

Project: Hate Speech detection using Transformers (Deep Learning)

Group Name: Itvn_sergey

Name: Litvinov Sergey

Email: Itvn.sergey.work@gmail.com

Country: Russia

Company: NVI Solutions

Specialization: NLP

Problem description

Hate speech detection is a problem of sentiment classification. The task is to classify certain piece of text as the one with hate speech or not, by training deep learning model on labeled data.

The term hate speech is understood as any type of verbal, written or behavioral communication that attacks or uses derogatory or discriminatory language against a person or group based on what they are, in other words, based on their religion, ethnicity, nationality, race, color, ancestry, sex or another identity factor.

Data understanding

There are two **csv** format datasets in available:

- **train.csv** - with **id**, **tweet** with text and **label** - 1 for hate-speech and 0 - otherwise
- **test.csv** - with **id**, **tweet** with text

EDA performed

1. Check for NA values
2. Check for duplicates
3. Investigating class balance
4. Visual text analysis

Data cleansing and transformation done on the data.

1. Normalize casing
2. Lower case
3. Remove **URLs**
4. Remove mentions
5. Remove special characters
6. Expand contractions
7. Tokenize sentences into individual terms, using tokenizers
8. Lemmatization
9. Remove stop words