



The usefulness of machine learning analysis for predicting the presence of depression with the results of the Korea National Health and Nutrition Examination Survey

Sang Won Kim¹, Min Cheol Chang²

¹Medical Research Center, College of Medicine, Yeungnam University, Daegu, Republic of Korea; ²Department of Rehabilitation Medicine, College of Medicine, Yeungnam University, Daegu, Republic of Korea

Contributions: (I) Conception and design: Both authors; (II) Administrative support: Both authors; (III) Provision of study materials or patients: Both authors; (IV) Collection and assembly of data: Both authors; (V) Data analysis and interpretation: Both authors; (VI) Manuscript writing: Both authors; (VII) Final approval of manuscript: Both authors.

Correspondence to: Min Cheol Chang, MD. Department of Rehabilitation Medicine, College of Medicine, Yeungnam University, 317-1, Daemyungdong, Namku, Daegu 705-717, Republic of Korea. Email: wheel633@gmail.com.

Background: Depression is a major public health concern, with an estimated 10.8% of adults experiencing depression. Depression can have a significant impact on an individual's quality of life, social function, and productivity. Early diagnosis of depression is important in preventing its progression. Several tools, such as the Patient Health Questionnaire-9 (PHQ-9) and Beck Depression Inventory, are used to screen patients for depression. We investigated the potential of machine learning in predicting the presence of depression using the results of a national survey.

Methods: We collected the data of 5,420 patients from the 2020 Korea National Health and Nutrition Examination. The presence of depression was defined as ≥ 5 PHQ-9. We categorized output variables into the presence of depression (PHQ-9, ≥ 5) and absence of depression (PHQ-9, < 5). We used 20 variables related to sociodemographic characteristics, health behavior, and presence of chronic disease for the development of three machine learning algorithms [random forest, logistic regression, and deep neural network (DNN)]. Eighty-seven decision trees were used for the random forest model. Linear regression algorithm shows a linear relationship between various input and output variables. For the DNN model, three layers with 16-32-64 neurons, Adam optimizer, and rectified linear unit (ReLU) activation were used. Of the included samples, 70% and 30% were randomly divided into the training and test sets, respectively.

Results: The area under the curve (AUC) of the test dataset for the random forest model was 0.803 [95% confidence interval (CI), 0.776–0.829], 0.812 (95% CI, 0.787–0.837) for the logistic regression model, and 0.805 (95% CI, 0.780–0.831) for the DNN model.

Conclusions: Our study demonstrated the potential of machine learning for the development of models for predicting the presence of depression based of various health-related data. Machine learning models can potentially overcome the limitations of traditional diagnostic methods for depression by incorporating a wide range of objective variables to accurately identify patients with depression, thus avoiding the subjectivity and potential diagnostic errors associated with the subjective interpretation of symptoms observed by a clinician. Further efforts to increase the accuracy of machine learning models by utilizing more variables and data needed to detect depression.

Keywords: Depression; machine learning; deep learning; prediction; South Korea

Submitted Jan 19, 2023. Accepted for publication May 09, 2023. Published online Jun 19, 2023.

doi: 10.21037/apm-23-78

View this article at: <https://dx.doi.org/10.21037/apm-23-78>

Introduction

Depressive is a major public health concern in the general population (1). Depression can significantly impact an individual's quality of life, social functioning, and productivity (2-4). The global prevalence of depression was reported to be as high as 10.8% (5). Depression can be cured with prompt treatment; however, if left untreated, it can lead to suicidal behavior (6). Therefore, screening and early diagnosis of depression is important for preventing the progression of the condition and the development of related symptoms.

To date, there have been several studies on the prediction of the presence of depression (7-10). Most of these previous studies screened or diagnosed depression based on clinical symptoms and imaging modalities, and these studies had small sample sizes (7-10). In addition, owing to the limitations of existing traditional statistical methods, these studies used a limited number of variables for determining the presence of depression in recruited individuals.

Various factors, such as female sex, old age, marital status, low socioeconomic status, unemployment, social isolation, poor housing, problems with alcohol use, stress, and underlying diseases, have been reported to be risk factors for depression (11). The risk of each of these factors for depression has been individually investigated (12-16). The consideration of all the variables related to the occurrence of depression can increase the accuracy of detecting the presence of depression. However, when multiple variables are considered for analysis, the use of conventional statistical analysis methods is inappropriate (17).

