引文格式：费腾,付康钰,卞萌,等. 基于情绪计算的GIS硕士面试成绩预测[J]. 地理空间信息, 2024, 22(4): 128-132. doi:10. 3969/j.issn. 1672-4623.2024.04.031

基于情绪计算的GIS硕士面试成绩预测

费　腾[1]，付康钰[1]，卞　萌[2]，杨妙玲[1]
（1. 武汉大学 资源与环境科学学院，湖北 武汉 430079；
2. 武汉大学 遥感信息工程学院，湖北 武汉 430079）

摘　要：针对组织GIS专业学生面试带来人力资源与时间成本大的问题，本研究基于真实面试的视频数据和专家综合打分数据，通过人脸情绪识别与构建特征工程，使用包括偏最小二乘回归、支持向量回归、随机森林回归和梯度提升回归树的4种机器学习模型进行回归模型构建，探究情绪特征与面试成绩的关系。结果表明偏最小二乘回归模型的稳定性和泛化能力在本应用中更好，模型的决定系数为0.486。进一步分析发现，基本情绪中，我们定义的“负面情绪”与面试成绩呈显著负相关；而中性情绪、积极情绪与面试成绩呈显著正相关。本研究为GIS人才面试评估的自动化做出了有益的探索。

关键词：特征工程；情绪提取；机器学习；GIS专业
中图分类号：P208　　**文献标志码：**B　　**文章编号：**1672-4623（2024）04-0128-05

Interview Scores Prediction for GIS Master Based on Emotional Computing

FEI Teng[1], FU Kangyu[1], BIAN Meng[2], YANG Miaoling[1]
(1. School of Resource and Environmental Science, Wuhan University, Wuhan 430079, China;
2. School of Remote Sensing and Information Engineering, Wuhan University, Wuhan 430079, China)

Abstract: In response to the problem of organizing interviews for GIS specialties that bring large human resources and time costs, based on real interview video data and expert scoring data, we explored the association between emotional features and interview scores through face emotion recognition and construction feature engineering using four machine learning models including partial least squares regression, support vector regression, random forest regression and gradient boosting regression trees for regression model construction. The results show that the stability and generalization ability of partial least squares regression model are better in this application, and the determination coefficient of model is 0.486. Further analysis reveals that among the basic emotions, our definition of “negative emotions” is significantly negatively correlated with interview performance, while neutral emotions and positive emotions are significantly positively correlated with interview performance. This study can make a meaningful exploration for the automation of GIS talent interview assessment.

Key words: feature engineering, emotion extraction, machine learning, GIS specialty

近年来，我国普通高校中地理信息科学（geographic information science，GIS）相关专业人才培养的规模不断扩大，为适应地理信息行业发展需求，对GIS人才培养方法进行创新具有重要意义。GIS专业研究生选拔除了考查学生的专业基础知识，更重要的是考查学生是否具有科学研究的综合素质[1-2]。相对笔试而言，面试可以弥补笔试的不足，即在较短的时间内考查学生的语言组织能力，应变能力，表达能力等研究素质。传统面试需要多名经验丰富的考官在现场考核，人力资源与时间成本大。随着信息技术的发展，信息技术和学习管理系统在教学和评估中的兴起，为GIS人才面试的快速评估提供了一种选择[3]。

“价值-控制”理论认为学业成就情绪会影响学业行为与表现[4]，国内外学者对情绪与学生成绩表现的相关性进行了研究，结果表明情绪是有效学习和解决问题的一个关键组成部分，对学生成绩和表现是重要的影响因素[5]。此外，面试中的情绪表达可能对面试评分存在重要影响。研究表明，对求职者面试表现的主观印象是决定面试官评价的一个重要因素[6]，情绪的自我调节和表现会影响面试的结果，具有更高社交能力的个体善于传递情感和社交信息，从而在结构化面试中获得更好的评价[7]。因此在教学与面试中进行情绪的挖掘和成绩预测研究具有实践价值。

从目前的研究中发现机器学习算法在教学领域多

收稿日期：2022-10-27。
项目来源：国家自然科学基金资助项目（42271476）。
第一作者简介：费腾（1981—），博士，副教授，研究方向为城市地理大数据、生态遥感应用，E-mail：feiteng@whu.edu.cn。

