Allen Manoj **(**190501015**)**
Deepak K V **(**190501030**)**
Janani K **(**190501045**)**

Batch No   : **08**
Name of the Guide : **Ms. V. Radha**
Date of Review   : **10/11/2022**
Domain   : **Machine Learning**

TABLE OF CONTENTS

1. Problem Statement
2. Abstract
3. Issues and Challenges
4. Detailed Architecture diagram
5. System Requirements and Tools
6. Modules & its Description
7. Implementation Screenshots
8. References
9. Future Works

DETECTION, PREDICTION, AND ANALYSIS OF MENTAL WELL-BEING USING MACHINE LEARNING

Mental health issues in people have shown drastic increases affecting almost all age groups. Most of these issues remain uncured and neglected due to lack of awareness and being considered a taboo concept in our society. To be aware of susceptibilities towards mental health, early detection is an essential procedure.

Abstract

Mental health is a veritable topic of concern in most parts of the world. Mental disorders refer to a wide range of health conditions that affect your mood, thinking, and behavior which include depression, anxiety disorders, schizophrenia, eating disorders, and addictive behaviors. The increase in mental health problems and the need for effective medical health care have led to an investigation of machine learning that can be applied to mental health problems. Our software is proposed to help build understanding of the user's mental health, and prevent mishaps. Our software conducts a survey and the collected information is used as test dataset in our machine learning model for prediction. The prediction will be displayed with respect to their answers in the survey stating the condition of their mental health. Our proposed solution is expected to provide **80.423%** accuracy. The **end goal** of our software is to create awareness for mentally unwell people before they reach vulnerability.

Since most people with **mental illness** do not reach out for help on their own considering what people say, they tend to keep everything to themselves and restrict their circle. Social exposure is confined which makes it harder to cure the problem.

This **issue can be solved** with our project solution, as it is easily accessible by anyone and we maintain the confidentiality of every user's results. Challenges in the development would include preprocessing of our dataset with regards to working aspects of the user's response.

Architecture Diagram
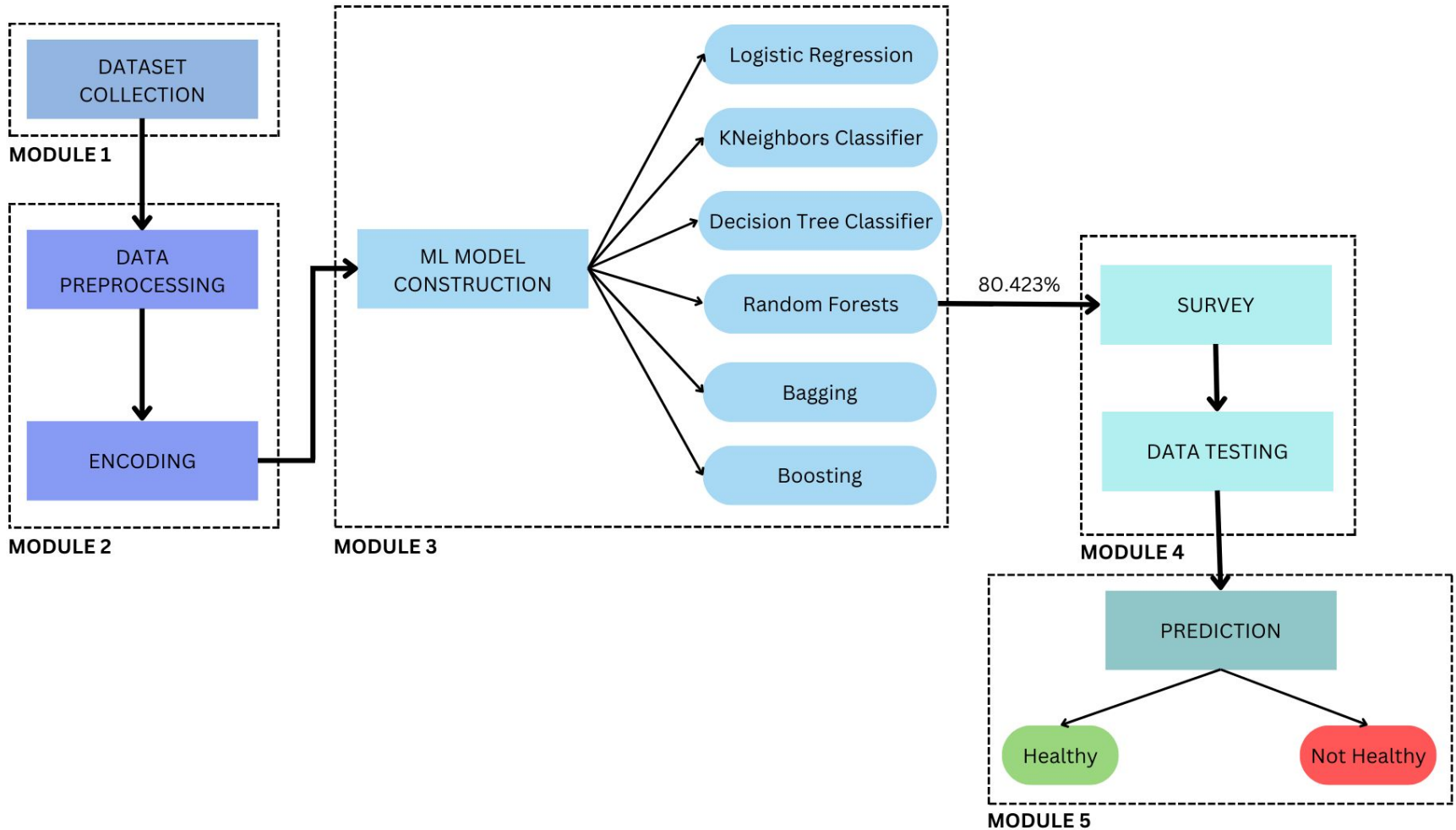
Architecture diagram for Mental Health Prediction using Machine Learning

