Learning a Loss Function for Text Simplification

Rebekah Cramerus, Paul Opuchlich, Malte Klingenberg
28.06.2018

Learning a Loss Function for Text Simplification

problem

- training a neural network using gradient descent requires a differentiable loss function
- but: typical text simplicity features defined stepwise, not smooth

Learning a Loss Function for Text Simplification

problem

- training a neural network using gradient descent requires a differentiable loss function
- but: typical text simplicity features defined stepwise, not smooth

- idea: create „helper" network to calculate difficulty score

Learning a Loss Function for Text Simplification

datasets

- British Council „LearnEnglish“, News in levels etc.
- but: only small number of difficulty levels

- hopefully: Newsela (quasi-continuous difficulty scale)

Learning a Loss Function for Text Simplification

defining difficulty score

- lexical features: average word length, ratio of basic english words, type-token ratio...

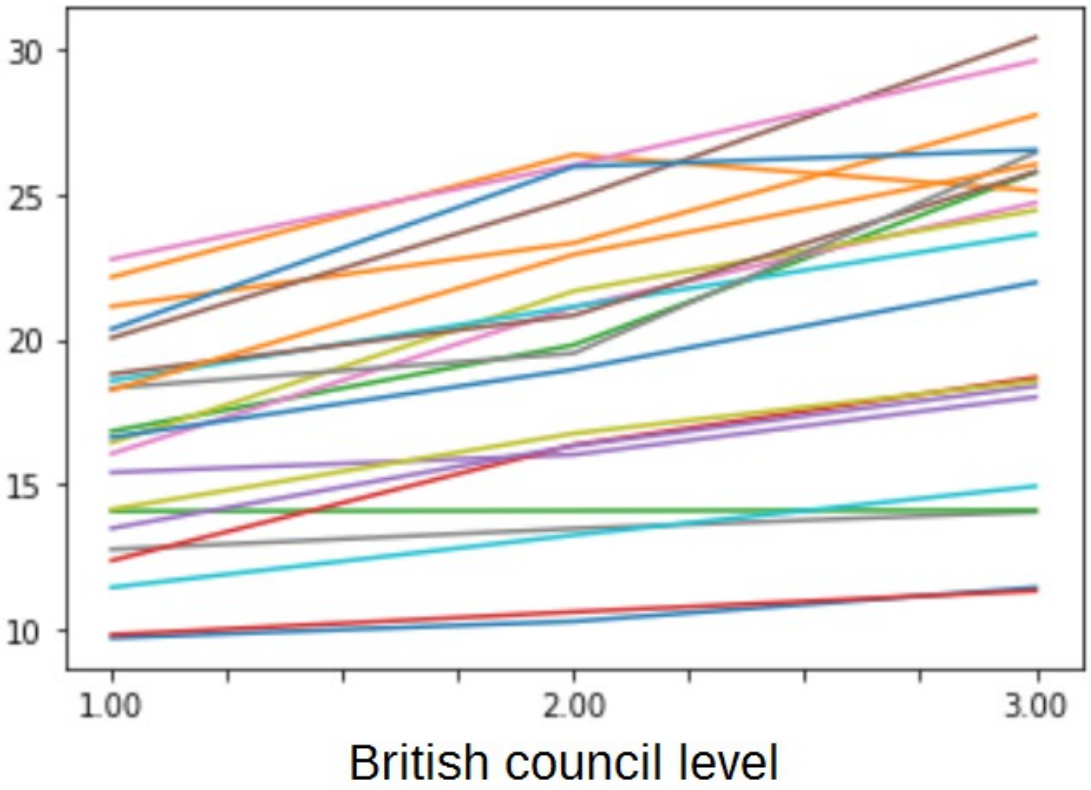
Learning a Loss Function for Text Simplification

defining difficulty score

- lexical features: average word length, ratio of basic english words, type-token ratio...
- syntactical features: number of subordinate clauses, prepositional phrases, passive voice...

