

Article

Explainable Deep-Learning-Based Depression Modeling of Elderly Community after COVID-19 Pandemic

Hung Viet Nguyen  and Haewon Byeon * 

Department of Digital Anti-Aging Healthcare (BK21), Inje University, Gimhae 50834, Republic of Korea

* Correspondence: bhwpuma@naver.com; Tel.: +82-10-7404-6969

Abstract: The impact of the COVID-19 epidemic on the mental health of elderly individuals is causing considerable worry. We examined a deep neural network (DNN) model to predict the depression of the elderly population during the pandemic period based on social factors related to stress, health status, daily changes, and physical distancing. This study used vast data from the 2020 Community Health Survey of the Republic of Korea, which included 97,230 people over the age of 60. After cleansing the data, the DNN model was trained using 36,258 participants' data and 22 variables. We also integrated the DNN model with a LIME-based explainable model to achieve model prediction explainability. According to the research, the model could reach a prediction accuracy of 89.92%. Furthermore, the F1-score (0.92), precision (93.55%), and recall (97.32%) findings showed the effectiveness of the proposed approach. The COVID-19 pandemic considerably impacts the likelihood of depression in later life in the elderly community. This explainable DNN model can help identify patients to start treatment on them early.

Keywords: deep learning; deep neural network; LIME; explainable AI; depression; post-COVID-19

MSC: 68T01; 68T09; 68T07



Citation: Nguyen, H.V.; Byeon, H. Explainable Deep-Learning-Based Depression Modeling of Elderly Community after COVID-19 Pandemic. *Mathematics* **2022**, *10*, 4408. <https://doi.org/10.3390/math10234408>

Academic Editors: James Liou and Artūras Kaklauskas

Received: 14 October 2022

Accepted: 21 November 2022

Published: 23 November 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

As the infectious disease caused by the coronavirus (COVID-19) rapidly spread around the world, in March 2020 the World Health Organization declared a new infectious disease pandemic [1]. COVID-19 is a respiratory infection syndrome caused by infection with the SARS-CoV-2 virus. COVID-19 is spread through droplets from coughing or sneezing or by touching objects contaminated with the virus and then touching the eyes, nose, or mouth. However, COVID-19 is a high-risk infectious disease that causes severe respiratory illnesses, such as pneumonia, or death. In particular, it is perilous for people over 65 with underlying medical conditions, such as high blood pressure [2], diabetes [3], chronic cardiovascular disease [4], and obstructive pulmonary disease [5]. To protect against COVID-19, limiting face-to-face contact with others is the best way to reduce the spread of the virus.

However, the elderly are suddenly cut off from the outside environment and social relationships due to social distancing [6]. As the COVID-19 situation continues, older adults experience more limitations in going out and in daily activities, and depression increases as a result [7]. In particular, since the elderly are classified as a high-risk group for infection, they are undoubtedly more susceptible to exposure to the COVID-19 virus [8]. They are also restricted from daily activities because they are more fearful of COVID-19 than populations such as adolescents and adults. However, while containment of the physical environment is an effective way to prevent infectious diseases, social distancing increases psychological problems for which older people need to receive emotional support. Anxiety, stress, and life activities were all identified as contributing reasons for geriatric depression during the COVID-19 epidemic [9]. However, there is a lack of research on how the spread of COVID-19 affects the elderly regarding their subjective health, stress levels,

daily routines, and social distancing. Therefore, we suggest a deep neural network (DNN) model to predict the depression of the elderly during the pandemic period based on social factors related to stress, health status, daily changes, and physical distancing.

Artificial intelligence (AI) algorithms are frequently viewed as mysterious black boxes that make illogical choices. The idea that a machine-learning model and its output can be explained in a way that “makes sense” to a human being at an acceptable level is known as explainability (also called “interpretability”) [10]. While they may be less effective, some classes of algorithms, such as more conventional machine-learning algorithms, tend to be easier to understand. While some are more efficient than others, such as deep-learning systems, they are still challenging to understand. The ability to better understand AI systems is still a topic of active research. There are numerous approaches for making the output of a black-box classification model more interpretable [11–14]. One strategy is providing a global explanation by cataloging traits that are often more significant when generating a prediction [15]. However, it is preferable to have an instance-specific justification in healthcare. Consequently, patient can receive more tailored treatment as a result of increased autonomy in decision making.

In order to obtain an explanation for one particular occurrence, any black-box classification model may be utilized with a newly created framework known as local interpretable model-agnostic explanation (LIME) [16]. This method works by providing local justification for the classification and identifying the fewest characteristics that maximally contribute to the likelihood of a single-class result for a single observation. Despite the fact that LIME has previously been used with healthcare classification models, nothing is known about how well healthcare professionals can understand and accept LIME explanations.