Software Requirement:

It denotes the requirement of the software specifications. The requirement of the Software to function efficiently in the given Hardware are as follows:

- **Operating Systems**   :   Microsoft Windows 7 or later, MacOS Montenary
- **Python Version**   :   3.8.3
- **Development Tools**   :   Anaconda Framework
- **Compatible Tools**   :   Microsoft Visual Studio Code
- **Python Libraries**   :   Pandas, NumPy, Scikit-Learn, Streamlit

Hardware Requirement

It denotes the requirement of the hardware devices. The requirement of the Hardware Devices to function optimally for the given task is that as follows.

- **Random Access Memory**   :   8GB Minimum
- **Processors**   :   Intel Core i3 Processor or above
- **Microphone**   :   50 Hz - 15 kHz

The modules and their description for this project is as follows:

- **MODULE - I**   :   Data Collection & Loading

- **MODULE - II**   :   Data cleaning and encoding

- **MODULE - III**   :   Model Construction Phase

- **MODULE - IV**   :   Data Entry

- **MODULE - V**   :   Predictions

MODULE 1:

The below used datasets measure attitudes towards mental health and frequency of mental health disorders in a tech workplace.

- OSMI Mental Health in Tech Survey 2016

- Survey on Mental Health in the Tech Workplace in 2014

- Kaggle.com

- data.world

In the data collection phase our primary focus is to gather information on surveys of people with various personalities.

MODULE 2:

The dataset is being processed to clean NaN values and encode it.

Missing data, assigning default value and creation of list containing the following attributes:

```python
# Create lists by data type
intFeatures = ['Age']
stringFeatures = ['Gender', 'Country', 'self_employed', 'family_history', 'treatment', 'work_interfere',
                  'no_employees', 'remote_work', 'tech_company', 'anonymity', 'leave', 'mental_health_consequence',
                  'phys_health_consequence', 'coworkers', 'supervisor', 'mental_health_interview', 'phys_health_interview',
                  'mental_vs_physical', 'obs_consequence', 'benefits', 'care_options', 'wellness_program',
                  'seek_help']
floatFeatures = []
```

	Total NULL
Age	0
Gender	0
obs_consequence	0
mental_vs_physical	0
phys_health_interview	0
mental_health_interview	0
supervisor	0
coworkers	0
phys_health_consequence	0
mental_health_consequence	0
leave	0
anonymity	0
seek_help	0
wellness_program	0
care_options	0
benefits	0
tech_company	0
remote_work	0
no_employees	0
work_interfere	0
treatment	0
family_history	0
self_employed	0
age_range	0

