Dante Razo

LING-X 490

10/22/2019

Assignment 4

# Research Questions

1. Why would we want to use biased datasets when boosted random sampling nets better results (i.e. higher F1 scores)?
2. There are many papers focusing on explicit abusive language detection. What can be done to improve implicit abusive language detection?

# Summary

In the paper we read last week (Wiegand, Ruppenhofer, Kleinbauer 2019), multiple datasets were compared in abusive language classification. The highest F1 scores were earned by *boosted random sampling* datasets. The *Founta* dataset lacked the topic bias seen in *Waseem*; this is evident in Table 2 where the words with the strongest correlation with abusive posts are all considered vulgar. *Waseem*, on the other hand, classified typically benign words (e.g. football, mankind, driver) as abusive due to the focus / topic bias on sports. The words were flagged because of the environments they appeared in, meaning that a biased dataset might be good for working with implicit abusive language, but isn’t always reliable when dealing with explicitly abusive material.