Kalpa Publications in Computing

Volume 19, 2024, Pages 53–60

Proceedings of 6th International Conference
on Smart Systems and Inventive Technology

Suicide Rate Analysis using Extra Trees Algorithm

Ignisha Rajathi G[1], Sneharajalakshmi P[2], Midhun Kumar V N[2], Vikash Kumar S[2]

[1]Associate Professor, Department of Computer Science and Business Systems, Sri Krishna College of Engineering and Technology, Coimbatore-641008. ignishaelton@gmail.com

[2]UG Scholar, Department of Computer Science and Business Systems, Sri Krishna College of Engineering and Technology, Coimbatore-641008.

ignishaelton@gmail.com, sneharajalakshmi@gmail.com

Abstract

In recent times, suicide has emerged as a significant global public health concern, causing profound harm to individuals, families, and communities. Despite intensified efforts to address mental health challenges, suicide rates continue to surge in numerous regions worldwide. To effectively tackle this escalating problem, there is an urgent requirement for a thorough examination and understanding of the factors driving suicide occurrences. The primary objective is to conduct a comprehensive analysis of suicide rates utilizing advanced analytics techniques. This analysis endeavors to uncover the intricate network of socio-economic, demographic, and psychological factors shaping suicide trends across various countries and regions. Through meticulous investigation and interpretation of these factors, this research aims to gain insights that can inform targeted interventions and policies aimed at mitigating the prevalence of suicide and promoting mental well-being on a global scale.

Keyword: Suicide Rate Analysis, Predictive Model, Machine Learning, Data Analysis, Mental Health, Suicide Prevention, extra trees regressor.

1 Introduction

This research embarks on a comprehensive exploration of suicide, recognizing its intricate nature influenced by individual, societal, and economic dynamics. The significant global variability in suicide rates among diverse countries propels the mission to conduct an in-depth analysis and construct a predictive model. Harnessing the power of machine learning, the approach involves predicting global suicide rates using a vast dataset spanning over a hundred countries from 1987 to 2014 with a total of 27000+ rows of data that roots from United Nations Development Program (2018), Human development index (HDI), World Bank. (2018), World development indicators: GDP (current US$) by

country:1985 to 2016, Suicide in the Twenty-First Century and Suicide in the Twenty-First Century that was collected from Kaggle as an amalgamation. The dataset incorporates critical variables such as country name, age, sex, population, Gross Domestic Product (GDP), and generation. Each of these factors contributes uniquely to the overall suicide rate, highlighting the nuanced interplay of socio-economic and demographic elements. This endeavor seeks not only to understand the complex nature of suicide but also to contribute valuable insights for potential preventive measures on a global scale.

2 Literature Survey

The research by Okada. M, et al., [1] looks at the effect of regional unemployment rates, GDP per capita, and financial assistance for suicide prevention programs on suicide mortality in Japan. This study looks at the relationship between these variables and suicide rates in various geographic areas using official statistics data. The findings indicate that higher unemployment rates and lower GDP per capita are associated with an increase in suicide mortality. Furthermore, regions with more financial resources for suicide prevention programs have lower suicide rates. These findings highlight the relevance of economic stability and focused suicide prevention programs in lowering regional suicide rates in Japan. In this study by Matsumoto, R and others [2] focuses on working-age individuals in Japan from 2007 to 2022 are examined for patterns in suicide mortality and reasons. Using data from a variety of sources, the study investigates trends in suicide rates and the causes of suicides during this era. The results show variations in the rates of suicidal death over time, emphasizing the necessity of ongoing surveillance and countermeasures. The study also investigates the motivations for suicides, offering light on the elements that contribute to suicidal behaviour among Japan's working-age population. These findings can help to inform targeted treatments and support services to address the complicated issue of suicide in this demographic.

