Abstract—According to the 2021 Report from the World Health Organization (WHO), more than 700,000 people have taken their life. Suicide can be prevented but so far most of the efforts to do so have fallen short. However, the use of machine learning and artificial intelligence offers new opportunities to increase the accuracy level of prediction and aid the goal of suicide prevention. This paper reviews literature concerning the machine learning methods used to help identify various risk factors and help prevent suicide. This paper also presents our research and analysis findings which were used to identify various suicide risk factors and additional analysis of whether there are any correlations or variations in the risk factors from pre- and post-pandemic datasets regarding suicide rates. This is especially important during times of high stress, such as a worldwide pandemic and quarantine. The dataset(s) obtained from WHO suggest that high levels of risk factor identification are possible and This paper and the analysis serve as supporting research and guide to aid in the continued ambitious goal of suicide prevention worldwide.

Index Terms—Suicide Prevention, Python, R, Post-COVID, Quarantine, Mental Health, Volume Analysis, Algorithm, Suicide, Machine Learning, Risk Factors, AI

I. INTRODUCTION

Suicide is a complex global public health problem. According to the World Health Organization (WHO), every year, more than 700,000 people take their own life and there are many more who attempt to do so. Only 38 countries report having a national strategy for suicide prevention. Suicide and suicide attempts have a devastating ripple effect that impacts the community, families, friends, coworkers, and societies. The COVID pandemic has created a new sense of urgency to analyze and engage all countries and communities in aiding the prevention and finding an achievable and sustainable solution to the growing problem. The quarantine lockdowns and restrictions over the past couple of years, may have led to a heightened sense of loneliness, anxiety, fear, and depression in many people across the globe. Suicides are preventable and a lot can be done to improve the international strategy efforts for suicide prevention. Our mission is to identify the key drivers of suicide, without underscoring that suicide is determined by multiple factors. Advanced Brain MRI techniques can show the signs of depression [1]. Also Serum Expression Level of High-Mobility Group Box 1 (HMGB1) in Multiple Sclerosis Patients after COVID-19 can show a Relationship with Physical and Psychological Status [2].

Early identification of individuals who are at higher risk of suicide is vital in preventing suicide. According to multiple studies, machine learning and the advantage of Big Data are becoming the new promising approach in this “early identification” objective.

When understanding the motive behind suicide there are many factors that must be noted:

- Geographic Location
- Mental Health Conditions
- Life Stressors
- Relationship Problems
- Financial Problems
- Cultural/Societal Expectations
- Current Events

II. BACKGROUND

Death by suicide is a very complex issue that causes torment and distress for hundreds of thousands of people globally. But with timely intervention and prediction, suicides can be prevented. This is especially important for times of high stress, economic uncertainty, pandemic or natural disasters, and other stress factors that can have a serious effect on mental health and morale.

The WHO compiles and disseminates data on death and morbidity annually. This data is reported by all participating “Member States”, according to the WHO mandates. These Member States are foreign countries, which at first started out as only 11 back in 1950, right after the inception of WHO, then growing to 74 in the year 1985, and reaching 183 in 2019. Thus, this is one of the most complete and thorough mortality data banks that are publicly available, and reported suicide deaths are an integral part of this data bank. However, although this global data bank is from a reputable organization and with specific data classification requirements, it still offers some challenges. Only about 80 Member States provide good-quality, reliable data that can be used to estimate the risk factors associated with suicide deaths. “This problem of poor-quality mortality data is not unique to suicide but given the sensitivity of suicide – and the illegality of suicidal behavior in some countries – it is likely that under-reporting and misclassification are greater problems for suicide than for most other causes of death.” [3]. In addition, the official statistics relating to attempted suicide are even more skewed, which makes it difficult to
analyze and correlate national global suicide trends to suicide attempt trends. “...suicide may be hidden and underreported for several reasons, e.g. as a result of prevailing social or religious attitudes. In some places, it is believed that suicide is underreported by a percentage between 20% and 100%.” [3]. AI models can be used to analyze COVID-19 data models [4]–[9] [10]–[26].

In spite of these dataset collection and classification challenges, the WHO mortality dataset is one of the most comprehensive and reliable data banks in the world. An important role of big data gained from sources such as the World Health Organization (WHO), Electronic Health Records (EHR) [27], or social media regarding suicide, is supporting the objective of prevention. “The potential in this area is tremendous. The suicidal phenotype is characterized by extreme heterogeneity, and potentially suicidal individuals are very often excluded from any clinical trials. Big data could help by combining very complex and large data samples to detect patterns, signaling suicidal inclinations [28] With such a powerful tool as big data available research teams and analysts should be able to predict high risk factors with a high degree of certainty. So far there have been cases where scientists were successful in implementing tools to survey and predict risk, with 91% accuracy [29], but unfortunately this level of accuracy is only due to the smaller and more controlled groups, such as adolescents in the state of Utah. The challenge still exists, not only for a nationwide population for all age groups, but also a worldwide population and especially during high stress time periods caused by economic downturns.

