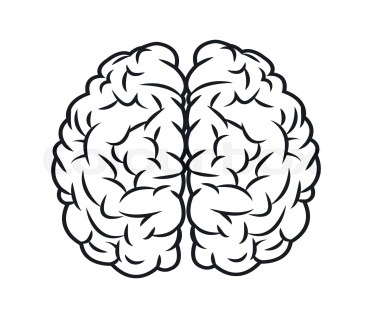
   
 ***Menoufia University***

***Faculty of Computers & information***

|  |
| --- |
| *Depression Detection* |



***Prepared by:***

**Walaa Khaled Ibrahim Eliwa**

**Fatma Ali Ali Elshihna**

**Sohila Essam Mohamed Salem**

**Mariam Abdeen Abdelmaksoud mazroa**

**Salma Ayman Mohamed Lashien**

**Supervised by:**

­

**Dr /Hyame Mousa Elhussine**

**Bioinformatics Department**

**4th Year Graduation Project 2022-2023**

|  |  |
| --- | --- |
| *TABLE OF CONTENTS* |  |
| ***03 Team Members***  ***04 Acknowledgment***  ***05 Abstract***  ***06 list OF Contents*** | |
| ***08 List of Figures***  ***101 Conclusion***  ***102 References*** | |
|  | |

# Team Members

* ***Fatma Ali Ali Elshihna***
* ***Walaa Khaled Ibrahim Eliwa***
* ***Sohila Essam Mohamed Salem***
* ***Mariam Abdeen Abdelmaksoud mazroa***
* ***Salma Ayman Mohamed Lashine***

# Acknowledgment:

We express our heartfelt gratitude to Allah for His unwavering support and guidance throughout our academic journey.

We extend our sincere appreciation to **Menoufia** **University** for providing us with a conducive learning environment and equipping us with the necessary skills and knowledge.

Our academic supervisor, “Dr/Hyame Mousa Elhussine”, deserves special recognition for his invaluable mentorship and encouragement that propelled us towards excellence.

We acknowledge the contributions of all the professors and assistant staff that have played a significant role in shaping our academic growth.

We, also, give many thanks to Eng: Hussien Abd Elraouf who helped us a lot by giving us advice and guiding us through the whole project.

We would like to express our gratitude to our families and friends for their constant support and encouragement throughout our academic journey. Their love and motivation have been a source of strength for us, and we are forever grateful for their unwavering support.

Finally, I would like to thank everyone who has helped and has been on my side forever.

# Abstract:

Depression is a common mental health disorder that affects millions of people worldwide. Early detection and treatment of depression can significantly improve the quality of life for those affected. It can have a significant impact on an individual's quality of life, leading to social isolation, decreased productivity, and even suicide.

In recent years, there has been growing interest in using technology to detect depression, including analyzing text and facial expressions. This project presents an overview of the current state-of-the-art in depression detection using these modalities.

This project reviews the literature on text analysis techniques, such as sentiment analysis and discusses their potential for detecting depression from written communication their potential for detecting depression from written communication. We also examine recent advances in facial expression analysis, including machine learning algorithms that can identify subtle changes in facial expressions associated with depression .So, We aim to develop a depression detection tool that analyze data from text patterns and Facial expression. The tool will provide clinicians with an objective measure of depression severity and aid in the early identification of individuals at risk for developing depression.

The benefits of this project include, reduced healthcare costs associated with untreated depression, and increased awareness and understanding of mental health disorders in society.

Overall, this project has the potential to make a significant impact on the lives of individuals affected by depression and their families.

* List of Contents

**Chapter One: Introduction10**

1.1. Introduction11

1.2. Problem Definition11

1.3. Project Importance and Objectives12

1.4 motivations14

1.5. Traditional solutions15

1.6. Technological solutions16

1.7. Expectations and Achievements17

**Chapter Two: Dataset and Preprocessing**21

2.1. Dataset22

2.1.1 Image Dataset22

2.1.2 Text Dataset22

2.2. Preprocessing23

2.2.1 Median equalization-23

2.2.2 ImageDataGenerator24

2.2.3 clean text25

**Chapter three: System Overview26**

3.1. Competitors Applications27

3.2. Depression & Not Depression System27

3.3. System Architecture27

3. 4. Functional Requirements29

3. 5. Non-Functional Requirement-30

**Chapter four: ML&DL31**

4.1. Abstract32

4.2. Introduction33

4.3. Two different Tests to classify34

4.4. Machine learning techniques35

4.4.1. Support Vector Machine35

4.4.2. K-Nearest Neighbors35

4.4.3. Random forest36

4.4.4. Logistic regression36

4.4.5. Naïve Bayes37

4.5. Deep learning techniques38

4.5.1 Convolutional Neural Network “CNN”-39

4.5.2Transfer learning techniques42

VGG-16 Results43

VGG-19 Results43

**Chapter five: Discussion and results48**

5. 1. Discussion-49

5.2. Result57

5.2.1. Result of ML &DL57

**Chapter six: Web Application83**

6.1. Home page84

6.2. Sign-up Page85

6.3. log-in Page87

6.4. Test image Page89

6.5. Test Text Page91

6.6. Depression (blog) pages92

6.6.1 All about Depression page93

6.6.2 Depression affects Relation &Work page93

6.6.3 Hotline Numbers page 94

6.6.4 Your Questions Answered about Depression page 94

6.6.5 Treatments page95

6.6.6 Phrases page96

6.6.7 About US page97

**Chapter seven: analysis and Diagrams98**

7. 1. Sequence Diagrams99

7.2. Activity Diagrams100

* ***List of Figures***

[1.1. **Figure1**: First test of classification 34](#_Toc140427094)

[1.2. **Figure2**: second test of classification 3](#_Toc140427095)4

[1.3. **Figure3**: CNN for image 39](#_Toc140427096)

[1.4. **Figure4:** CNN for text- 41](#_Toc140427097)

[1.5 Figure5: Transfer Learning Techniques 42](#_Toc140427098)

[1.6. Figure6: VGG-16 45](#_Toc140427099)

[1.7. Figure7: VGG-19 46](#_Toc140427100)

[1.8. Figure8: SVM Images ROC graph for train and test 58](#_Toc140427101)

[1.9. Figure9: KNN Images ROC graph for train and test 60](#_Toc140427102)

[1.10. Figure10: RF Images ROC graph for train and test 62](#_Toc140427103)

[1.11. Figure11: LR Images ROC graph for train and test 64](#_Toc140427104)

[1.12. Figure12: NB Images ROC graph for train and test 66](#_Toc140427105)

[1.13 Figure 13: CNN Images Accuracy and Loss graph 68](#_Toc140427106)

[1.14. Figure14: vgg16 Images ROC graph for train and test- 71](#_Toc140427107)

[1.15. Figure15: vgg19 Images ROC graph for train and test 74](#_Toc140427108)

[1.16. Figure16: KNN Text ROC graph for train and test 77](#_Toc140427109)

[1.17. Figure17: RF Text ROC graph for train and test 78](#_Toc140427110)

[1.18. Figure18: LR Text ROC graph for train and test 8](#_Toc140427111)0

[1.19. Figure19: NB Text ROC graph for train and test 81](#_Toc140427112)

[1.20 Figure 20: CNN Text Accuracy and Loss graph 82](#_Toc140427113)

[1.21. Figure21: Home page 84](#_Toc140427114)

[1.22. Figure22: Sign-up Page 86](#_Toc140427115)

[1.23 Figure 23: Log-in Page 88](#_Toc140427116)

1.24. Figure24: Test image Page89

1.25. Figure25: Test Text Page91

1.26. Figure26: depression (blog) pages 92

1.27. Figure27: All about Depression page 93

1.28. Figure28: Depression affects Relation &Work page93

1.29. Figure29: Hotline Numbers page 94

1.30. Figure03: Your Questions Answered about Depression page 94

1.31. Figure03: Treatments page 95

1.32. Figure03: phrases page 96

1.33. Figure00: About Us page 97

***Chapter One***

# *INTRODUCTION*



# Introduction

In this chapter, we will discuss the problem at hand and highlight its significance and motivation for being addressed. We will also examine previous attempts to solve this problem and identify their limitations. Furthermore, we will present our proposed application and its development process. Finally, we will provide definitions for the key terms used in this chapter.

# 1.2. Problem definition

Depression is a common mental health disorder that affects millions of people worldwide. It is characterized by persistent feelings of sadness, hopelessness, and loss of interest in activities that were once enjoyable. Depression can have a significant impact on an individual's quality of life, relationships, and work productivity.

Depression is a complex disorder that can be caused by various factors such as genetics, environmental factors, and life events. It can also co-occur with other mental health disorders such as anxiety and substance abuse.

# 1.3. Project Importance and Objectives

The importance of addressing depression lies in its prevalence and impact on individuals and society as a whole. According to the World Health Organization (WHO), depression is the leading cause of disability worldwide. It affects people of all ages, genders, and backgrounds and it can lead to severe consequences such as suicide if left untreated. Therefore, it is crucial to develop reliable and accurate methods for detecting depression in its early stages.

A web system to predict depression detection is important for several reasons:

1. *Early detection:* A web system can help in detecting depression at an early stage, which is crucial for effective treatment and management of the condition.
2. **Accessibility:** A web-based system can be accessed from anywhere, making it easier for people to get help and early detection for depression.
3. ***Privacy***: Many people may feel uncomfortable discussing their mental health issues with others. A web-based system provides a level of privacy that can encourage more people to seek help.
4. ***Cost-effective:*** A web-based system can be more cost-effective than traditional methods of depression detection, such as face-to-face consultations with mental health professionals.
5. ***Data analysis:*** A web-based system can collect data on a large scale, which can be used to analyze trends and patterns in depression rates and treatment outcomes.
6. ***Reduced stigma:*** By detecting depression early, we can reduce the stigma associated with mental health conditions by promoting awareness and understanding of the condition.
7. ***Improved productivity:*** Depression can significantly impact a person's ability to function at work or school. Early detection and treatment can help individuals maintain their productivity and reduce absenteeism.

Overall, a web system to predict depression detection is an important tool in the fight against mental illness. It can help improve access to care, reduce stigma, and ultimately save lives.

# 1.4. Motivation

Depression is a serious mental health condition that affects millions of people worldwide. It can lead to a range of negative outcomes, including impaired functioning, decreased quality of life, and even suicide. Early detection and treatment of depression are crucial in preventing these negative outcomes.

Depression detection is important because it allows individuals to receive timely and appropriate treatment. This can include therapy, medication, or a combination of both. Early intervention can help individuals manage their symptoms and improve their overall well-being.

Additionally, depression detection can help reduce the stigma associated with mental illness. By raising awareness about depression and its symptoms, individuals may be more likely to seek help when they need it.

Overall, the motivation behind depression detection is to improve the lives of those who are affected by this condition. By identifying depression early and providing appropriate treatment, we can help individual’s lead happier and healthier lives.

# 1.5. Traditional solutions

Previous solutions to this problem include self-assessment questionnaires, clinical interviews, and psychological assessments. However, these solutions have their limitations. Self-assessment questionnaires may not be accurate, clinical interviews can be time-consuming and expensive, and psychological assessments may not be accessible to everyone. Previous Solutions to the Problem There are various treatments available for depression such as medication, psychotherapy, and lifestyle changes. However, these treatments have limitations.

Medication can have side effects and may not be effective for everyone. Psychotherapy requires time commitment and may not be accessible or affordable for everyone. Lifestyle changes such as exercise and healthy eating can be challenging to implement for individuals with depression.

Moreover, there is a stigma associated with seeking help for mental health disorders which can prevent individuals from seeking treatment.

There are various ways to detect depression, including self-assessment questionnaires, clinical interviews, and psychological assessments. Some commonly used tools for depression screening include the Patient Health Questionnaire (PHQ-9), Beck Depression Inventory (BDI), and Hamilton Depression Rating Scale (HDRS). It is important to note that these tools are not diagnostic but can help identify individuals who may need further evaluation by a mental health professional. Additionally, machine learning algorithms have been developed to analyze speech patterns and facial expressions to detect signs of depression.

# 1.6. Technological solutions

There are several technological solutions for detecting depression using text and image datasets. Some of them are:

1. ***Image Analysis***: Image analysis techniques can be used to detect facial expressions that are associated with depression, such as sadness or hopelessness. Machine learning algorithms can be trained on a dataset of images labeled as depressed or non-depressed to identify patterns that distinguish between the two.

2. ***Deep Learning***: Deep learning is a subset of machine learning that uses neural networks to learn from large datasets. It has been used to develop models for detecting depression from text and image data. For example, deep learning models have been trained on text dataset to predict depression symptoms with high accuracy.

3. ***Sentiment Analysis***: Sentiment analysis is a technique for identifying the emotional tone of text data. It can be used to detect signs of depression in written communication such as emails or chat logs. Sentiment analysis algorithms can identify words or phrases that indicate negative emotions and use them to predict the likelihood of depression.

Overall, these technological solutions offer promising avenues for detecting depression using text and image datasets, which could help improve early diagnosis and treatment outcomes for individuals with this condition.

# 1.6. Expectations and Achievements

To complete the project, we first thoroughly studied and comprehended the problem at hand, recognizing its impact on society and daily life. We conducted an analysis of the necessity for a website that can detect depression with ease. Based on this, we researched and selected appropriate technologies to develop our project. Considering that our project is a national one, we opted for a web application as our platform of choice.

## Expectations:

* Increased awareness and education about depression
* Early detection and intervention for individuals at risk of depression
* Improved access to mental health resources and support
* Reduced stigma surrounding mental health issues

## Achievements:

* Increased screening and identification of individuals with depression
* Improved accuracy in identifying symptoms of depression
* Increased access to mental health resources and support for those in need
* Reduced stigma surrounding mental health issues through increased education and awareness

Overall, creating a website to detect depression has the potential to make a significant impact on the lives of individuals struggling with mental health issues. By increasing awareness, improving early detection, and reducing stigma, we can work towards a society that prioritizes mental wellness.

After studying the web application we found that there are frameworks such as Angular which helped us to build a good web application.

In the first stage, it was the work of analyzing and understanding the complete problem and how to solve it through the application and avoid the problems of the previous solutions.

On the other hand, our ***objectives*** are:-

* Design using an easy interface to use by users.
* Applying an easy way to find detect depression.
* Finding the best machine learning and deep learning algorithms that will help us.

Finally, we started working on the back-end side of the web application using Flask framework so that the project can work properly.

# *Table 1*

Shows some definitions about the terms mentioned above

|  |  |
| --- | --- |
| ***Term*** | ***Definition*** |
| ***Face Expression*** | Facial expression refers to the various changes and movements of the face that convey feelings, attitudes, and intentions.  Facial expressions can convey a wide range of emotions such as happiness or sadness |
| ***Text pattern*** | They are words that a person enters and through which his condition is identified if he suffers from depression or not |
| ***Machine learning*** | Machine learning is a branch of artificial intelligence concerned with developing systems and technologies that enable computers to learn and adapt from data rather than being explicitly programmed |
|  | |
| ***website*** | A group of World Wide Web pages usually containing hyperlinks to each other and made available online by an individual, company, government, or organization |
| ***flask*** | Flask is a lightweight web framework written in Python that allows developers to build web applications quickly and efficiently |
| ***Back end*** | Refers to any part of a website or software program that users do not see. It contrasts with the frontend, which refers to a program's or website's user interface. |

***Chapter Two***

# *Dataset and Prepressing*



# 2.1 Dataset

We collected two types of data **image** and **text**:

## 2.1.1 Image Dataset

Image datasets for people face to detect expression the data contain train file that has 7164 images for Non Depressed ,and 4938 images for Depressed, and validation file that has 1825 images for Non Depressed(), and 1139 images for Depressed. This data was collected in 2021, average image size of 48\*48 pixels. The images are in JPG format

## 2.1.2 Text Dataset

Text dataset, this dataset is a collection of posts from the "Suicide Watch" and "depression" subreddits of the Reddit platform. The posts are collected using Push shift API. All posts that were made to "Suicide Watch" from Dec 16, 2008(creation) till Jan 2, 2021, were collected while "depression" posts were collected from Jan 1, 2009, to Jan 2, 2021. All posts collected from Suicide Watch are labeled as suicide, while posts collected from the depression subreddit are labeled as

# 2.2 Pre-processing techniques

## 2.2.1 Median equalization

Median equalization, also known as histogram equalization, is a technique used in image processing and computer vision to enhance the contrast and improve the overall appearance of an image. The goal of median equalization is to redistribute the pixel values of an image in such a way that the resulting histogram becomes more evenly distributed across the available range of pixel values.

The process of median equalization involves the following steps:

1. Compute the histogram of the input image, which represents the frequency distribution of pixel values.

2. Calculate the cumulative distribution function (**CDF**) of the histogram. The **CDF** provides information about the cumulative probability of pixel values occurring in the image.

3. Normalize the CDF to the desired output range (e.g., 0 to 255 for an 8-bit gray scale image) by mapping the minimum and maximum pixel values of the original image to the desired range.

4. Apply the CDF transformation to the input image by mapping each pixel value to its corresponding normalized value using the CDF lookup table.

The result of median equalization is an image with improved contrast and enhanced details. It can be particularly useful in situations where an image has uneven lighting conditions or a narrow range of pixel values, leading to a loss of visual information. Median equalization helps to stretch the histogram of the image to cover a wider range, making it visually more appealing and easier to analyze or process further.

***2.2.2*** ImageDataGenerator

The Image Data Generator can be used for the following purposes:

Data Augmentation: Data augmentation is a technique used to increase the size and diversity of the training dataset by applying random transformations to the existing images. These transformations can include rotation, zooming, shifting, flipping, shearing, and more. By generating augmented images, the model can learn to be more robust to variations and improve its generalization ability.

**Real-Time Data Preprocessing**: The ImageDataGenerator can also be used to preprocess the image data in real-time during training. It provides options to rescale the pixel values to a specific range, apply feature-wise standardization or normalization, convert images to gray scale, or apply other custom preprocessing functions.

**Batch Generation**: The ImageDataGenerator can generate batches of augmented images on-the-fly, allowing for efficient memory utilization. It automatically loads and processes images in batches, which can be directly fed into the deep learning model during training.

