonyma](#) nebo navrhnout další slova pro další část věty. Word2vec reprezentuje každé odlišné slovo určitým seznamem čísel, kterému se říká [vektor](#). Vektory jsou pečlivě vybrány tak,

aby zachycovaly sémantické a syntaktické vlastnosti slov; jednoduchá

aby zachycovaly sémantické a syntaktické vlastnosti slov; jednoduchá

us

us

sémantické podobnosti

GloVe

reprezentaci, přičemž [BERT](#) bere v úvahu kontext daného slova. Zatímco vektor pro slovo anglické "run" z word2vec je stejnou reprezentací z word2vec pro oba jeho výskyty "He is running a company" a "He is running a marathon", BERT poskytnou reprezentaci slova "run", která se bude lišit v závislosti na větě. Nově se používají novější modely jako [BERT](#) či [GPT-3](#).^[2]

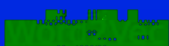
slovo ve slovníku jedinečnou reprezentací. Slovo "run" v kontextu "He is running a company" a "He is running a marathon" bude mít stejnou reprezentaci. BERT poskytnou reprezentaci slova "run", která se bude lišit v závislosti na větě. Nově se používají novější modely jako BERT či GPT-3.^[2]

Odkazy

Reference

Preklad textu z článku Word2vec na Wikicitate

V tomto článku byl použit překlad textu z článku Word2vec na anglické Wikipedii.



- [Oficiální stránka projektu](#)

Citováno z „<https://cs.wikipedia.org/w/index.php?title=Word2Vec&oldid=24823261>“

Kategorie:

- [Počítačová lingvistika](#)
- [Zpracování přirozeného jazyka](#)
- [Umělé neuronové sítě](#)

[Monitoring:Články přeložené z enwiki](#)

- [Monitoring:Autoritní kontrola s 0 identifikátory](#)

Hledání

Speciální:Hledání

Hledat

Word2Vec

18 jazyků

[Přidat téma](#)