In this paper, we provide a combination model that combines a LIME-based explainability model with a deep-learning model. The integrated model can explain a DNN classifier’s predictions understandably and accurately. Additionally, our explainability model may help psychologists and psychiatrists identify depression in seniors following the COVID-19 epidemic. The construction of this study is as follows: Chapter 2 explains the sources of data, analyzed variables, and model evaluation and explains the procedure of the model development. Chapter 3 compares the results of the developed prediction model with those of ML classifier models, such as logistic regression, gradient-boosting classifier, extra-gradient-boosting classifier, K-neighbors classifier, and support vector machine with kernel radial basis function. Lastly, chapter 4 presents directions for future studies.

2. Materials and Methods

2.1. Materials

This study used raw data from the Korea Centers for Disease Control and Prevention’s 2020 Community Health Survey. Since 2008, the local social health survey has been conducted at public health centers across the country with the goal of identifying local health statistics in order to establish and evaluate local health and medical plans, as well as to ask health-related questions at the regional level. As of July 2020, the target population resided in the Republic of Korea. The survey’s target population was adults aged 19 and up who lived in a chosen sample household at the time of the survey. A detailed description of sampling in the Community Health Survey is described in Kang et al. [17]. The community health survey was conducted by a trained researcher visiting the sample households directly from 16 August to 31 October 2020. The researcher went to 18 different locations to conduct face-to-face computer-assisted personal interviewing consisting of 142 questions. In the 2020 Community Health Survey, 229,269 people took part, with 97,230 of them being over the age of 60. However, 36,258 people with the presence of 22 variables were used for the final analysis.

2.2. Data Preprocessing

The dataset needed to be preprocessed before being fitted into the model. The dataset was preprocessed in this study to handle missing values, rebalance variables, and encode

labels. In order to handle missing values, rows with missing values and columns with 50% null values were removed from the dataset. The dataset was reduced to 42,788 samples and 22 variables after removing missing values. As the target variable, we chose the feature called “dep,” which meant “experienced depression” (yes or no).

Only 2501 of the 42,788 samples in the dataset experienced depression, whereas the total number of individuals without experiences of depression was 40,287 (94% of the target feature). As a result, we had to rebalance the dataset using both the oversampling and undersampling methods. The minority samples’ size was increased at random by 30% of the majority goal, and then the majority samples’ size was lowered until the number of majority samples was equal to twice the minority in the current dataset. After rebalancing the dataset, the number of “experienced depression” samples was 12,086, and the number of “no experience with depression” samples was 24,172. Finally, the dataset had a total of 36,258 samples with 22 variables encoded, as shown in Table 1.

Table 1. Variables and their descriptions.

Variables	Description	Field Type
covid_car	Changes in public transportation use since COVID-19	Categorical: increase, similar, decrease
life_satis_N	COVID-19 life satisfaction	Categorical: dissatisfied, satisfied
sex	Gender	Categorical: male, female
dong_ty_code	Urban residence type	Categorical: city, rural
edu_N	Educational level	Categorical: elementary school graduation or less, middle school graduation, high school graduation, college graduations or higher
income_N	Household income	Categorical: less than 1 million, 1–3 million, 3–5 million, 5 million or more
smoke_N	Smoking	Categorical: current smoker, past smoker, non-smoker
exer_N	Moderate physical activity per week	Categorical: none, yes
Sub_health_N	Subjective health	Categorical: good, average, bad
anx_dis	Concerned about COVID-19 infection	Categorical: yes, neutral, no
anx_death	Concerned about COVID-19 death	Categorical: yes, neutral, no
anx_sham	Concerned about criticism regarding COVID-19	Categorical: yes, neutral, no
anx_exp	Concerned about COVID-19 infection for vulnerable people	Categorical: yes, neutral, no
anx_econ	Concerned about economic damage from COVID-19	Categorical: yes, neutral, no
covid_phy	Change in number of encounters since COVID-19	Categorical: increase, similar, decrease
covid_sleep	Change in sleep-time activity after COVID-19	Categorical: increase, similar, decrease
quarantine	Isolation (hospitalization) due to COVID-19	Categorical: yes, no
help	People who can be asked for help	Categorical: none, 1–2 people, 3–5 people, 6 people or more
age_60n	Age	Categorical: 60–74 years, 75 years and older
stress	Subjective stress level	Categorical: very much, much, feeling a little, rarely
indigence	Recipient of basic livelihood	Categorical: current recipient, past recipient, no
dep	Experience of depression	Categorical: yes, no

2.3. Development of Deep Neural Network Model

In this study, we used a Keras sequential model wrapped in a SciKeras API to create a DNN model. In the Python programming language, Keras is a high-level neural network API [18]. SciKeras is an easy-to-use API for working with Keras deep neural networks similar to Scikit-Learn API in machine learning (ML). This was accomplished by wrapping Keras with a Scikit-Learn interface. Then, we could use these wrapped instances as Scikit-Learn ML model instances and invoke methods such as fit(), predict(), and score(). The functions of GridSearchCV() and predict_proba() were mainly what we needed to apply for tuning hyperparameters and the LIME explanation in the following steps.

