

# THE MAHARAJA SAYAJIRAO UNIVERSITY OF BARODA

## Faculty of Science



### Internship Report

*on*

**Emotion Recognition System: Data collection-preprocessing and Integration of Tkinter  
with OpenCV to store emotions into Database**

*for*

**Rishabh Software Pvt. Ltd**

*Submitted By:-*

**Jayati Goswami**

*In fulfillment of the degree*

*Of*

**Bachelors of Computer Applications**

*In*

**Department of Computer Applications**

July, 2020

**External Guide:**

Dr. Falguni Ranadive

Mr. Gagan Dubey

**Internal Guide:**

Mr. Kshitij Tripathi

**The Maharaja Sayajirao University of Baroda  
Faculty of Science  
Department of Computer Applications**

# INTERNSHIP REPORT

**Emotion Recognition System: Data collection-preprocessing and Integration of Tkinter with OpenCV to store emotions into Database**

**Date :- 7<sup>th</sup> July, 2020**

**Jayati Goswami**  
Seat Number: 322021

**Invigilator Signature: \_\_\_\_\_**

# Acknowledgment

It is with great pleasure that I present this document as a report of my work done at Rishabh Software Pvt. Ltd.

Success cannot be achieved single-handedly. For this reason, I would like to express my deepest gratitude to all those involved in making this project a success. This project would not have been successfully completed without the help of my mentors.

Firstly, I thank Prof. R.S. Srivastava (H.O.D.), associate coordinator of BCA program, who gave me this opportunity which will add to my experience.

I would also like to thank Mr. Kshitij Tripathi who provided me guidelines throughout the course of the internship. Not missing to thank my own C.A. Department, for giving me this opportunity to gain practical experience and enable me to understand the demand of the experience and professional environment. I have gained valuable experience which has given me confidence and has enhanced my professional skill for my competent career.

I would also like to thank all our teachers and family members and friends for helping, bearing, encouraging and boosting our morale, throughout our studies, especially during the project period. Last but not the least we would like to express our sincere gratitude to those who helped us in completing our project directly or indirectly.

I would like to thank Dr. Falguni Ranadive, Mr. Gagan Dubey, Sahil Bali, Paramtap Mewada, Swetika Sharma, and Balaji KS who have facilitated me and explained me the core concepts behind Machine Learning. They were ready to help me whenever I needed.

Special thanks to my friend, Vraj Kotwala, for helping and training and testing the model to identify emotions and to execute various statistical operations for further analysis on stored emotions. Without his help this project would not be possible

## Contents

1. Cover page	1
2. College certificate	2
3. Acknowledgement	3
4. Abstract	5
5. Introduction	6-12
5.1 Objective	7
5.2 Organization	8-12
5.2.1 Type of Organization	8
5.2.2 Products and Services provided by Organization	9-10
5.2.3 The Customers/Clients of the Organization	11
5.2.4 Major competitors of the Company	12
5.2.5 Major Strengths of the Organization	12
6. Internship Assignments	13-47
6.1 Method/Resources used to find this Internship	14
6.2 Departments that I worked for	15
6.3 Description of each Assignment with screenshot	16-
6.3.1 Assignment 1: Emotion Recognition	16-43
6.3.2 Assignment 2: Data Mining	44-45
6.4 Work Distribution	46
6.5 Challenges or difficulties faced during each assignment	47
6.6 Correlating with the Academics	48
7. Corporate Culture	49-55
7.1 Working Atmosphere	50
7.2 Culture within the Organization	51-55
7.2.1 Cultural Norms	51-53
7.2.2 Cultural Norms I picked up	54
7.2.3 Insights gained about the corporate culture	55
8. Communication	56-59
8.1 Ways in which I communicated within and outside the organization	57
8.2 Examples of good communication skills utilized by the employees	58
8.3 Methods to improve the effectiveness of communication within the organization	59
9. Career Reflection	60-62
9.1 Internship related to field of study	61
9.2 List of technologies learned during the Internship	62
10. Conclusion	63
11. References and Bibliography	64-65

## **Abstract**

During the period of my internship at Rishabh Software Pvt Ltd, I was given a proper training regarding the development of a project in professional and standardized manner by my guides, Dr. Falguni Ranadive and Mr. Gagan Dubey at the organisation.

This report contains the work done by me at the organisation from the period from Jan 2020 to June 2020. It reflects the challenges faced during the period of internship in terms of working on projects and mini assignments.

In this collaborative project of Emotion Recognition System, I collected and processed data also integrated Tkinter module with openCV library and modeled and connected the application with database to store emotions.

The internship report describes my objective of doing this internship and thereby meeting the goals I had set for myself as an individual in view of learning and matching up to the need of working as a professional in a corporate environment.

Some of the details of the project are not to be disclosed in this document as per company's policy.

- Emotion recognition is the process of identifying human emotion. The software detects seven emotions (Angry, Sad, Happy, Disgusted, Fearful, Neutral and Surprised).
- The data collection and data preprocessing for the emotions is the lengthiest process as it required precision.
- webcam captures your face and identifies the emotions and stores the same

# INTRODUCTION

## **Objective:**

One of the main objectives of choosing this internship was to expose myself to a particular job and a profession or industry. I had an idea about what a job is like, but getting into a real job is far more difficult than what we think as a student, so before entering into the corporate world, it's always preferred to get an internship to get the real time experience of what you'll be getting into in the near future.

Along with experience comes various aspects like acquiring practical skills, getting an opportunity to establish relationships with employers and employees in IT field. Choosing internship over the project helped me get familiar with the corporate culture. The corporate exposure which I received during the course of this internship is invaluable and it taught me discipline. This internship also helped me develop skills in the practical work situations.

## Organisation:



<https://www.rishabhsoft.com/>

## Type of Organisation:

Rishabh Software provides software, recruitment and engineering services and solutions to mid-sized enterprises globally. Through Rishabh's experience in mobile, web, cloud and analytics solutions, our technology clients get the best value for their IT budget. The company has been appraised for CMMi Level-3, ISO9001 and ISO27001 as well as been awarded for its Innovation in Gujarat. Since its inception in 1999, Rishabh Software has proudly delivered over 1000+ successful projects to its customers across 23 countries including USA, UK, Europe, Canada, Australia. It is both product-based and service-based company.



