

文章编号: 1003-0077(2021)01-0096-08

## 基于生活日志的情绪识别

王鹏宇<sup>1,2,3</sup>, 张 敏<sup>1,2,3</sup>, 马为之<sup>1,2,3</sup>, 刘奕群<sup>1,2,3</sup>, 马少平<sup>1,2,3</sup>

(1. 清华大学 计算机科学与技术系, 北京 100084;

2. 清华大学 人工智能研究院, 北京 100084;

3. 清华大学 北京信息科学与技术国家研究中心, 北京 100084)

**摘 要:** 抑郁症日益成为影响现代人生活幸福程度的重要因素。实时有效地识别用户情绪的方法对于抑郁症潜在患者的发现和治疗十分有意义。用户情绪的状态及变化会体现在其生活日志数据上。该文从使用可穿戴设备收集的用户生活日志数据出发, 对数据进行了特征方面的分析; 进一步地, 使用以回归树为弱学习器的集成学习模型, 设计了使用全部数据、仅用户自身数据以及仅他人数据进行训练的三组实验构成的对比实验框架, 以比较使用不同用户数据对识别结果的影响。实验结果表明, 基于生活日志数据的集成学习模型可以有效地识别用户的情绪状态。同时, 基于实验结果提出了用户认知不一致的猜想, 对于心理学上的抑郁分析也有一定的启发作用。该工作是目前所知第一个利用用户生活日志信息进行情绪识别以及抑郁症患者分析的工作, 为后续进一步扩大实验规模和改良实验设计提供了思路。

**关键词:** 生活日志数据; 情绪识别; 集成学习

**中图分类号:** TP391

**文献标识码:** A

## Emotion Recognition Based on Life Log Data

WANG Pengyu<sup>1,2,3</sup>, ZHANG Min<sup>1,2,3</sup>, MA Weizhi<sup>1,2,3</sup>, LIU Yiqun<sup>1,2,3</sup>, MA Shaoping<sup>1,2,3</sup>

(1. Department of Computer Science and Technology, Tsinghua University, Beijing 100084, China;

2. Institute for Artificial Intelligence, Tsinghua University, Beijing 100084, China;

3. Beijing National Research Center for Information Science and Technology, Tsinghua University, Beijing 100084, China)

**Abstract:** Depression is increasingly becoming an important factor affecting the happiness of modern people. The real time and effective identification of emotions is of great significance for the discovery and treatment of potential patients with depression. This paper analyzes the characteristics of the user's life log data collected by wearable devices. Further, in this paper, we apply an ensemble learning model with regression tree as weak learner on three groups of experiments designed to use all data, only user's own data and only others data for training, respectively. The experimental results show that the ensemble learning model based on life log data can effectively identify the emotional state of users. Meanwhile, the experimental results indicate a conjecture that users' cognition is inconsistent, serving as a potential inspiration for psychological depression analysis as well.

**Keywords:** life log data; emotion recognition; integrated learning

## 0 引言

抑郁症已经成为威胁当代人精神和生理健康的重要因素。其不仅是导致精神障碍和自杀的主要原

因之一,而且还与某些生理疾病相关。调查和分析<sup>[1]</sup>表明,抑郁症与生理疾病(如中风)的发病率和死亡率风险的显著增高有一定的相关性。因此,如何有效防治抑郁症已经成为现代医学一个重要的课题。

收稿日期: 2019-12-12 定稿日期: 2020-12-31

基金项目: 国家重点研发计划(2018YFC0831900);国家自然科学基金(61672311, 61532011);清华大学国强研究院资助项目(2019GQG0004)

目前对于抑郁症的诊察和治疗手段主要侧重于对患者的确诊和治疗,而缺乏对于潜在抑郁症用户的检测手段。由于用户对情绪倾向的自检不一定准确,潜在抑郁症用户并不容易察觉自己的情绪状况。此外,缺乏对患者情绪的实时监测手段,依靠定时给药等治疗方式难以应对突发的激烈情绪波动。同时,抑郁症存在一定的复发率,对于已治愈者,还缺少长期有效的观测手段。

抑郁症的主要特征包括心情低落、缺乏兴趣等情绪表现,故实时有效地判断和识别用户情绪的方法对于抑郁症的发现和治理大有意义。首先,实时的情绪识别有利于用户自检,便于潜在患者及时寻求医疗帮助;其次,医生可以通过对患者的情绪监测来判断是否应该改变治疗方式;最后,对于一度治愈抑郁症的人群,同样可以确认其是否再次发病。

对用户的情绪识别,尤其是实时的多模态情绪识别是近年来情感识别领域的热点话题。由于人们展现情绪的手段包括文字、表情、语音、动作、生理信号等多模态的方式,将用户的多模态的生活日志信息应用于情感预测的方式也已经被认为是主流的情绪识别手段。Busso 等人<sup>[2]</sup>的实验也表明,融合了用户多方面生活日志的模型在情绪识别方面优于仅使用单一数据的模型。要做到实时、准确地识别情绪,采集用户多种生活日志数据可以说是势在必行。

