Quest Journals
Journal of Research in Humanities and Social Science
Volume 13 ~ Issue 8 (August 2025) pp: 120-127
ISSN(Online):2321-9467
www.questjournals.org

Research Paper

Yoga practice in the United States: Are people who have used mental health counseling services more likely to practice it?

Elaine Su

Abstract
Objective: *This study aims to assess whether people who have used mental health counseling services are more likely to practice yoga among adults in the United States .*
Methods: *Data from the Adult Complementary Health Supplement (ACH) in the National Health Interview Survey (NHIS) were used. Logistic regression analysis and artificial neural network are applied to train and then test the model. Area Under the Curve (AUC) and ROC curves are estimated.*
Results: *The Area Under the Curve (AUC) is about 0.80for training sample according to above logistic regression, meaning that a randomly selected individual from the positive group has a test value larger than that for a randomly chosen individual from the negative group 80percent of the time. After logistic regression and network analysis were conducted in the training sample, we used the outputs from both models to predict the likelihood in the testing sample. The areas under the receiver operating characteristic curves were 81% and 82% for the logistic model and the neural network, respectively. There were no significant differences in predictive ability between the two approaches.*
Conclusion:
In this study, we built a predictive model for adult practice of Yoga by using Logistic regression analysis and neural networking model, and then evaluated the performance of the models. In our research, both logistic regression and neural network models did a good job of predicting the probability of adult practice of Yoga. We confirmed our hypothesis that people who have used mental health counseling services are more likely to practice yoga among adults in the United States .

Key words: *predictive model; Logistic regression analysis; artificial neural network*

*Received 08 Aug., 2025; **R**evised 19 Aug., 2025; **A**ccepted 21 Aug., 2025 © The author(s) 2025.*
P**ublished with open access at **www.questjournas.org

I. INTRODUCTION

In this study, we aimed to explore whether people who have used mental health counseling services are more likely to practice yoga, using data of a nationally representative sample from the United States.

II. RESEARCH METHODS

2.1 Data source

We used data from the Adult Complementary Health Supplement (ACH) in the National Health Interview Survey (NHIS) https://www.cdc.gov/nchs/nhis/index.htm.

NHIS is conducted by the Centers for Disease Control and Prevention (CDC) of the United States and has monitored health of the nation since 1957. It covers a broad range of health topics collected through personal household interviews. Survey data are made open to researchers, and results have been instrumental in providing data to track health status, and health care access.

At the time of this study, the most recent data was from year 2022. In this year, a sample adults aged 18 and olderfrom the whole survey were asked about their age, gender marital status, health status, practice of yoga, meditation etc. in the past 12 months. These data are called the Adult Complementary Health Supplement (ACH).

2.2. Variables of interest

2.1practice of Yoga**:**

Participants were asked the following question:

- DURING THE PAST 12 MONTHS, did you practice yogaas part of a class or on your own?

A variable "practice" was created, with a value of 1 if they answered yes, and a value of 0 if they answered no.

Personal attributes

A total of 13 variables are included as potential predictors which are largely based on literature search[1,2].

1. Mental mental_consulting: based on question "During the past 12 months, did you receive counseling or therapy from a mental health professional such as a psychiatrist, psychologist, psychiatric nurse, or clinical social worker?"
2. Age: ranging from 18 to 99 years
3. Gender
4. Race
5. Marital status: married/partnered vs. other status
6. Region of residence: 1 Northeast 2 Midwest 3 South 4 West
7. Has the person ever worked
8. Cancer: based on the question "Ever told by a doctor or health professional that you had cancer"
9. current smoking status: based on question 'Do you NOW smoke cigarettes every day, some days or not at all?'.we categorized responses into yes vs. no
10. drinking status: based on question "During the past 12 months, how many days per week, per month or per year did you drink any type of alcoholic beverage?". Responses ranged from 0 to 365
11. BMI category
12. Health status
13. Meditation practice

2.3 Statistical Analysis

Overall, we used both Logistic regression and artificial neural network (ANN) to develop/train and then to test/validate the model. Data was split into two random samples: a 75% training sample for developing the model, and a 25% testing sample for validating the model.

