Khazanah Journal of Religion and Technology
Universitas Islam Negeri Sunan Gunung Djati Bandung, Indonesia
https://journal.uinsgd.ac.id/kjrt
Online ISSN: 2987-6060
Volume 1 Issue 1 (2023)
DOI: https://doi.org/10.15575/kjrt.v1i1.161

The Influence of the Use of Social Media on the Intensity of Worshipping the Millenial Generation using Linear Regression

Fakhri Faisal Rochdiana
Program and Partnership
Google Developer Student Clubs
Bandung, Indonesia
fakhri.fr30@gmail.com

Muhammad Rizki Isa Darmawan
Program and Partnership
Google Developer Student Clubs
Bandung, Indonesia
riskiisa23@gmail.com

Nurul Aulia Dewi
Program and Partnership
Google Developer Student Clubs
Bandung, Indonesia
nurulad8@gmail.com

Wafda Zulfa Husna
Program and Partnership
Google Developer Student Clubs
Bandung, Indonesia
wafdazulha@gmail.com

Zulfa Dwi Audina
Program and Partnership
Google Developer Student Clubs
Bandung, Indonesia
zulfaaudina11@gmail.com

Nunik Destria Arianti
Department of Informatics
Universitas Nusaputra
Sukabumi, Indonesia
nunik@nusaputra.ac.id

Abstract— Worship is an activity carried out by every religious community. Sometimes there are activities that can interfere with worship activities, especially in the current era of globalization. In the current era of globalization, many people use their smartphones or gadgets for their daily needs. The millennial generation is one of the generations that has begun to be introduced to this gadget or smartphone, so that many of these millennial generations cannot be separated from what is called social media. During the COVID-19 pandemic, almost all our daily time is used to view social media. The purpose of this study was to see the effect of using social media on the intensity of worship of the millennial generation. The method used to assist this research is to use linear regression. The result of this research shows that the use of gadgets for a long time is accompanied by an extraordinary intensity of worship. This research also concludes that there is no influence given using social media on the intensity of worship of the millennial generation.

Keywords-linear regression, millenial, social media, worship

I. INTRODUCTION

Nowadays, almost everyone knows social media. Not only being a source of entertainment, social media is also often an important source of information for its users. This makes social media one of the things that is widely used by all humans. Because there are many things we can do on social media, of course this has a negative and positive impact on the influence of Social Media. From 4,66 billion internet user in the world [1] in 2021, including 202 million Indonesian people [2], there are 59,32% social media users [3], and this number continues increasing. The positive impact that we can immediately feel is that we can easily connect with people who are far away via virtual, but the negative impact that is often the main trigger is addiction to playing Social Media so

that we ignore obligations that should be carried out, one of which is the obligation to worship.

In the era of globalization, there are many things that can reduce worship activities. The use of devices is also an important factor in this matter. The various features provided are not only attractive, but sometimes also helpful in everyday life. One example is in the use of social media. social media is widely used by children, teenagers, and parents. The existence of COVID-19 really affects the use of gadgets and worship activities because in everyday life the use of gadgets is definitely involved by using social media. The millennial generation is a generation that uses a lot of gadgets and social media in their daily life [4], [5]. Excessive use of Social Media can take time, so there is a possibility of reduced worship activities. By taking this into account, it is necessary to predict whether the use of social media affects worship activities or not. To predict the effect of using social media on worship activities, machine learning technology can be utilized, one of which is the regression technique.

