Systematic Review and Meta-Analysis

Medicine

OPEN

An app for predicting patient dementia classes using convolutional neural networks (CNN) and artificial neural networks (ANN)

Comparison of prediction accuracy in Microsoft Excel

Sam Yu-Chieh Ho, MD[a,b], Tsair-Wei Chien, MBA[c], Mei-Lien Lin, MS[d], Kang-Ting Tsai, MD[b,e,f,*]

Abstract

Background: Dementia is a progressive disease that worsens over time as cognitive abilities deteriorate. Effective preventive interventions require early detection. However, there are no reports in the literature concerning apps that have been developed and designed to predict patient dementia classes (DCs). This study aimed to develop an app that could predict DC automatically and accurately for patients responding to the clinical dementia rating (CDR) instrument.

Methods: A CDR was applied to 366 outpatients in a hospital in Taiwan, with assessments on 25 and 49 items endorsed by patients and family members, respectively. The 2 models of convolutional neural networks (CNN) and artificial neural networks (ANN) were applied to examine the prediction accuracy based on 5 classes (i.e., no cognitive decline, very mild, mild, moderate, and severe) in 4 scenarios, consisting of 74 (items) in total, 25 in patients, 49 in family, and a combination strategy to select the best in the aforementioned scenarios using the forest plot. Using CDR scores in patients and their families on both axes, patients were dispersed on a radar plot. An app was developed to predict patient DC.

Results: We found that ANN had higher accuracy rates than CNN with a ratio of 3:1 in the 4 scenarios. The highest accuracy rate (=93.72%) was shown in the combination scenario of ANN. A significant difference was observed between the CNN and ANN in terms of the accuracy rate. An available ANN-based app for predicting DC in patients was successfully developed and demonstrated in this study.

Conclusion: On the basis of a combination strategy and a decision rule, a 74-item ANN model with 285 estimated parameters was developed and included. The development of an app that will assist clinicians in predicting DC in clinical settings is required in the near future.

Abbreviations: ANN = artificial neural networks, CC = correlation coefficient, CDR = clinical dementia rating, CNN = convolutional neural networks, DC = dementia class, EMR = electronic medical record, ML = machine learning, MMSE = Mini Mental State Examination.

Keywords: accuracy rate, artificial neural networks (ANN), clinical dementia rating (CDR), convolutional neural networks (CNN), dementia class (DC)

1. Introduction

Dementia is a progressive disease that worsens over time as cognitive abilities deteriorate. Effective preventive interventions require early detection.[1] Recent surveys indicate that dementia is underdiagnosed. When using electronic health records to investigate demographic characteristics or clinical associations of an illness (e.g., dementia), International Classification of Diseases 10th Revision codes alone cannot serve as a reliable gold standard.[2]

The authors have no funding to disclose.

The authors have no conflicts of interest to disclose.

The datasets generated during and/or analyzed during the current study are publicly available.

The Research Ethics Review Board of the Chi-Mei Medical Center approved and monitored this study.

Supplemental Digital Content is available for this article.

[a] Department of Emergency Medicine, Chi Mei Medical Center, Tainan, Taiwan, [b] Department of Geriatrics and Gerontology, Chi Mei Medical Center, Tainan, Taiwan, [c] Department of Medical Research, Chi Mei Medical Center, Tainan, Taiwan, [d] Department of Examination Room, Chi Mei Medical Center, Tainan, Taiwan, [e] Center for Integrative Medicine, Chi Mei Medical Center, Tainan, Taiwan, [f] Department of Nursing, Chung Hwa University of Medical Technology, Tainan, Taiwan.*

* Correspondence: Kang-Ting Tsai, Department of Geriatrics and Gerontology, Chi-Mei Medical Center, 901 Chung Hwa Road, Yung Kung Dist., Tainan 710, Taiwan (e-mail: codingpaperabc@gmail.com).

How to cite this article: Ho SY-C, Chien T-W, Lin M-L, Tsai K-T. An app for predicting patient dementia classes using convolutional neural networks (CNN) and artificial neural networks (ANN): Comparison of prediction accuracy in Microsoft Excel. Medicine 2023;102:4(e32670).

Received: 15 November 2022 / Received in final form: 26 December 2022 / Accepted: 27 December 2022

http://dx.doi.org/10.1097/MD.0000000000032670

Key points:

1. An app for patients predicting dementia classes was developed and demonstrated in this study.
2. The 2 models of convolutional neural networks and artificial neural networks were compared in the prediction accuracy based on 5 classes (i.e., no cognitive decline, very mild, mild, moderate, and severe) in 4 scenarios.
3. A decision rule to design the artificial neural networks app was built, particularly in double-checking the dementia classes based on additional family assessments or radar plots if the class is predicted beyond the 2 classes of no cognitive decline and very mild.

1.1. Dementia detection and prediction in the literature

Although dementia cannot be diagnosed by a single diagnostic test, clinicians employ a variety of tools and tests to detect dementia, whether it is caused by Alzheimer disease or by another disease.[3] One of the most common tests for assessing cognitive impairment is the mini mental state examination (MMSE).[4] The diagnostic accuracy of the MMSE has been investigated at various cut points for dementia in people >65 years of age.[5] The authors[5] noted that the MMSE is helpful in diagnosing dementia in settings with a low prevalence of dementia and suggested that the MMSE should not be used alone to confirm or exclude the presence of a disease. A second approach is to use the clock drawing test,[6] which provides a simple scoring system for rapid screening for mild cognitive impairment in patients with dementia.[7]

A variety of automatic speech-based tools have been developed to detect dementia more recently. These approaches typically employ machine learning (ML) classifiers trained with various vocal features derived from recorded data resulting from standard spoken tasks.[8,9] In addition, automatic speech-based tools can utilize lexical and conversation analysis-inspired features that are derived from transcripts of recorded data.[10] Furthermore, CogAware[11] analyzes transcripts of individuals describing the "cookie theft" image[12] to determine whether they originate from a patient with dementia or from a cognitively normal individual.

An underutilized but powerful tool for detecting dementia quickly and automatically is the patient's electronic medical record (EMR). There are increasing numbers of EMRs available that contain large quantities of heterogeneous data.[3] ML models have been trained using this information (e.g., images that demonstrate cortical atrophy, demographic and clinical information, as well as performance on cognitive tests) to detect the presence and severity of dementia.[13–15] Models of this type have also been used to assess the risk of a person converting to dementia from a stage of mild cognitive impairment.[16,17]

Another study[3] conducted and analyzed by demographic and clinical data from EMRs to determine individual patients' risk scores for dementia based on both structured and unstructured data from EMRs. A similar retrospective study was performed[18] using structured data, including medical diagnoses, primary care tests, investigations, lifestyle information, and prescription data. The study[3] analyzed structured clinical data from an elderly population cohort and found that their prediction models reached average F1-scores of 0.93 and 0.81 for both basic and severity diagnoses, respectively.

In this study, 2 research questions were raised, including no such research predicted classifying patients into >4 classes (no cognitive decline, very mild, mild, moderate, and severe[19]) to predict dementia class (DC) (i.e., rather than the binary classes of absence and presence in dementia as in traditional ML studies) and developed an app to assist clinicians in predicting DC in clinical settings.

1.2. Dementia assessments in clinical practice

The clinical dementia rating (CDR) is one of the most frequently used tools for evaluating dementia severity[20] in clinical settings. The CDR is a problem-oriented questionnaire completed by the patient and their family members to assess the extent of the patient's dementia.[21] However, it takes considerable time to complete all items on the 8 subscale domains, including memory, orientation, judgment, community affairs, hobbies at home, personal care, personality and behavioral problems, and language.[22] For hospital technicians, CDR assessments are time-consuming, tedious, and subjective.