用于学习成绩预测与考核任务的自动评分，但对于面试的评估研究较少。此外，在心理学、教育、医学等各个领域中，机器学习的方法已被证明能从教育数据中寻找规律以达到良好的预测效果[8]。因此，本文聚焦GIS领域人才选拔方法的创新，基于GIS专业硕士面试过程的视频数据，拟从受试者面部表情出发，训练并使用卷积神经网络实时识别人脸情绪，建立基于时间序列的情绪特征，进而采用4种机器学习模型分析情绪特征与面试成绩的关系，并通过比较模型精度指标，选择最优模型，了解在多大程度上面试者的面部情绪信息可被用来预测面试成绩，旨在推动GIS人才培养方法的改革创新。

1　实验流程和数据获取

1.1　研究对象

某双一流大学的某年的46名大四学生参与了本研究，有效被试为44名，其中女生23名，男生11名。学生中最大年龄为23岁，最小年龄为20岁，平均年龄为21.5岁，他们分别来自地理科学类专业、土地资源管理专业。

1.2　实验流程

本实验为选拔性综合面试，考核内容为英语能力与专业知识掌握情况。每位面试者面试时长为10~15 min。获取每位面试者固定时长的视频数据，在该时间内面试者将完成2项任务，分别是英语自我介绍、第一个GIS专业问题回答。

1.3　数据获取

为了避免不相关因素的干扰，面试视频数据采集经过严格控制，录制视频时在光线充足，明亮的环境，固定摄像头角度进行正面拍摄，为了有效捕捉受试者情绪变化，采集帧率设置为30帧/s[9]，采集时长为面试开始的前200 s，后续人脸识别将无法识别情绪或人脸的视频帧进行剔除。

为了评估面试表现，聘请5名行业专家、研究生导师根据英语能力、专业知识回答、心理素质3项指标以百分制的方式对学生进行打分，评分表如表1所示，每个学生的最终得分为5位专家评分的均值，英语能力、专业知识、心理素质分别占总分的30%、50%和20%。

2　研究方法

2.1　mini_Xception面部情绪识别

mini_Xception架构是一个全卷积神经网络，它由

表1　专家评分表

评分指标	英语能力	专业知识	心理素质
评价标准	语言流畅性与连贯性、语法与词汇表达	专业知识储备、思考深度与广度	情绪调控能力、自信心
总分	30	50	20

3个普通卷积层和4个残差模块以及3个池化层组成，如图1所示，作为全卷积神经网络的主流框架，与其他模型相比，有更好的性能[10]。该架构使用Dlib库进行人脸检测，其预训练模型在FER-2013数据集上测试情感分类任务获得了的65.9%的准确率。本研究对Fer-2013数据集进行一系列的扩充，将图像进行旋转、变形、归一化等图像增强操作，然后利用扩充后的数据集进行训练并保存为HDF5文件，最终其精度达到70.00%。除了“中立”情绪外，6种情绪可从面部表情中识别并量化：“开心”、“悲伤”、“厌恶”、“惊讶”、“恐惧”和“生气”。本研究使用训练后的mini_Xception模型对视频中面试者的情绪进行提取，使用Harr分类器[11]获取人脸区域坐标从而提取人脸图片，进行人脸面部特征的识别和人脸表情的分类，得到每一张人脸对应的每一种表情的概率值预测的集合，并将分类结果记录在数据库中。

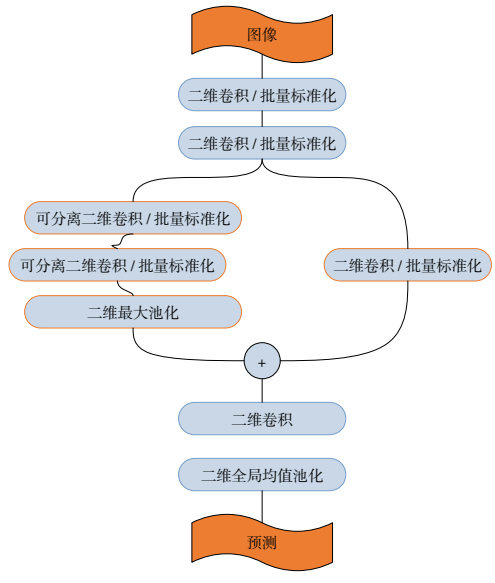


图1　mini_Xception架构

2.2　生成潜在有效的情绪特征组合

通过机器学习模型提取的情绪参数为时域上的数据集，该数据集为44×200×9的矩阵。其中，44为面试者人数，200为采样数（1 Hz ×200 s）；9列数据中除了第1列和第2列为样本ID和采集时间，其他7列为不同情绪的预测概率值。Tsfresh是基于Python的开

