Technical Sciences, 2022, 25, 193–203

UNIWERSYTET WARMIŃSKO-MAZURSKI W OLSZTYNIE

Yearbook peer-reviewed scientific journal
ISSN 1505-4675
e-ISSN 2083-4527
TECHNICAL SCIENCES
Homepage: https://czasopisma.uwm.edu.pl/index.php/ts/


DOI: https://doi.org/10.31648/ts.8571

# MEASURING SADNESS INDEX BASED ON COUNTRY STATISTICS

*Artur Samojluk*[1], *Bartosz Nowak*[2], *Karolina Papiernik*[3]

[1]ORCID: 0000-0001-5822-2210
[2]ORCID: 0000-0002-2577-4470
[3]ORCID: 0000-0002-4948-1680
Faculty of Mathematics and Computer Science
University of Warmia and Mazury in Olsztyn

Received 25 December 2022, accepted 28 December 2022, available online 30 December 2022.

K e y w o r d s: happiness index, sadness index, *k-nn*, regression, machine learning.

A b s t r a c t

The article studied topics related to measuring people's sadness. For this purpose, the question was asked which factor: social, economic or climate, matters most. The paper analyzed, using machine learning, statistical data related to the number of suicides against the factors: level of Internet access, average income, temperature in a country and, in addition, population density. The method used was correlational statistical analysis using the *K*-nearest neighbor (KNN) method and also Pearson's correlation. The results were visualized in the form of graphs, then subjected to final analysis and included in the form of final conclusions.

## Introduction

The question of under what conditions a person is sadness is a complex question. A person's life is made up of many factors, and each of these factors is important in the subjective feeling of happiness (RAJNOHA et al. 2021, KAUR et al. 2019, IVANOVÁ et al. 2022, DREHMER 2018, BOGOMOLOV et al. 2013). The happiness question is one that has been asked many times before and

Correspondence: Artur Samojluk, Katedra Metod Matematycznych Informatyki, Wydział Matematyki i Informatyki, ul. Słoneczna 54, 10-710 Olsztyn, e-mail: artur.samojluk@uwm.edu.pl.

has been measured mainly using survey methods (SKIDELSKY 2014). In order to eliminate the psychological factor of respondents and various technicalities, for the purposes of this paper, survey research has been omitted. The authors assume that acts that cause final and irrevocable consequences (suicides), are more significant than words (survey declarations). The paper focuses on direct statistical factors. The suicide rate was taken as the baseline, extreme factor. Then four reference factors were selected, against which calculations were made using machine learning (IBNAT et al. 2021). Data for the calculations were taken from public databases, then correlated with each other and tabulated (*Business and economic…* 2022, *List of Countries by Average Temperature* 2022, *List of countries by suicide rate* 2022).

## Statistical research by countries

In this paper, the authors studied whether there is a linear relationship between the four factors and the level of sadness, and tested how effectively the level of unhappiness in a country can be predicted by considering only one of the factors. To check the linear relationship, the popular Pearson's correlation coefficient (DHARANEESHWARAN 2017) was used.

To measure the effectiveness and relevance of individual factors in predicting the level of sadness, a machine learning (POLKOWSKI, ARTIEMJEW 2015, CPALKA et al. 2007, POLKOWSKI 2022) method called *K*-nearest neighbor (KNN) (COVER, HART 1967, NOWICKI et al. 2014, MA et al. 2016) method was used. Efficiency in this case means how effectively with a single factor the algorithm can predict the value of the sadness level. The higher the efficiency, the lower the final RMSE (root-mean-square error) measured by comparing the true values with those predicted by the *k-nn* algorithm. *K-nn* is a simple, popular and effective method dedicated to statistical machine learning calculations (NOWICKI 2014, KUMAR 2015). It is extremely easy to understand and usually gives very good results, especially when the number of attributes processed simultaneously is small (BEYER et al. 1999). Its disadvantages is need to store a set of reference samples, while its huge advantage is the lack of need to create a knowledge base, since it is derived directly from the set of reference samples. Using this method, or another one for approximation or regression (MARVIN et al. 1990, QI et al. 2022), it is possible to predict which factors usually more accurately and which usually less accurately can predict the level of sadness.