## **Products and services offered by Rishabh Software:**

### **Mobile App Development Services**

1. IOS App Development
2. Android App Development
3. Xamarin App Development
4. Cross Platform App Development

### **Enterprise Application Development Services**

1. App Re-Engineering + Modernization
2. Enterprise Application Integration Services
3. Application Maintenance

### **Cloud Development**

1. SAAS App Development
2. Cloud Integration

### **Web Application Development**

1. Drupal Development
2. Ecommerce
3. Web Development Services
4. Magento Development
5. Zen Cart Development
6. Portals + Collaboration

### **PHP Development Services**

1. Laravel Development

### **Microsoft Development Services & Solutions**

1. Net Development
2. Sharepoint Development
3. Dynamics CRM

### **Front End Development**

1. Angular Development

## **Business Intelligence Services**

## **Java Development Services & Solutions**

## **QA & Software Testing Services**

## **Business Process Outsourcing (BPO)**

1. Recruitment Process Outsourcing
2. Payroll Processing
3. Tax Form Processing
4. Accounting and Bookkeeping

## **Engineering Design Services:**

1. Engineering Services
2. CAD Design Services
3. Stress Analysis Services
4. Pre-Project Planning

## **Industry:**

1. Retail
2. Education
3. Finance
4. Healthcare
5. Digital Enterprise
6. Logistics & Supply Chain
7. Telecom

## Some of their clients:



Figure 1: Rishabh Software's Clients

- |                       |                                      |                          |
|-----------------------|--------------------------------------|--------------------------|
| 1. Zero Chaos         | 10. Dhara                            | 19. Mole Productions     |
| 2. Bausch + Lomb      | 11. Far Heap Solutions Inc           | 20. Interactyx           |
| 3. Siemens            | 12. Fisher outdoor Leisure           | 21. Innovative Creations |
| 4. ABB                | 13. Xciton                           | 22. Health Screen        |
| 5. Absolutely New     | 14. The George Washington University | 23. GHG Corporation      |
| 6. Aditya Birla Group | 15. Supreme                          | 24. Genera interactive   |
| 7. Alembic            | 16. Solid Source                     | 25. Garden               |
| 8. Bayer              | 17. Omnia Consulting                 | 26. Texas Applications   |
| 9. Bupa               | 18. National Career Fairs            | 27. Vipana               |

## **Major Competitors of the Organization:**

1. TCS
2. Sanovi
3. Ncode Technologies
4. Healthera
5. Accenture
6. Infosys
7. Wipro
8. Cognizant
9. Appiqo Technologies
10. Xicom Technologies Ltd.
11. Mindinventory Company
12. WebSenor InfoTech

## **Major Strengths of the Organization**

1. Nice work environment
2. Spread over 100,000 sq. ft. infrastructure
3. Good and open work culture
4. The company has CMMI DEV3 certifications and,
5. ISO 9001 Bureau Veritas Certifications
6. Provides software for most industries.

## **INTERNSHIP ASSIGNMENTS**

## Resources used to find this Internship:

I had collected a list of IT companies in vadodara from the internet. Also, had few references of companies.

My friend and the co-partner in the project suggested me to visit Rishabh Software. I also gave interviews in few companies before Rishabh software was suggested to me.

After visiting Rishabh Software, they conducted a personal interaction session after which I decided to work for them as a trainee/intern. Rishabh software also develops python based software that is what I was searching for, hence I started working for Rishabh Software.

I created the following list of companies in Vadodara before joining as an intern at rishabh software:

Adri IT solutions	088662 44200	contact@adriitsolutions.com
Aimtron Electronic Pvt. Ltd.	096876 32057	hrindia@aimtron.com
Arth technology	093758 20283	contact@arthtechnology.com
Bharti soft pvt ltd	2656544312	-
Byte Graffix 0265 653 7968 -	0265 653 7968	-
Byte Technologies	093281 83331	info@byte-technology.com
Collabera	-	-
Digiflux IT solutions	96017 20466	info@digiflux.io
E-Intelligence	090330 04622	careers@e-intelligence.in
GeekConnects	094032 89972	geekconnects@gmail.com
Helios solutions	0265 298 0949	info@heliossolutions.co
JDM IT Solutions	+91 7984877461	contact@jdmit.in
Mastercard	-	-
PRIZMA web solutions	91 9409 701 233	-
Rishab Software Pvt. Ltd	085111 22697	career@rishabhsoft.com
ScrumBeesTechnologies	942 733 0294	hr@scrumbees.com
Softcom Technologies	098790 21944	support@siliconbrain.co.in
Webmyne Systems	0265 233 6907	-
WhirlWind	0265 231 1166	saumil@whirlwind.co.in

## **Departments of Organization in which I worked:**

I worked under the RISE department. Rishabh Integrated Skill Enhancement (RISE) program is a skill enhancement initiative. The department offer handcrafted courses for students and working professionals.

The RISE comprises of academic internships in IT domains across Python, Java, PHP, .Net, Mobile Application Development (Android & iOS), Quality Assurance, Graphical user Interface, Data Analysis and many more.

RISE follows a 70-20-10 % rule to improve the efficiency and effectiveness to help the students. The 70% content emphasis is on Experiential learning, with 20% on Social and 10% on Formal education.

My first project was the part of Rishabh's selection process, where I was given a project on Trip Analysis which was based on Java.

The second project (main project after selection) was a collaborative project on Emotion Recognition System which was based on Python. I worked as a software developer, where I learnt everything from scratch regarding the project development in a professional environment.

## **Description of each Assignment with screenshot**

### **Assignment 1: Emotion Recognition a collaborative project – Emotion Recognition with Convolutional Network includes:**

1. Dataset collection.
2. Data Pre-processing
3. Classification of Dataset
4. LOGO Design (done by my collaboration partner)
5. Graphical User Interface design
6. Training the model (done by my collaboration partner)
7. Testing the model (done by my collaboration partner)
8. Modelling and creating the Database to store emotions
9. Statistical Analysis of Stored Emotions (done by my collaboration partner)

### **Testing Tools Used:**

1. Anaconda to maintain dependencies.
2. Jupyter IDE

### **Technologies and Packages/Modules used:**

1. Python 3.7
2. NumPy
3. Open-CV 2
4. MySQL-Connector
5. Time
6. Tkinter
7. PIL (Python Imaging Library)



## **Project definition:**

This project aims to identify emotions of a person using image recognition from the predefined seven emotions: happy, sad, angry, disgusted, fearful, neutral and surprised, using deep convolutional neural networks. The model is trained and tested using supervised learning from the collected dataset. The detected emotions are stored in the database for further statistical analysis.

