Natural Language Processing

nltk

Natural Language Toolkit
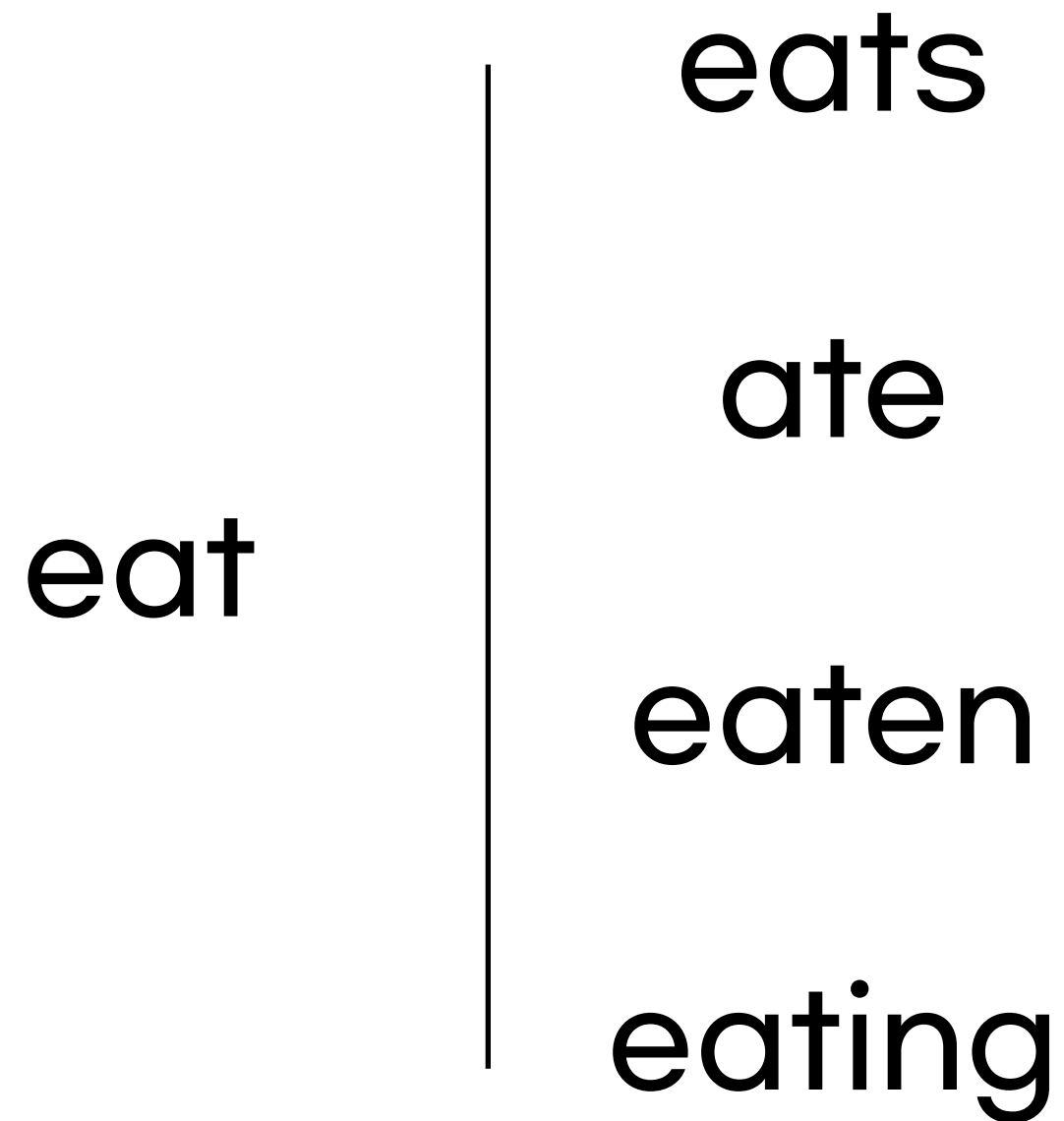
Tokenization

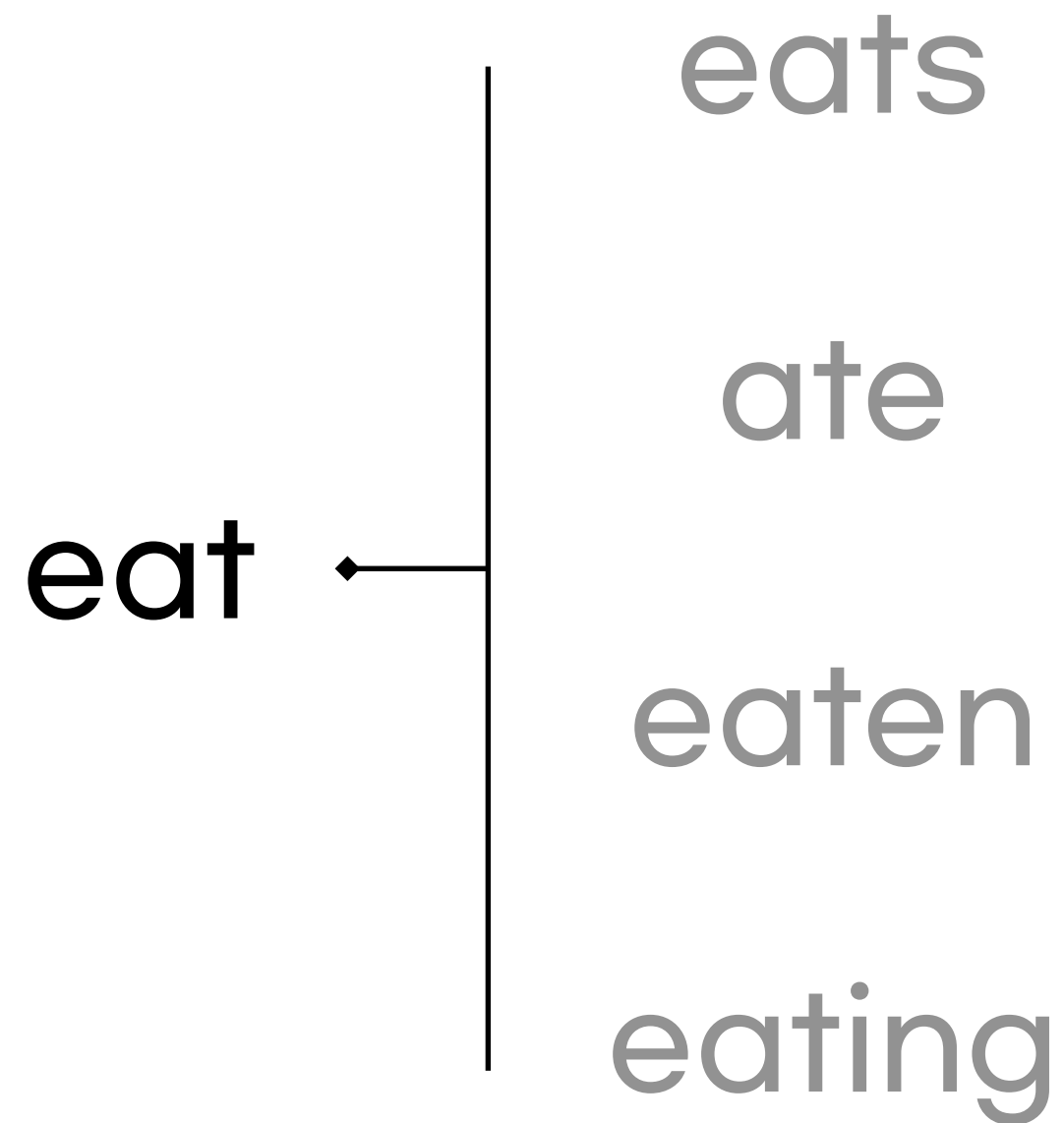The cat purrs.

Tokenization

The | cat | purrs

Normalization

eat | eats

ate

eaten

eating

Normalization

eat   ◄—| eats

  ate

  eaten

  eating

Stemming

eat~~ing~~ ⟶ eat

Lemmatization

morphology

ate ⟶ eat

POS Tagging

DT	determiner
JJ	adjective
NN	noun
NNS	plural noun
VBP	present verb
VBG	verb with -ing
IN	preposition