Machine learning is a technique of artificial intelligence

in which a system learns patterns and rules from given information (18-22). Machine learning has several advantages regarding the detection of possible interactions between a large number of variables (18-22). The usefulness of machine learning for the diagnosis of various disorders, using a large number of variables as input data, has been demonstrated (23-26). We hypothesized that machine learning techniques can effectively detect the presence of depression based on a large number of factors related to the occurrence of depression.

Recently, health and lifestyle surveys in many countries and organizations have been conducted to elucidate appropriate health policies. The Korea National Health and Nutrition Examination Survey (KNHANES) is an ongoing surveillance system that began in 1998 to provide nationwide statistics on the health status and behavior of the South Korean population (27). The data obtained from KNHANES are utilized as evidence for the evaluation and formulation of health policies (27). In addition, the data are publicly available to researchers, and have thus been analyzed and studied by many researchers in South Korea. In the present study, we obtained various health and lifestyle-related data from 1998 to 2021 from KNHANES; we collected data from a cumulative total of 127,545 cases (27). Therefore, data from the KNHANES are being utilized in big data and machine learning studies.

Herein, we investigate the potential of machine learning to predict the presence of depression with the national survey data (KNHANES). We present this article in accordance with the STARD reporting checklist (available at <https://apm.amegroups.com/article/view/10.21037/apm-23-78/rc>).

Methods

Data collection

We collected data from the 2020 KNHANES, which is an ongoing surveillance system that was initiated to provide nationwide statistics regarding the health behavior, health status, and food and nutrient intake of the Korean population (28). To evaluate the health and nutritional status of the general South Korean population, a nationwide sampling method (clustered, multistage, stratified, and randomized) was applied for proportional distribution according to sex, age, and geographical area. The survey participants differed every year and were not serially monitored, resulting in random sampling every

Highlight box

Key findings

- Our study showed that machine learning may be useful for the development of models for predicting depression based on various health-related data.

What is known and what is new?

- The usefulness of machine learning for the diagnosis of various disorders is well-demonstrated.
- The ability of machine learning-based models for predicting the presence of depression was excellent.

What is the implication, and what should change now?

- The use of machine learning can improve the accuracy of depression screening in community residents.

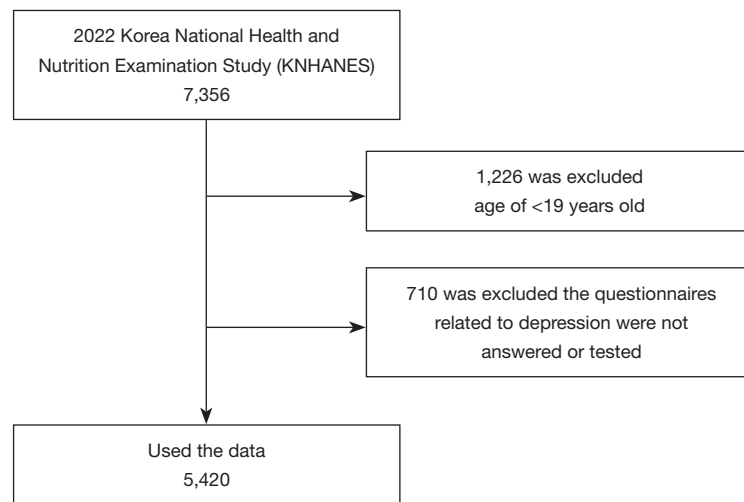


Figure 1 Flow chart depicting sample selection.

year. The KNHANES evaluates data from three sources: health questionnaires, health and physical examinations, and nutrition questionnaires administered by experienced interviewers, registered nurses, and laboratory technicians (27–29). The response rate was 70–80% (27).

Participants

In the present study, individuals who were ≥ 19 years old and responded to questions that focused on depression were included (Figure 1). Of the 7,356 participants in the 2020 KNHANES, we excluded 1,936 participants (1,226 were excluded due to the age < 19 years and 710 because they did to answer questions related to depression). The study was conducted in accordance with the Declaration of Helsinki (as revised in 2013), and was approved by the ethics board of Yeungnam University Hospital (No. 2021-12-009). Patient consent was waived due to the retrospective nature of the study.

