

大学生人格特质对抑郁水平的影响：基于随机森林模型分析

侯世玲, 肖放, 牛佳伟, 周燚森

西南交通大学心理研究与咨询中心, 四川 成都

收稿日期: 2025年2月8日; 录用日期: 2025年3月3日; 发布日期: 2025年3月17日

摘要

目的: 运用随机森林算法研究大学生人格特质对抑郁水平的影响。方法: 采用中国大五人格问卷极简版、抑郁自评量表对1436名大学生进行测量。使用回归模型分析大学生人格特质是否影响其抑郁水平, 通过R语言构建随机森林模型分析大学生人格特质对抑郁水平的重要性。结果: 回收得到有效问卷1368份, 问卷有效率为95.3%。回归模型显示, 大学生人格特质对其心理健康的影响有统计学意义。随机森林模型显示, 大学生人格特质对抑郁水平的影响按重要性排名分别是神经质、责任心、宜人性、开放性、外倾性。结论: 人格特质是影响大学生心理健康中的抑郁水平的重要因素。测量大学生的人格特质, 是识别和干预大学生抑郁水平的有效途径。

关键词

心理健康, 抑郁, 随机森林模型, 大学生

The Influence of College Students' Personality Traits on Depression Levels: An Analysis Based on Random Forest Model

Shiling Hou, Fang Xiao, Jiawei Niu, Yisen Zhou

Psychological Research and Consulting Center of Southwest Jiaotong University, Chengdu Sichuan

Received: Feb. 8th, 2025; accepted: Mar. 3rd, 2025; published: Mar. 17th, 2025

Abstract

Objective: The random forest algorithm was used to explore the influence of personality traits on

文章引用: 侯世玲, 肖放, 牛佳伟, 周燚森(2025). 大学生人格特质对抑郁水平的影响:基于随机森林模型分析. *心理学进展*, 15(3), 262-269. DOI: 10.12677/ap.2025.153152

depression level of college students. **Methods:** 1436 college students were measured by the Chinese Big Five Personality Questionnaire Inventory Brief 15-item Version and Self-rating Depression Scale. Regression model was used to analyze whether college students' personality traits affected their depression level, and R language was used to construct a random forest model to analyze the importance of college students' personality traits to depression level. **Results:** 1368 valid questionnaires were collected, the effective rate was 95.3%. The regression model shows that the influence of college students' personality traits on their mental health has statistical significance. Random Forest model shows that the influence of personality traits on depression level of college students is neuroticism, conscientiousness, agreeableness, openness and extraversion. **Conclusion:** Personality trait is an important factor affecting the level of depression in college students' mental health. Measuring the personality traits of college students is an effective way to identify and intervene in their levels of depression.

Keywords

Mental Health, Depression, Random Forest Algorithm, College Students

Copyright © 2025 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

抑郁症作为一种常见的心理障碍,不仅影响着个体的身心健康,也对家庭、社会产生了深远的影响。抑郁症的主要特征表现为情绪低落、兴趣缺乏、失去动力等,严重影响患者的日常生活和社交功能(张琪等, 2017)。大学生患抑郁症的比例明显高于普通人群(Ibrahim et al., 2013)。抑郁症是大学生最常见的心理问题之一,研究显示,在大学学习期间, 25% 的学生经历过抑郁症(Sheldon et al., 2021)。在一项元分析研究中显示,中国大学生抑郁症状的总检出率为 24.71%(王蜜源等, 2020)。大学生的心理健康工作,是大学教育开展的重要一环,影响未来社会的发展和进步。

人格特质是指个人在不同情境下表现出的相对稳定而持久的特性,是构成一个人的内在情感和外在行为表现的统合模式。大学生属于埃里克森人格发展阶段的中后期,其人格特质具有一定的稳定性。关于大五人格的研究指出,人格的这些维度与抑郁症存在相关,研究人员倾向于认为其中相关性最强的是神经质(De Moor et al., 2015)。

