

# 基于机器学习构建社会支持与大学生心理健康的预测模型

张瑾<sup>1\*</sup>, 汪晓攀<sup>2#</sup>, 靳雅鑫<sup>1</sup>, 谢永香<sup>2</sup>, 杨大航<sup>2</sup>, 凌兴厚<sup>1</sup>

<sup>1</sup>湖北中医药大学护理学院, 湖北 武汉

<sup>2</sup>湖北省中西医结合医院护理部, 湖北 武汉

收稿日期: 2024年9月15日; 录用日期: 2024年10月10日; 发布日期: 2024年10月21日

## 摘要

目的: 运用机器学习法构建社会支持对大学生心理健康影响的预测模型。方法: 采用方便抽样法, 选取2024年1月至2024年6月在武汉市高校就读的在校大学生开展调查。调查工具包括一般情况调查表以及社会支持量表和症状自评量表。使用多因素线性回归分析社会支持是否为大学生心理健康的影响因素, 构建随机森林模型计算模型的稳健性以及影响因素的权重。结果: 共回收问卷520份, 有效问卷503份, 问卷有效率为96.7%。回归模型结果表明, 社会支持( $t = -3.656, p = 0.000 < 0.01$ )等六项因素对大学生心理健康水平的影响具有统计学意义。随机森林模型的结果显示, 模型预测效果较好, 均方误差(MSE)为0.365。权重占比前三分别为社会支持、家庭的月收入以及是否吸烟。结论: 社会支持是影响大学生心理健康的重要因素。建议家长关注孩子的校园生活, 学校促进学生交流。定期开展心理辅导活动, 培养健康的生活方式, 减轻学业负担, 降低压力的影响。

## 关键词

心理健康, 社会支持, 机器学习, 随机森林模型, 大学生

# Construct a Predictive Model of Social Support and Mental Health of College Students Based on Machine Learning

Jin Zhang<sup>1\*</sup>, Xiaopan Wang<sup>2#</sup>, Yaxin Jin<sup>1</sup>, Yongxiang Xie<sup>2</sup>, Dahang Yang<sup>2</sup>, Xinghou Ling<sup>1</sup>

<sup>1</sup>School of Nursing, Hubei University of Chinese Medicine, Wuhan Hubei

<sup>2</sup>Department of Nursing, Hubei Hospital of Integrated Traditional Chinese and Western Medicine, Wuhan Hubei

Received: Sep. 15<sup>th</sup>, 2024; accepted: Oct. 10<sup>th</sup>, 2024; published: Oct. 21<sup>st</sup>, 2024

\*第一作者。

#通讯作者。

文章引用: 张瑾, 汪晓攀, 靳雅鑫, 谢永香, 杨大航, 凌兴厚. 基于机器学习构建社会支持与大学生心理健康的预测模型[J]. 护理学, 2024, 13(10): 1405-1412. DOI: 10.12677/ns.2024.1310201

## Abstract

**Objective:** To construct a predictive model of the impact of social support on college students' mental health by using machine learning. **Methods:** A convenient sampling method was used to select college students in Wuhan from January 2024 to June 2024. The survey tools included the general situation questionnaire, the social support scale and the symptom self-rating scale. The multifactor linear regression model was used to analyze whether social support was an influencing factor for college students' mental health, and the robustness of the model and the weight of the influencing factors were calculated by building a random forest model. **Results:** A total of 520 questionnaires were collected, 503 were valid, and the effective rate was 96.7%. The results of regression model showed that social support ( $t = -3.656, p = 0.000 < 0.01$ ) and other six factors had statistical significance on the mental health level of college students. The results of random forest model showed that the model predicted better, with a Mean Square Error (MSE) of 0.365. The top three weight shares were social support, monthly family income, and whether or not smoking. **Conclusion:** Social support is an important factor affecting college students' mental health. It is suggested that parents pay attention to their children's school life, and the school promotes student exchanges. Regular psychological counseling activities should be carried out to cultivate a healthy lifestyle, reduce the academic burden and reduce the impact of stress.