Input variables

We used the following variables for the development of machine learning algorithms: sociodemographic characteristics (age, sex, marital status, income, education level, region, town, housing type, and job), health behavior (stress, drinking, obesity, use of medication, sleep time, and activity restriction), and presence of chronic disease (hypertension, dyslipidemia, myocardial infarction or angina, diabetes mellitus, and arthritis). Accordingly, 20

variables were used as input data.

Presence of depression (output variables)

The nine-item version of the Patient Health Questionnaire-9 (PHQ-9) was used to measure depression. The PHQ-9 is a depression module of PHQ that consists of the nine items upon which the diagnosis of DSM-IV depressive disorder is based. Each item is scored from 0 (not at all) to 3 (nearly every day), and the PHQ-9 score ranges from 0 to 27 (30). The presence of depression was defined as a score of 5 or higher on the PHQ-9 (30). We categorized output variables as follows: presence of depression (PHQ-9 score ≥ 5) and absence of depression (PHQ-9 score < 5).

Machine learning algorithms

We used three machine learning algorithms: random forest, logistic regression, and deep neural network (DNN). The machine learning models were trained with all variables as inputs to determine the presence or absence of depression. The random forest algorithm comprises several decision trees that consist of multiple true or false conditions using input variables (31). The sum of the decisions made by the decision trees is used for the final classification. For our random forest model, 87 decision trees were used. Linear regression algorithm shows a linear relationship between various input and output variables (32). The DNN is composed of a series of artificial neurons that are interconnected through multiple layers (33). The DNN

is designed based on the biological neuron and receives multiple inputs multiplied by weights, and outputs the sum of the inputs. For the DNN model, three layers with 16-32-64 neurons, Adam optimizer, and rectified linear unit (ReLU) activation were used. To prevent overfitting, we used only three layers, applied dropout regulation and early stopping, and withheld the validation and test datasets to check potential overfitting.

Of the 5,420 included samples, 70% (n=3,794) and 30% (n=1,626) were randomly divided into the training and test sets, respectively. TensorFlow version 1.1.0 (Google, Mountain View, CA, USA) and scikit-learn toolkit version 0.18.1 (Google) were used to train the machine learning models.

Statistical analysis

Statistical analyses were performed using Python 3.7.9 and scikit-learn version 0.23.2. Receiver operating characteristic curve analysis was employed, and the area under the curve (AUC) was calculated. The confidence interval (CI) for the AUC was calculated using the method employed by DeLong *et al.* (34).

Results

Finally, we used the data of 5,420 participants. All participants received a full explanation of the aims and protocol of the KNHANES and they provided written informed consent. *Table 1* shows the details of the study participants. Of the 5,420 included individuals, 4,138 did not have depression and 1,282 had depression. The AUC of the test dataset was 0.803 (95% CI, 0.776–0.829) in the random forest model, 0.812 (95% CI, 0.787–0.837) for the logistic regression model, and 0.805 (95% CI, 0.780–0.831) for the DNN model (*Table 2, Figure 2*).

Discussion

In the present study, we developed machine learning-based models for predicting the presence of depression. The machine learning models were built using random forest, linear regression, and DNN, and the AUCs were 0.803, 0.812, and 0.805, respectively. Considering that AUC values of 0.7–0.8 are acceptable, 0.8–0.9 are excellent, and

Table 1 Characteristics of the included individuals

Characteristics	N (%) or mean \pm SD
Sex	
Male	2,438 (45.0)
Female	2,982 (55.0)
Age (years)	51.4 \pm 17.1
Marital status	
Married	4,293 (79.2)
Not married	1,127 (20.8)
Income	
Lowest	725 (13.4)
Low	955 (17.6)
Middle	1,123 (20.7)
High	1,288 (23.8)
Highest	1,329 (24.5)
Education	
Elementary school	919 (17.0)
Middle school	523 (9.6)
High school	1,900 (35.1)
University	2,078 (38.3)
Town	
Urban	4,336 (80.0)
Rural	1,084 (20.0)
Housing type	
General housing	2,470 (45.6)
Apartments	2,950 (54.4)
Job	
Yes	3,240 (59.8)
No	2,180 (41.2)
Stress recognition	
Yes	1,496 (27.6)
No	3,924 (72.4)
Depression	
Yes	1,282 (23.6)
No	4,138 (77.4)

