EAST WEST UNIVERSITY

Course Code: CSE475

Course Name: <u>Machine Learning</u>

Spring-2020

Section: 02

Project Title: Suicide Attempt Prediction

Group Members' Name:

1) MD. Serajus Salekin Khan

   ID: 2017-1-60-006

2) Tanvir Hasan

   ID: 2017-1-60-102

3) Rajib Hawladar

   ID: 2017-1-60-104

Department of Computer Science and Engineering

Date: 03-04-2020

Project Submission Declaration

(Student 1)

ID	2017-1-60-006
Name	MD. Serajus Salekin Khan
My contribution in doing this project (in percentage) in the group	33.33%
Description of my contribution in this project in the group	In this project I have implemented Random Forest Classifier using scikit-learn and showed accuracy of this classifier for our dataset. I have tried to visualize decision trees which helps to understand clearly.

(Student 2)

ID	2017-1-60-102
Name	Tanvir Hasan
My contribution in doing this project (in percentage) in the group	33.33%
Description of my contribution in this project in the group	I did some research to establish the model for the project. I brought some new ideas which made this project more efficient. I helped in some parts of the code and debugged in some cases when I found any error.

(Student 3)

ID	2017-1-60-104
Name	Rajib Hawladar
My contribution in doing this project (in percentage) in the group	33.33%
Description of my contribution in this project in the group	I helped my team members to established the whole project and also followed our planned instructions when researching for the project. I also helped my members to debug the code.

Introduction of the project:

The purpose of this study is to conduct a step-by-step analysis of the data set and identify factors that affect suicide attempts among respondents.

Objectives:

To determine whether a particular person will attempt to suicide or not based on previous record.

Motivations:

At the present era, many peoples are committing suicide for various reason. According to a report by the World Health Organization 19,697 people committed suicide in Bangladesh in 2011. According to Police Headquarters 11,095 people committed suicide in Bangladesh in 2017. According to a report by Shaheed Suhrawardy Medical College Hospital, Dhaka, published in 2010, around 6,500,000 people of Bangladesh are prone to suicide. The rate is 128.08 people per 100,000 commit suicide in Bangladesh every year [1]. If we can predict for a particular person that he/she will commit for suicide or not then this will be helpful to do something for these peoples. Also, we will be informed about which categorized peoples are in this list.

Model Description:

For this project we have selected Random Forest Classifier to predict for suicide attempts. We have implemented RF classifier using scikit-learn and visualized some decision trees. Also, we have tried with DT classifier and then compared both of two classifier which resulted that RF classifier has better accuracy. Besides this, RF classifier will gather better prediction as it compared multiple decision trees output.

Pseudocode:

1. Select randomly features from dataset
2. For each item in features
   a. Calculate information gain
   b. Select node which has highest information gain
   c. Split node into sub-nodes
   d. Repeat step a,b and c to construct the tree until reach minimum number of samples required to split.
3. Repeat steps 1,2 for N times to build forest of N trees
4. Compare all of the trees and predict output

Reference:

[1] https://en.m.wikipedia.org/wiki/Suicide_in_Bangladesh