By utilizing the ImageDataGenerator, researchers and practitioners can effectively enhance their image datasets, increase the diversity of training examples, and improve the performance and robustness of deep learning models for image-related tasks, such as image classification, object detection, and image segmentation.

## 2.2.3 Clean Text

The function "clean text” is function created by developers or data analysts to preprocess and clean textual data. It is used to remove unwanted characters, symbols, or noise from the text and perform various text cleaning operations.

Here are some common operations that the "clean text" function may include:

**Removing Punctuation:** This involves removing punctuation marks such as periods, commas, question marks, etc., which are not relevant to the analysis or modeling task.

**Converting to Lowercase:** This operation converts all the text to lowercase. It helps to standardize the text and avoid treating the same word with different cases as different entities.

**Removing Stop Words**: Stop words are commonly used words in a language (e.g., "a," "the," "is") that do not carry significant meaning and can be removed from the text to reduce noise

**Handling Special Characters**: Some text data may contain special characters, emesis, or non-standard symbols. The "clean text" function may include operations to handle or remove these special characters appropriately.

***Chapter Three***

# *System Overview*



# 3.1 Competitors Applications

There are numerous web applications available for use. In this analysis, we will highlight some of these applications, discuss their main features, and finally examine their advantages and disadvantages.

# 3.2 Depression & Not Depression system

With this developed web application, people who think they are suffering from depression, anyone can register via his e-mail address and take the first test by entering his photo to determine if he is happy or sad. If the result is sad, he will go to the second test. Enter a text to know if he is suffering from depression or not.

**3.3 System Architecture**

The system architecture is a system with different functions related to depression detection such as entering the patient a picture of his face to know facial expressions, and also entering text through it to determine whether he suffers from depression or not.

**It is important to note that depression is a complex condition, and detecting it solely through self-reporting or screening tools may not provide a definitive diagnosis. Professional assessment by qualified healthcare providers is crucial for accurate diagnosis and appropriate treatment planning so we do this system.**

## Table2:

|  |  |
| --- | --- |
| ***Project need*** | The main objectives of this project are to detect depression disease. |
| ***System Requirement*** | The depression website offers users easy and quick access to the system, with a user-friendly interface. The system includes features that allow users to   * upload photos, * enter text, * Learn about methods for preventing depression.   Additionally, the website provides a directory of doctors for booking appointments and communicating with them. |
| ***special issues and constraints*** | The project team requires a high level of security to safeguard against unauthorized access and protect sensitive information from unauthorized users. |
| **System value** | Our objective is to develop a system that enables a large number of patients to detect the disease and also to prevent its occurrence among people. |

**3.4 Functional Requirements:** Table 3

|  |  |
| --- | --- |
| ***Requirement Name*** | ***Requirement Description*** |
| ***Log in*** | This function will enable patients to login the system. |
| ***Change Password*** | This function will enable patients to change the password. |
| ***Insert Photo*** | This function will enable patients to insert photo to detect facial expression |
| ***Insert Text*** | This function will enable patients to insert text to identify feelings |
| ***guidance and information*** | This function will provide patients with comprehensive guidance and information to ensure they have access to the necessary knowledge and support. |

# 3.5 Non Functional Requirements

The non-functional requirements of the breast cancer website are described below

1. **Security:**

To ensure protection against unauthorized access, the system will implement a username and password authentication mechanism. If the entered username and password are incorrect, an error message will be displayed. Each user will be required to log in to the system by entering their unique username and password.

1. **Performance:**

The system is designed to provide a quick response time of few seconds when processing and verifying user-entered usernames and passwords.

***Chapter four***

# *ML&DL*



# Machine Learning

## 4.1 Abstract

The use of Machine Learning (ML) and Deep Learning (DL) techniques in detecting depression offers several important benefits:

Depression is a prevalent global disease, and early and accurate diagnosis plays a crucial role in rehabilitation and treatment. However, traditional detection methods face uncertainties. Machine Learning (ML) and Deep Learning (DL) techniques offer promising solutions for developing effective tools to aid physicians in early detection and diagnosis of depression. This research paper compares five widely used ML techniques: Support Vector Machine (SVM), Random Forest (RF), K-Nearest Neighbor (KNN), Logistic Regression (LR), and Naive Bayes (NB). Additionally, it compares five popular DL techniques: Convolutional Neural Network (CNN), Transfer Learning (VGG-16), and ( VGG-19). By evaluating these techniques, the study aims to determine their effectiveness in depression detection and diagnosis.

***4.2 Introduction:***

Depression is a mental health disorder characterized by persistent feelings of sadness, loss of interest or pleasure in activities, changes in appetite and sleep patterns, low energy, and difficulty concentrating. It is a complex condition that affects people of all ages and backgrounds, often significantly impacting their daily lives, relationships, and overall well-being.

Machine Learning (ML) and Deep Learning (DL) are branches of Artificial Intelligence (AI) that enable machines to learn and make predictions without explicit programming. They learn from large datasets and gain knowledge through experience. In recent years, ML methods have gained significant popularity in various fields, including the development of predictive models to facilitate informed decision-making.

In the context of depression research, ML and DL techniques can be utilized to analyze extensive datasets and uncover patterns that may not be apparent through traditional statistical approaches. By applying these techniques to depression data, researchers can identify meaningful associations, risk factors, and predictive markers.

ML algorithms can help classify individuals into different groups based on various characteristics and symptoms. They can also be used for early detection and diagnosis of depression by analyzing patterns in data such as behavioral patterns, . This can lead to more accurate and timely identification of individuals at risk, enabling early intervention and treatment.

DL in particular, with its deep neural networks, can learn intricate representations of data and extract complex features. This makes it well-suited for analyzing diverse data types, such as text or images, and detecting subtle patterns in the context of depression.

* 1. **Two different Tests to classify**

## 4.3.1 First Test:

Multiclass “2 classes into classification” to all data set is shown in next

***Figure1: First test of classification***

## 4.3.2 Second Test:

Depend on binary classification*:*

Show in Figure 2: Second approach of classification

“Non-Depressed” (non-patient) /“Depressed class “(patient).

If the result of first approach is depressed do second approach **text** test

***Figure 2: Second test of classification***

* 1. **Machine learning techniques**

Now we will discuss some of important machine learning techniques as: SVM, KNN, RF, LR and NB in brief.

**4.4.1 Support Vector Machine**

Support Vector Machine (SVM) is a supervised machine learning algorithm that is widely used for both classification and regression tasks. However, it is primarily applied to classification problems. In our study, we will employ SVM to analyze the previously filtered data and determine the optimal values for accuracy, precision, recall, and F1 score. For more detailed information on the performance of this model, please refer to the results section, specifically Table 1 to Table 5. These tables will provide insights into the effectiveness of the SVM algorithm in addressing the classification challenges at hand.

**4.4.2 K-Nearest Neighbor**

K-Nearest Neighbor (K-NN) is a simple and popular machine learning algorithm that falls under the category of supervised learning. It is based on the assumption of similarity between new and existing data points and assigns the new data point to the category that is most similar to the available categories. The K-NN algorithm stores all the available data points and classifies a new data point based on their similarity.

The K-NN algorithm is known for its simplicity and effectiveness in classification tasks. It is often referred to as a "lazy learner" algorithm because it does not learn from the training set immediately. Instead, it stores the dataset and performs the classification task when new data points need to be classified.

In our study, we have applied the K-NN algorithm to our dataset and evaluated its performance. To gain a deeper understanding of this model and its effectiveness in addressing the classification challenges, please refer to the results section, specifically Table 6 to Table 10. These tables provide detailed insights into the performance metrics such as accuracy, precision, recall, and F1 score, which can help assess the performance of the K-NN algorithm in our specific context.

**4.4.3 Random forest**

Random Forest is a classification algorithm that utilizes multiple decision trees on various subsets of the given dataset to improve the predictive accuracy. Unlike relying on a single decision tree, Random Forest takes predictions from each tree and predicts the final output based on the majority votes of predictions. The more trees in the forest, the higher the accuracy and it also prevents over fitting issues.

**4.4.4 Logistic regression**

Logistic regression is a popular Supervised Learning algorithm used for predicting categorical dependent variables based on independent variables. It predicts the outcome of a categorical variable, which can be either Yes or No, 0 or 1, true or false, etc. Instead of giving exact values of 0 and 1, it provides probabilistic values between 0 and 1.

**4.4.5 Naïve Bayes**

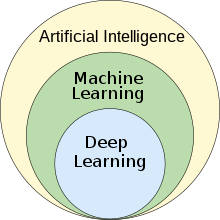
The Naïve Bayes algorithm is a popular supervised learning algorithm that is based on Bayes' theorem and is commonly used for solving classification problems. It is a probabilistic classifier, meaning it predicts the class membership of an object based on the probability of it belonging to a particular class.

Naïve Bayes is known for its simplicity and speed, making it a popular choice for classification tasks. It assumes that the presence of a particular feature in a class is independent of the presence of other features, which is referred to as the "naïve" assumption. Despite this simplifying assumption, Naïve Bayes often performs well and can be used for both binary and multi-class classifications.

# Deep Learning

## 4.5 Deep learning techniques

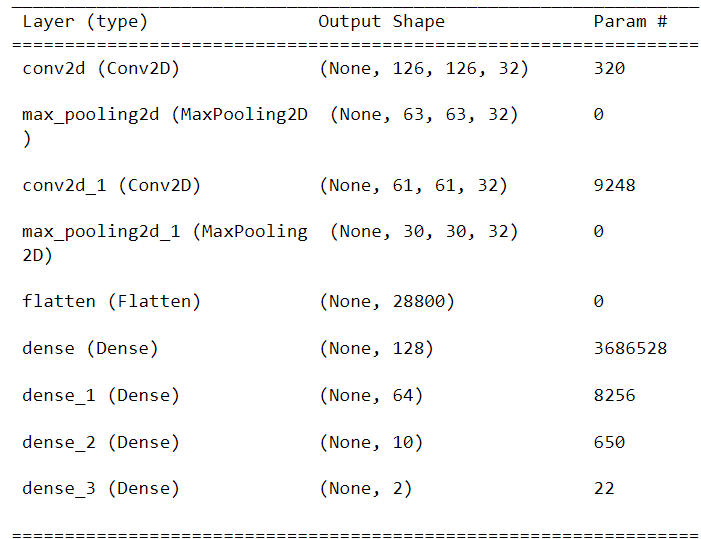
Deep learning is a type of machine learning that falls under the umbrella of artificial intelligence. It involves using multiple layers of non-linear processing units to extract and transform features, with each layer building upon the output of the previous one. Compared to traditional machine learning methods, deep learning models are capable of producing superior results. One key difference between deep learning and machine learning is that deep learning utilizes hidden layers for feature extraction. In this project, we will be working with a dataset consisting of ultrasound images and masks, training our model on the images to achieve optimal result.



***4.5.1 Convolutional Neural Network “CNN”***

***CNN for image:***

CNN, or Convolutional Neural Network, is a powerful deep learning model that excels in processing data with a grid-like structure, such as images. It is specifically designed to learn and identify spatial patterns and features in a hierarchical manner, starting from low-level features to high-level patterns. This adaptive learning approach allows CNNs to automatically extract meaningful information from complex data sets and make accurate predictions or classifications.

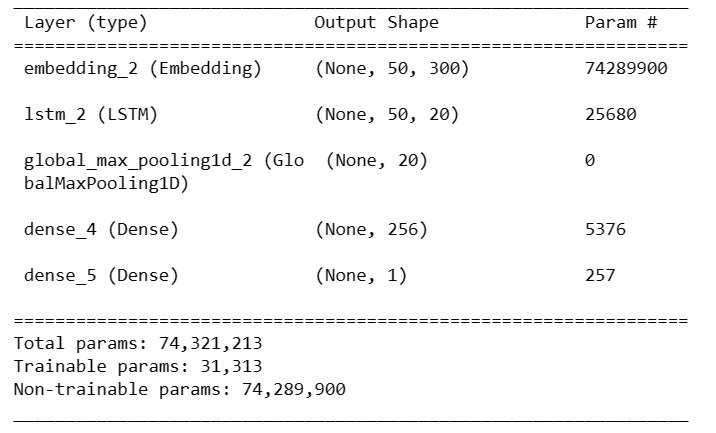
  
***Figure 3: “CNN” for image***

***CNN for Text:***

We apply CNN, or Convolutional Neural Network In the context of text classification, a Convolutional Neural Network (CNN) model is designed to capture local patterns or features within the input text. Unlike traditional CNN models used for image classification, which operate on 2D spatial data, CNN models for text classification typically operate on 1D sequence of words or characters.

Here's a general architecture of a CNN model for text classification:

1. **Input layer**: It specifies the input shape, which corresponds to the length of the input sequences or the number of words in a sentence.
2. **Embedding layer**: It maps the input words or characters to dense vectors of fixed size. This layer helps to represent the textual data in a continuous vector space, capturing semantic relationships between words.
3. **Dense layers:** These layers are fully connected layers that take the flattened feature vectors as input. They capture higher-level representations and perform classification based on the extracted features.
4. **MaxPooling layers**: These layers down sample the feature maps by selecting the maximum value within each local region. MaxPooling reduces the dimensionality of the feature maps while preserving the most important features.



**Figure4:”CNN” for text**

The model is trained using a labeled dataset where the input text is associated with binary labels indicating the presence or absence of depression. The model is optimized using appropriate loss functions (e.g., binary cross-entropy) and evaluated based on performance metrics such as accuracy, precision, recall, and F1-score.

It's important to preprocess the input text data by tokenizing the sentences, converting words to numerical representations (e.g., word embedding’s), and padding the sequences to a fixed length. Additionally, techniques like dropout and regularization can be applied to prevent over fitting.

Overall, the CNN model for text classification brings together the strengths of convolutional operations, hierarchical representation learning, efficient parameter sharing, automatic feature extraction, scalability, and generalization. These qualities make it an effective tool for detecting depression based on textual data, offering potential for early identification and intervention to support individuals who may be experiencing mental health challenges.

**Depression**

**CNN**

“I always hate life and want to die”

***Input text***

**Not Depression**

***4.5.2* Transfer Learning Techniques**

Transfer learning is a powerful technique in deep learning that involves leveraging the knowledge and learned features from a pre-trained model on one task and applying it to a different but related task. In our case, we will be using the VGG-16 and VGG-19 models for image classification.

The VGG-16 and VGG-19 models are deep convolutional neural networks (CNNs) that were originally trained for image classification tasks on the Image Net dataset.

***Figure 5:* Transfer Learning Techniques**

***VGG-16 & VGG19***:

VGG-16 and VGG-19 are popular deep convolutional neural network (CNN) architectures that were developed by the Visual Geometry Group (VGG) at the University of Oxford. These models are widely used for various computer vision tasks, including image classification and object detection.

***VGG-16:***

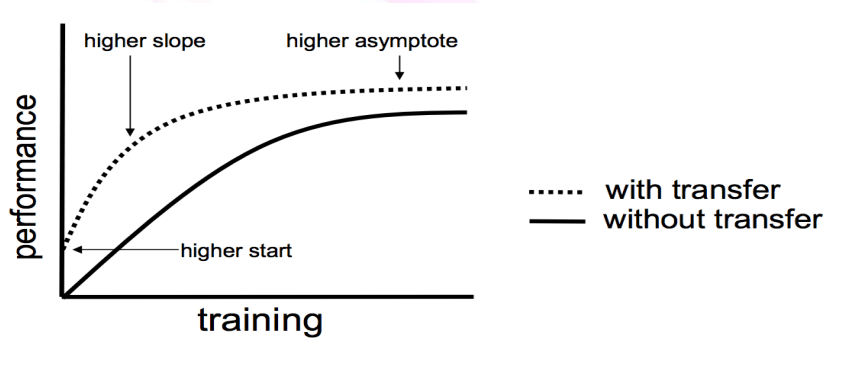
VGG-16 is a CNN architecture that consists of 16 layers, hence the name. The key characteristic of VGG-16 is its simplicity and uniformity. It stacks multiple convolutional layers with small 3x3 filters, followed by max-pooling layers to reduce spatial dimensions. The network gradually increases the number of filters as the layers go deeper. The final layers of VGG-16 are fully connected layers that perform classification based on the extracted features. VGG-16 achieved high accuracy in the Image Net Large-Scale Visual Recognition Challenge (ILSVRC) 2014 competition and has become a popular architecture for transfer learning and feature extraction.

***VGG-19:***

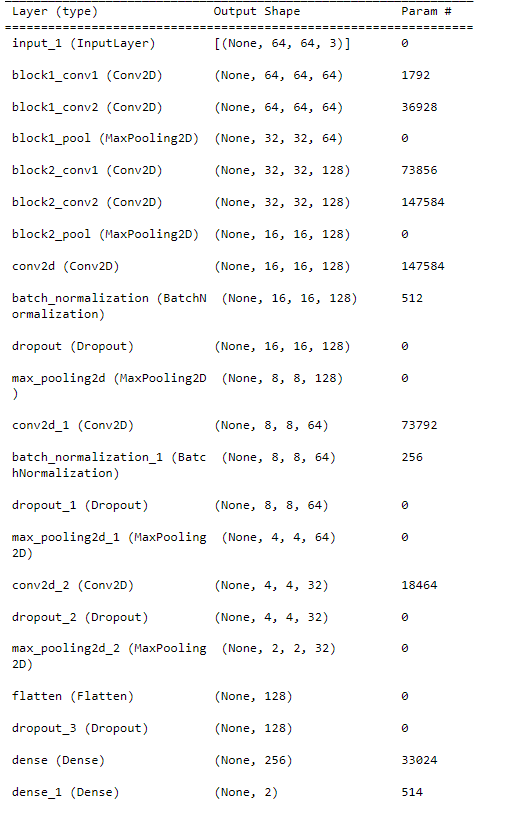
VGG-19 is an extension of VGG-16 and has 19 layers. It follows a similar architecture as VGG-16, but with additional convolutional layers. The extra layers enable VGG-19 to capture more complex features and increase the representational power of the network. However, the additional layers also make the model deeper and computationally more expensive to train and evaluate.

Both VGG-16 and VGG-19 have a relatively simple and straightforward architecture compared to more recent CNN architectures like Reset, Inception, or Efficient Net. Despite their simplicity, VGG models have shown excellent performance in various computer vision tasks and serve as a benchmark for evaluating newer architectures. These models are often used as pre-trained models for transfer learning, where the learned features are leveraged for different tasks or datasets.

