LIFE! Annotation Guidelines for Suicide prevention corpus

Project Summary

Suicide, depression and Bullying, are problems that involve a huge rage of population but, according to OMS, they can be prevented with the suitable tools.

Life! is a project developed by the department of language and computer systems (DLSI) that it belongs to the University of Alicante (UA). It wants to help and try to prevent these problems by using social networks as a tool of fast detection in order to create a suitable protocols. All this related with the technologies of human language and automatic learning.

Task Summary

Your task will consist in categorize your assigned documents according to two different classifications as well as the language of these documents. The phenomena are explained in 'General considerations' section. It should be mentioned that during the training round the documents are pre-categorized. So you just need to correct one classification, only if you don't agree with it.

These documents belong to a corpus described in the Corpus description section.

Your working languages will be both Spanish and English and you will be annotating using Gate 8.2 tool.

Corpus Description

A corpora, in general, needs as much representative expressions as It can be included. So, in suicide field, the fast access of social networks is a handicap but an advantage too if We had suitable tools so as to It lends us to catch the most representative expressions in that field and, train learning machines for that.

That is the reason why all the information here included belongs to social networks as Twitter, Facebook and other sites as Reddit, Blogger or Wattpad.

But the process is complex, so, we designed a method in order to obtain an objective corpus.

Step 1. Looking for seeds related with suicide in the most common social networks.

Step 2. Obtaining new words after having the seeds located. For that task, we looked for Snippets, messages in Twitter browser, Google browser, etc.

Step 3. All the words obtained, they were searched in facebook public groups, Facebook and Twitter lists, public blogs and public profiles.

Step 4. In excel document, putting all the information in order to have the urls and, If It were necessary, doing some comments.

It's important to remark that we only observed the active profiles, forums, etc. so as to get the representative aim of our corpus.

Specific Considerations

In our corpus, we have one annotation "SuicidalClassification" that has four properties: "Language", "AlertLevel", "PrimaryMessageType" and "SecondaryMessageType". The first three are mandatory, whereas the latter is optional.

Language

The first property, "Language", includes two classes (i.e. EN, for English; and ES for Spanish).

In some cases, we have both languages so the selected one will be the most substantial.

AlertLevel

The second one, "AlertLevel", deals with four categories, namely: Immediate, Urgent, Possible and No risk.

This property indicates the necessity or not that an user has to be treated or derived to health services. Obviously, If we notice that there are necessity, we'll mark Urgent or Immediate.

1. Inmediate

Immediate

Description This is the highest level of alert because the people include here

needs an immediate derivation of health services. They express clearly the idea of being a suicidal. Of course suicidal groups and instigator profile are suitable here.

Detection hints Description of auto-lesive behaviour with a determined expressions

as now or a complex process described are suitable here.

Examples "In two hours my life doesn't make sense for anyone, I'll be with God"

2

"Mom, Dad I love You and I'm sorry for everything, I''ll be better there rather than here. If you are reading this, It means that I won't be with you"

2. Urgent

Urgent

Description The suicidal thought is remarkable but not in such as immediate way.

Repetitive depressive thoughts are related to this category.

Detection hints Urgent. Here we would apply the next cases:

- If we have ways of killing yourself but in the texts we can

observe supportive things such as loved relatives for whom they don't want to commit suicide. - Phrases as "I feel that I'm not anyone" - Description of auto-lesive behaviour.

Examples "I don't find anything interesting so, I felt tired and exhausted. I don't

want to continue anymore but I've got one child and he doesn't deserves that I''ll be passed away"

3. Possible

Possible

Description Not every message has the same risk. So as to a possible message indicates a temporal episode of depression, sadness, dissatisfaction with life, etc.

Detection hints Possible is related to cases in which we have doubts about If we have

risk or not.

Examples "Life is a shit"

"I have been looking for the truly meaning of life"

4. No risk

3

No risk

Description Here we include texts in which the suicide is not commented in any

way. Also, there are positive thoughts and encouragement positive messages too.

Detection hints

Examples "It's a beautiful day I can't stop myself from smiling"

"The social networks are a good way to connect with the people"

MessageType

There are two properties that we use to categorize the type of text:

"PrimaryMessageType" and "SecondaryMessageType". Both of them share the same

categories but only the former is mandatory. Specifically, there are 12 categories:

Auto-Pro-Suicide, No-Auto-Prosuicide, Irony, Mental Illness or Disorder, Suicide Curiosity,

Euthanasia, Mysticity, Undefined, Instigator, Sadness/Melancholy, Citation and Depression.

1. Auto-Pro-Suicide

Auto-Pro-Suicide

Description The human behaviour is complex. That's the reason why one person

with suicidal thoughts or with depression is suitable here. Also we have to include similar feelings like desire of dying, negative thoughts (in general) and, of course, all the people who has talk about suicide attempts.

Detection hints Depression facts that It matches with Auto-Pro-Suicide behaviours

are included here so, we don't have to mark in the property TypeMesssage the category Depression. Minimun, in the property TypeMesssage this category has to be marked as Urgent or Immediate. In case of agreement with the death topic but It is not mentioned as a possibility, It can be marked in the TypeMessage Possible. The autolesive behaviours in order to reveal a concrete expression of angry but not related with suicide are not included here.

Examples 2. Auto-Pro-Suicide

3. Auto-No-Pro-Suicide

Auto-No-Pro-Suicide

Description Not everybody agrees with the idea of suicide, so we had to create a

category in which this people can be included. The suicidology movement, positive thoughts (in general), supportive texts and people who had overcome suicidal thoughts and depressive disorder.