## Keywords

Mental Health, Social Support, Machine Learning, Random Forest Model, College Student

Copyright © 2024 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

健全的心理状态对于个人的全面发展而言,是不可或缺的先决条件和基础支撑,特别是在大学时期。据研究显示,大学生群体是心理疾病的易感人群[1]。学者汪立夏等人[2]对江西省多所高校进行了为期十年的调查研究,结果表明,约有20%的大学生存在不同程度的心理健康问题。学业压力[3]、人际关系[4]以及对未来就业前景的考量[5]等因素均会对大学生的心理健康产生影响。此外,部分大学生因背井离乡求学,置身于陌生的环境与人群之中,可能会体验到社会支持的减少,从而引发不同程度的心理问题[6]。社会支持指的是个体从社会结构或人际网络中获取的援助与支持,涵盖物质资源与精神资源两个方面[7]。目的在于协助个体应对生活中遭遇的难题与危机,确保其日常生活的平稳运行[8]。刘莉等人[9]的研究指出,社会支持能直接影响大学生的心理状态。武颖等研究者在其研究[10]中指出,社会支持程度较高的人群,其抑郁症的发病率相对较低。健全的社会支持体系对于大学生应对逆境、减轻心理负担以及对成年期发展产生深远影响具有重要作用[11][12]。当前,针对社会支持与大学生心理健康之间关系的研究,主要集中在社会支持作为幸福感知能力等变量的中介效应上。相比之下,直接比较社会支持与大学生心理健康之间相互作用的研究还较少。本研究旨在深入分析社会支持与大学生心理健康之间的关联,以期为提升大学生的心理健康水平提供理论依据和实践指导。

## 2. 研究对象与方法

### 2.1. 研究对象

选取2024年1月至2024年6月期间,在武汉市高校就读的全日制在校大学生作为研究对象。纳入

标准包括：① 在校大学生；② 年龄不得低于 18 岁；③ 有良好的沟通、认知和理解能力；④ 明确表示知情并同意参与本研究。

## 2.2. 研究方法

### 2.2.1. 抽样方法

采用方便抽样法，选取 2024 年 1 月至 2024 年 6 月在武汉市高校就读的在校大学生开展调查。

### 2.2.2. 调查工具

采用问卷调查法通过问卷星发放问卷进行匿名在线调查。研究工具包括：1) 一般资料调查表：包括性别、年级、居住城市、是否独生子女、学习成绩、年级、吸烟史、饮酒史、锻炼情况、家族精神疾病史、父母婚姻状况及家庭月收入等；2) SCL-90 症状自评量表[13]-[15]：由 90 个项目构成，其中又分为躯体化、强迫症状、人际关系敏感性、抑郁、焦虑、敌对、恐怖、偏执、精神病性以及其他附加因子等 10 个因子。在本次研究中，我们选取了强迫症状、人际关系敏感性(主要涉及个体的不适感和自卑感)、抑郁、偏执和焦虑这五个维度，共计 36 个问题。评分采用 Likert 五级量表法，评分范围从“没有”到“严重”，分别对应 1 至 5 分。当总均分  $\geq 2.5$  时，认为被试者表现出症状，即存在心理问题，且随着分数的升高，症状的明显程度亦会增加；3) 社会支持评定量表：该量表由肖水源等人[16]所编制，包含 10 个项目，涵盖客观支持、主观支持以及对支持的利用度三个维度。得分愈高，意味着社会支持状况愈佳。量表在各个维度上的内部一致性系数介于 0.89~0.94，重测信度为 0.92，能够全面反映个体所获得的社会支持状况。

### 2.2.3. 统计学方法

问卷回收后，利用 SPSS 25.0 与 Excel 软件对所收集的数据进行录入和整理。通过单因素分析，筛选潜在的影响因素，并对那些在统计学上具有显著性的变量执行多因素线性回归[17]-[20]，以确定可能的预测因子。最后，基于筛选出的预测因子，采用随机森林(Random Forest Algorithm, RF) [21] [22]机器学习的方法对模型进行评价并对变量权重进行排序。

## 3. 结果

### 3.1. 一般资料调查情况

通过问卷星平台，共收集到 520 份问卷。在剔除无效问卷后，最终获得有效问卷 503 份，问卷的有效回收率为 96.7%。其中，心理状况得分的总均分  $\geq 2.5$  的问卷 128 份，比例为 25.4%。男性占 16.3%，女性占 83.7%，独生子女占 25.84%，非独生子女占 74.16%。具体情况见表 1。