由于用户情绪是用户对于外在事物的生理反应,所以用户生理和行为上的各种信息就成了用户情绪的直观表现之一。这些信息包括心率、血压、步数、活动类型和活动强度等,对用户生活日志的采集和分析,将会给预测用户情感和保障用户健康带来莫大帮助。近年来,智能可穿戴设备的普及,使得从用户处全天候大量采集生理活动信息数据成为可能,相关的生活记录信息已被应用到心理健康领域的研究中,这为我们的工作提供了思路。

## 1 研究目标和内容

本工作的目标为试图提出一种基于用户生活日志信息的情绪识别方法,并检验将其应用于一般用户和抑郁症患者用户的效果。

由于本工作为将基于生活日志的情绪识别方法应用在抑郁症领域的初次尝试,没有相应的成熟数据集,故本工作中,我们将从收集和整理用户生活日志和情感标签数据出发,构建数据集,进而对生活日志数据的特征进行统计分析。

另外,由于基于生活日志的抑郁症相关研究还属于新的领域,所以我们的工作将更加侧重于对情绪识别可行性和可扩展性的研究上,而非提出新模型。本工作尝试将生活日志数据应用在一般用户和抑郁症患者用户具体的情绪预测工作中,检验使用生活日志对抑郁症患者情绪预测的可行性,并进行相应的对比实验,对结果及其意义进行讨论。

## 2 相关工作

### 2.1 用户生理特征和情绪的关系

心理学研究已经对用户的一系列生理特征诸如血压、心率、运动情况和用户情绪之间的关系展开了探索。其中对于血压的研究中,Davydov 等人的实验<sup>[3]</sup>指出,心血管收缩压的强度与用户的消极情绪唤起度有着直接的联系;Southard 等人关于 28 位青少年血压和情感的研究<sup>[4]</sup>表明,血液的动态收缩压与焦虑、敌对、抑郁和紧张的情绪呈正相关关系,而动态舒张压与青少年敌意、抑郁和不安情绪以及其对环境的敌意感知相关;而在关于静息收缩压的研究中,Johnson 对 32 名男性志愿者的压力和唤起度的测量和分析<sup>[5]</sup>显示,静息收缩压水平和部分用户的焦虑情绪有相关性。关于心率的研究中,Brosschot 等人的实验<sup>[6]</sup>发现,相比积极情绪,消极情绪对于心率的影响时间更长;而 Joel 等人的研究<sup>[7]</sup>揭示,在紧张性刺激的应激源影响下的高频心率变化和抑郁情绪有关。

除了循环系统方面的研究,还有一部分研究人员对用户的脑电图(EEG)和皮肤电流反映(GSR)曲线进行了研究。Li Yang 等人<sup>[8]</sup>针对脑电情感识别问题提出了一种新的回归模型,并在 14 名受试者的脑电情感数据上进行了验证;Pawel 等人<sup>[9]</sup>将 EEG 和 GSR 信息结合起来,对 27 名被试者的观影情绪进行了识别;在 Vahey 等人的一篇综述<sup>[10]</sup>中陈述了这一事实:情绪障碍的人群的 GSR 曲线更加平缓,同时平均曲线高度也低于正常人群。

用户运动状态和睡眠质量的研究也已经有所进展,Thayer 等人发现<sup>[11]</sup>,人的情绪倾向与每天走的步数有关,步数越多,越容易产生积极的情绪倾向;Sano 等人对大学生睡眠时间和睡眠不规则性的分析<sup>[12]</sup>显示,悲伤情绪和不良健康相关因素之间有统计学意义的关联,睡眠规律是比睡眠时间更重要的情绪鉴别器。

## 2.2 基于生活日志信息的分析与挖掘

通过智能设备对用户进行全天候的信息采集,可以使我们方便地获得更加丰富的数据供学习和训练。同时,用户的生活记录信息受到包括自然语言理解和情感识别等领域在内的各个研究领域的重视,NTCIR-13<sup>[13]</sup>也将用户生活记录信息的处理作为其核心任务之一。

在有关用户生活日志数据的研究中,情感相关的研究取得了多样的成果。Rika Mochizuki 等人提出了基于用户生活日志比较的情感交流<sup>[14]</sup>和情感分享<sup>[15]</sup>模型;Nomiya 等人<sup>[16]</sup>提出了基于面部表情在生活日志视频中探测情感场景的算法。

此外,生活日志信息也被用于人格分析和用户的个性化建模领域,同样取得了相当可靠的结果。Soleimaninejadian 等人的实验<sup>[17]</sup>中,将用户的性别、情绪变化、心率和房间的信息用作材料,分析了用户人格方面的特征;Mafrur 等人<sup>[18]</sup>则基于智能手机传感器收集的信息,对用户进行了行为识别以及个性化的行为建模,并提出了数据缺失的解决方案。

