Zhang Y. (2024) DEVELOPING A BIG DATA-DRIVEN ASSOCIATION MODEL LINKING ADOLESCENT PHYSICAL EXERCISE BEHAVIOR AND MENTAL HEALTH. Revista Internacional de Medicina y Ciencias de la Actividad Física y el Deporte vol. 24 (97) pp. 59-71.
DOI: https://doi.org/10.15366/rimcafd2024.97.005

ORIGINAL

DEVELOPING A BIG DATA-DRIVEN ASSOCIATION MODEL LINKING ADOLESCENT PHYSICAL EXERCISE BEHAVIOR AND MENTAL HEALTH

Yuancai Zhang

College of Physical Education and Health, East China Normal University, Shanghai 200241, China
E-mail: zhangyuancai666@163.com

Recibido 05 de diciembre de 2023 **Received** December 05, 2023
Aceptado 04 de julio de 2024 **Accepted** July 04, 2024

ABSTRACT

This study examines the correlation between physical exercise habits and mental health outcomes in adolescents using a regression-based Multilayer Perceptron (MLP) model. By utilizing a substantial dataset, the model successfully captures intricate, non-linear connections between several exercise-related characteristics and mental health scores. The MLP model exhibited robust prediction accuracy, characterized by elevated R^2 values and diminished MSE, underscoring its promise for practical implementations. By integrating big data, a thorough analysis was conducted, which demonstrated that engaging in regular physical activity is a substantial indicator of mental well-being. The results emphasize the significance of engaging in physical activity to enhance mental well-being and showcase the effectiveness of machine learning and big data in health evaluations. Subsequent research could investigate the integration of supplementary lifestyle elements and more sophisticated deep learning architectures to further improve prediction abilities.

KEYWORDS: Physical Exercise, Mental Health, Adolescents, Multilayer Perceptron, Big Data, Machine Learning, Health Assessment

