Arabic based chatbot using Natural Language processing for analyzing and detecting suicidal tendencies in the workplace.

Bachelor Thesis

German University in Cairo

Faculty of Management technology

Business Informatics

Author:

Mohaned Mashaly

Supervisor:

Dr.Mohammed Abdel Megeed

TABLE OF CONTENTS

Page

1 Introduction 1

1. INTRODUCTION

The main focus of this research is to highlight the importance of mental health in the workplace by analyzing and showing the impact of suicidal inclinations on Employee productivity. We are also interested in proposing a solution to detect the mental health and suicidal tendencies of the employees, which can help in the early diagnosis of mental health disorders like depression, anxiety.

First, we will talk briefly about the problem and its complications. Mental health is a global problem. According to [4] suicide was the second leading cause of death for age group ranging between 15-29 in 2015, contributing to 1.5 percent of the world casualties.It is hard for employers or managers to believe that mental disorders can affect employees productivity such as the physical illness can do. Organizations and companies need to become more conscious of their employees mental health status.

Suicidal tendencies are correlated in most patients with mental disorders like depression and anxiety. These mental disorders can drive the patient to commit suicide if he did not receive proper medical treatment. According to [2] American Foundation for suicide prevention reported that suicide factors are classified into three categories: health factors like Mental disorders, physical pain, trauma, and Environmental factors like stressful life events, exposure to constant stress, and historical factors like previous suicide attempts, family history of suicide, childhood abuse. Suicide cases happen because of one or more of these factors.

Mental disorders such as depression, ptsd, anxiety are serious problems that can unfortunately develop to suicidal tendencies, if the patient did not receive proper medical treatment. According to [1] 90 percent of suicidal attempts cases were suffering

from mental disorders, mental health should not be taken lightly because they have colossal impact on the society not only the person suffering from these disorders, in this section we will shed light on the importance of mental health awareness in workplace and how it must be an integral part of the workplace, stakeholders in a certain organisation might think that an employee's mental health are none of their business or interfering in such situations is an exception not the norm, while on the other hand stakeholders are the ones who would lose the most if some employees suffered from such mental disorders, for example if a person who is occupying an important managerial position in a certain organisation was suffering from depression, or any other mental disorder, his/her performance is going to degrade which will also affect the performance of people under his/her supervision which will cost the company a lot of resources and money, these points can be confirmed by a study published by [5], this study confirms that mental disorders in workplace promotes to absenteeism and decrease in productivity level which will negatively affect the organisations.

The solution suggested by this research paper is to implement a chatbot using natural language processing that helps in detecting suicidal tendencies and make it accessible for all employees. The chatbot considers two factors when trying to detect suicidal tendencies. These factors are text which is simply a paragraph or sentence written by the employee and, the variables that will act as features. These features will determine the likelihood of a person committing suicide or having suicidal thoughts. Some of these variables are age group, Gender, income level and educational level. We will explore more variables later as we go on. These variables are keys for constructing a persona to identify the person who has a high probability of having suicidal thoughts or committing suicide.

Sentiment analysis is one of the natural language processing techniques that will help in solving the problem. According to [3] sentiment analysis uses Natural language processing to process the text or words using various techniques to get insights

from the text and classify these insights into one of three categories: positive, negative, neutral, and some models, classify into positive and negative.

Before talking more about NLP and Machine Learning, we will look at how we are going to approach the problem, for example, what are the steps needed to create such solutions, how much data do we need, identifying the user or stalk-holder, the problem causing pain to, in our case, it is employees, do we have available data to reach the desired level of accuracy, do we have the right parameters to solve the problem? all these questions will be answered in the next paragraphs.

The first step to solve any ML/NLP problem is the data, how much data is available, is the available data is verified, should any feature engineering be done to the data before the training phase, this paragraph will discuss the data point view in brief since it will be discussed more deliberately in the potential data-sources section, there's no accurate estimation for the amount of data do we need since it depends on a variety of factors like how hard the problem is, type of algorithms used, usually nonlinear algorithms require more data than linear because of the complex nature the variables exhibits, so we won't be able to assume precisely how much data we need until we define our problem.

After cleaning the data and performing feature engineering on it, we should check if there's a strong correlation between the variables we choose to work which are called independent variables, and the dependent variable that is going to be the degree to this person will commit suicide or not, the variables we are interested in the problem can be classified into five categories, demographic factors like age, gender, ethnicity, race and martial status, and socio-economic factors, Religios factors and biological factors. In this research, we are interested in measuring the correlation between suicide ideation, attempt, and these factors. We will deploy a variety of methods like Pearson correlation or any other correlation algorithms which measures the degree of

correlation between these independent variables and the dependent variable.

Our chatbot will adapt a hybrid model approach, where we will utilize both pattern matching and generative model techniques that mainly utilizes natural language processing techniques and deep learning, we will dive more deep in the chatbot and its architecture and implementation in the chatbot part.

One of the limitations we have is data as the chatbot supports the Arabic language. Finding corpora in Arabic since Arabic is not as developed as English in terms of NLP and sentiment analysis research. Also, Arabic is more complex than English since it has many dialects and types.

Bibliography

[1] Louise Brådvik. *Suicide risk and mental disorders*. 2018.

[2] Shaoxiong Ji et al. "Suicidal ideation detection: A review of machine learning methods and applications". In: *IEEE Transactions on Computational Social Systems* (2020).

[3] Shrija Madhu. "An approach to analyze suicidal tendency in blogs and tweets using Sentiment Analysis". In: *Int. J. Sci. Res. Comput. Sci. Eng.* 6.4 (2018), pp. 34–36.

[4] World Health Organization et al. *Depression and other common mental disorders: global health estimates*. Tech. rep. World Health Organization, 2017.

[5] T Rajgopal. "Mental well-being at the workplace". In: *Indian journal of occupational and environmental medicine* 14.3 (2010), p. 63.

all