The neurons in the human brain serve as the basis for DNN design. It is a deeper-layered artificial neural network. When an artificial neural network, also known as a feedforward neural network, has more than one hidden layer, it is referred to as being “deep” [19]. The phrase can also refer to a graded structure in which the depth designates the number of layers. The features computed in one hidden layer may be reused in subsequent hidden layers by deep neural networks. With fewer weights and units, a deep neural network (DNN) can approximate many natural functions by taking advantage of the compositional structure in a function [20]. Similar to a lookup table, a shallow neural network must piece together the function it approximates, whereas a deep neural network can benefit from its hierarchical structure. By increasing the precision with which a function can be approximated on a fixed budget of parameters, a more complex architecture can improve generalization after learning new examples [20].

The architecture of a DNN model and the activation function to be utilized must be decided before the model can be built. Therefore, a significant portion of deep learning involves hyperparameter optimization. The reason for this is the complexity and difficulty of configuring neural networks, which require setting a large number of parameters. Additionally, training time for individual models can be increased.

In this work, we tuned the hyperparameters of the DNN model using the grid search technique. Grid search is a method for optimizing model hyperparameters. This approach is available in Scikit-Learn via the GridSearchCV class. We could also use GridSearchCV for the Keras sequential model wrapped in SciKeras because the SciKeras API was similar to Scikit-Learn. One model was then built and assessed using the GridSearchCV procedure for each set of parameters. Each model was assessed using cross-validation, and we employed 5-fold cross-validation. We obtained the best model, which used three hidden layers, a ReLU activation function for hidden layers, a Sigmoid activation function for the output layer, a 0.5 dropout rate, an Adam optimizer with 0.001 of learning_rate, and a “binary crossentropy” loss function, after tuning the DNN model’s hyperparameters, as shown in Figure 1.

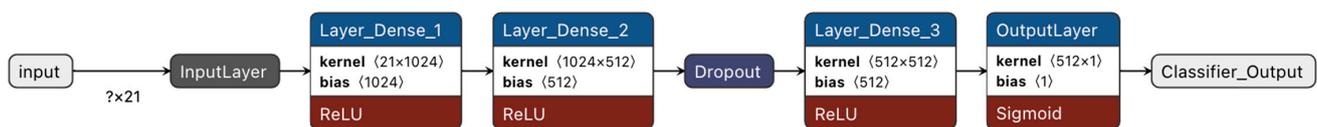


Figure 1. Flowchart of the DNN model proposed in this study.

A rectified linear unit (ReLU) activation function was employed for hidden layers, as already indicated. Any real value can be the input for a neuron with the ReLU activation function, but a neuron can only become active when the input is larger than 0. However, the Sigmoid function, which ranges from 0 to 1, makes it an excellent choice for output layers in the context of binary classification. In order to complete the necessary preparations for the model to learn, the dataset was split into two sets, one for training and one for testing, with an 80:20 split, respectively.

2.4. Model Evaluation

Evaluating the effectiveness of the proposed DNN model against competing methods for classification tasks is a crucial step after implementation. We used the loss function and the accuracy score to assess the model’s effectiveness. In addition to accuracy, the most well-known terms for a binary classification test include precision, recall, and F1-score, all of which provide statistical measurement of a classifier model’s efficacy. The calculations of the accuracy, precision, recall, and F1-scores are shown in Equations (1)–(4), respectively.

$$\text{Accuracy} = (\text{True}_{\text{positive}} + \text{True}_{\text{negative}}) / (\text{True}_{\text{positive}} + \text{True}_{\text{negative}} + \text{False}_{\text{positive}} + \text{False}_{\text{negative}}) \quad (1)$$

$$\text{Precision} = \text{True}_{\text{positive}} / (\text{True}_{\text{positive}} + \text{False}_{\text{positive}}) \quad (2)$$

$$\text{Recall} = \text{True}_{\text{positive}} / (\text{True}_{\text{positive}} + \text{False}_{\text{negative}}) \quad (3)$$

$$\text{F1 score} = 2 \times (\text{Precision} * \text{Recall}) \times (\text{Precision} + \text{Recall}) \quad (4)$$