It should also be recognized that as big data gains a more prominent role in psychiatry, issues of governance and security will need to be clearly considered, and that there must be a thorough and open public dialogue on ethical issues.” [28] Even still, big data in the world of psychiatry can offer significant benefits to help with not only treatment of the patients but also comprehension of their disorders and early diagnosis and prevention of suicide or self-harm. Although, it’s important to keep in mind that machine learning, predictive algorithms, and big data needs to also run hand in hand with additional efforts and factors to be fully effective in identifying risk factors and preventing suicide. Medical and psychiatric professionals need to recognize the potential of big data technological advancements, while at the same time also being careful to maintain a high standard of professionalism, as well as the traditional doctor-patient relationship which is based on trust and confidentiality.

III. RELATED WORK

As mentioned in the introduction as well as in the background, there are various reasons, risk factors, and warning signs that lead to suicide and suicidal thoughts. Many studies have been conducted over the years that focus on the various factors in an attempt to better understand and hopefully lower suicidal rates.

In this paper, we reviewed six literatures that were published between 2012 and 2020. Of the two that were published pre COVID-19, each focused on suicide rates in the USA during economic recession during a different period. The third pre COVID-19 paper discussed the national cost of suicides and suicide attempts in the United States in 2013. The remaining three literatures observed trends and discussed whether the COVID-19 pandemic resulted in suicide rate increases.

In the report titled, “Increase in state suicide rates in the USA during economic recession” by Reeves et al., it mentions evidence from European countries that show a rise in suicides during economic recessions. Among the worst being Greece, were suicides have risen more than 60% since 2007. Using data on suicide mortality rates from 1999 to 2010 from the Centers for Disease Control and Prevention, along with ‘unemployment’ data from the Bureau of Labor Statistics, the authors extended their previous analyses of recessions and suicides in Europe to assess trends in all 50 US states. Their findings showed that in the years before the recession (from 1999 to 2007), the suicide mortality rate in the USA were rising on average at a rate of 0.12 per 100K per year. During the recession period (2008-2010), the suicide rate increased at a rate of 0.51 deaths per 100K per year. This difference in rate corresponds to an additional 1580 suicides per year [30].

Another study by Harper et al. explores suicide mortality rate in the USA over a 30 year period. Prior to this study, there were several studies that suggested strong associations between economic downturns and suicide mortality. This study aimed to provide more robust evidence by using a quasi-experimental design. The researchers analyzed 955K suicides that occurred in the USA from 1980 to 2010 and used a broad index of economic activity in each US state to measure economic conditions. Based on the quasi-experimental and fixed-effects design, and after accounting for secular trends, seasonality, and unmeasured fixed characteristics of states, they found that an economic downturn in magnitude to the 2007 Great Recession increased suicide mortality by 0.14 deaths per 100K population or around 350 deaths. The effects were also stronger for men than women and for those with less than 12 years of education [31].

The third paper reviewed addresses economic costs of suicides in the United States in 2013. Understanding that suicide would not be eliminated from our society anytime, we felt it was important to review some literature on its economic cost to better put things in perspective. According to the authors, there were three previous studies using different approaches to address the matter. For their study, to seek to improve on the previous by addressing the increase in the number of suicides since those publications by incorporating adjustments for underreporting and using additional data. Shepard et al. paper concluded that the national cost of suicides and suicide attempts in the United States in 2013 was $58.4 billion based on reported numbers. Lost productivity represented most (97.1%) of the cost. When adjusted for under-reporting, it increased the total cost to $93.5 billion [32].
Similar to understanding how economic downturns affected suicide mortality rates in the first two reviewed literature, we reviewed a few more literatures that were published during the COVID-19 pandemic to observe how and whether the pandemic affected the rate.

In the paper titled “Suicide risk and prevention during the COVID-19 pandemic”, Gunnell et al. seek to understand how suicide is likely to become a more pressing concern as the pandemic spreads and its effects on the general population, the economy, and vulnerable groups.

Deterioration in population mental was one of several factors that underpinned the concern that suicide rates may increase during the COVID-19 pandemic. Based on some widely reported studies modelling the effects of the pandemic on suicide rates predicted increases ranging from 1% to 145% [33]. In the study titled “Trends in suicide during covid-19 pandemic”, the authors tracked and reviewed relevant studies for a living systematic review. In the early months of the pandemic, reports suggest either no rise in suicide rates (Massachusetts, USA; Victoria, Australia; England) or a fall (Japan, Norway) in high income countries. Not much is known for low income countries.

IV. PROBLEM DESCRIPTION

The main purpose of this analysis is to identify risk factors in the hopes of mitigating suicide rates and help combat the worldwide phenomena. Identifying these drivers could help communities implement proactive protective practices and decrease the rate of suicide.

We are approaching this problem by identifying which factors effects the trigger of suicide in each individual like their country GDP or their Income or which place they are from and so on. We will be cleaning the data using R/Python and use scatter plot and ggplot packages to form comparison between the data from 1986 to 2016, and data from 2019 to current.

