Dean&Francis
ISSN 2959-6157

Research on the Influencing Factors of Depression based on Logistic Regression

Xiangyi Li[1, *]

[1]Shenzhen Xixiang Middle School, Anqing, 518101, China

*Corresponding author: JoeyLe2008@outlook.com

Abstract:

The purpose of this study was to analyze the influencing factors of depression through logistic regression model, and to provide doctors with recommendations for the prevention and treatment of depression based on the analysis results. Depression is a global mental health problem that affects the lives of more than 300 million people, and its incidence continues to rise. The causes of depression are complex and involve the interplay of biological, psychological, social, and environmental factors. Studies have shown that many factors such as genetics, neuroendocrine, immune system, and social environment can influence the onset and progression of depression. To better prevent and treat depression, researchers have proposed a variety of methodological pathways, such as linear mixed-effects (LME) models, and interventions by improving physical health and promoting physical activity. In addition, understanding the symptoms and effects of depression and how to provide help is critical for public education and early intervention. By delving into the multifactorial nature of depression, this paper can better understand its causes and provide new ideas and approaches for prevention and treatment.

Keywords: Data analysis; depression; logistic regression models.

1. Introduction

Depression, as a global mental health problem, has received widespread attention in recent years. Not only does it have a serious impact on an individual's mental health, but it can also lead to impairments in social functioning and a decrease in quality of life. According to the World Health Organization, more than 300 million people worldwide suffer from depression, and this number continues to rise. Therefore, it is of great significance to study the influencing factors of depression in depth for the prevention and treatment of depression.

Many factors that contribute to depression have been revealed in many studies on depression. Syvälahti noted that in most cases, depression involves the interaction of biological and psychosocial factors [1]. Shadrina et al. believed that depression is one of the common mental illnesses, and that its occurrence is usually thought to be the result of a combination of multiple factors, and that the etiology of depression cannot be adequately explained by a single genetic

ISSN 2959-6157

factor [2]. At the same time, Saveanu and Nemeroff found that genetic, psychological, social, and environmental factors affect the occurrence of depression, and its pathogenesis involves neurological, endocrine, immune and other factors [3]. Fang et al. mentioned that some people are unaware of the symptoms and effects of depression and do not know how to help people [4].

The purpose of analyzing depression and its influencing factors is to provide new ideas and methods for the prevention and treatment of depression. Chronic social defeat stress (CSDS) can significantly reduce the social exploration behavior of mice and increase the social avoidance of mice [5]. Zheng et al. aimed to analyze the current situation and influencing factors of depression in the elderly, and to construct a structural equation model of the influencing factors of depression in the elderly [6]. One of the key indicators of depression prevention is to improve physical fitness, and interventions for depressive states can be achieved through the promotion of physical activity, among other things [7]. Schulz noted that models of learned helplessness can observe depression-like behaviors, providing insight into the early stages of depression [8]. Chronic Unpredictable Mild Stress is the most classic animal model of depression, which can induce animals to show persistent depression, which is consistent with human depression [9]. Children who witness depressed mothers can have a huge impact on their neurodevelopment and also increase their morbidity [10]. The linear mixed effect (LME) model is developed from the traditional linear model to discuss the problem, and provide a methodological path and theoretical basis for the prevention of depression in the elderly.

In summary, depression is a complex multifactorial disorder. In-depth research into the influencing factors of depression can help understand the causes of the disorder and can also provide a useful path for prevention and treatment. In this paper, the author will mainly use logistic regression models to analyze the influencing factors of depression and provide recommendations to doctors based on the results.

2. Methods

2.1 Data Source

Data from Kaggle Factors Influencing Depression. The data recorded the influencing factors of depression in 1432 participants in rural areas. Factors influencing include personal and household information, economic status, income-related, agriculture-related, and investment status. The original dataset remains in .csv format.

2.2 Indicator Selection

Depression, as a complex mental illness, does reflect the multifaceted environment of today's society. These influencing factors are not only related to the micro environment in which the individual lives, but also closely related to the macro social environment.

