Application of K-Nearest Neighbor Algorithm for Classification of Mental Disabilities Patients Based on Age

Nurhidayati* Suhartini*
1,2Universitas Hamzanwadi, Lombok, Indonesia

*email : hidayati2188@gmail.com

Abstract. People with mental disabilities have the same position, rights and obligations as the non-disabled community. As part of Indonesian citizens, it is not appropriate for men and women with mental disabilities to be treated bad and experience discrimination. In Sukamulia Subdistrict, East Lombok, it is noted that the number of people with disabilities at the end of 2018 recorded 358 sufferers where people with disabilities are dominated by people over 18 years who are categorized as persons with disabilities. To find out the number of patients categorized as persons with disabilities, classification is needed. The purpose of this study was to classify people with disabilities based on age using the classification of KNN (K-Nearest Neighbor) Algorithm based on Age. Validity test is used to determine the best and most accurate level of accuracy. In this study the level of accuracy obtained is that in testing using K-Fold Validation 3 with a value of k = 3, the accuracy result is 97.49%.

Keyword : Mental disability, East Lombok Sukamulia District, K-Nearest Neighbor Classification.

1. Introduction
Namely mental mental disability, physical mental disability and multiple mental disabilities. Based on East Lombok Regency Social Service Data in 2018 especially in Sukamulia sub-district the number of mentally disabled people was 358 (three hundred and fifty-eight) sufferers. however, the amount of data to be processed as a test sample is only used by 99 people.[1]
By using a method that is a data mining approach by applying the K-Nearest Neighbor method to classify people with mental disabilities in Sukamulia District by classifying mental disabilities based on the number of people with mental and physical mental disabilities aged less than 18 years categorized as Children with Disabilities (AD) ) as well as those aged 18 years and over are categorized as Persons with Disabilities (PD) based on Age. [2]

2. Methods
The methods used in this study include:
Study of literature
At this stage the authors will search, learn from various kinds of literature and documents that support this research, especially those related to the k-nearest neighbor algorithm the emptied end which can be used after several stages and initial processing. for the verification of people with mental disabilities.[3]
Observation
Observing the data examined, conducting interviews with related parties in the social service office to find information needed in the preparation of this study such as the number of persons with disabilities in Sukamulia sub-district.[4]
Analysis of data that has been collected
The data used in this study are primary data obtained from the Office of the Social Office of East Lombok. The data of people with disabilities in this study were 358 records, 9 attributes and 1 attribute. [5]
System Design and Design
Below is a picture of the proposed system trial:
Analysis of data that has been collected

![Diagram of how the system works](image)

Figure 1: Schematic of how the system works.

3. Results and Discussion
Tests carried out to determine the effect of k on the level of accuracy.[6] The best k value in this algorithm depends on the type of data used. A good k value can be chosen with parameter optimization, for example by using cross-validation. In special cases, the classification is predicted based on the training data closest (in other words k = 1) so it is called the K-Nearest Neighbor algorithm. The k values (the closest neighbors) used in this study were 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 with cross validation 10.[7]

![Testing the K-Nearest Neighbor](image)

Figure 2. Testing the K-Nearest Neighbor

Here are the results of tests that have been carried out based on the picture above.
Figure 3 K-Nearest Neighbor (Training) Algorithm Test Results

From the testing process, the test results obtained with the highest value are at the value of $k = 3$, namely accuracy $97.48\%$, precision $94.00\%$, recall $87.00\%$ and AUC value $0.960$.[8]

1. Testing Phase
   a. K-Fold Validation 3
   In testing using K-Fold Validation 3 with a value of $k = 3$, we get an accuracy of $97.49\%$.

Figure 4. Accuracy Results with K-Fold Validation

Figure 5. Results of Precision K-Nearest Neighbor

Figure 6 Results of AUC K-Nearest Neighbor
The number of true positive (TP) is 308 records classified as disabled class and false negative (FN) of 3 records classified as disabled class but in reality the accuracy of the data analysis of mental disability patients in Sukamulia East Lombok Subdistrict using the K-Nearest Neighbor Algorithm. [9] The data analyzed is the patient data is a disability child. Next 41 positive false records (TN) are classified as children with disabilities class, and 6 false positive records (FP) are classified as disabled children but in fact are classified as people with disabilities. Based on Table 4.2 shows that the level of accuracy using the K-Nearest Neighbor algorithm with K-Fold Validation3 is 97.49%.[2]

The purpose of this study is to test mental disability. From the trial using k-fold validation 3 with a value of k = 3, the values of sensitivity, specificity, Ppv and Npv were found.[10]

| Table 1. Result of he tasting |
|-----------------------------|
| **K-Fold Validation 3**     |
| Accuracy: 97.49%            |
| Sensitivity: 99.04%         |
| Specificity: 87.23%         |
| Ppv: 98.09%                 |
| Npv: 93.18%                 |