It's worth noting that due to the depth and number of parameters in VGG-19, it requires more computational resources and training time compared to VGG-16. The choice between VGG-16 and VGG-19 depends on the specific task requirements, available computational resources, and trade-offs between model complexity and performance.

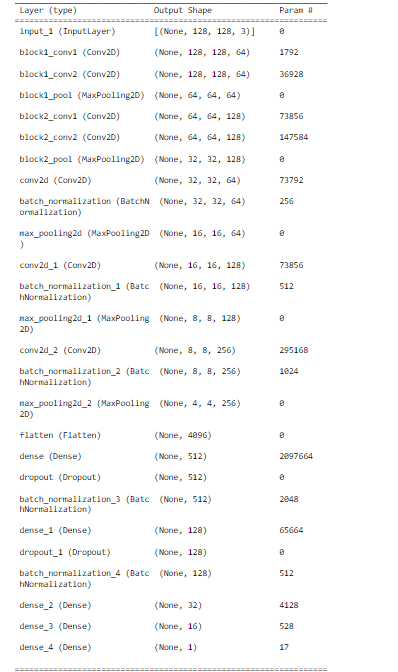


To know about VGG-16 model and its result see Figure

****

**Figure6**: **VGG-16**

To know about VGG-19 model and its result see Figure?



**Figure7: VGG-19**

To enhance the accuracy of the multi-class classification using the **VGG-16** model, it has been observed that utilizing images along with the target label yields better results compared to using the actual images directly. The previous paper demonstrated promising outcomes using the **VGG-16** model; hence we will adopt this model for our task.

By employing images, we can focus the model's attention on the relevant regions or features of the images that are crucial for classification. Masking involves identifying and highlighting specific regions of interest while suppressing irrelevant background information. This process helps in reducing noise and enhancing the discriminative features, leading to improved accuracy in the classification task.

The **VGG-16** model, with its deep architecture and large number of parameters, has shown strong performance in various computer vision tasks. Leveraging its learned features can further boost the model's ability to extract meaningful and discriminative information from the input data.

**Depression**

**Not Depression**

**VGG16**



***Input image***

***Chapter Five***

***Discussion and Results***



* 1. **Discussion**

***5.1.1 What are the symptoms of depression?***

Depression can manifest in various ways, and it is important to note that individuals may experience different symptoms and severity levels. Here are some common symptoms of depression:

1. **Persistent Sadness:** Feeling sad, empty, or hopeless most of the time.
2. **Loss of Interest or Pleasure:** Losing interest in activities, hobbies, or things that were previously enjoyable.
3. **Fatigue and Lack of Energy:** Feeling constantly tired, lacking energy, or experiencing a significant decrease in productivity.
4. **Changes in Sleep Patterns**: Experiencing insomnia (difficulty sleeping), hypersomnia (excessive sleepiness), or disrupted sleep patterns.
5. **Changes in Appetite or Weight**: Significant changes in appetite, leading to weight loss or weight gain.
6. **Difficulty Concentrating:** Experiencing problems with focus, memory, decision-making, or concentration.
7. **Feelings of Guilt or Worthlessness:** Persistent feelings of guilt, worthlessness, or self-blame, even for minor issues.
8. **Irritability or Restlessness:** Feeling easily irritable, agitated, or restless without any specific reason.
9. **Withdrawal from Social Activities**: Withdrawing from social interactions, avoiding friends, family, or social events.
10. **Physical Symptoms:** Experiencing unexplained physical symptoms such as headaches, digestive problems, or chronic pain.

It's important to note that experiencing a few of these symptoms for a short duration may not necessarily indicate clinical depression. However, if these symptoms persist for an extended period (usually two weeks or longer) and significantly affect daily life.

## 5.1.2 What causes depression?

Depression is a complex condition influenced by various factors. The exact cause of depression is often not attributed to a single factor but rather a combination of biological, genetic, environmental, and psychological factors. Here are some common factors that can contribute to the development of depression:

***Biological Factors*:** Imbalances in certain brain chemicals, such as neurotransmitters like serotonin, norepinephrine, and dopamine, can play a role in depression. Hormonal imbalances, changes in the functioning of the brain regions involved in mood regulation, and genetic predisposition can also contribute to the risk of developing depression.

***Genetics*:** Family history of depression can increase the likelihood of developing the condition. Individuals with a close relative, such as a parent or sibling, who has experienced depression, may be at a higher risk.

***Environmental Factors*:** Certain life events or circumstances can trigger or contribute to depression, including traumatic experiences, childhood abuse or neglect, chronic stress, major life changes (such as loss of a loved one, divorce, or job loss), social isolation, financial difficulties, or living in an unsupportive or abusive environment.

***Medical Conditions***: Some medical conditions, such as chronic illnesses (e.g., cancer, diabetes, heart disease), hormonal disorders (e.g., thyroid problems), or neurological conditions, may increase the risk of depression.

***Substance Abuse***: Substance abuse and addiction can co-occur with depression. The misuse of drugs or alcohol can contribute to the development or worsening of depressive symptoms.

***Psychological Factors***: Certain personality traits, such as low self-esteem, pessimism, or a tendency towards negative thinking, can make individuals more vulnerable to depression. Additionally, having a history of other mental health disorders, such as anxiety or eating disorders, can increase the risk.

It is important to note that not everyone with these risk factors will develop depression, and some individuals may experience depression without identifiable risk factors. The interplay of these factors can vary for each individual, and a comprehensive assessment by a healthcare professional or mental health provider is crucial to understanding the specific causes and developing an appropriate treatment plan.

‘

**5.1.3 Should a depression patient be tested? What tests does he need?**

Yes, depression can come back after a period of remission. Depression is often characterized by recurrent episodes, and individuals who have experienced depression in the past are at an increased risk of future episodes.

The likelihood of a depressive episode recurring varies from person to person. Some individuals may experience a single episode of depression and never have another, while others may have multiple episodes throughout their lifetime. Research suggests that about 50% to 80% of individuals who have experienced a major depressive episode are likely to have another episode in the future.

1. **Physical Examination**: A physical examination is performed to rule out any underlying medical conditions that may be causing or contributing to the symptoms of depression. The healthcare provider may ask about the patient's medical history and conduct necessary tests if needed.
2. **Psychological Evaluation**: The healthcare provider will conduct a detailed psychological assessment to gather information about the patient's symptoms, emotions, thoughts, behavior patterns, and any potential triggers or stressors. This assessment may involve the use of standardized questionnaires or interviews to assess the severity and duration of symptoms.
3. **Diagnostic Criteria:** The healthcare provider will refer to the Diagnostic and Statistical Manual of Mental Disorders (DSM-5), which outlines specific criteria for diagnosing depressive disorders. The patient's symptoms will be evaluated against these criteria to determine if they meet the diagnostic threshold for depression.
4. **Medical and Family History:** Gathering information about the patient's medical history, including any prior episodes of depression or mental health disorders, as well as their family history of depression or other psychiatric conditions, can provide additional insights into the diagnosis.
5. ***Laboratory Tests***: While there are no specific laboratory tests to diagnose depression, certain tests may be conducted to rule out underlying medical conditions or to assess overall health. These tests may include blood tests to check thyroid function, rule out vitamin deficiencies, or identify other medical conditions that may mimic depressive symptoms.

It's important to remember that the diagnosis of depression is based on a comprehensive evaluation of the patient's symptoms, history, and clinical judgment. The goal is to assess the severity, duration, and impact of the symptoms on the individual's daily functioning and well-being.

***5.1.4 Are there different types of depression?***

Yes, there are different types of depression. Depression is a complex mental health condition that can manifest in various ways, and different classifications and diagnostic systems recognize different types or subtypes of depression. Here are some commonly recognized types:

1. ***Major Depressive Disorder (MDD):*** Also known as clinical depression or unipolar depression, MDD is the most common form of depression. It involves persistent feelings of sadness, loss of interest or pleasure in activities, changes in appetite or weight, sleep disturbances, fatigue, feelings of worthlessness or guilt, difficulty concentrating, and recurrent thoughts of death or suicide.

2. ***Persistent Depressive Disorder (PDD):*** PDD, previously known as dysthymia, is a chronic form of depression lasting for at least two years. People with PDD experience a depressed mood most of the time, along with symptoms that may be less severe than in MDD but still significantly impact daily functioning.

***3. Bipolar Disorder:*** Bipolar disorder is characterized by alternating periods of depression and mania (elevated mood, increased energy, racing thoughts, impulsivity). During depressive episodes, individuals experience symptoms similar to those of major depressive disorder.

4. ***Seasonal Affective Disorder (SAD):*** SAD is a type of depression that occurs seasonally, typically during the fall and winter months when there is less natural sunlight. Symptoms include low mood, lack of energy, oversleeping, overeating (especially carbohydrates), and social withdrawal.

5. ***Postpartum Depression (PPD):*** PPD is a type of depression that occurs after childbirth. It involves intense feelings of sadness, anxiety, and exhaustion that may interfere with the ability to care for oneself or the newborn.

6. ***Psychotic Depression:*** This type of depression is characterized by severe depressive symptoms accompanied by psychotic features, such as hallucinations (seeing or hearing things that are not there) or delusions (holding false, irrational beliefs).

These are some of the recognized types of depression, but it's important to note that each individual's experience of depression can be unique, and symptoms may vary. Diagnosis and treatment should be conducted by a qualified mental health professional that can assess the specific symptoms and provide appropriate care.

**5.1.5 Do depressions come back after remission?**

Yes, depression can come back after a period of remission. Depression is often characterized by recurrent episodes, and individuals who have experienced depression in the past are at an increased risk of future episodes.

The likelihood of a depressive episode recurring varies from person to person. Some individuals may experience a single episode of depression and never have another, while others may have multiple episodes throughout their lifetime. Research suggests that about 50% to 80% of individuals who have experienced a major depressive episode are likely to have another episode in the future.

There are several factors that can contribute to the recurrence of depression:

1. ***History of depression***: Having experienced depression in the past is a significant risk factor for future episodes.

2. ***Inadequate or discontinued treatment***: Stopping or not fully completing treatment for depression, including medication, therapy, or a combination of both, can increase the chances of a relapse.

3. ***Presence of ongoing stressors***: Stressful life events, such as relationship difficulties, work-related stress, or financial problems, can contribute to the reemergence of depressive symptoms.

4. ***Co-occurring mental health conditions***: The presence of other mental health disorders, such as anxiety disorders or substance use disorders, can increase the risk of depression recurrence.

5***. Lack of effective coping strategies***: Inadequate coping skills to deal with stress, negative thinking patterns, and difficulty managing emotions can make individuals more vulnerable to depression relapse.

It's important for individuals who have experienced depression to be aware of the potential for recurrence and to have a plan in place to manage their mental health. This may involve ongoing therapy, medication, lifestyle changes, and developing effective coping strategies. Regular check-ins with a mental health professional can also be beneficial to monitor and address any early signs of depression relapse.

* 1. **Results**

## 5.2.1. Result of ML &DL in image

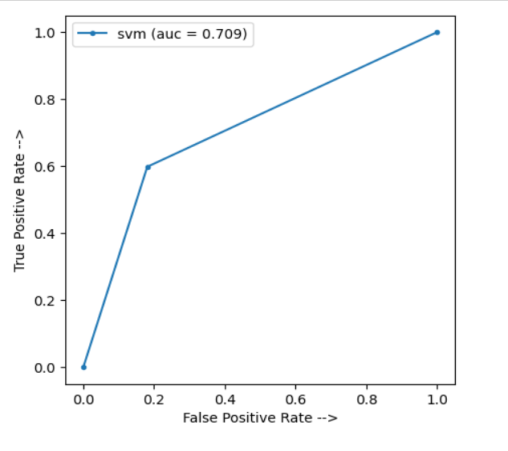
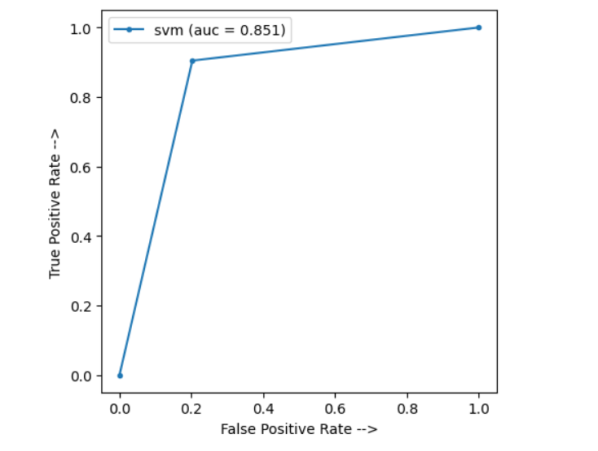
***First in Machine learning for image:***

1. ***Support Vector Machine***

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| ***Table4. Without pre-processing filters*** | | | | | |
|  | **Classifier** | **Precision** | **Recall** | **Accuracy** | **f1\_score** |
| ***Depressed/***  ***Non depressed*** | *SVM* | 0.755203 | 0.755523 | 0.755523 | 0.753885 |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| ***Table5. With pre-processing filters for train Image*** | | | | | |
|  | **Classifier** | **Precision** | **Recall** | **f1\_score** | **Support** |
| ***Non Depression*** | SVM | 0.87 | 0.80 | 0.83 | 5772 |
| **Depression** | SVM | 0.85 | 0.90 | 0.87 | 7140 |
| **Accuracy** | SVM |  |  | 0.86 | 12912 |
| **Macro Avg** | SVM | 0.86 | 0.85 | 0.85 | 12912 |
| **Weighted Avg** | SVM | 0.86 | 0.86 | 0.86 | 12912 |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| ***Table 6. With pre-processing filters for test Image*** | | | | | |
|  | ***Classifier*** | ***Precision*** | ***Recall*** | ***f1\_score*** | ***Support*** |
| ***Non Depression*** | *SVM* | 0.79 | 0.67 | 0.72 | 1443 |
| ***Depression*** | *SVM* | 0.76 | 0.85 | 0.80 | 1786 |
| ***Accuracy*** | *SVM* |  |  | 0.77 | 3229 |
| ***Macro Avg*** | *SVM* | 0.77 | 0.76 | 0.76 | 3229 |
| ***Weighted Avg*** | *SVM* | 0.77 | 0.77 | 0.77 | 3229 |



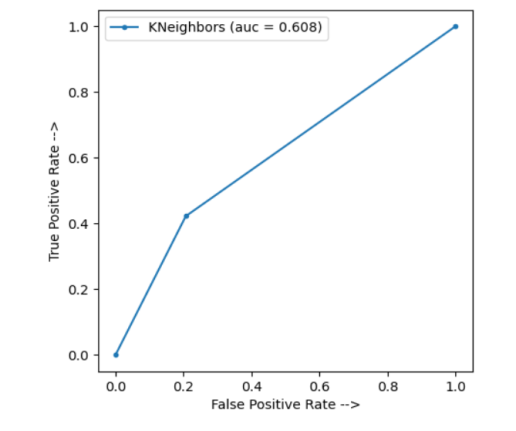
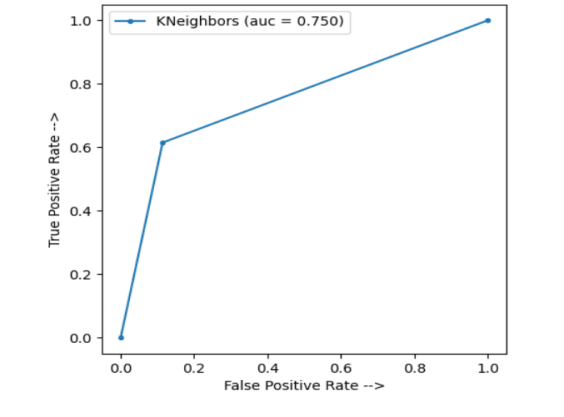
***Figure 8: SVM Images ROC* graph for train and test**

1. ***K-Nearest Neighbor***

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| ***Table 7. Without pre-processing filters*** | | | | | |
|  | ***Classifier*** | ***Precision*** | ***Recall*** | ***Accuracy*** | ***f1\_score*** |
| ***Depression/***  ***Non depression*** | ***KNN*** | 0.640134 | 0.634731 | 0.634731 | 0.635757 |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Table 8. With pre-processing filters for train Image** | | | | | |
|  | **Classifier** | **Precision** | **Recall** | **f1\_score** | **Support** |
| **Non Depression** | KNN | 0.74 | 0.77 | 0.75 | 5772 |
| **Depression** | KNN | 0.81 | 0.78 | 0.79 | 7140 |
| **Accuracy** | KNN |  |  | 0.77 | 12912 |
| **Macro Avg** | KNN | 0.77 | 0.77 | 0.77 | 12912 |
| **Weighted Avg** | KNN | 0.78 | 0.77 | 0.77 | 12912 |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| ***Table 9. With pre-processing filters for test Image*** | | | | | |
|  | ***Classifier*** | ***Precision*** | ***Recall*** | ***f1\_score*** | ***Support*** |
| ***Non Depression*** | KNN | 0.60 | 0.64 | 0.62 | 1443 |
| ***Depression*** | KNN | 0.69 | 0.65 | 0.67 | 1786 |
| ***Accuracy*** | KNN |  |  | 0.65 | 3229 |
| ***Macro Avg*** | KNN | 0.64 | 0.65 | 0.64 | 3229 |
| ***Weighted Avg*** | KNN | 0.65 | 0.65 | 0.65 | 3229 |