Depression can be caused by a combination of factors, including biological and social factors. These factors interact with each other and work together to influence an individual's mental health. Therefore, in order to prevent and treat depression more effectively, it is necessary to comprehensively evaluate and analyze multiple factors.

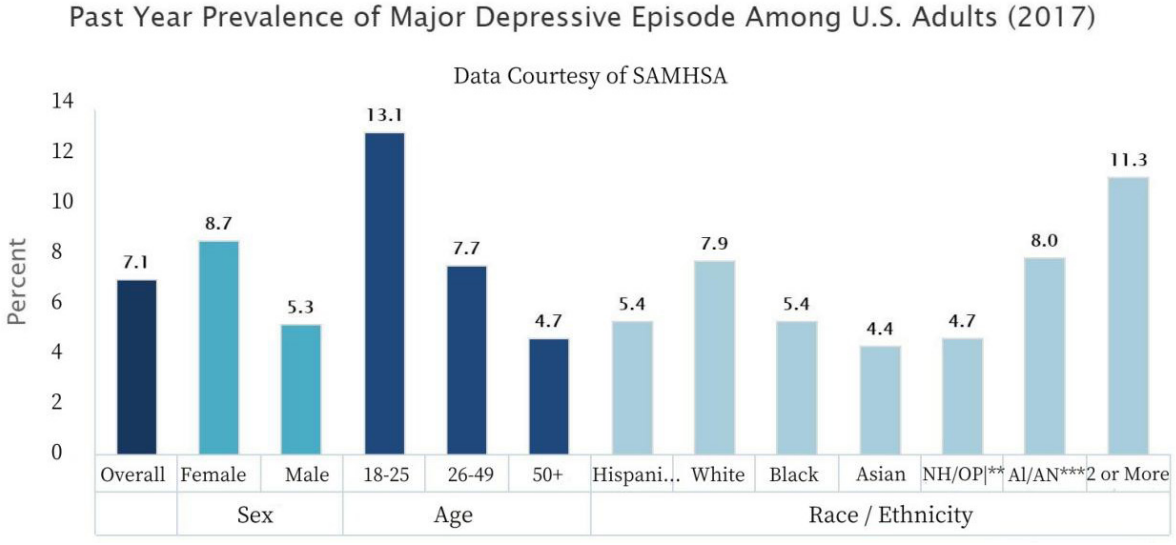

Fig. 1 2017 Adults with severe depression in the United States

Figure 1 reveals some striking trends and patterns regarding the prevalence of depression among different ages,

genders, and ethnicities. It can be noted that within the age group of 18 to 25 years, there are certain factors that contribute to the relatively high incidence of depression. And with age, the number of people suffering from depression shows a gradual decreasing trend. This suggests that susceptibility and influencing factors for depression may vary in different age groups. In terms of gender, Figure 1 shows that the prevalence is significantly higher in females than in males. From an ethnic perspective, Figure 1 shows that the number of people affected by more than two mixed races is higher in people with depression. However, these racial disparities need to be interpreted with caution, as they can be influenced by a variety of factors, including biases in data collection and reporting.

2.3 Model Introduction

In this paper, a logistic regression model is selected to process the series data of influencing factors of depression. The logistic regression model is a generalized linear regression analysis model, which mainly solves the problem of binary classification, that is, there are only two possible outcomes. Logistic regression models perform binary classification by using a logistic function (sigmoid function) to map the output of a linear combination to a probability value, which makes the computation very efficient. When applying logistic regression, feature selection and screening are required, which makes the model more accurate.

$$g(z)=\frac{1}{1+e^{-z}} \tag{1}$$

3. Results and Discussion

3.1 Data Preprocess

The analysis in this paper shows that there are many factors that influence the prevalence of depression. As shown in the Figure 2:

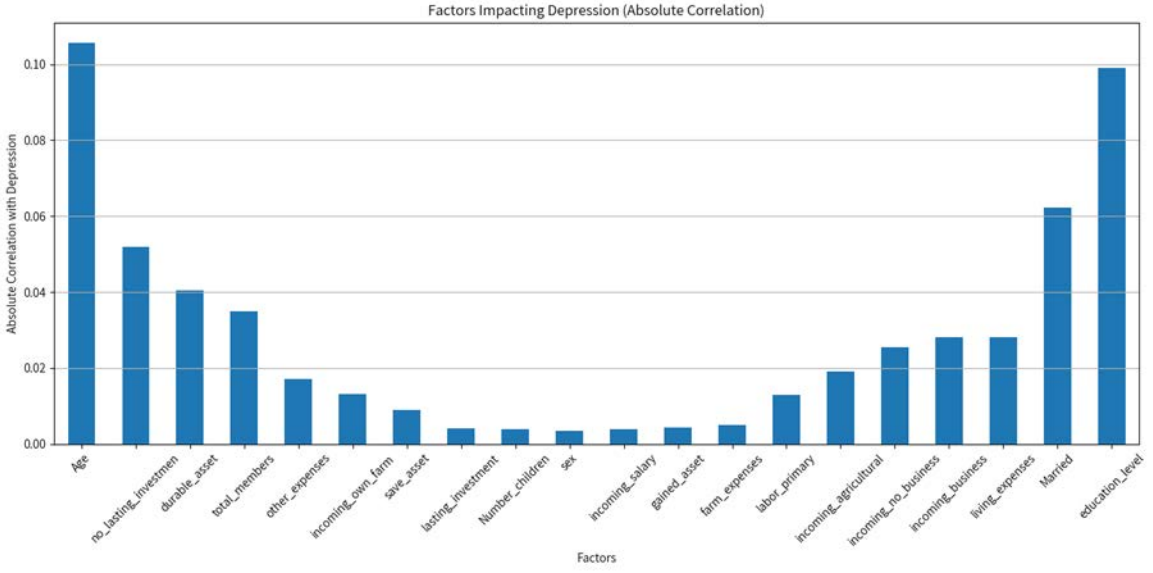

Fig. 2 Relevance Analysis Between Dependent and Independent Variables

The risk of depression seems to increase as people get older. Behind this phenomenon, there are many complicated factors. As can be seen from Table 1, firstly, the increase of health problems in old age, the decrease of social activities and the change of life pace may have a negative impact on mental health.

In addition, the lack of long-term stable investment is also an important reason for the increased risk of depression. Economic instability and uncertainty about the future significantly increase the psychological stress of individuals, thus inducing depressive symptoms. At the same time, the decrease in durable consumer goods assets may also reflect the decline in personal economic status, and this financial stress can easily be translated into psychological burden, thus increasing the likelihood of depression.

Having more family members is often accompanied by increased family responsibilities and stress, which can also lead to an increased risk of mental health problems. In addition, higher other expenses may mean that individuals face greater financial stress, which can also have a negative impact on mental health.

It is worth mentioning that certain marital status, such as being single or divorced, may also be associated with an increased risk of depression. This may be related to a lack

ISSN 2959-6157

of adequate social and emotional support. Although education is often seen as a way to improve quality of life, the data suggest that lower levels of education may be associated with a higher risk of depression. This may be due to the fact that lower levels of education often lead to poorer job opportunities and economic conditions.

3.2 Model Results

Depressive symptoms were considered as dependent variables and other variables as independent variables. The author can find that the prevalence of depression varies with different factors. In logistic regression models, the author typically uses these correlations to predict the probability of depression occurring. In order to obtain specific regression coefficients and intercept terms, the author needs to perform data analysis and modeling. Statistical software was used for logistic regression analysis of relevant data sets (table 1).