The baseline rate (the number of suicides per 100,000 people in a single year) was compared to three completely different factors. The first is an economic factor in the form of the average income of a citizen in a country. The second is a social factor, measuring the percentage of the population with access to the Internet. The third is a climatic factor closely related to the temperature

in a country (region). The fourth factor analyzed is the population density per square kilometer in the country under study.

For the purpose of this research, data was collected from 164 countries from around the world.

It has been assumed that in countries where people are most unhappy the measuring factor is the number of suicides per 100,000 people. Hence, it can be assumed that in countries where the number of suicides is the lowest, these are the happiest countries and people there are satisfied with their lives. This factor also appears to be a fairly objective expression of satisfaction, independent of the perception of happiness.

The authors additionally decided to test if by knowing only the value of one of the factors it is possible to predict the level of happiness in that country. In this way, it is also possible to examine the level of correlation between the studied factors and happiness. For this purpose, the well-known, popular and easy to understand *k-nn* algorithm (COVER, HART 1967, NOWICKI et al. 2014) was used. However, this algorithm, if used incorrectly, could return useless results due to a phenomenon called overfitting. If the algorithm were used on the same test samples as the reference samples and using the appropriate control parameter, it could almost always indicate that it is possible to predict the value of the luck level completely accurately. The authors decided to eliminate this problem by using Leave-One-Out Cross-Validation. In addition, the Nested Cross-Validation technique, also in combination with Leave-One-Out, was used to neglect the impact of the *k-nn* algorithm's control parameter on the results.

It was assumed that a single sample describes selected features of a single country. Each reference sample is weighted proportionally to the number of peoples. The set of sample input attributes used depends on the assumed factor, while the output attribute is always the suicide rate (per 100,000 people in a single year). For the purposes of the study, the *k-nn* algorithm was used to predict the level of happiness based on the values of available attributes. This task in Machine Learning terminology is called approximation or regression (QI et al. 2022, NOWICKI et al. 2014).

A single use of the *k-nn* algorithm will provide a supposed output value for the given test sample (predicted suicide rate). The test sample contains the values of the characteristics under test, which are also called input attributes. The algorithm needs a set of test samples to work. The operation of the algorithm can be described in several steps:

1. Count the similarity of each reference sample to the test sample. When measuring similarity, only the values of the selected samples' input attribute (feature) are considered. Similarity is determined by considering only the absolute value of the difference between the value of the test sample and the reference sample of the attribute (feature) under consideration.

2. Sort the reference samples according to their similarity counted in the previous step. The most similar samples should come first.

3. Select the $k$-first reference samples. The value of $k$ is chosen to be the smallest number such that the sum of the weights of the input samples is greater than or equal to the sum of the weights of all input samples multiplied by the control parameter $\text{var}_{\text{weight\_rate}}$ of $k$-$nn$. In this way, the number $k$ can vary for each test sample. Such solution is unusual and is different from most implementations of the $k$-$nn$ algorithm. This change is due to the need to weight the reference samples.

4. Count the average output value for the first $k$ reference samples. When determining this average, their weight is not taken into account, because surprisingly, this resulted in better results. The value of this average is considered as the predicted output value provided for this test sample.

Below is an example of how this algorithm works. Let the test sample have an input value of 12, the control parameter $\text{var}_{\text{weight\_rate}} = 18\%$ and a set of reference samples $(x_{\text{ref},1}, \ldots, x_{\text{ref},8})$ described by weight, input value, output value:

| reference sample | $x_{\text{ref},1}$ | $x_{\text{ref},2}$ | $x_{\text{ref},3}$ | $x_{\text{ref},4}$ | $x_{\text{ref},5}$ | $x_{\text{ref},6}$ | $x_{\text{ref},7}$ | $x_{\text{ref},8}$ |
|---|---|---|---|---|---|---|---|---|
| weight | 3 | 1 | 2 | 1 | 1 | 1 | 2 | 4 |
| value of input | 5 | 6 | 8 | 10 | 10 | 11 | 15 | 15 |
| value of output | -5 | -2 | -1 | 0 | -2 | 1 | 3 | -1 |