人格特质是影响心理健康的重要因素之一(Buelow & Cayton, 2020)。个体的人格特质会影响到我们心理和行为的各个方面(Chamorro-Premuzic & Furnham, 2008; Furnham & Chamorro-Premuzic, 2004: p. 2008),有研究指出人格特质和抑郁症复发有关(李戛等, 2019)。人格特质对个体抑郁症状可以起到有效的预测作用,抑郁与大五人格的开放性呈显著的负相关(Gong et al., 2020)。

2. 研究对象与方法

2.1. 研究对象

2024 年 12 月,本研究采用方便抽样的方法,通过问卷星和纸质问卷相结合的方式在西南地区某大学发放调查问卷。在调查之前设置知情同意书,被试确认并同意后进行测试,发放问卷共 1436 份,剔除明显乱答、漏答的问卷,最终回收有效问卷 1368 份,有效率为 95.3%,调查了西南地区某大学的大学生 1368 人。

2.2. 调查工具

2.2.1. 中国大五人格问卷极简版(Chinese Big Five Personality Inventory Brief 15-Item Version, CBF-PI-15)

本研究采用王孟成编制的中国大五人格量表。由于西方大五人格量表在中国文化情景下适用性不佳, (王孟成等, 2011)等在中国文化背景下编制了中国大五人格量表(CBF-PI), 在简化后编制了一个更加简短的 CBF-PI-B 的 15 个项目的版本。包含 5 个维度: 神经质、外向性、开放性、宜人性和责任性, 每个维度 3 道题, 采用五点计分。该量表五个维度的 Cronbach’s α 系数为 0.611~0.803。

2.2.2. 抑郁自评量表(Self-Rating Depression Scale, SDS)

本研究选用了 Zung 于 1965 年设计的抑郁自评量表, 适用于具有抑郁症状的成年人, 是我国目前使用范围最广的抑郁自评量表之一。该量表共分为 20 个条目, 使用 Likert4 点计分, 量表的内部一致性系数 Cronbach’s α 系数为 0.862。判定标准按照中国常模进行, 标准总分在 53~62 分之间为轻度抑郁, 63~72 分之间为中度抑郁, 72 分以上为重度抑郁。

2.3. 统计方法

使用 SPSS27.0 统计软件对数据进行相关分析、回归分析。使用 R4.4.1 软件进行统计分析和随机森林分析。本研究使用最小二乘法作为回归分析模型, 以抑郁自评量表得分为被解释变量, 以大五人格特质各维度得分为解释变量, 性别、生源地等为控制变量进行回归分析, 考察大五人格特质 5 个维度: 神经质、外向性、开放性、宜人性和责任心与抑郁水平之间的关系。计算随机森林方差解释率, 进而得出随机森林对各变量的重要性排序。

3. 结果

3.1. 关于大学生人口学变量、抑郁的描述性分析结果

调查结果显示, 在有效回收问卷 1368 份中, 男生 915 (66.89%)人, 女生 453 (33.11%)人, 城镇学生 976 (71.35%)人, 农村学生 392 (28.65%)人。汉族 1285 (93.92%)人, 少数民族(6.08%)人。19.96%的大学生抑郁得分处于 53~62 分之间, 即存在轻度抑郁情绪; 2.78%的大学生抑郁得分处于 63~72 分之间, 即存在中度抑郁情绪; 0.58%的大学生抑郁得分大于 72 分, 即存在重度抑郁情绪。详见表 1。变量赋值详见表 2。

Table 1. Survey results on demographic variables and depression levels among college students
表 1. 大学生人口学变量、抑郁水平调查结果

项目	类别	例数	占比(%)
性别	男	915	66.89
	女	453	33.11
生源地	城镇	976	71.35
	农村	392	28.65
民族	汉族	1285	93.92
	少数民族	83	6.08
抑郁水平(分)	轻度抑郁(53~62)	273	19.96
	中度抑郁(63~72)	38	2.78
	重度抑郁(>72)	8	0.58