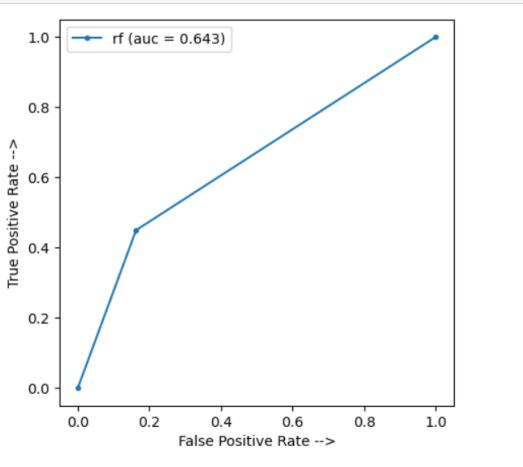
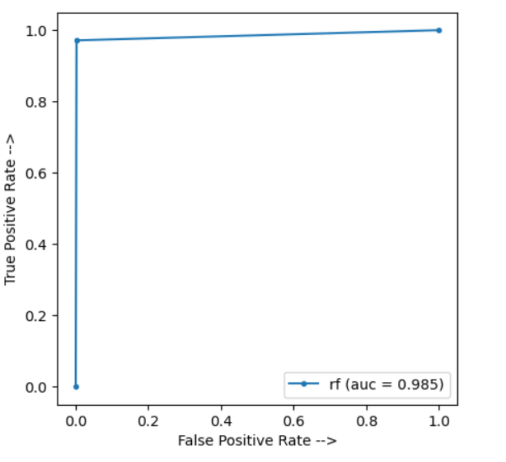
***Figure 9: KNN Images ROC* graph for train and test**

1. ***Random Forest***

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| ***Table 10. Without pre-processing filters*** | | | | | |
|  | **Classifier** | **Precision** | **Recall** | **Accuracy** | **f1\_score** |
| ***Depression/***  ***Non depression*** | RF | 0.69686 | 0.693991 | 0.693991 | 0.686069 |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| ***Table 11. With pre-processing filters for train Image*** | | | | | |
|  | ***Classifier*** | ***Precision*** | ***Recall*** | ***f1\_score*** | ***Support*** |
| ***Non Depression*** | RF | 0.98 | 1.00 | 0.99 | 5772 |
| ***Depression*** | RF | 1.00 | 0.99 | 0.99 | 7140 |
| ***Accuracy*** | RF |  |  | 0.99 | 12912 |
| ***Macro Avg*** | RF | 0.99 | 0.99 | 0.99 | 12912 |
| ***Weighted Avg*** | RF | 0.99 | 0.99 | 0.99 | 12912 |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| ***Table12. With pre-processing filters for test Image*** | | | | | |
|  | ***Classifier*** | ***Precision*** | ***Recall*** | ***f1\_score*** | ***Support*** |
| ***Non Depression*** | RF | 0.67 | 0.67 | 0.72 | 1443 |
| ***Depression*** | RF | 0.73 | 0.73 | 0.73 | 1786 |
| ***Accuracy*** | RF |  |  | 0.70 | 3229 |
| ***Macro Avg*** | RF | 0.70 | 0.70 | 0.70 | 3229 |
| ***Weighted Avg*** | RF | 0.70 | 0.70 | 0.70 | 3229 |



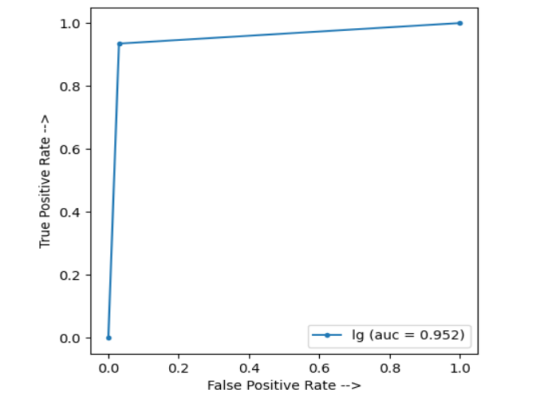
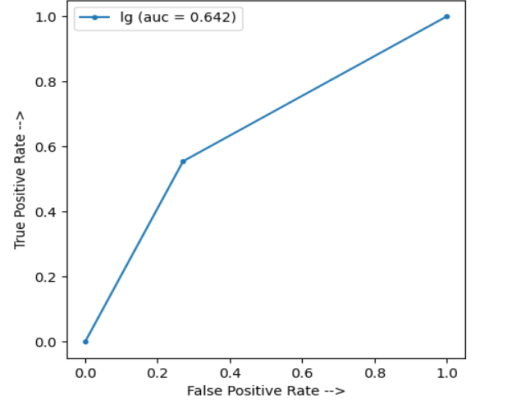
***Figure 10: RF Images ROC* graph for train and test**

1. ***Logistic regression***

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| ***Table 13. Without pre-processing filters*** | | | | | |
|  | ***Classifier*** | ***Precision*** | ***Recall*** | ***Accuracy*** | ***f1\_score*** |
| ***Depression/***  ***Non depression*** | LR | 0.645324 | 0.645261 | 0.645261 | 0.645292 |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| ***Table 14. With pre-processing filters for train Image*** | | | | | |
|  | ***Classifier*** | ***Precision*** | ***Recall*** | ***f1\_score*** | ***Support*** |
| ***Non Depression*** | LR | 0.86 | .83 | 0.85 | 5772 |
| ***Depression*** | LR | 0.87 | 0.89 | 0.88 | 7140 |
| ***Accuracy*** | LR |  |  | 0.87 | 12912 |
| ***Macro Avg*** | LR | 0.87 | 0.86 | 0.86 | 12912 |
| ***Weighted Avg*** | LR | 0.87 | 0.87 | 0.86 | 12912 |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| ***Table 15. With pre-processing filters for test Image*** | | | | | |
|  | ***Classifier*** | ***Precision*** | ***Recall*** | ***f1\_score*** | ***Support*** |
| ***Non Depression*** | LR | 0.64 | 0.60 | 0.62 | 1443 |
| ***Depression*** | LR | 0.69 | 0.73 | 0.71 | 1786 |
| ***Accuracy*** | LR |  |  | 0.67 | 3229 |
| ***Macro Avg*** | LR | 0.66 | 0.66 | 0.66 | 3229 |
| ***Weighted Avg*** | LR | 0.67 | 0.67 | 0.67 | 3229 |



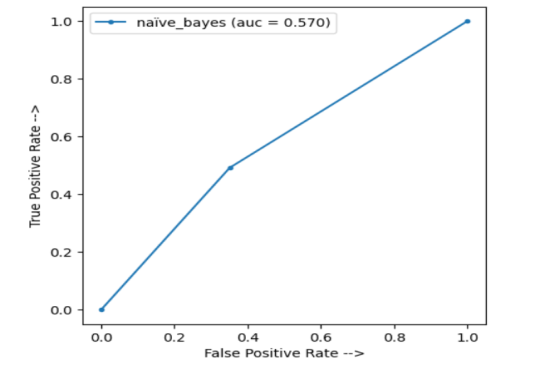
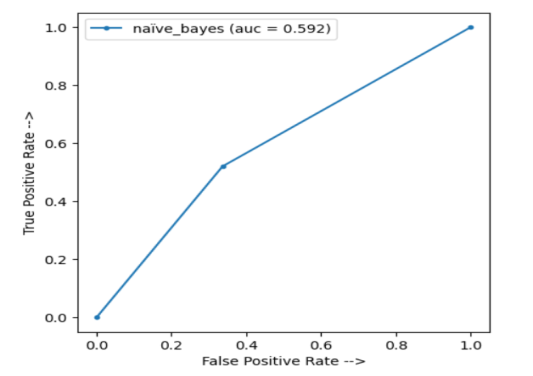
***Figure 11: LR Images ROC* graph for train and test**

1. ***NaïveBaye***

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| ***Table 16. Without pre-processing filters*** | | | | | |
|  | ***Classifier*** | ***Precision*** | ***Recall*** | ***Accuracy*** | ***f1\_score*** |
| ***Depression/***  ***Non depression*** | NB | 0.590524 | 0.575263 | 0.575263 | 0.574594 |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| ***Table 17. With pre-processing filters for train Image*** | | | | | |
|  | ***Classifier*** | ***Precision*** | ***Recall*** | ***f1\_score*** | ***Support*** |
| ***Non Depression*** | NB | 0.52 | .68 | 0.59 | 5772 |
| ***Depression*** | NB | 0.66 | 0.50 | 0.57 | 7140 |
| ***Accuracy*** | NB |  |  | 0.58 | 12912 |
| ***Macro Avg*** | NB | 0.59 | 0.59 | 0.58 | 12912 |
| ***Weighted Avg*** | NB | 0.60 | 0.58 | 0.58 | 12912 |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| ***Table 18. With pre-processing filters for test Image*** | | | | | |
|  | *Classifier* | *Precision* | *Recall* | *f1\_score* | *Support* |
| ***Non Depression*** | NB | 0.51 | 0.66 | 0.57 | 1443 |
| ***Depression*** | NB | 0.63 | 0.48 | 0.55 | 1786 |
| ***Accuracy*** | NB |  |  | 0.56 | 3229 |
| ***Macro Avg*** | NB | 0.57 | 0.57 | 0.56 | 3229 |
| ***Weighted Avg*** | NB | 0.58 | 0.56 | 0.56 | 3229 |



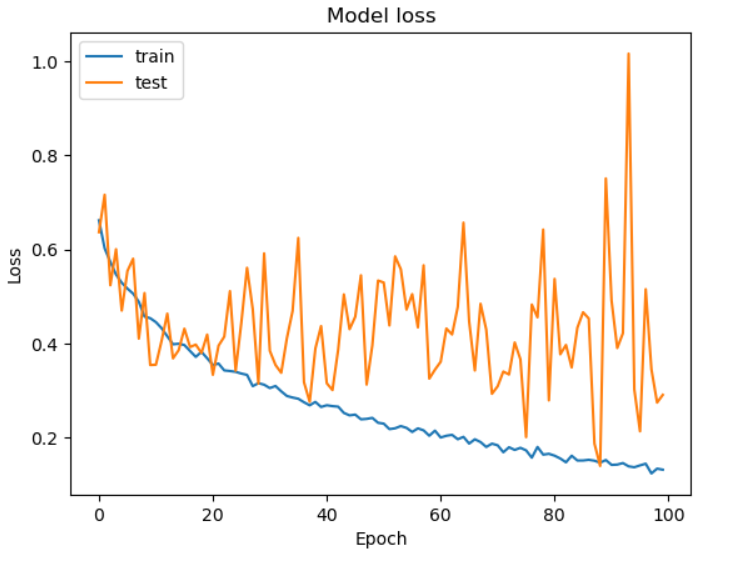
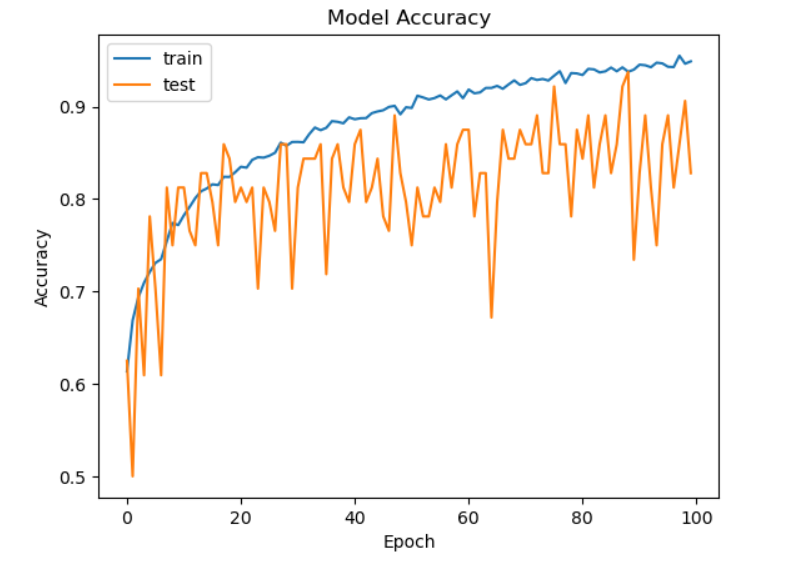
***Figure 12: NB Images ROC* graph for train and test**

***Second in Deep learning for image:***

***1-CNN***

|  |  |
| --- | --- |
| ***Table 19: CNN*** | |
|  | **Image** |
| ***Classifier accuracy on***  ***the training set*** | 93.75% |
| ***Time of CNN model in seconds*** | 74 |

|  |  |  |
| --- | --- | --- |
| ***Table 20: The parameters used in CN model*** | | |
| **Layer (type)** | **Output Shape** | **# Parameters** |
| conv2d\_4 (Conv2D) | (None, 126, 126, 32) | 320 |
| max\_pooling2d\_4 (MaxPooling2D) | (None, 63, 63, 32) | 0 |
| conv2d\_5 (Conv2D) | (None, 61, 61, 32) | 9248 |
| max\_pooling2d\_5 (MaxPooling2D) | (None, 30, 30, 32) | 0 |
| flatten\_2 (Flatten) | (None, 28800) | 0 |
| dense\_8 (Dense) | (None, 128) | 3686528 |
| dense\_9 (Dense) | (None, 64) | 8256 |
| dense\_10 (Dense) | (None, 10) | 650 |
| dense\_8 (Dens | (None, 2) | 22 |

******

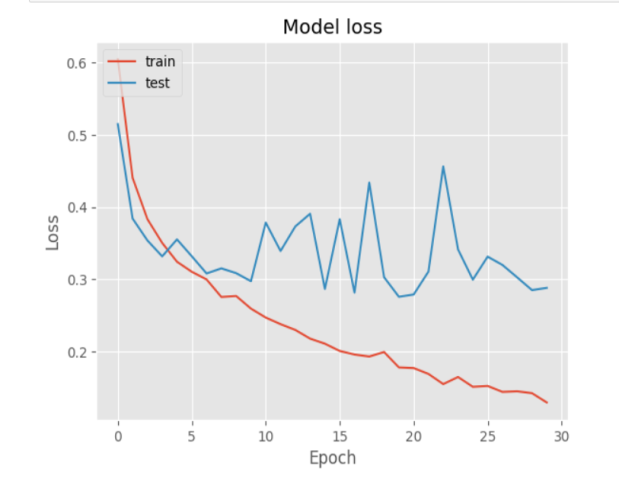
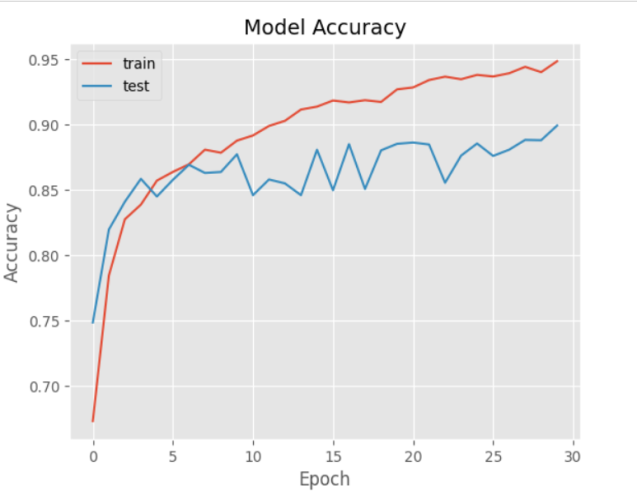
***Figure 13: CNN Images Accuracy and Loss graph***

***2-VGG-16***

|  |  |
| --- | --- |
| ***Table 21: VGG-16*** | |
|  | **Image** |
| ***Classifier accuracy on***  ***the training set*** | 90 |
| ***Confusion matrix of testing*** |  |
| ***ROC Curve Of Testing*** |  |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| ***Table 22: Classification report of image using VGG16*** | | | | |
|  | ***Precision*** | ***Recall*** | ***f1\_score*** | ***Support*** |
| ***Depression*** | 0.87 | 0.89 | 0.88 | 1630 |
| ***Non Depression*** | 0.92 | 0.90 | 0.91 | 2364 |
| ***Accuracy*** |  |  | 0.90 | 3994 |
| ***Macro Avg*** | 0.89 | 0.90 | 0.90 | 3994 |
| ***Weighted Avg*** | 0.90 | 0.90 | 0.90 | 3994 |

|  |  |  |  |
| --- | --- | --- | --- |
| ***Table 23: The parameters used in VGG-16 model*** | | | |
| **Layer (type)** | **Output Shape** | | **# Parameters** |
| block1\_conv1 (Conv2D) | (None, 64, 64, 64) | | 1729 |
| block1\_conv2 (Conv2D) | (None, 64, 64, 64) | | **36928** |
| block1\_pool (MaxPooling2D) | (None, 32, 32, 64) | | 0 |
| block2\_conv1 (Conv2D) | (None, 32, 32, 128) | | **73856** |
| block2\_conv2 (Conv2D ) | (None, 32, 32, 128) | | **147584** |
| block2\_pool (MaxPooling2D) | (None, 16,16, 128) | | 0 |
| conv2d (Conv2D) | (None, 16,16, 128) | | **147584** |
| batch\_ normalization (Batch Normalization) | (None, 16,16, 128) | | 512 |
| dropout (Dropout) | (None, 16,16, 128) | | 0 |
| max\_pooling2d (MaxPooling2D) | (None, 8, 8, 128) | | 0 |
| conv2d\_1 (Conv2D) | *(None, 8, 8, 64)* | | 73792 |
| batch\_normalization\_1(Batch Normalization) | *(None, 8, 8, 64)* | | 256 |
| dropout\_1 (Dropout) | *(None, 8, 8, 64)* | | 0 |
| max\_pooling2d\_1 (MaxPooling2D) | *(None, 4,4, 64)* | | 0 |
| conv2d\_2 (Conv2D) | *(None, 4 ,4, 32)* | | 18464 |
| dropout\_2 (Dropout) | *(None, 4 ,4, 32)* | | 0 |
| max\_pooling2d\_2 (MaxPooling2D) | *(None, 2, 2, 32)* | 0 | |
| flatten (Flatten) | *(None, 128)* | 0 | |
| dropout\_3 (Dropout) | *(None128,)* | 0 | |
| dense (Dense) | *(None, 256)* | 33024 | |
| Dense\_1 (Dense) | *(None, 2)* | 514 | |