Additionally, the CDR assessment result is required for a variety of purposes, including payment for Alzheimer disease medicines (such medications are regulated by the Taiwan government insurance institute) and employment of foreign caregivers hired by patients' families.[21] Therefore, CDR certification has become increasingly in demand in Taiwan, where the number of dementia patients has reached 124,263, or 0.54% of Taiwan's 23 million residents.[23] To reduce the burden on technicians in the administration of CDRs, an app developed to predict DC for patients is urgently needed.

1.3. ML-based app for predicting DC

To identify undetected dementia in primary care patients in the UK, a variety of ML models were compared with baseline epidemiological approaches.[18] Although logistic regression and random forest algorithms allow for the exposure of important features for achieving this prediction task in dementia and ML algorithms have been shown to be effective in predicting progression to dementia in memory clinic patients,[24] no such app has been developed in use for clinicians. Developing an app to predict DC for patients is thus necessary.

ML algorithms are typically used as black boxes without the ability to interpret individual predictions.[3] To provide clinicians with a clear understanding of the rationale for using ML, one emerging challenge is to make it easy and clear to understand the process of ML. Microsoft Excel is a familiar program to many researchers. We are motivated to compare the accuracy of 2 popular ML models (i.e., convolutional neural networks, [CNN] and artificial neural networks [ANN]) that have been demonstrated in the literature,[25–32] but with binary classes only. In general, CNN is considered to be a more powerful and accurate method of solving classification problems. When datasets are limited and image inputs are not needed, ANNs remain the most effective method.[33]

1.4. Study aims

This study aimed to develop an app that could predict the DC automatically and accurately for patients responding to the CDR instrument.

Two hypotheses were made for verification, including that the ANN has higher predictive accuracy than the CNN (because datasets are limited and image inputs are not required in this study[33]) and that an app can be developed and designed to help clinicians predict DC.

2. Methods

2.1. Data source

A CDR scale was applied to 366 outpatients diagnosed with dementia at a 1236-bed medical center in Taiwan from June to September 2013. All CDR data were collected, including questionnaires completed by both patients and family members,[34] with 25 and 49 items, respectively, upon request from the authors[22] (see data deposited in Supplemental Digital Content S1, http://links.lww.com/MD/I326). It is generally accepted that the severity of dementia can be classified into 5 degrees: healthy (CDR 0), very mild (CDR 0.5), mild (CDR 1), moderate (CDR 2), or severe (CDR 3).[35,36]

The Research Ethics Review Board of the Chi-Mei Medical Center approved and monitored this study.[22] The demographic data were collected anonymously and deposited in Supplemental Digital Content S1, http://links.lww.com/MD/I326 without identifying information about the participants.

2.2. Study results displayed in tables and visualizations

2.2.1. Part I: Descriptive statistics of the data. A contingency table was used to report the relationship between CDR scores and frequency in a sample of 366 patients. A chord diagram[37] was used to display the correlation coefficient (CC) between the summation of CDR scores and each item in the patient and family assessments. Those subjects whose scores were endorsed by patients and family members were displayed on a radar plot[38,39] based on positively standardized scores from 0 to 10. Those that are more severe are dispersed far from the origins of the radar plot. We examined the difference between proportional counts equally distributed on the radar plot and raw counts in the sample using the χ^2 test.

2.2.2. Part II: Model building.

2.2.2.1. ANN algorithm and model building. 2.2.2.1.1. Parameters in neuron stems. There are k neuron stems (such as those that are used by all patients, for example, $Y_k = a_1 \times x_1 + a_2 \times x_2 + \ldots + a_L \times x_L + bias_k$ providing k values for each patient in Equation 1; using the MS Excel function of SUMPRODUCT, it is possible to obtain the sum of multiplication for each pair of elements, a and x, in Equation 1). For each neuron stem, the sigmoid function in Equation 2 transformed the probability (yielded in Equation 1) into a value between 0 and 1.0.

$$Neuron\ Stem_k = \sum_{j=1}^{L} a_{kj} \times x_j + bias_k, \tag{1}$$

$$Sigmoid\ function_k = S_k = \frac{1}{1+\exp(-1 \times Neuron\ Stem_k)} = \frac{\exp(Neuron\ Stem_k)}{1+\exp(Neuron\ Stem_k)}, \tag{2}$$

In ANN, L represents the length of the item, and k represents the number of neuron stems. The parameter a_{kj} is combined with the observed score of item j in the kth neuron stem. There are $(k \times L + k)$ parameters to be estimated in these k neuron stems.

2.2.2.1.2. Parameters in neuron filters. There are m neuron filters based on m classes we are going to predict in the ANN model (e.g., 5 strata in dementia for patients, m = 5: no cognitive decline, very mild, mild, moderate, and severe). Each filter has k + 1 parameters (e.g., $Y_m = b_1 \times S_1 + b_2 \times S_2 + bias_m$, when k = 2 referred to Equation 3). As such, there are $(\mathrm{k} \times m + m)$ parameters to be estimated in these m neuron filters. The sigmoid function in Equation 4 transformed the probability (yielded in Equation 3) into a value between 0 and 1.0.

$$Neuron\ Filter_{mk} = \sum_{i=1}^{k} b_{mi} \times S_k + bias_m, \tag{3}$$

$$Sigmoid\ function_m = S_m = \frac{1}{1+\exp(-1 \times Neuron\ Filter_m)} = \frac{\exp(Neuron\ Stem_m)}{1+\exp(Neuron\ Stem_m)}, \tag{4}$$

where S_m is the mth probability of classes provided by the ANN model.

2.2.2.1.3. Model parameters to be estimated. The number of model parameters can be computed by the formula (= $(k \times L + k)$+ (k $\times m + m$) in stems and filters; for example, 53 parameters are involved in 10 items, 3 stems, and 5 filters based on 5 classes to be predicted.

2.2.2.2. CNN algorithm and model building. 2.2.2.2.1. Fundamental concept of CNN. This type of artificial neural network accepts image-type data as inputs (e.g., a 144-pixel image has 144 scores and 16 subimages, each containing 9 pixels). For example, the patient in a dementia assessment has 30 responses that could be fully incorporated into these image-type data sequentially, that is, 144 responses with some repetition. Before conducting CNN, we should determine some elements in the manipulated scenario:

(1) What is the number of classes that should be predicted (e.g., 5 strata in dementia for patients with m = 5: no cognitive decline, very mild, mild, moderate, and severe). In a CNN model with m = 5, there are 2 types of neuron stems and filters with 5 sets each.
(2) What is the number of pixels that contains the patient's response in an image, as well as the number of subimages within an image (e.g., 144 pixels in an image, n = 144) and 16 subimages each with 9 pixels, h = 9) and q = 16 = 144 ÷ 9 = the number of subimages).
(3) What is the number of parameters (denoted by β) in a neuron filter, which is dependent on the number of feature maps (denoted by δ) projected in the pooling layer If δ is 4, β is 20 + 1 because there are 80 (=16 × 5 = q × m points that should be projected by the feature map with 4 points each (i.e., 20 = 80 ÷ 4 by adding another parameter of bias in the neuron filter).

Therefore, 3 elements are determined now, namely, m = 5, n = 144, h = 9, q = 16 (=144 ÷ 9), β = 21, and δ = 4.