II. RELATED WORKS

There are several related studies that discuss about prediction or forecasting research with machine learning, especially regression method, such as:

1. Using sequential minimal optimization with logistic regression and fuzzy rough nearest neighbour to detect credit card fraud [6]. This study discovered that logistic regression (LR) can enhance the final outcome of prediction. With detection rates of 84.90% and 76.30%, the comparison with seven different algorithms shows that the ensemble model can successfully identify credit card fraud.
2. An analysis and prediction model for crimes using many variables [7]. This study forecasts the frequency (count) of crimes at the beat-day level in the city of Chicago in an effort to determine when they will occur. Forecasting crimes aids in the development of ways for preventing them, and the frequency of crimes aids in focusing on the particular type of crime. This innovative work is a partnership between computer science and criminal justice that aims to create a data mining process that can hasten the resolution of crimes. The author concentrated on the daily variables contributing to crime rather than those contributing to crime's occurrence, such as political animosity, the criminal history of the perpetrator, etc.
3. Using machine learning instead of linear regression modelling can provide accurate predictions of ozone concentrations [8]. Linear Regression, Neural Network, and Boosted Decision Tree are the Machine Learning techniques investigated in this study. Wind speed, humidity, Nitrogen Oxide, Carbon Monoxide, and Nitrogen Dioxide all had a major impact on ozone generation, according to the findings. For all stations, Boosted Decision Tree outperformed Linear Regression and Neural Network techniques.
4. Exponential regression prediction of COVID-19 vaccination target achievement [9]. This study predicts that the achievement of the national COVID-19 immunization objective in Indonesia will be difficult due to a variety of current impediments. Predictions using exponential regression modeling suggest that the vaccine objective will be met 100 percent on January 18, 2022, but only 80 percent on December 31, 2021. According to recent data, greater acceleration is required, especially if it is expected to be completed by December 2021, as determined by President Joko Widodo, there will be a 20 percent shortage based on prediction findings.
5. The variant of regression and decision tree which is regression tree has conducted for several case, such as to detect complication of malaria and to classify open unemployment [10], [11].

III. RESEARCH METHODS

A. *Research activities*

Figure 1 describes the research activities that begin with problem identification, data collecting, data pre-processing, linear regression modelling, model evaluation, and conclusion result.

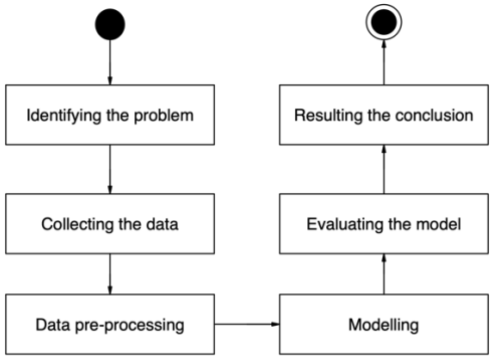

Fig 1. Research activities

The problem identification stage is an attempt to explain the problem and make a measurable explanation. This identification was carried out as the first step of research. So, if there is a problem, we can conduct research to find a solution to the problem. This literature study stage was carried out by research from several journals that have the same research that we did.

The data used in this study is data via the Google form which we distribute through family groups with a target age of the millennial generation (25-40 years). The data collection was carried out within one week. This data contains those related to social media and worship activities including name, age, social media that is most often used, duration of using social media per day, how often to perform the obligatory prayers per day, how often to perform the midnight prayer in a week, how often to perform the Duha prayer in a week, how

many times to read Al-Qur'an in a week and the duration of reading Al-Qur'an in a week.

In data pre-processing is a process that transforms raw data into a form that is easier to understand. This study uses a simple linear regression method to make predictions according to the data we have obtained. In predicting calculations, the model is used to train training data and validate data testing. Accuracy calculations are carried out at the training and testing stages, so that you can see the difference in accuracy from the two stages. Conclusions are drawn based on the predicted data which will then be compared with the raw data.

B. Linear Regression Algorithm

Linear regression is a simple analysis model with interval data types [12], [13]. With this analysis, it is done by predicting based on predetermined data. In general, linear regression is used to determine whether the independent variables studied have a significant correlation with the dependent variable and find out which variables have a significant effect on the dependent variable [8], [14], [15].

Basically, there are two types of linear regression, including:

1. Simple Linear Regression

Simple linear regression is a linear one that is used to find out the correlation between the independent and dependent variables. In simple linear regression, there is one independent variable and one dependent variable. The formula for simple linear regression is available in formula (1).

$$Y = a + b \quad (1)$$

Where Y is dependent variable, a in an intercept value, b is slope of the line, and X is independent variable.