## **Pros and Cons :**

### **Pros:**

1. Emotions are identified using image recognition which makes the prediction more accurate than the emotions identified on the basis of written text.
  - a. For instance, if a person tweets about his bad experience of watching a movie, he might use words like disgusting, not good, sad etc. which might confuse the model to predict the accurate emotion because of the overload of emotions in a single tweet. On the contrary if he expresses his experience through a video, it becomes easier to identify his real emotions without any confusion.
2. Emotions of a person throughout the day can be detected and the data can be used for statistical analysis.
  - a. For instance, if an employee at work is being monitored throughout his working hours, his emotion cycle throughout the day can be predicted and also how satisfied the employee was at the end of the day, such kind of analysis can be helpful to the companies for human resource management.

### **Cons:**

1. Inner expressions or emotions cannot be identified as it traces emotions on the basis of facial expressions. Hence, predictions are not necessarily accurate.
2. Being a supervised model it fails to recognize emotions other than the predefined emotions.

## **1.Data Collection:**

Data collection is the process of gathering and measuring information from countless different sources. In order to use the data we collect to develop practical artificial intelligence (AI) and machine learning solutions, it must be collected and stored in a way that makes sense for the business problem at hand.

Data collection is the single most important step in solving any machine learning problem. However, it is also a critical roadblock for many researchers and data scientists. An inordinate amount of time is usually spent on data collection, which largely consists of data acquisition, data labeling, and improvement of existing data or models. Diving head first into projects without considering the right data collection process often don't lend required results. Fortunately, there are many data collection tools to help prepare training datasets quickly and at scale.

The best data collection tools are easy to use, support a range of functionalities and file types, and preserve the overall integrity of data.

**Considering various data collection techniques and keeping the requirements in mind following techniques have been used for data collection:**

### **Open-Source Datasets**

Another way to obtain raw data for machine learning is to obtain pre-built, publicly available datasets on the internet. There are thousands of publicly available datasets spanning a wide range of industries and use cases.

Various open source datasets are available publically for research purpose and for various competitions held in the field of machine learning and data science. Considering the requirements of project consisting of 7 emotions:angry, sad, happy,disgusted,surprised,fearful and neutral open source datasets for kaggle have been of great use.40-50% of the data has been extracted from famous kaggle datasets like FER-2013 and CK-48 based on the availability of required emotions. These datasets havn't been updated over time and were not accurate enough to start modelling on their basis.

Links of the open source datasets:

<https://www.kaggle.com/shawon10/ckplus?>

<https://www.kaggle.com/ahmedmoorsy/facial-expression>

## **Data Scraping**

Web scraping describes the automated, programmatic use of an application to extract data or perform actions that users would usually perform manually, such as social media posts or images. Images can be scrapped from videos to gather dataset

All the open-source datasets have been collected either using scrapping, augmentation or by using synthetic data generators. First hand researchers use these techniques to collect dataset as per their requirements and upload them publically for use.

After collecting data from kaggle references :FER-2013 and CK-48 ,data scrapping from video Is used to extract required images for various emotions.as the already available data was not sufficient enough and also outdated containing lots of constraints. Videos were downloaded from web and images were extracted using extraction code while applying grayscaling and resizing functions provided by openCV.

### **Face detection using haarcascade file provided by OpenCV and saving the extraxted data into dataset folder:**

OpenCV comes with a trainer as well as detector. If you want to train your own classifier for any object like car, planes etc. OpenCV already contains many pre-trained classifiers for face, eyes, smile etc. Those XML files are stored in `opencv/data/haarcascades/` folder.

First XML classifiers are loaded. Then the input image (or video) is loaded in grayscale mode .Once the faces are identified

