

Slovenski NLTK označevalnik

Niko Colnerič Nejc Banič

Fakulteta za Računalništvo in Informtiko
Univerza v Ljubljani

January 16, 2012

This is a short introduction to Beamer class.

- *Trigram tagger*
- *Brill tagger*
- classifier based *Naive Bayes tagger*

Postopek treniranja označevalnikov

- 1 Transformacija *.XML* korpusa v *.pos*
- 2 Treniranje z skripto iz NLTK-trainer
- 3 Ocenjevanje točnosti
- 4 Ocenjevanje hitrosti

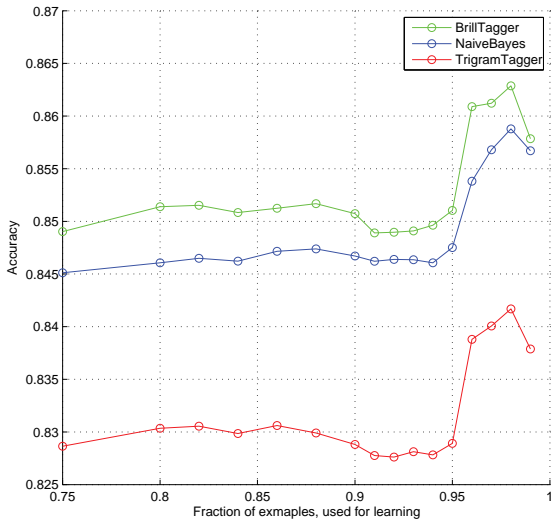
Lep je dan, vse diši že po pomladi

- (*Lep* — *PPNMEIN*) - pridevnik
- (*je* — *GP-STE-N*) - glagol
- (*dan* — *SOMEI*) - samostalnik
- (*,* — *,*) - ni razlage
- (*vse* — *ZC-SEI*) - zaimek
- (*diši* — *GGNSTE*) - glagol
- (*že* — *L*) - členek
- (*po* — *DM*) - predlog
- (*pomladi* — *SOZEM*) - samostalnik
- (*!* — *!*) - ni razlage

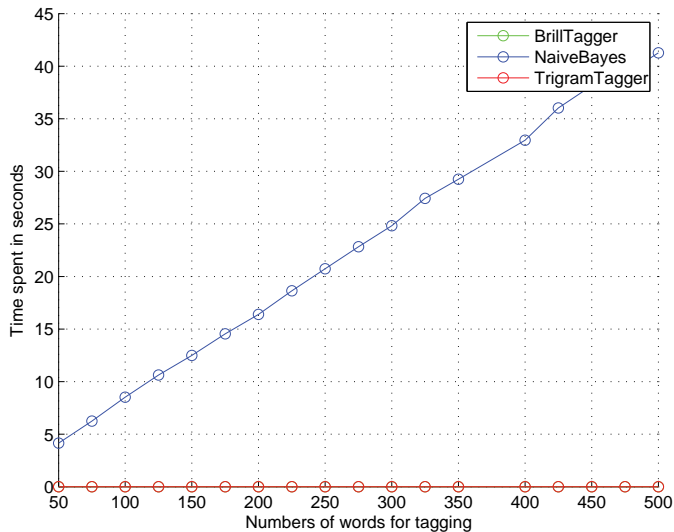
pridevnik

vrsta	splošni
spol	moški
število	ednina
sklon	imenovalnik
živost	0
vid	0
oblika	0
oseba	0
nikalnost	0
stopnja	nedoločeno
določnost	ne
število_svojine	0
spol_svojine	0
naslonskost	0
zapis	0

Rezultati natančnost



Rezultati hitrost



Zaključek in izboljšave