After calculation, the absolute value of the differences between the reference samples and the test sample equals, considering only the input attribute:

| reference sample: | $x_{\text{ref},1}$ | $x_{\text{ref},2}$ | $x_{\text{ref},3}$ | $x_{\text{ref},4}$ | $x_{\text{ref},5}$ | $x_{\text{ref},6}$ | $x_{\text{ref},7}$ | $x_{\text{ref},8}$ |
|---|---|---|---|---|---|---|---|---|
| abs. of difference: | 7 | 6 | 4 | 2 | 2 | 1 | 3 | 3 |

Then, after sorting the reference samples taking into account the similarity to the test sample:

| reference sample: | $x_{\text{ref},6}$ | $x_{\text{ref},5}$ | $x_{\text{ref},4}$ | $x_{\text{ref},7}$ | $x_{\text{ref},8}$ | $x_{\text{ref},3}$ | $x_{\text{ref},2}$ | $x_{\text{ref},1}$ |
|---|---|---|---|---|---|---|---|---|
| abs. of difference: | 1 | 2 | 2 | 3 | 3 | 4 | 6 | 7 |

After the calculation, the sum of the weights of the reference samples multiplied by the control parameter $\text{var}_{\text{weight\_rate}}$ is $15 \cdot 0.18 = 2.7$. Therefore, taking the most similar reference samples, such that the sum of their weights is greater than or equal to 2.7:

| reference sample: | $x_{\mathrm{ref},6}$ | $x_{\mathrm{ref},5}$ | $x_{\mathrm{ref},4}$ | $x_{\mathrm{ref},7}$ | $x_{\mathrm{ref},8}$ | $x_{\mathrm{ref},3}$ | $x_{\mathrm{ref},2}$ | $x_{\mathrm{ref},1}$ |
|---|---|---|---|---|---|---|---|---|
| abs. of difference: | 1 | 2 | 2 | 3 | 3 | 4 | 6 | 7 |
| weight | 1 | 1 | 1 | 2 | 4 | 2 | 1 | 3 |

Finally, to determine the output value for the test sample, it is necessary to calculate the average value of the outputs, for the selected, most similar reference samples selected in the previous step:

| reference sample: | $x_{\mathrm{ref},6}$ | $x_{\mathrm{ref},5}$ | $x_{\mathrm{ref},4}$ |
|---|---|---|---|
| value of output | 1 | -2 | 0 |

After calculations, it comes out that the *k-nn* algorithm set the output value as the average $\{1; -2; 0\} = -\frac{1}{3}$.

The used algorithm *k-nn* has one control parameter $\mathrm{var}_{\mathrm{weight\_rate}}$. It affects how many samples are taken into account when determining the output value for a single test sample. As this parameter increases, the average number of $k$ samples taken into account when predicting the value of a test sample increases. Thus, the value of this control parameter has a great impact on the obtained results. Often, many researchers manually set the values of the control parameters of the tested models so that the obtained results are the best. Such a solution does not seem very fair, because it leads to the so-called "information leakage", in other words, the value of the control parameters is determined on the basis of the results they affect. This can be compared to a situation such as if the composition of the jury was determined by one of the sides in a judicial conflict. A solution that addresses this type of problem is Nested Cross Validation.

The Nested Cross Validation algorithm selected one of the following variants of the value of the control parameter of the *k-nn* algorithm: 25, 10, 5, 2.5, 1 percent.