### Code snippet:

```
import cv2
scale_factor = 1.3
min_neighbors = 3
min_size = (100, 100)
webcam=False
def detect(path):
    cascade = cv2.CascadeClassifier(path)
    video_cap = cv2.VideoCapture("videoplayback.mp4")
    i=0
    while True:
        # Capture frame-by-frame
        ret, img = video_cap.read()

        #converting to gray image for faster video processing
        gray = cv2.cvtColor(img, cv2.COLOR_BGR2GRAY)
        #gray = cv2.cvtColor(img, cv2.COLOR_BGR2RGB)

        #rects = cascade.detectMultiScale(gray, scaleFactor=scale_factor,
minNeighbors=min_neighbors,
#minSize=min_size)
        rects = cascade.detectMultiScale(gray, scaleFactor=scale_factor,
minNeighbors=min_neighbors, minSize=min_size)
        # if at least 1 face detected
        if len(rects) >= 0:
            # Draw a rectangle around the faces
            for (x, y, w, h) in rects:
                face = gray[y:y+h,x:x+w]
cv2.rectangle(img, (x, y), (x + w, y + h), (0, 255, 0), 2)

cv2.imwrite('C:/Users/jayat/OneDrive/Desktop/EmotiBot/VidToImage/Dataset/'+str(i)+
'.jpg',face)

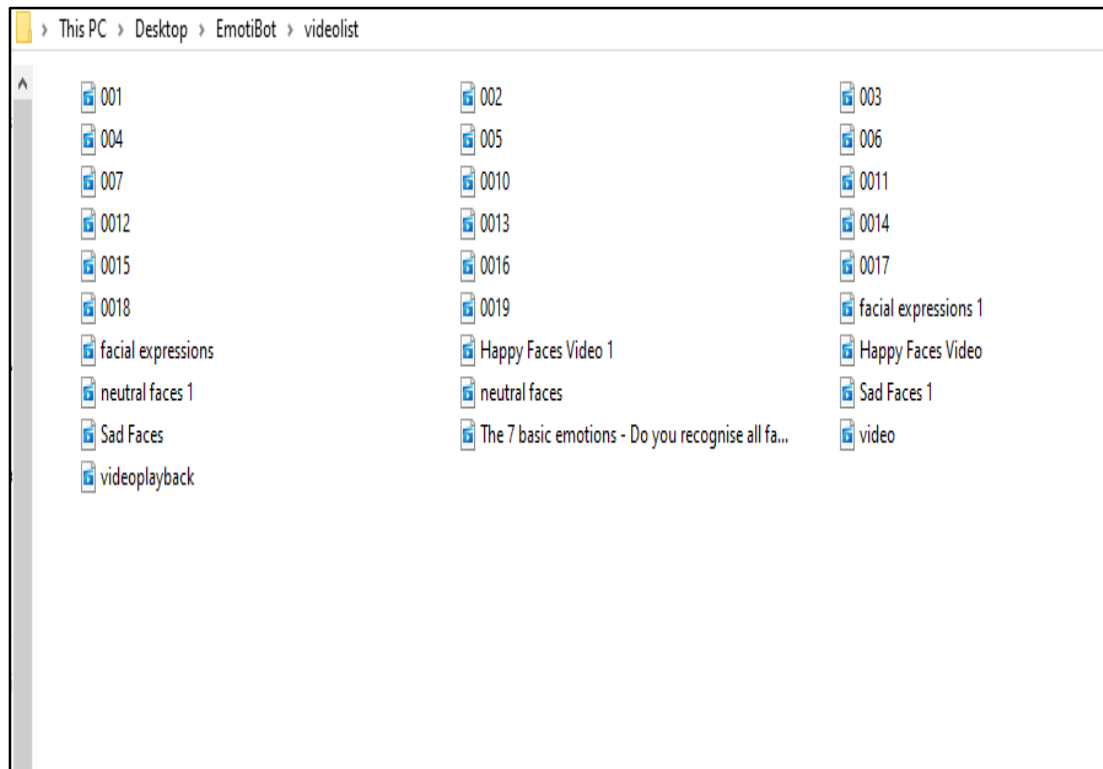
            i+=1
            # Display the resulting frame
            cv2.imshow('Face Detection on Video', img)

            #wait for 'c' to close the application
            if cv2.waitKey(1) & 0xFF == ord('c'):
                break
        video_cap.release()
def main():
    cascadeFilePath="C:/Users/jayat/OneDrive/Desktop/EmotiBot/VidToImage/haarcas
cade.xml"
    detect(cascadeFilePath)
    cv2.destroyAllWindows()

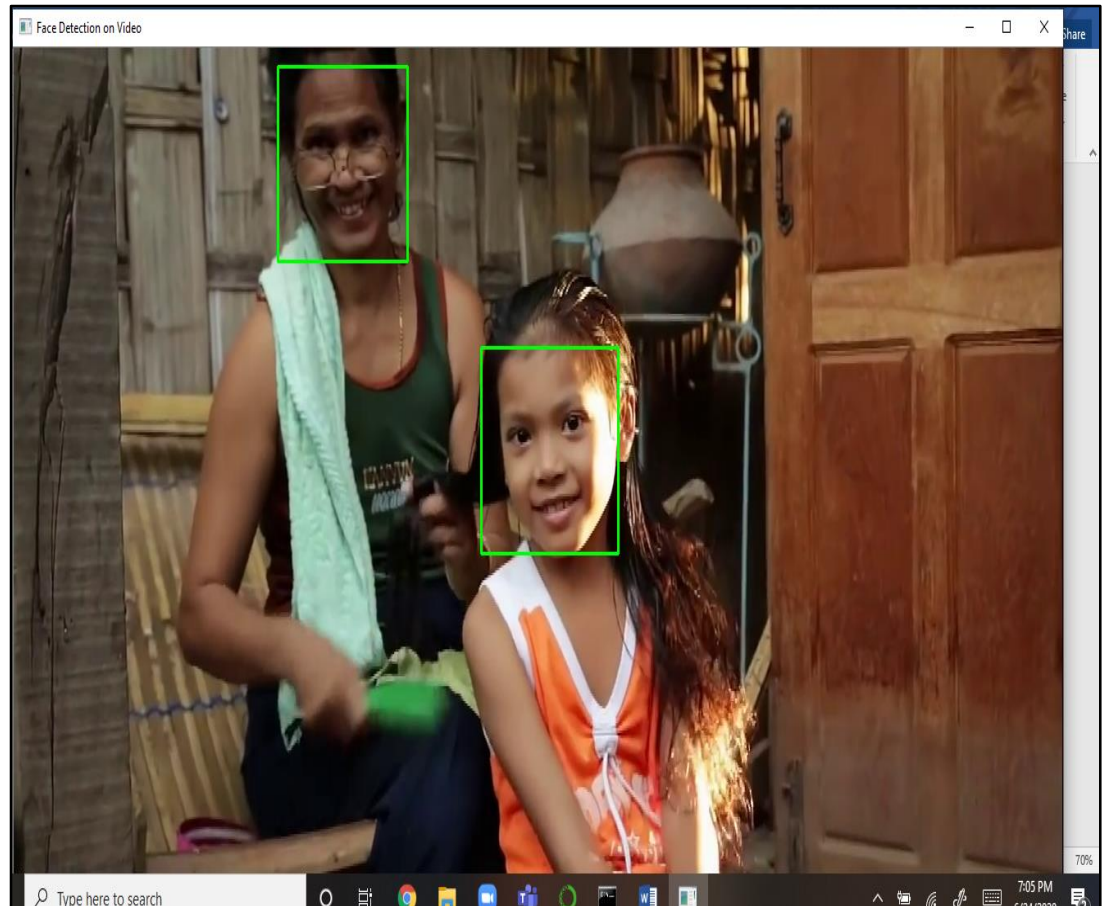
if __name__ == "__main__": main()
```

## Screenshots:

### 1. Video files:



## 2. Runtime:



### 3. Collected raw data:



## 2.Data preprocessing:

### Why Data Preprocessing?

When we talk about data, we usually think of some large datasets with huge number of rows and columns. While that is a likely scenario, it is not always the case — data could be in so many different forms: Structured Tables, Images, Audio files, Videos etc..Machines don't understand free text, image or video data as it is, they understand 1s and 0s. So it probably won't be good enough if we put on a slideshow of all our images and expect our machine learning model to get trained just by that! Hence data preprocessing comes into picture

### What is Data Preprocessing?

Data preprocessing is a data mining technique that involves transforming raw data into an understandable format. Real-world data is often incomplete, inconsistent, and/or lacking in certain behaviors or trends, and is likely to contain many errors. Data preprocessing is a proven method of resolving such issues. Data preprocessing prepares raw data for further processing.