A loss function, sometimes called a cost function, is a statistical measure that considers the degree of dissimilarity between a prediction and an actual value. As a result, we can observe a model's performance in more detail. Additionally, the test's accuracy was determined by its capacity to distinguish between depression and non-depression instances. Calculating the proportion of true-positive and true-negative results in all the analyzed cases was necessary to estimate the test's accuracy. Meantime, precision was determined by dividing the number of correctly identified positive samples by the total number of positive samples. Precision gauges how accurately a sample is classified as positive by a model. On the other hand, the recall was determined by dividing the total number of correctly categorized positive samples by the overall number of positive samples. Recall gauges a model's capacity to identify positive samples. The more positive samples identified, the larger the recall. Consequently, the F1-score equalized the weights of recall and precision by being the harmonic mean of the two. It is useful when both recall and precision are important. In addition to performance evaluation based on the aforementioned measures, the model's improvement percentage was also computed. The improvement was calculated by comparing the accuracy results from the DNN and SVM models.

2.5. Developing Model Explanations Using LIME

Explaining the results of a machine-learning (ML) system constitutes establishing a connection between the inputs to a system and the results the system produces in a way that humans can understand. This subject matter has become quite important in recent years [21]. Modern machine-learning systems (such as deep-learning-based prediction models) include highly parametric designs that make it challenging to understand the reasoning behind a model's conclusions. In this study, we used local interpretable model-agnostic explanations (LIME) to develop explanations for the output of the model, indicating the correlation between depression and other variables.

Using the LIME technique, any black-box machine-learning model can be approximated with a local, interpretable model to account for each distinct prediction. The inspiration for the concept came from a 2016 study in which the authors perturbed the initial datapoints, fed them into a black-box model, and then observed the related results. The algorithm then weighted the additional datapoints in relation to the original point. In the end, the dataset was fitted using a surrogate model of linear regression utilizing the sample weights. Each raw datapoint could then be explained using the trained explanation model. More precisely, we let $\Omega(g)$ be a measure of complexity. In the classification, $f(x)$ represented the probability (or binary indicator) that x belonged to a certain class. In order to determine the locality around x , we further used $\pi_x(z)$ as a closeness measure between an instance of z to x . Let $\mathcal{L}(f, g, \pi_x)$ be a measure of how inaccurately g approximated f in the locale denoted by π_x . We needed to keep $\Omega(g)$ low enough to be interpretable by people while minimizing $\mathcal{L}(f, g, \pi_x)$ to guarantee local fidelity and interpretability. As a result, a LIME explanation was obtained: $\xi(x) = \operatorname{argmin}_g \mathcal{L}(f, g, \pi_x) + \Omega(g)$. This formulation could be used with different explanation families G , fidelity functions L , and complexity measures Ω [22].

3. Results

3.1. Performance Evaluation of DNN Classifier Model

Fifty epochs of DNN model training were performed. Over the epochs, the classification model's accuracy steadily improved, scoring consistently greater than 89% for the last five epochs. The loss function, which showed how well the algorithm modeled the data over 50 iterations, also showed a steady decline, which means it did an excellent job of classifying.

After the results of the epochs were gathered, a model evaluation was also generated. This assessment further evaluated the classifier using precision, recall, and F1-score. The outcomes demonstrated that the DNN classifier's performance was excellent. When de-

termining whether or not older people had depression, the classifier had an accuracy rate of 89.92%. The model also achieved an F1-score of 0.92, a precision value of 93.55%, and a recall value of 97.32%. This shows that this model had a low true-positive rate when it came to diagnosing real cases of depression but a better ability to classify real cases of not being depressed.

The outcomes of the application of the DNN model were compared with several ML models in order to validate the performance of the suggested classifier. ML classifier models such as logistic regression (LR), gradient-boosting classifier (GBM), extra-gradient-boosting classifier (XGB), K-neighbors classifier (KNN), and support vector machine (SVM) with kernel radial basis function were compared with the performance of the DNN model in this study (RBF). The accuracy, precision, recall, and F1-scores of the ML models compared to the DNN model are displayed in Table 2, which shows that the DNN model improved the accuracy of classifying depression in the elderly in the range from 4% to 13%.

Table 2. Comparison with ML models.

	LR	GBM	XGB	KNN	SVM	DNN
Accuracy	76%	77%	84%	85%	77%	89.92%
Precision	74%	75%	82%	82%	75%	93.55%
Recall	69%	71%	81%	85%	71%	97.32%
F1-score	0.71	0.72	0.81	0.83	0.72	0.92

Logistic regression = LR, gradient-boosting classifier = GBM, extra-gradient-boosting classifier = XGB, K-neighbors classifier = KNN, support vector machine = SVM.