The entire research can be described in the following steps:

1. $X_{\mathrm{all}}$ = the entire set of samples.
2. Var = {25%, 10%, 5%, 2.5%, 1%} = the set of considered values for the control parameter of the *k-nn* algorithm.
3. $\mathrm{Errors}_{\mathrm{outer}} = \{\}$ = empty set for errors obtained for all test samples.
4. Program loop, let $s_{\mathrm{outer}} = 1 \ldots \|X_{\mathrm{all}}\|$:
   a – $x_{\mathrm{outer,tst}}$ $X_{\mathrm{all},s_{\mathrm{outer}}}$ = test sample, the $s_{\mathrm{outer}}$-th sample in $X_{\mathrm{all}}$;
   b – $X_{\mathrm{outer,vallrn}} = X_{\mathrm{all}} - x_{\mathrm{outer,tst}}$ = set of validation and learning samples;
   c – $\mathrm{Eval}_{\mathrm{var}} = \{\}$ = empty set containing the evaluations obtained for individual variants of the *k-nn* control parameters;

d – Programming loop, let var_nr = 1 ‖Var‖:
- $\text{Errors}_{\text{inner}} = \{\}$ empty set of errors for *k-nn* algorithm used for var_nr variant of control parameters,
- Program loop. let $s_{\text{inner}} = 1\ldots \|X_{\text{outer,vallrn}}\|$:
  - $x_{\text{inner,val}} = X_{\text{outer,vallrn},s_{\text{inner}}}$ = validation sample,
  - $X_{\text{inner,lrn}} = X_{\text{outer,vallrn}} - x_{\text{inner,val}}$ = set of learning samples,
  - $\text{out}_{\text{correct}}$ = the correct output value for the current validation sample ($x_{\text{inner.val}}$),
  - $\text{out}_{\text{actual}} = k\text{-}nn(X_{\text{inner,lrn}}, x_{\text{inner\_val}}; \text{Var}_{\text{var\_nr}})$ = the result obtained for the validation sample $x_{\text{inner,val}}$ using the *k-nn* algorithm, taking $X_{\text{inner,lrn}}$ as reference samples and $\text{Var}_{\text{var\_nr}}$ as the *k-nn* control parameter ($\text{var}_{\text{weight\_rate}}$),
  - $\text{Errors}_{\text{inner}} = \text{Errors}_{\text{inner}} \cup (\text{out}_{\text{correct}} = \text{out}_{\text{actual}})$,
- $\text{Eval}_{\text{var,var\_nr}} = \sqrt{\frac{\sum_{s_{\text{inner}}=1}^{\|X_{\text{outer,vallrn}}\|}\left(\text{Errors}_{\text{inner},s_{\text{inner}}}\right)^2}{\|X_{\text{outer,vallrn}}\|}}$ = rating assigned for variant , Root-Mean-Square Error;

e – var_best_nr = $\text{minarg}_{\text{var\_nr}=1\ldots\|\text{Var}\|}(\text{Eval}_{\text{var,var,nr}})$ = number of the best variant obtained for the set of samples $X_{\text{outer,vallrn}}$;

f – $\text{out}_{\text{correct}}$ = correct output value for the current test sample ($x_{\text{outer,tst}}$);

g – $\text{out}_{\text{actual}} = k\text{-}nn\ (X_{\text{outer,vallrn}}, x_{\text{outer,tst}}; \text{Var}_{\text{var\_best\_nr}})$ = the result obtained for the test sample $x_{\text{outer,tst}}$ using the *k-nn* algorithm, taking $X_{\text{outer,vallrn}}$ as reference samples and the *k-nn* control parameter ($\text{var}_{\text{weight\_rate}}$) equal to $\text{Var}_{\text{var\_best\_nr}}$;

h – $\text{Errors}_{\text{outer}} = \text{Errors}_{\text{outer}} \cup (\text{out}_{\text{correct}} = \text{out}_{\text{actual}})$, add the actual error obtained for the test sample $x_{\text{outer,tst}}$.

5. $\text{Eval}_{\text{global}} = \sqrt{\frac{\sum_{s_{\text{outer}}=1}^{\|X_{\text{all}}\|}\left(\text{Errors}_{\text{outer},s_{\text{outer}}}\right)^2}{\|X_{\text{all}}\|}}$ = global rating given for the used algorithm and chosen input attributes, Root-Mean-Square Error.