Questions we aim to answer:

- Is a prior suicide attempt a clear risk factor and indicator of another attempt?
- Is there a correlation between a country’s GDP and suicide rate?
- Is there a correlation between gender and suicide rates?
- Is there a correlation between the year and suicide events? If so, is this due to any global events at the time?
- Do the countries that invest more in mental health see a lower suicide rate?
- Do countries that are more impacted by the opioid epidemic see a higher rate of suicide?
- On average how many combined factors does a suicide victim have?
- Do the countries that have easier access to suicide methods (substances, firearms, etc.) have a higher suicide rate?

V. METHODOLOGY

A. How we collected the data

The team used secondary data gathered from the “1986 - 2016 Suicide Rates” [34] dataset which we obtained from Kaggle. This dataset is a combination of four datasets, linked by time and place to help identify risk factors worldwide. The four datasets were:


B. Preliminary Results

- Is a prior suicide attempt a clear risk factor and indicator of another attempt?
- Is there a correlation between a county’s GDP and suicide rate?
- Is there a correlation between gender and suicide rates?
- Is there a correlation between the year and suicide events? If so, is this due to any global events at the time?
- Do the countries that invest more in mental health see a lower suicide rate?
- Do countries that are more impacted by the opioid epidemic see a higher rate of suicide?
• On average how many combined factors does a suicide victim have?
• Do the countries that have easier access to suicide methods (substances, firearms, etc.) have a higher suicide rate?

VI. CHALLENGES

Due to limitation of recent and extended periods of worldwide suicidal data from credible sources like WHO or CDC (who only publish suicide figures up to 2019), it is difficult to make clear comparisons of pre and post-pandemic suicidal rates and factors. From our review of recent articles and publication, most do not draw a conclusion that suicidal rates increased during the pandemic, but state that previous pandemics have been associated with increases in suicide rate. The U.S. reported an increase in suicides during the Spanish Flu (1918-19) [35]. Africa did experience an increase in suicides during the Ebola epidemic spanning 2013 to 2016 [36]. Hong Kong did observe an increase in elderly suicides during 2003 SARS outbreak [37]. One common denominator among these publications is that they were all completed years after the epidemic/outbreak or after multiple years of data was available for analysis. There are multiple factors that lead researchers to believe that the current COVID-19 pandemic will share similarities with past events. By the end of 2020, over 76 million people worldwide were infected by the COVID virus, SARS-CoV-2 [38]. Shutdown of businesses and business activities because of lockdown measures affected those living in poverty, relied on hourly wage positions, and many that relied on government programs. As mentioned in the literature review that examined US suicide rates during 2008 as well as a publication by Oyesanya et al. [39], both conclude that economic stress has been associated with higher suicide rates. Another effect of the COVID lockdown is social isolation. Studies such as the one by Christensen et al. have documented that social isolation is associated with increased suicidal thoughts. A 2016 study by VanderWeele et al. show that participating in religious communities is associated with lower suicide rates, however, with churches and community centers closed during lockdowns, social isolation possibly increased suicidal thoughts. One factor that differs greatly from the current pandemic to previous epidemic/outbreaks is the amount of media coverage. With 24-7 news coverage, social media outlets, unlimited SMS service, etc., it’s possible that anyone with preexisting mental health conditions will experience intensified anxiety and fear.

VII. CONCLUSION

Similar to the report by John et al. that examined suicide rates during the earlier stage of COVID (lockdown, stay-at home period), other reports by Faust et al., Applyby et al., and Qin et al., all [40], [41], concluded that suicide rate for the same period compared to a year ago were similar or did not show significant difference from previous years. Despite these similar conclusions, we find that it is too early for a study to compare pre and post pandemic suicidal rates due to the lack of data and that the pandemic is still ongoing.

VIII. FUTURE WORK

For a follow on study to be more conclusive, it is best to be done after WHO declares the end of the pandemic and when most countries’ citizens have returned to a lifestyle similar to pre pandemic. In the follow on study, instead of focusing on worldwide figures, we suggest to examine certain groups that may be more vulnerable to the effects of the pandemic and experienced increased suicide rates.

• Unemployed: In early 2020, the International Labour Organization (ILO) predicted the pandemic cost 25 million jobs worldwide. Studies of the Great Recession in the early 2000s found an increase of suicide risk by 20-30% between 2000 and 2011 with a peak during 2008 [42].
• Mentally ill: individuals with preexisting mental health conditions were likely affected by interruption in treatment, and experienced increased isolation and intensified anxiety and fear due to the pandemic [43].
• Healthcare workers: In the early stage of the pandemic, medical staff have reported increased hopelessness, guilt, and insomnia. All which can increase the risk for suicide [44].
• Racial minorities: Racial minorities who owned small businesses or worked at hourly waged positions were affected more by lockdowns and shutdown of business activities.
• Youth: Preliminary data from England suggest that child suicide deaths may have increased during the early stages of lockdown, possibly due to disruptions to education, outside activities, and support services [45].
• Elderly: Similar to youth, elderly suffered greatly from social disconnectedness during the pandemic. Social and self-isolation affects elderly who do not have close family and friends, or who have decreased literacy in or access to digital resources [46].

[47]–[52]

REFERENCES


M. Heidari, “Nlp approach for social media bot detection/fake identity detection” to increase security and trust in online platforms,” 2022.