Logistic regression analysis

With the training data, stepwise technique in logistic regression analysis is performed to select variables. Logistic regression is a widely used statistical model for analyzing binary outcomes, and it can make the prediction of the odds and the related probability of an outcome or event from a set of predictor variables. In this study, the outcome is "if the participant practices Yoga". The predictors can be either continuous variables, categorical variables, or both. More explanation of the logistic regression model is provided below:

- The general formula of logistic regression is: $\ln(\text{odds of an event occurring}) = \ln\left(\frac{\mathrm{P}}{\mathrm{P}-1}\right) = \beta + \beta_1 * X_1 + \beta_2 * X_2 + \cdots + \beta_n * X_n$.P is the probability of an event, which is convertible with odds.
- X_n is a predictor variable, and β_n is a regression coefficient. The relationship between the odds ratio and the coefficients is $\mathrm{OR} = \mathrm{e}^{\beta}$. If the coefficient β of a variable X_n is larger than 0, X_n is related to a higher odds/probability of the event. The odds ratio related to X_n is above 1 in this case.
- If the coefficient of a variable X_n is equal to 0, X_n is not related to the event. The odds ratio related to X_n is equal to 1 in this case.
- If the coefficient of a variable X_n is smaller than 0, X_n is related to a lower odds/probability of the event. The odds ratio related to X_n is below 1 in this case.

Artificial neural network (ANN)

ANN, often just called a "neural network" (NN), is a mathematical model or computational model based on biological neural networks, in other words, is an emulation of biological neural system. It consists of an interconnected group of artificial neurons and processes information using a connectionist approach to computation. In more practical terms neural networks are non-linear statistical data modeling tools. ANN is widely used these days to model complex relationships between inputs and outputs or to find patterns in data. This model was done using R software 'neuralnet' package.

Model validation

For model validation, a ROC curve is generated which is a measure of classifier performance. [3]It is a graphic that shows the diagnostic ability of a model in predicting binary outcome as its discrimination threshold is varied. Ultimately, the focus is the area under the ROC curve, or AUROC. AUROC ranges from 0.50 to 1.00,

and the higher AUROC, the better the model is. It should be noted, however, that in this study, the purpose is to compare which model has a higher AUROC, instead of the absolute values of AUROC.