Table 1 (continued)

Table 1 (continued)

Characteristics	N (%) or mean \pm SD
Drinking	
Yes	2,782 (51.3)
No	2,638 (48.7)
Obesity	
Low weight	210 (3.9)
Normal	1,883 (35.2)
Obesity	2,036 (60.9)
Use of medication	
Yes	385 (7.1)
No	4,838 (89.2)
Not applicable	197 (3.7)
Activity restriction	
Yes	399 (7.4)
No	5,020 (92.6)
Hypertension	
Yes	1,368 (25.2)
No	4,052 (74.8)
Dyslipidemia	
Yes	1,190 (22.0)
No	4,230 (78.0)
Myocardial infarction or angina	
Yes	180 (3.3)
No	5,240 (96.7)
Diabetes mellitus	
Yes	591 (10.9)
No	4,829 (89.1)
Arthritis	
Yes	709 (13.1)
No	4,711 (86.9)

SD, standard deviation.

those above 0.9 are outstanding, the ability of the models to predict the presence of depression was excellent (35).

Random forest is an ensemble learning method that forms multiple decision trees (31). The input data are entered into and pass through each decision tree

Table 2 Outcomes of three developed models

Machine learning model	Details
Random forest	87 estimators Mean test accuracy score: 81.1% Mean training accuracy score: 89.0% Test AUC 0.803 (95% CI, 0.776–0.829) Sensitivity: 0.814, specificity: 0.779
Linear regression	Mean test accuracy score: 81.0% Mean training accuracy score: 80.1% Test AUC 0.812 (95% CI, 0.787–0.837) Sensitivity: 0.809, specificity: 0.786
DNN	3 layers with 16-32-64 neurons, Adam optimizer, and ReLU activation Mean test accuracy score: 81.2% Mean training accuracy score: 79.9% Test AUC 0.805 (95% CI, 0.780–0.831) Sensitivity: 0.810, specificity: 0.787

AUC, area under the curve; CI, confidence interval; DNN, deep neural network; ReLU, rectified linear unit.

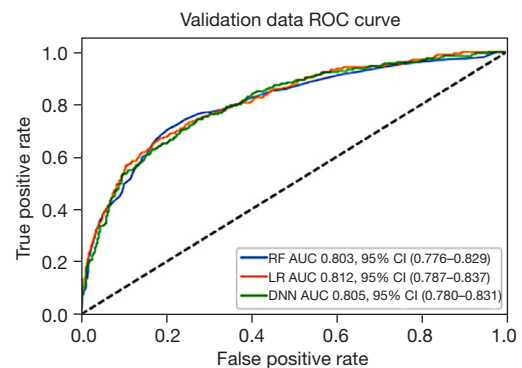


Figure 2 Receiver operating characteristic curve for the models for test datasets. ROC, receiver operating characteristic; RF, random forest; AUC, area under the curve; CI, confidence interval; LR, linear regression; DNN, deep neural network.

simultaneously, and the most accurate or appropriate result is selected as the final decision (31). The model shows high performance in classification tasks. In addition, it can provide information about the importance of each feature used in the model, which can be helpful in understanding the underlying relationships between input and output

variables (31). Linear regression is a statistical method for modeling the relationship between input and target variables (32). It provides both classification and probability, and improves the understanding of the contribution of each variable to the final fit (33). Furthermore, it quickly analyzes the relationship among the various variables. By showing the formula used, linear regression allows researchers to relatively easily understand the process that produced the result after the analysis. DNN includes multiple hidden layers between input and output layers (33). The multiple hidden layers form complex networks, which are efficient for representing the complex characteristics of input data (33). DNN can adapt to new data or new tasks by adjusting weights and parameters, a feature that makes DNN models flexible and able to handle a wide range of machine learning tasks (33). Our developed DNN model may recognize variables related to depression and weighted variables that were more strongly related to the occurrence of depression.

