ATLANTIS PRESS

Advances in Social Science, Education and Humanities Research (ASSEHR), volume 156

2nd International Seminar on Education Innovation and Economic Management (SEIEM 2017)

Predict Physics Achievement in Middle School Education by Big Five Model and Artificial Neural Network

Meng-Meng Yang
School of Computer Science and Technology
Nanjing Normal University
Nanjing, Jiangsu 210023, P R China
1320981274@qq.com

Soriya Aok
Faculty of Science and Technology
Angkor University
Borey Seang Nam, Siem Reap Province, Cambodia

John Liu
College of Engineering
Michigan State University
East Lansing, MI 48823, USA

***Abstract*—In order to predict physics achievement in middle school, this paper proposed a new method based on big five model. First, we collected 300 samples, in which 150 passed and the other 150 failed the final physics examination. Then, we submitted the five demographic features and five big-five personality trait features to the artificial neural network (ANN). Third, we used back propagation algorithm to train the ANN. The cross validation results show that our method yielded a sensitivity of 83.00± 2.09%, a specificity of 82.73± 4.12%, and an accuracy of 82.87± 2.75%.**

***Keywords*—big five model; physics achievement; middle school; artificial neural network; back propagation neural network;**

I. INTRODUCTION

The big five personality traits [1], containing openness, conscientiousness, extraversion, agreeableness, and neuroticism (abbreviated as OCEAN) as shown in Fig. 1, are an important psychological model to describe personality for non-only human but also animals like chimpanzees.

Scholars have used the big five personality trait model to predict human behaviors. For example, Nagle [2] used big five to predict video game mechanics based on individual personality. Nishimura [3] studied the satisfaction and frustration in Japan. León [4] investigated the experiences in public sectors, and induced they were related to personality traits.

In this study, we used a five-attribute basic demographic features and a five-attribute big five features, and stretch it out to a ten-entry vector. We then used a pass/fail as the output. We establish an artificial neural network (ANN) to establish mapping between the input and the output.

The structure of this paper is as follows. Section II describes the collected materials. Section III describes the methodology used. Section IV contains the experiments and results. Section V offers the concluding remarks.

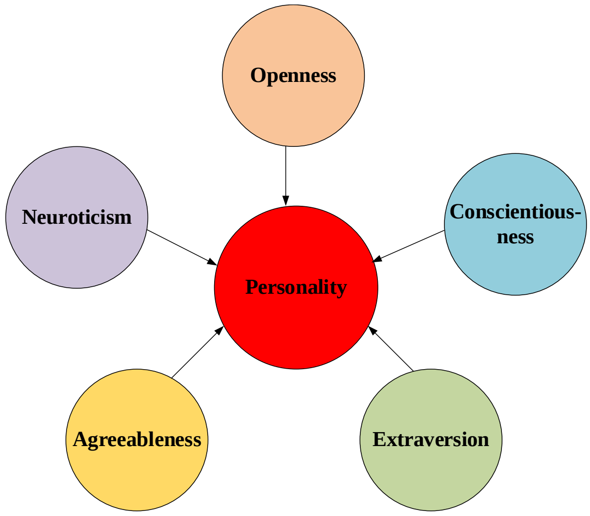

Fig. 1. Big Five Personality Trait Model

II. MATERIALS

A. Subjects

We enrolled in total 300 junior middle school students (100 in Grade 7, 100 in Grade 8, and 100 in Grade 9). We also record their genders, ages, one-child attribute, and locations. The possible values of each attribute were listed in Table 1.

TABLE I. DEMOGRAPHICS OF 60 SUBJECTS FROM A JUNIOR MIDDLE SCHOOL

Characteristics	Value
Gender	Male, Female
Age	11, 12, 13, 14, 15, 16
One-Child	No, Yes
Location	City, Suburb
Grade	Grade 7, Grade 8, Grade 9

B. Big Five Test

We used the shortened version, 60-item inventory, NEO-Five Factor Inventory (NEO-FFI) to test the personal trait of each student. Table 2 shows the personality dimension plot of big five model.