For each of the 164 reference samples (loop 4), first it was determined which of the 5 variants of the control parameter $\text{var}_{\text{weight\_rate}}$ was the best ($164 \cdot 163 \cdot 5$ calls of the *k-nn* algorithm, loop 4 d and step 4 e) and then it was examined what answer the best-matched parameter gave (164 runs of the *k-nn* algorithm, step 4 g). The leave-one-out cross-validation scheme was thus used to determine the accuracy of the algorithm's performance (loop 4) and to determine the accuracy of the performance of each control parameter (loop 4 d).

The level of complexity of the entire examination process can be accurately estimated as $\theta(n^2)$, which, considering 164 samples, is still acceptable, and ensures reliable results with cross-validation. The *k-nn* algorithm itself has a worst-case complexity $\theta(n^2)$, due to the presence of sorting operations in it.

Using the above algorithm to determine the rating of each feature, the *k-nn* algorithm has as been executed 164(163 · 5 + 1) times. The results for different attributes are shown in the Table 1.

Table 1

Results for examined features

| Used input attribute (feature) | *k-nn* algorithm, Root Mean Square Error | Pearson's correlation coefficient |
|---|---|---|
| GPD per capita | 5.98 | −0.1053 |
| Access to Internet | 5.73 | −0.1698 |
| Avg. temperature | 6.17 | −0.2077 |
| Avg. temp. in the coldest month | 5.90 | −0.1737 |
| Avg. temp. in the hottest month | 6.47 | −0.2389 |
| Density of population | 6.07 | −0.0882 |

The results of the calculations are shown below in Figures 1, 2, 3, 4, using the best variant of parameter controlling *k-nn* algorithm and all samples. Countries have been plotted on the charts, and their size has been weighted against the size of the population (see legend on the chart). The first graph (Fig. 1) shows the relationship between the suicide rate and the income ratio GDP (Gross Domestic Product). The second graph (Fig. 2) shows results measuring the relationship between a country's suicide rate and access to the Internet in the study country. The third graph (Fig. 3) analyzes the relationship between a country's suicide rate and the average annual temperature that occurs in that country. The fourth graph (Fig. 4) studied the relationship between suicide rates and population density per square kilometer. In the next section the research results are analyzed.