Digital image processing is the use of computer algorithms to perform image processing on digital images. As a subfield of digital signal processing, digital image processing has many advantages over analogue image processing. It allows a much wider range of algorithms to be applied to the input data — the aim of digital image processing is to improve the image data (features) by suppressing unwanted distortions and/or enhancement of some important image features so that our AI-Computer Vision models can benefit from this improved data to work on.

An image is nothing more than a two-dimensional array of numbers(or pixels) ranging between 0 and 255. It is defined by the mathematical function  $f(x,y)$  where  $x$  and  $y$  are the two co-ordinates horizontally and vertically. The value of  $f(x,y)$  at any point is giving the pixel value at that point of an image.

To train a network and make predictions on new data, your images must match the input size of the network. If you need to adjust the size of your images to match the network, then you can rescale or crop your data to the required size.



Following steps are required can be followed for data preprocessing:

1. Read the image
2. Resize the image
3. Grayscale of image
4. Remove noise(Denoise)

### 1. Read the image:

In this step, we store the path to our image dataset into a variable to load folders containing images into arrays.

#### Code snippet:

```
img=cv2.imread('C:/Users/jayat/OneDrive/Desktop/EmotiBot/VidToImage/
Dataset/'+str(i)+'.jpg')
```

### 2. Resize the image:

#### Why do we resize our image during the pre-processing phase?

Some images captured by a camera and fed to our AI algorithm vary in size, therefore, we should establish a base size and scale factor for all images fed into our AI algorithms.

#### Code snippet:

```
scale_factor = 1.3
min_neighbors = 3
min_size = (100, 100)
rects = cascade.detectMultiScale(gray, scaleFactor=scale_factor,
minNeighbors=min_neighbors, minSize=min_size)
```

### 3. Grayscale of image:

A **grayscale** (or graylevel) **image** is simply one in which the only colors are shades of gray. The reason for differentiating such **images** from any other sort of color **image** is that less information needs to be provided for each pixel

### Why Grayscale is required?

An RGB Image consists of 3 layers R,G,B as it is clearly see through its name. It's a 3 dimensional matrix, for example, 3 consecutive pages in your book. where grayscale image is of only 2 dimensions, and the values ranges between 0–255 (8-bit unsigned integers).Therefore, some algorithms can only applied on 2-D image effeciently rather than 3-D, hence we convert an RGB image into a grayscale image, for instance, Black and White conversion of an image, convolution of an image, etc.

### Code snippet:

```
gray = cv2.cvtColor(img, cv2.COLOR_BGR2GRAY)
```

#### 4. Remove Noise(Denoise):

One of the fundamental challenges in image processing and computer vision is Image denoising. What denoising does is to estimate the original image by suppressing noise from the image. Image noise may be caused by different sources ( from sensor or from environment) which are often not possible to avoid in practical situations. Therefore, image denoising plays an important role in a wide range of applications such as image restoration, visual tracking, image registration, and image segmentation. While many algorithms have been proposed for the purpose of image denoising, the problem of image noise suppression remains an open challenge, especially in situations where the images are acquired under poor conditions where the noise level is very high.

**Gaussian noise:** "Each pixel in the image will be changed from its original value by a (usually) small amount. A histogram, a plot of the amount of distortion of a pixel value against the frequency with which it occurs, shows a normal distribution of noise. While other distributions are possible, the Gaussian (normal) distribution is usually a good model, due to the central limit theorem that says that the sum of different noises tends to approach a Gaussian distribution."

**cv2.fastNlMeansDenoisingColored()** function is the implementation of **Non-local Means Denoising** algorithm for colored images but when we apply gray scale function, **cv2.fastNlMeansDenoising (img, None ,10,7,21)**, it seems to be working better denoising black and white picture with Gaussian noise:

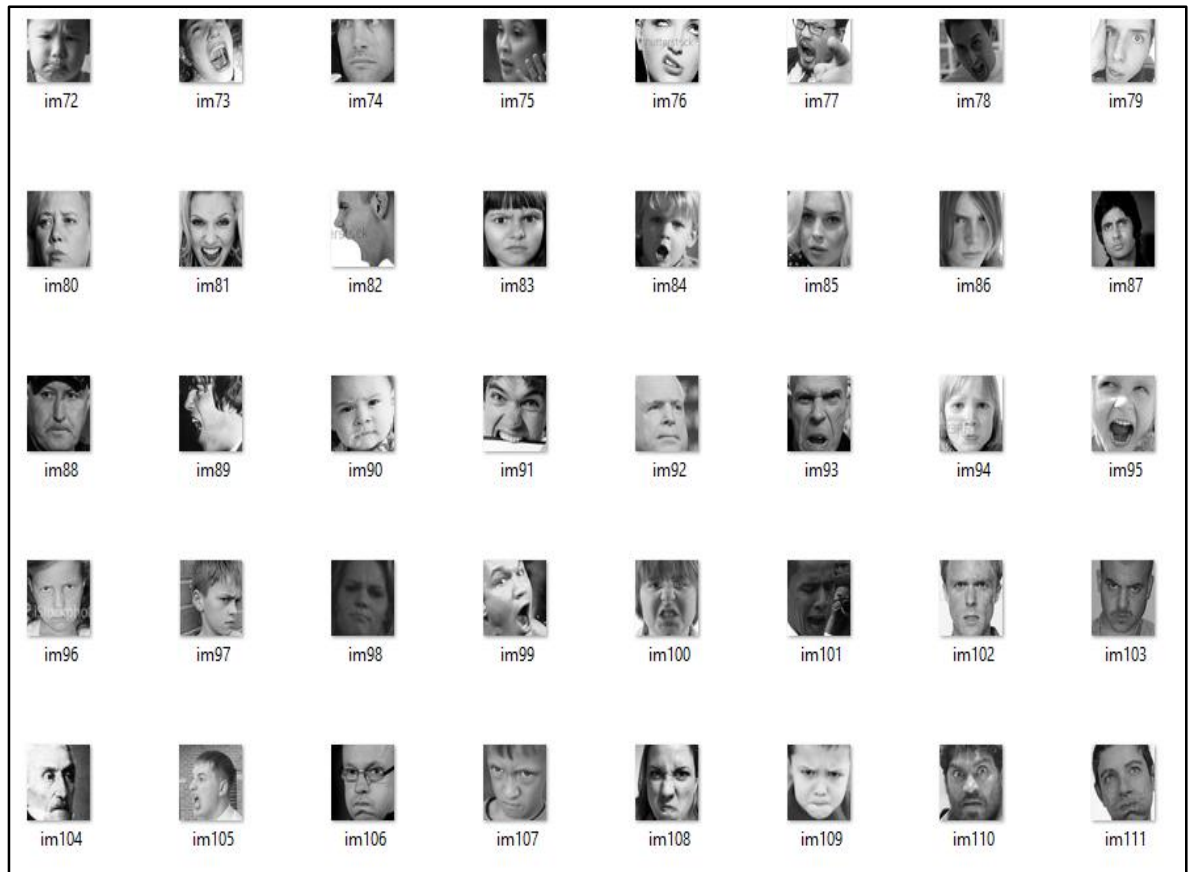
**code snippet:**

```
import numpy as np
import cv2
from matplotlib import pyplot as plt
img=cv2.imread('Desktop/EmotiBot/VidToImage/Dataset/1.jpg')
b,g,r = cv2.split(img)          # get b,g,r

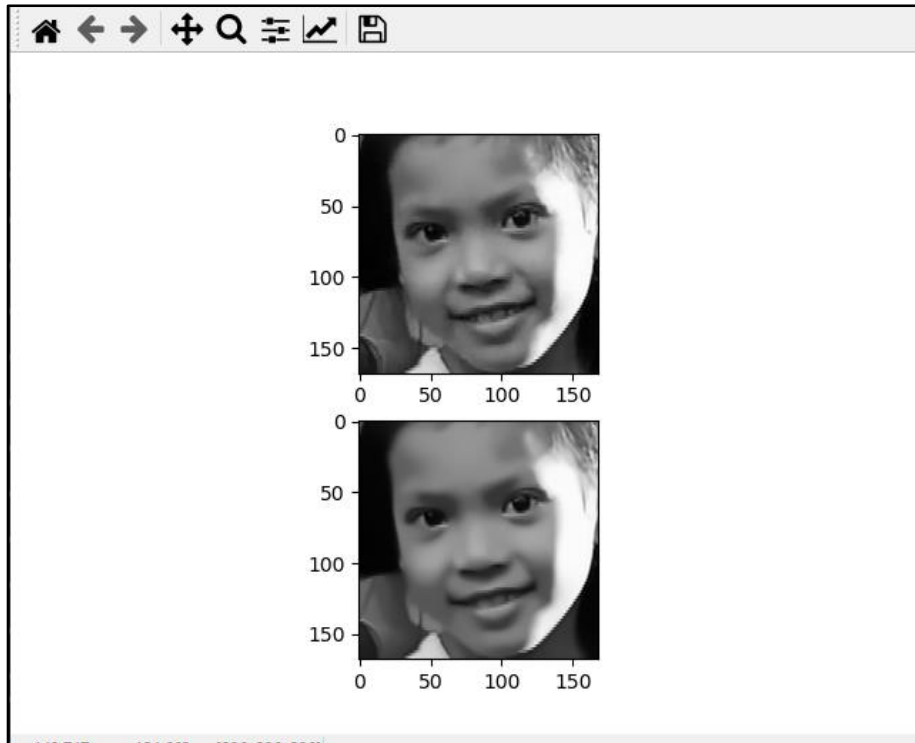
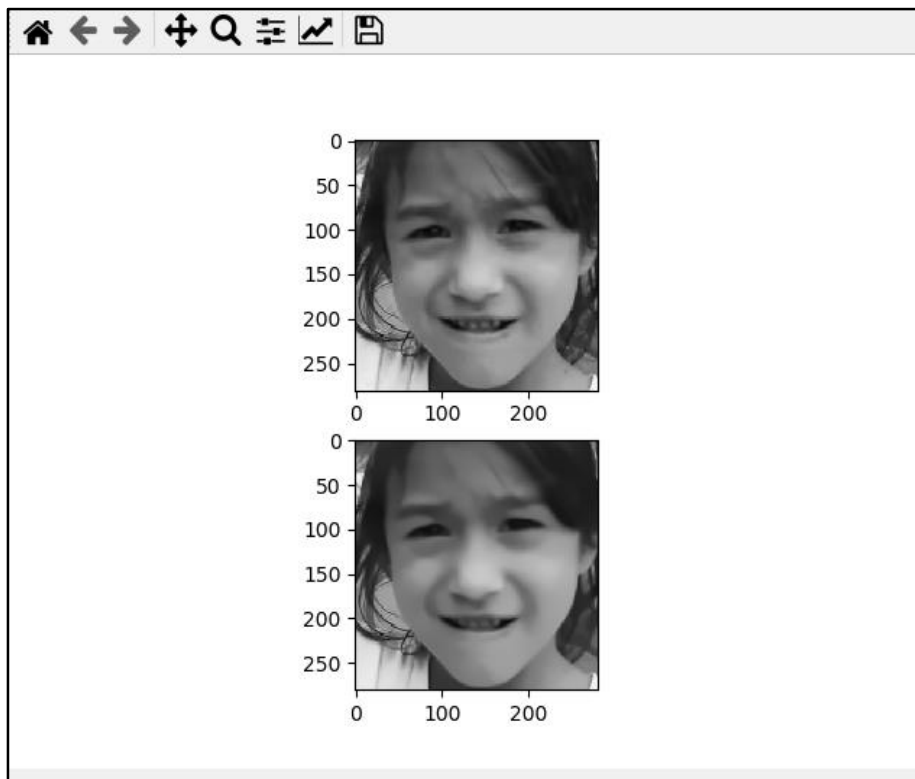
rgb_img = cv2.merge([r,g,b])    # switch it to rgb
#Denoising
#dst = cv2.fastNlMeansDenoisingColored(img,None,10,10,7,21)
dst=cv2.fastNlMeansDenoising(img,None,10,7,21)
b,g,r = cv2.split(dst)         # get b,g,r
rgb_dst = cv2.merge([r,g,b])   # switch it to rgb
plt.subplot(211),plt.imshow(rgb_img)
plt.subplot(212),plt.imshow(rgb_dst)
plt.show()
```