TABLE II. PERSONALITY DIMENSION

Personality dimension	Facets
Openness to experience	Aesthetics, Fantasy, Feelings, Ideas, Values, Actions,
Neuroticism	Hostility, Anxiety, Depression, Impulsiveness, Self-Consciousness, Vulnerability to Stress
Extraversion	Activity, Warmth, Assertiveness, Excitement Seeking, Gregariousness, Positive Emotion
Conscientiousness	Order, Competence, Dutifulness, Achievement Striving, Deliberation, Self-Discipline,
Agreeableness	Modesty, Trust, Altruism, Compliance, Tendermindedness, Straightforwardness,

C. Physics Achievement

In this study, we used the scores of final examination of corresponding grades. For simple, we set the scores over than 60 as pass, and set scores less than 60 as fail. To balance the dataset, we chose 150 students who passed the final exam, and chose another 150 students who failed the final exam.

TABLE III. CATEGORY OF SCORES

Scores	Category
>=60	Pass
<60	Fail

III. METHODOLOGY

There are many methods used for classification. For example, the logistic regression [5] is an improved version of basic linear classifier. Decision tree [6-8] can be regarded as using rules on each attribute, the support vector machine (SVM) [9-17] is a linear classifier using samples lying in the hyperplanes, i.e., support vectors. Fuzzy SVM [18-21] integrated the fuzzy membership function into SVM. Artificial neural network (ANN) [22-26] is an extension of perceptron, and can be trained in back propagation style with descent gradient algorithm. Extreme learning machine [27-29] has a similar structure of ANN, but part of its weights are randomly generated and fixed.

On the other hand, deep learning is the hottest field in classification of very large dataset. The convolutional neural network [30-32] and auto encoders [33-35] are two common models in deep learning.

Nevertheless, our collected 300-sample dataset is too small to use deep learning. Hence, we turn to use traditional artificial neural network due to the universal approximation theory.

A. Artificial Neural Network

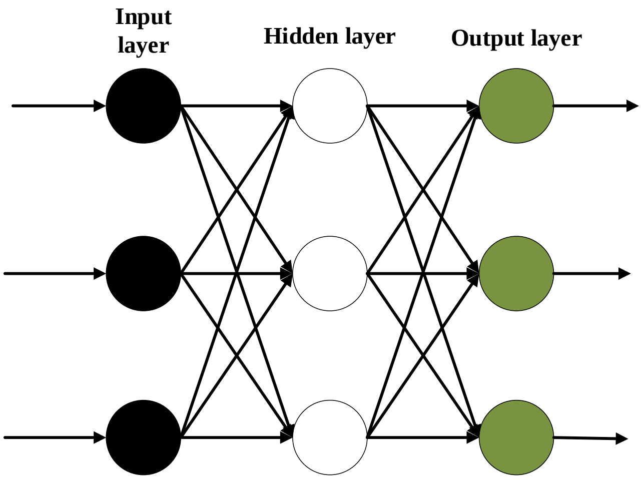

Fig. 2. Architecture of ANN

The commonly used artificial neural network (ANN) contains three layers. The input layer receives the fed features extracted from each samples. The hidden layer contains several hidden neurons. The number is usually determined beforehand. The output layer usually contains the same number of classes, i.e., gives scores for each class [36, 37]. Finally, the class with the maximum score will be selected as the predicted class. The activations for both hidden layer and output layer are Sigmoid function defined as

$$S(x) = \frac{1}{1+\exp(-x)} \tag{1}$$

In this study, we chose the structure of ANN is 10-15-2. That means: we have 10 input neurons, 15 hidden neurons, and 2 output neurons.

B. Cross Validation

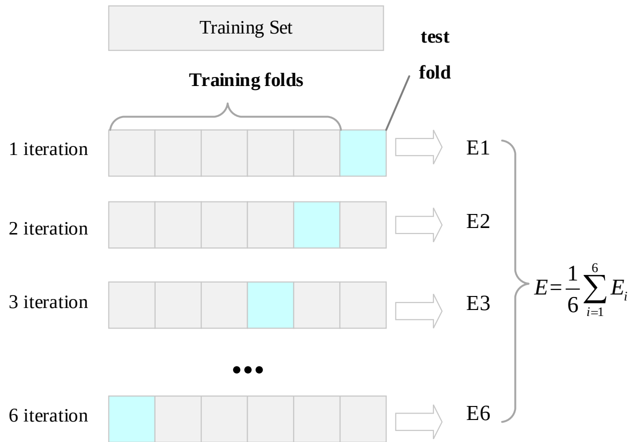

Fig. 3. Diagram of 6-fold cross validation (E represents the error)

6-fold cross validation was used to get the unbiased error. That means, we segment the dataset into six parts, each containing 50 samples. Then, in each trial, we used five folds as training, and the rest for test. We repeated above trial six times, and each time a new fold is used as the test set. Fig. 3 shows a diagram of 6-fold cross validation method.