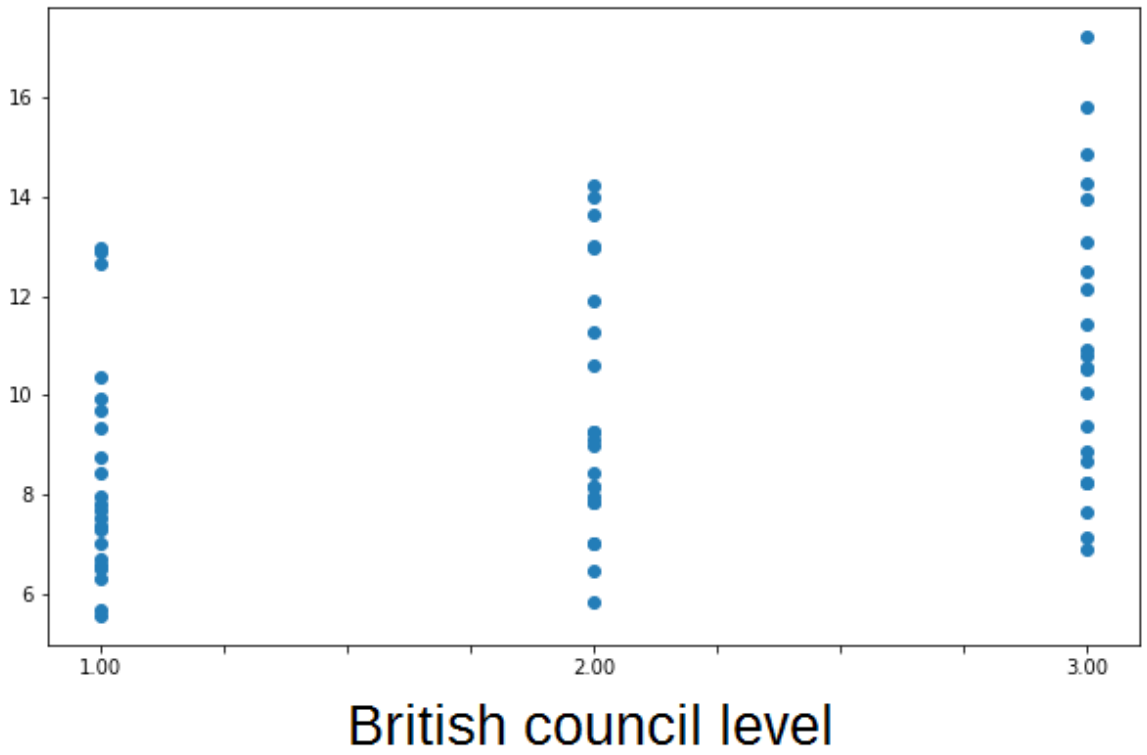
Learning a Loss Function for Text Simplification

type-token ratio



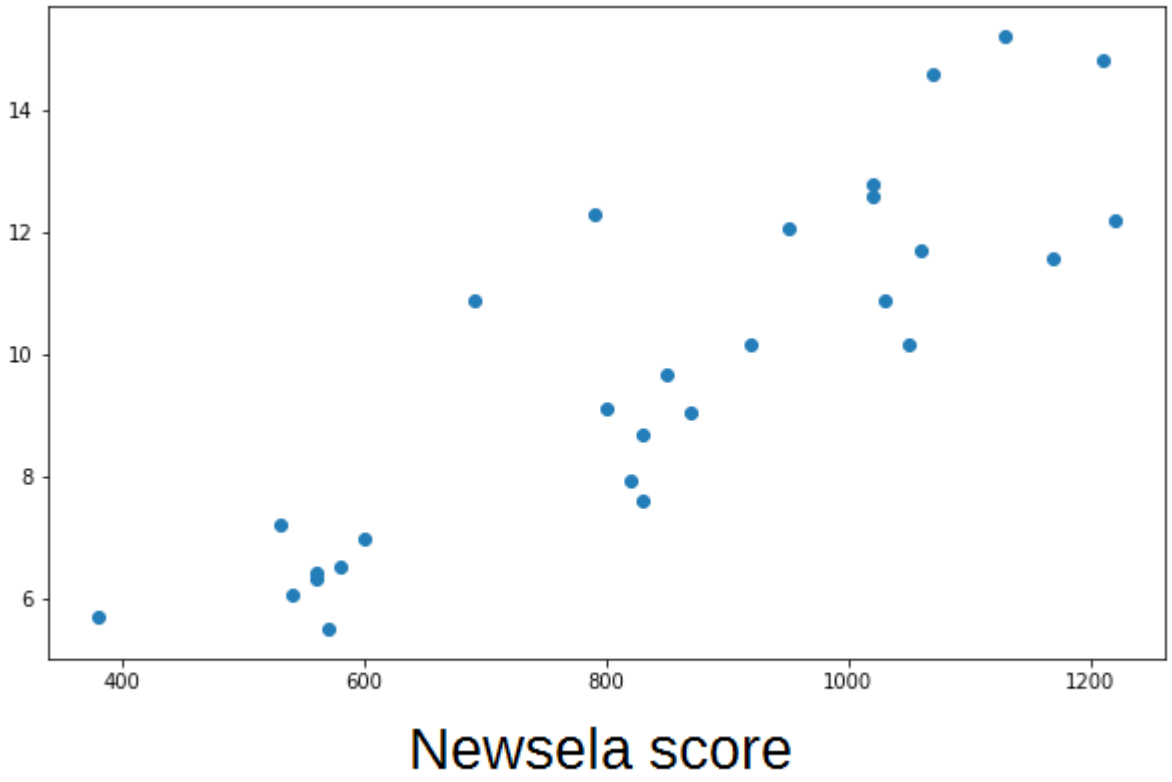
Learning a Loss Function for Text Simplification

all features combined



Learning a Loss Function for Text Simplification

all features combined



Learning a Loss Function for Text Simplification

next steps

- cross fingers about Newsela corpus
- train neural network on the difficulty scores