1. INTRODUCTION

In recent years, there has been a growing focus on the physical and mental well-being of teenagers in China. This is due to worries about the decreasing levels of physical exercise and the increasing occurrence of related health problems. Chinese teenagers are experiencing a worrisome decrease in their physical fitness, with problems including obesity, early development of nearsightedness, and overall poor physical health becoming more common (Dang et al., 2024). These developments are indicative of more general lifestyle shifts, which include a decline in regular physical exercise among young individuals. The importance of this decrease is underscored by the results of a nationwide survey conducted in 2005, which indicated that more than half of Chinese students failed to partake in the recommended one hour of daily physical activity, and a quarter of students abstained from any form of physical exercise entirely. Despite some efforts to address the issue, recent research comparing the 2005 and 2019 national student physical fitness surveys shows that the prevalence of obesity, overweight, and vision impairment among teenagers are still increasing. The results emphasize the immediate requirement for successful interventions to improve the physical fitness and general health of Chinese teenagers (Z. Chen et al., 2022). Physical exercise is universally recognized as a key factor in maintaining and improving the physical health of adolescents (Ortega, Ruiz, Castillo, & Sjöström, 2008). Regular participation in physical activities, such as running, swimming, and team sports like basketball and soccer, is crucial for enhancing cardiovascular health, strengthening muscles, and improving flexibility. These activities also play a vital role in preventing lifestyle-related diseases, such as obesity and cardiovascular conditions, which are increasingly being observed among younger populations. Moreover, the benefits of physical exercise extend beyond physical health; it is also a powerful tool for mental well-being (PARRA et al., 2020). Engaging in regular physical activity has been shown to significantly reduce levels of stress, anxiety, and depression among adolescents, offering a natural and effective way to manage the psychological pressures associated with academic life. Furthermore, physical exercise fosters the development of essential social skills, providing adolescents with opportunities to work as part of a team, develop leadership abilities, and build self-confidence and self-esteem. Despite the well-documented benefits of physical exercise, the current state of adolescent physical activity in China remains a cause for concern. Existing research indicates that while some forms of exercise, such as running, basketball, badminton, and table tennis, are popular among Chinese adolescents (P. Chen et al., 2020), there are notable gender differences in the types of activities chosen, and the overall frequency and duration of exercise are often insufficient. For instance, while many students report engaging in physical exercise, a significant number do not meet the recommended frequency or duration necessary to reap the full health benefits. The findings of various studies show that a substantial percentage of students exercise less than three times per week, and when they do exercise, the duration often falls short of the recommended one hour per session. Moreover, only a small percentage of students engage in physical exercise for an hour each day, with many exercising for less than 30 minutes. This insufficient exercise duration is compounded by the fact that a large proportion of students experience significant fatigue after just 30 minutes of exercise, indicating a lack of physical fitness and endurance (Loy, O'Connor, & Dishman, 2013). This fatigue can discourage further participation in physical activity, creating a cycle of inactivity that is detrimental to both physical and mental health. The determinants affecting the involvement of adolescents in physical exercise are intricate and diverse, with personal characteristics such as individual motivation and self-efficacy playing crucial roles. Intrinsic motivation, specifically, plays a crucial role in determining whether an adolescent will participate in and maintain regular physical activity. Studies in the field of exercise psychology have demonstrated that motivation plays a vital role in connecting individual personality features with exercise behavior. Adolescents that possess elevated levels of intrinsic motivation are more inclined to engage in consistent physical activity, since they are motivated by internal incentives such as personal fulfillment and enjoyment of the activity. On the other hand, those who have lower levels of inherent drive may find it difficult to consistently engage in exercise routines, especially if they do not see immediate advantages from their exertions. Self-efficacy, which refers to an individual's confidence in their capability to achieve success in particular conditions or complete a task, is a significant determinant of exercise behavior in adolescents (Pajares, 2006). Increased levels of self-efficacy have been linked to increased engagement in physical activity, as those who have confidence in their ability to succeed are more inclined to tackle challenges and persevere in the presence of obstacles. Adolescents who cultivate a robust belief in their ability to succeed in physical exercise are more likely to engage in regular physical activity. This, in turn, leads to improved physical and mental health outcomes (Robbins, Pender, Ronis, Kazanis, & Pis, 2004). Research has demonstrated that self-efficacy has a direct impact on exercise behavior and also acts as a mediator between other psychological characteristics, such as health attitudes, and exercise habits. Adolescents who possess a strong belief in their own abilities are more inclined to embrace and sustain consistent exercise regimens, resulting in notable enhancements in both physical fitness and mental well-being. Given the significant health challenges faced by Chinese adolescents, there is a pressing need to explore the relationship between physical exercise behavior and mental health outcomes in this population. Advances in big data analysis provide a powerful tool for investigating these complex relationships (Provost & Fawcett, 2013), enabling researchers to construct models that can elucidate the associations between various factors influencing physical activity and mental health. By leveraging large datasets, researchers can gain insights into the patterns and predictors of physical exercise behavior among adolescents (Colmenarejo, 2020), identify key factors that promote or hinder participation in physical activity, and develop targeted interventions aimed at improving both physical and mental health outcomes (Liang, Zheng, & Zeng, 2019). The objective of this study is to investigate the complex connections between physical exercise habit and mental health in Chinese adolescents, with a specific emphasis on the influence of motivation and self-efficacy. This research aims to develop a model using big data analysis to gain a thorough understanding of how many factors combine to impact exercise habits and mental health outcomes. The results of this study have the capacity to enhance the current knowledge in the area of adolescent health, providing valuable perspectives that can guide the creation of efficient public health initiatives and interventions (Marsch, 2021). Through the promotion of consistent physical exercise and the improvement of motivation and self-confidence among teenagers, these treatments can effectively tackle the urgent health issues that this group faces and contribute to the overall well-being of Chinese young in the long run (Provost & Fawcett, 2013).

2. Methodology

2.1 Participants Description

The study employed a combination of stratified sampling and cluster sampling techniques to gather data from students at six middle schools. The poll included a sample of 360 eighth-grade pupils. Among these, a total of 330 questionnaires were effectively gathered, resulting in a response rate of 91.67%. After conducting a comprehensive screening process, a total of 319 questionnaires were determined to be legitimate, resulting in an effective rate of 88.61%. Table 1 presents the comprehensive breakdown of the sample's distribution throughout the schools, as well as the number of questionnaires that were delivered, collected, and considered legitimate.

