Predicting Anxiety, Depression and Stress

Abstract

The Predicting Anxiety, Depression and Stress project aims to develop a model that can accurately predict the likelihood of an individual experiencing symptoms of anxiety, depression, and stress. The model will be trained on a dataset of psychological assessments and will use various features such as age, gender, personality traits, and environmental factors to predict the probability of the three mental health conditions. The goal of this project is to provide individuals with an early warning system for potential mental health issues and improve access to timely mental health care.

Introduction

The Taylor Manifest Anxiety Scale was first developed in 1953 to identify individuals who would be good subjects for studies of stress and other related psychological phenomena. Since then, it has been used as a measure of anxiety as a general personality trait. Anxiety is a complex psychological construct that includes a multiple of different facets related to extensive worrying that may impair normal functioning. The test has been widely studied and used in research however there are some concerns that it does not measure a single trait but instead measures a basket of loosely related ones and so the score is not that meaningful.

Method

The survey was open to anyone, and people were motivated to take it to get personalized results. At the end of the test, they also were given the option to complete a short research survey. This dataset comes from those who agreed to complete the research survey and answered yes to the question "Have you given accurate answers, and may they be used for research?" at the end.

This data was collected 2017 - 2019.

The following questions were included in the survey:

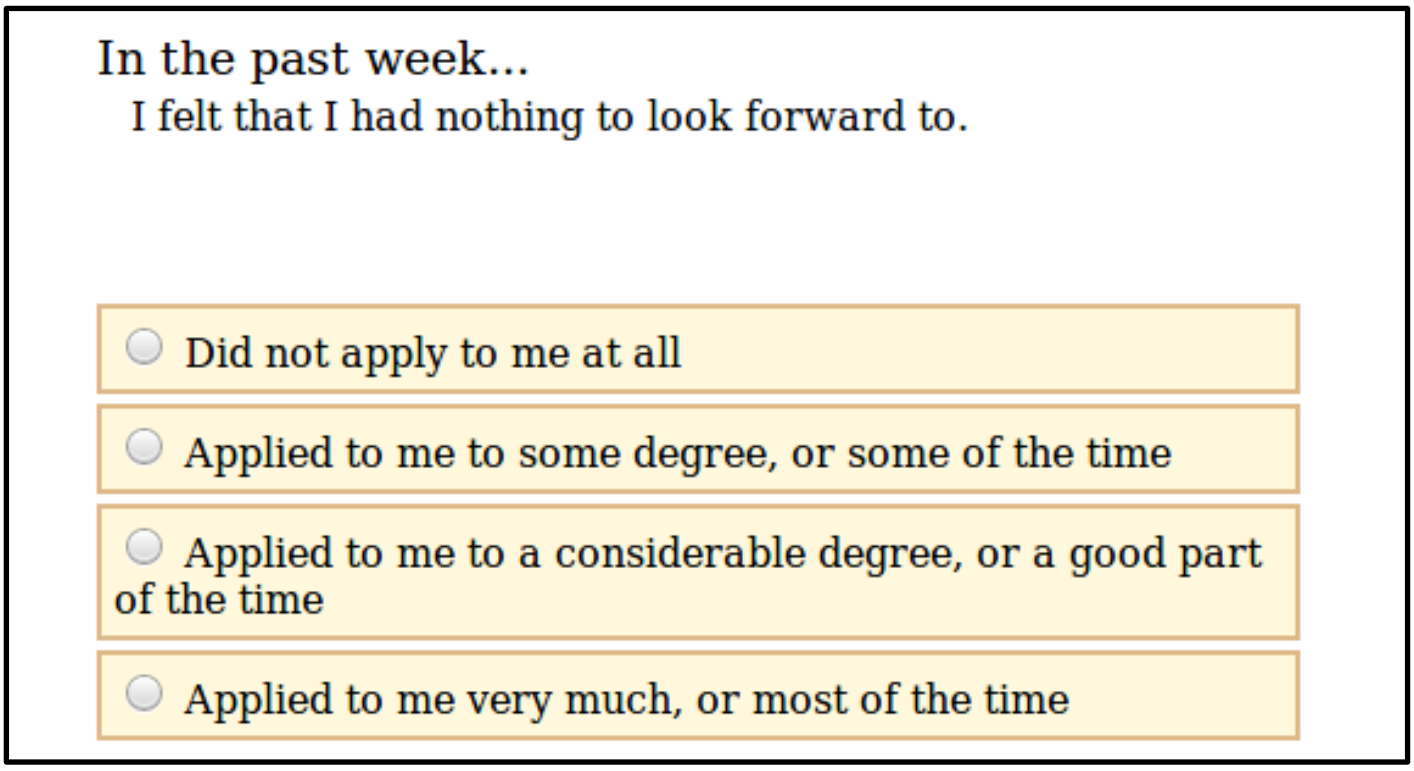
1. I found myself getting upset by quite trivial things.
2. I was aware of dryness of my mouth.
3. I couldn't seem to experience any positive feeling at all.