Table 2. Variable assignment encoding
表 2. 变量赋值编码

变量	赋值
性别	0 = 男, 1 = 女
民族	0 = 少数民族, 1 = 汉族
生源地	0 = 农村, 1 = 城镇

3.2. 大学生人格特质和抑郁的相关/回归

本研究以抑郁自评量表得分为被解释变量,以大五人格特质各维度得分为解释变量,性别、生源地、民族为控制变量进行回归分析。结果显示,大五人格特质对抑郁水平呈现显著影响,其中神经质($\beta = 0.442$, $t = 19.984$, $P < 0.001$)、责任心($\beta = -0.200$, $t = -9.283$, $P < 0.001$)、外倾性($\beta = -0.087$, $t = -3.851$, $P < 0.001$)、宜人性($\beta = -0.169$, $t = -7.759$, $P < 0.001$)、责任心($\beta = -0.091$, $t = -4.122$, $P < 0.001$)。性别($\beta = 0.014$, $t = 0.699$, $P = 0.485$)、民族($\beta = 0.011$, $t = 0.554$, $P = 0.580$)、生源地($\beta = -0.025$, $t = -1.236$, $P = 0.217$)对大学生抑郁水平无显著影响。回归分析结果见表 3。

Table 3. Regression analysis of influencing factors of depression levels among college students
表 3. 大学生抑郁水平影响因素的回归分析

变量	B	SE	β	t	P
性别					
男*	0.293	0.420	0.014	0.699	0.485
女					
民族					
少数民族*	0.458	0.827	0.011	0.554	0.580
汉族					
生源地					
城镇*	-0.543	0.439	-0.025	-1.236	0.217
农村					
神经质	1.547	0.077	0.442	19.984	<0.001
责任心	-0.962	0.104	-0.200	-9.283	<0.001
外倾性	-0.302	0.078	-0.087	-3.851	<0.001
宜人性	-0.799	0.103	-0.169	-7.759	<0.001
开放性	-0.357	0.087	-0.091	-4.122	<0.001

注: *为对照组。

3.3. 大学生抑郁水平影响因素的重要性分析

随机森林(Random Forest)算法是一种集成学习方法(Breiman, 2001),它通过构建多个决策树并输出平均结果来降低模型的方差和误差,从而有效提高预测的准确性,适用于处理回归和分类问题(Pargent et al., 2023)。在本研究中,我们设定随机种子为 123,并从有效样本中随机抽取了 1000 份样本构建训练集,而将剩余的 368 份样本用作测试集。基于这 1000 棵树构成的随机森林模型,我们计算得出模型的平均均方

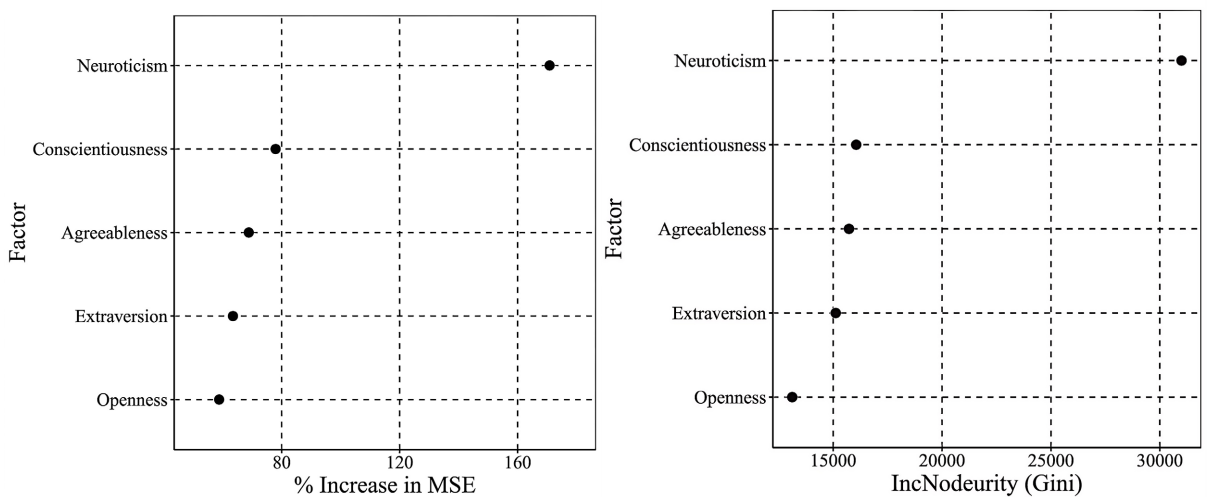
误差(Mean Squared Error, MSE)为 0.468, 表明模型具有较好的预测性能。