For the development of the machine learning models, sociodemographic characteristics, health behavior, and presence of chronic disease were used as input data. We believe that each input data would have played an important role in the development of depression, although the degree of involvement in the occurrence of depression was different. If more variables are used as input data for the development of machine learning models, the ability of the models to screen depression can be improved. Early screening and detection of depression among community dwellers are important, and many countries are focusing on screening for the presence of depression in community settings (36,37). The KNHANES is a surveillance system for investigating health-related aspects, with a focus on the community dwellers of South Korea (27). We developed the machine learning models for predicting depression by using the KNHANES data as input data, and the three models showed high accuracies.

Since the KNHANES provides nationally representative health survey data, machine learning models created using the data ensure the representativeness of the entire population's overall health status (38). In addition, KNHANES collects various variables such as health, nutrition, and lifestyle; therefore, the use of this data to develop machine learning models can guarantee diversity in predictive and independent variables (38).

The clinical relevance of the outcome of our study lies in the potential of machine learning models to improve the accuracy and speed of diagnosis of depression. Traditional

methods for diagnosing depression rely on the subjective interpretation of symptoms observed by a clinician, which can lead to diagnostic errors and delays in the management of depression. Machine learning models have the potential to overcome these limitations by incorporating a wide range of objective variables to accurately identify patients with depression.

Recently, several studies have evaluated the capacity of machine learning to detect depression. Gil *et al.* predicted the presence of depression using machine learning based on 171 family and individual factors, consisting of a set of demographic and health-related behaviors, in 513 individuals (39). Gil *et al.* used three machine learning models: sparse logistic regression, support vector machine, and random forest. The predictive accuracies of these models were 0.784, 0.804, and 0.863, respectively (39). Lee *et al.* developed prediction models for depression among 8,628 individuals with hypertension using various machine learning models, such as DNN, random forest, AdaBoost, stochastic gradient boosting, XGBoost, and support vector machines (40). They developed models using data (sociodemographic, behavioral, and clinical data) obtained from the National Health and Nutrition Examination Survey conducted in the United States. The supported vector machine model showed the highest accuracy (0.771) in predicting the presence of depression. In addition, Miao *et al.* evaluated the capacity of machine learning to predict depression and anxiety based on gait patterns (41). A digital camera recorded each patient's gait pattern during walking, and positional and temporal data from 18 key body points were obtained. The accuracies of the machine learning algorithms that predicted the presence of depression and anxiety were 86% and 78%, respectively. Miao *et al.* reported that individuals with depression and anxiety had different walking movement patterns from individuals without depression and anxiety.

Conclusions

In conclusion, we believe that our study showed that machine learning can be useful for the development of models for predicting the presence of depression on the basis of various health-related data. In addition, we believe that the use of machine learning can enhance the accuracy of screening for the presence of depression among community dwellers. However, our study has some limitations. First, our study utilized a limited set of input variables. The inclusion of additional data, such as

interpersonal relationships, could potentially enhance the accuracy of the machine learning model. Second, we did not use a variety of machine learning models. We developed the algorithms using random forest, linear regression, and DNN because they are the most representative machine learning models. However, the use of newly developed models would help improve the predictive ability of the machine learning model. Third, we used PHQ-9, a screening tool, to determine the presence or absence of depression. Strictly, depression cannot be definitely confirmed with PHQ-9. Fourth, the inclusion and exclusion criteria were not strictly applied in this study. For example, people with cognitive impairment or mental illness may not have responded faithfully. Further studies addressing these limitations are thus warranted.

Acknowledgments

Funding: None.

Footnote

Reporting Checklist: The authors have completed the STARD reporting checklist. Available at <https://apm.amegroups.com/article/view/10.21037/apm-23-78/rc>

Conflicts of Interest: Both authors have completed the ICMJE uniform disclosure form (available at <https://apm.amegroups.com/article/view/10.21037/apm-23-78/coif>). The authors have no conflicts of interest to declare.

Ethical Statement: The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved. The study was conducted in accordance with the Declaration of Helsinki (as revised in 2013), and was approved by the ethics board of Yeungnam University Hospital (No. 2021-12-009). Patient consent was waived due to the retrospective nature of the study.