4. I experienced breathing difficulty (e.g., excessively rapid breathing, breathlessness in the absence of physical exertion).
5. I just couldn't seem to get going.
6. I tended to over-react to situations.
7. I had a feeling of shakiness (e.g., legs going to give way).
8. I found it difficult to relax.
9. I found myself in situations that made me so anxious I was most relieved when they ended.
10. I felt that I had nothing to look forward to.
11. I found myself getting upset rather easily.
12. I felt that I was using a lot of nervous energy.
13. I felt sad and depressed.
14. I found myself getting impatient when I was delayed in any way (e.g., elevators, traffic lights, being kept waiting).
15. I had a feeling of faintness.
16. I felt that I had lost interest in just about everything.
17. I felt I wasn't worth much as a person.
18. I felt that I was rather touchy.
19. I perspired noticeably (e.g., hands sweaty) in the absence of high temperatures or physical exertion.
20. I felt scared without any good reason.
21. I felt that life wasn't worthwhile.
22. I found it hard to wind down.
23. I had difficulty in swallowing.
24. I couldn't seem to get any enjoyment out of the things I did.
25. I was aware of the action of my heart in the absence of physical exertion (e.g., sense of heart rate increase, heart missing a beat).
26. I felt down-hearted and blue.
27. I found that I was very irritable.
28. I felt I was close to panic.
29. I found it hard to calm down after something upset me.
30. I feared that I would be "thrown" by some trivial but unfamiliar task.
31. I was unable to become enthusiastic about anything.
32. I found it difficult to tolerate interruptions to what I was doing.
33. I was in a state of nervous tension.
34. I felt I was pretty worthless.
35. I was intolerant of anything that kept me from getting on with what I was doing.
36. I felt terrified.

37.I could see nothing in the future to be hopeful about.
38.I felt that life was meaningless.
39.I found myself getting agitated.
40.I was worried about situations in which I might panic and make a fool of myself.
41.I experienced trembling (e.g., in the hands).
42.I found it difficult to work up the initiative to do things.

Each item was presented one at a time in a random order for each new participant along with a 4-point rating scale asking the user to indicate how often that had been true of them in the past week where.

1. Did not apply to me at all.
2. Applied to me to some degree, or some of the time.
3. Applied to me to a considerable degree, or a good part of the time.
4. Applied to me very much, or most of the time.



This response is stored in variable A (e.g., Q1A). Also recorded was the time taken in milliseconds to answer that question (E) and that question's position in the survey (I).

These other durations were also recorded (measured on the server's side):

- introelapse : The time spent on the introduction/landing page (in seconds)
- testelapse : The time spent on all the DASS questions (should be equivalent to the time elapsed on all the individual questions combined)

- surveyelapse : The time spent answering the rest of the demographic and survey questions

On the next page was a generic demographics survey with many different questions.

The Ten Item Personality Inventory was administered (see Gosling, S. D., Rentfrow, P. J., & Swann, W. B., Jr. (2003). A Very Brief Measure of the Big Five Personality Domains. Journal of Research in Personality, 37, 504-528.):

- TIPI1 Extraverted, enthusiastic.
- TIPI2 Critical, quarrelsome.
- TIPI3 Dependable, self-disciplined.
- TIPI4 Anxious, easily upset.
- TIPI5 Open to new experiences, complex.
- TIPI6 Reserved, quiet.
- TIPI7 Sympathetic, warm.
- TIPI8 Disorganized, careless.
- TIPI9 Calm, emotionally stable.
- TIPI10 Conventional, uncreative.

