

CogniCare - Mental Health Classifier

P7: Manav Hemantkumar Patel, Manan Manojkumar Tiwari, Anshul Khairnar,
Department of Computer Science, North Carolina State University
Raleigh, NC 27695
(mhpatel4, mtiwari3, akhairn)@ncsu.edu

1 Data Set

Our project will classify different types of mental health conditions. It contains over 2000 surveys and each survey has a text and nine questions linked to that survey, complied by Arxiv

2 Project Idea

Our project goal is to classify and predict mental health problems in US college students. The project will help to classify mental health disorders which are impacting college students. We will compare the performance of our model using our testing dataset to determine its accuracy.

3 Software to Write

We will write our program in Python and evaluate the results using our major frameworks - Scikit-learn/ TensorFlow. We would also use Pandas, NumPy, and Matplotlib for data manipulation. We will use MySQL, MongoDB as our database.

4 Relevant Papers

- [1] Jetli Chung, Jason Teo, "Mental Health Prediction Using Machine Learning: Taxonomy, Applications, and Challenges", *Applied Computational Intelligence and Soft Computing*, vol. 2022, Article ID 9970363, 19 pages, 2022.
- [2] Pedrelli, P., Nyer, M., Yeung, A., Zulauf, C., & Wilens, T. (2015). College Students: Mental Health Problems and Treatment Considerations. *Academic psychiatry : the journal of the American Association of Directors of Psychiatric Residency Training and the Association for Academic Psychiatry*, 39(5), 503–511.
- [3] Sahlan, Fadhluddin & Mohammad Nizam, Faris Hamidi & Misrat, Muhammad & Zamzuri, Muhammad & Wani, Sharyar & Gulzar, Yonis. (2021). Prediction of Mental Health Among University Students. *International Journal on Perceptive and Cognitive Computing*. 7. 85-91.

5 Division of Work

Manan - Research relevant paper ideas to build models and reference papers for testing purposes.
Manav - Implement preprocessing of our dataset and divide dataset into training and test dataset.
Anshul- Use NLP techniques to tokenize the surveys and generate feature matrix for each survey.
All- Build different models to find the most accurate way to classify mental health problems.

6 Midterm Milestone

We will implement the preprocessing of our dataset and divide the dataset into training and test datasets. We will utilize NLP techniques to analyze the textual content of the surveys. We will also combine NLP features to create a feature matrix for each survey. We plan to begin training multiple ML models to classify mental health conditions based on features.