2.2.2.2.2. Parameters in neuron stems. Five neuron stems (such as those that are used by all patients, for example, $Y_m = a_1 \times x_1 + a_2 \times x_2 + \ldots + a_h \times x_h + bias_m$ providing m × q values for each patient in Equation 5). For each neuron stem, the sigmoid function in Equation 6 transformed the probability (yielded in Equation 5) into a value between 0 and 1.0.

$$Neuron\ Stem_m = \sum_{i=1}^{m} \sum_{j=1}^{h} \sum_{r=1}^{q} a_{mj} \times x_{h\times(j-1)+r} + bias_m, \tag{5}$$

$$Sigmoid\ function_m = S_m = \frac{1}{1+\exp(-1 \times Neuron\ Stem_m)} = \frac{\exp(Neuron\ Stem_m)}{1+\exp(Neuron\ Stem_m)}, \tag{6}$$

In CNN, all responses in items for a patient have been fully filled into the image-type dataset (e.g., 16 responses with 9 times repeatedly). In this case, 80 (=144 ÷ 9 × 5) probabilities (denoted by S_m) are obtained in Equation 6. There are $(m \times h + m)$ parameters to be estimated in these m neuron stems.

2.2.2.2.3. Feature maps used to project those 80 probabilities. Based on the 80 probability results generated in the previous section, the feature map is used to select the maximum values with 4 points in the neuron stems (note: the average or minimum values are also suggested for use in CNN). Therefore, 20 values that are maximum in each feature map for each patient are found. The probabilities in the pooling layer can be produced using Equation 7 to generalize 5 sets of probabilities (against neuron filters in the next section) for predicting the DC.

$$Sigmoid\ function_F = S_F = \frac{1}{1+\exp(-1 \times Feature\ Map_F)} = \frac{\exp(Feature\ Map_F)}{1+\exp(Feature\ Map_F)}, \tag{7}$$

2.2.2.2.4. Parameters in neuron filters. There are m neuron filters based on m classes we are going to predict DC in the CNN model. Each filter has 21(= β mentioned in section 2.1.2.1) parameters to be estimated. Equation 8 can be expressed by the one: $Y_m = b_1 \times S_1 + b_2 \times S_2 + \ldots + b_{20} \times S_{20} + bias_m$, when m = 5 owing to 5 DCs to be chosen). As such, there are $(20 \times m + m)$ parameters to be estimated in these m neuron filters. The sigmoid function in Equation 9 transformed the probability (yielded in Equation 7) into a value between 0 and 1.0.

$$Neuron\ Filter_m = \sum_{i=1}^{k} b_{mi} \times S_F + bias_m, \tag{8}$$

A ANN

1.Buildind model:
Set STEM(L+1, k), L=item length, k= # of sterm
Stem(k) = MMULT(items,STEM), column=k
S(k) = Sigmoid(Stem(k)), column=k
Set FILTER(k+1, m), k= # of sterm, m=# of classes
Filter(m) = MMULT(Prob,FILTER), m=class number
S(m) = Sigmoid(Filter(m)), m=class number
#Class = at{Max(S(m))}, Predicting class
Set total RESIDUAL = SUMXMY2(#Class, TRUE)/2
2.Parameter estimation:
Solver in MSExcel to minimize total RESIDUAL
# of parameters =(k×L+k)+(k×m+m) in STEM & FILTER
3.Prediction:
Using model parameters to predict #Class for each patient

B CNN

1. Prepareness:
Set features, eg,m=5, n=144, h=9, q=16(=144÷9=# of sub-image), β=21, and δ=4(=feature map)
2.Buildind model:
Set STEM(h+1, m), h=size of sub-image, m=# of classes
Stem(q×m) = MMULT(h items each,STEM), loop=m, points=q×m
S(q×m) = Sigmoid(Stem(q×m)), S(q×m) within 0 and 1
Set Feature maps in a hidden pooling layer
MAX(q×m÷δ)= Max(δ points in S(q×m)), mapping each sequentially, points=q×m÷ δ
Set FILTER(β, m), β= # of parameters in FILTER, m=# of classes
Filter(m) = MMULT(MAX(q×m÷δ),FILTER), m=class number
S(m) = Sigmoid(Filter(m)), m=class number
#Class = at{Max(S(m))}, Predicting class
Set total RESIDUAL = SUMXMY2(#Class, TRUE)/2
2.Parameter estimation:
Solver in MSExcel to minimize total RESIDUAL
# of parameters =(h×m+m)+(β×m) in STEM & FILTER
3.Prediction:
Using model parameters to predict #Class for each patient

Note. MMULT(A,B): Product(A,B) + bias c
TRUE: true class
SUMXMY2: funtion in MSExcel

Figure 1. Model building and parameter estimation in ANN and CNN (Note: More information about the 2 model refers to the modules with MS Excel in Supplemental Digital Contents S3 and S4). ANN = artificial neural networks, CNN = convolutional neural networks.

$$Sigmoid\ function_m = S_m = \frac{1}{1 + \exp\left(-1 \times Neuron\ Filter_m\right)} = \frac{\exp(Neuron\ Filter_m)}{1 + \exp(Neuron\ Filter_m)}, \quad (9)$$

where S_m is the mth probability of classes provided by the CNN model.

2.2.2.3. Parameter estimation and class prediction. The number of model parameters can be computed by the formula (=[m × h + m] [q × m ÷ δ] × m + m) in stems and filters; for example, 155 parameters are involved in 5 stems and 5 filters based on 5 classes to be predicted.

We set all parameters (i.e., a, b and bias in stems and filters) with randomized values in a normal distribution before performing parameter estimation. The model residual is set using the function *SUMXMY2* in MS Excel via Equation 10.

$$Residual = \frac{SUMXMY2\left(\{true\ classes\}, \{S_m\ from\ 1\ to\ m\}\right)}{2}, \quad (10)$$

where {true classes} is the data string in m classes (e.g., 10000, 01000, 00100, 00010, and 00001 represent the true class as no cognitive decline, very mild, mild, moderate, or severe, from 1 to 5, when $m = 5$).

Solver add-in in MSExcel was applied to estimate model parameters (see the 2 models with MS Excel modules in Supplemental Digital Contents S2–S4, http://links.lww.com/MD/I327: http://links.lww.com/MD/I328: http://links.lww.com/MD/I329 for details).

The dementia grade is determined by the selection of class at the maximum probability in S_m; see Equation 11.

$$\#Class_i = at\{\max(S_m)\}, \quad (11)$$

where S_m is the mth probability of classes provided via Equation 9.

2.2.2.4. Parameter estimation and class prediction. Model building and parameter estimation in ANN and CNN are briefly described in Figure 1. Details about them are deposited in Supplemental Digital Contents S2 to S4, http://links.lww.com/MD/I327: http://links.lww.com/MD/I328: http://links.lww.com/MD/I329.

2.2.3. Part III: Predicting DC for patients.

2.2.3.1. Comparison of model prediction in accuracy between ANN and CNN.

Three scenarios were manipulated to examine

Table 1

Distribution of sample size across CDR categories.

CDR category	Description	n	%
0	No cognitive decline	10	2.73
0.5	Very mild	24	6.56
1	Mild	97	26.5
2	Moderate	141	38.52
3	Severe	84	22.95
4	Very severe	10	2.73
5	Late stage	0	0
	Total	366	100

CDR = clinical dementia rating.