## 2.3 现有工作中的不足

在现有研究用户生理与情绪相关性的心理学实验中,多限定测量用户在相对特殊条件下的情感,或者使用不便携带的仪器测量用户的生理信号,难以做到一般性和便携性;而现有的生活日志相关研究中,多数情况下只采用某一种生理信息作为参考,并且难以做到实时预测。另外,在抑郁症相关的生活日志采集与相应的情感识别方面,目前还是空白。而本文的工作将直接以用户在一般生活中的多种生理数据为依据,尝试实时预测情感。同时,我们的工作也将验证基于生活日志的情绪识别在抑郁症相关场景下是否可行。

## 3 数据的采集与分析

### 3.1 生活日志数据的采集与整合

由于本工作是生活日志应用于抑郁症患者情绪识别的第一个工作,没有成熟的抑郁症患者相关生活日志和情绪标签的数据集。因此,本工作必须从收集一般用户和患者用户的生活日志和情感标签数据出发,同时整合已有的其他方面的生活日志数据集,构建本工作所需的抑郁症相关的生活日志与情

感数据集。

#### 3.1.1 手环数据的采集与整合

本工作中,使用小米手环 3 及与其配套的开源安卓应用 Gadgetbridge 作为数据的采集手段,对 1 位一般用户和两位抑郁倾向用户进行了数据采集。通过定期将用户的手环与 Gadgetbridge 应用进行同步,采集并保存用户的生活日志数据,三位用户均全天佩戴手环。与此同时,本工作还整合了开源的 LifeMusic 数据集中 13 位一般用户的数据。

手环数据的采集频率为每分钟一次,采集的生活日志数据包括 4 种特征:用户在采集数据时的活动类型、用户当时的活动强度、用户的心率、以及该分钟内用户走过的步数。

#### 3.1.2 用户情感标签的采集与整合

用户在佩戴手环采集生活日志数据的同时,也按照要求记录自己每日的活动。其记录的信息包括:活动开始时间和结束时间、活动的类型、活动起止时的情绪。完整的情感标签采集表格及示例数据如表 1 所示,其中“情绪”一栏为从表 2 中展示的 12 种预选项中选择一项。

同时,为减少用户记录压力,我们仅要求用户约每小时记录一次情感标签数据。将采集到的情感标签添加到生活日志数据上,构建原始的生活日志-情感标签数据集。

表 1 情感标签采集表格及数据示例

|      |      |       |    |    |
|------|------|-------|----|----|
|      | 年份   | 月份    | 日期 | 小时 |
| 活动开始 | 2019 | 3     | 19 | 16 |
| 活动结束 | 2019 | 3     | 29 | 16 |
|      | 分钟   | 情绪    |    | 活动 |
| 活动开始 | 17   | 平静    |    | 工作 |
| 活动结束 | 36   | 无聊或厌烦 |    |    |

表 2 情绪可选项

|    |       |       |
|----|-------|-------|
| 平静 | 愉悦    | 抑郁或苦闷 |
| 镇静 | 放松    | 伤心    |
| 激动 | 不安    | 生气    |
| 高兴 | 无聊或厌烦 | 发怒    |

### 3.2 数据集规模与无效数据筛除

数据采集过程中,由于手环佩戴时不可避免地存在偶发的接触不良问题,因此采集到的数据中存在 31.3% 的无效数据,将这些数据筛除后,得到数

据集规模如下：

生活日志数据：总条数为 218 330 条,包含 16 人共 241 天次的的数据,其中患者用户数据为15 044 条,时间跨度 17 天。

原始情感标签数据：共 290 条原始数据,其中患者用户共 41 条原始记录。

### 3.3 情感标签的补充

在原始数据集中,情感标签数据仅为 290 条,难以作为训练数据进行实验,因此,需要考虑合适的标签补充方式,基于已有的情感标签数据,合理地补充情感标签,以使情感标签数据数量达到足以进行训练的程度。在本文实验中,采用了以下两种方法对情感标签进行补充：

(1) 基于活动的情感填充方法：用户在进行同一种活动时,如果时间跨度不大,且活动开始与结束时情感标签一致,则有理由认为进行该活动时用户的情感标签与该活动起止时间的情感标签一致。按此规则对原数据集进行填充,取填充后数据集中所有带有情感标签的数据点,构成新数据集,称为活动填充数据集。

(2) 基于原有标签的情感标签填充：由于用户记录情感标签数据的时间点往往是在其情绪较为明显的时候,而在此前后,情绪的变化需要时间完成,所以可以认为时间上紧邻原始数据的情感标签和该情感标签具有一致性。按此规则对原数据集进行填充,取填充后数据集中所有带有情感标签的数据点,构成新数据集,称为标签填充数据集。