The study by Blasco-Fontecilla H, et al., [3] looks into how economic cycles have affected suicide rates globally over the course of three decades, taking into account regional variances in terms of development. Using a mixed-effect model, the study examines data to determine how economic variations affect suicide rates around the world. The findings indicate subtle variations depending on a nation's economic level, with differing consequences on suicide patterns in periods of economic recession and expansion. The results underscore the intricate connection between economic cycles and suicide rates, stressing the need for customized interventions and support networks that take into account the economic circumstances of various areas in order to lessen the negative effects on mental health outcomes. The study done by Kato, R and others [4] investigates how financial assistance might lower the suicide death rate. Using data analysis and statistical methodologies, the study analyses if increasing financial aid is associated with lower suicide rates. The findings point to a possible relationship between reduced suicide mortality and financial support, suggesting that social welfare initiatives and sufficient financial resources may help avoid suicides. These results highlight the necessity of ongoing research and funding for social welfare programs in order to enhance public health outcomes. They also emphasize the significance of support networks and stable economic conditions in addressing mental health issues and lowering suicide rates.

The impact of economic crises on suicide rates is examined in this research by Miret, M., et al., [5] also explores the connection between shifting suicide trends and economic downturns. Using data and analysis, the study analyses how economic crises affect suicide behaviour in various communities. Results show a complex interaction between economic downturns and suicide rates, with the former frequently causing the latter, especially in disadvantaged populations. The study emphasizes the importance of comprehensive mental health care and crisis responses during economic downturns to

reduce the risk of suicide. It is imperative to comprehend these dynamics in order to put into practice focused methods that treat the mental health issues that are made worse by economic downturns. The impact of regional features on suicidal behavior is the main focus of the study's investigation done by Oka, M., and others [6], of the relationship between geography and suicide rates. The research investigates relationships between regional characteristics and suicide rates in various locations through statistical analysis and data inspection. The results indicate that a number of regional factors, including socioeconomic status, environmental factors, and the availability of mental health treatments, are important in influencing the suicide rate. Comprehending the impact of geography on mental health issues and the reduction of suicide rates in certain areas necessitates the development of tailored interventions and support networks.

In order to forecast possible suicide instances in Aguascalientes, Mexico, the study by Ramirez López, et al.,[7] provides a geographical analysis utilizing Empirical Bayesian Kriging (EBK). Geographical data and statistical modelling are employed to determine high-risk locations for suicide by taking into account a number of variables, including socioeconomic status, mental health care accessibility, and demographic data. The study's goal is to shed light on regional patterns of suicidal behaviour, allowing for more targeted intervention tactics and resource allocation in areas with high suicide rates. The research enables the development of proactive strategies for mental health support and suicide prevention at the local level by fusing geospatial analysis with predictive modelling techniques. The main objective of the study done by Swain PK and others [8] is to use empirical research to forecast suicide rates in India. It makes predictions about future trends in suicide rates based on a range of sociodemographic and economic parameters by using statistical methodologies and data analysis. The study's goal is to create models that can properly predict suicide rates in various regions of India by analysing historical data and patterns. Insights from the research can guide focused preventative tactics and mental health interventions to reduce the risk of suicide. It also advances our understanding of the many variables impacting suicidal behaviour in the nation.

The research by Kim, J. W. and others [9] examines suicide rates in South Korea based on age, gender, and marital status. The research analyses trends and variations in suicide rates over time among various demographic groups using statistical techniques and data analysis. The findings show differences in suicide rates among age groups, genders, and marital statuses, indicating distinct patterns and risks. It is important to address unique vulnerabilities and lower suicide rates among various demographic categories. By understanding these dynamics, one can get insights into the intricacies of suicidal behaviour in South Korea and emphasize the need of customized preventative measures and targeted treatments. The study done by Dedić, G. J., & Tepšić-Ostojić, V., [10] looks at how the suicide rates in Serbia changed by gender between 2016 and 2020. The study uses data analysis and statistical approaches to analyse patterns and variances in suicide rates among men and women throughout this time period. Conclusions highlight higher suicide rates in men than in women and point to possible contributing variables. They also show clear patterns and risk factors. Comprehending the disparities in suicide rates between genders is imperative in order to formulate focused preventive approaches and interventions that cater to the distinct difficulties and susceptibilities encountered by men and women in Serbia. The ultimate goal is to lower the suicide rate and enhance mental health consequences.

