ISSN 2616-5775 Vol. 7, Issue 5: 222-228, DOI: 10.25236/AJCIS.2024.070529

A Study on the Influence of Mothers' Physical and Mental Indicators on Infants' Behavioral Characteristics and Sleep Quality Based on Decision Tree and Genetic Algorithm

Weiyi Zhang[1,*]

[1]*School of Mathematics and Physics, Southwest University of Science and Technology, Mianyang, China*
[*]*Corresponding author: zwy1063208702@outlook.com*

Abstract: *The purpose of this study was to investigate the effects of mothers' physical and mental health on infant development. Through mathematical modeling and data analysis, correlations were found between mothers' physical and psychological indicators and infants' behavioral characteristics and sleep quality. Correspondence analysis and statistical methods were used to determine the degree of correlation between the two. The relationship model between maternal indicators and infant behavioral characteristics was established, and the decision tree model was used to predict infant behavioral characteristics. Through integer optimization problems and genetic algorithms, methods were proposed to adjust the psychological indicators of mothers to improve infant behavioral characteristics and minimize treatment costs. The results of the study provide a scientific basis for developing intervention strategies and promoting the healthy growth of infants.*

Keywords: *Decision Tree, Genetic Algorithm, Integer Optimization*

1. Introduction

In the early stages, the physical and mental health of the mother is crucial to the development of the infant [1]. The physical and mental state of the mother is directly related to the development and overall health of the infant [2]. Particularly in the context of breastfeeding and emotional communication, the health status of the mother plays a key role in the nutritional intake, immune protection and emotional development of the infant. However, poor physical and mental states may affect the quality and availability of breast milk, as well as the quality of mother-infant emotional interactions, which in turn have a profound impact on infants' cognitive, emotional, and social skills.

In order to thoroughly investigate the effects of mothers' physical and mental health on infant development, this study used data processing, analysis and modeling tools to focus on the effects of mothers' physical and psychological indicators on infants' behavioral characteristics and sleep quality. Through data preprocessing, correlation analysis and modeling, we aim to reveal the association between mothers' indicators and infants' behavioral characteristics, and to provide a scientific basis for the development of future intervention strategies. In addition, we focus on the effects of mothers' psychological indicators on infants' behavioral characteristics, and use decision classification tree models and genetic algorithms to optimize the problem and explore how to change infants' behavioral characteristics by adjusting the mothers' psychological state, so as to maximize the healthy growth of infants. Through this research, we expect to provide new insights and methods to improve the overall development of infants and contribute to the promotion of healthy mother-infant relationships and family well-being.

2. Correlations analysis

In order to visualize the correspondence between each group of data, this paper uses SPSS as a tool to analyze the correspondence between each group of variables.

First, the data were categorized according to international standards, as shown in Table. 1.

After that, the correlation between the variables was quantitatively analyzed by using various

ISSN 2616-5775 Vol. 7, Issue 5: 222-228, DOI: 10.25236/AJCIS.2024.070529

calculation methods, which were categorized into the chi-square test and the calculation of Spearman correlation coefficient.

Table 1: Indicator data clusters

Mother's age	Early childbirth	< 23 age	Optimal fertility	$26 \sim 30$ age	Delayed childbirth	> 30 age
Gestation time	Prematurity	< 30 week	Preterm labor	$30 \sim 37$ week	Normal childbirth	$37 \sim 43$ week
CBTS	Normal	< 11 mark	Slight depression	> 13 mark	Severe depression	< 12 mark
EPDS	Normal	$11 \sim 13$ mark	Slight depression	$12 \sim 13$ mark	Severe depression	> 13 mark
HADS	Normal	< 8 mark	Slight depression	$8 \sim 11$ mark	Severe depression	≥ 13 mark

Spearman's correlation coefficient can be used to measure the correlation of fixed class data, and it can be regarded as Pearson's correlation coefficient between ranked variables [3]. If the correlation coefficient is defined in terms of ρ, for the sample with sample size n, n initial data are transformed into ranked data, and the specific formula is as follows:

$$\rho = \frac{\sum_i (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_i (x_i - \bar{x})^2 \sum_i (y_i - \bar{y})^2}} \tag{1}$$

The spearman correlation coefficients were calculated in SPSS software and the results were obtained as shown in Table. 2.