Table 1: Distribution of Survey Participants

SCHOOL NAME	GRADE	DISTRIBUTED	COLLECTED	VALID
SCHOOL 1	8	60	55	53
SCHOOL 2	8	60	58	55
SCHOOL 3	8	60	55	53
SCHOOL 4	8	60	59	58
SCHOOL 5	8	60	56	54
SCHOOL 6	8	60	47	46
TOTAL		360	330	319

2.2 Measurement of Physical Exercise Behavior

The study employed the Physical Activity Rating Scale (PARS) (Teixeira, Carraça, Markland, Silva, & Ryan, 2012) to evaluate the individuals' physical exercise behavior. This scale is specifically created to assess many dimensions of physical exercise behavior, encompassing factors such as the level of exertion, the length of time engaged in exercise, and the number of exercise sessions completed during the previous month. The PARS is composed of three components, each evaluated using a five-point rating system:

Exercise Intensity: Rated from 1 (very light) to 5 (very intense), with corresponding scores ranging from 1 to 5 points. Exercise Duration: Measured by the duration of each exercise session, rated from 1 (less than 10 minutes) to 5 (more than 60 minutes). The scoring for duration is slightly different, ranging from 0 to 4 points. Exercise Frequency: Assessed by the number of exercise sessions per week, rated from 1 (less than once per week) to 5 (more than five times per week), with scores from 1 to 5 points. The overall physical activity level was calculated using the formula:

$$Exercise\_Volume = Intensity \times Duration \times Frequency \qquad (1)$$

Based on this formula, the total exercise volume was classified into three categories: Low Exercise Volume: 0-19 points, Moderate Exercise Volume: 20-42 points, High Exercise Volume: 43-100 points. Table 2 and Figure 1 below presents the distribution of participants across these categories, while the accompanying bar chart visually represents these data.

Table 2: Distribution of Participants by Exercise Volume Category

EXERCISE VOLUME CATEGORY	NUMBER OF PARTICIPANTS
LOW (0-19)	134
MODERATE (20-42)	110
HIGH (43-100)	47

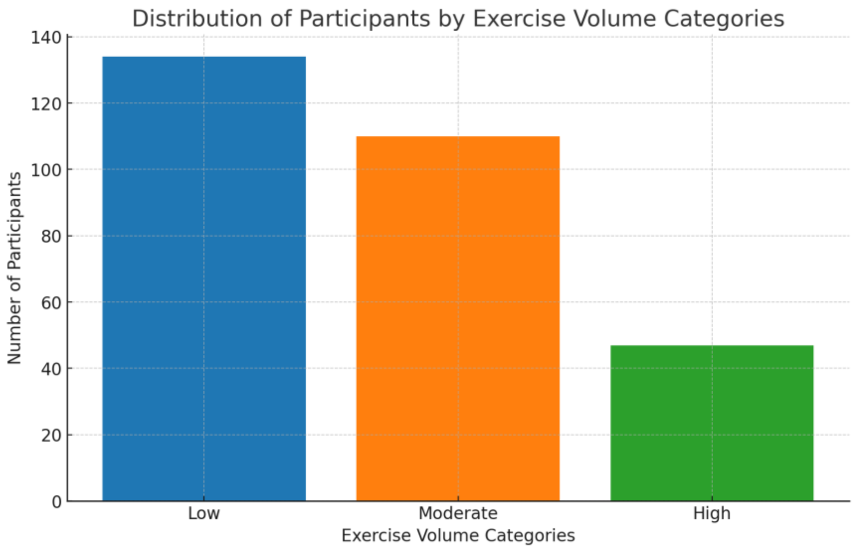

Figure 1: Distribution of Participants by Exercise Volume Categories

The data indicate that the majority of participants fall into the Low and Moderate exercise volume categories, with fewer participants in the High category. Specifically, 134 participants reported low exercise volumes, 110 participants reported moderate exercise volumes, and only 47 participants engaged in high exercise volumes. This distribution suggests that while some students are moderately active, a significant proportion engage in minimal physical activity, highlighting the need for interventions to encourage higher levels of exercise among adolescents.

2.3 Measurement of Mental Health

To assess the mental health of participants, the study utilized a standardized Mental Health Inventory (MHI) (Thorsen, Rugulies, Hjarsbech, & Bjorner, 2013), which evaluates various aspects of psychological well-being. The MHI consists of several subscales, typically including measures for anxiety, depression, positive well-being, emotional ties, and general distress. The MHI assesses each issue using a Likert scale, often ranging from 1 (Never) to 5 (Always). The overall mental health score is determined by applying the following formula:

$$Mental\_Health\_Score = \frac{\sum(Positive\_Items) - \sum(Negative\_Items)}{Number\_of\_Items} \quad (2)$$

where $Positive\_Items$ represent items that contribute to a positive mental health outcome (e.g., feelings of well-being, satisfaction). $Negative_Items$represent items that reflect negative mental health outcomes (e.g., anxiety, depression). Scores are categorized as follows: Low Mental Health (0-1.99 points), Moderate Mental Health (2-3.99 points), High Mental Health (4-5 points). Table 3 and Figure 2 below presents the distribution of participants across these categories, while the accompanying bar chart visually represents these data.