Based on table 1 above, it can be concluded that KNN (K-Nearest Neighbor) has a good ability in solving data mining problems. Experiments using the best KNN (K-Nearest Neighbor) method resulted in an accuracy value of 97.49% and AUC of 0.960%.[9] The results were obtained by the K-Fold Validation 3 method with a value of k = 3, where the data is divided into 3 parts for training and testing. From this success it can be seen that the success of KNN (K-Nearest Neighbor) is strongly influenced by the selection of the right attributes by not using too many criteria that can cause overlap.[11]

**4. Conclusion**

From the results of the research that has been done, it can be concluded that the impact of mental disabilities sufferers is analyzed and evaluated by utilizing data mining techniques using the K-Nearest Neighbor algorithm.[12] from the data of mentally disabled persons it can be seen that the average disability sufferer is dominated by patients over 18 years namely the category of persons with disabilities.[13] Where this test is carried out by calculating in accordance with the performance of the K-NN (K-Nearest Neighbor) algorithm to produce a model so that it is included in the excellent classification category,[14] with proof of the experiments that have been carried out it can be known the results of the K-Fold Validation value3 with the accuracy obtained by 97.49% and AUC of 0.960[15]. Thus the K-Nearest Neighbor algorithm method is well used and is able to classify patients categorized as disabled persons and children, so in analyzing the data this method is accurate enough to analyze patients categorized as disabled persons and children.

**5. References**

[1] A. Jagati and P. Chowdhary, “Dermatologia of institutional records Ñ’, Ñ” An. Bras. Dermatol., no. xx, pp. 1–5, 2020.

[2] V. Robles-bykbaev, M. López-nores, and J. J. Pazos-arias, “Maturation assessment system for speech and language therapy based on multilevel PAM and KNN,” Procedia Technol., vol. 16, pp. 1265–1270, 2014.

[3] E. Stern, I. Van Der Heijden, and K. Dunkle, “How people with disabilities experience programs to prevent intimate partner violence across four countries,” Eval. Program Plann., vol. 79, no. July 2019, p. 101770, 2020.
[4] K. Riester, L. Kappos, K. Selmaj, S. Lindborg, and I. Lipkovich, “Impact of informative censoring on the treatment effect estimate of disability worsening in multiple sclerosis clinical trials,” *Mult. Scler. Relat. Disord.*, vol. 39, no. July 2019, p. 101865, 2020.

[5] S. Hong and D. Shin, “Affiliation:,” *Med. Hypotheses*, p. 109562, 2020.

[6] N. Rennó, G. Carvalho, and R. Peixoto, “Wage structure differential and disability in Brazil — Underperformance or discrimination?,” 2019.

[7] J. Erickson, K. Abbott, and L. Susienka, “Automatic address validation and health record review to identify homeless Social Security disability applicants,” *J. Biomed. Inform.*, vol. 82, no. January, pp. 41–46, 2018.

[8] “1-s2.0-S1877050915000320-main.pdf.”

[9] M. P. Volpato et al., “Single Cupping Therapy Session Improves Pain, Sleep, and Disability in Patients with Nonspecific Chronic Low Back Pain,” no. xxxx, 2019.

[10] S. Matsuyama et al., “Association between improved adherence to the Japanese diet and incident functional disability in older people: The Ohsaki Cohort 2006 Study,” *Clin. Nutr.*, no. xxxx, 2019.

[11] T. L. King, M. Shields, T. Shakespeare, A. Milner, and A. Kavanagh, “SSM - Population Health: An intersectional approach to understandings of mental health inequalities among men with disability,” *SSM - Popul. Heal.*, vol. 9, no. December 2018, p. 100464, 2019.

[12] C. G. Hudson and N. J. Doogan, “SSM - Population Health: The impact of geographic isolation on mental disability in the United States,” *SSM - Popul. Heal.*, vol. 8, no. June, p. 100437, 2019.

[13] T. Heller, H. M. Scott, M. P. Janicki, and P. W. Caregiving, “Caregiving, intellectual disability, and dementia: Report of the Summit Workgroup on Caregiving and Intellectual and Developmental Disabilities,” vol. 4, pp. 272–282, 2018.

[14] A. Kaehne, J. Kiernan, and J. Ridley, “Heliyon: Systematic review of study designs and methods in health transition research for young people with intellectual disabilities,” *Heliyon*, vol. 5, no. July, p. e02750, 2019.

[15] M. Hendel, A. Benyettou, and F. Hendel, “Informatics in Medicine Unlocked: Hybrid self-organizing map and probabilistic quadratic loss multi-class support vector machine for mental tasks classification,” *Informatics Med. Unlocked*, vol. 4, pp. 1–9, 2016.