### 3.4 数据集部分特征统计与分析

首先对两类用户的活动强度进行占比统计,得到图 1 和图 2。由这两个图可知,一般用户的较高强度活动的时间占比高于患者用户,而对两类用户的步数进行统计则发现,绝大多数时间内,两类用户的步数均为 0。

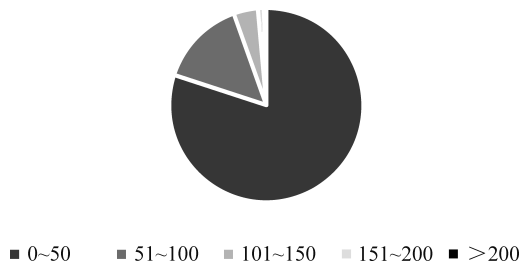


图 1 一般用户活动强度占比

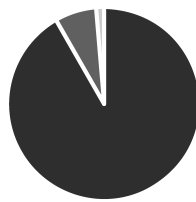


图 2 患者用户活动强度占比

将两类用户的活动强度和步数分别按照频数进行统计,得到图 3 和图 4,其中图 4 的分布曲线经过局部放大。由这两图可知,尽管一般用户的最大活动强度和步数大于患者用户,两类用户的活动强度和步数频数分布曲线趋势基本相同,在频数分布的意义上,两类用户无法得到明确区分。

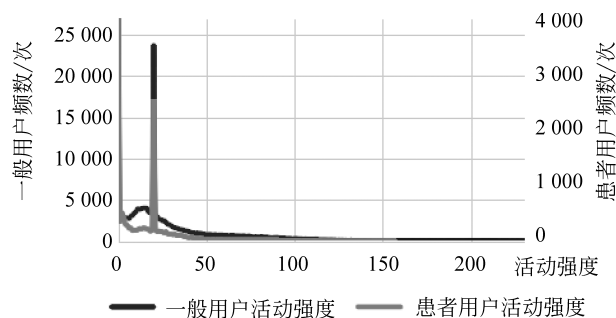


图 3 按频数统计的活动强度曲线图

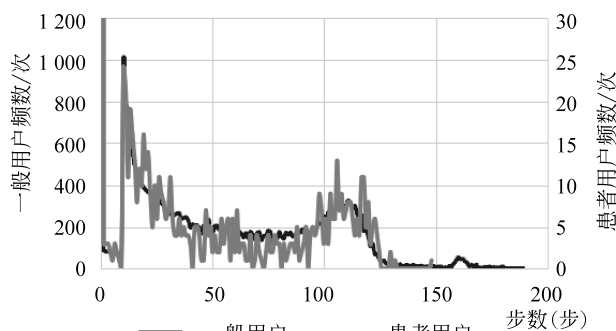


图 4 按频数统计的步数曲线图

## 4 基于生活日志的情绪预测

### 4.1 任务描述

本节中,我们对一般用户和抑郁症患者用户进行情绪预测。情感预测任务使用生活日志和情感标签训练模型,使模型能够预测给定的生活日志特征对应的情感标签。由于本文工作首次涉及抑郁症患者情绪识别问题,因而侧重于验证既有情绪识别模



型的可行性,观察不同用户的效果和用户间差异,同时探究方法本身的可扩展性。

统计分析的结果表明,单纯从数据的频数和分布来看,难以区分两类用户。

## 4.2 特征与模型设计

### 4.2.1 生活日志特征设计

本文选取从手环采集到的用户活动类型、活动强度、心率和步数这 4 个维度的特征用作实验数据。在收集到的生活日志数据中,活动类型数据在形式上与其他 3 种数据有较大差异:首先其数值为离散数据,其次其数值代表的含义与数值大小没有关系。因而对所有出现的活动类型、频数进行统计,结果如表 3 所示。

表 3 对活动类型的频数统计

|    |        |        |        |        |
|----|--------|--------|--------|--------|
| 类型 | 80     | 112    | 96     | 1      |
| 频数 | 93 416 | 39 965 | 30 799 | 21 210 |
| 类型 | 90     | 122    | 其他     |        |
| 频数 | 12 358 | 8 645  | 1 1937 |        |

统计结果显示,排名前 6 的活动类型频数和其他活动类型频数之和处于同一数量级范围内,所以可将活动类型以 one\_hot 的形式拆分成 7 维向量,以弥补原始用户生活日志特征维数过少的缺陷,在实验中我们也将该方法与不拆分用户活动类型的方法进行了对比。

### 4.2.2 多维度情感模型与情感标签合并

两个较重要的多维度情感模型是 Russel 等人于 1980 年提出的 Arousal-Valence 模型<sup>[19]</sup>与 Mehrabian 等人于 1996 年提出的 PAD 模型(Pleasure Arousal Dominance, PAD)<sup>[20]</sup>。模型使用多维连续值描述具体的情感,可以更加细致地对情感进行区分或对情感定义距离。下面分别介绍两个模型的概念。