The study by Innamorati M, et al., [11] looks at differences in the suicide rates among the senior citizens of the European Union between 1980 and 2006. The study uses data analysis and macroeconomic elements to analyse trends and the impact of various socioeconomic variables on suicide rates. Results show differences in the rates of suicide among senior citizens in several European nations, underscoring the impact of social and economic variables. Understanding these disparities is critical for establishing targeted interventions and policies to reduce suicide rates among the elderly while also addressing the socioeconomic factors that contribute to suicidal behaviour in this

demographic. In the framework of neural networks, the study by M. Fernández-Delgado and others [12] provides a thorough experimental survey of different regression approaches. The study uses rigorous experimentation to assess the efficacy of several regression algorithms across a variety of datasets. To identify the benefits and drawbacks of each approach, important factors including accuracy, robustness, and computational efficiency are examine. The results offer insightful information about the performance of regression models inside neural network frameworks, which helps practitioners and researchers choose the right methods for the jobs at hand. This comprehensive survey advances our understanding and optimization of regression algorithms in neural networks.

The impact of public service announcements (PSAs) on attitudes toward requesting help is examined in this research by Klimes-Dougan Bonnie and others [13]. The study investigates how varied messages in PSAs influence people's willingness to seek help for mental health concerns related to suicide ideation. Results imply that views toward asking for help are strongly impacted by the message's substance and framing. PSAs that emphasize hope, coping skills, and available support services are more effective at encouraging help-seeking behaviour than those that only focus on risk factors or warning signs. These findings emphasize the necessity of properly developing messages in suicide prevention efforts in order to favourably influence attitudes and increase help-seeking behaviour. In this research by Shu-Sen Chang and others [14], we investigate whether rising suicide rates in East/Southeast Asian nations are related to the 1997–1998 economic crisis. Based on a time-trend analysis of Japan, Hong Kong, South Korea, Taiwan, Singapore, and Thailand, the study looks into the relationship between suicide rates and economic downturns. Results imply that although there were some variations in suicide rates during the crisis, there were national differences in the association between economic downturns and suicide rates. The prevalence of suicide was also shaped by various environmental factors, cultural variations, and social support networks. The study emphasizes the complexities of factors influencing suicide rates during economic crises in various locations.

The research by Kyla Thomas and others [15] investigates the suicide rates in Wales and England between 1861 and 2007 through a thorough time-trend analysis. The study investigates long-term trends and variations in suicide rates over more than a century using statistical techniques and data analysis. The findings show considerable swings and trends in suicide rates throughout time periods, with notable increases in particular decades. Insights on the changing terrain of suicide epidemiology in England and Wales over the studied period are also provided by the research, which also reveals differences in suicide rates by age, gender, and techniques utilized. This report on a study by Turecki, G, et al., [16] gives a thorough examination of suicide and risk factors. It covers a number of topics, including the epidemiology of suicide, protective factors, and risk factors such as genetic, environmental, and psychological influences. The writers address the intricate interaction of biological, social, and environmental elements that influence suicide risk. In order to prevent suicide, the research also explores the significance of comprehending and addressing mental health concerns, early risk factor identification, and successful intervention measures. The detailed analysis is an invaluable resource for healthcare practitioners, politicians, and suicide prevention researchers alike. This research by S. Mary Fabiola and others [17] investigates how artificial intelligence (AI) can be used to mitigate the psychological effects of the COVID-19 pandemic. It examines how AI technology can be utilized to address mental health issues created by the epidemic, including as stress, anxiety, and depression. The promise of AI in delivering individualized mental health support, early mental health issue detection, and the creation of successful intervention strategies is highlighted by the authors. The study uses AI tools and methodologies to provide unique solutions for assisting individuals' psychological well-being during times of crisis, such as the COVID-19 pandemic.

3 Prediction System

The proposed system is an anticipatory model designed for the analysis of suicide rates. It evaluates various factors such as country, year, individual gender, population, gross domestic product (GDP), GDP per capita, and the generation to which the individual belongs. Although suicides are inherently individual acts, observed correlations among these attributes allowed for a refined analysis, particularly concerning the gender of the individuals. Notably, significant correlations were identified between factors such as the year and the GDP of the country, influencing the prevalence of suicides on a considerable scale. The workflow of the model states the steps involved (figure 1).