Open Access Statement: This is an Open Access article distributed in accordance with the Creative Commons Attribution-NonCommercial-NoDerivs 4.0 International License (CC BY-NC-ND 4.0), which permits the non-commercial replication and distribution of the article with the strict proviso that no changes or edits are made and the original work is properly cited (including links to both the

formal publication through the relevant DOI and the license). See: <https://creativecommons.org/licenses/by-nc-nd/4.0/>.

References

1. Lee J, Kim H, Hong JP, et al. Trends in the Prevalence of Major Depressive Disorder by Sociodemographic Factors in Korea: Results from Nationwide General Population Surveys in 2001, 2006, and 2011. *J Korean Med Sci* 2021;36:e244.
2. Brenes GA. Anxiety, depression, and quality of life in primary care patients. *Prim Care Companion J Clin Psychiatry* 2007;9:437-43.
3. Shin J, Cho E. Trajectories of depressive symptoms among community-dwelling Korean older adults: findings from the Korean longitudinal study of aging (2006-2016). *BMC Psychiatry* 2022;22:246.
4. Woo JM, Kim W, Hwang TY, et al. Impact of depression on work productivity and its improvement after outpatient treatment with antidepressants. *Value Health* 2011;14:475-82.
5. Lim GY, Tam WW, Lu Y, et al. Prevalence of Depression in the Community from 30 Countries between 1994 and 2014. *Sci Rep* 2018;8:2861.
6. Halfin A. Depression: the benefits of early and appropriate treatment. *Am J Manag Care* 2007;13:S92-7.
7. Bifulco A, Brown GW, Moran P, et al. Predicting depression in women: the role of past and present vulnerability. *Psychol Med* 1998;28:39-50.
8. Cohen SE, Zantvoord JB, Wezenberg BN, et al. Magnetic resonance imaging for individual prediction of treatment response in major depressive disorder: a systematic review and meta-analysis. *Transl Psychiatry* 2021;11:168.
9. Frässle S, Marquand AF, Schmaal L, et al. Predicting individual clinical trajectories of depression with generative embedding. *Neuroimage Clin* 2020;26:102213.
10. Shoib S, Das S. Factors predicting the presence of depression in obstructive sleep apnea. *Ind Psychiatry J* 2020;29:29-32.
11. King M, Walker C, Levy G, et al. Development and validation of an international risk prediction algorithm for episodes of major depression in general practice attendees: the PredictD study. *Arch Gen Psychiatry* 2008;65:1368-76.
12. Bruce ML, Hoff RA. Social and physical health risk factors for first-onset major depressive disorder in a community sample. *Soc Psychiatry Psychiatr Epidemiol*

