**Course: Computer Science Module: 6006CEM Machine Learning and Related Applications**

**Student Name: Saad Iftikhar**

**Student ID: 9789180**

**Coventry GitHub Repository URL** or **Coventry OneDrive URL** (mandatory):

< <https://github.coventry.ac.uk/iftikhars/9789180-SI-s1>>

**Table of contents**

[Table of figures 2](#_Toc84718354)

[Introduction 3](#_Toc84718355)

[**Problem statement** : Predicting the Suicide rates of males and females at different ages 3](#_Toc84718356)

[Motivation: 3](#_Toc84718357)

[Related work 3](#_Toc84718358)

[Existing approaches or methods and their results 3](#_Toc84718359)

[Similarities and differences between this work and the existing works 3](#_Toc84718360)

[Dataset 5](#_Toc84718361)

[Pre-processing 6](#_Toc84718362)

[Implementation 7](#_Toc84718363)

[Results 8](#_Toc84718364)

[Evaluation 9](#_Toc84718365)

[Conclusion 10](#_Toc84718366)

[References 11](#_Toc84718367)

[Appendix A 11](#_Toc84718368)

[Appendix B 13](#_Toc84718369)

# Table of figures

# Introduction

## **Problem statement** : Suicide risk prediction

## Motivation:

In research from Brådvik (2018, p. 1) the suicide mortality rate is 1.4 percent of all global deaths. Biological, psychological, clinical, social, and environmental variables all have a role in the development of suicide risk. Because several risk variables are involved in defining a person's risk of suicide, determining a person's risk of suicide is difficult. Using computer testing and genetic screening, for example, to improve risk assessment is a topic of continuing study (Turecki et al., 2019, pp. 1).

This project will focus on implementing different machine learning algorithms and techniques to predict a person’s likelihood of committing suicide. This study will use country, age , gender, population, internet usage per percent Population, total labour force, cellular subscription per 100 people, GDP per year, GDP per capital, Expense’s percent of GDP, physicians per 1000 people and population of refugees by country or territory of origin to determine the probability persons to comic suicide. To lower the number of suicide fatalities, prevention is essential (Turecki et al., 2019, pp. 1). By predicting high risk variables that increase a person’s chance of suicide.

## Related work

## Existing approaches or methods and their results

## Similarities and differences between this work and the existing works

# Dataset

Removing null values 30 % rule

<https://analyticsindiamag.com/5-ways-handle-missing-values-machine-learning-datasets/>

30 % of our data set = **4,533-1**

Our data that will be removed = 4,044

Data removal is best option

# Pre-processing

# Implementation

# Results

# Evaluation

# Conclusion

# References

Turecki, G., Brent, D. A., Gunnell, D., O’Connor, R. C., Oquendo, M. A., Pirkis, J., & Stanley, B. H. (2019). Suicide and suicide risk. *Nature reviews Disease primers*, *5*(1), 1-22.

Brådvik, L. (2018). Suicide Risk and Mental Disorders. *International Journal of Environmental Research and Public Health*, *15*(9), 1–4. https://doi.org/10.3390/ijerph15092028

Belsher, B. E., Smolenski, D. J., Pruitt, L. D., Bush, N. E., Beech, E. H., Workman, D. E., Morgan, R. L., Evatt, D. P., Tucker, J., & Skopp, N. A. (2019). Prediction Models for Suicide Attempts and Deaths. *JAMA Psychiatry*, *76*(6), 642–651. https://doi.org/10.1001/jamapsychiatry.2019.0174

# Appendix A

< A suggested checklist for you, for full details please refer to the coursework brief >

1. The following naming convention is used for the Coventry GitHub Repository and Coventry OneDrive

StudentID-Initials-s1

For example, for a student Alan Turing whose student ID was 1234567, it should be

1234567-AT-s1

Failing to follow the naming convention may delay the release of marks and feedback for your coursework.

1. **Coventry** GitHub Repository URL **or** **Coventry** OneDrive URL: added to the top of this report
   1. Coventry GitHub Repository includes:

* URL of the selected dataset(s) included in README
* The original selected dataset(s)
* Source-code (.ipynb)
* Demonstration video (.mp4)
  1. Coventry OneDrive folder includes:
* URL of the selected dataset(s) included in a separated text file
* The original selected dataset(s)
* Source-code (.ipynb)
* Demonstration video (.mp4)

1. Source-code added **as text** in Appendix B (below)
2. Submission in the form of a **Word** document. *\*\*Other format is not accepted.*

# Appendix B

< **Replace** this instruction with all the Programming Code for the coursework.

Make sure you have highlighted and referenced any code not written by you >

< **DO NOT** use screenshots of your code here. Your code should be presented **as text**.

There are many good tools to help you format your code such as <http://hilite.me> >

< You can select and copy **all code at once** in a notebook by:

1. Graphical user interface, text, application

   Description automatically generatedclicking in any cell of the notebook, the cell will be highlighted in green as below
2. Graphical user interface, text, application

   Description automatically generatedthen press Esc on your keyboard, the selected cell will be highlighted in blue as below
3. now you can Ctrl+A to select all cells of the notebook

Graphical user interface, text, application

Description automatically generated

1. and then copy and paste as normal to some tool, such as hilite.me above, make sure you select a correct language (Python), then click Hightlight

Graphical user interface, text, application

Description automatically generated

1. you now can select the text in the Preview and copy and paste it over to this Appendix

Graphical user interface, text, application, email

Description automatically generated

1. finally, remember to remove all text in this instruction for this Appendix >