Table 1. General information questionnaire

表 1. 一般资料调查表

名称	选项	人数	百分比(%)
性别	女	421	83.7
	男	82	16.3
是否为独生子女	否	373	74.16
	是	130	25.84
居住地为几线城市	一线	151	30.02
	三线	99	19.68
	二线	73	14.51

续表

	五线	118	23.46
	四线	62	12.33
年级	大一	4	0.8
	大三	253	50.3
	大二	190	37.77
	大五	13	2.58
	大四	43	8.55
学习成绩	一般	382	75.94
	好	81	16.1
	差	40	7.95
是否吸烟	否	475	94.43
	是	28	5.57
是否喝酒	否	335	66.6
	是	168	33.4
平均每周锻炼时长	10 小时以上	14	2.78
	2 小时以下	311	61.83
	2~4 小时	111	22.07
	4~7 小时	48	9.54
	7~10 小时	19	3.78
精神疾病家族史	否	483	96.02
	是	20	3.98
家庭月收入	12000 以上	50	9.94
	3000~6000	168	33.4
	3000 以下	93	18.49
	6000~9000	113	22.47
	9000~12,000	79	15.71
家庭结构	其他	6	1.19
	单亲	47	9.34
	双亲	433	86.08
	重组	17	3.38
合计		503	100

### 3.2. 大学生心理健康影响因素逐步回归分析

将社会支持得分、性别、是否为独生子女、居住地为几线城市、年级、认为自己学习成绩如何、是否

吸烟、是否喝酒、每周花在体育锻炼的时长为多少小时、家人是否曾患有精神方面的疾病、家庭的月收入、家庭结构作为自变量，而将大学生心理健康状况作为因变量进行逐步回归分析，经过模型自动识别，最终余下社会支持( $t = -3.656, p = 0.000 < 0.01$ )、性别( $t = 3.146, p = 0.002 < 0.01$ )、认为自己学习成绩如何( $t = 4.564, p = 0.000 < 0.01$ )、是否吸烟( $t = 7.374, p = 0.000 < 0.01$ )、每周锻炼时长为多少小时( $t = -2.223, p = 0.027 < 0.05$ )、家庭的月收入( $t = 2.355, p = 0.019 < 0.05$ )。一共 6 项因素对大学生心理健康产生影响。R 方值为 0.197，模型通过 F 检验( $F = 20.324, p = 0.000 < 0.05$ )。如表 2 所示。

**Table 2.** Results of linear regression analysis

**表 2.** 线性回归分析结果

	非标准化系数		标准化系数	t	p
	B	标准误	Beta		
常数	0.738	0.250	-	2.950	0.003**
社会支持	-0.017	0.005	-0.151	-3.656	0.000**
性别	0.286	0.091	0.130	3.146	0.002**
学习成绩	0.244	0.053	0.185	4.564	0.000**
是否吸烟	1.090	0.148	0.309	7.374	0.000**
每周锻炼时长	-0.089	0.040	-0.093	-2.223	0.027*
家庭月收入	0.061	0.026	0.096	2.355	0.019*
R <sup>2</sup>			0.197		
调整 R <sup>2</sup>			0.188		
F			F(6496) = 20.324, p = 0.000		
D-W 值			0.331		