源时序数据特征提取工具包，内置了众多统计学的特征计算函数，例如计算均值、标准差等常规统计量，与传统特征提取方法比，其提取效率高和范围广，能自动计算出大量的时间序列特征[12]，部分特征如表2所示。情绪特征工程的流程如图2所示，最终对情绪时序数据提取了5 509×44维时序特征。

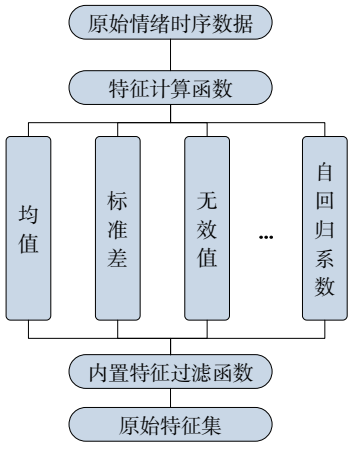


图2 Tsfresh特征提取流程

表2 Tsfresh特征工程包含的部分特征

特征名称	简单描述
abs_energy	时序数据的平方总和
absolute_sum_of_changes	一阶差分结果的绝对值之和
ar_coefficient	自回归系数，衡量时序数据的周期性和波动性
change_quantiles	给定区间的时序数据描述统计
fft_coefficient	快速傅里叶变换算法计算一维离散傅立叶变换系数
fft_aggregated	绝对傅里叶变换频谱的平均值、方差、偏度和峰度
friedrich_coefficients	Langevin模型拟合多项式系数
standard_deviation	时序数据标准偏差

2.3 特征选择

本研究对提取的众多时序特征进行初步筛选，先剔除对面试分数预测无关的特征，包括单一取值的特征、低相关性的特征。研究设置了情绪特征与面试综合分数的相关性阈值，对其绝对值大于0.45，并且通过 $P<0.05$ 显著性检验的情绪特征进行保留。为了避免多重共线性造成模型失真，在进行模型回归前对特征进行多重共线性检验，筛选合适的特征，保证回归模型的可靠性。通过计算方差膨胀因子(variance inflation factor，VIF)进行共线性检验[13]，VIF评估当预测因子相关时，回归系数的方差会增加多少，该值越大意味着共线性越强，回归系数难以估计准确。本研究对VIF大于10的特征进行剔除，确定特征数据集，最终得到满足要求的12个情绪特征，进行模型回归。

2.4 回归模型构建与分析

偏最小二乘回归（partial least square regression，PLSR）以及支持向量回归（support vector regression，SVR）是传统典型的统计方法和机器学习模型。随机森林回归模型（random forest regression，RFR）、梯度提升回归树（gradient boosting regression tree，GBRT）作为经典的集成模型，对于小样本的回归具有较强的泛化能力。为了比较不同机器学习模型在情绪评分上应用的效果，本文采用PLSR、SVR、RFR、GBRT进行回归模型的构建与预测分析。

2.4.1 模型评价指标

本文采用均方根误差（root mean square error，RMSE）、决定系数 R^2 来衡量模型的预测精度。均方根误差取值越小，模型的拟合性能越好。决定系数 R^2 越接近1，因变量的方差可以被独立变量所解释的部分越多，模型的效果越好。决定系数和均方根误差的公式如式（1）、（2）所示：

$$R^2=1-\frac{\sum_{i=1}^{n}(y_i-y_{pi})^2}{\sum_{i=1}^{n}(y_i-y_m)^2} \tag{1}$$

$$\mathrm{RMSE}=\sqrt{\sum_{i=1}^{n}\frac{(y_i-y_{pi})^2}{n}} \tag{2}$$

式中，y_i 为真实预测值；y_{pi} 为模型预测值；y_m 为真实预测值的平均值。

在调整模型的参数时，常用的方法是交叉验证（cross validation，CV）。为了解决数据集数据量不够大的问题，本文采用留一交叉验证（least one out cross validation，loocv）计算模型的评价指标的平均结果，更好地反映模型的预测精度。

2.4.2 排列重要性

排列重要性（permutation importance，PI）是与模型无关的衡量特征重要性的有效方法[14]，能够研究不同特征对模型预测结果的影响，排列重要度越高说明特征对模型预测的准确度影响越大。排列重要性的原理是改变数据集中某一个特征数据的排列顺序，其他特征保持不变，对比变化前后模型预测精度的变化大小，就认定为该特征的排列重要度。本文计算模型的PI值，绘制影响因素的排列重要性图，分析影响表现的情绪特征。