- 1994;29:165-71.
13. Stansfeld SA, Fuhrer R, Shipley MJ, et al. Work characteristics predict psychiatric disorder: prospective results from the Whitehall II Study. *Occup Environ Med* 1999;56:302-7.
 14. Weich S, Lewis G. Material standard of living, social class, and the prevalence of the common mental disorders in Great Britain. *J Epidemiol Community Health* 1998;52:8-14.
 15. Weich S, Lewis G. Poverty, unemployment, and common mental disorders: population based cohort study. *BMJ* 1998;317:115-9.
 16. Weich S, Sloggett A, Lewis G. Social roles and gender difference in the prevalence of common mental disorders. *Br J Psychiatry* 1998;173:489-93.
 17. Choo YJ, Chang MC. Use of Machine Learning in Stroke Rehabilitation: A Narrative Review. *Brain Neurorehabil* 2022;15:e26.
 18. Choo YJ, Kim JK, Kim JH, et al. Machine learning analysis to predict the need for ankle foot orthosis in patients with stroke. *Sci Rep* 2021;11:8499.
 19. Kim JK, Choo YJ, Chang MC. Prediction of Motor Function in Stroke Patients Using Machine Learning Algorithm: Development of Practical Models. *J Stroke Cerebrovasc Dis* 2021;30:105856.
 20. Kim JK, Choo YJ, Park IS, et al. Deep-Learning Algorithms for Prescribing Insoles to Patients with Foot Pain. *Appl Sci* 2023;13:2208.
 21. Kim JK, Lv Z, Park D, et al. Practical Machine Learning Model to Predict the Recovery of Motor Function in Patients with Stroke. *Eur Neurol* 2022;85:273-9.
 22. Sarker IH. Machine Learning: Algorithms, Real-World Applications and Research Directions. *SN Comput Sci* 2021;2:160.
 23. Kumar N, Narayan Das N, Gupta D, et al. Efficient Automated Disease Diagnosis Using Machine Learning Models. *J Healthc Eng* 2021;2021:9983652.
 24. Lee GW, Shin H, Chang MC. Deep learning algorithm to evaluate cervical spondylotic myelopathy using lateral cervical spine radiograph. *BMC Neurol* 2022;22:147.
 25. Myszczyńska MA, Ojamies PN, Lacoste AMB, et al. Applications of machine learning to diagnosis and treatment of neurodegenerative diseases. *Nat Rev Neurol* 2020;16:440-56.
 26. Ya Y, Ji L, Jia Y, et al. Machine Learning Models for Diagnosis of Parkinson's Disease Using Multiple Structural Magnetic Resonance Imaging Features. *Front Aging Neurosci* 2022;14:808520.
 27. Kweon S, Kim Y, Jang MJ, et al. Data resource profile: the Korea National Health and Nutrition Examination Survey (KNHANES). *Int J Epidemiol* 2014;43:69-77.
 28. Oh K, Kim Y, Kweon S, et al. Korea National Health and Nutrition Examination Survey, 20th anniversary: accomplishments and future directions. *Epidemiol Health* 2021;43:e2021025.
 29. Park HJ, Choi JY, Lee WM, et al. Prevalence of chronic low back pain and its associated factors in the general population of South Korea: a cross-sectional study using the National Health and Nutrition Examination Surveys. *J Orthop Surg Res* 2023;18:29.
 30. Han C, Jo SA, Kwak JH, et al. Validation of the Patient Health Questionnaire-9 Korean version in the elderly population: the Ansan Geriatric study. *Compr Psychiatry* 2008;49:218-23.
 31. Louppe G. Understanding random forests: From theory to practice. *arXiv preprint, arXiv:1407.7502*, 2014.
 32. Schneider A, Hommel G, Blettner M. Linear regression analysis: part 14 of a series on evaluation of scientific publications. *Dtsch Arztebl Int* 2010;107:776-82.
 33. Alzubaidi L, Zhang J, Humaidi AJ, et al. Review of deep learning: concepts, CNN architectures, challenges, applications, future directions. *J Big Data* 2021;8:53.
 34. DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. *Biometrics* 1988;44:837-45.
 35. Kim JK, Chang MC, Park D. Deep Learning Algorithm Trained on Brain Magnetic Resonance Images and Clinical Data to Predict Motor Outcomes of Patients With Corona Radiata Infarct. *Front Neurosci* 2022;15:795553.
 36. Eack SM, Singer JB, Greeno CG. Screening for anxiety and depression in community mental health: the beck anxiety and depression inventories. *Community Ment Health J* 2008;44:465-74.
 37. Tai SY, Ma TC, Wang LC, et al. A community-based walk-in screening of depression in Taiwan. *ScientificWorldJournal* 2014;2014:184018.
 38. Oh T, Kim D, Lee S, et al. Machine learning-based diagnosis and risk factor analysis of cardiocerebrovascular disease based on KNHANES. *Sci Rep* 2022;12:2250.
 39. Gil M, Kim SS, Min EJ. Machine learning models for predicting risk of depression in Korean college students: Identifying family and individual factors. *Front Public Health* 2022;10:1023010.
 40. Lee C, Kim H. Machine learning-based predictive

- modeling of depression in hypertensive populations. *PLoS One* 2022;17:e0272330.
41. Miao B, Liu X, Zhu T. Automatic mental health identification method based on natural gait pattern. *Psych J* 2021;10:453-64.

Cite this article as: Kim SW, Chang MC. The usefulness of machine learning analysis for predicting the presence of depression with the results of the Korea National Health and Nutrition Examination Survey. *Ann Palliat Med* 2023;12(4):748-756. doi: 10.21037/apm-23-78