3.2. Evaluation of LIME-Based DNN Model

We chose a very complex instance to analyze in order to show how the LIME model worked with the DNN model. A description of an older person's experience with depression is shown in Figure 2. Figure 2a summarizes the patient's state and contributing circumstances. The patient was a male between the ages of 60 and 74 with depression-related symptoms, such as high stress levels, lack of support from others, low annual income, and COVID-19 life dissatisfaction. Additionally, he was concerned about the economic harm caused by COVID-19, as well as COVID-19 infection, COVID-19 death, COVID-19 criticism, and COVID-19 infection of weaker people. Our DNN model predicted that the patient would have severe depression. Figure 2b illustrates the LIME approach. The blue bars show the states and variables that significantly contributed to the prediction's support, while the orange bars show the variables that contributed to the prediction's negation. According to the explanation, at the time of the prediction, stress, income, age, COVID-19 life satisfaction, quarantine, moderate physical activity per week, concern about COVID-19 infection, concern about COVID-19 death, concern about COVID-19 infection of vulnerable people, concern about economic damage from COVID-19, changes in sleep-time activity after COVID-19, smoking, concern about criticism regarding COVID-19, and the amount of people who could be asked for help were the target's main factors and states that most contributed to the prediction.

We assessed the relative contributions of variables to the prediction of depression in the elderly after applying LIME to all the testing data in case a person had depression. With a weight of $17.7 + 0.01\%$, stress values contributed the most to model prediction, while subjective health (Sub_health_N) contributing $6.7 + 0.1\%$. Gender (sex), changes in sleep-time activity after COVID-19 (covid_sleep), and age (age_60n) were responsible for 5.2%, 4.8%, and 4.1% of the variance, respectively. As seen in Figure 3, the top five variables were responsible for almost 40% of the prediction.



Figure 2. Example of an elderly individual with depression experience.

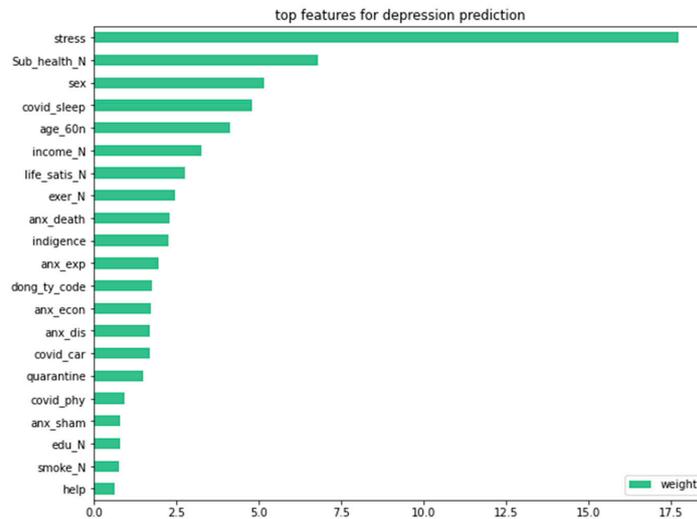


Figure 3. LIME top features for depression prediction cases.

4. Discussion

4.1. Risk of Depression

Depression is a phenomenon that manifests itself noticeably in instances of disaster or calamity [23]. Because infectious diseases, such as COVID-19, have higher morbidity and mortality rates than natural catastrophes or acts of terrorism, they have greater negative effects on the overall populace. Researchers from all over the world are gravely concerned about the rise in mental health problems in both vulnerable communities and the general population. Wang et al. (2020) [24] evaluated psychological impact, stress, anxiety, and depression in a study about the general population’s mental health during the COVID-19 outbreak in China. Regarding anxiety and depression, 8.1% of respondents to the initial poll reported moderate-to-severe stress. In a different study from China, Liu et al. (2020) [25]

examined 253 people from one of the areas most impacted by the COVID-19 pandemic and found that 7% of them had post-traumatic stress symptoms one month after the pandemic's breakout. According to Özdin et al. (2020) [26], who examined the elements that contributed to depression in the Turkish population during the COVID-19 pandemic, 45.1% of the population and 23.6% of the population, respectively, scored over the depression cut-off mark and with depression, respectively. In line with this study's concerns, Hyland et al. (2020) [27] found that the general Irish population's 65-year-old and older residents experienced the highest levels of anxiety specifically related to the COVID-19 pandemic.