Table 2: Characterization of infant behavior

	Mother's age	Marital status	Education	Gestation period	Mode of delivery	CBTS	EPDS	HADS	Infant behavioral traits
Mother's age	1 (0.000^{***})	0.011 (0.833)	0.183 (0.000^{***})	−0.124 (0.016^{**})	0.039 (0.453)	−0.033 (0.522)	−0.087 (0.090^{*})	−0.069 (0.178)	0.087 (0.089^{*})
Marital status	0.011 (0.833)	1 (0.000^{**x})	0.11 (0.032^{**})	0.004 (0.946)	0.023 (0.661)	−0.008 (0.882)	−0.023 (0.662)	−0.066 (0.203)	−0.005 (0.919)
Education	0.183 (0.000^{***})	0.11 (0.032^{**})	1 (0.0000^{***})	0.026 (0.615)	−0.016 (0.760)	−0.072 (0.163)	−0.114 (0.026^{**})	−0.1 (0.051^{*})	−0.041 (0.426)
Gestation period	−0.124 (0.016^{**})	0.004 (0.946)	0.026 (0.615)	1 (0.000^{***})	−0.18 (0.000^{**x})	−0.077 (0.133)	−0.05 (0.336)	−0.106 (0.039^{**})	0.002 (0.972)
Mode of delivery	0.039 (0.453)	0.023 (0.661)	−0.016 (0.760)	−0.18 (0.000^{***})	1 (0.000^{***})	−0.021 (0.683)	−0.02 (0.700)	−0.066 (0.201)	−0.003 (0.956)
CBTS	−0.033 (0.522)	−0.008 (0.882)	−0.072 (0.163)	−0.077 (0.133)	−0.021 (0.683)	1 (0.000^{***})	0.781 (0.000^{***})	0.71 (0.000^{***})	−0.114 (0.027^{**})
EPDS	−0.087 (0.090^{*})	−0.023 (0.662)	−0.114 (0.026^{**})	−0.05 (0.336)	−0.02 (0.700)	0.781 (0.000^{***})	1 (0.000^{***})	0.784 (0.000^{***})	−0.132 (0.010^{***})
HADS	−0.069 (0.178)	−0.066 (0.203)	−0.1 (0.051^{*})	−0.106 (0.039^{**})	−0.066 (0.201)	0.71 (0.000^{***})	0.784 (0.000^{***})	1 (0.000^{***})	−0.123 (0.017^{**})
Infant behavioral traits	0.087 (0.089^{*})	−0.005 (0.919)	−0.041 (0.426)	0.002 (0.972)	−0.003 (0.956)	−0.114 (0.027^{**})	−0.132 (0.010^{***})	−0.123 (0.017^{**})	1 (0.000^{***})

As can be seen from Table. 2 above, infant behavioral traits have strong correlations with all three psychological indicators of mothers, while for physical indicators of mothers, the correlation is high only with mother's age, and the correlation with the rest of the physical indicators is not significant.

As shown in Table. 3 below, the sleep quality of infants was also strongly correlated with the three psychological indicators of mothers, while the physical indicators of mothers were weakly correlated with the age and gestation time of mothers, and the correlations with the rest of the physical indicators were not significant.

The chi-square test is a test used to measure the degree of deviation between the actual observed value and the theoretical inferred value of a statistical sample, the scope of application is categorical variables, this paper uses SPSS to carry out chi-square test on the physical and psychological indicators of mothers, and the obtained results are shown in Table. 4.