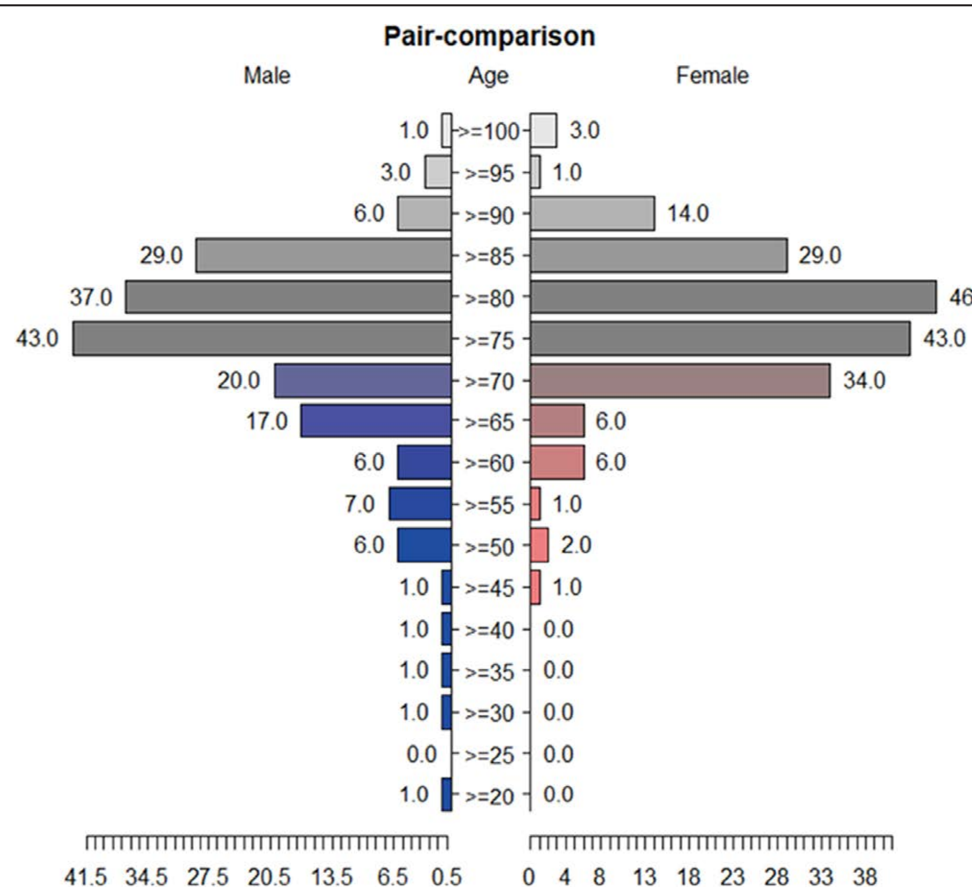

Figure 2. Distribution of frequency in ages between groups of males and females (Note: The plot was drawn in R with the code in Multimedia Supplemental Digital Content S2).

the difference in model accuracy between ANN and CNN, including all 74 items, 25 items for patients, 49 items for family members, and a combination of selected dementia grades based on the highest accuracy rate in each category. Accuracy is defined as the number of true predicted outcomes divided by the total number of patients (=366).

Comparisons of accuracy rates in those 4 scenarios were made using forest plots.[40–42] The Freeman–Tukey double arcsine transformation of proportions was applied to stabilize the variances.[43] The significance level of Type I error was set at $\alpha = 0.05$.

2.2.3.2. Decision rule for increasing the accuracy rate. In scenarios manipulated in the previous section, the combination scenario should have the highest accuracy rate, since the highest accuracy rates in each DC are selected and applied. The decision rule was designed and displayed on a Sanky diagram.[44–46]

2.2.3.3. An app designed for predicting DC. Based on the decision rule outlined in the previous section, an online application was developed for predicting the DC. A dementia grade classification (with the highest probability according to Equation 11) is displayed immediately on the website after the 74 responses have been collected from both patient and family assessments.

2.3. Creating dashboards on Google Maps

All graphs were drawn by author-made modules in Excel (Microsoft Corp). We created pages of HTML used for the network graph and forest plots on Google Maps.

The graphs on Google Maps can be zoomed in and out with a link to the website. The method of how to conduct this study is deposited with a PDF file in Supplemental Digital Content S2, http://links.lww.com/MD/I327. The 2 modules of the CNN and ANN models are deposited in Supplemental Digital Contents S3, http://links.lww.com/MD/I328 and S4, http://links.lww.com/MD/I329. A list of 74 items can be found in Multimedia Supplemental Digital Content S5, http://links.lww.com/MD/I330. The MP4 videos on CNN, ANN, and App are shown at the references.[47–49]

3. Results

3.1. Descriptive statistics of the data

This study included 366 patients, of whom 179 were male (49%) and 187 were female (51%). The majority fell into the moderate category (141, 38.52%). The male to female ratios were 34:20, 17:6, and 7:1 for those aged between 70 and 75, 65 and 70, and 55 and 60, respectively (Fig. 2). Frequency was highest in moderate categories (38.52%), followed by mild categories (26.50%), and severe categories (22.95%). There was an equal number of cases in the 2 classes of no cognitive decline and very severe (=10, 2.73%). There were no cases observed in the late stage category (Table 1).

Two chord diagrams display the CCs between the summation of CDR scores and each item in the patient and family assessment[35] (Fig. 3). Supplemental Digital Content S5, http://links.lww.com/MD/I330 contains all items with the respective symbols. Based on the 3 statistics of maximum, minimum, and median, we can see that the CCs are similar between the 2 assessments.

Those subjects whose scores have been endorsed by patients and family members are displayed on a radar plot (Fig. 4). It can be seen that patients who are more severe are dispersed far from the origins of the radar plot, such as the case of patient number 043, with a family score of 70 and a patient score of 30, at the top of the radar plot. As a result of the χ^2 test, there was no difference between proportional counts distributed on the radar plot and raw counts in the sample (χ^2 = 4.82, P = .58).

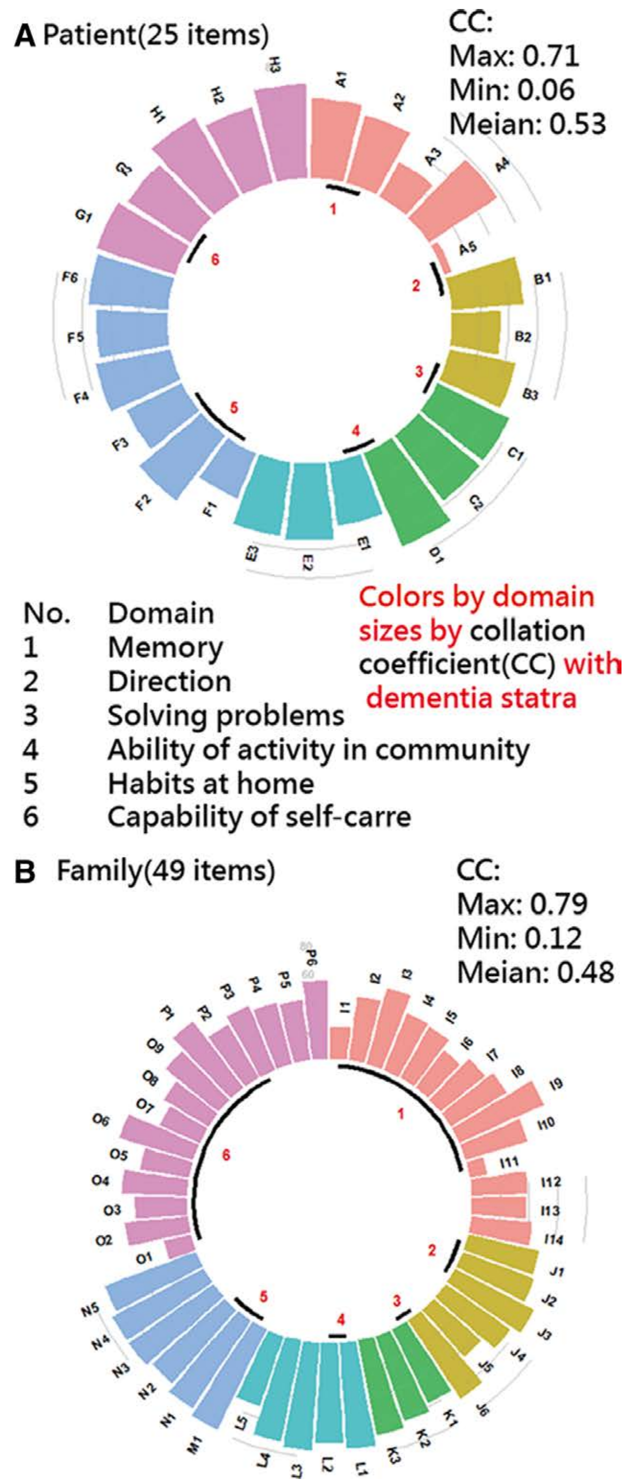

Figure 3. Comparisons of correlation coefficients between the CDR score and items in the 2 assessments (Note: The plot was drawn in R with the code in Supplemental Digital Content S2). CDR = clinical dementia rating.