Detection hints

Examples "I understand you, but life is beautiful so, If you need talk, I'm here for

you" "Few months ago I was in your situation and thought about suicide myself but I have just discovered the meaningful meaning of life"

4. Irony

Irony

Description The irony consists of a figurative use of language. That's the reason

why we had to use a cognitive operations in an adequate context. For that goal, the receptor have to do a comparison, contrast and eco labours (Uaje, José, Mendoza et. al. 2012). In our corpus, the Irony is showed.

Detection hints

Examples "It's like choosing soup over the salad"

5. Mental Illness or Disorder

Mental Illness or Disorder

Description Manifestations of mental illnesses or disorders are included here.

5

Detection hints Here we ONLY include mental illnesses that had been

diagnosticated previously.

Examples "I've been taken pills for major depression diagnosed few years ago"

6. Suicide Curiosity

Suicide curiosity

Description The people included here express a kind of interest of suicide topic.

Detection hints

Examples "I don't know why but I feel a bit curiosity about the idea of suicide, If I

had all the tools needed, I would try"

7. Euthanasia

Euthanasia

Description There are some cases in which the assisted suicide is an profitable

option according to the state of a health problem. That is showed terminal or strange illnesses so, we had to create one category in order to include that kind of cases.

Detection hints

Examples "I agree with Euthanasia"

8. Mysticity

Mysticity

Description Writing it can be a risk or a protective factor but for our goal (to have a representative corpus of suicide) this kind of stories it can be a way of express things that in other ways It couldn't be possible.

6

Also, we include religion facts.

Detection hints The type of texts included express, in a fantastical way, the idea of diying buy it's not mentioned specifically. In this category, we don't have to mark the second property SecondaryMessageType.

Examples "Only I'll be happy When I lived with God"

"She was like an angel and our soul were in touch in that night"

9. Undefined

Undefined

Description In this section we will put the cases that can not be defined in the other categories.

Detection hints If the message is neither suicide nor depressive the AlertLevel is Not

Risk.

Examples "The only solution is to close the eyes, I have a problem"

10. Instigator

Instigator

Description This category include the people that encourage to other people to commit suicide.

Detection hints Imperative forms are very common in this category.

Examples "Try to disappear of this ruin world"

"The life is so bored that It doesn't make sense being alive"

10. Sadness/melancholy

Sadness/melancholy

Description Sadness that lasts (depend on the person) an undetermined period of time, often without any obvious reason.

Detection hints Poetic and without-context texts are suitable here. Also, we can

include texts in which the key words It doesn't belong to another category. We understand that the messages included here only they have in the property AlertLevel, the category No Risk.

Examples "All finished without a beginning"

"I'm broken inside and empty. I need a coffee in a rainy day, hearing the sound of my soul"

11. Citation

Citation

Description When you express a thing in your own terms, you let the people know

with kind of message you can to transmit, but If you don't use your words, you want to express the same idea but the language used is not the same, so, It's not your natural way of expressing ideas.

Detection hints Use of single/double quotes.

In this category, we don't have to mark the second property SecondaryMessageType.

Examples ... " That which does not kill us makes us stronger" (Friedrich

Nietzsche).

12. Depression

Depression

Description Depression It can be a state of mind or a particular disorder that It

combines behaviours such as fatigue, lack of sleep, eating disorders, lack of concentration, etc.

Detection hints Words as everything or nothing make sense here if the context

combine with the topic. Empty indicates something that lasts on time for a long. Here we do not include depression that had been previously diagnosticated. Always, the property AlertLevel will be selected, minimum, with the Possible category if we mark this depression category.

Examples "I don't want to eat, I can't sleep, I don't want to do anything"

"no good for nothing" "I don't find rest during the night, If only I could sleep well"

General Considerations

- We have to assume that the texts showed here belong to specific forums or pages

related with suicide or deep emotions. So in case of doubt, we'll assume the negative way. - If we want to select two categories in one message type, the choice will be: to assign to PrimaryMessageType that the first one available on Gate 8.2, and to select for SecondaryMessageType the other one. - There are texts that can be categorised as "Mental Illness or Disorder" but also with

"Depression". When both classes are suitable for a particular text and it expresses depression signs that hadn't been diagnosed (i.e. not explicitly stated according to our issuer), the categorisation will be as follows: you should assign to PrimaryMessageType the category "Mental Illness or Disorder", and "Depression" should be assigned to SecondaryMessageType. - Adverbs as too, never, often, now, early, tomorrow, etc. Help to be sure the suitable

    category due to the fact that elevates the meaning of the message. - There are some cases that the AlertLevel is suitable for being Possible. Otherwise if

we include a concrete category such as SuicideCuriosity, then the alert changes too. In order not to change the AlertLevel, if AutoProSuicide and Undetermined are suitable options and the treatment it not necessary, we mark SuicideCuriosity.

Summary: External resources to use

Resource Name URL Type

Gate 8.2 http://gate.ac.uk/ Annotation

                                                         tool Gate 8.2 manual

https://gate.ac.uk/sale/tao/split.html

9

Guidelines Annotation Guidelines This file is titled: Guía de anotación del corpus sobre suicidios.pdf.

Urban Dictionary http://www.urbandictionary.com/ Online

Dictionaries

WordReference http://www.wordreference.com/es/

Linguee http://www.linguee.es/espanol-ingles