大五人格特质模型是心理学中用来描述人类个性差异的一个广泛接受的框架, 它被认为是人格结构的一个全面和有效的描述。通过将这些人格特质作为解释变量, 我们旨在探究它们与大学生抑郁水平之间的关系。

对大学生抑郁水平随机森林模型的解释变量的重要性进行排序。我们可以通过模型的输出结果来评估各个解释变量对抑郁水平的影响程度。具体来说, 模型利用均方误差的增加量来衡量每个变量对抑郁水平方差的解释能力(Yarkoni & Westfall, 2017)。这种方法可以帮助我们理解在模型中哪些变量对抑郁水平的预测更为关键。根据 MSE 增加量的分析结果, 我们得到了以下各变量对抑郁水平方差解释率的排名, 从高到低依次为: 神经质、责任心、宜人性、外倾性和开放性。

进一步地, 我们还可以通过节点纯度(node purity)来衡量各变量在模型中对抑郁水平预测的相对重要性。节点纯度是指在随机森林模型中, 一个变量在分裂节点时能够减少多少不纯度(Fife & D'Onofrio, 2022)。一个变量的节点纯度越高, 说明它在模型中的作用越大, 对抑郁水平的预测贡献也越大。根据节点纯度的分析结果, 我们得到了各变量对抑郁水平相对重要性的排名, 从高到低依次为: 神经质、责任心、宜人性、外倾性和开放性。

综合 MSE 增加量和节点纯度的分析结果, 我们可以得出结论: 在评估大学生抑郁水平的随机森林模型中, 神经质和责任心是最重要的两个解释变量, 它们对抑郁水平的预测贡献最大(夏学仓等, 2024)。紧随其后的是宜人性, 它在模型中也扮演着重要的角色。外倾性和开放性虽然也有一定的影响, 但相较于前三者, 它们的重要性相对较低。见图 1。



注: Increase in MSE 为均方误差增加量; IncNodeurity (Gini)为节点纯度, Neuroticism 为神经质, Conscientiousness 为责任心, Agreeableness 为宜人性, Extraversion 为外倾性, Openness 为开放性。

Figure 1. Random forest model diagram of variance explanatory rate and relative importance
图 1. 方差解释率和相对重要性的随机森林模型图

在本研究中, 我们采用了一种 R 语言中的数据分析技术——随机森林算法中的 importance 函数, 来评估不同数据特征对模型预测结果的重要性。这个函数通过计算每个特征在模型中减少不纯度的能力, 为我们提供了一个量化的指标, 从而可以评估每个特征对模型预测的贡献度。这种方法特别适用于复杂模型, 因为它能够处理高维数据并识别出最具影响力的预测变量。

研究选择了大五人格特质模型中的五种参数, 分别是神经质、责任心、宜人性、外倾性和开放性。

通过运行 importance 函数,我们得到了变量重要性排序图,这个图表直观地展示了每个人格特质对抑郁水平预测的贡献程度。见图 2。结果显示,神经质在所有人格特质中对抑郁水平的预测最为重要,这与以往的研究结果相一致,即神经质较高的个体更有可能经历抑郁情绪。紧随其后的是责任心,这一特质在预测抑郁水平时也显示出了较高的重要性。宜人性、外倾性和开放性虽然在预测抑郁水平方面的重要性相对较低,但它们仍然对模型的预测能力有所贡献。