Previous studies have shown that the impact of COVID-19 has significantly impacted the mental health of people in different countries. As pandemics of new infectious diseases bring intensive infectious disease management policies, such as isolation, quarantining, social distancing, and cohort blockading, at the national level, the levels of social and economic anxiety in the local community may be greater than that of natural disasters. In particular, the elderly are more likely than younger individuals to experience major mental health problems [23]. These affect physical health, as well as the quality of life, of these individuals. Thus, it is essential to preserve individuals' mental health and to develop psychological interventions that can improve the mental health of vulnerable groups during the COVID-19 pandemic. However, not much extensive epidemiological research is available on elderly community residents' anxiety and mental health during the COVID-19 epidemic. Longitudinal studies are necessary to determine the traits of depression focusing on the high-risk COVID-19 depression group reported in this study.

4.2. Proposed Evaluation Framework

The accuracy of the DNN model was 89.92%, higher than that of the LR, GBM, XGB, KNN, and SVM models, which varied from 4% to 13%. This result supports the idea that a DNN is a good classifier for large amounts of data. The dataset used for this research had a very imbalanced initial set of class labels. Additionally, it had a sizable amount of missing values. Although random oversampling and random undersampling were utilized to circumvent this, it was apparent that these dataset restrictions could lead to bias. The fact that 66.67% of the samples were accepted as non-depression was likely why the test performance of our best model performed better for the larger non-depression group than for the smaller depression group.

This research offered a methodology for objectively assessing LIME with healthcare-related black-box machine-learning models. The interdependence of LIME's performance with that of a machine-learning model must be considered when assessing LIME or other model-agnostic interpretability methodologies. This is significant because the end user can consider LIME and a machine-learning model to be one system. This means that, if a model performs poorly, LIME may be rated unfairly. The most obvious answer to this problem is to improve a model's performance in order to reduce its detrimental impact on LIME evaluation. We employed the GridSearchCV approach in this study to tune the DNN model's hyperparameters. In order to improve the performance of the model, we plan to test other hyperparameter-tuning techniques in the future, such as Bayesian, hyperband, and hyperopt.

According to our LIME study, stress, subjective health, sex, changes in sleep-time activity after COVID-19, and age (in the order listed) made the greatest contributions to the prediction. In actual life, depression quite likely results from these conditions [28–31]. As a result, our DNN model could be utilized as a tool to help psychologists or psychiatrists treat patients, as well as to predict a patient's status using large-scale surveys.

4.3. Limitations

LIME has many benefits. It is flexible enough to be utilized in various machine-learning models, simple to deploy, quick to compute, and good at giving concise explanations that appeal to physicians on the go. However, there are some restrictions that make up the majority of the current LIME research. First, LIME's explanations are not always stable or

consistent; this is because different samples are used, or different boundaries are drawn around which local datapoints are included for a local model. Furthermore, LIME's explanations of a simple, local linear model may not necessarily be consistent with the model's global logic for complicated datasets. Additionally, it is unclear how to choose the kernel width or the number of features for a local model, which forces a compromise between the complexity and interpretability of a local model.

Moreover, this study used raw data from the Korea Centers for Disease Control and Prevention's 2020 Community Health Survey. It is difficult to perform diagnostic tests for depression in large-scale epidemiological studies. For this reason, our study, which analyzed secondary data, used depression experience as an outcome variable. In future studies, we intend to include a medical diagnosis of depression as a variable in the outcome.

In the current study, our DNN model's hyperparameters were optimized to achieve the greatest performance. Ensemble learning, however, can combine several individual models to obtain a generalized performance [32]. As a result, we intend to improve our model in the future to increase its reliability for mental healthcare by developing an explainable deep-ensemble-learning model.

5. Conclusions

In conclusion, this study evaluated a LIME-based deep neural network model to explain depression and non-depression predictions produced by a black-box dee-learning model. The outcomes provided a reliable DNN model with good assessment scores in accuracy, precision, recall, and F1-score. The LIME explanation allowed us to observe that this DNN model made decisions comparable to those of humans based on extremely logical factors. More research into how to improve LIME, as well as characteristics that may increase its trust among physicians, are critical for the widespread adoption of black-box machine-learning-predictive technologies in healthcare.

Author Contributions: Conceptualization, H.V.N. and H.B.; methodology, H.V.N. and H.B.; software, H.V.N.; validation, H.V.N. and H.B.; formal analysis, H.B.; investigation, H.V.N. and H.B.; writing—original draft preparation, H.V.N.; writing—review and editing, H.B.; visualization, H.V.N.; supervision, H.B.; project administration, H.B.; funding acquisition, H.B. All authors have read and agreed to the published version of the manuscript.

