Authorship attribution with neural networks and multiple features

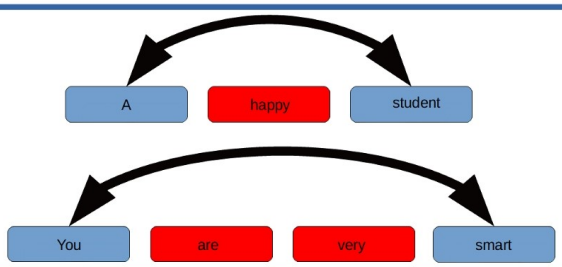
Łukasz Gągała

Abstract

However neural networks are receiving more and more attention from different fields of computer-aided research, application of this approach to stylometry and authorship attribution is still relatively infrequent in comparison to other domains of natural language processing.

In our paper we present our attempt to analyse frequencies of different types of linguistic data (part-of-speech, most frequent words, n-grams and skip-grams) with the means of simple neural networks.

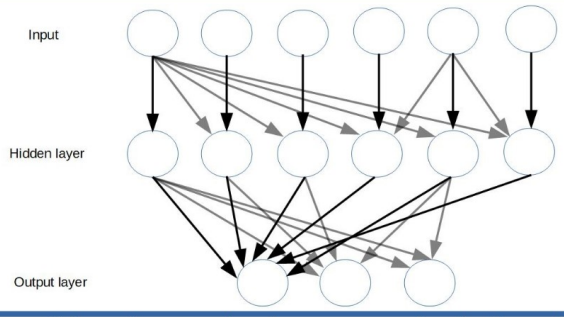
Skip-grams



Results of the initial experiment for different types of input data and for different network sizes

Type of input data	Number of features	Average F1 score
lemmas	1000	0.486
lemmas	2000	0.504
lemmas	3000	0.504
POS trigrams	1000	0.302
POS trigrams	2000	0.316
POS trigrams	3000	0.302
POS and 350 MFLemmas trigrams	1000	0.420
POS and 350 MFLemmas trigrams	2000	0.445
POS and 350 MFLemmas trigrams	3000	0.473

Since all neurons in subsequent layers are connected with each other this type of ANN-architecture is called "dense" or "fully connected" in the contrast to convolutional layers, which preselect data output from a previous layer by so-called filters.



The term "deep learning" refers to a method of stacking many layers of artificial neurons together, what improve their computational capabilities. In our approach we use simple architecture of so-called dense layers already proposed for stylometric analysis with n-grams of characters. We enhance this approach with various categories of features, since the authorial fingerprint is thought to be present across different types of text characteristics (most frequent words, function words, part-of-speech tags).

Features Selection

(1) Normal unprocessed text.

"The funeral had been a nightmare. Being who he is, the security kept away those who wished to sabotage the ceremony, and allowed only a minimal number
of people to enter the graveyard. It didn't stop some of the invited people from whispering in harsh tones insults that Mycroft chose to ignore."

(2) Character trigrams.

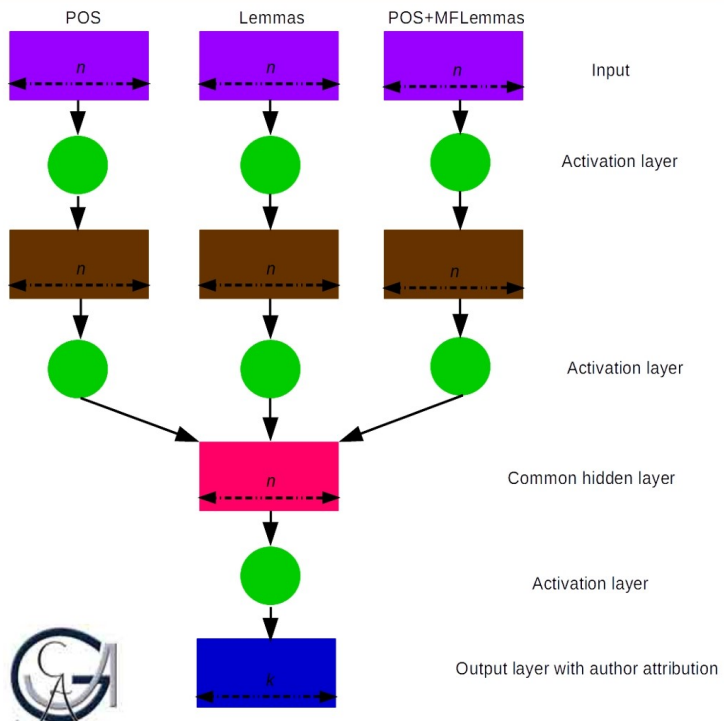
'_-t-h', 't-h-e', 'h-e-_', '_-f-u', 'f-u-n', 'u-n-e', 'n-e-r', 'e-r-a', 'r-a-l', 'a-l-_', '_-ha', 'h-a-d', 'a-d-_', '_-b-e', 'b-e-e', 'e-e-n', 'e-n-_', '_-a-_', '_-n-i', 'n-i-g', 'i-g-h', 'g-h-t', 'h-t-m', 't-m-a', 'm-a-r', 'a-r-e', 'r-e-_' (only first sentence showed)

(3) PoS-tags mixed with 150 most frequent words in the lemma form.

'the', 'DT', 'NN', 'have', 'VBD', 'be', 'VBN', 'a', 'DT', 'NN', 'NN', 'be', 'VBG', 'who', 'WP', '-PRON-', 'PRP', 'be', 'VBZ', 'the', 'DT', 'NN', 'keep', 'VBD', 'away', 'RB', 'DT', 'who', 'WP', 'VBD', 'to', 'TO', 'VB', 'the', 'DT', 'NN', 'and', 'CC', 'VBD', 'only', 'RB', 'a', 'DT', 'JJ', 'NN', 'of', 'IN', 'NNS', 'to', 'TO', 'VB', 'the', 'DT', 'NN', '-PRON-', 'PRP', 'didn', 'VBZ', 't', 'NN', 'stop', 'VB', 'some',
'DT', 'of', 'IN', 'the', 'DT', 'VBN', 'NNS', 'from', 'IN', 'VBG', 'in', 'IN', 'JJ', 'NNS', 'NNS', 'that', 'IN', 'NN', 'VBD', 'to', 'TO', 'VB'

(4) PoS-tags with corresponding lemma for each token.

"'the', 'DT', 'funeral', 'NN', 'have', 'VBD', 'be', 'VBN', 'a', 'DT', 'nightmare', 'NN', 'be', 'VBG', 'who', 'WP', '-PRON-', 'PRP', 'be', 'VBZ', 'the', 'DT', 'security', 'NN', 'keep', 'VBD', 'away', 'RB', 'those', 'DT', 'who', 'WP', 'wish', 'VBD', 'to', 'TO', 'sabotage', 'VB', 'the', 'DT', 'ceremony'"
(only first sentence showed)



gefördert durch
Cusanuswerk
Bischöfliche Studienförderung

GEORG-AUGUST-UNIVERSITÄT
GÖTTINGEN