ISSN 2616-5775 Vol. 7, Issue 5: 222-228, DOI: 10.25236/AJCIS.2024.070529

Table 3: Analysis of infant sleep quality

	Mother's age	Marital status	Education	Gestation period	Mode of delivery	CBTS	EPDS	HADS	Sleep time	Awake times	Sleep way
Mother's age	1 (0.000^{***})	0.011 (0.833)	0.183 (0.000^{***})	−0.124 (0.016^{**})	0.039 (0.453)	−0.033 (0.522)	−0.087 (0.090^{*})	−0.069 (0.178)	0.017 (0.747)	0.051 (0.323)	−0.088 (0.085^{*})
Marital status	0.011 (0.833)	1 (0.000^{***})	0.11 (0.032^{**})	0.004 (0.946)	0.023 (0.661)	−0.008 (0.882)	−0.023 (0.662)	−0.066 (0.203)	0.019 (0.706)	−0.036 (0.488)	−0.052 (0.313)
Education	83 (0.000^{***x})	0.11 (0.032^{**})	1 (0.000^{***})	26 (0.615)	.016 (0.760)	−0.072 (0.163)	−0.114 (0.026^{**})	−0.1 (0.051^{*})	−0.058 (0.256)	0.07 (0.172)	0.036 (0.487)
Gestation period	−0.124 (0.016^{**})	0.004 (0.946)	0.026 (0.615)	1 (0.000^{***})	−0.18 (0.000^{***})	−0.077 (0.133)	−0.05 (0.336)	−0.106 (0.039^{**})	0.07 (0.176)	0.087 (0.089^{*})	0.04 (0.441)
Mode of delivery	0.039 (0.453)	0.023 (0.661)	−0.016 (0.760)	−0.18 (0.000^{***})	1 (0.000^{***})	−0.021 (0.683)	−0.02 (0.700)	−0.066 (0.201)	0.015 (0.766)	−0.063 (0.221)	0.063 (0.218)
CBTS	033 (0.522)	−0.008 (0.882)	−0.072 (0.163)	−0.077 (0.133)	−0.021 (0.683)	1 (0.000^{***})	0.781 (0.000^{**x})	0.71 (0.000^{***})	−0.128 (0.013^{**})	0.074 (0.150)	0.054 (0.295)
EPDS	87 (0.090^{*})	−0.023 (0.662)	−0.114 (0.026^{**})	−0.05 (0.336)	−0.02 (0.700)	0.781 (0.000^{***})	1 (0.000^{***})	0.784 ($0.000^{**\pi}$)	−0.173 (0.001^{***})	0.112 (0.029^{**})	0.009 (0.863)
HADS	69 (0.178)	−0.066 (0.203)	−0.1 (0.051^{*})	−0.106 (0.039^{**})	−0.066 (0.201)	0.71 (0.000^{***})	0.784 (0.000^{***})	1 (0.000^{***})	−0.122 (0.017^{**})	0.07 (0.172)	0.059 (0.248)
Sleep time	17 (0.747)	0.019 (0.706)	−0.058 (0.256)	0.07 (0.176)	0.015 (0.766)	−0.128 (0.013^{**})	−0.173 (0.001^{***})	−0.122 (0.017^{**})	1 (0.000^{***})	−0.318 (0.000^{***})	0.232 (0.000^{***})
Awake times	0.051 (0.323)	−0.036 (0.488)	0.07 (0.172)	0.087 (0.089^{*})	−0.063 (0.221)	0.074 (0.150)	0.112 (0.029^{**})	0.07 (0.172)	−0.318 (0.000^{***})	1 (0.000^{***})	−0.255 (0.000^{***})
Sleep way	−0.088 (0.085^{*})	−0.052 (0.313)	−0.036 (0.487)	0.04 (0.441)	0.063 (0.218)	0.054 (0.295)	0.009 (0.863)	0.059 (0.248)	0.232 (0.000^{***})	−0.255 (0.000^{***})	1 (0.000^{***})

Table 4: Results of chi-square test

Item	Status	X^2	P
CBTS	Light depression Severe depression Normal	0.047	1.000
EPDS	Light depression Severe depression Normal	6.766	0.149
HADS	Normal Light depression Severe depression	2.302	0.680
Mother's age	Optimal fertility Late childbearing Early childbearing	8.927	0.063^{*}
Weeks of gestation	Super premature Preterm labor Normal labor	3.387	0.495
Mode of delivery	Cesarean section Natural parturition	1.198	0.549

Looking at the p-value of the chi-square test, we can see that there is no significant difference between the above indicators and infant behavioral characteristics.

Comparing the results of the two quantitative methods, we find that there is a strong correlation between the infant's behavioral characteristics and the mother's psychological indicators, while the mother's physical indicators are strongly correlated with the mother's age, gestation time, and method of parturition.

Therefore, it is reasonable to assume that the mother's psychological and physical indicators have an effect on the infant's behavioral characteristics and sleep quality.