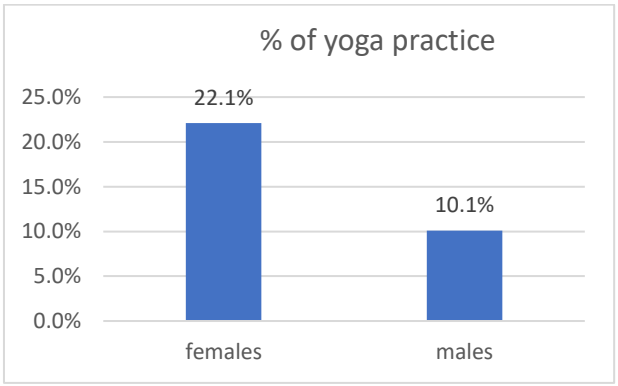
III. RESULTS

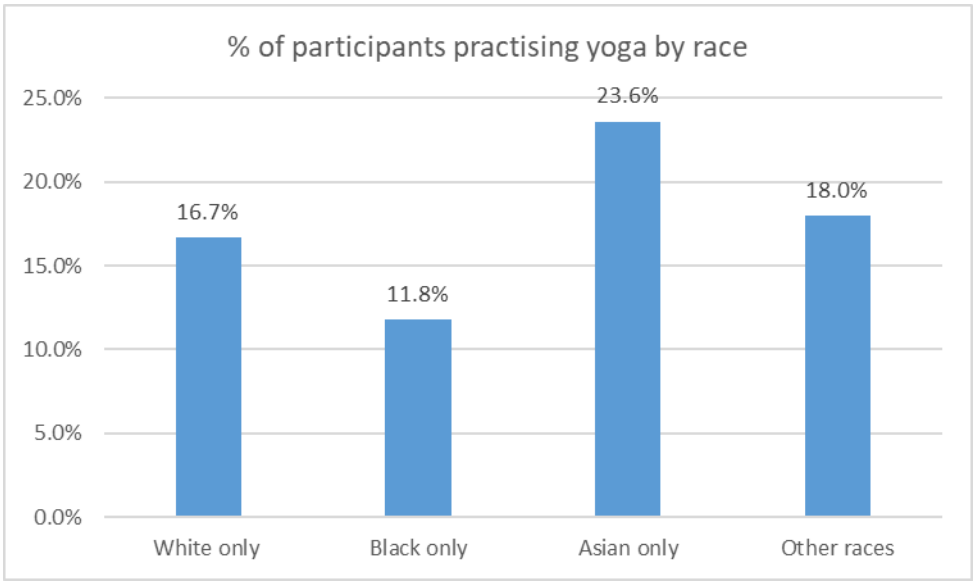
3.1 Demographic profile of the study sample

In the final sample of 27,651 participants who had complete data on the variables of interest, the average age was 53years, ranging from 18 to 99 years old. 46% are males.A majority are White Americans.

White only	78.0%
Black only	12.3%
Asian only	6.4%
Multiple race	3.3%

In the study sample, 16.6% ever practiced Yoga in the past year. The rate is higher among females than among males. There are differences in proportion of participants practicingYoga by race.