如图 5 所示,以唤起度和正负性为平面的两个维度,所有的常见情感以坐标系中的点的形式存在于平面内。其中唤起度描述情感的强度,即情感中含有的能量多少;而正负性代表情感的倾向,即情感是积极的还是消极的。唤起度和正负性的取值范围是区间 $[-1,1]$ ,在这个模型下,情感被表示为二元组 $(a, v)$ 的形式,两个维度分别对应 Arousal 和 Valence 的取值,情感积极性越高,唤醒度越强,其位置越靠近右上角,反之,则越靠近左下角。

在 PAD 模型中, P 维度起到衡量情感正负性的作用,而新引入的 D 维度则起到衡量产生情绪的个

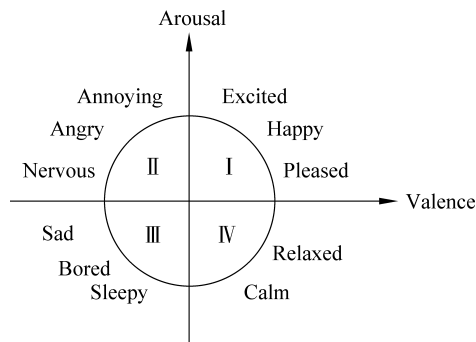


图 5 唤起度—正负性情感平面

体对当时情景和周围人的控制状态。在 PAD 三维模型下,一些仅凭唤起度和正负性不易区分的情感,如愤怒和恐惧,可以得到很好的区分。

在数据集中,实际出现的情感标签为除“发怒”外的 11 种。其中存在部分倾向相近的情绪。由于实际上我们更加关注的是情感本身的正负性和唤醒度相关的特征,所以可以将情感标签按照唤起度-正负性模型的四象限合并成表 4 中显示的 5 类。将此合并方法应用到“活动填充数据集”和“标签填充数据集”上,分别得到“活动填充-合并数据集”和“标签填充-合并数据集”,以便验证该方法的效果。

表 4 情感标签合并结果

|      |           |           |    |    |    |    |
|------|-----------|-----------|----|----|----|----|
| 情感标签 | 激动        | 高兴        | 愉悦 | 生气 | 不安 | 发怒 |
| 分类   | 1         | 1         | 1  | 2  | 2  | 2  |
| 情感标签 | 无聊或<br>腻烦 | 抑郁或<br>苦闷 | 伤心 | 镇静 | 放松 | 平静 |
| 分类   | 3         | 3         | 3  | 4  | 4  | 0  |

### 4.2.3 基于决策树的集成学习模型

对于数据量较小的生活日志数据集,提高模型正确率的方法选用集成学习方法。目前比较主流的集成学习方法之一,是以决策树作为弱分类器来实现集成学习的方法。而在本文中,我们用来预测用户情绪的模型为 Breiman 提出的随机森林(RF)算法<sup>[21]</sup>和 Friedman 提出的 GBDT 算法<sup>[22]</sup>,二者所使用的弱学习器均为 CART 树。

CART 树包含分类树和回归树,二者的区别为样本输出的连续性。其中前者基于基尼系数进行优化,后者则基于方差。按照特征值  $A$  进行划分的样本  $D$ ,其基尼系数如式(1)所示。

$$\text{Gini}(D, A) = \frac{|D_1|}{D} \text{Gini}(D_1) + \frac{|D_2|}{D} \text{Gini}(D_2) \quad (1)$$

CART 分类树的建立就是不断选择基尼系数最小的划分点和特征进行二分类,预测时选择样本点所在叶节点中概率最大的值。

基于回归树的集成学习算法,其本质是在预测时,考虑处于回归树森林中的多棵回归树的预测结果,最终优化模型整体的预测结果。其中,随机森林算法使用 CART 树作为弱学习器,在学习器的集成上使用了 bagging 原理,其中各个弱学习器之间没有关系。随机森林的特点在于,随机选取一部分样本特征作为决策树子树划分的依据,使得模型的泛化能力更强。在进行分类预测时,采用投票法(分类树)或取均值(回归树)作为输出。

而 GBDT 算法则采用了 boosting 原理,用前一轮迭代得到的强学习器和损失函数建立弱学习器,使得本轮的损失函数最小,进而构造更强的学习器。通过前一次的结果改变新增弱学习器的分类或回归方式,通过集成来不断尝试提高整体强学习器正确率  $P$  (Precision)。

4.3 情绪预测实验设置

我们基于用户不同日期间数据的独立性,设计了一个包含三个使用不同训练数据进行测试的对比实验框架,并且在这个框架下,验证不同情感标签补充方法合并情感标签与否的效果优劣。对比实验框架的设计思路是:选取某用户某天的数据作为测试集,在余下的数据中选择训练数据。训练手段采用 GBDT 和 RF 算法。框架内包含的三个实验如下:

**实验一** 基于全体数据的学习和预测,将其余所有数据作为训练集。

**实验二** 仅使用自身数据的学习和预测,将该用户的其余数据作为训练集。

**实验三** 仅使用他人数据进行预测,使用其他人的数据作为训练集。

上述三个实验均使用用户情感预测的正确率作为评价标准,分别计算数据集上的平均正确率和用户的平均正确率。

其中,实验一与实验二对比,可以验证他人数据的有无是否对情感预测有积极影响;实验一与实验三对比,可以验证在仅有他人数据的情况下是否能够进行相对准确的预测;三者对比,可以对不同人认知情感的偏差这一现象进行分析。

4.4 实验结果分析

实验一 基于全体数据的学习预测

使用 GBDT 和 RF 方法,在 4 个数据集上进行

基于全体数据的学习和预测,其平均正确率如表 5 所示。由表 5 可知,使用随机猜测方法的对照组命中率接近于用户情绪类别的倒数,而使用用户的生活日志数据进行情感预测可以得到相对较优的结果。其中,在使用全部数据时,GBDT 的表现要优于 RF,由 4 个数据集结果比较可知,无论是否合并标签数据,使用基于活动的情感标签填充方式的表现都要明显好于基于标签的填充方式,这表明使用基于活动的标签填充方式所补充的数据更加贴近于用户的实际情感数据。进一步表明,用户的情绪变化依赖于活动,在短时间内进行同一活动的情况下情绪倾向于稳定;而在活动状态进行转换时,用户的情绪也倾向于发生变化。此外,活动类别的分割对于预测正确率有一定提升。

表 5 是否分割活动类别结果对比(%)

| 活动类别 | 分割    | 不分割   |
|------|-------|-------|
| GBDT | 35.65 | 35.04 |
| RF   | 31.83 | 30.97 |

对于情感标签合并的效果而言,由原有的 11 分类问题简化为 5 分类问题后,正确率确实有所提高,但并没有显著提高正确率,这与原有数据集中情感标签并非均匀分布有关。

统计 GBDT 方法在活动填充-合并数据集上得到的各用户平均正确率,绘制图 6,对结果进行用户粒度的分析。在基于全体数据的学习和预测中,患者用户达到的最高预测正确率为 56.46%,而除 1 号、6 号、7 号之外的用户的平均正确率可达 50.51%。对于 5 号和 14 号用户,其正确率可以达到 93.26%和 75.21%。表明基于生活日志和情感标签,可以对部分用户进行有效的预测。

表 6 实验一不同数据集结果对比(%)

| 数据集  | 活动填充  | 标签填充  | 活动合并         | 标签合并  |
|------|-------|-------|--------------|-------|
| GBDT | 35.65 | 32.04 | <b>37.03</b> | 33.49 |
| RF   | 30.98 | 26.06 | 32.96        | 28.75 |
| 随机猜测 | 9.07  | 9.20  | 20.39        | 21.36 |

对于数据异常的 6 号用户,经比对预测结果和情感标签,发现该用户倾向于将大多数标签标注为“愉悦”等积极情感,而基于全体数据训练的模型对此给出的预测则大多数为“平静”。这说明 6 号用户的标注可能存在个人偏差,即情感标注与实际情绪不一致。

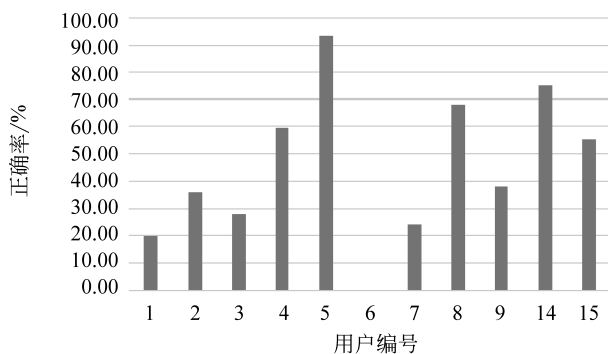


图6 实验一用户平均正确率

### 实验二 仅使用用户自身数据的学习预测

GBDT 和 RF 在仅使用用户自身数据时,用同样方法绘制用户粒度分析图(见图7)。由图7可见,大部分用户的数据中,实验一的平均正确率要高于实验二的平均正确率。然而也存在反常的用户,1号、6号和7号用户仅在使用自己的数据时,取得较高的正确率,而9号用户在使用自己的数据时,正确率反而偏低。对于这些反常用户进行预测结果和用户标签的比对后发现,1号和7号与6号相反,倾向于向数据集中标注“无聊”和“抑郁”;9号用户的标注则表现为每天的情感较为单一,但每天之间情感差异很大,导致实验二中,多数情况下测试集中几乎不包含训练集的数据。