## Screenshots:

1. Collected grayscale and resized images:



## 2. noise removal:



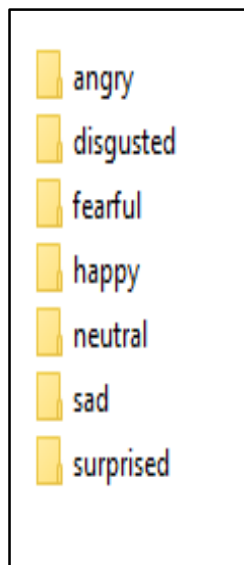
### 3.Data Classification/ Categorization:

After Data collection and pre-processing comes data classification/categorization. Data classification Data classification process comprises of two steps:

1. classifying and storing data into respective emotion folder:

40% of the data shortlisted after removing the unnecessary images from the dataset collected from kaggle were taken as reference to identify emotions of the images extracted from video. After the raw data was finalized, images were divided manually into 7 emotions as follows:

**Screenshot:**

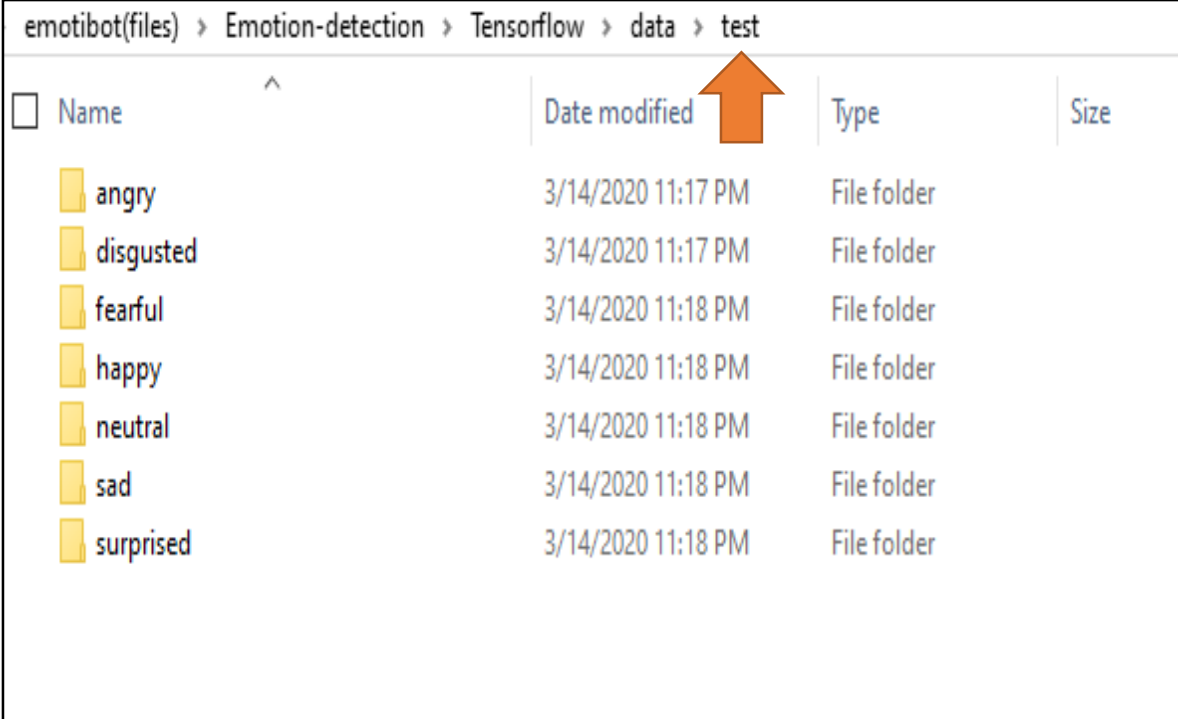









## 2. classifying or dividing data into train and test dataset:

After the classification of images into their respective emotion's folder, these folders were further divided into train and test folders to subdivide the images into 80-20 ratio for training and testing purpose.








### Screenshot:

#### 1.test dataset



emotibot(files) > Emotion-detection > Tensorflow > data > test				
<input type="checkbox"/> Name	Date modified	Type	Size	
 angry	3/14/2020 11:17 PM	File folder		
 disgusted	3/14/2020 11:17 PM	File folder		
 fearful	3/14/2020 11:18 PM	File folder		
 happy	3/14/2020 11:18 PM	File folder		
 neutral	3/14/2020 11:18 PM	File folder		
 sad	3/14/2020 11:18 PM	File folder		
 surprised	3/14/2020 11:18 PM	File folder		

## 2.train dataset

emotibot(files) > Emotion-detection > Tensorflow > data > train >			
<input type="checkbox"/> Name	Date modified	Type	Size
 angry	3/14/2020 11:18 PM	File folder	
 disgusted	3/14/2020 11:19 PM	File folder	
 fearful	3/14/2020 11:19 PM	File folder	
 happy	3/14/2020 11:19 PM	File folder	
 neutral	3/14/2020 11:20 PM	File folder	
 sad	3/14/2020 11:20 PM	File folder	
 surprised	3/14/2020 11:21 PM	File folder	



## 4.Database Modeling and connectivity:

### Purpose of database:

Database has been modeled to store the emotions and various details regarding emotions for further statistical analysis. A very basic database consisting of one table has been created to store the emotions predicted or identified by the application.