注：因变量 = 大学生心理健康状况评分。

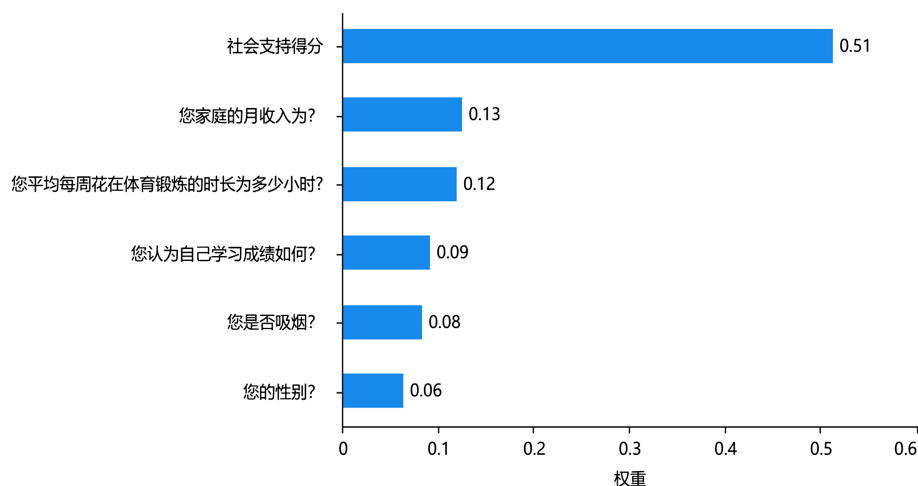
### 3.3. 基于随机森林分析大学生心理健康的影响因素

采取随机森林机器法：其是一种集成算法，包含多棵决策树的分类器，相对于单棵决策树来讲，随机森林算法会有更好的表现，并能有效防止过拟合现象。将社会支持、性别、认为自己学习成绩如何、是否吸烟、每周花在体育锻炼的时长为多少小时、家庭的月收入 6 项作为自变量，将大学生心理健康状况作为因变量进行随机森林建模。总共有 503 个样本参与分析。构建过程步骤：1) 选取 450 份样本作为训练样本，组成新的训练集；53 份样本作为测试集；2) 用新选取的样本进行决策树构造，在构造的时候，并不是使用全部特征而是采用部分特征，每次分裂时采用某种策略选择其中一个作为分裂属性；3) 重复执行第二步骤，直至分裂过程无法继续；4) 依次完成 1)~3) 的步骤，构建多棵决策树，这些决策树共同组成随机森林。本次研究的均方误差(Mean Squared Error, MSE)为 0.365，模型预测效果较好。如表 3 所示。

特征权重显示了各项对模型贡献的重要程度，其加和值为 1。由图 1 可知：社会支持所占比重为 45.04%，该特征的权重最高，对模型构建起着关键作用；家庭的月收入所占比重为 16.69%；是否吸烟所占比重为 12.46%；上述 3 项的比重总和占到了 74.19%；剩余 3 项认为自己学习成绩如何、每周花在体育锻炼的时长为多少小时、性别的比重则分别为 12.25%、7.77%、5.79%。

**Table 3.** Results of model evaluation**表 3.** 模型评估结果

指标	训练集时	测试集时
R 方值	0.814	0.421
平均绝对误差值 MAE	0.241	0.437
均方误差 MSE	0.122	0.365
均方误差根 RMSE	0.350	0.604

**Figure 1.** Weight ratio of influencing factors**图 1.** 影响因素权重占比

## 4. 讨论

本研究深入探讨了社会支持对大学生心理健康状态的影响。一项覆盖全国 12.6 万名在校大学生的调查显示，有高达 20.3% [23] 的学生群体面临着不同程度的心理问题。在本研究所针对的 503 名大学生样本中，发现存在心理健康问题的学生比例达到了 25.4%，凸显当前大学生心理健康问题的严峻。

本次调查结果显示，社会支持在大学生心理健康影响因素中占比权重最大，回归分析显示其与心理健康状况得分呈现负相关性。即社会支持得分越高，其所获得的社会支持越多，进而心理状况得分越低，心理越健康。前人研究表明[24]，社会支持程度较高的个体往往展现出较强的抗压能力和心理韧性，能够积极应对生活中的困难与挑战。相对而言，社会支持程度较低的个体常避免与父母进行沟通，缺乏紧密的同伴关系，未能构建起坚实的社会支持网络。在生活压力累积至一定程度后，他们可能难以承受，从而导致心理健康问题的出现。尤其是大学生迈入大学校园之际，新的生活环境和不同的学习方式使得他们更加需要人际交往。家长应当密切关注孩子的校园生活，及时给予他们来自家庭的关怀，以助其快速融入新环境。同时，学校亦应倡导学生重视与师长、同学之间建立并维护良好的人际关系，并促进他们与家人的交流互动。社会支持与人格特质之间也存在相关性[9]。积极自信的人与消极自卑的人在获得相同社会支持的前提下，前者感知到的社会支持大于后者，在面对困境时也更乐于接纳他人给予的支持。社会支持的增强有助于提升个体的自信水平，减轻负面情绪的影响[25]。因此，在大学校园环境，辅导员应当积极履行其职责，对性格内向的学生实施更为细致的关怀策略，以促进他们更快地适应校园生活，并增强其自信心，进而形成一个良好的循环。