2. Multiple Linear Regression

Multiple linear regression is the same as simple linear regression, the difference is that multiple linear regression is carried out to find out the correlation between the number of independent variables studied by more than one and the dependent variable. The equation of multiple linear regression is provided in formula (2).

$$Y = a + b_1X_1 + b_2X_2 + b_3X_3 + \cdots + b_nX_n \quad (2)$$

In advance, linear regression is developed to stepwise regression [16], ridge regression [17], lasso regression [18], elastic net regression [19], decision tree regression [20], support vector regression [21], random forest regression [22], and so on.

IV. RESULT AND DISCUSSION

A. Data Collecting

Data in this study were collected by preparing a questionnaire instrument containing the following questions:

1. How often do you pray obligatory prayers in a day (*Seberapa sering Anda sholat wajib dalam sehari*)? - X_1
2. How often do you pray Tahajud in a week (*Seberapa sering Anda sholat Tahajud dalam seminggu*)? - X_2
3. How often do you pray Dhuha in a week (*Seberapa sering Anda sholat Dhuha dalam seminggu*)? - X_3
4. How many times do you read the Koran in a week (*Berapa kali Anda membaca Al-Qur'an dalam seminggu*)? - X_4
5. How long do you use social media in a day (*Berapa durasi Anda menggunakan media sosial dalam sehari*)? - X_5
6. How long do you read the Qur'an in a week (*Berapa durasi Anda membaca Al-Qur'an dalam seminggu*)? - X_6

These questions will be the independent variables in this study. Research data was obtained by distributing questionnaires to relatives or colleagues of all researchers. While, the dependent value is a worship intensity value which is calculated based on the amount of time spent for worship compared to the amount of time spent on social media.

B. Data Pre-processing

Data pre-processing is the initial data mining technique for converting raw data into more efficient and useful formats and information. This process must be carried out, because raw data usually does not have an orderly format and by doing this stage, we can see the data according to our needs by displaying the correlation between one variable and another. This research conducts the cleaning and data transformation.

In the data cleaning stage, the raw data will be sorted. If there is data whose type is object, it will be converted into integer data, then filtering for empty data, and only taking data that is the millennial generation (25-40 years old). Variable X_6 is not used in the regression model development. Then, in the data transformation process, the worship variables (X_1 - X_4) are transformed become integer with worship intensity value calculation. For X_5 the duration of social media also change into numeric, duration less than 5 hours become 1, 5 to 10 hours become 2, while more than 10 hours become 3. Figure 1 provides the example of dataset.

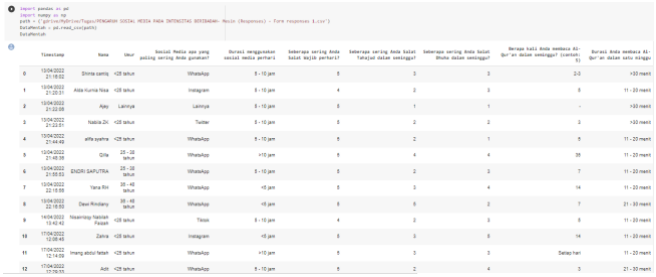

Fig 2. Dataset example

C. Modeling

The worship variables (X_1 - X_4) are calculated and become worship intensity value (Y). Table I shows the

example of relation between worship intensity value (Y) and duration of use of social media (X). The coefficient value of regression model in this research is 12.232.