Table 3: Distribution of Participants by Mental Health Category

MENTAL HEALTH CATEGORY	NUMBER OF PARTICIPANTS
LOW (0-1.99)	75
MODERATE (2-3.99)	180
HIGH (4-5)	64

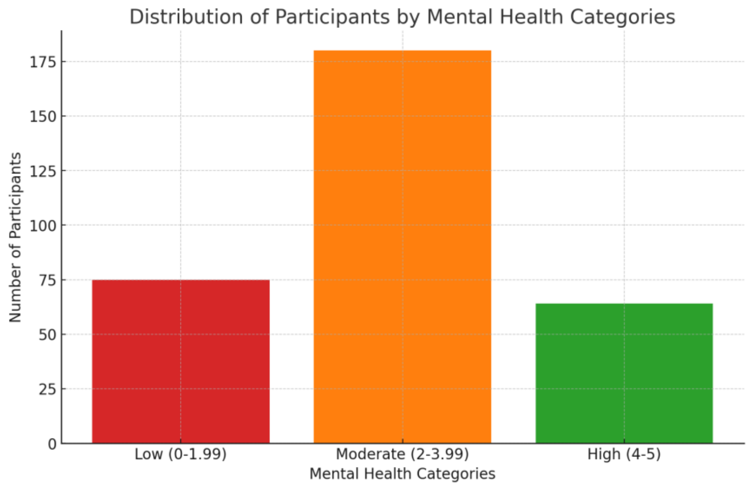

Figure 2: Distribution of Participants by Mental Health Categories

3. Construction of Association Model Using Deep Learning

In this study, a deep learning approach was employed to construct an association model that links physical exercise behavior and mental health among adolescents. Leveraging the power of big data, the model was designed to uncover complex, non-linear relationships between these variables, providing insights that traditional statistical methods may not capture.

3.1 Data Preprocessing

The data preprocessing phase was essential for preparing the dataset, gathered from 319 valid student responses, for analysis using deep learning. The primary steps involved data cleaning, transformation, feature engineering, and dataset splitting. The raw data contained issues such as missing values, outliers, and potential duplicates. Missing data, particularly in exercise duration and mental health scores, was addressed through imputation, using the mean for continuous variables and the mode for categorical variables. Outliers were identified using the Z-score method, with extreme values adjusted to minimize their impact. Duplicate records were removed to ensure data integrity. Categorical variables, such as exercise type and gender, were converted into numerical values using one-hot encoding. This transformation prevented any artificial hierarchy between categories.

To ensure uniformity across different variables, normalization was applied, scaling all features to a range between 0 and 1, which helped in balancing their influence on the model (Bilu et al., 2023). To enhance the model's predictive power, interaction features were created by combining key variables like exercise intensity and frequency. This allowed the model to capture more complex relationships. Correlation analysis was used to select the most relevant features, reducing noise and improving model efficiency. The dataset was divided into a training set, which accounted for 70% of the data, and a validation set, which accounted for the remaining 30%. This division was done in order to train the model and evaluate its capacity to generalize. The partitioning of the data was important in mitigating overfitting and guaranteeing the model's ability to generalize effectively to novel, unseen data.

3.2 Model Construction Using Multilayer Perceptron

This section provides a comprehensive explanation of how a Multilayer Perceptron (MLP) model, based on regression, is built to investigate the correlation between physical exercise behavior and mental health scores in adolescents. The Multilayer Perceptron (MLP) is a specific sort of neural network that excels in representing intricate and non-linear connections within organized data.

3.2.1 Structure of model

(1) Input Layer: The input layer is the first step in the MLP model and is responsible for receiving the features that have been preprocessed and engineered from the dataset. The features in this study include: Exercise-Related Variables: These variables include exercise intensity, duration, frequency, and interaction terms (e.g., intensity * frequency). These features provide a detailed representation of the physical exercise behavior of the participants. Mental Health Scores: Derived from the Mental Health Inventory (MHI), this continuous variable represents the mental health status of the participants, which is the target variable for the regression model. Demographic variables: such as age and gender, are incorporated to consider the impact of demographic factors on mental health. The variables are encoded suitably, with categorical variables converted into numerical representations using techniques such as one-hot encoding. The quantity of neurons in the input layer is directly proportional to the number of characteristics. For example, if the dataset has 15 characteristics, the input layer will consist of 15 neurons. Each neuron in the input layer corresponds to a single characteristic of the dataset.

