

COMP 4641 Project Proposal

Twitter Abusive Behaviour identification

HE Jianle

20536134

Chiu Chi Fun

20536110

1 Reaction Paper

1.1 Summary

1.1.1 Opinions and social pressure

The paper is written by S. Asch in 1955 about the majority affecting others' own opinions. People have different opinions on different topics when your own opinion is different from society's majority, there was great pressure on those people. In the paper, the author has done some experiments on testing people facing the society's majority, they will stand with their own opinions or follow the majority to run away from the pressure of being different from others'. In the experiment stated that 75 percent of people agree with the majority. Although the majority are wrong and misleading others, there were fewer people who could follow their own choice and even point out others' misleading. This paper pointed out the importance of the majority affecting the public choice and this phenomenon applies to the misleading majority would cause a bad influence on the public society.

This paper states that this phenomenon can also apply to social media nowadays. It can also be related to lecture 6: Centrality. The starting point of the majority can be out by the centrality of the social media network. If the

majority is something misleading the public can be easily found by it and fix them.

1.1.2 The Small World Problem

This is a paper studied by Stanley Milgram and published in 1967. From the research mentioned in the paper, a randomly selected American can mark down around 500 people's names who have had a conversation in the past 100 days with them. And conclude that there is a chance that two Americans knew each other by randomly picking from 200, 000 Americans. An experiment also mentioned in the paper, two randomly picked Americans asked to pass a file from one to another through their own social network. The result found that the average number of people needed to deliver the file is five.

The Small World is one of the chapters of the course. We learn this phenomenon with great understanding by applying the concept and technique such as graph theory. Analysing the diameter, shortest path and cluster of the network, are some of the examples mentioned in the lesson.

1.2 Critique

1.2.1 Strength

In the paper, Opinions and social pressure, there are many different experiments to prove the phenomenon exists inside of the society. Besides, the experiments were set at a different angle that can provide various views on the majority affecting others' opinions. Lastly, the paper provided some questions to readers to think more about the paper that allows thinking deeper to the paper's centre problem.

The Small World Problem provided a clear definition and attempt to address the reasons by using experimental data. While explaining the abstract concept, it gave many graphs and images to explain and to visualise. This provided a solid ground and direction for the analysis in the future.

1.2.2 Weakness

On the other hand, the sample size of Opinions and social pressure's experiment is too small to increase the credibility of the result. There were only 10-20 people in the whole experiment. Moreover, the experiment is too old to apply to the nowadays information network and society. As society is greatly different from it in the 1950s, the invention of the Internet has greatly changed the society. Therefore, the experiment can be changed a bit to fulfil social media on the Internet in order to increase credibility.

These issues also exist in The Small World Problem. There will be some dropout cases in the experiment. The new technology nowadays can greatly increase the willingness of experimental subjects. Sending a file by emails, instance messages are much

more convenient than passing by hand or by postal.

1.3 Brainstorming

This paper can help to find out the centrality of the hateful speech on twitter. We can build Twitter Abusive Behaviour identification to find out those hateful speech and ban or hide those tweets. Nowadays, Twitter is always spreading different information on the Internet by using the retweet function. When a piece of misleading information that can attack other people was retweeted by many users, it would become a kind of majority among the twitter users and affect other users to express their own opinions and also the victim would be hurt by those tweets.

2. Project Proposal

2.1 Introduction

In the recent year, there are tremendous arguments around the world, in such a fierce world, the number of hateful speech in the social network soaring which seriously damages the mental health of the public and even threatens the growth of the teenagers, make it become one of the most serious problems on the social network. To help make a better virtual social world, we are necessary filtering out most of the hateful speech, which imply we need to identify and screen those abusive behaviour and hateful speech out.

2.2 Dataset

To have better understanding on the problem, the data from Harvard University is used which is annotated tweets associated with Inappropriate speech like abusive and hateful speech, as well as Normal interactions and Spam.

2.3 Planing

2.3.1 Dataset

Analyst the dataset does find out useful metrics and hidden patterns by using different programming techniques, like finding the distance and time needed to spread a hateful speech worldwide. Also, using the mathematical technique to build a model as an identifier to distinguish the hateful speech among the ocean of public opinion.

2.3.2 Evaluation

To tackle and identify the hateful speech, which turns natural language into data, Natural Language

Processing(NLP) is required. For example, using recurrent neural networks with LSTM is one of the directions for training the NLP model. Moreover, we need to process the data through various mathematical and machine learning models. Building a classifier with a neural network for deep learning will be a good idea. Finally, visualisation, some python packages can be used to plot the clear graph and build table, as we all know, pandas and matplotlib.

With the given data, a supervised machine learning can be carried out, which imply we can use the method of evaluate the model as the way to evaluate our method, such as F1 score and accuracy, they are definitely helpful to evaluate and compare model, which able to give us a bright and clear way to achieve a functional and useful model.

2.4 Final Result

There are two things we are going to submit. First one, a hateful speech identifier that is well-trained and efficient. Also, a report with a clear picture on the problem and able to give good explanation and analysis on the pattern of the data.

Reference

S. Milgram, The small world problem,
Psychology today, 1967.

S. Asch, Opinions and social pressure,
Scientific American, 1955.