TABLE I. THE EXAMPLE DATASET BETWEEN WORSHIP INTENSITY VALUE (Y) AND DURATION OF USE OF SOCIAL MEDIA (X)

Use of Social Media Duration (X)		Worship Intensity Value (Y)
Less than 5 hours	1	31.07
More than 10 hours	3	78.57
Less than 5 hours	1	75
Less than 5 hours	1	62.5
5 to 10 hours	2	55.35
5 to 10 hours	2	46.42
Less than 5 hours	1	33.92
Less than 5 hours	1	41.42
Less than 5 hours	1	42.85
Less than 5 hours	1	32.14
Less than 5 hours	1	58.93

Based on the results of the regression model obtained from survey data, survey participants with a duration of using their gadgets for more than ten hours, which is represented by the number three, actually show an extraordinary intensity of worship, which after being converted into a score, gets a score of 78.57. On the other hand, survey participants with a gadget usage duration of less than five hours showed a very diverse intensity of worship, ranging from 31.07 to 75. This regression model shows that the duration of social media use does not directly affect the millennial generation's worship. This is shown from the duration of using social media, both briefly and for a long time, can have a high intensity value of the millennial generation's worship.

V. CONCLUSION

Based on the results of the analysis that has been carried out, from the data that has been collected which has passed the data cleaning and data transformation stages, the conclusion that can be drawn is that the higher the time spent using social media, the higher the intensity of the worship performed. So that in general there is no direct connection between the duration of social media use and the intensity of the millennial generation's worship. However, from the results of this analysis, there is a discrepancy between the time spent using social media and worship, because the higher the use of social media, the lower the time that can be used for worship. Therefore, further research needs to prepare better datasets and analyze the required variables more deeply.

REFERENCES

[1] M. I. Marsyaf, "Jumlah Pengguna Internet Sedunia Mencapai 4,66 Miliar," *sindonews.com*, 2021. https://tekno.sindonews.com/read/316920/207/jumlah-pengguna-internet-sedunia-mencapai-466-miliar-1611820860 (accessed Oct. 03, 2021).

[2] G. P. Riyanto, "Jumlah Pengguna Internet Indonesia 2021 Tembus 202 Juta," *Kompas.com*, 2021. https://tekno.kompas.com/read/2021/02/23/16100057/jumlah-pengguna-internet-indonesia-2021-tembus-202-juta (accessed May 15, 2021).

[3] Cindy Mutia Annur, "Ada Berapa Pengguna Internet dan Media Sosial di Seluruh Dunia?," *katadata.co.id*, Dec. 2022. https://databoks.katadata.co.id/infografik/2022/12/08/ada-berapa-pengguna-internet-dan-media-sosial-di-seluruh-dunia (accessed Mar. 08, 2023).

[4] A. Taylor, "A study of the information search behaviour of the millennial generation.," *Information research: an international electronic journal*, vol. 17, no. 1, p. n1, 2012.

[5] E. S. W. Ng, L. Schweitzer, and S. T. Lyons, "New generation, great expectations: A field study of the millennial generation," *J Bus Psychol*, vol. 25, pp. 281–292, 2010, doi: https://doi.org/10.1007/s10869-010-9159-4.

[6] A. S. Hussein, R. S. Khairy, S. M. M. Najeeb, and H. T. ALRikabi, "Credit Card Fraud Detection Using Fuzzy Rough Nearest Neighbor and Sequential Minimal Optimization with Logistic Regression.," *International Journal of Interactive Mobile Technologies*, vol. 15, no. 5, 2021.

[7] S. Shukla, P. K. Jain, C. R. Babu, and R. Pamula, "A multivariate regression model for identifying, analyzing and predicting crimes," *Wirel Pers Commun*, vol. 113, pp. 2447–2461, 2020.

[8] E. Jumin *et al.*, "Machine learning versus linear regression modelling approach for accurate ozone concentrations prediction," *Engineering Applications of Computational Fluid Mechanics*, vol. 14, no. 1, pp. 713–725, 2020, doi: 10.1080/19942060.2020.1758792.

[9] T. E. E. Tju, D. S. Maylawati, G. Munawar, and S. Utomo, "Prediction of the COVID-19 Vaccination Target Achievement with Exponential Regression," *JISA (Jurnal Informatika dan Sains)*, vol. 4, no. 2, pp. 179–182, 2021.

