

信息网络专题研究文献阅读报告（应用层）

一、 文献信息

- 1、 作者： Shanya Sharma(SAP LABS); Manan Dey(SAP LABS)
- 2、 论文题目： Assessing Viewer's Mental Health by Detecting Depression in YouTube Videos
- 3、 发表途径： 33rd Conference on Neural Information Processing Systems (NeurIPS 2019),
Vancouver, Canada
- 4、 发表时间： Sat Dec 14th, 2019

二、 问题意义

1、 研究问题

开发和测试机器学习技术的功效,使其应用于通过正式文本捕获的 YouTube 视频内容,并确定视频是否是抑郁的, 或者是否含有使人变得压抑郁闷的诱因。

2、 研究背景

众所周知,抑郁症是最为普遍的心理健康问题之一,它也被证实了是自杀的主要原因之一,这种疾病给家庭和社会都带来了沉重的经济负担。据世界卫生组织统计,世界上大约有 3 亿人患有抑郁症,但绝大多数都还没有进行治疗,原因是他们并不知道得了这种病。因此,更早地发现和诊断抑郁症至关重要,这篇文章提出了一种在早期能帮助诊断是否患抑郁症的方法,这是一个能够通过跟踪 YouTube 视频网站上用户的观看记录来分析视频的情感模式,从而预测用户的心理状态的模型。

与以往的研究不同,作者是从用户末端被动地收集数据,以便客观地了解用户的心理状态。这和从社交媒体平台上获取个人的自我描述型采访视频相比,更有说服力和公正性。

3、 研究意义

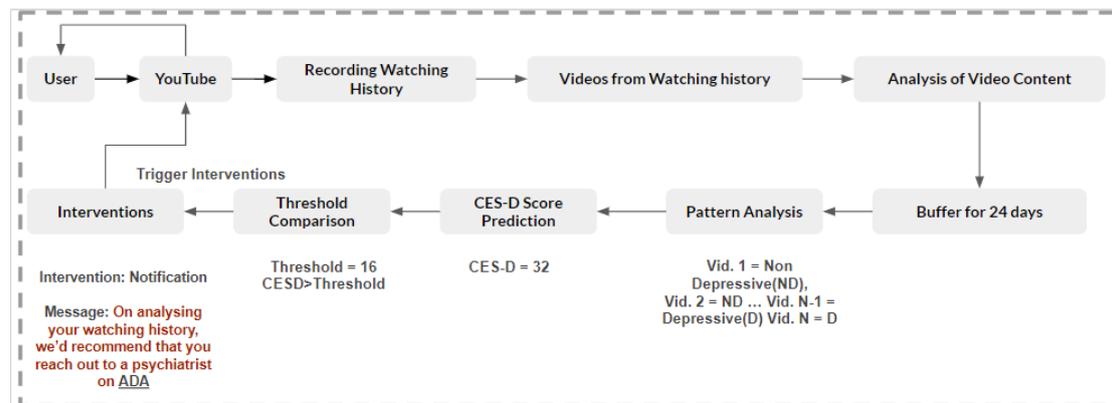
- (1) 构建了一个分类器,将视频分类为"抑郁/非抑郁",帮助了解视频的内容;
- (2) 为了而给出模型分类结果的真实确认,对发布在视频下的评论用 CES-D 分数作为分类准确性的度量。

三、 思路方法

1、 模型描述

管道式模型主要由三部分组成:数据采集,特征提取(个人视频及评论分析),分析随

时间变化观看视频的模式。如下图所示。因为抑郁症患者倾向于在很长一段时间内做相同的事，故相同情感内容的视频是重点分析的模式。



2、数据

总共 3000 份正式文本，分为抑郁和非抑郁两大类，其中抑郁的占 48%，非抑郁的占 52%。这些数据由搜索关键字得出，抑郁类的关键字有：

Button Poetry Depression、Self-harm confessions、Self-injury、The Dying Edits、Kurt Cobain Suicide Confessions、Confessions in depression、Depression short films、Anorexia short films、Broken short films、Famous suicide letters、Suicide short films、Self-harm short films 等，非抑郁的则对应 funny、educational 等关键词。

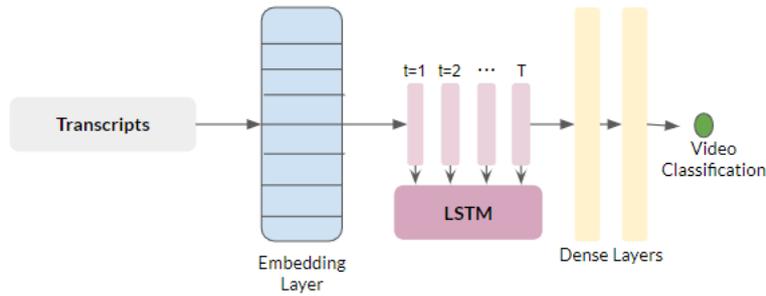
3、分类

作者测试了 3 个模型：

(1) N 元模型的朴素贝叶斯算法：使用从 Empath 模型中提取的特征构建多语种朴素贝叶斯 (NB) 分类器，分析 CES-D 中描述的症状所含的 seed-terms，由下表列出。

Questionnaire Symptoms	Seed-Terms
I DID NOT FEEL LIKE EATING	APPETITE
I FELT THAT I COULD NOT SHAKE OFF MY BLUES	SAD
I HAD TROUBLE KEEPING MY MIND ON WHAT I WAS DOING	DISTRACTION, ADHD
I FELT DEPRESSED	DEPRESSION
MY SLEEP WAS RESTLESS	INSOMNIA, NIGHTMARE
I FELT MY LIFE HAD BEEN A FAILURE	FAILURE, SELF-DOUBT, SELF-HATE
I FELT LONELY/PEOPLE WERE UNFRIENDLY	LONELY