(2) Hidden Layers: The hidden layers of the model are responsible for acquiring the ability to comprehend and represent the intricate connections between the input features and the goal variable. The performance of a regression-based MLP model is highly dependent on the configuration of its hidden layers. The setup that is currently being utilized is as follows: Layer Configuration: First Hidden Layer: Consists of 64 neurons with a Rectified Linear Unit (ReLU) activation function. The ReLU function introduces non-linearity into the model, allowing it to capture complex patterns in the data. Second Hidden Layer: Comprises 32 neurons, also using the ReLU activation function. This layer further processes the information passed from the first hidden layer, refining the model's understanding of the relationships between features. Third Hidden Layer: Includes 16 neurons with ReLU activation, depending on the complexity of the data and the model's performance in preliminary testing. This layer can help improve the model's ability to generalize by capturing additional patterns. The ReLU activation function is defined as:

$$ReLU(x) = max(0, x) \qquad (3)$$

This function outputs the input directly if it is positive; otherwise, it outputs zero. ReLU is advantageous because it avoids the vanishing gradient problem, which can occur with other activation functions like sigmoid or tanh.

(3) Output Layer: The output layer in this regression model is designed to generate a continuous value, which is the predicted mental health score based on the input features. Regression Output: The output layer contains a single neuron with a linear activation function. The linear activation function is

appropriate for regression tasks because it allows the model to output a continuous value across a range of possible outcomes. The linear activation function is defined as:

$$f(x) = x \qquad (4)$$

This function ensures that the output of the neuron is a direct linear combination of the inputs, suitable for predicting a continuous target variable like the mental health score.

(4) Loss Function: The loss function measures the difference between the predicted values and the actual values in the training data. For regression tasks, the Mean Squared Error (MSE) is commonly used:

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2 \qquad (5)$$

where y_i is the actual mental health score, $\hat{y}_i$is the predicted score, and n is the number of samples. The MSE penalizes larger errors more severely than smaller ones, which helps the model focus on minimizing large deviations between predictions and actual outcomes. The training process employs the Adam optimizer. Adam, short for Adaptive Moment Estimation, is a widely used optimization technique in the field of deep learning. It is favored due to its ability to incorporate the advantages of two additional stochastic gradient descent extensions, namely AdaGrad and RMSProp. Adam optimizes the learning rate during training by considering the gradients' moments, resulting in efficient and successful training of deep learning models.

3.2.2. Model Training

The Model Training process for the regression-based Multilayer Perceptron (MLP) model in this study involves several key steps aimed at predicting the mental health scores based on input features like physical exercise behavior and demographic data. The key steps in model training include: Data Preparation: The dataset is divided into a training set, which comprises 70% of the data, and a validation set, which comprises the remaining 30%. The training set is utilized to instruct the model, whereas the validation set is employed to evaluate the model's performance on unfamiliar data, confirming its ability to generalize well. Initialization: The model's parameters, including weights and biases, are initialized randomly. This initial setup is crucial as it can influence the model's convergence and the quality of the final predictions. Forward Propagation: The input data is passed through the network layers, and the output (predicted mental health score) is computed. Backward Propagation: The error between the predicted and actual scores is calculated using the MSE. The gradients of this error with respect to each weight in the network are computed, and the weights are updated using the Adam optimizer.

Epochs: The entire training dataset is passed through the network multiple times, with each pass called an epoch. The number of epochs typically ranges from 50 to 200, depending on how quickly the model converges.

4. Model Evaluation

The Model Evaluation process is a critical step in determining how well the trained regression-based Multilayer Perceptron (MLP) model performs on unseen data. This process involves assessing the model's predictive accuracy, generalization ability, and overall fit using specific evaluation metrics.

4.1 Performance Metrics

Two primary metrics are used to evaluate the performance of the MLP regression model: R-Squared (R^2) and Mean Squared Error (MSE).