3.2. Model building and prediction

As shown in Table 2, the CNN and ANN models were compared in 4 scenarios. With a ratio of 3:1, ANN had higher accuracy rates than CNN. The highest accuracy rate (=93.72%) was shown in the combination scenario of ANN, as shown at the bottom right of Table 1. A significant difference was observed between CNN and ANN in terms of accuracy rate, as shown in the forest plot (Fig. 5).

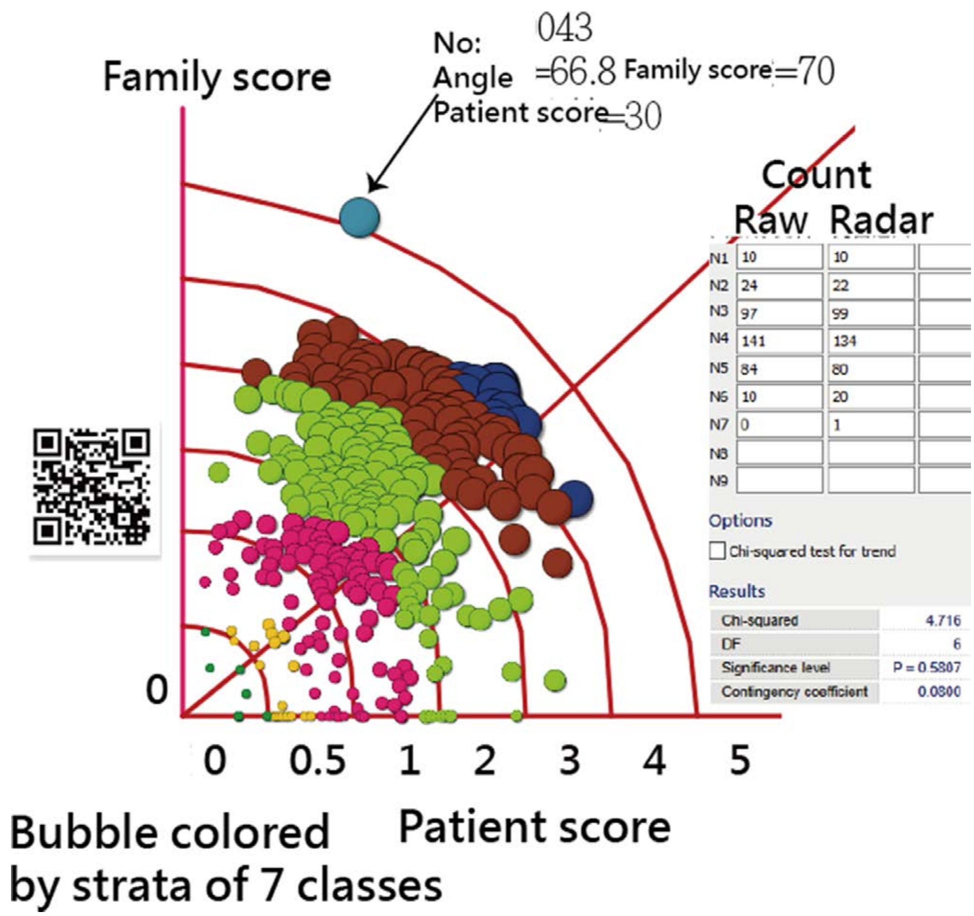

Figure 4. Patients dispersed on the radar plot based on the 2 summation scores of patient and family (Note: The plot was drawn and referred to the way in Supplemental Digital Content S2).

3.3. An app suggested for predicting DC

An algorithm was developed based on the combination of classes in ANNs to predict the DC, which is shown at the bottom of Table 1. Decision rules are based on the use of ANN and radar plots to determine DCs (Fig. 6). Double-checking of the dementia prediction is performed for each individual patient to ensure that additional family assessments or radar plots are required if the class is not predicted in the top 2 categories.

Figure 7 illustrates an online application of dementia classification. Once the QR code is scanned and 74 responses are pasted into the input box, a bar chart with the probability of each DC is displayed on the smartphone or website. Based on the ANN model, the class with the highest probability is predicted and suggested.

3.4. Online dashboards shown on Google Maps

All the QR codes in Figures[50–52] are linked to the dashboards. Readers are suggested to examine the displayed dashboards on Google Maps.

4. Discussion

4.1. Principal findings

In this study, we found that ANN had higher accuracy rates than CNN with a ratio of 3:1 in the 4 scenarios. The highest accuracy rate (=93.72%) was shown in the combination scenario of ANN. A significant difference was observed between the CNN and ANN in terms of the accuracy rate. An available ANN-based app for predicting DC in patients was successfully developed and demonstrated in this study.

Accordingly, the 2 hypotheses that[1] the ANN has higher predictive accuracy than the CNN and an app that can be developed and designed to help clinicians predict DC have been confirmed.

4.2. Additional information

As with all web-based technologies, advances in mobile communication technology are rapidly increasing. A multidimensional computerized adaptive test for the CDR scale has been demonstrated in the literature.[22] Even though multidimensional computerized adaptive test could reduce the burden on technicians who administer CDRs, we cannot guarantee that the prediction accuracy will be higher than that of ML. Nonetheless, the combination of ML and web-based Computerized Adaptive Testing[31] can be applied to the field of dementia in the future.

The 2 popular ML models (i.e., CNN and ANN) have also been demonstrated separately in MS Excel.[25–32] Many researchers are familiar with the use of Microsoft Excel. Until now, no research has been conducted comparing model accuracy between the 2 ML models under the MS Excel environment, particularly when the modules and Equations are detailed in Supplemental Digital Contents, as we did in this research.

To predict DC, multiple classes of predictive models are derived from patient/family questionnaires. Based on ML, there are substantial differences between traditional binary classes of prediction. We use 2 kinds of features in our models that vary with each scenario based on the assessment of the patient and the family. A decision rule was then formed with the ANN

Table 2

Comparisons of accuracy rates in scenarios between CNN and ANN models.