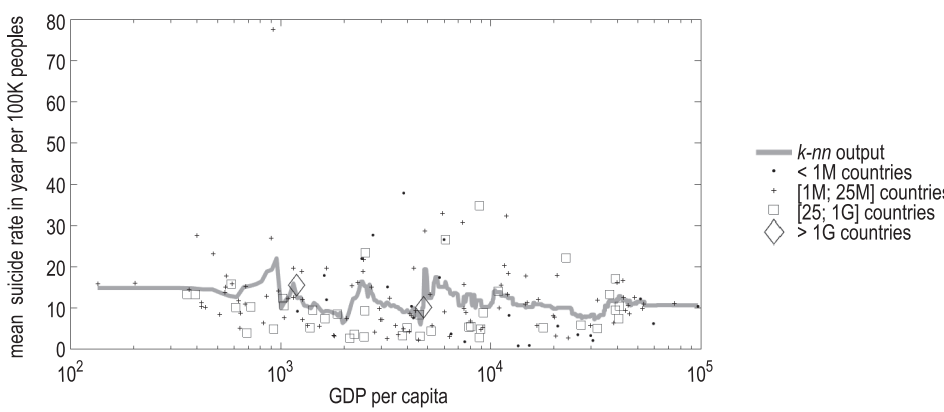


Fig. 1. Suicide rate by country vs GDP per capita

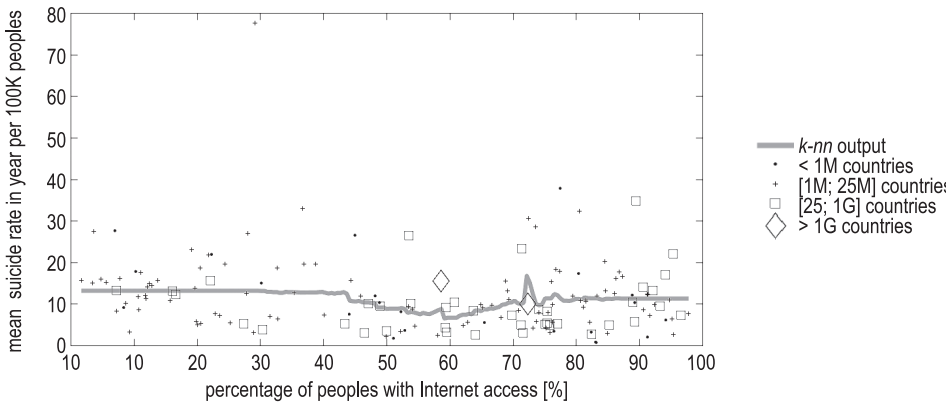


Fig. 2. Suicide rate by country vs Internet access

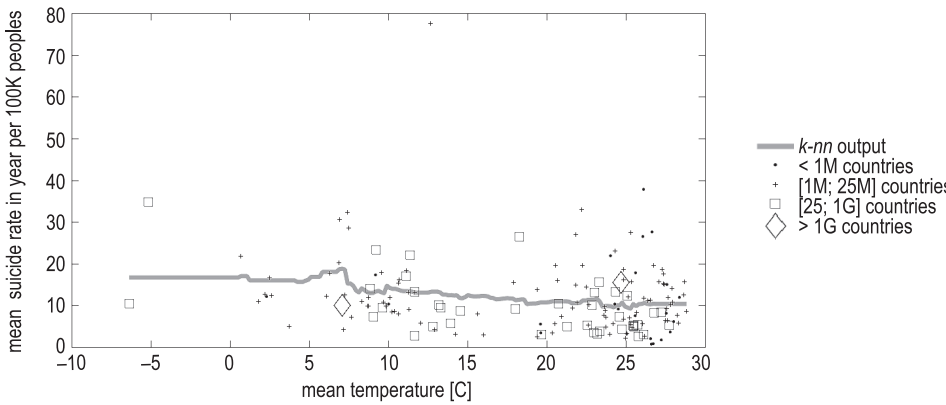


Fig. 3. Suicide rate by country vs avg. temperature

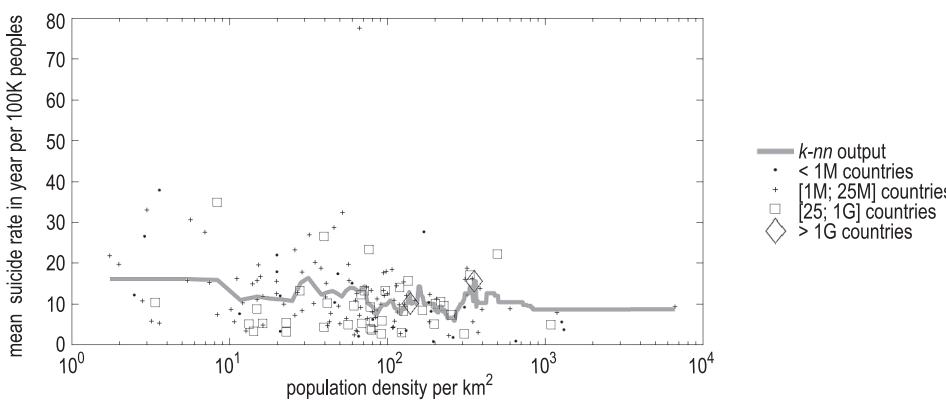


Fig. 4.

## Analysis of the results

We begin our analysis of the results using a graph (Fig. 1) that investigates the relationship between the suicide rate and a country's per capita income (GDP) ratio. To simplify the analysis of the results, our main referent will be the smoothed ratio. This method of analysis will be substituted as well, for the subsequent charts covered in this paper.