随机森林模型及逐步回归显示, 学生家庭月收入与是否吸烟和学习成绩均为重要影响因素。经济困难对大学生心理的影响主要体现在自卑、焦虑等方面[26]。在经济压力下, 一些大学生可能会更加努力地学习, 将压力转化为动力, 但也有部分大学生可能因为无法积极面对经济困难而产生自卑感, 影响自信心和人际交往能力[27]。面对此类问题, 可以尝试针对经济困难的学生群体进行心理辅导来积极应对, 通过团体心理辅导, 提升家庭经济困难学生的宿舍人际关系处理技巧和适应能力。此外, 有研究表明, 80%的大学生是在进入大学后开始吸烟[28]。其中, 有的是出于好奇跟风模仿, 有的是由于不能适应学校新环境, 心情压抑而寻求发泄方式。应采取有效且合理的方法教育高校大学生中的吸烟者, 维护其心理健康。同时, 采取恰当的宣传策略, 引导他们树立正确的价值观, 缓解大学生活中的压力, 培养积极健康的生活方式。高中学习压力影响, 部分学生初入大学时学习动力减弱, 出现迟到、早退、旷课等现象, 将精力转向娱乐, 成绩下滑, 而大学阶段学习对未来就业影响大, 学生仍面临较大压力, 故容易产生焦虑心态[29]。可以定期实施心理健康教育讲座, 同时, 组织学生参与抗挫折能力训练, 以此加强学生心理资本的构建与培养, 确保学生心理健康水平的持续提升。

综上所述, 机器学习的应用可以加强模型的稳健性, 且能够对模型构建的权重进行划分, 让我们更好地了解到影响因素重要性的排序。由此可见, 逐步回归分析与随机森林机器学习算法的结合能够更加全面地分析数据。

基于本研究的结论, 建议如下: 首先, 大学生初入校园之际, 家长应密切关注孩子的校园生活, 并适时提供关怀。同时, 学校方面应促进学生与教师、同学及家庭成员之间的交流; 其次, 对于经济条件不佳的学生, 建议开展集体心理辅导活动, 以增强他们处理宿舍人际关系的能力和适应环境的能力; 再次, 学校应实施有效的宣传策略, 引导学生树立正确的价值观, 并培养积极健康的生活方式; 最后, 建议适当减轻大学生的学业负担, 以降低由此产生的学业压力对学生的不良影响。

## 基金项目

湖北省教育厅 2023 年度指导性项目(23G119)。

## 参考文献

- [1] Kessler, R.C., Amminger, G.P., Aguilar-Gaxiola, S., Alonso, J., Lee, S. and Üstün, T.B. (2007) Age of Onset of Mental Disorders: A Review of Recent Literature. *Current Opinion in Psychiatry*, **20**, 359-364. <https://doi.org/10.1097/ycp.0b013e32816ebc8c>
- [2] 汪立夏, 舒曼. 大学生心理健康状况、负性事件及求助行为趋势分析——基于江西省十年大学生心理健康状况调查[J]. *教育学术月刊*, 2013(5): 24-27.
- [3] 林斐, 连榕, 蓝瑞铭, 黄雅婧. 新冠疫情常态化防控期大学生适应障碍和抑郁对心理健康的影响[J]. *教育评论*, 2022(8): 92-96.
- [4] 刘美涓, 余华. 大学生 SC1-90 测试结果的研究[J]. *心理科学*, 1995, 18(5): 295-298.
- [5] 舒曼玲, 廖心仪, 秦露露. 后疫情时代长沙市在校大学生心理健康状况及影响因素[J]. *中国健康心理学杂志*, 2021, 29(11): 1712-1717.
- [6] 马伟娜, 林飞. 大学生社会支持、自我效能与心理健康的关系[J]. *中国临床心理学杂志*, 2006, 14(6): 641-643.
- [7] 孟捷, 孙剑涛, 冯洁, 等. 女性居民感知压力、睡眠质量、社会支持与抑郁症状的关系[J]. *中国健康教育*, 2024, 40(9): 798-802.
- [8] 蔡娜娜. 大学生心理健康档案的价值与应用研究[J]. *山西档案*, 2024(6): 181-184.
- [9] 刘莉, 毕晓慧, 王美芳. 社会支持与大学生主观幸福感的关系: 公证世界信念的中介作用[J]. *中国临床心理学杂志*, 2015, 23(4): 715-717, 721.
- [10] 武颖, 李炜迪, 王宇星, 等. 医学生抑郁与人格特质和社会支持的关系分析及预测模型建立[J]. *中国健康教育*, 2024, 40(5): 455-460.