	TRUE in CNN						TRUE in ANN				
A. Patient & family (Item length = 74)											
Predict	1	2	3	4	5	Predict	1	2	3	4	5
1	20	3	2			1	32	1			
2	12	76	13			2	1	94	38		
3	1	20	110	8		3		4	85	1	
4			13	78	9	4			15	85	10
5					1	5					
77.87%	60.61	76.77	79.71	90.7	10	**80.65%**	96.97	94.95	61.59	98.84	0
66.86	(Residual)					72.26	(Residual)				
B. Patient (Item length = 25)								2	3	4	5
Predict	1	2	3	4	5	Predict	1	2	3	4	5
1	26	3	2			1	27	1			
2	6	71	13			2	4	81	10		
3	1	25	112	20		3	1	17	120	21	
4			11	66	10	4			6	65	10
5						5	1		2		
75.14%	78.79	71.72	81.16	76.74	0	**79.84%**	81.82	81.82	86.96	75.58	0
67.99	(Residual)					60.58	(Residual)				
C. Family (Item length = 49)											
Predict	1	2	3	4	5	Predict	1	2	3	4	5
1	19	5	1			1	1				
2	13	86	6			2	31	84	4		
3	1	8	124	16		3	1	15	122	10	
4			7	70	10	4			12	76	10
5						5					
81.69%	57.58	86.87	89.86	81.4	0	**77.11%**	3.03	84.85	88.41	88.37	0
55.05	(Residual)					60.61	(Residual)				
D. Combination (Item length = 74)											
Accuracy	26	86	124	78	10		32	94	122	85	10
Top%	78.79	86.87	89.86	90.7	100%	Top%	96.97	94.95	88.41	98.84	100%
88.52%				Combined		**93.72%**	122			Combined	
Strategy	Patient	Family	Family	All		Strategy	ALL	ALL	Family	ALL	
n = 366	33	99	138	86	10		33	99	138	86	10

Bold indicates larger values in horizontal rows.
ANN = artificial neural networks, CNN = convolutional neural networks.

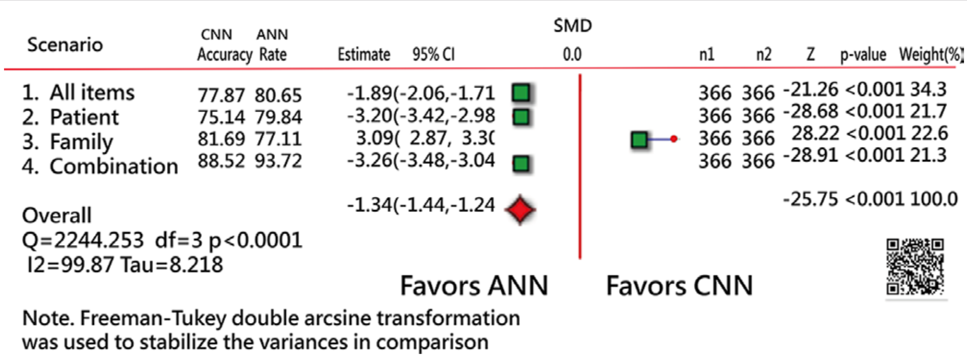

Figure 5. Comparison of accuracy rates in scenarios between CNN and ANN models (Note: The combination refers to the strategy at the bottom of Table 1. That is, a computer algorithm is designed in response to the classification based on its specific model derived from all items, patient or family items). ANN = artificial neural networks, CNN = convolutional neural networks.

algorithm incorporated with the radar plot (Fig. 4) to enhance the prediction accuracy (e.g., 93.72% in Table 2), which can be applied to future relevant studies.

There are numerous underlying causes (e.g., Alzheimer disease, which is the most prevalent at least in later life) in addition to the biology and pathophysiology of dementia, all of which are influenced by different factors (e.g., comorbidity, lifestyle, and genetics).[3] To differentiate dementia from patients' conditions, we require individualized and precise treatment for patients in the future. According to the approach described in this paper, numerous factors result in unpredictability that impacts the accuracy and efficiency of prediction. By using the app, this complexity could be reduced, resulting in improved prediction performance.

4.3. Implications and changes

There are 3 features that could make sense in prediction DC in future relevant studies, including an app for patients predicting DC should be developed, particularly with multiple classes to predict; the 2 models of CNN and ANN compared in the prediction accuracy based on 5 classes (i.e., no cognitive decline, very mild, mild, moderate, and severe) in 4 scenarios that can be applied to other algorithm comparisons in the future; a decision rule to

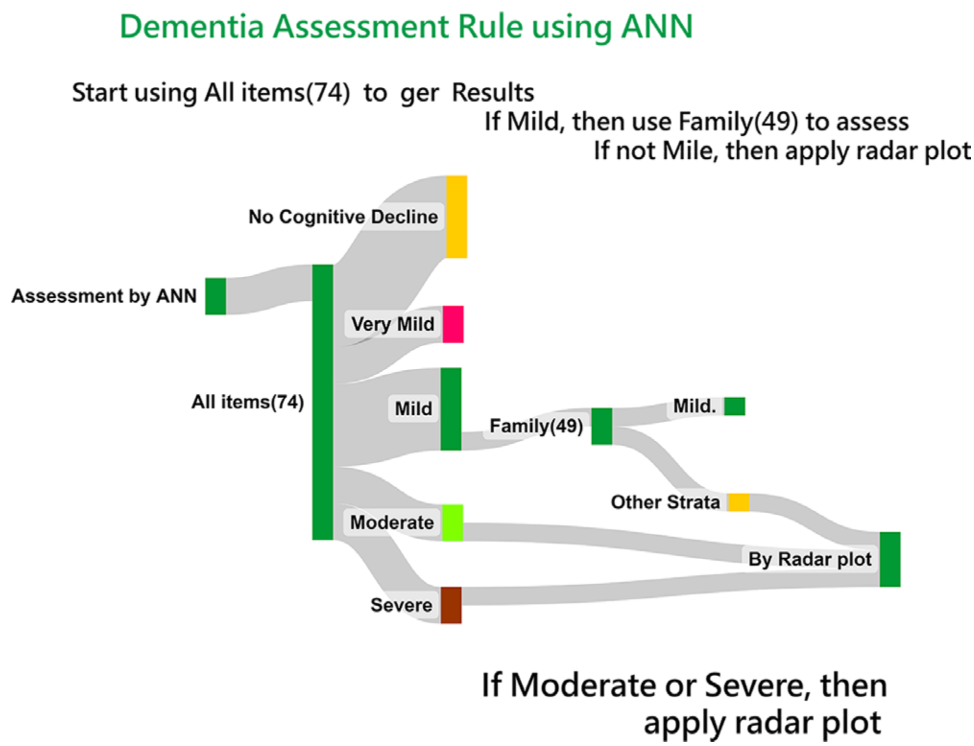

Figure 6. An algorithm design decision rule based on the use of ANN and the radar plot to determine dementia classifications (Note: To double check the dementia classification for an individual patient, additional family assessments or radar plots must be conducted if the class is not predicted in the top 2 classes). ANN = artificial neural networks.

design the app can be drawn on the Sankey diagram, particularly in double-checking the DC based on additional information (e.g., family assessments or radar plots if the class is predicted beyond the 2 classes of no cognitive decline and very mild).

In addition, many researchers are familiar with Microsoft Excel. To provide readers with a better understanding of the accuracy of the 2 popular ML models, a comparison of their accuracy was displayed on the forest plot, and modules and equations are included in the Supplemental Digital Contents and Figure 1.

4.4. Limitations and suggestions

There are a number of issues that need to be addressed in detail in further research. As a first concern, only the 2 ML models were compared. More algorithm models could be compared regarding dementia prediction in the future.

The second point is that although there are a few studies examining the use of multiple classes in prediction in the literature (for instance, a number of studies comparing prediction accuracy in binary classes using receiver operating characteristic curves), MS Excel's Solver add-in function does not allow for the estimation of more parameters(e.g., >=300 in an attempt).

Third, there were only a few neuron stems assigned to the ANN model, only 4 points were designed in feature maps, and 9 pixels were used in the CNN model. To improve prediction accuracy, future studies should enlarge these types of settings.

Fourth, the ANN model was proposed in this study based on our findings in Table 2. A greater sample size is suggested for future studies, which will be compared to the results of using ANNs in the future. Furthermore, it is possible to examine the prediction accuracy by separating the 2 sets of learning and testing samples in the future.