Modules & its Description

ENCODING

```
label_Age [18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 53, 54, 55, 56, 57, 58, 60, 61, 62, 65, 72]
label_Gender ['female', 'male', 'trans']
label_Country ['Australia', 'Austria', 'Belgium', 'Bosnia and Herzegovina', 'Brazil', 'Bulgaria', 'Canada', 'China', 'Colombia', 'Costa Rica', 'Croatia', 'Czech Republic', 'Denmark', 'Finland',
label_self_employed ['No', 'Yes']
label_family_history ['No', 'Yes']
label_treatment ['No', 'Yes']
label_work_interfere ["Don't know", 'Never', 'Often', 'Rarely', 'Sometimes']
label_no_employees ['1-5', '100-500', '26-100', '500-1000', '6-25', 'More than 1000']
label_remote_work ['No', 'Yes']
label_tech_company ['No', 'Yes']
label_benefits ["Don't know", 'No', 'Yes']
label_care_options ['No', 'Not sure', 'Yes']
label_wellness_program ["Don't know", 'No', 'Yes']
label_seek_help ["Don't know", 'No', 'Yes']
label_anonymity ["Don't know", 'No', 'Yes']
label_leave ["Don't know", 'Somewhat difficult', 'Somewhat easy', 'Very difficult', 'Very easy']
label_mental_health_consequence ['Maybe', 'No', 'Yes']
label_phys_health_consequence ['Maybe', 'No', 'Yes']
label_coworkers ['No', 'Some of them', 'Yes']
label_supervisor ['No', 'Some of them', 'Yes']
label_mental_health_interview ['Maybe', 'No', 'Yes']
label_phys_health_interview ['Maybe', 'No', 'Yes']
label_mental_vs_physical ["Don't know", 'No', 'Yes']
label_obs_consequence ['No', 'Yes']
label_age_range ['0-20', '21-30', '31-65', '66-100']
```
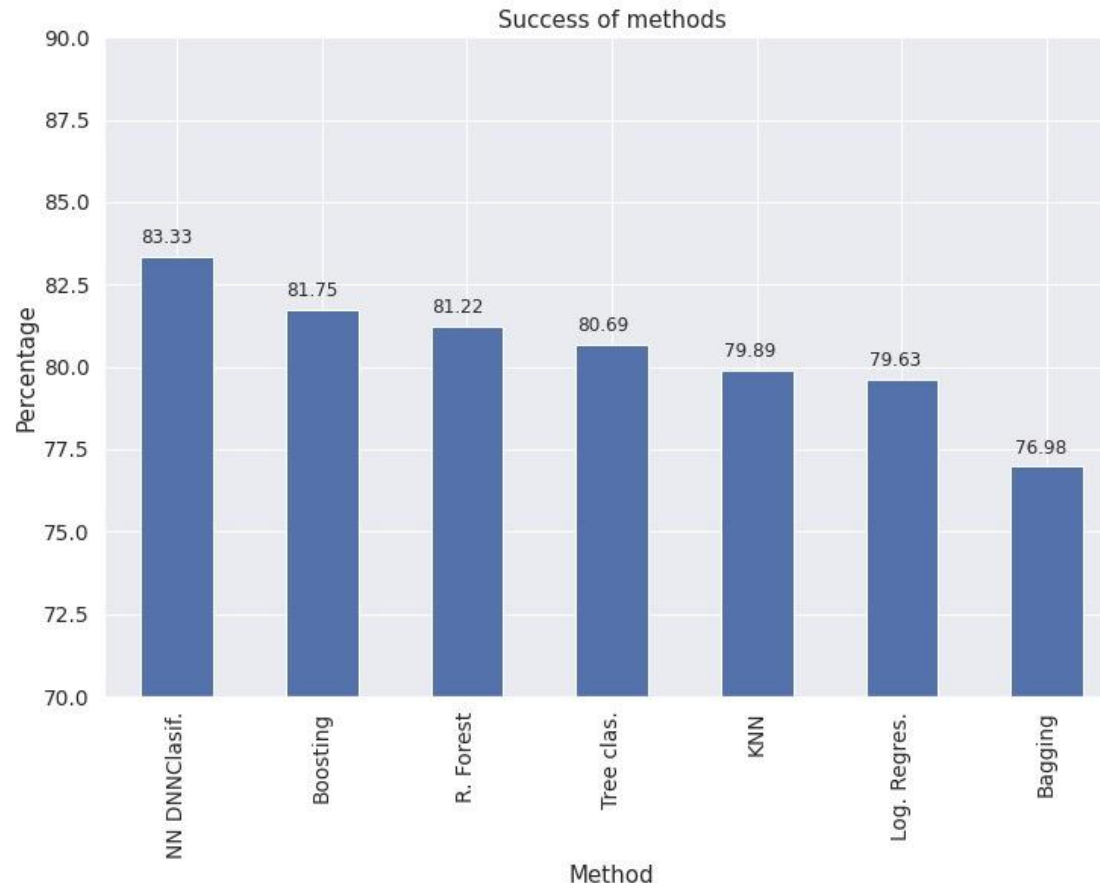