- [11] 何安明, 包灿灿, 惠秋平. 大学生情绪智力与心理健康的关系: 社会支持的中介作用和手机依赖的调节作用[J]. 心理发展与教育, 2020, 36(4): 494-501.
- [12] 贺子京, 丁晶, 梁芷君, 等. 家庭环境、物质主义价值观对大学生心理健康的影响[J]. 中国健康心理学杂志, 2020, 28(12): 1887-1892.
- [13] 周华. 农业院校硕士研究生心理健康状况与对策建议[J]. 黑龙江高教研究, 2018, 36(7): 140-145.
- [14] 毕晨虹, 姚玉红, 陈增堂. 大学新生适应与其人格特质和家庭人际动力特征的关系[J]. 中国心理卫生杂志, 2015, 29(4): 316-320.
- [15] 毛琦, 沃尔夫冈·马斯特纳克, 余倩. 舞蹈治疗联合音乐治疗改善高校大学生心境障碍的干预研究——一种多样联合表达疗法的模式建立与疗效初探[J]. 北京舞蹈学院学报, 2022(4): 132-139.
- [16] 肖水源, 杨德森. 社会支持对身心健康的影响[J]. 中国心理卫生杂志, 1987, 1(4): 183-187.
- [17] Dao-de, S. (2000) Selection of the Linear Regression Model According to the Parameter Estimation. *Wuhan University Journal of Natural Sciences*, 5, 400-405. <https://doi.org/10.1007/bf02850764>
- [18] Cameron, A.C. and Trivedi, P.K. (2005) *Microeconometrics: Methods and Applications*. Cambridge University Press.
- [19] 威廉·H·格林. 经济计量分析[M]. 北京: 中国社会科学出版社, 1998.
- [20] 张厚粲, 徐建平. 现代心理与教育统计学[M]. 第3版. 北京: 北京师范大学出版社, 2009.
- [21] Ho, T.K. (1995) Random Decision Forests. *Proceedings of 3rd International Conference on Document Analysis and Recognition*, Montreal, 14-16 August 1995, 278-282. <https://doi.org/10.1109/ICDAR.1995.598994>
- [22] 武晓岩, 李康. 随机森林方法在基因表达数据分析中的应用及研究进展[J]. 中国卫生统计, 2009, 26(4): 437-440.
- [23] 曾洪发, 刘芳, 周紫微. 大学生练习中国传统导引术的心理体验与功效[J]. 体育科技文献通报, 2021, 29(7): 130-133, 183.
- [24] 刘焕青. 大学生领悟社会支持与负性情绪的研究[J]. 社会心理科学, 2012, 27(11): 39-42, 57.
- [25] 张薇薇, 张玉柱. 领悟社会支持与大学生抑郁: 生命意义感和自我控制的作用[J]. 内蒙古师范大学学报(自然科学版), 2021, 50(1): 82-87.
- [26] 田勇, 文甄, 马文涵, 等. 高职大学生体育锻炼及运动动机与心理健康的关联[J]. 中国学校卫生, 2024, 45(9): 1300-1303, 1308.
- [27] 常瑞敏, 尤桦. 基于心理普测结果的家庭经济困难大学生心理健康影响因子分析[J]. 才智, 2022(14): 86-88.
- [28] 周佳瑾. 高校大学生烟民心理健康教育研究[J]. 科学咨询, 2022(13): 115-117.
- [29] 尤丽吐孜·吐尔逊. 学业压力影响大学生心理健康问题的相关探讨[J]. 深圳青年, 2024(5): 128-130.