The TIPI items were rated "I see myself as:" _ such that.

1. Disagree strongly.
2. Disagree moderately.
3. Disagree a little.
4. Neither agree nor disagree.
5. Agree a little.
6. Agree moderately.
7. Agree strongly.

The following items were presented as a checklist and subjects were instructed "In the grid below, check all the words whose definitions you are sure you know":

- VCL1 boat
- VCL2 incoherent
- VCL3 pallid
- VCL4 robot
- VCL5 audible
- VCL6 cuivocal
- VCL7 paucity
- VCL8 epistemology

- VCL9 florted
- VCL10 decide
- VCL11 pastiche
- VCL12 verdid
- VCL13 abysmal
- VCL14 lucid
- VCL15 betray
- VCL16 funny

A value of 1 is checked, 0 means unchecked. **The words at VCL6, VCL9, and VCL12 are not real words and can be used as a validity check.**

A bunch more questions were then asked:

education "How much education have you completed?", 1=Less than high school, 2=High school, 3=University degree, 4=Graduate degree

urban "What type of area did you live when you were a child?", 1=Rural (country side), 2=Suburban, 3=Urban (town, city)

gender "What is your gender?", 1=Male, 2=Female, 3=Other

engnat "Is English your native language?", 1=Yes, 2=No

age "How many years old are you?"

hand "What hand do you use to write with?", 1=Right, 2=Left, 3=Both

religion "What is your religion?", 1=Agnostic, 2=Atheist, 3=Buddhist, 4=Christian (Catholic), 5=Christian (Mormon), 6=Christian (Protestant), 7=Christian (Other), 8=Hindu, 9=Jewish, 10=Muslim, 11=Sikh, 12=Other

orientation "What is your sexual orientation?", 1=Heterosexual, 2=Bisexual, 3=Homosexual, 4=Asexual, 5=Other

race "What is your race?", 10=Asian, 20=Arab, 30=Black, 40=Indigenous Australian, 50=Native American, 60=White, 70=Other

voted "Have you voted in a national election in the past year?", 1=Yes, 2=No

married "What is your marital status?", 1=Never married, 2=Currently married, 3=Previously married

familysize "Including you, how many children did your mother have?"

major "If you attended a university, what was your major (e.g., "psychology", "English", "civil engineering")?"

The following values were derived from technical information:

country ISO country code of where the user connected from

screensize 1=device with small screen (phone, etc), 2=device with big screen (laptop, desktop, etc)

uniquenetworklocation 1=only one survey from user's specific network in dataset, 2=multiple surveys submitted from the network of this user (2 does not necessarily imply duplicate records for an individual, as it could be different students at a single school or different members of the same household; and even if 1 there still could be duplicate records from a single individual e.g. if they took it once on their Wi-Fi and once on their phone)

source how the user found the test, 1=from the front page of the site hosting the survey, 2=from google, 0=other or unknown

Modeling

After checking different models on the collected data we had the various accuracy for training and testing.

D/A/S	Sr. No.	Model Name	Training Accuracy (%)	Testing Accuracy (%)	Variance
D	**1**	**C & R**	**76.03**	**75.48**	**0.72**
	2	Neural Net	76.81	76.25	0.73
	3	CHAID	76.61	75.67	1.23
	4	Bayes Net	76.86	75.14	2.24
	5	C5.0	81.24	75.12	7.53
A	**1**	**C & R**	**86.87**	**86.15**	**0.83**
	2	CHAID	86.89	85.89	1.15
	3	Neural Net	87.23	85.78	1.66
	4	Bayes Net	86.92	84.68	2.58
	5	C5.0	88.63	85.73	3.27
S	**1**	**C & R**	**67.62**	**66.65**	**1.43**
	2	CHAID	67.32	66.25	1.59
	3	Neural Net	69.17	67.24	2.79
	4	Bayes Net	69.64	65.33	6.19
	5	C5.0	73.47	65.95	10.24

Evaluation

After checking different models accuracy, we find C&R model the most accurate. So will implement the same model using python code. So, we could deploy the code and run it on the live environment.

Future

We could deploy the created model using the python code and hosting it online so people would visit the hosted website and answer the questions and can have their feedback as soon as the finishes the survey.

Bibliography

- https://www.sciencedirect.com/science/article/pii/S1877050920309091
-