4.1.1 R-Squared (R^2)

The coefficient of determination, denoted as R^2, quantifies the percentage of variability in the dependent variable (mental health scores) that can be explained by the independent factors (exercise activity, demographics, etc.). It offers a measure of the degree to which the model's predictions align with the real data. The formula for R^2 is:

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2} \quad (6)$$

where, y_i denotes the actual mental health score for the i^{th} observation. $\hat{y}_i$ denotes the predicted mental health score for the i^{th} observation. $\bar{y}$ denotes the mean of the actual mental health scores. n denotes the number of observations. R^2 ranges from 0 to 1, where 1 indicates that the model perfectly predicts the target variable, and 0 indicates that the model does no better than simply predicting the mean of the target variable.

4.1.2 Mean Squared Error (MSE)

MSE measures the average squared difference between the actual and predicted values. It provides a direct measure of the prediction error, with lower values indicating better model performance. The formula for MSE is:

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2 \quad (7)$$

where y_i is the actual mental health score, $\hat{y}_i$ is the predicted score, and n is the number of samples. MSE is always non-negative, and values closer to 0 indicate better model performance.

4.2 Presentation of Results

The results of the model evaluation are typically presented in Table 4, summarizing the performance metrics for easy comparison and interpretation.

Table 4: Model Evaluation Metrics

METRIC	VALUE
R-SQUARED (R²)	0.85
MEAN SQUARED ERROR (MSE)	3.45

High R^2 Indicates that the model explains a large portion of the variance in mental health scores, suggesting strong predictive power. Low MSE Implies that the model's predictions are, on average, close to the actual values, indicating good accuracy. However, it's essential to consider both metrics together, as a model might have a high R^2 but still have significant prediction errors if the MSE is large. Conversely, a low MSE with a low R^2 might indicate that while predictions are close to actual values, the model might not be capturing all the underlying variance in the data.

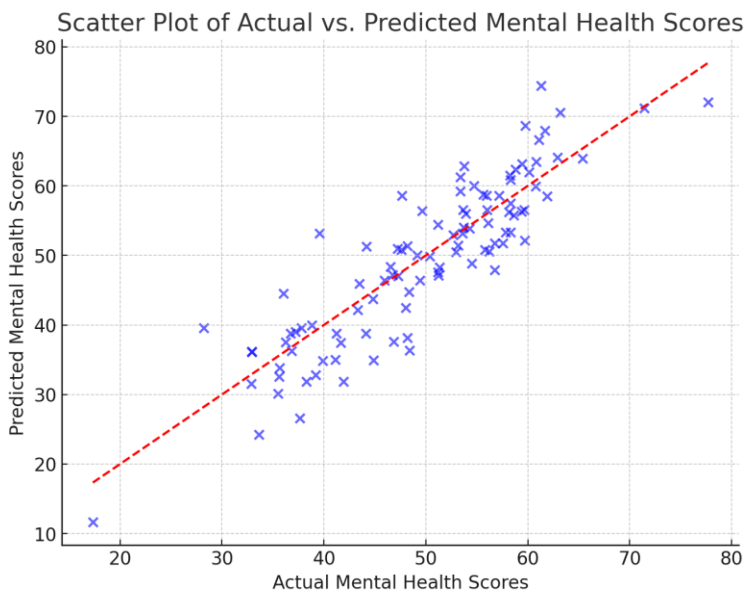

Figure 3: Scatter Plot of Actual vs. Predicted Mental Health Scores

Figure 3 is the scatter plot showing the relationship between actual and predicted mental health scores. The blue points represent the data, and the red dashed line indicates the line of perfect prediction (where the actual and predicted scores would be equal). The plot typically displays the actual mental health scores y_i on the x-axis and the predicted scores ($\hat{y}_i$) on the y-axis. A line of perfect prediction (where $y = \hat{y}$) is also plotted for reference. This visual representation helps assess how closely the model's predictions match the actual values. Points near the red line indicate better predictions. The Model Evaluation process is a comprehensive assessment that involves calculating key metrics like R^2 and MSE on the validation set. These metrics provide insights into the model's ability to predict mental health outcomes based on the input features. The results are typically summarized in a table and can be further analyzed using visual tools like scatter plots to ensure the model's robustness and accuracy in practical applications.