****

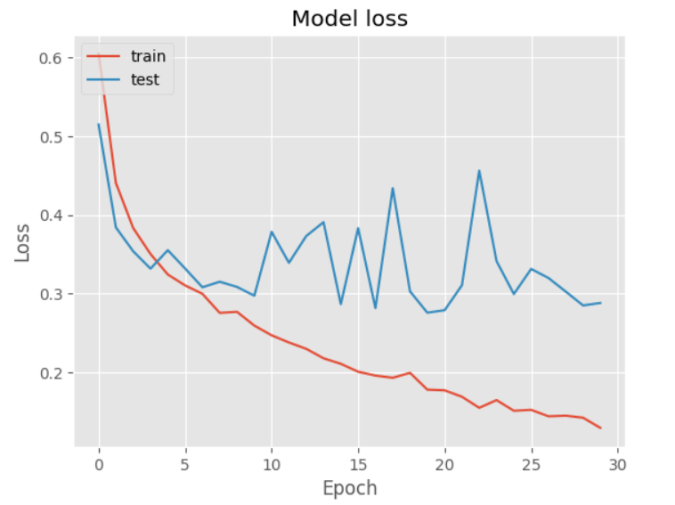
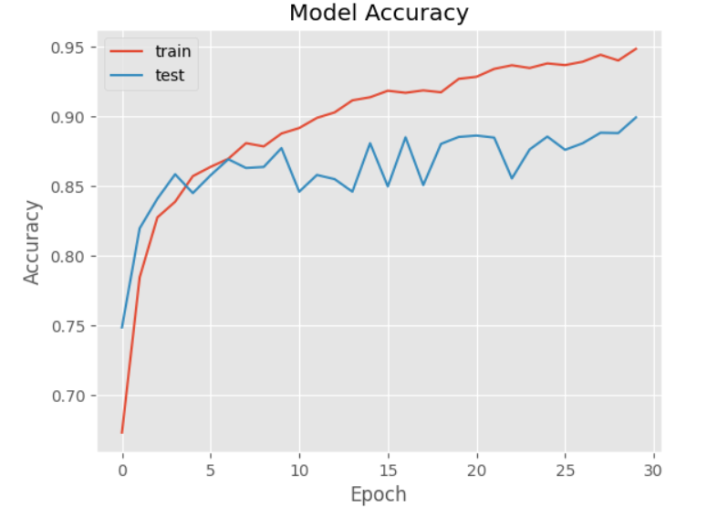
***Figure 14: VGG16 Images Accuracy and Loss graph***

***3-VGG-19***

|  |  |
| --- | --- |
| ***Table 24: VGG-19*** | |
|  | **image** |
| ***Classifier accuracy on***  ***the training set*** | 90 |
| ***Confusion matrix of testing*** |  |
| ***ROC Curve Of Testing*** |  |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| ***Table 25: Classification report of image using VGG-19*** | | | | |
|  | ***Precision*** | ***Recall*** | ***f1\_score*** | ***Support*** |
| ***Depression*** | 0.87 | 0.89 | 0.88 | 1630 |
| ***Non Depression*** | 0.92 | 0.90 | 0.91 | 2364 |
| ***Accuracy*** |  |  | 0.90 | 3994 |
| ***Macro Avg*** | 0.89 | 0.90 | 0.90 | 3994 |
| ***Weighted Avg*** | 0.90 | 0.90 | 0.90 | 3994 |

|  |  |  |  |
| --- | --- | --- | --- |
| ***Table 26: The parameters used in VGG-19 model*** | | | |
| **Layer (type)** | **Output Shape** | | **# Parameters** |
| block1\_conv1 (Conv2D) | (None, 64, 64, 64) | | 1729 |
| block1\_conv2 (Conv2D) | (None, 64, 64, 64) | | 36928 |
| block1\_pool (MaxPooling2D) | (None, 32, 32, 64) | | 0 |
| block2\_conv1 (Conv2D) | (None, 32, 32, 128) | | 73856 |
| block2\_conv2 (Conv2D ) | (None, 32, 32, 128) | | 147584 |
| block2\_pool (MaxPooling2D) | (None, 16,16, 128) | | 0 |
| conv2d (Conv2D) | (None, 16,16, 128) | | 147584 |
| batch\_ normalization (Batch Normalization) | (None, 16,16, 128) | | 512 |
| dropout (Dropout) | (None, 16,16, 128) | | 0 |
| max\_pooling2d (MaxPooling2D) | (None, 8, 8, 128) | | 0 |
| conv2d\_1 (Conv2D) | *(None, 8, 8, 64)* | | 73792 |
| dropout\_1 (Dropout) | *(None, 8, 8, 64)* | | 0 |
| max\_pooling2d\_1 (MaxPooling2D) | *(None, 4,4, 64)* | | 0 |
| conv2d\_2 (Conv2D) | *(None, 4 ,4, 32)* | | 18464 |
| dropout\_2 (Dropout) | *(None, 4 ,4, 32)* | | 0 |
| max\_pooling2d\_2 (MaxPooling2D) | *(None, 2, 2, 32)* | 0 | |
| flatten (Flatten) | *(None, 128)* | 0 | |
| dropout\_3 (Dropout) | *(None128,)* | 0 | |
| dense (Dense) | *(None, 256)* | 33024 | |
| Dense\_1 (Dense) | *(None, 2)* | 514 | |

******

***Figure 15: VGG-19 Images Accuracy and Loss graph***

***5.2.2. Result of ML &DL in text***

First in Machine learning for text:

1. ***Support Vector Machine***

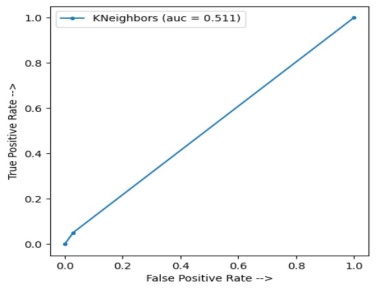
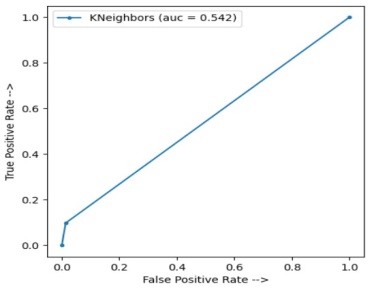
|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| ***Table27. With pre-processing filters for train Text*** | | | | | |
|  | **Classifier** | **Precision** | **Recall** | **f1\_score** | **Support** |
| ***Non Depression*** | SVM | 0.98 | 0.98 | 0.98 | 87092 |
| **Depression** | SVM | 0.98 | 0.98 | 0.98 | 87092 |
| **Accuracy** | SVM |  |  | 0.98 | 174184 |
| **Macro Avg** | SVM | 0.98 | 0.98 | 0.98 | 174184 |
| **Weighted Avg** | SVM | 0.98 | 0.98 | 0.98 | 174184 |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| ***Table 28. With pre-processing filters for test Text*** | | | | | |
|  | ***Classifier*** | ***Precision*** | ***Recall*** | ***f1\_score*** | ***Support*** |
| ***Non Depression*** | *SVM* | 0.93 | 0.95 | 0.94 | 29074 |
| ***Depression*** | *SVM* | 0.94 | 0.93 | 0.94 | 29074 |
| ***Accuracy*** | *SVM* |  |  | 0.94 | 58148 |
| ***Macro Avg*** | *SVM* | 0.94 | 0.94 | 0.94 | 58148 |
| ***Weighted Avg*** | *SVM* | 0.94 | 0.94 | 0.94 | 58148 |

1. **K-Nearest Neighbor**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Table 29. With pre-processing filters for train Text** | | | | | |
|  | **Classifier** | **Precision** | **Recall** | **f1\_score** | **Support** |
| **Non Depression** | KNN | 0.52 | 0.99 | 0.68 | 87092 |
| **Depression** | KNN | 0.87 | 0.07 | 0.13 | 87092 |
| **Accuracy** | KNN |  |  | 0.53 | 174184 |
| **Macro Avg** | KNN | 0.69 | 0.53 | 0.41 | 174184 |
| **Weighted Avg** | KNN | 0.69 | 0.53 | 0.41 | 174184 |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| ***Table 30. With pre-processing filters for test Text*** | | | | | |
|  | ***Classifier*** | ***Precision*** | ***Recall*** | ***f1\_score*** | ***Support*** |
| ***Non Depression*** | KNN | 0.51 | 0.98 | 0.67 | 29074 |
| ***Depression*** | KNN | 0.67 | 0.04 | 0.08 | 29074 |
| ***Accuracy*** | KNN |  |  | 0.51 | 58148 |
| ***Macro Avg*** | KNN | 0.59 | 0.51 | 0.37 | 58148 |
| ***Weighted Avg*** | KNN | 0.59 | 0.51 | 0.37 | 58148 |

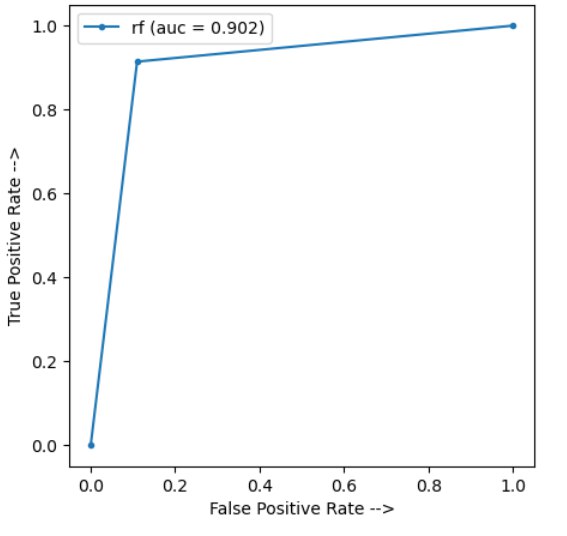
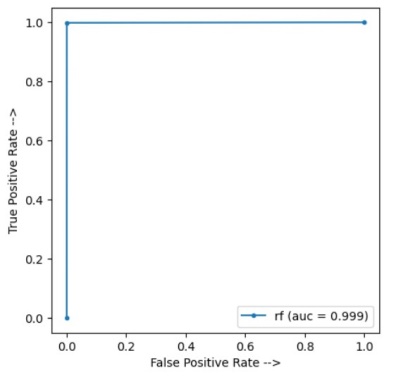
******

***Figure 16: KNN text ROC* graph for train and test**

1. ***Random Forest***

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| ***Table 31. With pre-processing filters for train Text*** | | | | | |
|  | ***Classifier*** | ***Precision*** | ***Recall*** | ***f1\_score*** | ***Support*** |
| ***Non Depression*** | RF | 1.00 | 1.00 | 1.00 | 87092 |
| ***Depression*** | RF | 1.00 | 1.00 | 1.00 | 87092 |
| ***Accuracy*** | RF |  |  | 1.00 | 174184 |
| ***Macro Avg*** | RF | 1.00 | 1.00 | 1.00 | 174184 |
| ***Weighted Avg*** | RF | 1.00 | 1.00 | 1.00 | 174184 |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| ***Table 32. With pre-processing filters for test Text*** | | | | | |
|  | ***Classifier*** | ***Precision*** | ***Recall*** | ***f1\_score*** | ***Support*** |
| ***Non Depression*** | RF | 0.91 | 0.89 | 0.90 | 29074 |
| ***Depression*** | RF | 0.89 | 0.91 | 0.90 | 29074 |
| ***Accuracy*** | RF |  |  | 0.90 | 58148 |
| ***Macro Avg*** | RF | 0.90 | 0.90 | 0.90 | 58148 |
| ***Weighted Avg*** | RF | 0.90 | 0.90 | 0.90 | 58148 |

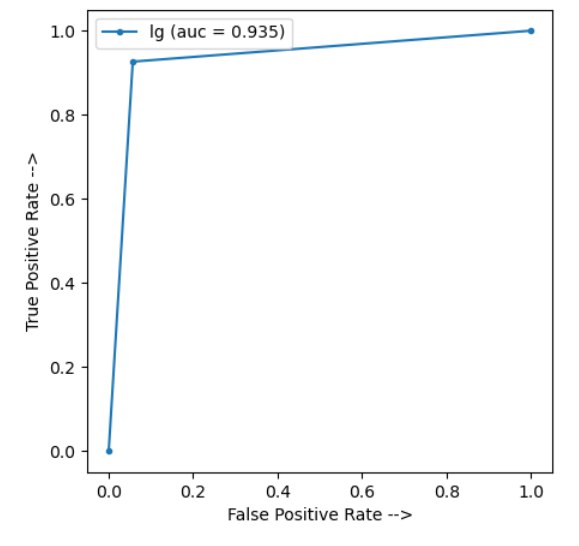
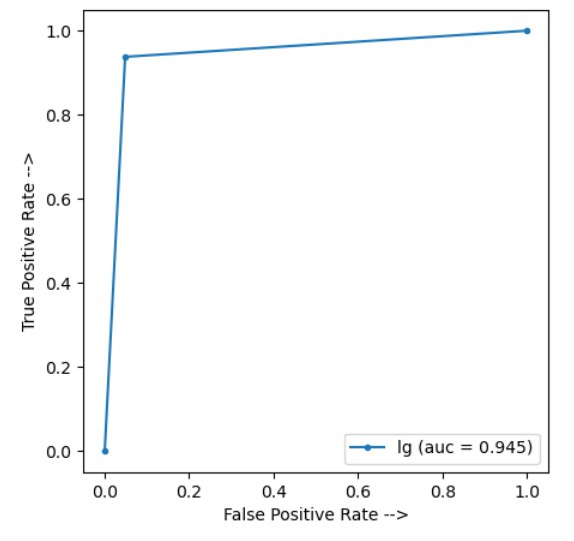
******

***Figure 17: RF text ROC* graph for train and test**

1. ***Logistic regression***

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| ***Table 34. With pre-processing filters for test Text*** | | | | | |
|  | ***Classifier*** | ***Precision*** | ***Recall*** | ***f1\_score*** | ***Support*** |
| ***Non Depression*** | LR | 0.96 | 0.95 | 0.96 | 29074 |
| ***Depression*** | LR | 0.95 | 0.96 | 0.96 | 29074 |
| ***Accuracy*** | LR |  |  | 0.96 | 58148 |
| ***Macro Avg*** | LR | 0.96 | 0.96 | 0.96 | 58148 |
| ***Weighted Avg*** | LR | 0.96 | 0.96 | 0.96 | 58148 |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| ***Table 34. With pre-processing filters for test Text*** | | | | | |
|  | ***Classifier*** | ***Precision*** | ***Recall*** | ***f1\_score*** | ***Support*** |
| ***Non Depression*** | LR | 0.93 | 0.94 | 0.94 | 29074 |
| ***Depression*** | LR | 0.94 | 0.93 | 0.93 | 29074 |
| ***Accuracy*** | LR |  |  | 0.93 | 58148 |
| ***Macro Avg*** | LR | 0.93 | 0.93 | 0.93 | 58148 |
| ***Weighted Avg*** | LR | 0.93 | 0.93 | 0.93 | 58148 |

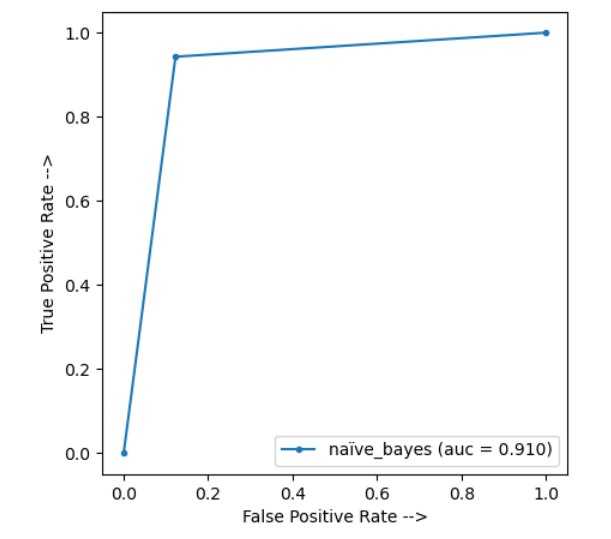
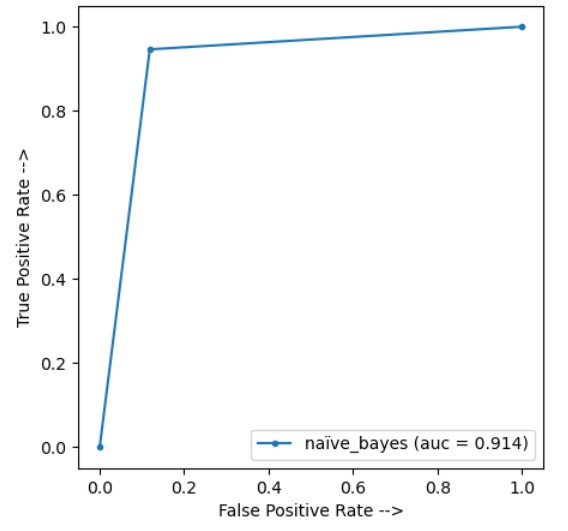
******

***Figure 18: LR text ROC* graph for train and test**

***5.NaïveBaye***

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| ***Table 35. With pre-processing filters for train Text*** | | | | | |
|  | ***Classifier*** | ***Precision*** | ***Recall*** | ***f1\_score*** | ***Support*** |
| ***Non Depression*** | NB | 0.94 | 0.88 | 0.91 | 87092 |
| ***Depression*** | NB | 0.89 | 0.95 | 0.92 | 87092 |
| ***Accuracy*** | NB |  |  | 0.91 | 174184 |
| ***Macro Avg*** | NB | 0.91 | 0.91 | 0.91 | 174184 |
| ***Weighted Avg*** | NB | 0.91 | 0.91 | 0.91 | 174184 |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| ***Table 36. With pre-processing filters for test Text*** | | | | | |
|  | *Classifier* | *Precision* | *Recall* | *f1\_score* | *Support* |
| ***Non Depression*** | NB | 0.94 | 0.88 | 0.91 | 29074 |
| ***Depression*** | NB | 0.88 | 0.94 | 0.91 | 29074 |
| ***Accuracy*** | NB |  |  | 0.91 | 58148 |
| ***Macro Avg*** | NB | 0.91 | 0.91 | 0.91 | 58148 |
| ***Weighted Avg*** | NB | 0.91 | 0.91 | 0.91 | 58148 |

******

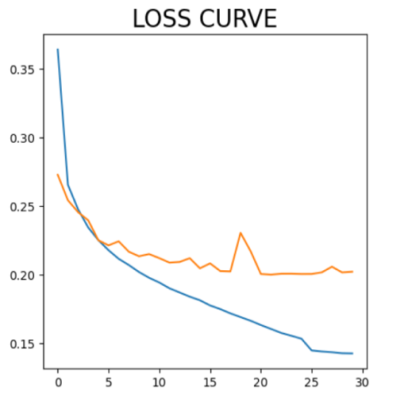
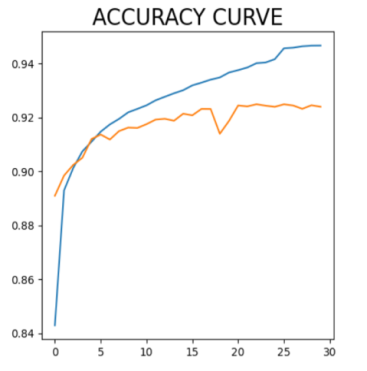
***Figure 19: NB text ROC* graph for train and test**

***Second in Machine learning for text:***

1. ***CNN model***

|  |  |  |
| --- | --- | --- |
| ***Table 37: The parameters used in CNN model*** | | |
| **Layer (type)** | **Output Shape** | **# Parameters** |
| embedding\_2 (Embedding) | (None, 50, 300) | 74289900 |
| lstm\_2 (LSTM) | (None, 50, 20) | 25680 |
| global\_max\_pooling1d\_2 (GlobalMaxPooling1D) | (None, 20) | 0 |
| dense\_4 (Dense) | (None, 256) | 5376 |
| dense\_5 (Dense) | (None, 1) | 257 |

The accuracy of text model is 0.9466 and val\_accuraccy is 0.9239



***Figure 20: CNN Model Accuracy and Loss graph***

***Chapter Six***

***Web Application***



* 1. The Home page

We do simple home page that, the benefit of the home page is that we can easily signup, login to website, explore and going to do test.