Feature importances
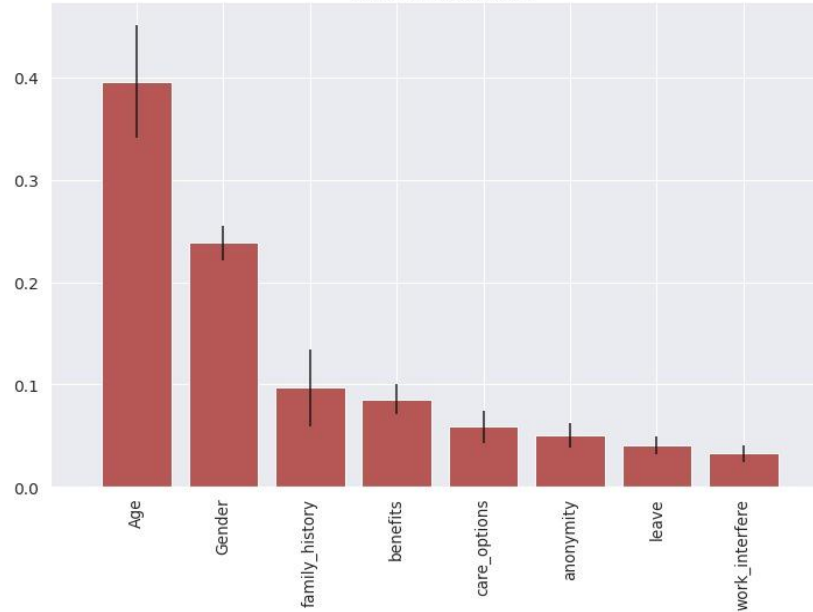
svce.ac.in

MODULE 3:

Model is evaluated and tuned by Logistic regression, KNeighbors Classifier, Decision Tree classifier, Random Forests, Bagging, and Boosting.

Success of methods

Modules & its Description

Accuracy of the model has been determined as **80.423%** using Random forests algorithm.

```
(base) Allens-MacBook-Pro:treat code allen$ python3 model.py
label_Age [18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 53, 54, 55, 56, 57, 58, 60, 61, 62, 65, 72]
label_Gender ['female', 'male', 'trans']
label_Country ['Australia', 'Austria', 'Belgium', 'Bosnia and Herzegovina', 'Brazil', 'Bulgaria', 'Canada', 'China', 'Colombia', 'Costa Rica', 'Croatia', 'Czech Republic', 'Denmark', 'Finland', 'France', 'Georgia', 'Germany', 'Greece', 'Hungary', 'India', 'Ireland', 'Israel', 'Italy', 'Japan', 'Latvia', 'Mexico', 'Moldova', 'Netherlands', 'New Zealand', 'Nigeria', 'Norway', 'Philippines', 'Poland', 'Portugal', 'Romania', 'Russia', 'Singapore', 'Slovenia', 'South Africa', 'Spain', 'Sweden', 'Switzerland', 'Thailand', 'United Kingdom', 'United States', 'Uruguay', 'Zimbabwe']
label_self_employed ['No', 'Yes']
label_family_history ['No', 'Yes']
label_treatment ['No', 'Yes']
label_work_interfere ["Don't know", 'Never', 'Often', 'Rarely', 'Sometimes']
label_no_employees ['1-5', '100-500', '26-100', '500-1000', '6-25', 'More than 1000']
label_remote_work ['No', 'Yes']
label_tech_company ['No', 'Yes']
label_benefits ["Don't know", 'No', 'Yes']
label_care_options ['No', 'Not sure', 'Yes']
label_wellness_program ["Don't know", 'No', 'Yes']
label_seek_help ["Don't know", 'No', 'Yes']
label_anonymity ["Don't know", 'No', 'Yes']
label_leave ["Don't know", 'Somewhat difficult', 'Somewhat easy', 'Very difficult', 'Very easy']
label_mental_health_consequence ['Maybe', 'No', 'Yes']
label_phys_health_consequence ['Maybe', 'No', 'Yes']
label_coworkers ['No', 'Some of them', 'Yes']
label_supervisor ['No', 'Some of them', 'Yes']
label_mental_health_interview ['Maybe', 'No', 'Yes']
label_phys_health_interview ['Maybe', 'No', 'Yes']
label_mental_vs_physical ["Don't know", 'No', 'Yes']
label_obs_consequence ['No', 'Yes']
label_age_range ['0-20', '21-30', '31-65', '66-100']
ACCURACY OF THE RandomForestClassifier MODEL:  80.42328042328042
```

svce.ac.in

MODULE 4:

Data Entry is done on the website of the application which includes the collection of information (test data) from the end user.

MODULE 5:

Creating input functions, defining feature columns, Instantiating an Estimator, Evaluating the trained model, Making predictions (inferring) from the trained model, and Creating predictions on test set are done using Neural Networks.
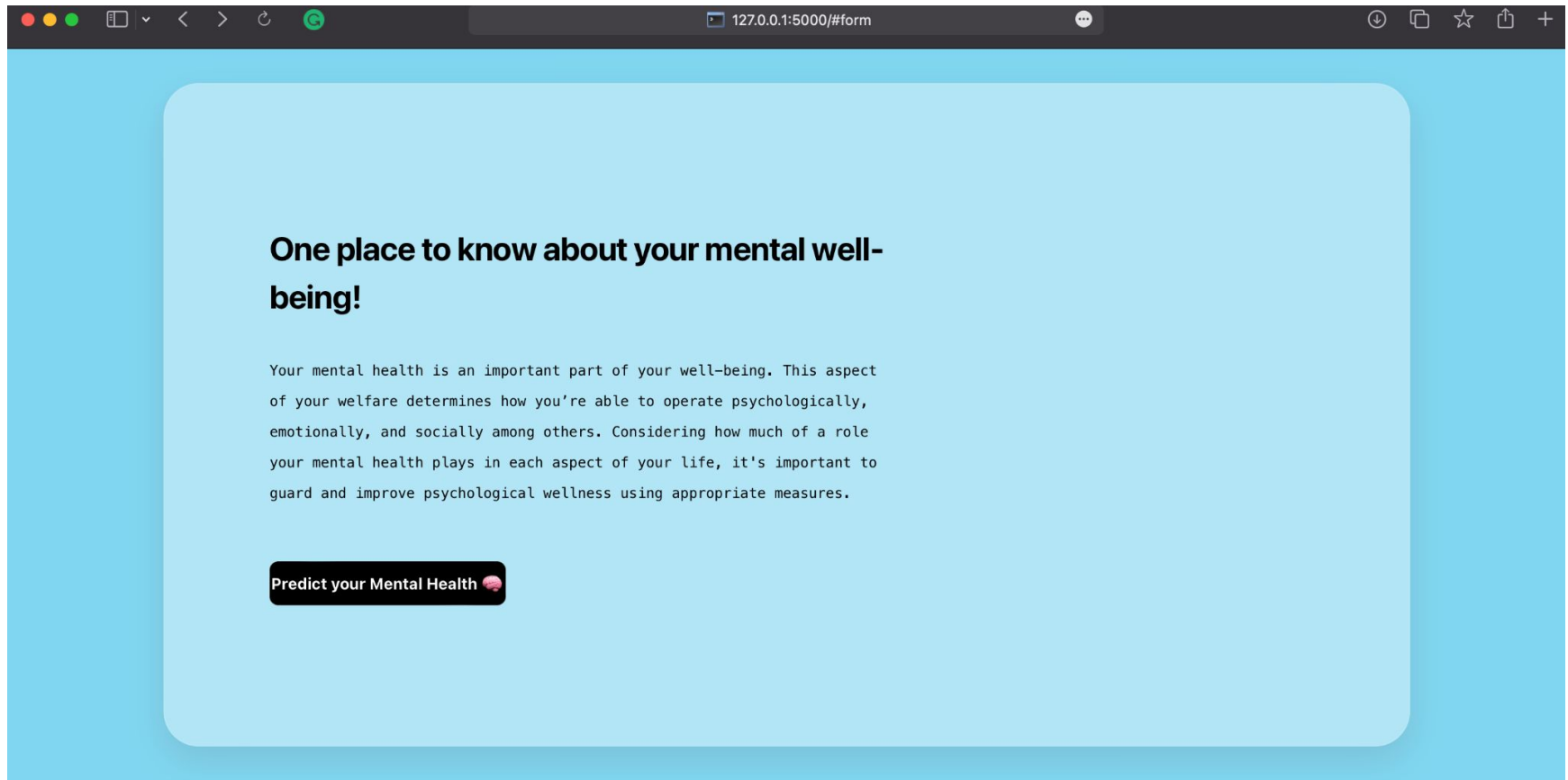
Evaluate the trained model

```
1   # Evaluate the model.
2 ∨ eval_result = model.evaluate(
3         input_fn=lambda:eval_input_fn(X_test, y_test, batch_size))
4
5   print('\nTest set accuracy: {accuracy:0.2f}\n'.format(**eval_result))
6
7   #Data for final graph
8   accuracy = eval_result['accuracy'] * 100
9   methodDict['NN DNNClasif.'] = accuracy
```

```
Test set accuracy: 0.80
```