Fifth, according to the radar plot in Figure 4 with an equal interval radius in ascending, patients who are more severe are dispersed far from the origins of the radar plot. There was no difference between proportional counts distributed on the radar plot and raw counts in the sample (χ^2 = 4.82, P = .58). The results of this study are worthy of further study in the future, as well as examination to determine whether they are consistent with other dementia-related studies.

Sixth, Freeman–Tukey double arcsine transformation of proportions[43] was used to visualize the pair-comparison in accuracy rates on the forest plot. There is a need to discuss whether the method can be used in other situations with a single accuracy rate and a variety of sample sizes for comparison in the future.

Finally, although the combination strategy is considered valuable and applicable with a substantially higher accuracy rate than other scenarios, some programming techniques are necessary due to the complexity of the situation with many loops that require double checking.

5. Conclusion

On the basis of a combination strategy and a decision rule, a 74-item ANN model with 285 parameters estimated was developed and contained. A detailed interpretation of the 2 ML models in MS Excel is provided in Supplemental Digital Contents to assist readers in understanding ML in action. The development of an app that will assist clinicians in predicting DC in clinical settings is required in the near future.

Acknowledgments

We are grateful for the data provided by authors[22] to help us complete this study.

Author contributions

Conceptualization: Sam Yu-Chieh Ho.
Investigation: Mei-Lien Lin, Kang-Ting Tsai.
Methodology: Tsair-Wei Chien.

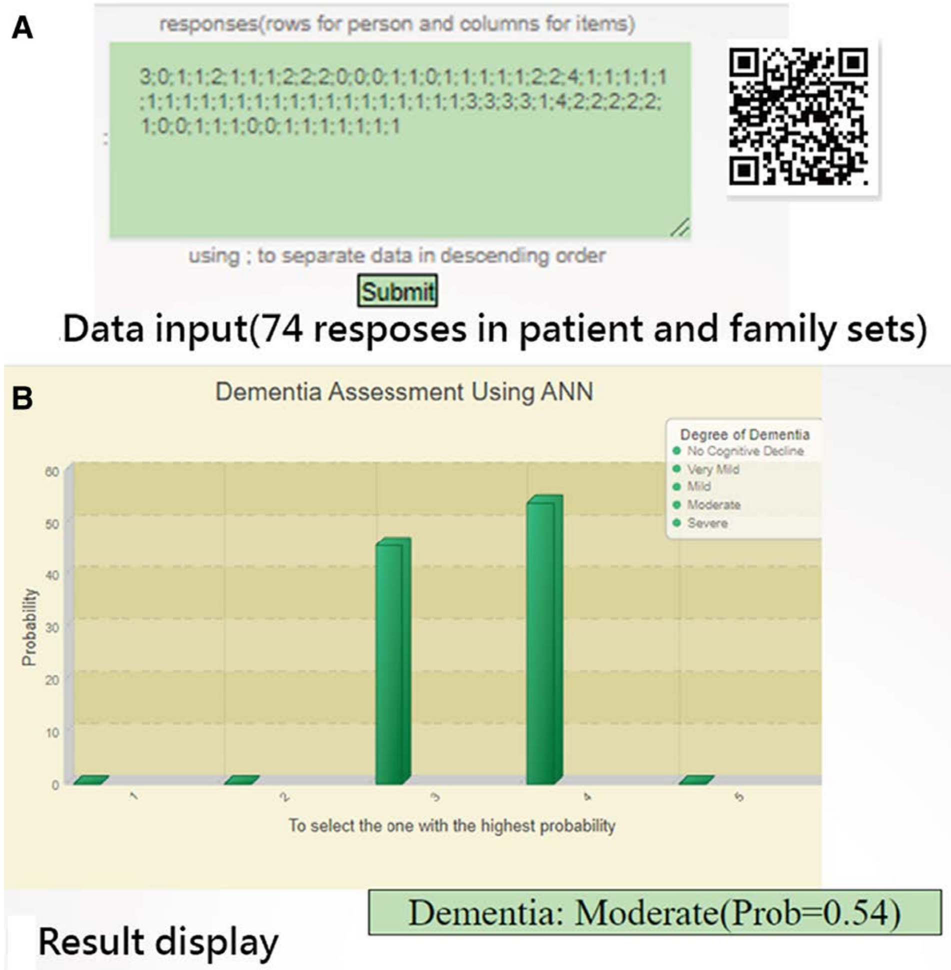

Figure 7. Snapshot of ANN assessment for dementia developed in this study (Note: With the QR-code to practice it online to know the details about the app). ANN = artificial neural networks.

References

[1] Lynch CA, Walsh C, Blanco A, et al. The clinical dementia rating sum of box score in mild dementia. Dement Geriatr Cogn Disord. 2006;21:40–3.

[2] Shao Y, Zeng QT, Chen KK, et al. Detection of probable dementia cases in undiagnosed patients using structured and unstructured electronic health records. BMC Med Inform Decis Mak. 2019;19:128.

[3] Vyas A, Aisopos F, Vidal ME, et al. Identifying the presence and severity of dementia by applying interpretable machine learning techniques on structured clinical records. BMC Med Inform Decis Mak. 2022;22:271.

[4] Arevalo-Rodriguez I, Smailagic N, I Figuls MR, et al. Mini-mental state examination (MMSE) for the detection of Alzheimer's disease and other dementias in people with mild cognitive impairment (MCI). Cochrane Database Syst Rev. 2015;3:CD010783.

[5] Creavin ST, Wisniewski S, Noel-Storr AH, et al. Mini-mental state examination (MMSE) for the detection of dementia in clinically unevaluated people aged 65 and over in community and primary care populations. Cochrane Database Syst Rev. 2016;1:CD011145.

[6] Borson S, Brush M, Gil E, et al. The clock drawing test: utility for dementia detection in multiethnic elders. J Gerontol Ser A Biomed Sci Med Sci. 1999;54:534–40.

[7] Rakusa M, Jensterle J, Mlakar J. Clock drawing test: a simple scoring system for the accurate screening of cognitive impairment in patients with mild cognitive impairment and dementia. Dement Geriatr Cogn Disord. 2018;45:326–34.

[8] Satt A, Sorin A, Toledo-Ronen O, et al. Evaluation of speech-based protocol for detection of early-stage dementia. Proc. Interspeech. 2013:1692–1696.

[9] Satt A, Hoory R, König A, et al. Speech-based automatic and robust detection of very early dementia. Proc. Interspeech. 2014:2538–42.

[10] Mirheidari B, Blackburn D, Walker T, et al. Dementia detection using automatic analysis of conversations. Comput Speech Lang. 2019;53:65–79.

[11] Rentoumi V, Paliouras G, Danasi E, et al. Automatic detection of linguistic indicators as a means of early detection of Alzheimer's disease and of related dementias: a computational linguistics analysis. In: 2017 8th IEEE international conference on cognitive infocommunications (CogInfoCom), 2017;000033–000038:IEEE International Conference on Cognitive Infocommunications (CogInfoCom).

[12] Goodglass H, Kaplan E, Barresi B. Boston Diagnostic Aphasia Examination Record Booklet. Philadelphia: Lippincott Williams & Wilkins. 2001.

[13] Möller C, Pijnenburg YA, van der Flier WM, et al. Alzheimer disease and behavioral variant frontotemporal dementia: automatic classification based on cortical atrophy for single-subject diagnosis. Radiology. 2016;279:838–48.

[14] Shankle WR, Mani S, Dick MB, et al. Simple models for estimating dementia severity using machine learning. MedInfo. 1998;52:472–6.