3　实验结果

在AMDRyzen74800 HCPU、RTX2060 GPU上，基于Python3.7完成人脸情绪识别与模型构建，主要采用的算法包为Keras 2.3.1、TensorFlow 2.2.0、TSfresh0.18.0和Scikit-learn 0.24.2。

对44名被试进行了面试专家评分和人脸情绪识别，评分分数分布在[75，92]，平均分为86.3分，方差为4.11，前200 s内统计的情绪结果如图3所示，其中虚线为平均值，空心圆为离群值。由图可知，面试过程中的中性情绪为学生表现的主体情绪，高兴和中性情绪表现占比超过50%。通过计算面试过程的平均情绪，其结果显示情绪为中性的占比为43.94%，其次是高兴情绪19.94%，悲伤、恐惧和愤怒情绪的占比分别为10.34%、10.43%、10.6%。本研究将惊讶和高兴定义为积极情绪，愤怒、厌恶、悲伤、恐惧定义为消极情绪。通过计算面试成绩与7种情绪平均值的Pearson相关系数，发现中性情绪、积极情绪和面试成绩呈现正相关，并且是显著的($P<0.01$)，愤怒、悲伤、恐惧与面试成绩呈现负相关，且也是显著的($P<0.01$)。

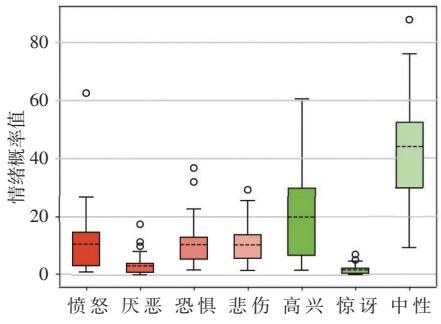


图3　7种情绪箱型图

经过特征选择确定了12个情绪特征，每个特征由不同的特征计算器函数得来。此外，在进行回归模型构建之前对12个特征计算方差膨胀因子，进行多重共线性检验，结果如下表3所示，所有特征对应的方差膨胀因子的值在[1.00，2.00]之间，均小于5，这表明特征选择出来的12个特征数据不存在严重多重共线性问题，如果只存在相对较为轻微的多重共线性，可保留作为模型构建的特征集。

采用支持向量回归、偏最小二乘回归、随机森林回归、梯度提升回归树以特征集作为输入，面试分数作为输出，构建回归模型。基于模型精度评价指标，对4种回归模型进行留一验证，表4列出了采用留一验证后的模型精度评价指标的平均值，图4是PLSR模型拟合效果图。由验证结果可知，RFR模型和GBRT模型的表现相似，其拟合能力较弱，RMSE和R^2的值非常接近，SVR模型的拟合能力和模型效果均优于RFR模型和GBRT模型。整体来看，PLSR模型的拟合能力、稳定性和精度表现最好，其R^2值为0.486，RMSE值为0.172，均高于前3个模型的预测精度。

表4　模型评价指标

模型	R^2	RMSE
偏最小二乘回归	0.486	0.172
支持向量回归	0.326	0.196
随机森林回归	0.236	0.209
梯度提升回归树	0.230	0.210

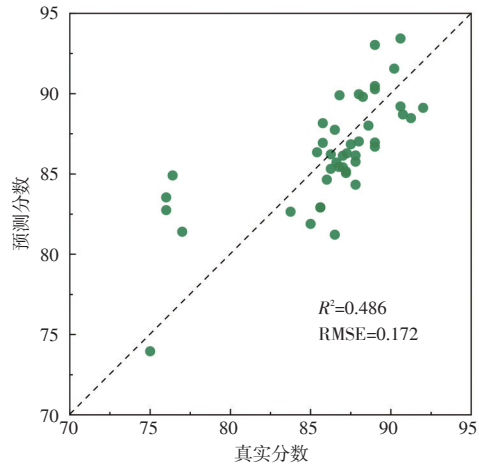


图4　PLSR模型预测效果

通过模型的训练与预测，本研究选择拟合效果更好的偏最小二乘回归模型，计算不同特征的PI并绘制影响特征的排列重要性图，如图4所示。排列重要性可以增加对模型的解释性，特征的重要性值越大，对模型的预测结果影响越大。从图中可以看出，愤怒情绪相关特征的排列重要性最高为0.123，意味着将该特征列进行随机打乱，会使模型的预测精度下降12.3%，影响相对较大。图4中权重的正值意味着该特征对预测结果有贡献，负值则表明该特征对预测结果贡献度不大，重要性接近于0。在重要性值排名前8位的情绪特征中，比例最高的是与负面情绪相关的特征（愤怒、悲伤、恐惧、厌恶）占62.5%，其次是中性情绪相关特征占25%，最后是积极情绪相关特征（高兴）占百分之12.5%。