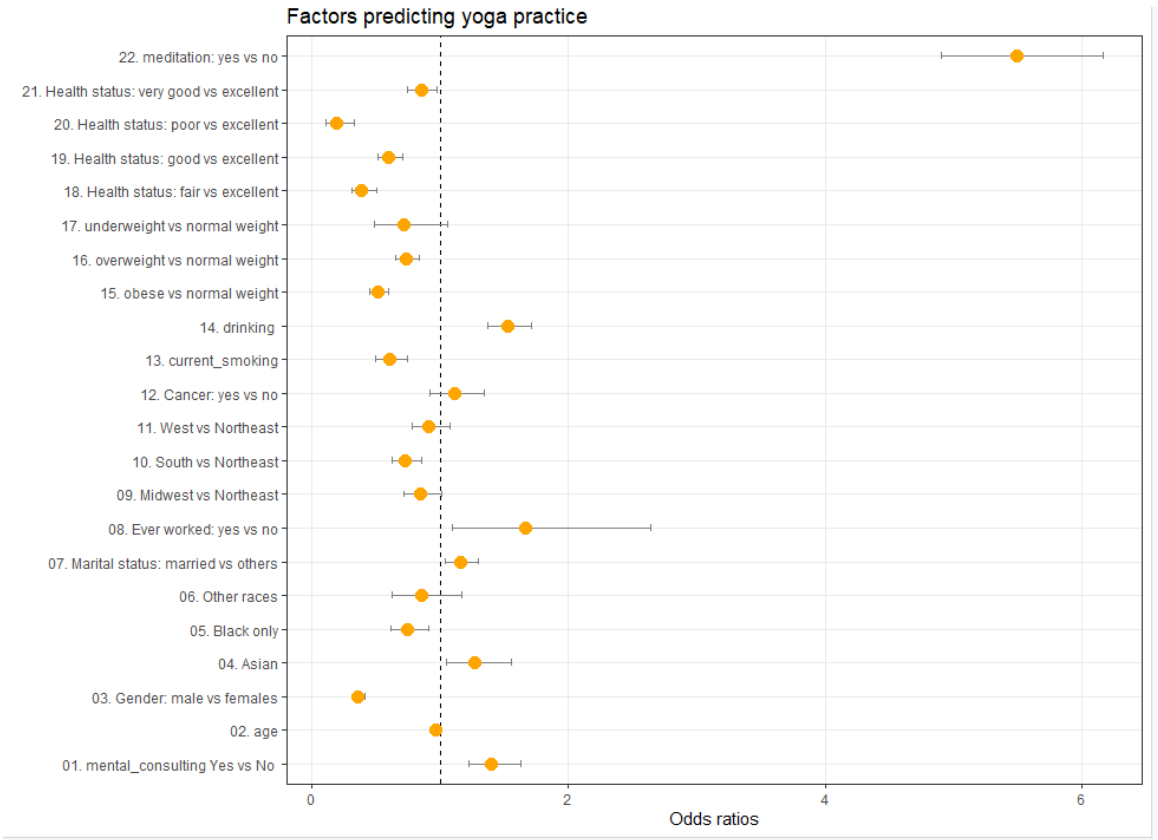


3.2 *Development of the prediction model*

From stepwise logistic regression analysis, all 13 variables are kept in the final prediction model.
The tables of coefficients and odds ratios are listed below

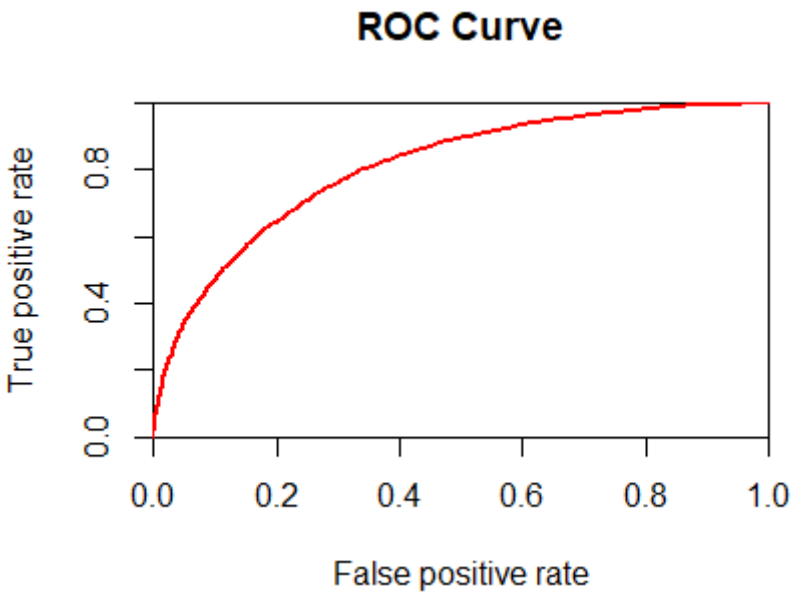
	Estimate	Std. Error	z value	P-value		Odds Ratio	Lower CI	Upper CI
(Intercept)	-0.57	0.25	-2.33	0.02	*	0.56	0.34	0.90
mental_consulting Yes vs No	0.34	0.07	4.62	<0.001	***	1.40	1.22	1.63
age	-0.02	0.00	-14.42	<0.001	***	0.97	0.97	0.98
Gender: male vs females	-1.01	0.06	-16.89	<0.001	***	0.36	0.33	0.41
Race								
Reference group: White								
Asian	0.25	0.10	2.42	0.02	*	1.27	1.05	1.56
Black only	-0.29	0.10	-2.91	0.00	**	0.75	0.62	0.91
Other races	-0.14	0.16	-0.91	0.36		0.86	0.63	1.17
Marital status: married vs others	0.15	0.06	2.66	0.01	**	1.16	1.04	1.30
Ever worked: yes vs no	0.51	0.22	2.31	0.02	*	1.67	1.10	2.64
Region								
reference: Northeast								
Midwest	-0.16	0.09	-1.85	0.06	.	0.85	0.72	1.01
South	-0.31	0.08	-3.90	<0.001	***	0.73	0.63	0.86
West	-0.08	0.08	-1.01	0.31		0.91	0.78	1.08
Cancer: yes vs no	0.11	0.10	1.12	0.26		1.11	0.92	1.34
current_smoking	-0.49	0.10	-4.77	<0.001	***	0.61	0.50	0.75
drinking	0.43	0.06	7.60	<0.001	***	1.53	1.37	1.71
BMI category								
reference: normal weight								
obese	-0.65	0.07	-8.92	<0.001	***	0.52	0.45	0.60
overweight	-0.30	0.07	-4.61	<0.001	***	0.74	0.65	0.84
underweight	-0.32	0.20	-1.62	0.11		0.72	0.49	1.06
Health status								
reference: excellent								
fair	-0.92	0.13	-7.26	<0.001	***	0.39	0.31	0.51
good	-0.50	0.08	-6.28	<0.001	***	0.60	0.52	0.71
poor	-1.62	0.28	-5.88	<0.001	***	0.19	0.11	0.33
very good	-0.15	0.07	-2.20	0.03	*	0.86	0.75	0.98
meditation: yes vs no	1.71	0.06	29.31	<0.001	***	5.50	4.91	6.17

