

Morfológiai elemzés deep learninggel

Ács Judit
judit@aut.bme.hu

BME AUT

2016. október 25.

Hogyan készítenénk egy magyar szólistát?

Hogyan készítenénk egy magyar szólistát?

- ▶ vegyünk egy „nagy” korpuszt,

Hogyan készítenénk egy magyar szólistát?

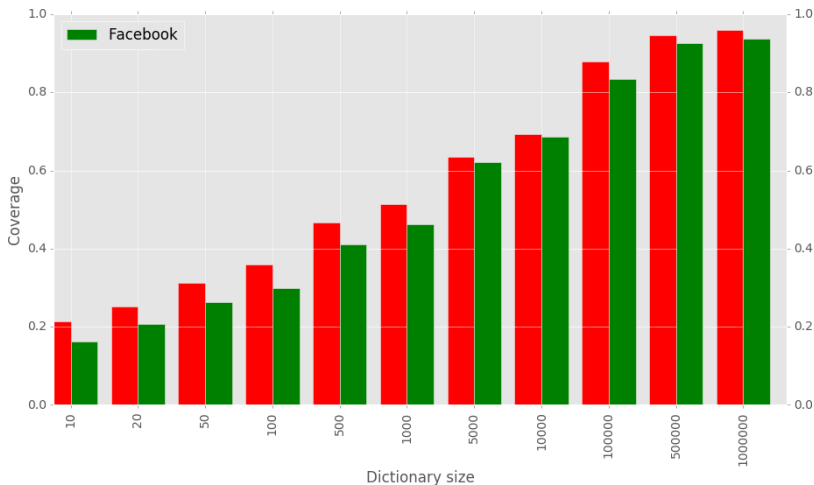
- ▶ vegyünk egy „nagy” korpuszt,
- ▶ bontsuk szavakra (tokenizáljuk),

Hogyan készítenénk egy magyar szólistát?

- ▶ vegyünk egy „nagy” korpuszt,
- ▶ bontsuk szavakra (tokenizáljuk),
- ▶ szótár!

Hogyan készítenénk egy magyar szólistát?

- ▶ vegyünk egy „nagy” korpuszt,
- ▶ bontsuk szavakra (tokenizáljuk),
- ▶ szótár!



Mi az a morfológiai elemzés?

- ▶ szavak belső szerkezetének feltárása

Mi az a morfológiai elemzés?

- ▶ szavak belső szerkezetének feltárása
- ▶ $házak = ház + (a)k$

Mi az a morfológiai elemzés?

- ▶ szavak belső szerkezetének feltárása
- ▶ *házak* = ház + (a)k
- ▶ zúzalékával?

Mi az a morfológiai elemzés?

- ▶ szavak belső szerkezetének feltárása
- ▶ *házak* = ház + (a)k
- ▶ *zúzalékával*?
- ▶ jelenleg főleg kézzel írt szabályokkal elemeznek,

Mi az a morfológiai elemzés?

- ▶ szavak belső szerkezetének feltárása
- ▶ $házak = ház + (a)k$
- ▶ *zúzalékával?*
- ▶ jelenleg főleg kézzel írt szabályokkal elemeznek,
- ▶ statisztikai módszerek, deep learning magyarra kevésbé elterjedt

Baseline: szófajok osztályozása szóalak alapján

Baseline: szófajok osztályozása szóalak alapján

- ▶ szóalak vs. ngram alapú

Baseline: szófajok osztályozása szóalak alapján

- ▶ szóalak vs. ngram alapú
- ▶ túltanulás! - gyakorlatilag egy memóriát építek

Baseline: szófajok osztályozása szóalak alapján

- ▶ szóalak vs. ngram alapú
- ▶ túltanulás! - gyakorlatilag egy memóriát építünk
- ▶ hogyan lehet általánosítani?

Fedettség angolra