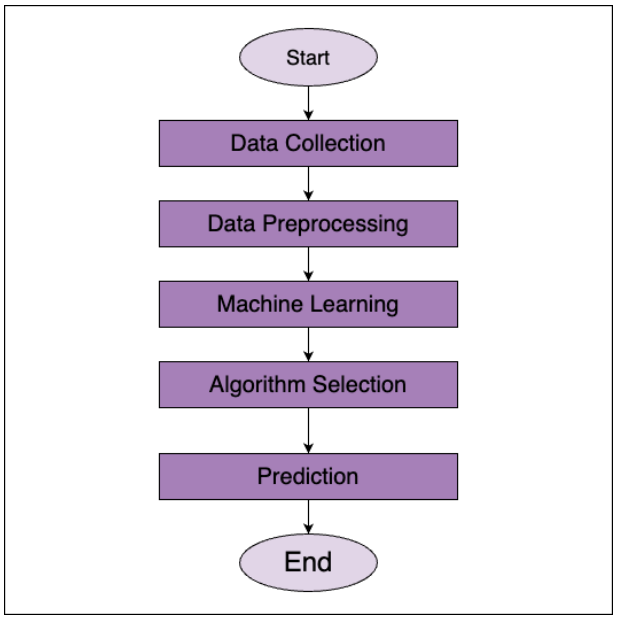

Figure 1: flowchart

To provide precision in the data, the dataset spans from 1987 to 2014, encompassing over 27,000 rows of data from 100 different countries. Data pre-processing is conducted to ascertain correlations among attributes. The primary objective of this model is to predict the number of suicides in a given year based on the considered attributes. The genesis of this research stemmed from the substantial disparity in suicide rates among individuals from different generations, affected by the cost of living prompting a need for comprehensive analysis. To improve the performance of the model the data undergoes feature engineering, hyperparameter tuning, ensemble learning, regularization and most importantly cross validation. Various factors like data nature, dataset size, problem complexity, and predictive accuracy guided model selection. Common algorithms (logistic regression, decision trees, random forests, gradient boosting) were compared, with Extra-trees showing superior accuracy and thus chosen for prediction.

4 Result and Discussion

The proposed predictive model for analysing suicide rates utilizes a wide array of attributes such as country, year, gender, population, GDP, GDP per capita, and generation. The data from United Nations Development Program [18], World Bank [19], Suicide in the Twenty-First Century [20] and World

Health Organization [21] has been amalgamated and retrieved from Suicide Rates Overview 1985 to 2016 [22]. By employing thorough data analysis and machine learning methods, the model identifies correlations between these attributes and suicide rates, facilitating the creation of a robust predictive tool (figure 2). Pre-processing is achieved by analyzing data for null values, categorical variables, and duplicates. Steps included dropping nulls, removing duplicates, encoding categorical variables, scaling numerical data, and performing train-test split for modeling. The correlation analysis uncovered significant links between different attributes and suicide rates, particularly highlighting strong associations between population and GDP as well as between population and the number of suicides, indicating that suicide rates are majorly affected by socio-economic factors.

	country	year	sex	age	suicides_no	population	gdp_per_capita ($)	generation	gdpyear
country	1.000000	0.022769	-0.000000	0.000025	0.118555	0.158896	0.053341	0.006069	0.163583
year	0.022769	1.000000	-0.000000	-0.000586	-0.004546	0.008850	0.339134	0.268301	0.094529
sex	-0.000000	-0.000000	1.000000	-0.000000	0.144629	-0.011242	-0.000000	-0.000000	-0.000000
age	0.000025	-0.000586	-0.000000	1.000000	-0.018256	-0.093361	-0.000157	0.214853	0.000061
suicides_no	0.118555	-0.004546	0.144629	-0.018256	1.000000	0.616162	0.061330	-0.070624	0.430096
population	0.158896	0.008850	-0.011242	-0.093361	0.616162	1.000000	0.081510	-0.067963	0.710697
gdp_per_capita ($)	0.053341	0.339134	-0.000000	-0.000157	0.061330	0.081510	1.000000	0.082899	0.303405
generation	0.006069	0.268301	-0.000000	0.214853	-0.070624	-0.067963	0.082899	1.000000	0.021251
gdpyear	0.163583	0.094529	-0.000000	0.000061	0.430096	0.710697	0.303405	0.021251	1.000000