From the Odds Ratios, …. have a lower likelihood of Yoga practice (Odds Ratio<1), while….are associated with a higher likelihood of practice (OR>1):

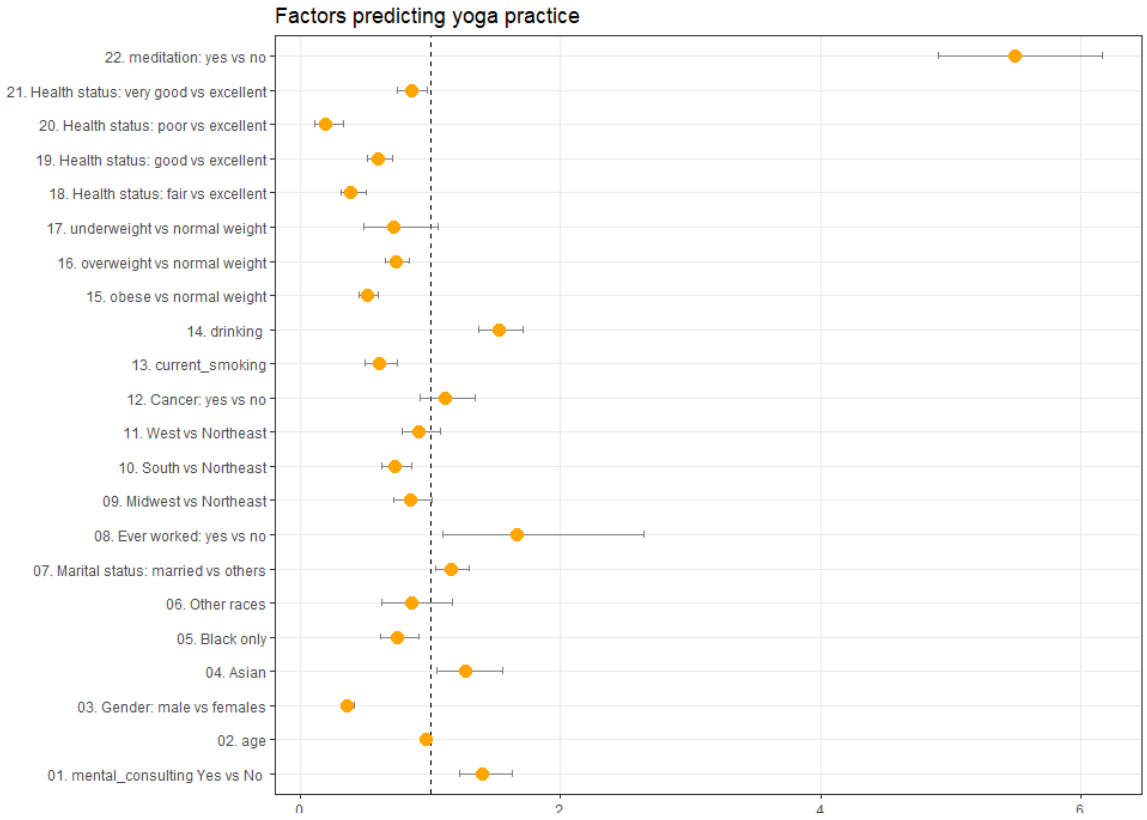


The Area Under the Curve (AUC) is 0.80for training sample according to above logistic regression. AUC is also referred to as *concordance index*, and it is a widely accepted traditional performance metric for a ROC curve. The higher the area under the curve the better prediction power the model has. *ACU=0.80*can be