4　结　语

根据上述分析，我们的研究发现通过分析面试视

频前200 s的表情数据，可以预测出面试成绩，可解释占比为48.6%。研究比较了PLSR、SVR、RFR和GBRT 4种模型在成绩预测中的应用，其中PLSR模型的预测精度与稳定性高于其他模型。

特征重要性排序结果表明，占比更高的消极情绪和中性情绪对面试成绩预测准确性的影响更大。结合情绪与面试成绩的相关系数，消极情绪与面试成绩呈显著负相关，可能是因为学生在面对专家提出的问题，出现无法解答的情况下，容易呈现负面情绪的表情如焦虑，愤怒；而中性情绪与成绩呈现正相关，这体现了稳定的情绪表达更容易得到专家的肯定。

本研究对GIS研究生面试过程中的情绪与面试成绩的关系进行了探索，研究表明受试者面部表情时间序列数据可以作为面试成绩表现的重要预测变量，特征挖掘进一步提高模型预测精度，采用深度学习分析情绪可以作为面试官打分的辅助依据。然而本文仅基于GIS专业硕士面试数据，结果未必具有跨专业的普适性，因此在后续的研究中，我们将扩大专业范围，测试本研究的普适性以及情绪识别模型精度的优化。

参考文献

[1] 曾洪云,刘慈,孙正宝,等. GIS专业创新创业教育与专业教育融合策略探析[J]. 地理空间信息,2022,20(9):166-168

[2] 俞艳,袁艳斌,黄解军. GIS拔尖创新人才培养课程体系设计[J]. 地理空间信息,2015,13(2):177-178

[3] Sabri N, Musa N H, Mangshor N N A, et al. Student Emotion Estimation Based on Facial Application in E-learning During COVID-19 Pandemic[J]. International Journal of Advanced Trends in Computer Science and Engineering, 2020(8):576-582

[4] Pekrun R. The Control-value Theory of Achievement Emotions: Assumptions, Corollaries, and Implications for Educational Research and Practice[J]. Educational Psychology Review, 2006, 18(4):315-341

[5] Pekrun R, Lichtenfeld S, Marsh H W, et al. Achievement Emotions and Academic Performance:Longitudinal Models of Reciprocal Effects[J]. Child Development, 2017, 88(5):1 653-1 670

[6] Gilmore D C, Ferris G R. The Effects of Applicant Impression Management Tactics on Interviewer Judgments[J]. Journal of Management,1989,15(4):557-564

[7] Hannah A, Lim B T, Ayers K M S. Emotional Intelligence and Clinical Interview Performance of Dental Students[J]. Journal of Dental Education,2009,73(9):1 107-1 117

[8] Huang J C, Ko K M, Shu M H, et al. Application and Comparison of Several Machine Learning Algorithms and Their Integration Models in Regression Problems[J]. Neural Computing and Applications,2020,32(10):5 461-5 469

[9] Yan W J, Wu Q, Liang J, et al. How Fast Are the Leaked Facial Expressions:The Duration of Micro-expressions[J]. Journal of Nonverbal Behavior,2013,37(4):217-230

[10] Fatima S A, Kumar A, Raoof S S. Real Time Emotion Detection of Humans Using Mini-exception Algorithm[C]// IOP Conference Series:Materials Science and Engineering. IOP Publishing,2021,1 042(1):012027

[11] Padilla R, Costa F C F F, Costa M G F. Evaluation of Haar Cascade Classifiers Designed for Face Detection[J]. International Journal of Computer and Information Engineering, 2012, 6(4):466-469

[12] Christ M, Braun N, Neuffer J, et al. Time Series Feature Extraction on Basis of Scalable Hypothesis Tests (Tsfresh-a PythonPackage)[J]. Neurocomputing, 2018, 307(9): 72-77

[13] Akinwande M O, Dikko H G, Samson A. Variance Inflation Factor: As a Condition for the Inclusion of Suppressor Variable (s) in Regression Analysis[J]. Open Journal of Statistics,2015,5(7):754

[14] Altmann A, Toloşi L, Sander O, et al. Permutation Importance:A Corrected Feature Importance Measure[J]. Bioinformatics, 2010, 26(10):1 340-1 347