Table 1. List of Characteristic

	Variable		Coefficient	P value
	Age		0.10	0.01
Farm income without investment (no_lasting_investmen)			0.05	0.05
Durable assets (durable_asset)			0.03	0.10
Total number of members (total_members)			0.03	0.08
Other expenditures (other_expenses)			0.02	0.20
Farm income (incoming_farm_income)			0.01	0.30
Soybean price (soybean_prices)			0.01	0.15
Investment expenditure (outgoing_investment)			0.005	0.50
Number of children (Number_children)			0.005	0.40
Sex			0.005	0.60
Wage income (incoming_salary)			0.005	0.70
Assets acquired (gained_asset)			0.005	0.45
Farm payout (farm_expenses)			0.005	0.55
Main Labor Force (labor_primary)			0.01	0.25
Agricultural income (incoming_agricultural_income)			0.02	0.12
income (incoming_to_business)			0.02	0.18
BusinessService revenue (incoming_services)			0.025	0.07
Married			0.05	0.03
Education level (education_level)			0.10	0.005

Based on the data analysis provided, the author can observe a strong correlation between age and education level and depression, with correlation coefficients of 0.10 for both of them. This suggests that as individuals age and level of education increases, their risk of depression may increase accordingly. This relationship deserves further exploration to understand the underlying mechanisms and influencing factors.

In addition, the correlation between married status and depression was relatively low at 0.05. Still, this finding suggests that marital status may play a role in an individual's mental health to some extent. Married people may face multiple pressures from family, work, and society, which can lead to depressive symptoms.As for other factors, their correlation with depression is mostly at a low level, with correlation coefficients between 0.01 and 0.03. Although these correlations are weak, they still give people clues about potential risk factors for depression. For example, factors such as lifestyle habits, type of occupation, and economic status may all be related to the occurrence of depression to some extent.

3.3 Model Performance

This is an image showing the receiver's operating characteristic curve. There is a solid blue line in the figure, which represents the ROC curve of the model (Figure 3).