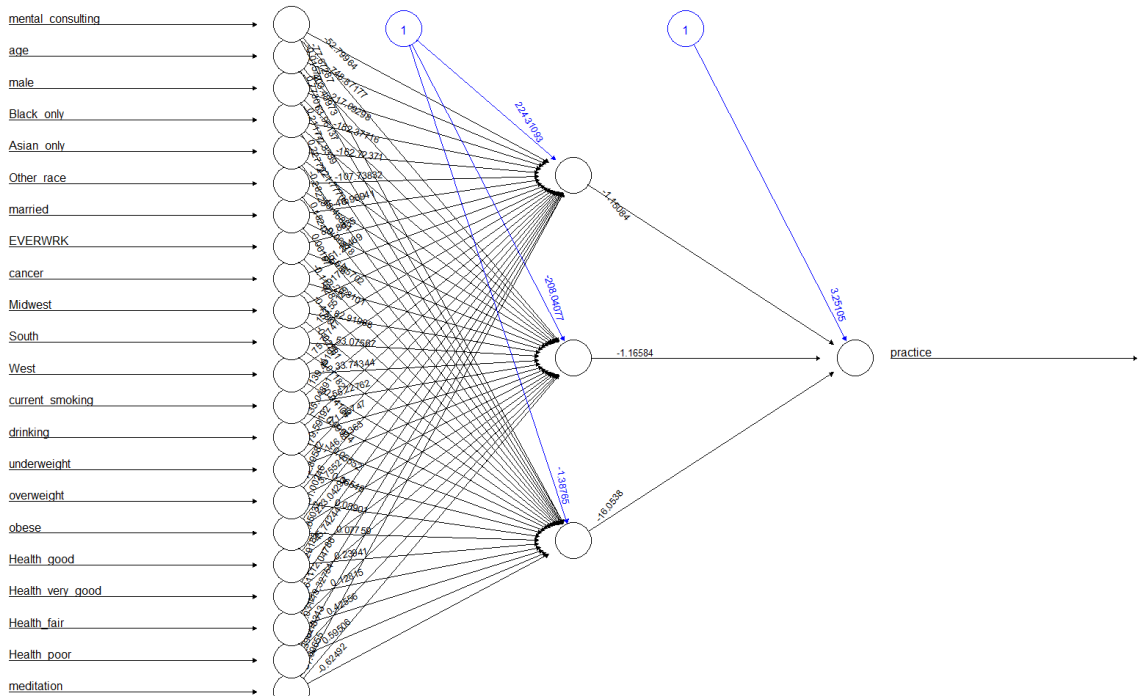
interpreted that a randomly selected individual from the positive group has a test value larger than that for a randomly chosen individual from the negative group 80 percent of the time.



Meanwhile, we calculated the Kolmogorov-Smirnov (KS) statistic of the Logistic model. KS statistic is a commonly used model evaluation metric for models predicting binary outcomes. It tests if the logistic model separates (i.e., discriminates between) events and non-events. KS ranges from 0% to 100%, and a higher value indicates a better model fit. According to an article, 0.41–0.6 represents moderate agreement (Landis and Koch, 1977.).In this case, KS is 0.47, indicating a good performance of the model.