**Figure 21**: Home page

* 1. The Sign-up page

Enables the user to create an account in order to access this application and fulfill their requirements. The user is prompted to provide their, name, email address, password and phone.

* **The name** is the name of the user that consists of characters and we make validation to it that must min Length of name is 3 and Name is required.

And our app guarantees that, the name is not empty and the length is at least 3 letters Allows the user to complete his information

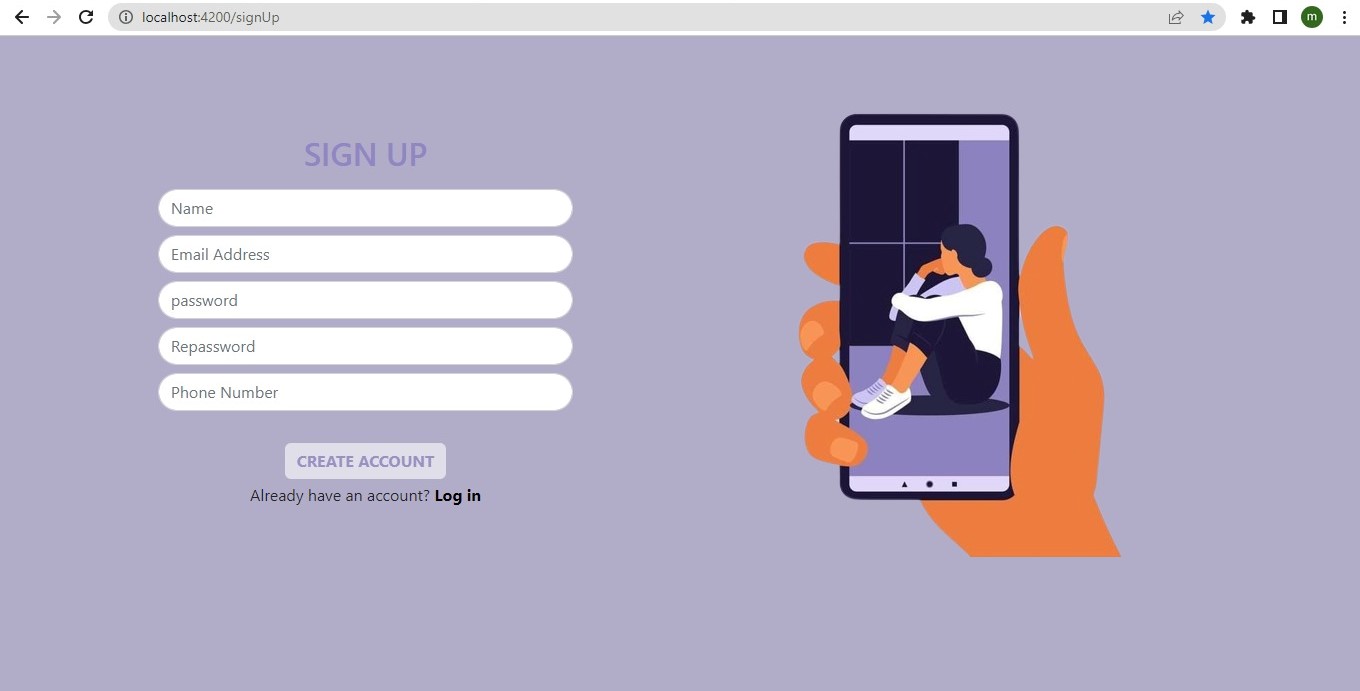
If the name is not found, do not allow the user to complete it

* **The email address** is a combination of characters, numbers, and special characters. It must be validated to ensure it contains the "@" character.
* **The password** must meet certain criteria, including the inclusion of numbers, characters, or special characters. It can contain all or any two of these types. Additionally, the password must have a minimum length of 8 characters.

Furthermore, on this page, all fields must be completed in order for the user to proceed. If any field is left blank, our application will not allow them to continue.

* **RePassword** In it, the password is entered from it must match with password, it is required.
* **Phone in** this field user must enter his/her phone number.

User mustenter Valid Egyptian Number and Your Phone is required.



**Figure 22**: Sign-up Page

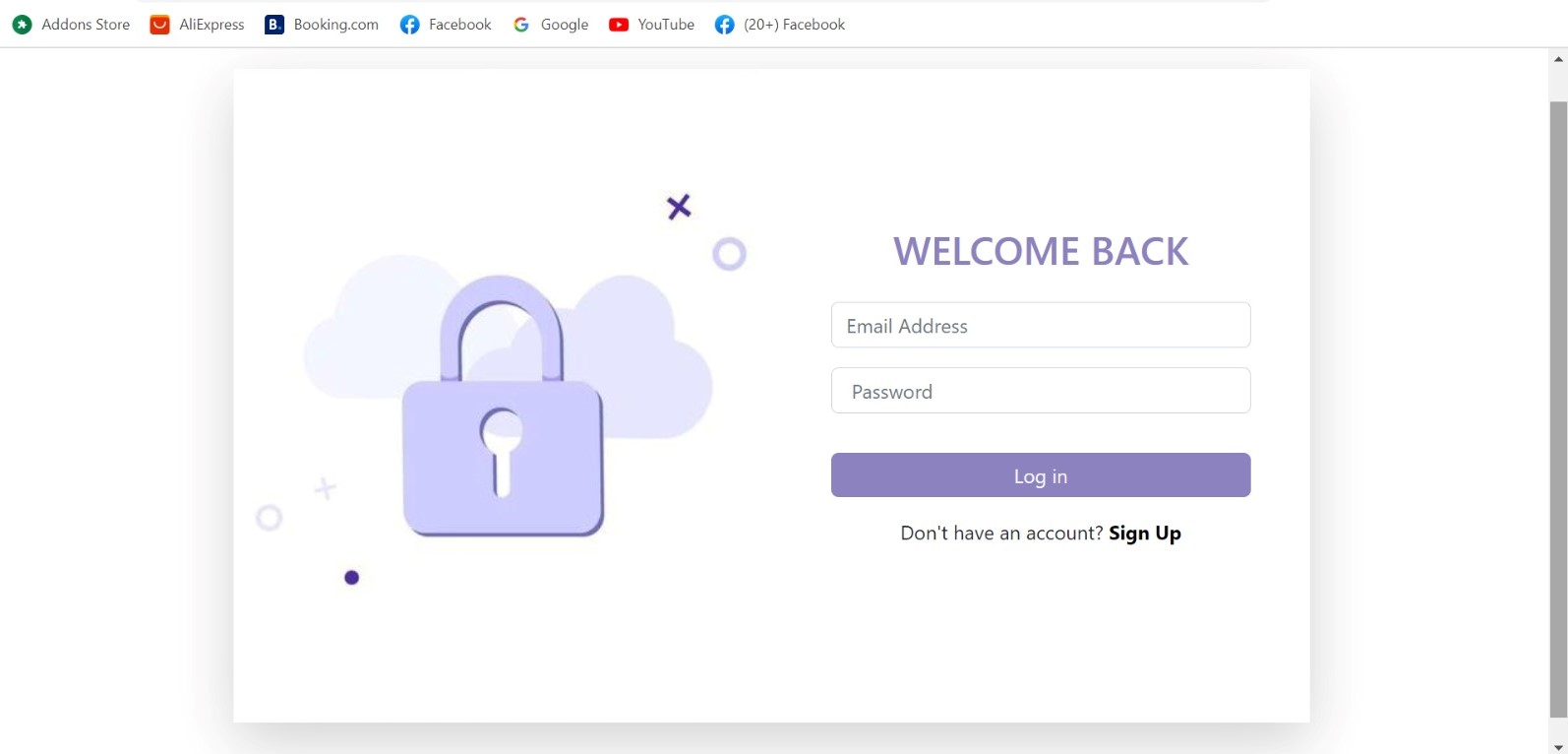
* 1. The Login page (figure 24)

The authentication process requires the user to provide their email address and password in order to verify their identity.

* **The email address** is a combination of characters, numbers, and special characters, and it typically follows a specific format. One common validation rule for an email address is that it must contain the "@" character.
* **The password** for the sign-up page must meet certain criteria. It should contain at least one number, one character, and one special character. Alternatively, it can contain all of these elements or any two of them. The password must also have a minimum length of 8 characters.

To access the application, users need to provide their username and password on the sign-up page. The application will store these authorized users to ensure that only authorized individuals can use it.

Additionally, we have implemented validation on this page. If a user fails to complete any field on the sign-up page, our application will not allow them to proceed until all fields are filled out correctly.



**Figure 23: Log-in page.**

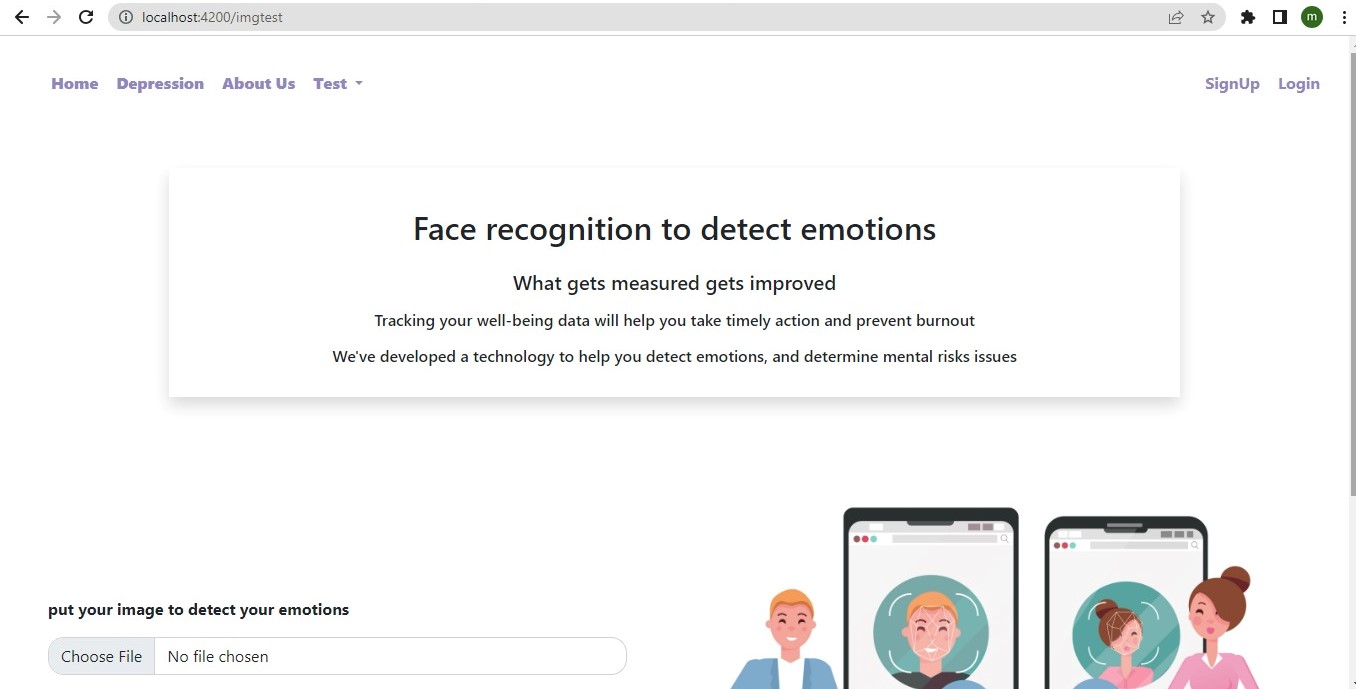
* 1. The Test Image Page

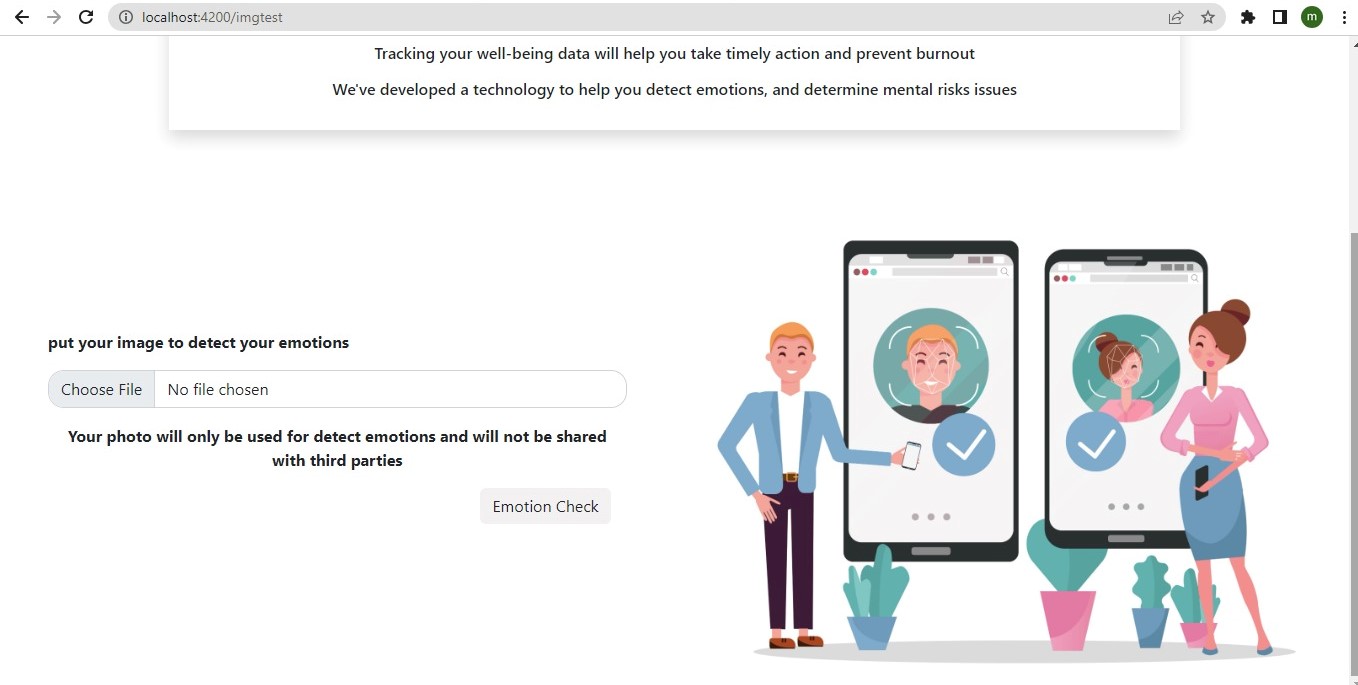
To perform facial recognition and determine whether a photo reflects a happy or sad expression, you can follow these steps:

1. Get the image: Put a picture of the face you want to analyze.

Click on upload button to upload your image

1. Face Detection and Alignment

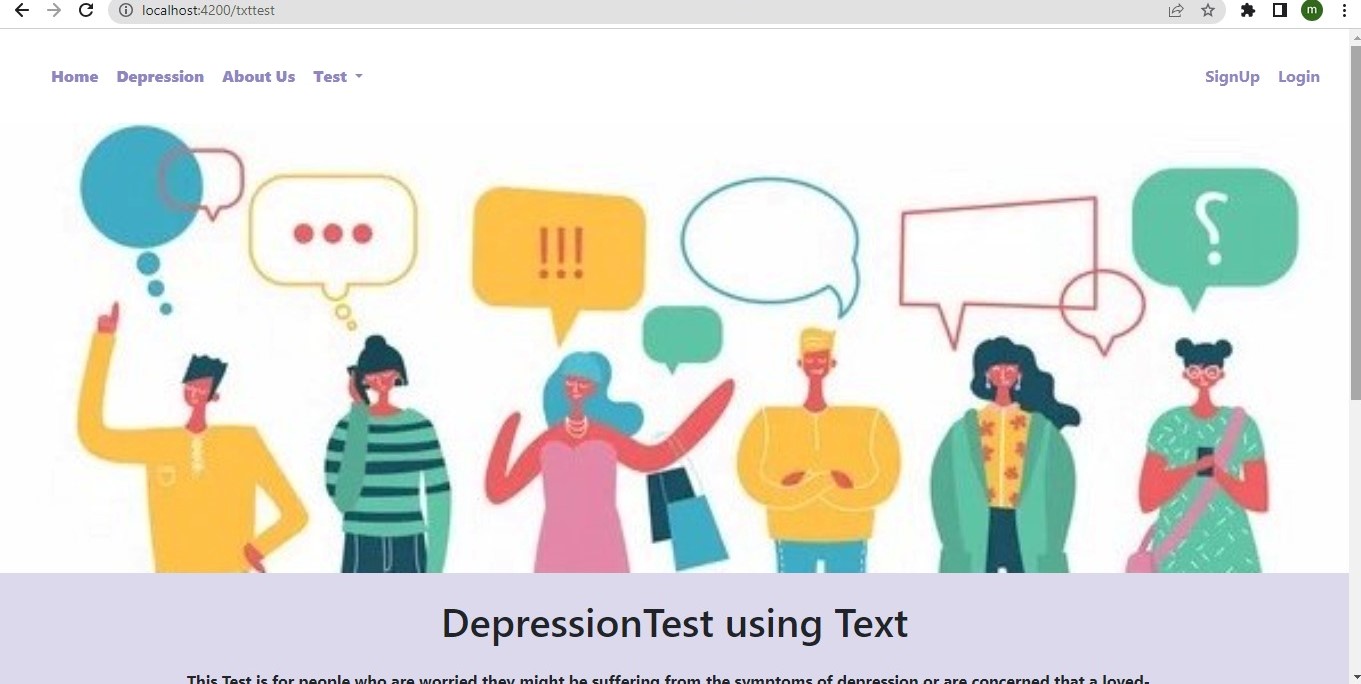


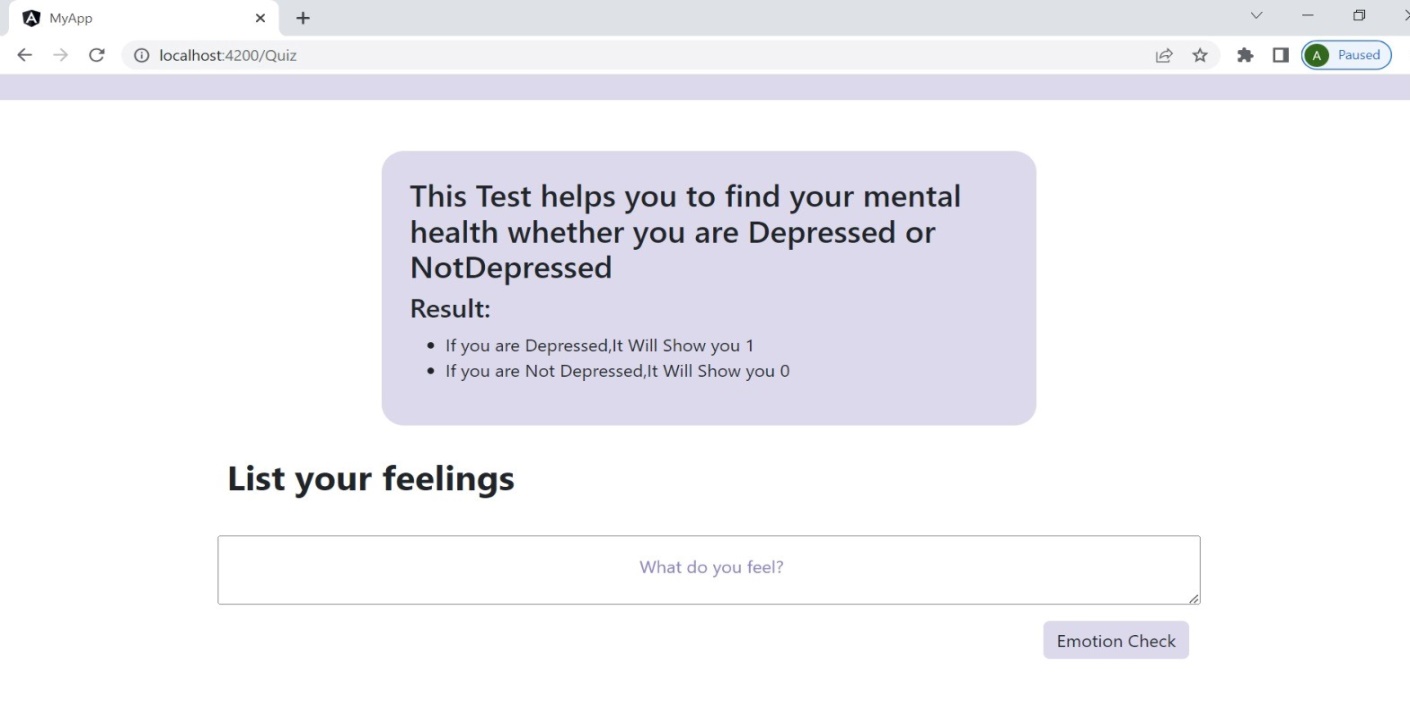


**Figure 24 Test Image page**

If the result of the first test, which is the image test, indicates that the image is portraying a sad expression, you can proceed to the second test, which is the text test.

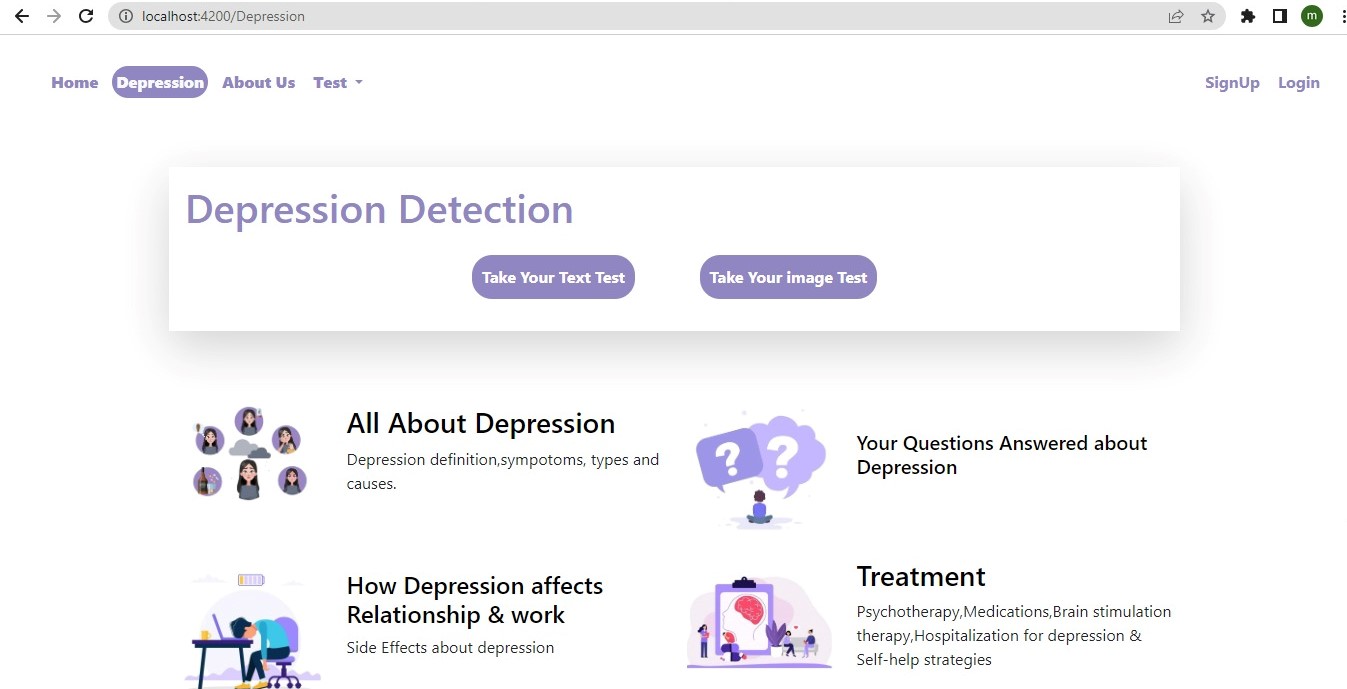
* 1. The Test Image Page





**Figure25: Test Text page**

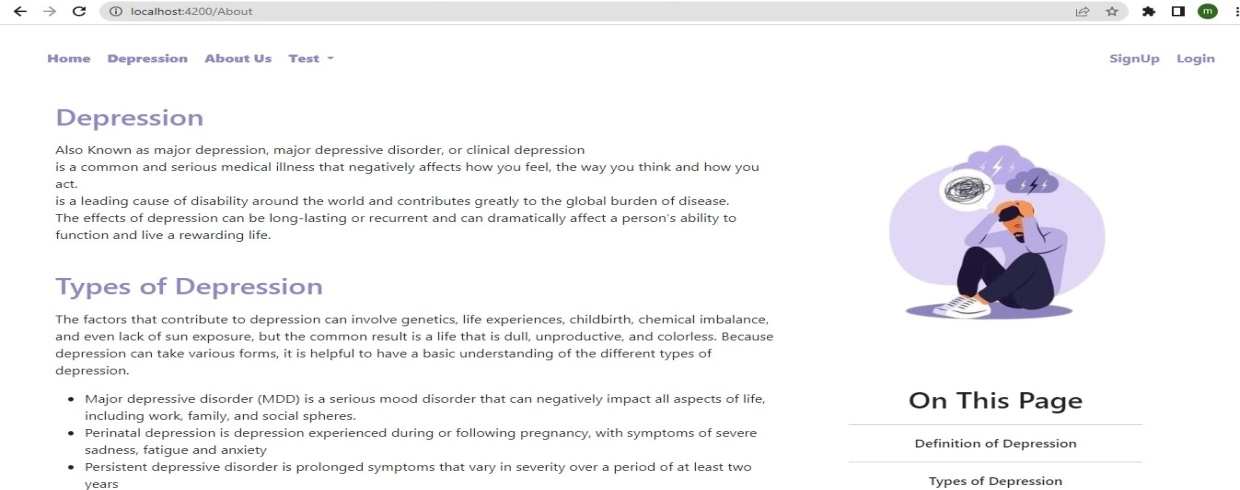
* 1. depression (blog) page



**Figure 26: depression (blog) pages**

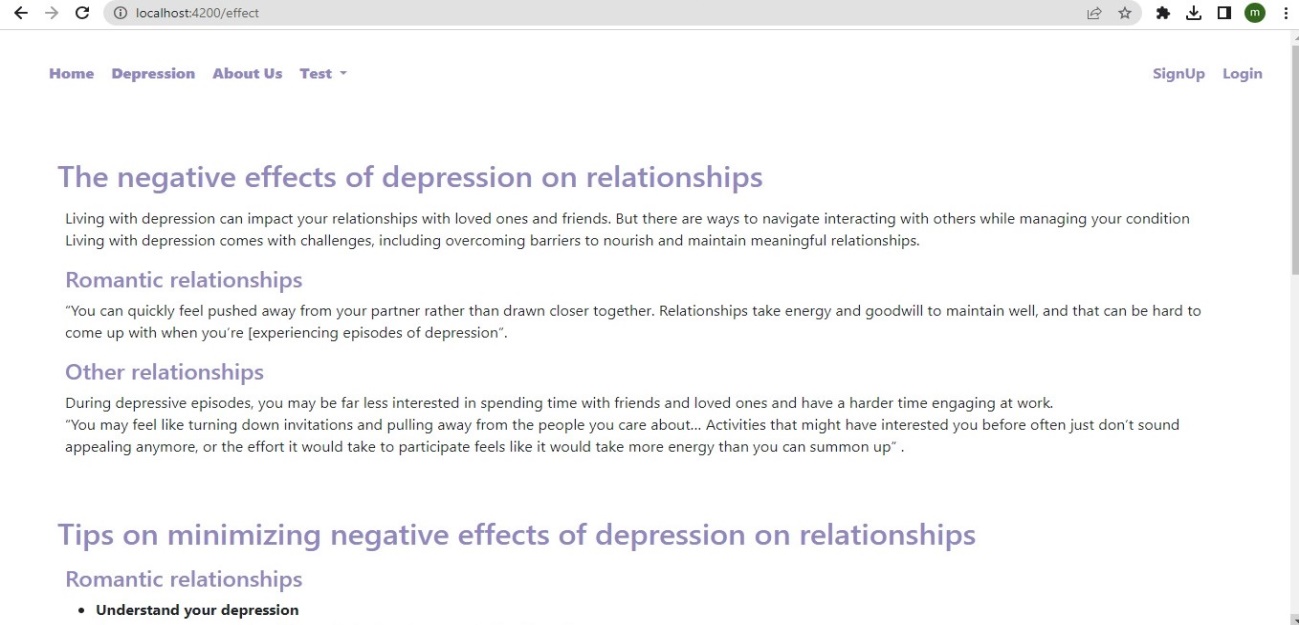
6.6.1 All about Depression page

This page has information about depression such as depression definition, symptoms, types and causes.



**Figure 27:** All about Depression page

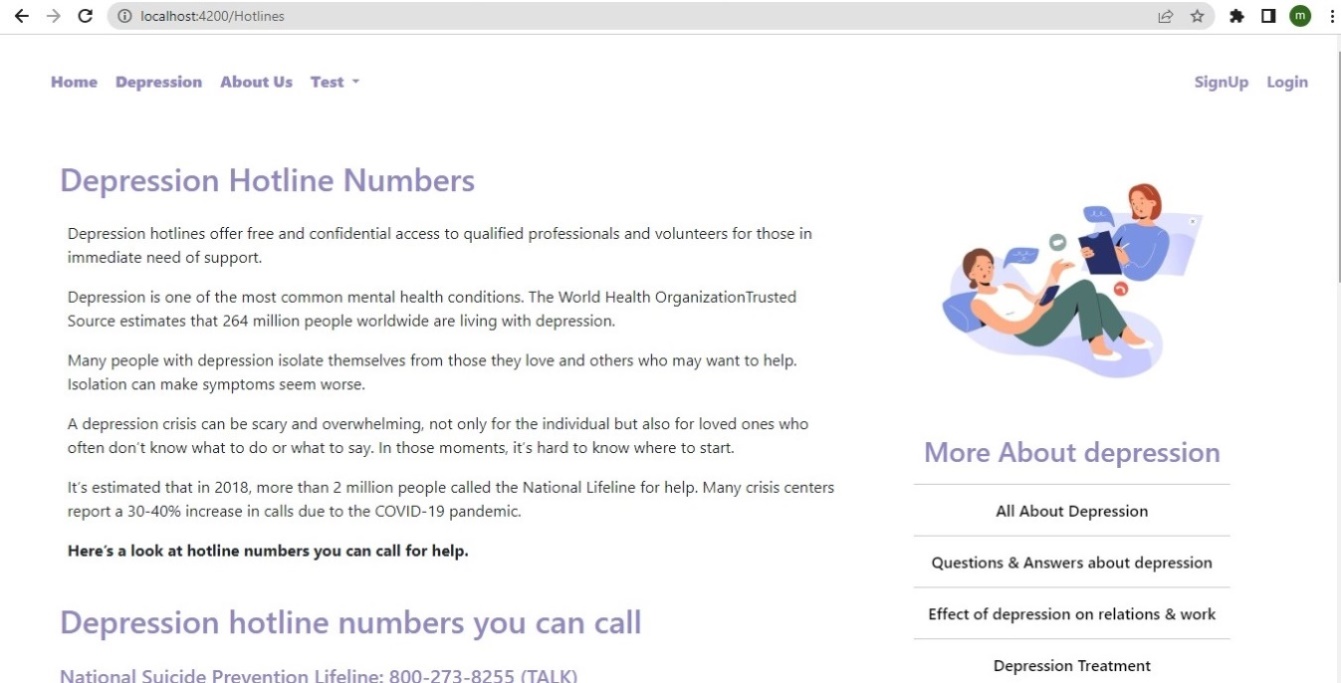
6.6.2 How Depression affects Relation &Work page



**Figure 28:** Depression affects Relation &Work page

6.6.3 Hotline Numbers page

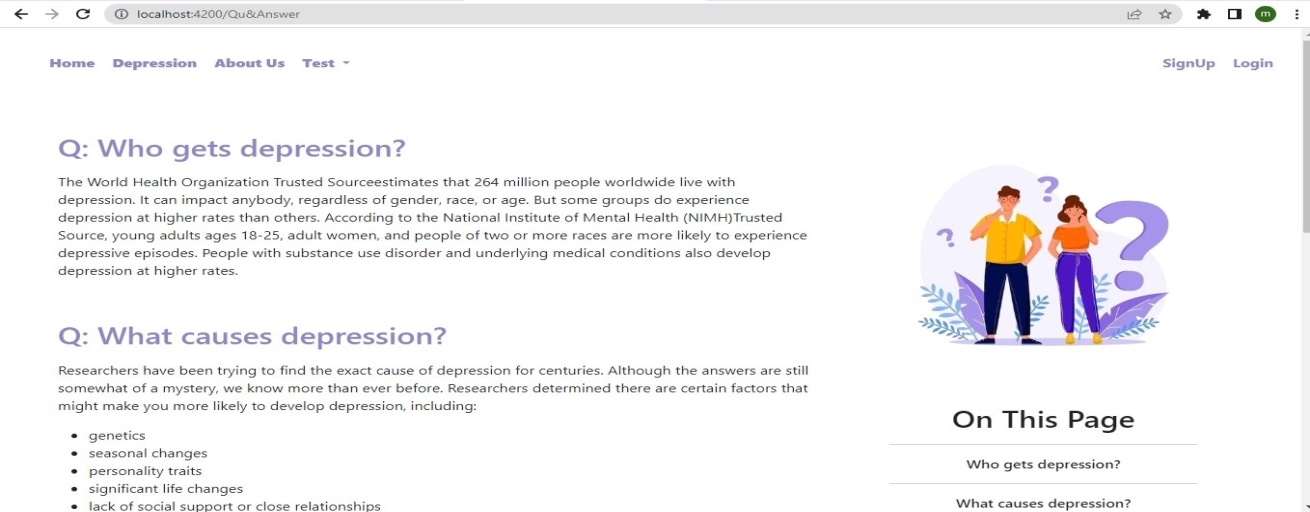
This page talks about depression hotline numbers you can call for help & when should you call them