[15] Vyas A, Aisopos F, Vidal M-E, et al. Calibrating mini-mental state examination scores to predict misdiagnosed dementia patients. Appl Sci. 2021;11:8055.

[16] Grueso S, Viejo-Sobera R. Machine learning methods for predicting progression from mild cognitive impairment to Alzheimer's disease dementia: a systematic review. Alzheimer's Res Ther. 2021;13:1–29.

[17] Katsimpras G, Aisopos F, Garrard P, et al. Improving early prediction of dementia progression using machine learning methods. ACM Trans Comput Healthc (HEALTH). 2022;3:1–16.

[18] Ford E, Rooney P, Oliver S, et al. Identifying undetected dementia in UK primary care patients: a retrospective case–control study comparing machine-learning and standard epidemiological approaches. BMC Med Inform Decis Mak. 2019;19:248.

[19] Aplaceformom. The 7 stages of dementia and symptoms. Available at: https://www.aplaceformom.com/caregiver-resources/articles/dementia-stages. [access date November 3,2022].
[20] Hughes CP, Berg L, Danziger WL, et al. A new clinical scale for the staging of dementia. Br J Psychiatry. 1982;140:566–72.
[21] Lin KN, Liu HC. Clinical dementia rating (CDR). Acta Neurol Taiwan. 2003;12:154–65.
[22] Lee YL, Lin KC, Chien TW. Application of a multidimensional computerized adaptive test for a Clinical Dementia Rating Scale through computer-aided techniques. Ann Gen Psychiatry. 2019;18:5.
[23] Kao YH, Hsu CC, Yang YH. A nationwide survey of dementia prevalence in long-term care facilities in Taiwan. J Clin Med. 2022;11:1554.
[24] James C, Ranson JM, Everson R, et al. Performance of machine learning algorithms for predicting progression to dementia in memory clinic patients. JAMA Netw Open. 2021;4:e2136553e2136553.
[25] Yan YH, Chien TW, Yeh YT, et al. An app for classifying personal mental illness at workplace using fit statistics and convolutional neural networks: survey-based quantitative study. JMIR Mhealth Uhealth. 2020;8:e17857.
[26] Chou PH, Chien TW, Yang TY, et al. Predicting active NBA players most likely to be inducted into the basketball hall of famers using artificial neural networks in microsoft excel: development and usability study. Int J Environ Res Public Health. 2021;18:4256.
[27] Tey SF, Liu CF, Chien TW, et al. Predicting the 14-day hospital readmission of patients with pneumonia using artificial neural networks (ANN). Int J Environ Res Public Health. 2021;18:51105110.
[28] Chen HC, Chien TW, Chen L, et al. An app for predicting nurse intention to quit the job using artificial neural networks (ANNs) in Microsoft Excel. Medicine (Baltim). 2022;101:e28915.
[29] Lin CY, Chien TW, Chen YH, et al. An app to classify a 5-year survival in patients with breast cancer using the convolutional neural networks (CNN) in Microsoft Excel: development and usability study. Medicine (Baltim). 2022;101:e28697.
[30] Lee YL, Chou W, Chien TW, et al. An app developed for detecting nurse burnouts using the convolutional neural networks in Microsoft Excel: population-based questionnaire study. JMIR Med Inform. 2020;8:e16528.
[31] Ma SC, Chou W, Chien TW, et al. An app for detecting bullying of nurses using convolutional neural networks and web-based computerized adaptive testing: development and usability study. JMIR Mhealth Uhealth. 2020;8:e16747.
[32] Yang TY, Chien TW, Lai FJ. Web-based skin cancer assessment and classification using machine learning and mobile computerized adaptive testing in a Rasch model: development study. JMIR Med Inform. 2022;10:e33006.
[33] Meel V. ANN and CNN: analyzing differences and similarities. Available at: https://viso.ai/deep-learning/ann-and-cnn-analyzing-differences-and-similarities/. [access date November 5, 2022].
[34] Lin ML, Chen YC, Lin GJ, et al. An analysis of the feasibility of using a computer-aided technique to score the clinical dementia rating scale. J Healthc Manage. 2014;15:306–26.
[35] Cedarbaum JM, Jaros M, Hernandez C, et al. Rationale for use of the clinical dementia rating sum of boxes as a primary outcome measure for Alzheimer's disease clinical trials. Alzheimer's Dementia. 2013;9(1 Suppl):S45–55.
[36] Williams MM, Storandt M, Roe CM, et al. Progression of Alzheimer's disease as measured by clinical dementia rating sum of boxes scores. Alzheimer's Dementia. 2013;9(1 Suppl):S39–44.
[37] Finnegan A, Sao SS, Huchko MJ. Using a chord diagram to visualize dynamics in contraceptive use: bringing data into practice. Glob Health Sci Pract. 2019;7:598–605.
[38] Castagna C, Consorti G, Turinetto M, et al. Osteopathic models integration radar plot: a proposed framework for osteopathic diagnostic clinical reasoning. J Chiropr Humanit. 2021;28:49–59.
[39] Shao Y, Chien TW, Jang FL. The use of radar plots with the Yk-index to identify which authors contributed the most to the journal of Medicine in 2020 and 2021: a bibliometric analysis. Medicine (Baltim). 2022;101:e31033e31033.
[40] Wu JW, Chien TW, Tsai YC, et al. Using the forest plot to compare citation achievements in bibliographic and meta-analysis studies since 2011 using data on PubMed Central: a retrospective study. Medicine (Baltim). 2022;101:e29213e29213.
[41] Yan YH, Chien TW. The use of forest plot to identify article similarity and differences in characteristics between journals using medical subject headings terms: A protocol for bibliometric study. Medicine (Baltim). 2021;100:e24610.
[42] Wang LY, Chien TW, Lin JK, et al. Vaccination associated with gross domestic product and fewer deaths in countries and regions: a verification study. Medicine (Baltim). 2022;101:e28619.
[43] Nyaga VN, Arbyn M, Aerts M. Metaprop: A Stata command to perform meta-analysis of binomial data. Arch Public Health. 2014;72:39.
[44] Li MJ, Chien TW, Liao KW, et al. Using the Sankey diagram to visualize article features on the topics of whole-exome sequencing (WES) and whole-genome sequencing (WGS) since 2012: bibliometric analysis. Medicine (Baltim). 2022;101:e30682e30682.
[45] Lee YL, Chien TW, Wang JC. Using Sankey diagrams to explore the trend of article citations in the field of bladder cancer: Research achievements in China higher than those in the United States. Medicine (Baltim). 2022;101:e30217e30217.
[46] Wang HY, Chien TW, Kan WC, et al. Authors who contributed most to the fields of hemodialysis and peritoneal dialysis since 2011 using the hT-index: bibliometric analysis. Medicine (Baltim). 2022;101:e30375e30375.
[47] Chien TW. CNN model in MS Excel in this study. Available at: https://youtu.be/ugSIPaIwk1E. [access date December 23, 2022].
[48] Chien TW. ANN model in MS Excel in this study. Available at: https://youtu.be/4B2oU4HnqSE. [access date December 23, 2022].
[49] Chien TW. App demonstrated for and used in this study. Available at: https://youtu.be/ygr3k4-ATA4. [access date December 23, 2022].
[50] Chien TW. Figure 4 in this study. Available at: http://www.healthup.org.tw/gps/cdr22.htm. [access date December 23, 2022].
[51] Chien TW. Figure 5 in this study. Available at: http://www.healthup.org.tw/gps/edmentialforest.htm. [access date December 23, 2022].
[52] Chien TW. Figure 7 in this study. Available at: http://www.healthup.org.tw/irs/anndementia.asp. [access date December 23, 2022].