Inspection of the smoothed result for (Fig. 1) shows that in countries where there is a higher per capita income ratio, the number of suicides statistically decreases. This is not a conclusive indicator, because, as we see, there are strong fluctuations, which are influenced by other social factors. The general conclusion, for the income ratio, can be taken as the conclusion that where poverty is less, people tend to be happier.

In the next part of the study, we considered suicide rate data, in relation to Internet access. Here the results are no longer so conclusive, although we can observe a strong decline in suicides in countries where Internet access is becoming more common. However, we can also see an upward shift in suicide rates in countries where Internet access is higher than 60%. This is an interesting phenomenon, and can potentially be explained by social burdens and more complicated relationships in a society with widespread Internet access. The third result, is the one the authors find most interesting, relates to the relationship associated with average temperature. In this case, we measure the suicide rate against the average temperature of the country. Here we observe a trend of decreasing suicide rates in countries where the average annual temperature is increasing. The supposition is that in warmer countries people are more satisfied with their lives, regardless of the level of access to the Internet and the income index.

## Conclusion

For research purposes, the authors analyzed data for 164 countries from around the world. The authors mainly tried to answer the question: which factor matters most: social, economic or climate. The main reference indicator to which the other indicators were compared was the suicide rate per hundred thousand people. The paper compared the benchmark indicator relative to: the income index (GDP), Internet access, average temperature in country, average temperature in the coldest and the warmest month, and population density.

The results of the study show a significant impact of the amount of income on the reduction of suicide in a country. This is likely due to the fact that as poverty decreases, people live better. This observation is a well-known result, while the authors in this paper examine the effect of the level of internet access

and average temperature on happiness. An analysis of the suicide rate relative to internet accessibility, showed an interesting phenomenon. A low level of access correlates with a high suicide rate, which tends to be related to the level of poverty in a country. Then, as accessibility increases, then the suicide rate decreases to reach its lowest value for access levels around 60%. This is followed by a mild reversal of the trend and the suicide rate slowly rises as Internet accessibility increases.

Subsequently analyzed indicators related to the impact of a country's temperature, on the suicide rate. The authors suggest that increasing life satisfaction may be due to an increase in the country's annual temperature. It can be clearly seen that successive local maxima of the smoothed average temperature index are getting lower and lower, which means that as the temperature increases, the suicide rate decreases significantly. Which is very evident in countries that are located in very cold regions of the planet.

The last factor is the effect of density on the density. Unfortunately, this factor un- surprisingly showed a trend that is not very pronounced. It can be seen that the level of happiness in people in the most densely populated countries is higher than in those with a sparser population.

Interpreting the values of Pearson's coefficients, it is possible to learn a simple linear relationship between the studied factors and the level of sadness. Using it can be seen that as each of the studied factors increases, the level of sadness tends to decrease.

## References

BEYER K., GOLDSTEIN J., RAMAKRISHNAN R., SHAFT U. 1999. *When Is "Nearest Neighbor Meaningful?"*. Lecture Notes in Computer Science, 1540.

BOGOMOLOV A., LEPRI B., PIANESI F. 2013. *Happiness Recognition from Mobile Phone Data*. International Conference on Social Computing, 08-14 September 2013, Alexandria, VA, USA. https://doi.org/10.1109/SocialCom.2013.118.

*Business and economic data for 200 countries*. 2022. The Global Economy. https://www.theglobaleconomy.com/rankings/gdp_per_capita_current_dollars/.

COVER T., HART P. 1967. *Nearest neighbor pattern classification*. IEEE Transactions on Information Theory, 13(1): 21–27.

CPALKA K., NOWICKI R.K., RUTKOWSKI L. 2007. *Rough-Neuro-Fuzzy Systems for Classifaction*. The First IEEE Symposium on Foundations of Computational Intelligence (FOCI'07).