Figure 2: correlation matrix

```
Linear Regression 0.45513827884911817
Random Forest 0.9839568507990769
Gradient Boosting 0.822946760456091
Extra Trees 0.9920747767399787
Bagging 0.9669252192876822
Decision Tree 0.9691594975974251
```

Figure 3: level of accuracy

The level of accuracy between various algorithms are obtained during testing and with the highest level of accuracy of 99.2% extra trees algorithm has been opted for the predictive analysis (figure 3). Utilizing the trained Extra Trees regressor algorithm, the model accurately forecasted the number of suicides in Colombia for the year 2023 based on input parameters such as population, GDP per capita, and GDP (figure 4).

```
Predicted value of no.of.sucidies in Colombia in year 2023 is 206.37.
```

Figure 4: result

Overall, the proposed predictive model for suicide rate analysis serves as a potent tool for comprehending and predicting suicide rates, drawing from a diverse set of socio-economic and demographic attributes. To achieve high prediction model accuracy, all steps of the machine learning

pipeline were closely monitored, including data collection, pre-processing, feature engineering, model selection, and evaluation. Continued refinement and validation of the model, along with ongoing enhancements to include additional data and features, can further elevate its accuracy and usefulness in informing public health policies, intervention strategies, and support services aimed at reducing suicide rates and fostering mental well-being on a global scale.

5 Conclusion

This comprehensive approach, considering demographic and economic factors, contributes to a nuanced understanding of suicide trends, potentially aiding in preventive strategies. The model's success in revealing intricate correlations underscores its potential as a valuable tool for forecasting and addressing suicide rates globally. Utilizing this model allows for the incorporation of diverse attributes including mental health enabling a more targeted analysis of specific factors that could collectively aid individuals in overcoming thoughts of self-harm. Prospective refinements may encompass themes like mental health awareness and the level of support provided by government, community, family, and friends, thereby enhancing the overall standard of assistance available.

References

1. Okada M, Hasegawa T, Kato R, et al Analysing regional unemployment rates, GDP per capita and financial support for regional suicide prevention programme on suicide mortality in Japan using governmental statistical data BMJ Open 2020;10:e037537. doi: 10.1136/bmjopen-2020-037537
2. Matsumoto, R.; Motomura, E.; Onitsuka, T.; Okada, M. Trends in Suicidal Mortality and Motives among Working-Ages Individuals in Japan during 2007–2022. Eur. J. Investig. Health Psychol. Educ. 2023, 13, 2795-2810. https://doi.org/10.3390/ejihpe13120193
3. Blasco-Fontecilla H, Perez-Rodriguez MM, Garcia-Nieto R, et al Worldwide impact of economic cycles on suicide trends over 3 decades: differences according to level of development. A mixed effect model study BMJ Open 2012;2:e000785. doi: 10.1136/bmjopen-2011-000785
4. Kato, R.; Okada, M. Can Financial Support Reduce Suicide Mortality Rates? Int. J. Environ. Res. Public Health 2019, 16, 4797. https://doi.org/10.3390/ijerph16234797
5. Miret, M., López-García, P. (2016). Effect of Economic Crises on Suicide Rates. In: Courtet, P. (eds) Understanding Suicide. Springer, Cham. https://doi.org/10.1007/978-3-319-26282-6_15
6. Oka, M., Kubota, T., Tsubaki, H. and Yamauchi, K. (2015), Geographic factors and suicide rate. Psychiatry Clin Neurosci, 69: 375-382. https://doi.org/10.1111/pcn.12254
7. Ramirez López, Carlos Manuel, et al. "Geospatial Situation Analysis for the Prediction of Possible Cases of Suicide Using EBK: A Case Study in the Mexican State of Aguascalientes." Handbook of Research on Natural Language Processing and Smart Service Systems, edited by Rodolfo Abraham Pazos-Rangel, et al., IGI Global, 2021, pp. 327-346. https://doi.org/10.4018/978-1-7998-4730-4.ch015
8. Swain PK, Tripathy MR, Priyadarshini S, Acharya SK (2021) Forecasting suicide rates in India: An empirical exposition. PLoS ONE 16(7): e0255342. https://doi.org/10.1371/journal.pone.0255342
9. Kim, J. W., Jung, H. Y., Won, D. Y., Noh, J. H., Shin, Y. S., & Kang, T. I. (2019). Suicide Trends According to Age, Gender, and Marital Status in South Korea. OMEGA - Journal of Death and Dying, 79(1), 90-105. https://doi.org/10.1177/0030222817715756