[10] R. Irmanita, S. S. Prasetiyowati, and Y. Sibaroni, "Classification of Malaria Complication Using CART (Classification and Regression Tree) and Naive Bayes," *Jurnal RESTI (Rekayasa Sistem Dan Teknologi Informasi)*, vol. 5, no. 1, pp. 10–16, 2021.

[11] F. E. Pratiwi and I. Zain, "Klasifikasi Pengangguran Terbuka Menggunakan CART (Classification and Regression Tree) di Provinsi Sulawesi Utara," *Jurnal Sains dan Seni ITS*, vol. 3, no. 1, pp. D54–D59, 2014.

[12] S. I. Bangdiwala, "Regression: simple linear," *Int J Inj Contr Saf Promot*, vol. 25, no. 1, pp. 113–115, 2018.

[13] K. Park, R. Rothfeder, S. Petheram, F. Buaku, R. Ewing, and W. H. Greene, "Linear regression," in *Basic Quantitative Research Methods for Urban Planners*, 2020, pp. 220–269. doi: 10.4324/9780429325021-12.

[14] D. Maulud and A. M. Abdulazeez, "A Review on Linear Regression Comprehensive in Machine Learning," *Journal of Applied Science and Technology Trends*, vol. 1, no. 4, pp. 140–147, 2020, doi: 10.38094/jastt1457.

[15] S. Rath, A. Tripathy, and A. R. Tripathy, "Prediction of new active cases of coronavirus disease (COVID-19) pandemic using multiple linear regression model," *Diabetes and Metabolic Syndrome: Clinical Research and Reviews*, vol. 14, no. 5, pp. 1467–1474, 2020, doi: 10.1016/j.dsx.2020.07.045.

[16] X. Li *et al.*, "Evaluation model of fabric transient cooling sensation based on multiple stepwise regression analysis," *J Eng Fiber Fabr*, vol. 18, p. 15589250221144014, 2023, doi: https://doi.org/10.1177/15589250221144014.

[17] M. Arashi, M. Roozbeh, N. A. Hamzah, and M. Gasparini, "Ridge regression and its applications in genetic studies," *PLoS One*, vol. 16, no. 4, p. e0245376, 2021, doi: ttps://doi.org/10.1371/journal.pone.0245376.

[18] Y. Qu, C. Pan, S. Guo, and H. Wu, "Dietary Intake and Asthma in Preschoolers: A Logistic Lasso Regression Analysis," *Front Pediatr*, vol. 10, 2022, doi: https://doi.org/10.3389%2Ffped.2022.870529.

[19] Y. Wu *et al.*, "Detection of functional and structural brain alterations in female schizophrenia using elastic net logistic regression," *Brain Imaging Behav*, pp. 1–10, 2022, doi: https://doi.org/10.1007/s11682-021-00501-z.

[20] E. Jumin, F. B. Basaruddin, Y. B. M. Yusoff, S. D. Latif, and A. N. Ahmed, "Solar radiation prediction using boosted decision tree regression model: A case study in Malaysia," *Environmental Science*

and Pollution Research, vol. 28, pp. 26571–26583, 2021, doi: https://doi.org/10.1007/s11356-021-12435-6.

[21] D. R. Prado, J. A. López-Fernández, M. Arrebola, M. R. Pino, and G. Goussetis, "Wideband shaped-beam reflectarray design using support vector regression analysis," *IEEE Antennas Wirel Propag Lett*, vol. 18, no. 11, pp. 2287–2291, 2019, doi: https://doi.org/10.1109/LAWP.2019.2932902.

[22] X. Du, P. Wang, L. Fu, H. Liu, Z. Zhang, and C. Yao, "Determination of chlorpyrifos in pears by Raman spectroscopy with random forest regression analysis," *Anal Lett*, vol. 53, no. 6, pp. 821–833, 2020, doi: https://doi.org/10.1080/00032719.2019.1681439.