Implementation Screenshots

FRONT END

svce.ac.in

Mental Health Predictor

Gender

[Male]

Are you self-employed?

[No]

Do you have a family history of mental illness?

[No]

If you have a mental health condition, do you feel that it interferes with your work?

[Never]

Number of employees in the organisation?

[More than 1000]

Do you work remotely (outside of an office) at least 50% of the time?

[No]

Is your employer primarily a tech company/organization?

[Yes]

Would you bring up a physical health issue with a potential employer in an interview?

[Maybe]

Do you feel that your employer takes mental health as seriously as physical health?

[Don't Know]

Have you heard of or observed negative consequences for coworkers with mental health conditions in your workplace?

[No]

Does your employer provide mental health benefits?

[Yes]

Do you know the options for mental health care your employer provides?

[Not sure]

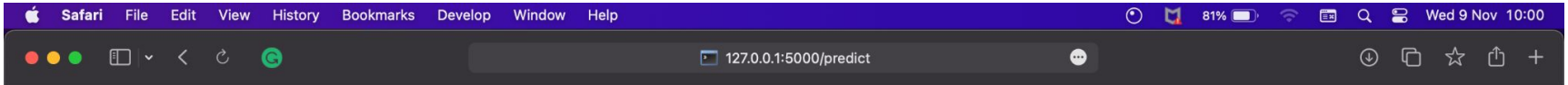**Has your employer ever discussed mental health as part of an employee wellness program?**

[Don't Know]

Does your employer provide resources to learn more about mental health issues and how to seek help?

[Don't Know]

[Predict]

SRI VENKATESWARA COLLEGE OF ENGINEERING
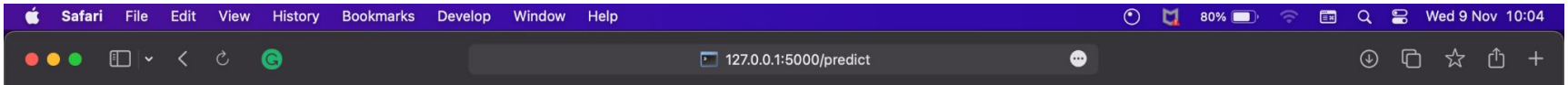
OUTPUT

PREDICTION RESULT

You are healthy

Your mental health is alright

PREDICTION RESULT

You are not healthy

You need medical guidance

Future Works

- Based on the prediction, we can analyse the intensity of mental health.
- Visualizations can be provided to develop each skill set of individuals.
- Initial help such as video, seminar, and event links related to mental wellbeing can be provided by our software.
- The software can maintain a mental health tracking technique which monitors the scores for a period of time incorporating the use of cloud technology for data storage.
- Based on the results, the individuals will be provided with helpline guidance.
- Psychiatrist and psychologist consultations can be booked through the software.
- We can develop the software as a service by supplying mental health related medication which are prescribed by the doctors.

References:

1. R. A. Rahman, K. Omar, S. A. Mohd Noah, M. S. N. M. Danuri and M. A. Al-Garadi, "Application of Machine Learning Methods in Mental Health Detection: A Systematic Review," in IEEE Access

2. Mallick, Sumitra, and Mrutyunjaya Panda. "Mental Health Illness Disease Prediction Using Most Common Feature in Tech Survey Dataset.",International Journal of Engineering Research and Applications

3. O. Oyebode, F. Alqahtani and R. Orji, "Using Machine Learning and Thematic Analysis Methods to Evaluate Mental Health Apps Based on User Reviews," in IEEE Access

Thank you