3. Prediction of infant behavioral characteristics

3.1 Modeling

First, we quantitatively coded the behavioral characteristics of infants according to Table. 5.

After obtaining the quantified data, since the scales of the three psychological indicators are the same, we utilize the entropy weighting method for the psychological indicators of mothers [4].

Table 5: Coding table

Type	Code
Quiet	3
Medium	2
Ambivalent	1

(1) Entropy weighting method to calculate weights

Set the value of the $i-th$ indicator as A_i, normalize the data,

$$A_i^* = \frac{A_i - \min\limits_k A_k}{\max\limits_k A_k - \min\limits_k A_k} \tag{2}$$

Calculate the probability matrix,

$$P_i = \frac{A_i^*}{\sum_{k=1}^{n} A_k^*} \tag{3}$$

Calculating Information Entropy, Entropy Vector, Information Utility Value,

$$e_i = -\frac{1}{\ln n}\sum_{k=1}^{n} P_k \ln(P_k) \tag{4}$$

$$d_i = 1 - e_i \tag{5}$$

Calculate the entropy weights of the three factors,

$$\omega_i = \frac{d_i}{\sum_{i=1}^{n} d_i} \tag{6}$$

Three entropy weights are calculated by substituting the data to obtain the weight vector.

$$\omega = [0.36,0.19,0.45]^T \tag{7}$$

The new psychological indicators of mothers' scores were obtained after treatment and were calculated as follows.

$$Psychological\ scores = 0.45 \times CBTS + 0.36 \times EPDS + 0.19 \times HADS \tag{8}$$

(2) Decision tree model

Decision tree is a classification algorithm based on the tree structure to make decisions, it is a top-down, sample data tree classification process, which consists of nodes and directed edges [5]. It consists of nodes and directs edges, where each internal node represents a feature or attribute, and the leaf nodes represent the categories, and a feature is selected from the features of the training data as the splitting criterion of the current node during the feature selection. To solve a classification problem with a decision tree, starting from the root node, a feature of the instance is tested, and according to the test result, the instance is taken as the corresponding child node, at this time, each child node corresponds to a value of the feature, recursively test the instance and assign it until it reaches the leaf node, and finally, the instance is divided into the class of the leaf node. The most common decision tree algorithms are ID3, C4.5, Classification and Regression Tree (CART), of which ID3 can solve the discrete attribute samples, from the given Annex 1 can be understood that the infant behavioral characteristics of the data for the discrete type of data, so in the solution of Problem 2, the choice of ID3 to solve the problem of discrete attribute samples classification.

The ID3 algorithm is based on information theory and takes information entropy and information gain as the criteria to realize the inductive classification of data. It is based on Occam's razor: the smaller the decision tree, the better the decision tree (besimple simple theory). The core idea of ID3 algorithm is to use the information gain to measure the attribute selection, to find a certain attribute has the largest information gain to the target attribute, so that the entropy of the research target decreases the most and is used as the attribute of the current decision tree node.

(3) Information entropy

Information entropy is used to measure the expected value of the occurrence of a random variable [6]. If the uncertainty of information is greater, the value of entropy is also greater, and the occurrence of various situations will be more. Information entropy is related to the probability distribution of an event, the more uniform the probability distribution, the greater the information entropy. When all probabilities are equal, the information entropy is maximum. For the sample set D, the number of categories is K. The empirical entropy of the data set D is expressed as follows.

ISSN 2616-5775 Vol. 7, Issue 5: 222-228, DOI: 10.25236/AJCIS.2024.070529

$$H(D) = -\sum_{k=1}^{k} \frac{|C_k|}{|D|} \log_2 \frac{|C_k|}{|D|} \tag{9}$$

Where the empirical conditional entropy $H(D \mid A)$ of a feature A for a dataset D is expressed as.

$$H(D \mid A) = \sum_{i=1}^{n} \frac{D_i}{D} \left(-\sum_{k=1}^{k} \frac{|D_{ik}|}{|D|} \log_2 \frac{|D_{ik}|}{|D|} \right) \tag{10}$$

Information Gain (IG): It is used to measure the ability of a feature to distinguish between samples, the higher the information gain of a feature, the more concise the tree built with that feature as a node. Therefore, the attribute with the highest information gain in ID3 will be used as a categorizing attribute. The information gain can then be expressed as the difference between.

$$IG(X) = H(c) - H(c \mid X) \tag{11}$$

The solution using the decision tree ID3 algorithm can be roughly divided into the following three steps:

(1) Initialize the attribute set and the data set, calculate the information entropy S of the data set and all the attributes.