IV. Experiments and Results

A. *Statistical Analysis of Proposed Method*

We used a 10x6-fold cross validation for checking the classification performance our method. Each fold contains 50 samples. The sensitivity, specificity, and accuracy results are shown in Table 4, Table 5, and Table 6, respectively. Those tables show that our method yielded a sensitivity of 83.00± 2.09%, a specificity of 82.73± 4.12%, and an accuracy of 82.87± 2.75%.

TABLE IV. Sensitivity of our Method

Run	Fold 1	Fold 2	Fold 3	Fold 4	Fold 5	Fold 6	Total
1	84	92	80	84	84	84	84.67
2	88	88	88	84	84	84	86.00
3	88	80	84	84	84	80	83.33
4	72	84	72	88	84	80	80.00
5	80	80	88	84	88	80	83.33
6	80	88	76	88	84	80	82.67
7	88	80	84	92	84	88	86.00
8	84	80	76	80	76	96	82.00
9	84	80	80	76	80	84	80.67
10	84	76	84	84	84	76	81.33
AVG	83.00± 2.09						

TABLE V. Specificity of our Method

Run	Fold 1	Fold 2	Fold 3	Fold 4	Fold 5	Fold 6	Total
1	76	84	88	80	84	92	84.00
2	92	88	92	84	92	88	89.33
3	76	80	80	88	84	88	82.67
4	76	84	76	80	80	76	78.67
5	84	84	84	76	76	84	81.33
6	88	88	80	88	84	80	84.67
7	84	80	92	88	80	76	83.33
8	76	68	80	80	76	76	76.00
9	88	80	80	72	76	80	79.33
10	88	88	88	88	88	88	88.00
AVG	82.73± 4.12						

TABLE VI. Accuracy of our Method

Run	Fold 1	Fold 2	Fold 3	Fold 4	Fold 5	Fold 6	Total
1	80	88	84	82	84	88	84.33
2	90	88	90	84	88	86	87.67
3	82	80	82	86	84	84	83.00
4	74	84	74	84	82	78	79.33
5	82	82	86	80	82	82	82.33
6	84	88	78	88	84	80	83.67
7	86	80	88	90	82	82	84.67
8	80	74	78	80	76	86	79.00
9	86	80	80	74	78	82	80.00
10	86	82	86	86	86	82	84.67
AVG	82.87± 2.75						

B. *Confusion Matrix*

The confusion matrixes of proposed method and ideal situation were plotted in Fig. 4. We can obverse that both confusion matrixes have a summation of 3,000, which is the 10 times of the size of dataset. For the ideal confusion matrix, 0 errors were made. By contrary, our proposed method misclassified 255 pass samples to fail, and misclassified 259 fail samples to pass.

$$\begin{bmatrix} 1245 & 255 \\ 259 & 1241 \end{bmatrix}$$

(a) proposed method

$$\begin{bmatrix} 1500 & 0 \\ 0 & 1500 \end{bmatrix}$$

(b) ideal

Fig. 4. Confusion Matrix

C. *Effect of Features*

In this experiment, we used only demographic features and only big five features. We measured the classifier by accuracy; the results are depicted in Table 7 and Table 8, respectively.

TABLE VII. Accuracy using only Demographic Features

Run	Fold 1	Fold 2	Fold 3	Fold 4	Fold 5	Fold 6	Total
1	52	56	52	50	52	52	52.33
2	52	58	56	56	58	52	55.33
3	54	58	62	56	58	58	57.67
4	58	56	66	52	58	54	57.33
5	54	58	54	58	52	58	55.67
6	58	60	60	64	60	58	60.00
7	60	56	60	60	60	56	58.67
8	58	62	62	58	60	60	60.00
9	56	56	54	60	58	56	56.67
10	60	58	58	56	58	58	58.00
AVG	57.17± 2.32						

TABLE VIII. Accuracy using only Big Five Features

Run	Fold 1	Fold 2	Fold 3	Fold 4	Fold 5	Fold 6	Total
1	72	74	70	70	72	70	71.33
2	78	74	76	78	82	78	77.67
3	82	78	78	80	82	78	79.67
4	80	74	78	78	80	74	77.33
5	72	72	76	76	70	70	72.67
6	72	72	76	76	72	76	74.00
7	70	74	76	76	72	68	72.67
8	70	78	76	82	80	78	77.33
9	76	80	72	76	74	80	76.33
10	78	82	86	84	84	82	82.67
AVG	76.17± 3.53						

Here we observe that the accuracy significantly decreased to only 57.17± 2.32% if only using the five demographic features, and slightly decreased to 76.17± 3.53% if only using the five big five features. Hence, we can conclude that the big five is more efficient in predicting physics achievement than demographic features.