5. Conclusion

This study leveraged big data and a regression-based Multilayer Perceptron (MLP) model to explore the relationship between physical exercise behavior and mental health among adolescents. By analyzing a large dataset, the model was able to capture complex, non-linear relationships between various exercise-related features and mental health outcomes, demonstrating strong predictive accuracy with high R-Squared (R^2) values and low Mean Squared Error (MSE). The integration of big data allowed for a more comprehensive analysis, enabling the model to generalize well across diverse samples and providing robust insights into how physical activity influences mental health. The findings suggest that regular physical exercise is a significant predictor of mental health, and the model's robustness supports its potential application in real-world settings. The utilization of big data in this study emphasizes its efficacy in revealing intricate patterns and correlations that may not be discernible in smaller datasets. This methodology enhances the precision of forecasts and allows for the creation of tailored interventions to enhance the mental well-being of adolescents. Subsequent investigations could improve the model by integrating supplementary variables like food, sleep, and social contacts, or by investigating more sophisticated deep learning structures. This study emphasizes the crucial significance of physical exercise in promoting mental well-being and the revolutionary capacity of big data in evaluating and implementing health assessments and interventions.

Reference

Bilu, Y., Flaks-Manov, N., Bivas-Benita, M., Akiva, P., Kalkstein, N., Yehezkelli, Y., . . . Lerner, U. (2023). Data-driven assessment of adolescents' mental health during the COVID-19 pandemic. *Journal of the American Academy of Child & Adolescent Psychiatry, 62*(8), 920-937.

Chen, P., Wang, D., Shen, H., Yu, L., Gao, Q., Mao, L., . . . Zhang, Y. (2020). Physical activity and health in Chinese children and adolescents: expert consensus statement (2020). *British journal of sports medicine, 54*(22), 1321-1331.

Chen, Z., Chi, G., Wang, L., Chen, S., Yan, J., & Li, S. (2022). The combinations of physical activity, screen time, and sleep, and their associations with self-reported physical fitness in children and adolescents. *International journal of environmental research and public health, 19*(10), 5783.

Colmenarejo, G. (2020). Machine learning models to predict childhood and adolescent obesity: a review. *Nutrients, 12*(8), 2466.

Dang, J., Cai, S., Zhong, P., Liu, Y., Shi, D., Chen, Z., . . . Song, Y. (2024). Associations of school physical activity environment with comorbid obesity and myopia in children and adolescents: Findings from a Chinese national follow-up study. *Scandinavian Journal of Medicine & Science in Sports, 34*(1), e14562.

Liang, Y., Zheng, X., & Zeng, D. D. (2019). A survey on big data-driven digital phenotyping of mental health. *Information Fusion, 52*, 290-307.

Loy, B. D., O'Connor, P. J., & Dishman, R. K. (2013). The effect of a single bout of exercise on energy and fatigue states: a systematic review and meta-analysis. *Fatigue: Biomedicine, Health & Behavior, 1*(4), 223-242.

Marsch, L. A. (2021). Digital health data-driven approaches to understand human behavior. *Neuropsychopharmacology, 46*(1), 191-196.

Ortega, F. B., Ruiz, J. R., Castillo, M. J., & Sjöström, M. (2008). Physical fitness in childhood and adolescence: a powerful marker of health. *International journal of obesity, 32*(1), 1-11.

Pajares, F. (2006). Self-efficacy during childhood and adolescence. *Self-efficacy beliefs of adolescents, 5*(1), 339-367.

PARRA, E., ARONE, A., AMADORI, S., MUCCI, F., PALERMO, S., & MARAZZITI, D. (2020). Impact of physical exercise on psychological well-being and psychiatric disorders. *Journal for ReAttach Therapy and Developmental Diversities, 3*(2), 24-39.

Provost, F., & Fawcett, T. (2013). Data science and its relationship to big data and data-driven decision making. *Big data, 1*(1), 51-59.

Robbins, L. B., Pender, N. J., Ronis, D. L., Kazanis, A. S., & Pis, M. B. (2004). Physical activity, self-efficacy, and perceived exertion among adolescents. *Research in Nursing & Health, 27*(6), 435-446.

Teixeira, P. J., Carraça, E. V., Markland, D., Silva, M. N., & Ryan, R. M. (2012). Exercise, physical activity, and self-determination theory: a systematic review. *International journal of behavioral nutrition and physical activity, 9*, 1-30.

Thorsen, S. V., Rugulies, R., Hjarsbech, P. U., & Bjorner, J. B. (2013). The predictive value of mental health for long-term sickness absence: the Major Depression Inventory (MDI) and the Mental Health Inventory (MHI-5) compared. *BMC medical research methodology, 13*, 1-7.