(2) 长短期记忆网络 (LSTM)：对每一条经过预处理的评论，让它依次经过嵌入层、LSTM 网络、2 层致密层得到它所属的类别。使用了二元交叉熵损失和 Adam 优化算法。如下图所示。



(3) N元模型的朴素贝叶斯：在(1)分类器的基础上，对 ngram 语言模型使用词频逆文本频率(TF-IDF)分析。结果为：Empath 模型与 TF-IDF 相结合的分类精度与 LSTM 几乎相同。

4、评估

(1) 为每一条视频的每条评论的抑郁程度打分 CES-D score，它形容了在给定文本中有负面含义的各症状（如失眠、自恨、食欲等）的程度。用 CES-D（流行病学研究中心抑郁测评表）度量一般人群中抑郁症状的紧急性，用结合了 NLP 技术的 Empath Model 度量症状的出现频率。

(2) 每个视频的每条评论的分数计算方法如下

Algorithm 1 CES-D Score

```

Require: data text
set ( categoryTerms )
termFreq  $\leftarrow$  0
while term in categoryTerms do
  if term in text then
    termFreq  $\leftarrow$  termFreq + 1
  end if
end while
termFreq  $\leftarrow$   $\frac{\text{termFreq}}{\text{len}(\text{text})}$ 
neg  $\leftarrow$  Empath.analyze(text, negative)
pos  $\leftarrow$  Empath.analyze(text, positive)
connotation  $\leftarrow$  (neg - pos)
CESD  $\leftarrow$  termFreq  $\times$  connotation
return CESD

```

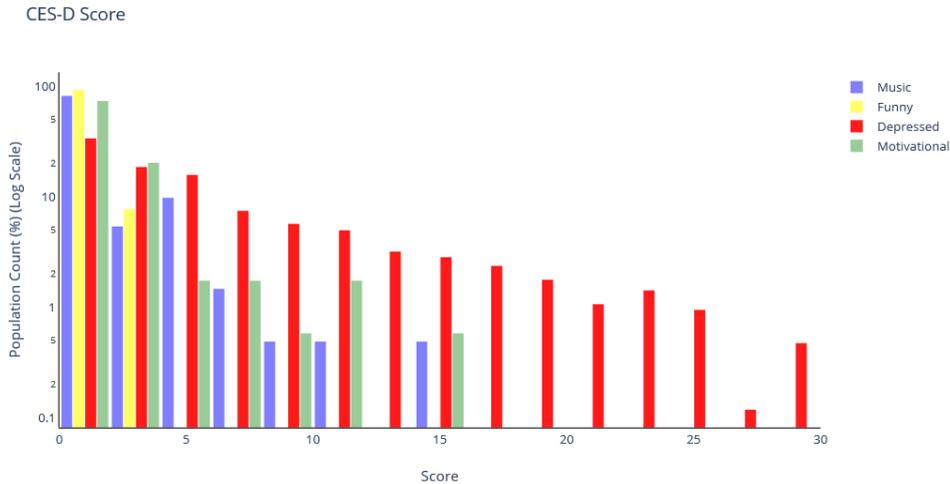
每条评论有自己的 seed-terms，每个 seed-term 生成一个类别，所有类别组成一个集合。最开始 termFreq 为 0，对文本进行遍历，如果文本中出现一个集合中的类别，该类别的频次 termFreq 就加 1。遍历结束后对频次进行归一化得到最终的 termFreq，CESD=termFreq \times connotation。

只考虑评论中 neg 一面对分数的影响，如 sad、lonely。所以如果一条评论的 pos 大于 neg，则该条评论的 CESD=0。

(3) 评论的分析

视频下的评论内容在很大程度上揭露了一个人的心理精神状态，作者将其分成四类：音

乐的、抑郁的、有趣的和自我鼓励的。比较抑郁和非抑郁类视频的 CES-D 分数，可以发现两者明显的区别。



(4) 评估结果

在为训练目的收集的抑郁视频中，得出了将 CES-D 分数平均为 20 的阈值的结论。基于此方法的分类器的准确率为 84%，用混淆矩阵可以评估。

四、 实验结论

1、 主要实验方法

使用文本特征提取的方法，借助机器学习完成分类任务，其分类效果由混淆矩阵来评估。

2、 实验结论

论证了一种分析视频中抑郁含量的方法，也证实了抑郁患者倾向于反复观看有压抑感情的视频以寻找认同感和缓解孤独感的现象。该实验的结果可以用于检测和限制基于观看历史的视频推荐策略，帮助人们尽早认清自己的心理状况，调节好自身情绪，必要时指导他们寻求专业的帮助。

五、 启发思考

机器学习在当下是一个非常热门的话题，此次学习让此前徘徊在门口的我终于跨进了它的大门，揭开了它的神秘面纱，认识到其真正内容。我觉得，一项好的技术是要能运用到生活中的，给人和社会带来益处的。无论其自身理论有多么高深，最终的使命都是帮助世界变得更好。机器学习同样如此，现今，它在各个方面被开发和挖掘，代替人力做到了许多曾经连想都不敢想的事情，比如在医学领域上的对抑郁症患者的帮助。我们在学习生活工作中做事也要带着这种想法，不是为了事情本身而做事，而是做这件事我们能从中收获到什么，带着这样的心态去做事情才是最有意义的。