(2) Select the attribute with the largest information gain as the test attribute, divide the samples with the same value of the test attribute into a sub-sample set. If the category attributes of the sample set contain only a single attribute, then branch into leaf nodes, determine the value of the attribute and label it with the corresponding symbols, and then return to the calling place, or else call the algorithm recursively on the sub-sample set.

(3) Finally output a decision tree, complete; to avoid overfitting, will be pruned to reduce the structure and size of the tree.

3.2 Model solving

The model is solved using the processed data, and the accuracy of the model at different depths is shown in Figure 1.

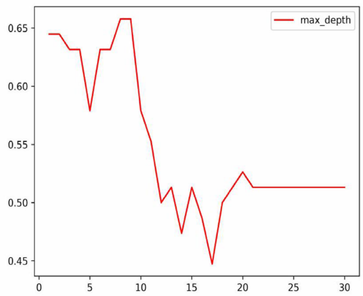

Figure 1: Accuracy graph

Depth 9 was chosen for practicality and usefulness of the model, with an accuracy of 65.78947%.

The obtained predictions are summarized in Table. 6 below.

Table 6: Projected results

No.	391	392	393	394	395	No.	391
Type	Medium	Medium	Medium	Quiet	Medium	Type	Medium
No.	396	397	398	399	400	No.	396
Type	Medium	Medium	Medium	Medium	Medium	Type	Medium
No.	401	402	403	404	405	No.	401
Type	Medium	Medium	Medium	Quiet	Medium	Type	Medium

4. Treatment options and costs

4.1 Model construction

First for CBTS, EPDS, and HADS, the rate of change of treatment costs relative to the degree of prevalence is all proportional to the treatment costs, and we list the following differential equation.

ISSN 2616-5775 Vol. 7, Issue 5: 222-228, DOI: 10.25236/AJCIS.2024.070529

$$\frac{dx}{dt} = kx \tag{12}$$

Where x represents the cost of treatment and t represents the degree of illness. Solving the differential equation yields.

$$x = ce^{kt} \tag{13}$$

Taking into account the relationship between the cost of the given treatment and the degree of illness, the unknown parameter of the above equation is solved as follows.

$$\begin{cases} c_1 = 200 \\ c_2 = 500 \\ c_3 = 300 \\ k_1 = 0.38266 \\ k_2 = 0.28875 \\ k_3 = 0.32395 \end{cases} \tag{14}$$

The final equation for the relationship between the cost of treatment and the degree of illness is as follows.

$$\begin{cases} x_1 = 200e^{0.38t_C} \\ x_2 = 500e^{0.29t_E} \\ x_3 = 300e^{0.32t_H} \end{cases} \tag{15}$$

Where x_1, x_2, x_3 represent the cost of treatment for CBTS, EPDS, and HADS, respectively. The cost of indirectly influencing the infant's behavioral traits by reducing the mother's three psychological indicators is defined as.

$$ce^{kt} - ce^{kt-\Delta t} \tag{16}$$

Thus, our objective function is.

$$\min\left[200e^{0.38t_{C_0}} - 200e^{0.38(t_{C_0}-\Delta t_C)} + 500e^{0.29t_{E_0}} - 500e^{0.29(t_{E_0}-\Delta t_E)} + 300e^{0.32t_{H_0}} - 300e^{0.32(t_{H_0}-\Delta t_H)}\right] \tag{17}$$

Since in the previous section of the model we combined the three psychometric indicators and defined a new psychometric score, we know that the value of this item indirectly affects the measure of infant behavioral characteristics as shown in Table. 7.