DHARANEESHWARAN, NITHYA S., SRINIVASAN A., SENTHILKUMAR M. 2017. *Calculating the user-item similarity using Pearson's and cosine correlation*. International Conference on Trends in Electronics and Informatics (ICEI), p. 1000-1004. https://doi.org/10.1109/ICOEI.2017.8300858.

DREHMER J.E. 2018. *Sex differences in the association between countries' smoking prevalence and happiness ratings*. Public Health, 160: 41-48. https://doi.org/10.1016/j.puhe.2018.03.027.

IBNAT F., GYALMO J., ALOM Z., AWAL M.A., AZIM M.A. 2021. *Understanding World Happiness using Machine Learning Techniques*. International Conference on Computer, Communication, Chemical, Materials and Electronic Engineering (IC4ME2).

IVANOVÁ M., KLAMÁR R., FECKOVÁ-ŠKRABULÁKOVÁ E. 2022. *Identification of factors influencing the quality of life in European Union countries evaluated by Principal Component Analysis*. Geographica Pannonica, 26: 13-29. https://doi.org/10.5937/gp26-34191.

KAUR M., DHALARIA M., SHARMA P.K., PARK J.H. 2019. *Supervised Machine-Learning Predictive Analytics for National Quality of Life Scoring*. Applied Sciences, 9: 1613. https://doi.org/10.3390/app9081613.

KUMAR T. 2015. *Solution of Linear and Non Linear Regression Problem by K Nearest Neighbour Approach: By Using Three Sigma Rule*. IEEE International Conference on Computational Intelligence & Communication Technology, p. 197-201. https://doi.org/10.1109/CICT.2015.110.

*List of Countries by Average Temperature*. 2022. List First. https://listfist.com/list-of- countries-by-average-temperature.

*List of countries by suicide rate*. 2022. Wikipedia. https://en.wikipedia.org/wiki/List_of_countries_by_suicide_rate.

MA W., TAN K., DU Q., DING J., YAN Q. 2016. *Estimating soil heavy metal concentration using hyperspectral data and weighted K-NN method*. 8th Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing (WHISPERS). https://doi.org/10.1109/WHISPERS.2016.8071813.

MARVIN H., GRUBER J. 1990. *Regression Estimators. A Comparative Study*. Academic Press, Stanford University, California.

NOWICKI R.K., NOWAK B.A., WOZNIAK M. 2014. *Rough k nearest neighbours for classification in the case of missing input data*. In: *Proceedings of the 9th International Conference on Knowledge, Information and Creativity Support Systems, Limassol, Cyprus, November 6-8, 2014*. Ed. G.A. Papadopoulos. University of Cyprus, Nicosia, Cyprus.

NOWICKI R.K., NOWAK B.A., STARCZEWSKI J.T., CPAŁKA K. 2014. *The Learning of Neuro-Fuzzy Approximator with Fuzzy Rough Sets in Case of Missing Features*. International Joint Conference on Neural Networks, p. 3759-3766.

POLKOWSKI L., ARTIEMJEW P. 2015. *Granular Computing in Decision Approximation. An Application of Rough Mereology*. Series: Intelligent Systems Reference Library, 77. https://doi.org/10.1007/978-3-319-12880-1_1.

POLKOWSKI L.T. 2022. *What Logics for Computer and Data Sciences, and Artificial Intelligence*. Studies in Computational Intelligence (SCI), p. 992.

QI X., GAO Y., LI Y., LI M. 2022. *K-nearest Neighbors Regressor for Traffic Prediction of Rental Bikes*. 14th International Conference on Computer Research and Development (ICCRD), p. 152-156, https://doi.org/10.1109/ICCRD54409.2022.9730527.

RAJNOHA R., LESNÍKOVÁ P., VAHANČÍK J. 2021. *Sustainable economic development: The relation between economic growth and quality of life in V4 and Austria*. Economics and Sociology, 14(3): 341-357. https://doi.org/10.14254/2071-789X.2021/14-3/18.

SKIDELSKY E. 2014. *What Can We Learn From Happiness Surveys?* Journal of Practical Ethics, 2(2): 20-32.