NLP: Conceptual Development Report

Project Initiation

For the development of this project we have elected to work in Python. We are currently researching what the best library to use is - options include Tensorflow, Keras, etc. We are taking an iterative development approach.

The Bitbucket repository where the project will be located is at the following link: https://bitbucket.org/karamihalevm/nlp-project/src/master/

Team roles

Marin Karamihalev - programmer, help w/ statistical analysis
Siyana Ivanova - scheduling, statistical analysis, write the paper, design GUI
Jakub Blaszkowski - data gathering & pre-processing
Jae Sun Lee - code review, testing & quality assurance
Wojciech Drezek - programmer

Rough schedule for the semester

13/03	Submit conceptual development report
Until 31/03	Gather and process data
	Research implementation methods
Month of April	Test different machine learning models
	GUI development
	Begin report
Month of May	Finalise selected model
	Testing and QA
	Finalise report and analysis
29/05	Present project
30/05-04/06	Polishing and final touches
05/06	Project deadline

Initial Schedule: First 3 weeks of development

Week 1:

13/03	Submit conceptual development report	Siyana
14-19/03	Collect data set	Jakub
	Begin data preprocessing	Jakub
	Research implementation methods	Marin

Week 2:

20/03	Consultation on issues (if any arise)	
21-26/03	Data preprocessing	Jakub
	Research implementation methods	Marin
	Further scheduling	Siyana

Week 3:

27/03-02/04	Finish data preprocessing	Jakub
	Begin writing the code	Marin & Wojciech
	Begin writing documentation	Siyana
	Further scheduling	Siyana

Business analysis

Objectives of the project

- Create a machine learning model which distinguishes and classifies transphobia in English-language tweets
- Perform a statistical analysis on the results obtained by the classifier
- Study the prevalence of transphobia on English-language Twitter by answering the following research questions
    - How prevalent is transphobia on Twitter?
    - What are the most common ways in which people express transfobia?
    - In what context does transphobic speech usually occur?
    - What locations are transphobic tweets most often sent from?

Functionalities of the program

- Run the processed data through the model and classify the transphobic content of a tweet on a scale of -1 to 1
    - -1 - transphobic
    - 0 - neutral
    - 1 - trans positivity
- Output a range of statistics on the dataset of tweets
    - Prevalence by country
    - Prevalence by hashtag
    - Percentage of tweets falling into each tier
    - Most commonly used transphobic slurs
- Gather tweets with a given hashtag and output ratings of transphobic, neutral or trans-positive content
- Receive a single tweet as input and output an evaluation of its content

Scope of the paper

- Discussion of the method that was used to implement the above functionalities
- Statistical breakdown and analysis of the results
- Answers to the research questions that were posed at the beginning of the project

Reflection on the method

Our goal is essentially to perform sentiment analysis on a set of tweets. The first step is to get approved for the use of Twitter API and collect a dataset of tweets with relevant words and phrases. Then the tweets will be ranked on a scale of -1 to 1. The resulting dataset will be processed (tokenized, part-of-speech tagged, stop words will be removed, etc.) and then run through a sentiment classifier developed by us. We will collect the relevant statistics and use them to draw our conclusions, which will be described in the final paper. We expect this project to result in a sentiment classifier which can be reused on more tweets, and the aforementioned report discussing our method and analysis.

Motivation

We've chosen this project topic due to our interest in how online hate speech works and in trans rights. According to a study from the University of Arizona published in 2018, suicide among transgender adolescents is much higher than among their cisgendered peers. In particular, among young trans men the rate of suicide attempts is as high as fifty percent. The researchers claim that a likely reason for this is the marginalisation and harassment that these individuals experience in their day to day lives. In this day and age, a large portion of communication happens on social media, on platforms like Twitter, which is especially notorious for hate speech. This is why we would like to examine the occurrence of transphobia (but also of positive sentiments on trans topics) on Twitter and study the circumstances in which it happens. We are also curious to see if clusters form in specific locations or around specific hashtags. We hope this may shed a little bit of light on this issue and provide some information about transphobia.

A link to an article discussing the study in question with the main author:
https://www.reuters.com/article/us-health-transgender-teen-suicide/trans-teens-much-more-likely-to-attempt-suicide-idUSKCN1LS39K