Funding: This research was supported by the Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (NRF-2018R1D1A1B07041091 and NRF-2021S1A5A8062526), as well as the 2022 Development of Open-Lab based on 4P in the southeast zone.

Institutional Review Board Statement: The study was conducted in accordance with the Declaration of Helsinki and approved by the Institutional Review Board (or Ethics Committee) of the Korea Workers' Compensation and Welfare Service (protocol code: 0439001; date of approval: 01/31/2018).

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Ciotti, M.; Ciccozzi, M.; Terrinoni, A.; Jiang, W.C.; Wang, C.B.; Bernardini, S. The COVID-19 pandemic. *Crit. Rev. Clin. Lab. Sci.* **2020**, *57*, 365–388. [[CrossRef](#)] [[PubMed](#)]
2. Tadic, M.; Cuspidi, C. The influence of diabetes and hypertension on outcome in COVID-19 patients: Do we mix apples and oranges? *J. Clin. Hypertens.* **2021**, *23*, 235. [[CrossRef](#)] [[PubMed](#)]
3. Sardu, C.; Gambardella, J.; Morelli, M.B.; Wang, X.; Marfella, R.; Santulli, G. Hypertension, thrombosis, kidney failure, and diabetes: Is COVID-19 an endothelial disease? A comprehensive evaluation of clinical and basic evidence. *J. Clin. Med.* **2020**, *9*, 1417. [[CrossRef](#)] [[PubMed](#)]
4. Bansal, M. Cardiovascular disease and COVID-19. *Diabetes Metab. Syndr. Clin. Res. Rev.* **2020**, *14*, 247–250. [[CrossRef](#)]
5. Olloquequi, J. COVID-19 Susceptibility in chronic obstructive pulmonary disease. *Eur. J. Clin. Investig.* **2020**, *50*, e13382. [[CrossRef](#)]
6. Choi, E.Y.; Farina, M.P.; Wu, Q.; Ailshire, J. COVID-19 social distancing measures and loneliness among older adults. *J. Gerontol. Ser. B* **2022**, *77*, e167–e178. [[CrossRef](#)]