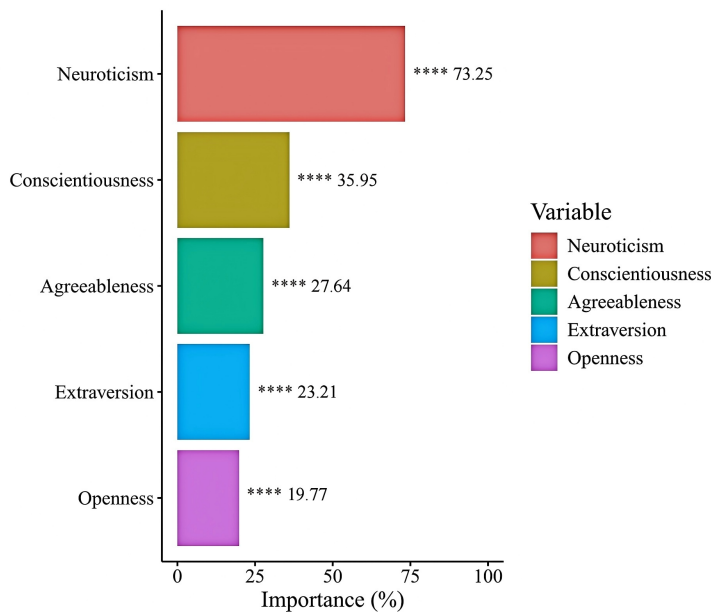


Figure 2. Feature importance ranking
图 2. 特征重要性排序

4. 讨论

本研究讨论了大学生人格特质对大学生心理健康的影响,筛查出抑郁倾向的大学生 319 名,占到总体水平的 23.32%。我们需要进一步探究这背后的原因,并积极寻求有效的解决措施。

随机森林的方差解释率和预测变量相对重要性程度的结果显示,神经质、责任心、宜人性、开放性、外倾性各维度得分对于预测大学生抑郁水平十分重要。不同人格特质可能会使大学生产生抑郁等不良情绪,从而影响大学生心理健康。

人格特质的各个维度都与抑郁有不同程度的关联。根据随机森林重要性作图,发现神经质在人格特质与抑郁关系中占据最为突出的地位。具有高神经质倾向的个体通常更容易感受到焦虑、沮丧等负面情绪的影响,并且对压力的反应更为强烈,情绪波动较大。这种性格特征导致他们在面对生活挑战和压力时,更容易陷入悲观情绪,从而增加了患抑郁症的可能性。因此,神经质被认为是一个重要的抑郁症预测因素,在抑郁症的预防和干预中应予以重视。责任心这个特质在对抑郁水平的预测中也显示出较高的重要性,过高的责任心可能导致个体承担过多的责任和义务,长期处于高压状态,难以放松和休息,从而增加抑郁风险。宜人性低的个体可能在人际关系中遇到挑战,缺乏同情心和合作精神,这可能导致社交孤立和人际关系紧张,从而增加抑郁风险(杨春等, 2025)。外倾性高的个体可能在社交活动中投入更多精力,但当社交活动未能带来预期的满足感时,可能会感到失望和孤独,增加抑郁风险。开放性较高的个体可能对新事物和变化更为敏感,这可能导致他们在面对不确定性和变化时感到焦虑和压力,增加抑郁风险。

总的来说,通过随机森林的应用,我们不仅得到了各变量对抑郁水平相对重要性的排名及其影响程度,而且还能够深入理解这些人格特质是如何影响大学生的抑郁水平的。这些发现为心理健康工作者提供了宝贵的信息,可以帮助更有效地识别和干预大学生中的抑郁问题。

本项研究揭示了神经质、责任心、宜人性、开放性、外倾性等人格特质在抑郁发生中的重要性占比。这些成果不仅丰富了我们对于抑郁症发病机制的理解,也为抑郁症的预防、识别和干预提供了新的思路(周莉等, 2024)。未来,我们可以继续探索如何通过干预人格特质来减少抑郁风险,并依据个体差异定制心理健康干预策略。同时,还需深入研究这些人格特质的相互影响及其在不同群体和文化背景下与抑郁的关联,以深化我们对人格特质与抑郁关系的理解。随着研究的持续推进,我们有望更有效地预防和治疗抑郁症,提升公众的心理健康水平,进而促进社会的稳定与进步。