V. Conclusions

In this study, we proposed a new method to predict physics achievements in middle school by big five model and artificial neural network. The results showed the effectiveness of proposed method.

In the future, we shall try to collect more data. In addition, we shall test other classifiers and try advanced optimization methods [38].

ACKNOWLEDGMENT

The paper is financially supported by Postgraduate Research & Practice Innovation Program of Jiangsu Province (SJCX17_0342).

REFERENCES

[1] Stajkovic, A.D., A. Bandura, et al., Test of three conceptual models of influence of the big five personality traits and self-efficacy on academic performance: A meta-analytic path-analysis. Personality And Individual Differences, 2018. 120: pp. 238-245

[2] Nagle, A., P. Wolf, et al., Towards a system of customized video game mechanics based on player personality: Relating the Big Five personality traits with difficulty adaptation in a first-person shooter game. Entertainment Computing, 2016. 13: pp. 10-24

[3] Nishimura, T. and T. Suzuki, Basic Psychological Need Satisfaction and Frustration in Japan: Controlling for the Big Five Personality Traits. Japanese Psychological Research, 2016. 58(4): pp. 320-331

[4] León, F.R., Is experience in the public sector associated with Big Five personality aspects? Liberabit, 2017. 23(1): pp. 103-109

[5] Zhan, T.M. and Y. Chen, Multiple Sclerosis Detection Based on Biorthogonal Wavelet Transform, RBF Kernel Principal Component Analysis, and Logistic Regression. IEEE Access, 2016. 4: pp. 7567-7576

[6] Zhou, X.-X., Comparison of machine learning methods for stationary wavelet entropy-based multiple sclerosis detection: decision tree, k-nearest neighbors, and support vector machine. Simulation, 2016. 92(9): pp. 861-871

[7] Zhang, Y., Binary PSO with mutation operator for feature selection using decision tree applied to spam detection. Knowledge-Based Systems, 2014. 64: pp. 22-31

[8] Dong, Z., Classification of Alzheimer disease based on structural magnetic resonance imaging by kernel support vector machine decision tree. Progress In Electromagnetics Research, 2014. 144: pp. 171-184

[9] Chen, Y. and X.-Q. Chen Sensorineural hearing loss detection via discrete wavelet transform and principal component analysis combined with generalized eigenvalue proximal support vector machine and Tikhonov regularization. Multimedia Tools and Applications, 2016, DOI: 10.1007/s11042-016-4087-6 (Online).

[10] Zhou, X.-X., J.-F. Yang, et al., Combination of stationary wavelet transform and kernel support vector machines for pathological brain detection. Simulation, 2016. 92(9): pp. 827-837

[11] Gorriz, J.M. and J. Ramírez, Wavelet entropy and directed acyclic graph support vector machine for detection of patients with unilateral hearing loss in MRI scanning. Frontiers in Computational Neuroscience, 2016. 10, Article ID: 160

[12] Chen, M., Morphological analysis of dendrites and spines by hybridization of ridge detection with twin support vector machine. PeerJ, 2016. 4, Article ID: e2207

[13] Yang, M., Dual-Tree Complex Wavelet Transform and Twin Support Vector Machine for Pathological Brain Detection. Applied Sciences, 2016. 6(6), Article ID: 169

[14] Chen, S., J.-F. Yang, et al., Magnetic resonance brain image classification based on weighted-type fractional Fourier transform and nonparallel support vector machine. International Journal of Imaging Systems and Technology, 2015. 25(4): pp. 317-327

[15] Liu, A., Magnetic resonance brain image classification via stationary wavelet transform and generalized eigenvalue proximal support vector machine. Journal of Medical Imaging and Health Informatics, 2015. 5(7): pp. 1395-1403

[16] Yang, J., Preclinical diagnosis of magnetic resonance (MR) brain images via discrete wavelet packet transform with Tsallis entropy and generalized eigenvalue proximal support vector machine (GEPSVM). Entropy, 2015. 17(4): pp. 1795-1813