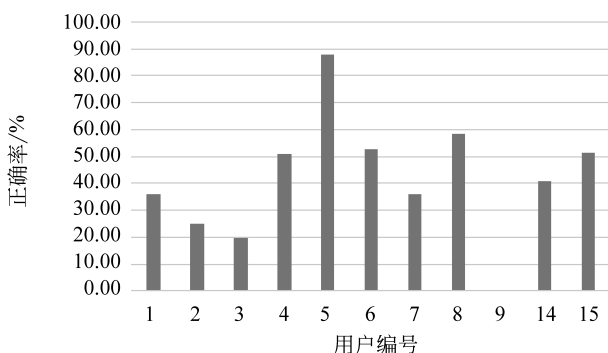


图7 实验二用户平均正确率

综合实验一和实验二的结果,我们提出了用户认知不一致现象的猜想。该现象的表现有两个,一是1号、6号和7号展现的用户对自己情绪的认知与实际情绪之间存在不一致;二是用户之间对于相同情绪的认识不尽相同。当算法的预测,尤其是基于全部用户数据的多种方法预测接近且与用户的标签不一致时,不能简单地当作错误来认识,应考虑用户认知不一致现象发生的可能性。此时的预测结果有其特别的意义,代表着判断出的用户未能察觉或者认识错误的潜在情绪。

### 实验三 仅使用他人数据的学习预测

本实验的用户平均正确率如图8所示。

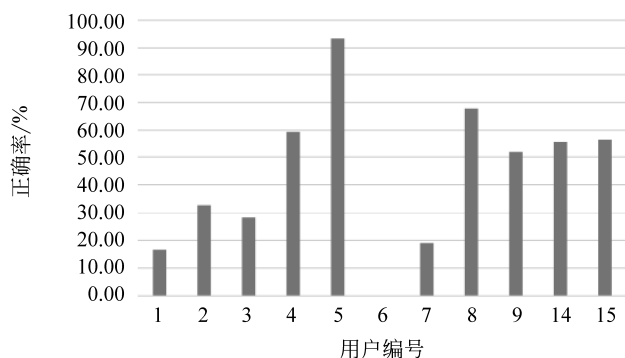


图8 实验三用户平均正确率

由图8可知,对于大部分新用户,仅使用他人数据预测也有相当高的正确率,对于编号15的患者用户而言,使用一般用户数据进行情感预测的正确率可以达到56.46%。表明患者的情绪可以使用一般数据进行正常预测。其意义在于:一方面说明患者情绪和一般用户情绪与生活日志数据的对应规律大体相同,并非不可预测;另一方面,训练相应的情绪预测模型,不需要采集大量的患者数据,使用一般用户数据即可得到相对较好的结果。另外,对比实验二和实验三的数据可以发现,对于大多数用户,完全使用用户自己的数据进行预测的准确性要略高于仅使用他人数据,这说明个人数据所带有的偏差确实会影响情绪预测的正确率。

## 5 总结与未来工作

本文是首个利用生活日志数据进行抑郁症用户情绪识别和分析的工作,对基于生活日志的情绪识别方法在此方面的应用可行性进行了验证。

这项工作中,我们从零开始收集和整合用户的生活日志和情感标签数据,构建了包含两类用户的生活日志情感数据集,并进行了数据的统计分析。我们以GBDT和RF作为训练模型,设计了包含使用全体数据、用户自身数据和仅使用他人数据的三个对比实验的框架进行了实验,比对了情感标签不同补充方式的效果和不同用户的训练数据之间的偏差对于识别正确率的影响。提出了有关用户认知不一致问题的猜想,并对其进行了分析。该现象的发现对于心理抑郁分析有一定的启示作用,但彻底确认还需更多数据支持。

未来我们将收集更大规模的数据集以供进一步

的实验,从而探究用户认知不一致现象和提高模型的可靠性。同时,需要发掘更加准确的情感记录手段,使用户的情感记录更加贴近实际。此外,还要结合心理学的研究与实践,改进模型效果,提出新模型。