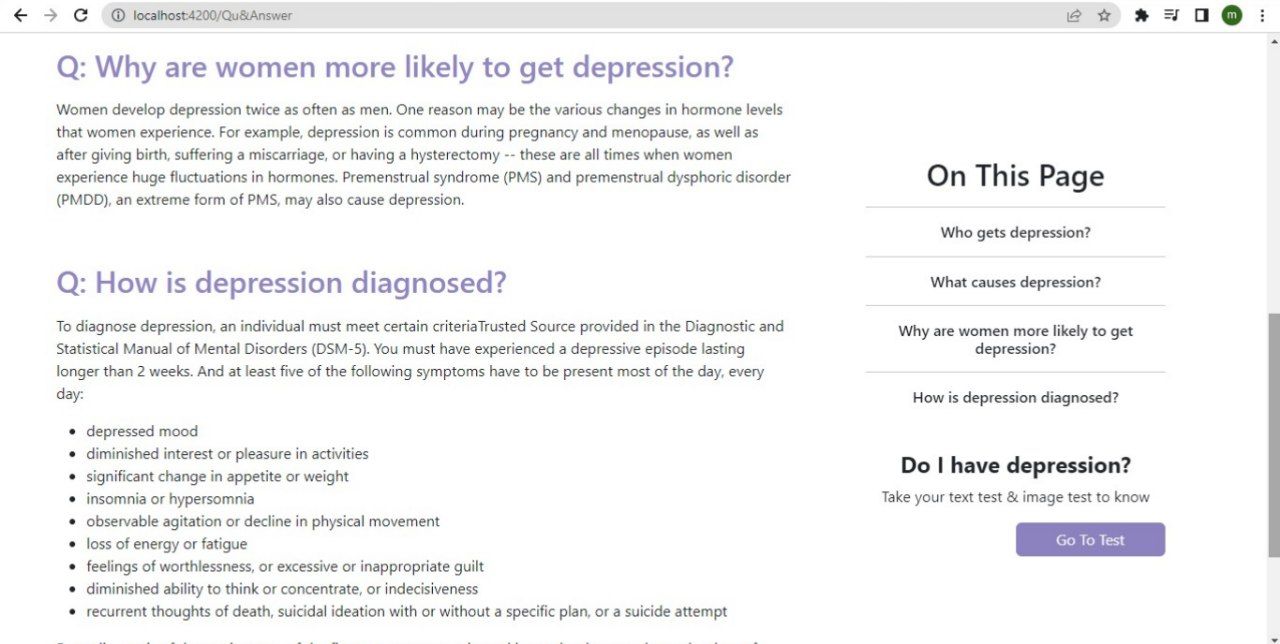


**Figure 29:** Hotline Numbers page

6.6.4 Your Questions Answered about Depression page

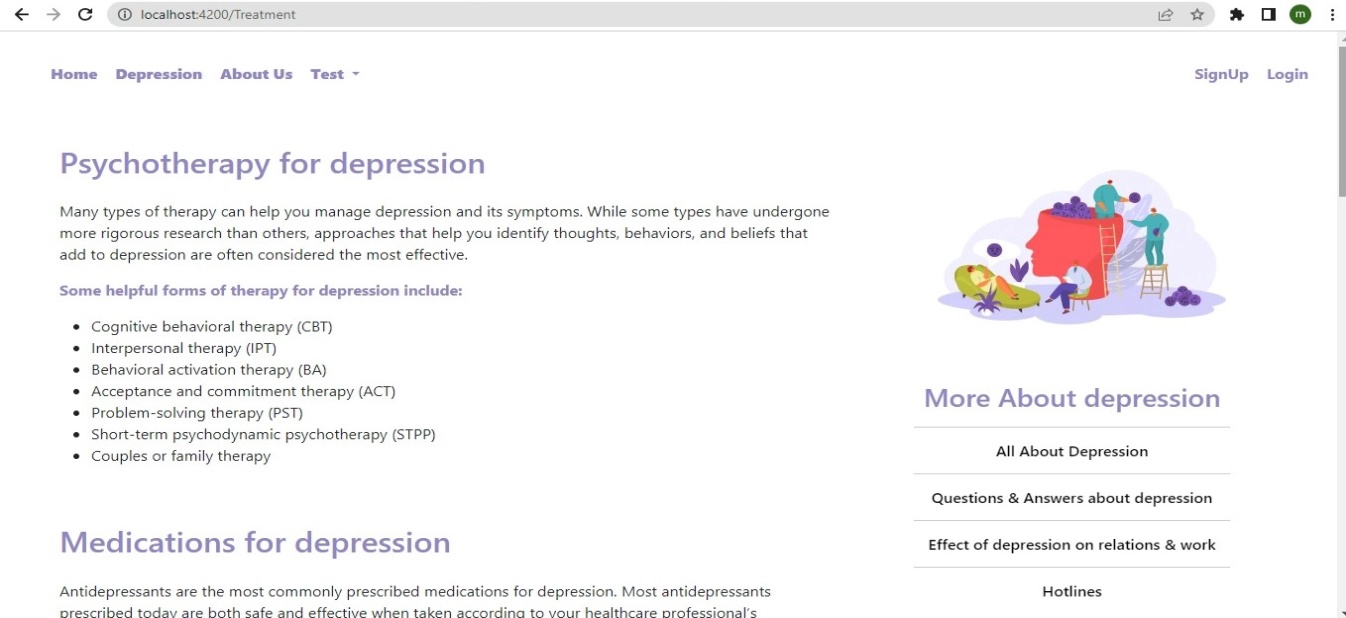
This page answers many questions about depression



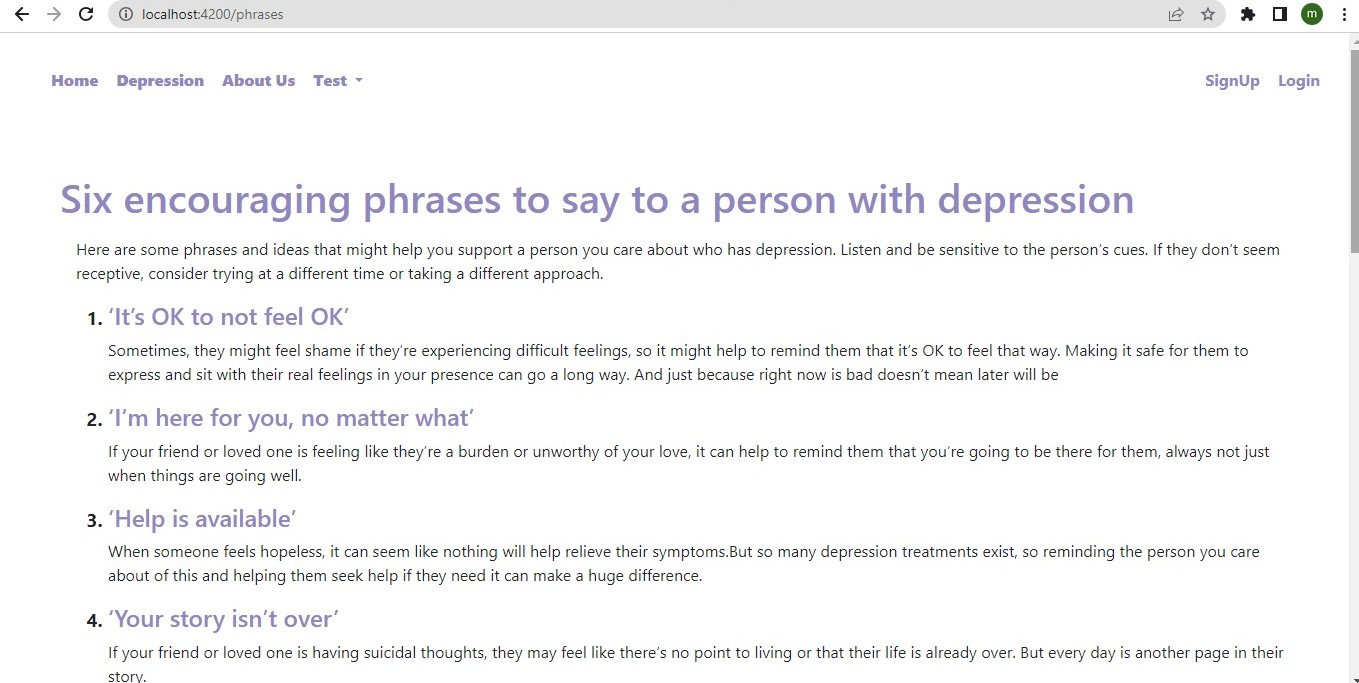


**Figure30:** You’re Questions Answered about Depression page

6.6.5 Treatment

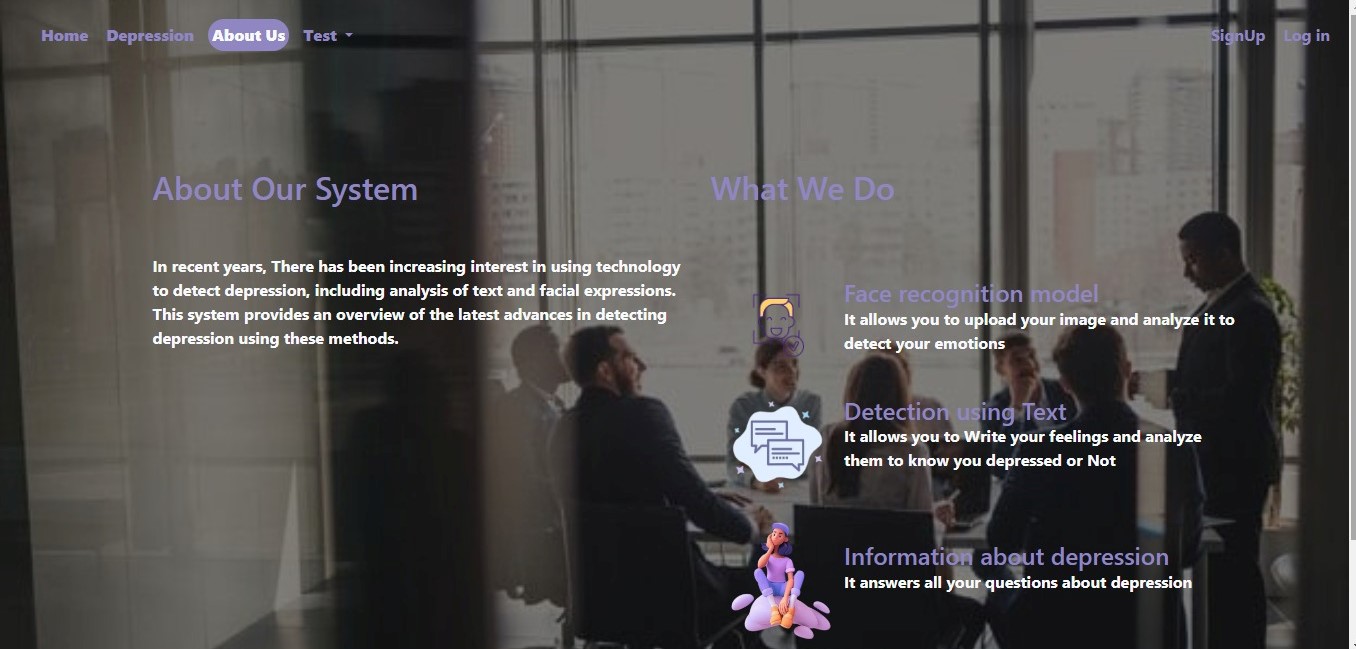
This page talks about ways to treat depression such as depression Psychotherapy, Medications,Brain stimulation therapy,Hospitalizat for depression & Self-help strategies **Figure31:** Treatment

6.6.6 Phrases page

This page talks about encouraging & discouraging phrases that you say or not to a person with depression

**Figure33:** Phrases page

* 1. About-Us page



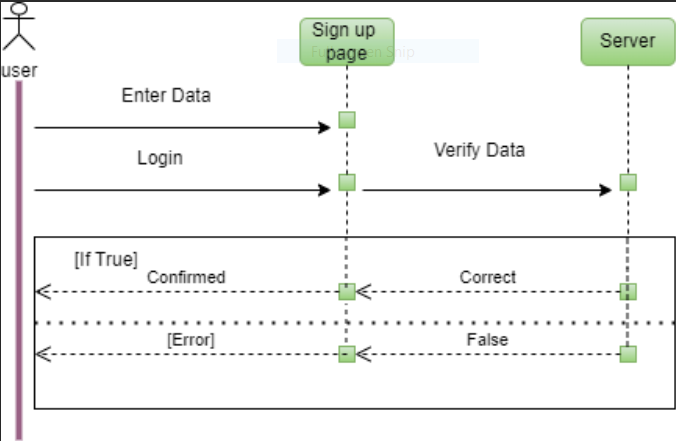
**Figure31:** About Us page

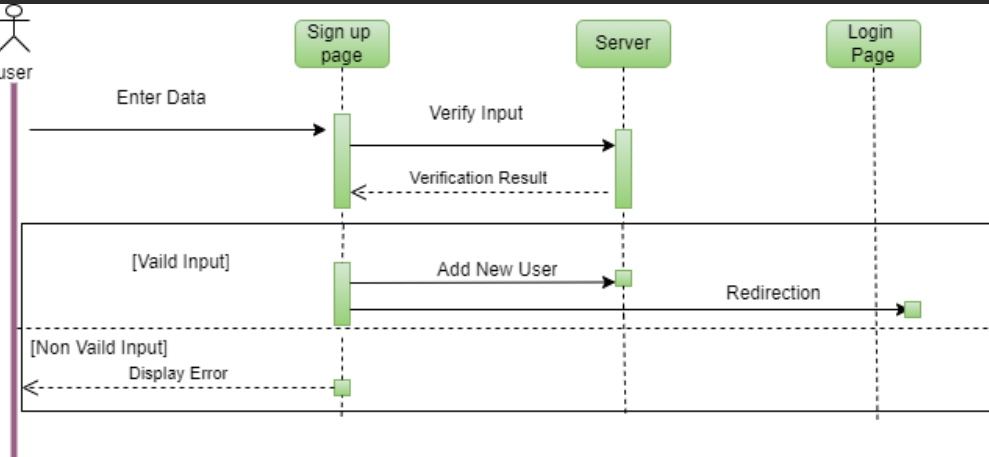
***Chapter seven***

# *Analysis and Diagrams*

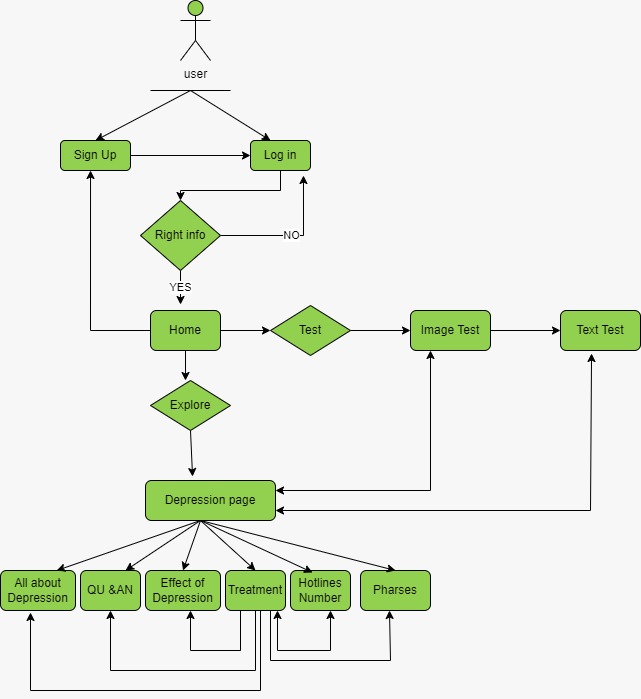


* 1. Sequence Diagram:





* 1. Activity Diagram:



* Conclusion

Depression is a prevalent mental health condition that affects millions of people worldwide. It can have a profound impact on an individual's overall well-being and quality of life. While the exact causes of depression are complex and multifaceted, there are various factors that can contribute to its development.

One significant factor is the presence of certain risk factors, such as a family history of depression, experiencing traumatic life events, or having certain medical conditions.

Advancements in technology have played a significant role in improving the diagnosis and treatment of depression. Machine learning and Deep learning algorithms have been developed to analyze vast amounts of data and identify patterns that may indicate the presence of depression. This technology has the potential to enhance early detection and intervention efforts.

Now we will comprise the results between machine learning and deep learning techniques in ml image the best result to SVM algorithm it is 73% with median Preprocessing filter in Depression/non Depression class , ml in text the best result to LR algorithm it is 96% in Depression/non Depression class ,In deep learning techniques in Image we comprised the CNN,Vgg16 and Vgg19 algorithms , .the accuracy CNN is 94 %,vgg\_16 is 90% , Vgg19 is 90% but we noticed that the best result of images is CNN. And deep learning techniques in Text we apply CNN algorithms and the accuracy of CNN is 92.39%.

* ***References***
* Image dataset from Kaggle <https://www.kaggle.com/datasets/muhammadusmansaeed/depression?select=train>
* Text dataset from kaggle <https://www.kaggle.com/datasets/nikhileswarkomati/suicide-watch>
* "SuicideWatch" from Dec 16, 2008(creation) till Jan 2, 2021
* [VGG Very Deep Convolutional Networks (VGGNet) - What you need to know - viso.ai](https://viso.ai/deep-learning/vgg-very-deep-convolutional-networks/)
* [World Health Organization (WHO)](https://www.who.int/)
* <https://www.analyticsvidhya.com/blog/2022/01/text-cleaning-methods-in-nlp/>
* <https://machinelearningmastery.com/clean-text-machine-learning-python/>
* <https://keras.io/examples/>
* <https://www.upgrad.com/blog/introduction-to-deep-learning-neural-networks-with-keras/>
* <https://arxiv.org/abs/1409.1556v6>
* <https://towardsdatascience.com/step-by-step-vgg16-implementation-in-keras-for-beginners-a833c686ae6c>
* <https://keras.io/examples/vision/image_classification_with_vision_transformer/>
* <https://medium.com/@mygreatlearning/everything-you-need-to-know-about-vgg16-7315defb5918>
* <https://medium.com/analytics-vidhya/depression-detection-using-machine-learning-28bbe0571d22>
* <https://flask.palletsprojects.com/>
* <https://www.youtube.com/playlist?list=PL-osiE80TeTs4UjLw5MM6OjgkjFeUxCYH>
* <https://www.researchgate.net/publication/358796994_Deep_Learning_for_Depression_Detection_from_Textual_Data>
* <https://www.frontiersin.org/articles/10.3389/fpsyt.2022.801301/full>
* <https://www.researchgate.net/publication/358612261_Depression_detection_using_emotional_artificial_intelligence_and_machine_learning_A_closer_review>
* <https://angular.io/>
* <https://getbootstrap.com/>
* <https://nodejs.org/en>
* <https://www.healthline.com/health/depression#treatment>
* <https://www.typescriptlang.org/>
* <https://material.angular.io/>