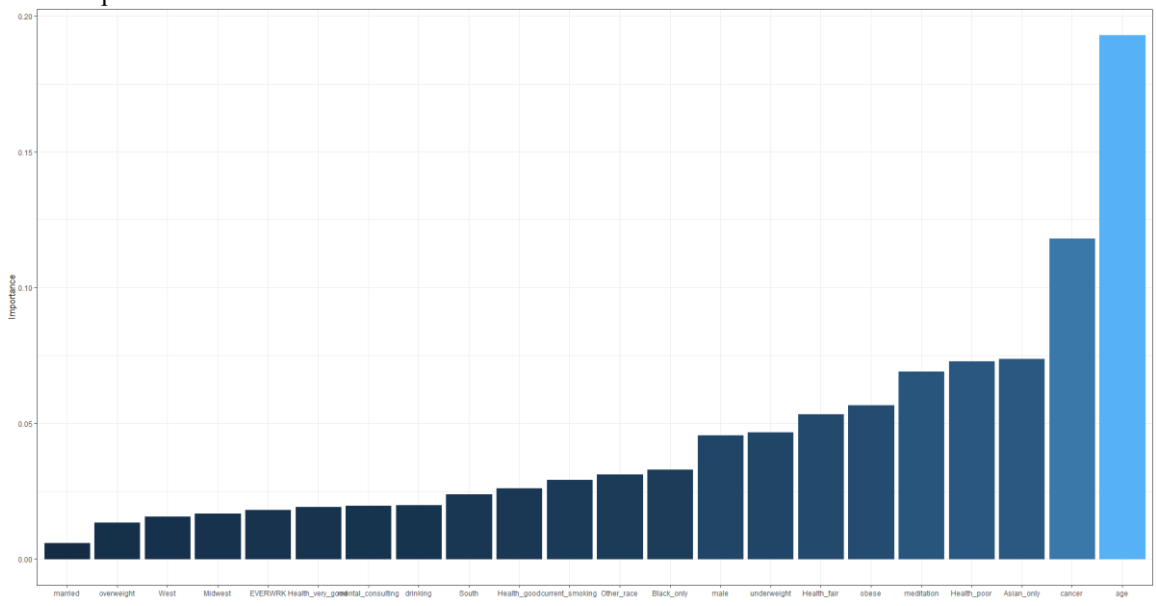


3.3 model training result from NN



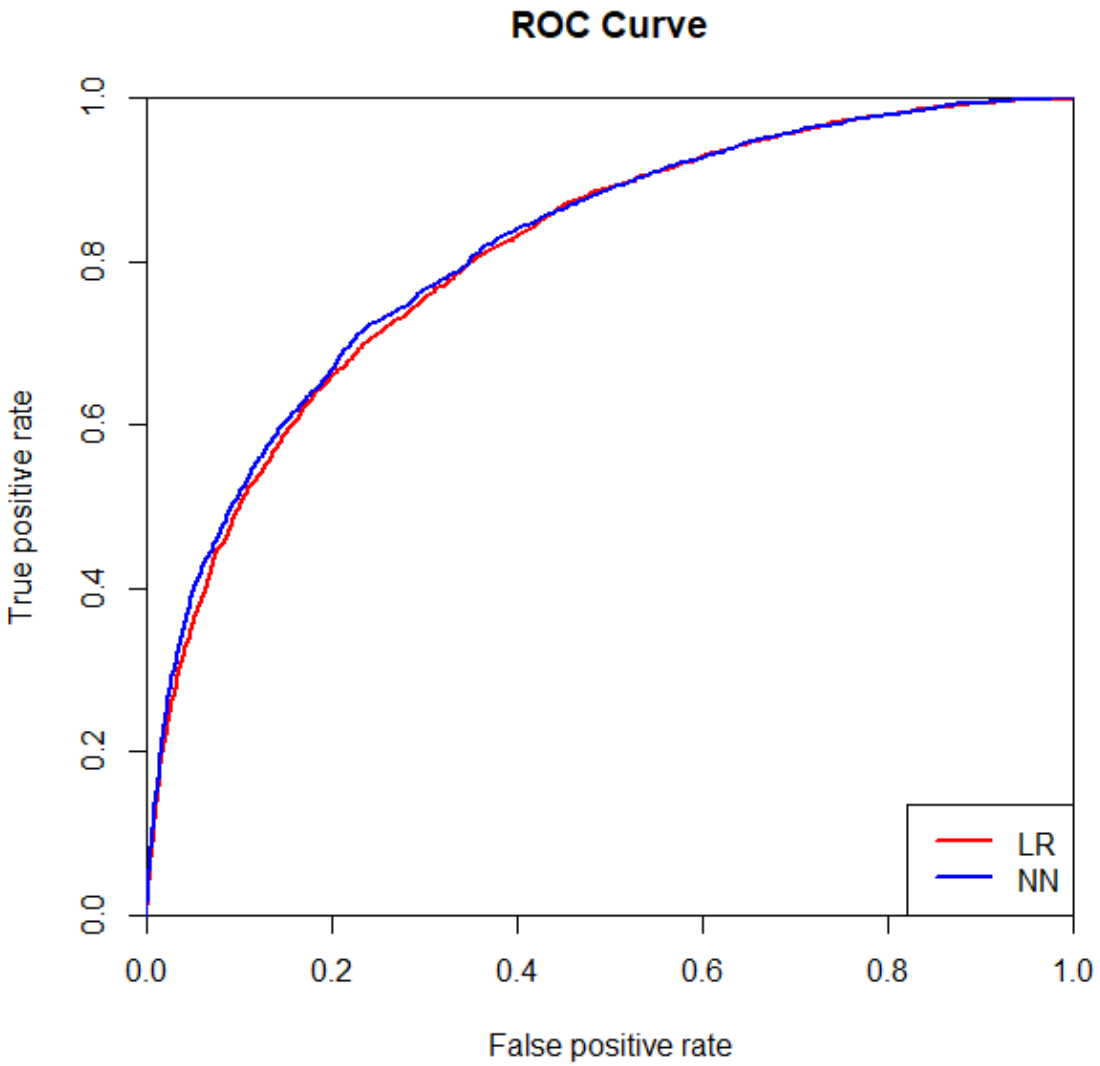
In the above plot, the net is essentially a black box so we cannot say that much about the fitting or the weights. However, it is sufficient to say that the training algorithm has converged and therefore the model is ready to be used.