Table 7: Measures of Infant Behavioral Characteristics

Psychological indicator scores (A)	Infant behavioral characteristics
$0 \le A \le 3.538$	Quiet
$3.538 < A \le 8.915$	Medium
$8.915 < A$	Ambivalent

The psychometric score of mothers 238 is 18.09. Therefore, we have developed an integer optimization model with constraints as follows.

$$\min\left[200e^{0.38t_{C_0}} - 200e^{0.38(t_{C_0}-\Delta t_C)} + 500e^{0.29t_{E_0}} - 500e^{0.29(t_{E_0}-\Delta t_E)} + 300e^{0.32t_{H_0}} - 300e^{0.32(t_{H_0}-\Delta t_H)}\right] \tag{18}$$

$$\text{s.t.} \begin{cases} A_1 \le 0.45\Delta t_C + 0.36\Delta t_E + 0.19\Delta t_H \le A_2 \\ 0 \le \Delta t_C \le t_{C_0} \\ 0 \le \Delta t_E \le t_{E_0} \\ 0 \le \Delta t_H \le t_{H_0} \end{cases} \tag{19}$$

Where, A_i represents the upper and lower bounds of different intervals of psychological index scores, $t_{C_0}, t_{E_0}, t_{H_0}, \Delta t_C, \Delta t_E, \Delta t_H$ denote the initial values of the CBTS, EPDS, and HADS scores and the amount of changes in the scores, respectively.

4.2 Model solving

A genetic algorithm is used to solve a constrained integer optimization model. Genetic algorithm (GA) is a global adaptive probabilistic search algorithm based on the principles of natural selection and genetic genetics, which draws on the natural selection mechanism of biological evolution and the genetic mechanism of genetic recombination and mutation in biological reproduction and evolution [7]. The main purpose of using genetic algorithm here is to avoid the situation that multiple AGVs go to the same storage node at the same time.

The results obtained using genetic algorithm are shown in Table. 8 and Table. 9.

Table 8: Treatment program I

Ambivalent	⇒	Medium
CBTS		$15 \longrightarrow 0$
EPDS		$22 \longrightarrow 22$
HADS		$18 \longrightarrow 5$
Treatment cost		1.5329×10^5

Table 9: Treatment program II

Ambivalent	⇒	Quiet
CBTS		$15 \longrightarrow 0$
EPDS		$22 \longrightarrow 0$
HADS		$18 \longrightarrow 18$
Treatment cost		3.5404×10^5

5. Conclusions

This study systematically examined the significant impact of mothers' physical and mental health on infant development. Through data processing, analysis and modeling, we found that mothers' physical and psychological indicators have a significant impact on infants' behavioral characteristics and sleep quality. Mothers' physical and mental health is directly related to the quality and supply of breast milk, and also affects the quality of mother-infant emotional interactions, which in turn has a profound impact on infants' cognitive, social and emotional development. The model developed reveals the relationship between mothers' physical and mental health and infants' development, providing a scientific basis for future intervention strategies. By adjusting the mother's psychological state, the infant's behavioral characteristics can be improved, and the overall development of the infant can be promoted. These findings are of great significance in promoting healthy mother-infant relationships and family well-being and provide a useful reference basis for building healthy family environments and promoting healthy infant growth.

References

[1] Newman L, Judd F, Olsson C A, et al. Early origins of mental disorder-risk factors in the perinatal and infant period [J]. BMC psychiatry, 2016, 16: 1-8.

[2] Eisenhower A S, Baker B L, Blacher J. Children's delayed development and behavior problems: Impact on mothers' perceived physical health across early childhood[J]. Social Science & Medicine, 2009, 68(1): 89-99.

[3] Jiang J, Zhang X, Yuan Z. Feature selection for classification with Spearman's rank correlation coefficient-based self-information in divergence-based fuzzy rough sets[J]. Expert Systems with Applications, 2024: 123633.

[4] Montirosso R, Riccardi B, Molteni E, et al. Infant's emotional variability associated to interactive stressful situation: A novel analysis approach with Sample Entropy and Lempel–Ziv Complexity[J]. Infant Behavior and Development, 2010, 33(3): 346-356.

[5] Fletcher S, Islam M Z. Decision tree classification with differential privacy: A survey[J]. ACM Computing Surveys (CSUR), 2019, 52(4): 1-33.

[6] Back A D, Angus D, Wiles J. Determining the number of samples required to estimate entropy in natural sequences [J]. IEEE Transactions on Information Theory, 2019, 65(7): 4345-4352.

[7] Verma V K, Kumar B. Genetic algorithm: an overview and its application[J]. International Journal of Advanced Studies in Computers, Science and Engineering, 2014, 3(2): 21.