7. Byeon, H. Exploring factors for predicting anxiety disorders of the elderly living alone in South Korea using interpretable machine learning: A population-based study. *Int. J. Environ. Res. Public Health* **2021**, *18*, 7625. [[CrossRef](#)]
8. Liu, K.; Chen, Y.; Lin, R.; Han, K. Clinical features of COVID-19 in elderly patients: A comparison with young and middle-aged patients. *J. Infect.* **2020**, *80*, e14–e18. [[CrossRef](#)]
9. Lee, K.; Jeong, G.C.; Yim, J. Consideration of the psychological and mental health of the elderly during COVID-19: A theoretical review. *Int. J. Environ. Res. Public Health* **2020**, *17*, 8098. [[CrossRef](#)]
10. Aslam, N.; Khan, I.U.; Mirza, S.; AlOwayed, A.; Anis, F.M.; Aljuaid, R.M.; Baageel, R. Interpretable Machine Learning Models for Malicious Domains Detection Using Explainable Artificial Intelligence (XAI). *Sustainability* **2022**, *14*, 7375. [[CrossRef](#)]
11. Papandrianos, N.I.; Feleki, A.; Moustakidis, S.; Papageorgiou, E.I.; Apostolopoulos, I.D.; Apostolopoulos, D.J. An Explainable Classification Method of SPECT Myocardial Perfusion Images in Nuclear Cardiology Using Deep Learning and Grad-CAM. *Appl. Sci.* **2022**, *12*, 7592. [[CrossRef](#)]
12. Moradi, M.; Samwald, M. Explaining black-box models for biomedical text classification. *IEEE J. Biomed. Health Inform.* **2021**, *25*, 3112–3120. [[CrossRef](#)] [[PubMed](#)]
13. Guidotti, R.; Monreale, A.; Ruggieri, S.; Turini, F.; Giannotti, F.; Pedreschi, D. A survey of methods for explaining black box models. *ACM Comput. Surv.* **2018**, *51*, 1–42. [[CrossRef](#)]
14. Pedreschi, D.; Giannotti, F.; Guidotti, R.; Monreale, A.; Ruggieri, S.; Turini, F. Meaningful explanations of black box AI decision systems. In Proceedings of the AAAI Conference on Artificial Intelligence, Honolulu, HI, USA, 27 January–1 February 2019; Volume 33, pp. 9780–9784.
15. Magesh, P.R.; Myloth, R.D.; Tom, R.J. An explainable machine learning model for early detection of Parkinson’s disease using LIME on DaTSCAN imagery. *Comput. Biol. Med.* **2020**, *126*, 104041. [[CrossRef](#)]
16. Zafar, M.R.; Khan, N. Deterministic local interpretable model-agnostic explanations for stable explainability. *Mach. Learn. Knowl. Extr.* **2021**, *3*, 525–541. [[CrossRef](#)]
17. Kang, Y.W.; Ko, Y.S.; Kim, Y.J.; Sung, K.M.; Kim, H.J.; Choi, H.Y.; Jeong, E. Korea community health survey data profiles. *Osong Public Health Res. Perspect.* **2015**, *6*, 211–217. [[CrossRef](#)]
18. Gulli, A.; Pal, S. *Deep Learning with Keras*; Packt Publishing Ltd.: Birmingham, UK, 2017; ISBN 978-1-78712-842-2.
19. Svozil, D.; Kvasnicka, V.; Pospichal, J. Introduction to multi-layer feed-forward neural networks. *Chemom. Intell. Lab. Syst.* **1997**, *39*, 43–62. [[CrossRef](#)]
20. Kriegeskorte, N.; Golan, T. Neural network models and deep learning. *Curr. Biol.* **2019**, *29*, R231–R236. [[CrossRef](#)]
21. Roscher, R.; Bohn, B.; Duarte, M.F.; Garcke, J. Explainable Machine Learning for Scientific Insights and Discoveries. *IEEE Access* **2020**, *8*, 42200–42216. [[CrossRef](#)]
22. Ribeiro, M.T.; Singh, S.; Guestrin, C. Why should I trust you? Explaining the predictions of any classifier. In Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, San Francisco, CA, USA, 13–17 August 2016; pp. 1135–1144.
23. Fergusson, D.M.; Horwood, L.J.; Boden, J.M.; Mulder, R.T. Impact of a major disaster on the mental health of a well-studied cohort. *JAMA Psychiatry* **2014**, *71*, 1025–1031. [[CrossRef](#)]
24. Zhong, J.; Zhong, C.; Qiu, L.; Li, J.; Lai, J.; Lu, W.; Wang, S.; Zhong, J.; Zhao, J.; Zhou, Y. Risk and protective factors for anxiety during COVID-19 pandemic. *BMC Public Health* **2021**, *21*, 1063. [[CrossRef](#)] [[PubMed](#)]
25. Wang, C.; Pan, R.; Wan, X.; Tan, Y.; Xu, L.; McIntyre, R.S.; Choo, F.N.; Tran, B.; Ho, R.; Sharma, V.K.; et al. A longitudinal study on the mental health of general population during the COVID-19 epidemic in China. *Brain Behav. Immun.* **2020**, *87*, 40–48. [[CrossRef](#)] [[PubMed](#)]
26. Liu, N.; Zhang, F.; Wei, C.; Jia, Y.; Shang, Z.; Sun, L.; Wu, L.; Sun, Z.; Zhou, Y.; Wang, Y.; et al. Prevalence and predictors of PTSS during COVID-19 outbreak in China hardest-hit areas: Gender differences matter. *Psychiatry Res.* **2020**, *287*, 112921. [[CrossRef](#)] [[PubMed](#)]
27. Özdin, S.; Özdin, S.B. Levels and predictors of anxiety, depression and health anxiety during COVID-19 pandemic in Turkish society: The importance of gender. *Int. J. Soc. Psychiatry* **2020**, *66*, 504–511. [[CrossRef](#)] [[PubMed](#)]
28. Hyland, P.; Shevlin, M.; McBride, O.; Murphy, J.; Karatzias, T.; Bentall, R.P.; Martinez, A.; Vallières, F. Anxiety and depression in the Republic of Ireland during the COVID-19 pandemic. *Acta Psychiatr. Scand.* **2020**, *142*, 249–256. [[CrossRef](#)]
29. Cahoon, C.G. Depression in older adults. *Am. J. Nurs.* **2012**, *112*, 22–30. [[CrossRef](#)]
30. Wilkinson, P.; Ruane, C.; Tempest, K. Depression in older adults. *BMJ* **2018**, *363*, k4992. [[CrossRef](#)]
31. Byeon, H. Development of a nomogram for predicting depression in the elderly using Patient Health Questionnaire-9 among a nationwide sample of Korean elderly. *J. Pers. Med.* **2021**, *11*, 645. [[CrossRef](#)]
32. Ganaie, M.A.; Hu, M.; Tanveer, M.; Suganthan, P.N. Ensemble deep learning: A review. *Eng. Appl. Artif. Intell.* **2022**, *115*, 105151. [[CrossRef](#)]