Demagoguery and its Automatic Recognition in Texts

Evghenia Morozova, Victoria Bobicev

Technical University of Moldova, evghenia.morozova@iis.utm.md,
victoria.bobicev@ia.utm.md, 0000-0003-4450-3964, http://lilu.fcim.utm.md/

Abstract. The work is devoted to automatic text analysis in order to detect the
demagoguery. The demagoguery is used to play with the hopes and fears, and
especially the hatred, of the people, betraying their true interests for the sake
of one's own political popularity and power [1]. Automatic detection of
demagoguery in texts will help people to stay aware of their real goals and
avoid such publications.

Theoretical aspects of demagoguery such as types of demagoguery, its
subjects and objects, areas of application and basic techniques are
exhaustively studied and described in the theoretical part of the work.

The practical part describes the work of the dataset collection and
annotation. The texts with various types of demagoguery and without it had
been collected from various news sources in Russian language. The task of
demagoguery detection and classification was a difficult one even for human
annotators. Several persons evaluated the collected texts, then the
consistency of evaluation had been calculated using Cohen's Kappa [2]. The
resulting correlations depended on the annotator's experience and ranged
from 0.44, that is considered moderate agreement for the Cohen scale, to
0.84, that is almost perfect agreement.

Initially, five types of demagoguery have been defined; the total number
of six classes used for annotation included five types of demagoguery and texts
without demagoguery. Table 1 presents statistics for the collected dataset. It
is clear that some classes were presented by a very small number of texts; such
a small number was definitely not enough to train any machine learning

classifier. Thus, for machine learning experiments, a decision had been made to transform the annotation into binary one: 0 – no demagoguery, 1 – demagoguery is present.

Table 1. Class rates for multi-label classification

markers	Type of demagoguery	Number of texts	% percent of the whole set
0	Without any kind of demagogy	211	42.3
1	Appeal to the obvious	49	9.8
2	Appeal to the individual	127	25.5
3	Appeal to authority	44	8.8
4	Appeal to the majority	29	5.8
5	False dilemma	39	7.8

By merging all types of demagoguery in one class, an almost balanced dataset was obtained with 52.1% of text without any kind of demagoguery and 47.9% of texts with presence of demagoguery.

Several methods of machine learning had been tested for demagoguery classification. Among the classical ones, Logistics regression, Naïve Bayes and KNN had been selected [3]. In the experiments with all 6 classes, traditional models could detect with acceptable accuracy only large classes, 0 and 2. In binary classification, however, they demonstrated rather good results: F-measure was in the range of 0.86–0.89.

In the final set of experiments, all these models were tested on a separate list of 20 texts of different styles. Actual accuracy varied from 60% to 100% depending on the model and studied class.

References

[1] Ceaser, James W. "Demagoguery, Statesmanship, and Presidential Politics". Designing a Polity: America's Constitution in Theory and Practice. Rowman & Littlefield. (2011) pp. 75–118.
[2] Hayes, A.F.; Krippendorff, K. "Answering the call for a standard reliability measure for coding data". Communication Methods and Measures. 1 (1) (2007) pp. 77–89.
[3] K.P. Yadav, Sandeep Kulkarni, Machine Learning in Natural Language Processing. Cherry Publications India (Frateclat Private Limited) (February 11, 2022).