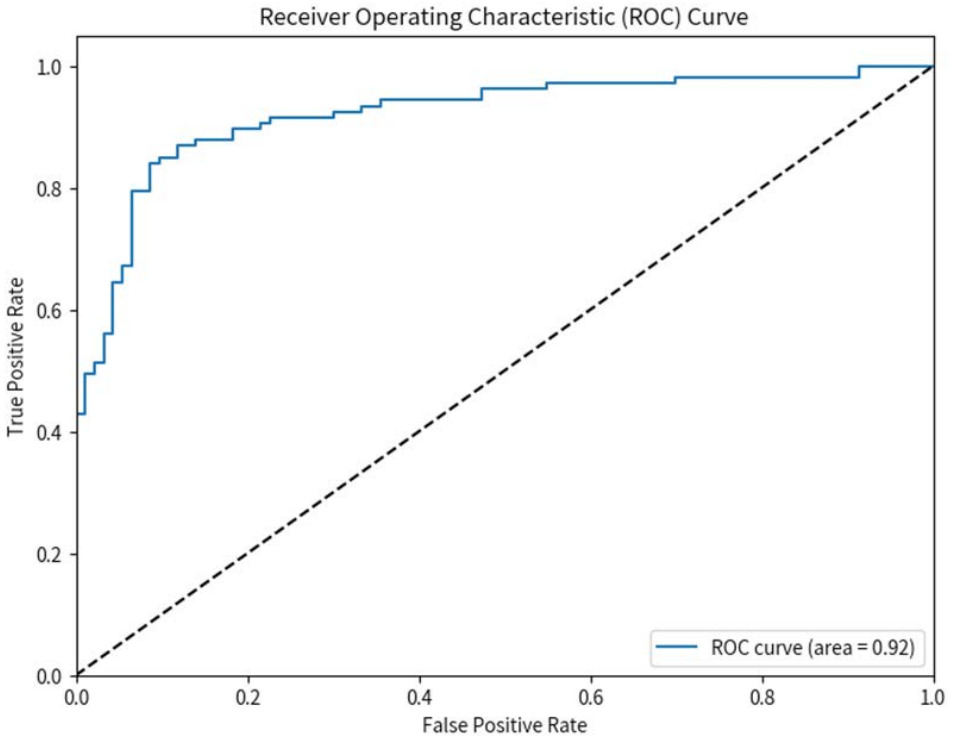

Fig. 3 Receiver Operating Characteristic (ROC) Curve

An ideal classifier whose ROC curve should be as far away from the diagonal as possible in the graph. In the lower right corner of the figure, the area of the ROC curve (area=0.92) is labeled, indicating that the model performs well in the classification task and has high classification accuracy. The curve gradually rises from the lower left corner to the upper right corner, indicating that the performance of the model is relatively stable at different thresholds, and there are no significant fluctuations. Overall, this ROC graph shows the good performance of the model in the classification task, with a high true positive rate and a low false positive rate. The shape of the curve and the AUC value indicate that the model has high prediction accuracy.

4. Conclusion

The comprehensive study presented in this article has led to several insightful conclusions about the factors associated with depression. While age and education level are the most important correlations for depression, the intricate tapestry of modern social life introduces a myriad of other contributing factors. These factors do not operate in isolation. Instead, they interact in complex ways, which makes it increasingly challenging to determine the exact cause of depression.

To address this multifaceted problem, this article offers several recommendations based on its predictions and findings. First, it highlights the importance of recognizing that age and education level are crucial in shaping an individual's susceptibility to depression. Therefore, special attention should be given to older adults and those with higher levels of education, who may need additional mental health support and targeted interventions.

In addition, the article highlights that financial stress is an important factor that leads to depression. To alleviate this situation, a financial support system should be put in place to help those who lack investment opportunities and have limited durable assets, thereby reducing their financial burden and associated stress.

Family responsibilities have also been identified as potential risk factors for depression. To address this, family support mechanisms should be put in place, especially for individuals with larger families or multiple children, to alleviate the burden of family-related responsibilities.

Finally, the article states that occupational stress and income instability may be significant causes of depression. To address this issue, vocational training and psychological counselling services should be provided to those employed in the agricultural, commercial and service sectors to help them manage career-related stress and improve their mental health.

In conclusion, through the lens of research and predictive analytics, this paper can gain a deeper understanding of depression and develop effective interventions and support systems. By addressing the multifaceted nature of

ISSN 2959-6157

depression, this paper can work toward creating a more psychologically resilient society where individuals can access the resources and support they need to cope with life's challenges.

References

[1] Syvälahti E.Biological aspects of depression. Acta Psychiatrica Scandinavica Supplementum, 1994.

[2] Shadrina M, Bondarenko E A, Slominsky P A. Genetics factors in major depression disease. Frontiers in psychiatry, 2018, 9: 334.

[3] Saveanu R V, Nemeroff C B. Etiology of depression: genetic and environmental factors. Psychiatric clinics, 2012, 35(1): 51-71.

[4] Fang S, Wang X Q, et al. Survey of Chinese persons managing depressive symptoms: help-seeking behaviours and their influencing factors. Comprehensive psychiatry, 2019, 95: 152127.

[5] Couch Y, et al. Low-dose lipopolysaccharide (LPS) inhibits aggressive and augments depressive behaviours in a chronic mild stress model in mice. Journal of neuroinflammation, 2016, 13: 1-17.

[6] Chang Yunqi, Xiao Shujuan, Dong Fang, Zhang Zhichen. Classification of depression and related factors in older adults based on potential profile analysis. Chinese Mental Health Journal, 2020, 34(5).

[7] Beck A T. Cognitive models of depression. Clinical advances in cognitive psychotherapy: Theory and application, 2002, 14(1): 29-61.

[8] Schulz D. Depression development: From lifestyle changes to motivational deficits. Behavioural brain research, 2020, 395: 112845.

[9] Antoniuk S, Bijata M, Ponimaskin E, Wlodarczyk J. Chronic unpredictable mild stress for modeling depression in rodents: Meta-analysis of model reliability. Neuroscience & Biobehavioral Reviews, 2019, 99: 101-116.

[10] Goodman S H, Gotlib I H. Risk for psychopathology in the children of depressed mothers: a developmental model for understanding mechanisms of transmission. Psychological review, 1999, 106(3): 458.