基金项目

四川省哲学社会科学基地心理健康教育研究中心:数据挖掘技术在大学生心理健康测评数据中的应用研究(XLJKJY2301A)。

参考文献

- 李戛,位东涛,孙江洲,等(2019).人格和抑郁症:理论模型与行为-脑研究综述. *生理学报*, 71(1), 163-172.
- 王孟成,戴晓阳,姚树桥(2011).中国大五人格问卷的初步编制III:简式版的制定及信效度检验. *中国临床心理学杂志*, 19(4), 454-457.
- 王蜜源,韩芳芳,刘佳,黄凯琳,彭红叶,黄敏婷,赵振海(2020).大学生抑郁症状检出率及相关因素的 meta 分析. *中国心理卫生杂志*, 34(12), 1041-1047.
- 夏学仓,高小惠,房少洁,刘超(2024).基于随机森林模型分析共情心理对大学生心理健康的影响. *中国健康教育*, 40(4), 321-325.
- 杨春,吴德华,刘晓艺,陈洋,袁中静,尹华站(2025).研究生应对方式和心理素质在抑郁症状与压力源关系中的作用. *中国心理卫生杂志*, 39(2), 180-185.
- 张琪,吴任钢,郝树伟,徐震雷,官锐园(2017).大学生自尊在领悟社会支持和抑郁间的中介作用. *中国健康心理学杂志*, 25(11), 1683-1687.
- 周莉,王宏霞,耿靖宇,雷雳(2024).大学生网络受欺负与抑郁的关系:心理资本和同伴支持的调节作用. *心理科学*, 47(4), 981-989.
- Breiman, L. (2001). Random Forests. *Machine Learning*, 45, 5-32. <https://doi.org/10.1023/a:1010933404324>
- Buelow, M. T., & Cayton, C. (2020). Relationships between the Big Five Personality Characteristics and Performance on Behavioral Decision Making Tasks. *Personality and Individual Differences*, 160, Article 109931. <https://doi.org/10.1016/j.paid.2020.109931>
- Chamorro-Premuzic, T., & Furnham, A. (2008). Personality, Intelligence and Approaches to Learning as Predictors of Academic Performance. *Personality and Individual Differences*, 44, 1596-1603. <https://doi.org/10.1016/j.paid.2008.01.003>
- de Moor, M. H. M., van den Berg, S. M., Verweij, K. J. H., Krueger, R. F., Luciano, M., Arias Vasquez, A. et al. (2015). Meta-Analysis of Genome-Wide Association Studies for Neuroticism, and the Polygenic Association with Major Depressive Disorder. *JAMA Psychiatry*, 72, 642-650. <https://doi.org/10.1001/jamapsychiatry.2015.0554>
- Fife, D. A., & D'Onofrio, J. (2022). Common, Uncommon, and Novel Applications of Random Forest in Psychological Research. *Behavior Research Methods*, 55, 2447-2466. <https://doi.org/10.3758/s13428-022-01901-9>
- Furnham, A., & Chamorro-Premuzic, T. (2004). Personality and Intelligence as Predictors of Statistics Examination Grades. *Personality and Individual Differences*, 37, 943-955. <https://doi.org/10.1016/j.paid.2003.10.016>
- Gong, Y., Shi, J., Ding, H., Zhang, M., Kang, C., Wang, K. et al. (2020). Personality Traits and Depressive Symptoms: The Moderating and Mediating Effects of Resilience in Chinese Adolescents. *Journal of Affective Disorders*, 265, 611-617. <https://doi.org/10.1016/j.jad.2019.11.102>
- Ibrahim, A. K., Kelly, S. J., Adams, C. E., & Glazebrook, C. (2013). A Systematic Review of Studies of Depression Prevalence in University Students. *Journal of Psychiatric Research*, 47, 391-400. <https://doi.org/10.1016/j.jpsychires.2012.11.015>
- Pargent, F., Schoedel, R., & Stachl, C. (2023). Best Practices in Supervised Machine Learning: A Tutorial for Psychologists.

-
- Advances in Methods and Practices in Psychological Science*, 6. <https://doi.org/10.1177/25152459231162559>
- Sheldon, E., Simmonds-Buckley, M., Bone, C., Mascarenhas, T., Chan, N., Wincott, M. et al. (2021). Prevalence and Risk Factors for Mental Health Problems in University Undergraduate Students: A Systematic Review with Meta-Analysis. *Journal of Affective Disorders*, 287, 282-292. <https://doi.org/10.1016/j.jad.2021.03.054>
- Yarkoni, T., & Westfall, J. (2017). Choosing Prediction over Explanation in Psychology: Lessons from Machine Learning. *Perspectives on Psychological Science*, 12, 1100-1122. <https://doi.org/10.1177/1745691617693393>