[17] Wu, L., Classification of fruits using computer vision and a multiclass support vector machine. Sensors, 2012. 12(9): pp. 12489-12505

[18] Li, Y., Detection of Dendritic Spines using Wavelet Packet Entropy and Fuzzy Support Vector Machine. CNS & Neurological Disorders - Drug Targets, 2017. 16(2): pp. 116-121

[19] Lu, H.M., Facial Emotion Recognition Based on Biorthogonal Wavelet Entropy, Fuzzy Support Vector Machine, and Stratified Cross Validation. IEEE Access, 2016. 4: pp. 8375-8385

[20] Liu, G., Pathological brain detection in MRI scanning by wavelet packet Tsallis entropy and fuzzy support vector machine. SpringerPlus, 2015. 4(1), Article ID: 716

[21] Yang, J., Identification of green, Oolong and black teas in China via wavelet packet entropy and fuzzy support vector machine. Entropy, 2015. 17(10): pp. 6663-6682

[22] Feng, C., Feed-forward neural network optimized by hybridization of PSO and ABC for abnormal brain detection. International Journal of Imaging Systems and Technology, 2015. 25(2): pp. 153-164

[23] Rao, R.V. and A. Liu, Abnormal Breast Detection in Mammogram Images by Feed-forward Neural Network trained by Jaya Algorithm. Fundamenta Informaticae, 2017. 151(1-4): pp. 191-211

[24] Ji, G., Fruit classification using computer vision and feedforward neural network. Journal of Food Engineering, 2014. 143: pp. 167-177

[25] Wei, L., Fruit classification by wavelet-entropy and feedforward neural network trained by fitness-scaled chaotic ABC and biogeography-based optimization. Entropy, 2015. 17(8): pp. 5711-5728

[26] Wu, J., Fruit classification by biogeography-based optimization and feedforward neural network. Expert Systems, 2016. 33(3): pp. 239-253

[27] Zhao, G. Smart Pathological brain detection by Synthetic Minority Oversampling Technique, Extreme Learning Machine, and Jaya Algorithm. Multimedia Tools and Applications, 2017, DOI: 10.1007/s11042-017-5023-0 (Online).

[28] Wu, Y., Extreme learning machine used for focal liver lesion identification. Journal of Gastroenterology and Hepatology, 2017. 32(S3): p. 168

[29] Lu, S. and X. Qiu, A Pathological Brain Detection System based on Extreme Learning Machine Optimized by Bat Algorithm. CNS & Neurological Disorders - Drug Targets, 2017. 16(1): pp. 23-29

[30] Muhammad, K. Image based fruit category classification by 13-layer deep convolutional neural network and data augmentation. Multimedia Tools and Applications, 2017, DOI: 10.1007/s11042-017-5243-3 (Online).

[31] Sun, J. Polarimetric synthetic aperture radar image segmentation by convolutional neural network using graphical processing units. Journal of Real-Time Image Processing, 2017, DOI: 10.1007/s11554-017-0717-0 (Online).

[32] Lv, Y.-D. and Y. Sui Alcoholism detection by data augmentation and convolutional neural network with stochastic pooling. Journal of Medical Systems, 2017, DOI: 10.1007/s10916-017-0845-x (Online).

[33] Jia, W., Three-Category Classification of Magnetic Resonance Hearing Loss Images Based on Deep Autoencoder. Journal of Medical Systems, 2017. 41, Article ID: 165

[34] Jia, W. Five-category classification of pathological brain images based on deep stacked sparse autoencoder. Multimedia Tools and Applications, 2017, DOI: 10.1007/s11042-017-5174-z (Online).

[35] Chen, H. Seven-lay deep neural network based on sparse autoencoder for voxelwise detection of cerebral microbleed. Multimedia Tools and Applications, 2017, DOI: 10.1007/s11042-017-4554-8 (Online).

[36] Sun, Y., A Multilayer Perceptron Based Smart Pathological Brain Detection System by Fractional Fourier Entropy. Journal of Medical Systems, 2016. 40(7), Article ID: 173

[37] Wang, S.-H. Single slice based detection for Alzheimer's disease via wavelet entropy and multilayer perceptron trained by biogeography-based optimization. Multimedia Tools and Applications, 2016, DOI: 10.1007/s11042-016-4222-4 (Online).

[38] Ji, G., A comprehensive survey on particle swarm optimization algorithm and its applications. Mathematical Problems in Engineering, 2015. 2015, Article ID: 931256