## 参考文献

- [1] Pan An, Sun Qi, Okereke Olivia I, et al. Depression and risk of stroke morbidity and mortality: A meta-analysis and systematic review[J]. The Journal of the American Medical Association, 2011, 306(11): 1241-1249.
- [2] Busso Carlos Deng Zhigang, Yildirim Serdar, et al. Analysis of emotion recognition using facial expressions, speech and multimodal information[C]//Proceedings of the 6th International Conference on Multimodal Interfaces, 2004: 205-211.
- [3] Davydov D M, Stewart R, Ritchie K, et al. Depressed mood and blood pressure: The moderating effect of situation-specific arousal levels[J]. International Journal of Psychophysiology, 2012, 85(2): 212-223.
- [4] Southard D R, Coates T J, Kolodner K, et al. Relationship between mood and blood pressure in the natural environment: An adolescent population[J]. Health Psychology, 1986, 5(5): 469-480.
- [5] Johnson D W, Anastasiades P. The relationship between heart rate and mood in real life[J]. Journal of Psychosomatic Research, 1990, 34(1): 21-27.
- [6] Brosschot J F, Thayer J F. Heart rate response is longer after negative emotions than after positive emotions [J]. International Journal of Psychophysiol, 2003, 50, 181-187.
- [7] Joel W Hughes, Catherine M Stoney. Depressed mood is related to high-frequency heart rate variability during stressors[J]. Psychosomatic Medicine, 2000, 62(6): 796-803.
- [8] Li Yang, Zheng Wenming, Cui Zhen, et al. EEG emotion recognition based on graph regularized sparse linear regression[J]. Neural Processing Letters, 2019, 49(2): 555-571.
- [9] Paweł Tarnowski, Marcin Kołodziej, Andrzej Majkowski, et al. Combined analysis of GSR and EEG signals for emotion recognition[C]//Proceedings of the 2018 International Interdisciplinary PhD Workshop, 2018: 137-141.
- [10] Vahey R, Becerra R. Galvanic skin response in mood disorders: A critical review [J/OL]. International Journal of Psychology and Psychological Therapy, 2015, 15(2): 275-304.
- [11] Thayer R E, Biakanja L, O'Hanion P, et al. Amount of daily walking predicts energy, mood, personality, and health [C]//Proceedings of the 2005 Annual Meeting of the American Psychological Association, 2005.
- [12] Sano A, Yu A Z, McHill A W, et al. Prediction of happy-sad mood from daily behaviors and previous sleep history[C]//Proceedings of the 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, 2015: 6796-6799.
- [13] Kato M P, Liu Y. Overview of NTCIR-13[C]//Proceedings of the 13th NTCIR Conference on Evaluation of Information Access Technologies, 2017: 1-5.
- [14] Rika Mochizuki, Tomoki Watanabe, Toru Kobayashi. Emotion communication model based on life-log comparison-mutual understanding through comparable experiences [C]//Proceedings of the 2012 IEEE/IPSJ 12th International Symposium on Applications and the Internet, 2012: 91-99.
- [15] Rika Mochizuki, Tomoki Watanabe. Emotion sharing model based on life-log comparison[J]. Information and Media Technologies, 2016, 11: 79-92.
- [16] Shota Sakaue, Hiroki Nomiya, Teruhisa Hochin. Estimation of emotional scene from lifelog videos in consideration of intensity of various facial expressions [M]. Roger Lee. Software Engineering, Artificial Intelligence, Networking and Parallel/Distributed Computing. Cham: Springer International Publishing, 2018: 121-136.
- [17] Soleimaninejad P, Wang Y, Tong H, et al. THIR2 at the NTCIR-13 lifelog-2 task: Bridging technology and psychology through the lifelog personality, mood and sleep quality[C]//Proceedings of the 13th NTCIR Conference on Evaluation of Information Access Technologies, 2017: 20-27.
- [18] Mafrur R, Nugraha I G D, Choi D. Modeling and discovering human behavior from smart-phone sensing life-log data for identification purpose[J]. Human-centric Computing and Information Sciences, 2015, 5(1): 1-18.
- [19] Russell J A. A circumplex model of affect[J]. Journal of Personality & Social Psychology, 1980, 39(6): 1161-1178.
- [20] Mehrabian A, et al. Pleasure-arousal-dominance: A general framework for describing and measuring individual differences in temperament[J]. Current Psychology, 1996, 14(4), 261-292.
- [21] Leo Breiman. Random forests[J]. Machine Learning, 2001, 45(1), 5-32.
- [22] Friedman J H. Greedy function approximation: A gradient boosting machine[J]. The Annals of Statistics, 2001, 29(5): 1189-1232.

(下转第 112 页)



[21] 赵希明. 基于深度学习的微博情感分析研究[D].北

京:北方工业大学硕士学位论文,2019.



陈玮(1964—),硕士,副教授,主要研究领域为模式识别与智能系统,智能信息处理。

E-mail: chenwei@usst.edu.cn



林雪健(1996—),通信作者,硕士研究生,主要研究领域为情感分析。

E-mail: xjbbqcl@126.com



尹钟(1988—),博士,副教授,主要研究领域为生物医学信号处理,情感计算。

E-mail: yinzhong@usst.edu.cn



(上接第 103 页)



王鹏宇(1997—),本科生,主要研究领域为人工智能,信息检索。

E-mail: wpymango@163.com



张敏(1977—),通信作者,博士,副教授,主要研究领域为个性化推荐,信息检索,用户建模。

E-mail: z-m@tsinghua.edu.cn



马为之(1992—),博士,助理研究员,主要研究领域为个性化推荐与用户建模。

E-mail: mawz12@hotmail.com