Relative importance of variables from NN model



Variable importance from the NN model is listed. The relative importance of a specific explanatory variable can be determined by identifying all weighted connections between the nodes of interest. It can be interpreted as the strength of association between an explanatory variable and the response variable. The number indicates relative importance with the absolute magnitude from zero to one.[4] From the model, it can be seen that some of the most important factors associated with Yoga practice includeage, cancer, Asian American. AUC is 82% for the NN model.

3.4 Model validation:
In testing sample, the ROC was 81% for the Logistic regression and 82% for the artificial neural network. Both indicate good performance.



IV. DISCUSSION

From Logistic regression analysis, females, Asian Americans, people who practice meditation, and people who have used mental health counseling serviceshave a higher likelihood of Yoga practice. This is consistent with previous research findings [1,2]. Meanwhile, our results are similar with previous finding that "Regarding health behaviors, yoga users were less likely to smoke and more likely to drink alcohol"[5].

Both Logistic regression and NN model have good performance with a similarly high AUC. The NHIS did not include information on participants' education level, which has been found to be associated with holistic movement practices including yoga[1,2]. If this information is available, the performance of the predictive model might be further improved.

V. CONCLUSION

In this study, we built a predictive model for Yoga practice among adults in the U.S. by using Logistic regression analysis and neural networking model, and then evaluated the performance of the models. In our research, both logistic regression and neural network models did a good job of predicting the probability of Yoga practice. We confirmed our hypothesis that people who have used mental health counseling serviceshave a higher likelihood of Yoga practice.

References

[1]. Birdee, G. S., Wayne, P. M., Davis, R. B., Phillips, R. S. & Yeh, G. Y. T'ai chi and qigong for health: patterns of use in the United States. *J. Altern. Complement. Med.***15,** 969–973 (2009).

[2]. Wang, C. C., Li, K., Choudhury, A. & Gaylord, S. Trends in Yoga, Tai Chi, and Qigong Use Among US Adults, 2002-2017. *Am. J. Public Health***109,** 755–761 (2019).

[3]. *Evaluation of Predictive Models. Decision Systems Group, Brigham and Women's Hospital Harvard Medical School*.

[4]. Garson, G. D. Interpreting neural network connection weights. *Artif. Intell. Expert***6,** 46–51 (1991).

[5]. More adults and children are using yoga and meditation. *National Institute of Health* (2018). Available at: https://www.nih.gov/news-events/news-releases/more-adults-children-are-using-yoga-meditation. (Accessed: 6th February 2018)