10. Dedić, G. J., & Tepšić-Ostojić, V. (2023). "Gender differences in suicide in Serbia within the period 2016-2020." *Vojnosanitetski pregled*, 80(1), 33-40. https://doi.org/10.2298/VSP210909015D

11. Innamorati M, Tamburello A, Lester D, et al. Inequalities in Suicide Rates in the European Union's Elderly: Trends and Impact of Macro-Socioeconomic Factors between 1980 and 2006. The Canadian Journal of Psychiatry. 2010;55(4):229-238. doi:10.1177/070674371005500406

12. M. Fernández-Delgado, M.S. Sirsat, E. Cernadas, S. Alawadi, S. Barro, M. Febrero-Bande, An extensive experimental survey of regression methods, Neural Networks, Volume 111, 2019, Pages 11-34, ISSN 0893-6080, https://doi.org/10.1016/j.neunet.2018.12.010.

13. Klimes-Dougan B, Wright N, Klingbeil DA. Suicide Prevention Public Service Announcements Impact Help-Seeking Attitudes: The Message Makes a Difference. Front Psychiatry. 2016 Jul 15;7:124. doi: 10.3389/fpsyt.2016.00124. PMID: 27471477; PMCID: PMC4945655.

14. Shu-Sen Chang, David Gunnell, Jonathan A.C. Sterne, Tsung-Hsueh Lu, Andrew T.A. Cheng, Was the economic crisis 1997–1998 responsible for rising suicide rates in East/Southeast Asia? A time–trend analysis for Japan, Hong Kong, South Korea, Taiwan, Singapore and Thailand, Social Science & Medicine, Volume 68, Issue 7, 2009, Pages 1322-1331, ISSN 0277-9536, https://doi.org/10.1016/j.socscimed.2009.01.010.

15. Kyla Thomas, David Gunnell, Suicide in England and Wales 1861–2007: a time-trends analysis, International Journal of Epidemiology, Volume 39, Issue 6, December 2010, Pages 1464–1475, https://doi.org/10.1093/ije/dyq094

16. Turecki, G., Brent, D.A., Gunnell, D. et al. Suicide and suicide risk. Nat Rev Dis Primers 5, 74 (2019). https://doi.org/10.1038/s41572-019-0121-0

17. S. Mary Fabiola, G. Ignisha Rajathi, R.Vedhapriyavadhana, L.R.Priya, "Artificial Intelligence to Combat the Sting of the Pandemic on the Psychological Realms of Human Brain", Springer Nature in March 2022. doi:10.1007/s42979-022-01038-8

18. United Nations Development Program. (2018). Human development index (HDI). https://hdr.undp.org/data-center/country-insights#/ranks

19. World Bank. (2018). World development indicators: GDP (current US$) by country:1985 to 2016. https://databank.worldbank.org/source/world-development-indicators

20. [Szamil]. (2017). Suicide in the Twenty-First Century [dataset]. https://www.kaggle.com/code/szamil/suicide-in-the-twenty-first-century/notebook

21. World Health Organization. (2018). Suicide prevention. https://www.who.int/health-topics/suicide#tab=tab_1

22. Suicide Rates Overview 1985 to 2016. https://www.kaggle.com/datasets/russellyates88/suicide-rates-overview-1985-to-2016