### Code snippet:

```
import time
import mysql.connector
#Create the connection object
myconn = mysql.connector.connect(host = "localhost",
                                user = "root", passwd = "", database = "EmotiBot")

#creating the cursor object

cur = myconn.cursor()
def insertdb(emotio, accu, timestmp):
    sql = "insert into logs(Emotion, Accuracy, TimeStamp)
           values (%s, %s, %s)"
    val = (emotio, accu, timestmp)
    try:
        #inserting the values into the table
        cur.execute(sql,val)
        #commit the transaction
        myconn.commit()
    except:
        myconn.rollback()
        print("failed")

#create a list
emotion_list = []
acc=[]
emotion_list.append(maxindex)
if (((emotion_list[len(emotion_list)-2]) != maxindex)
    or len(emotion_list)==1) and
    (((emotion_list[len(emotion_list)-2]) != maxindex) or
    len(emotion_list)==1):
    Insertdb(emotion_dict[maxindex], acc[maxindex], time.ctime())
```

**Screenshot:**

**Data Dictionary:**

6/25/2020

Print view - phpMyAdmin 4.8.4

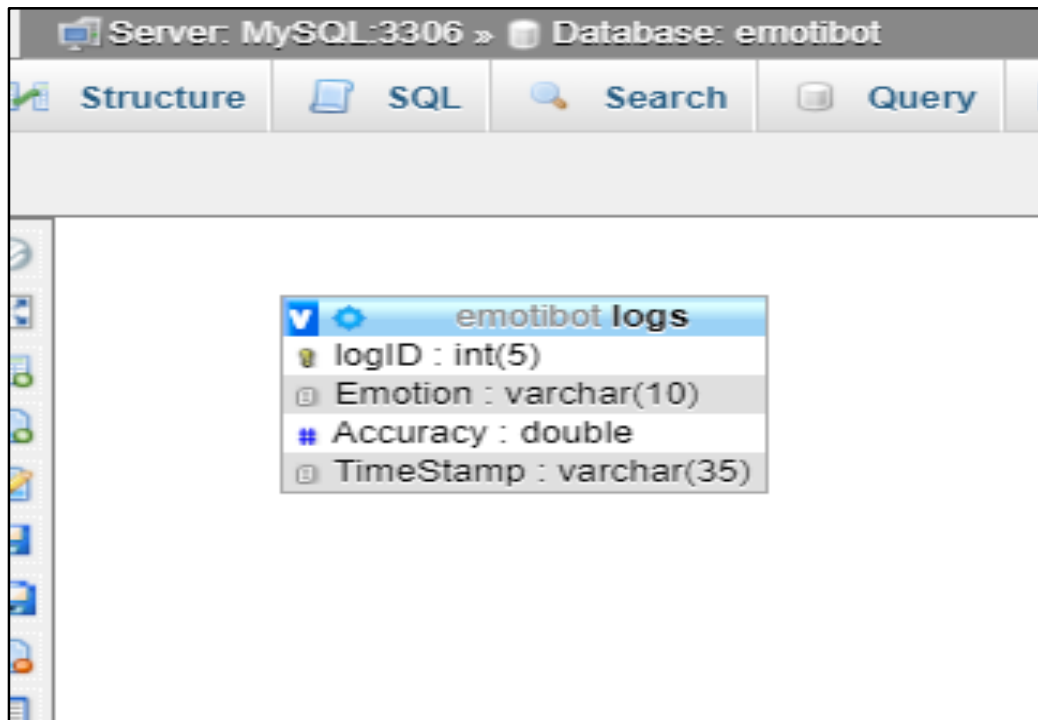
logs

Column	Type	Null	Default	Comments
logID ( <i>Primary</i> )	int(5)	No		
Emotion	varchar(10)	No		
Accuracy	double	No		
TimeStamp	varchar(35)	No		

Indexes

Keyname	Type	Unique	Packed	Column	Cardinality	Collation	Null	Comment
PRIMARY	BTREE	Yes	No	logID	788	A	No	

Designer view:



Department of Computer Application, The MS University of Baroda

36 | Page

## 5.Graphical User Interface:

Python provides various options for developing graphical user interfaces (GUIs).

Tkinter – Tkinter is the Python interface to the Tk GUI toolkit shipped with Python.

### Graphical User Interface using Tkinter:

Tkinter is the standard GUI library for Python. Python when combined with Tkinter provides a fast and easy way to create GUI applications. Tkinter provides a powerful object-oriented interface to the Tk GUI toolkit.

Creating a GUI application using Tkinter requires following steps –

- **Import the *Tkinter* module:**

Importing tkinter is same as importing any other module in the Python code.

```
Import tkinter
```

- **Create the GUI application main window:**

Tk(screenName=None, baseName=None, className='Tk', useTk=1): To create a main window, tkinter offers a method 'Tk(screenName=None, baseName=None, className='Tk', useTk=1)'. To change the name of the window, you can change the className to the desired one.

```
Main_window=tkinter.Tk()
```

- **Add one or more widgets to the GUI application and Geometry Management.**

Tkinter provides various controls, such as buttons, labels and text boxes used in a GUI application. These controls are commonly called widgets.

All Tkinter widgets have access to the specific geometry management methods, which have the purpose of organizing widgets throughout the parent widget area. Tkinter exposes the following geometry manager classes: pack, grid, and place.

```
w = widget_name ( master, option=value, ... )  
w.pack/grid/place()
```

- **Enter the main event loop to take action against each event triggered by the user.**

**mainloop():** There is a method known by the name `mainloop()` is used when application is ready to run. `mainloop()` is an infinite loop used to run the application, wait for an event to occur and process the event as long as the window is not closed.

`main_window.mainloop()`

#### code snippet:

```
from tkinter import *
from PIL import Image, ImageTk
import cv2
win = Tk()
win.title('EMOTIBOT')
win.config(background = "#D9D9D9")
win.resizable(width=FALSE, height=FALSE)

#frame1 for webcam
frame1 = Frame(win, width=600, bg="white",height=300, padx=10,
pady=10,highlightbackground="grey", highlightcolor="black",
highlightthickness=5)
frame1.pack(side=LEFT ,fill=Y,padx=10, pady=10)
cap = cv2.VideoCapture(0)
def show_frame():
    __, frame = cap.read()
    frame = cv2.flip(frame, 1)
    cv2image = cv2.cvtColor(frame, cv2.COLOR_BGR2RGBA)
    img = Image.fromarray(cv2image)
    imgtk = ImageTk.PhotoImage(image=img)
    lmain.imgtk = imgtk
    lmain.configure(image=imgtk)
    lmain.after(10, show_frame)

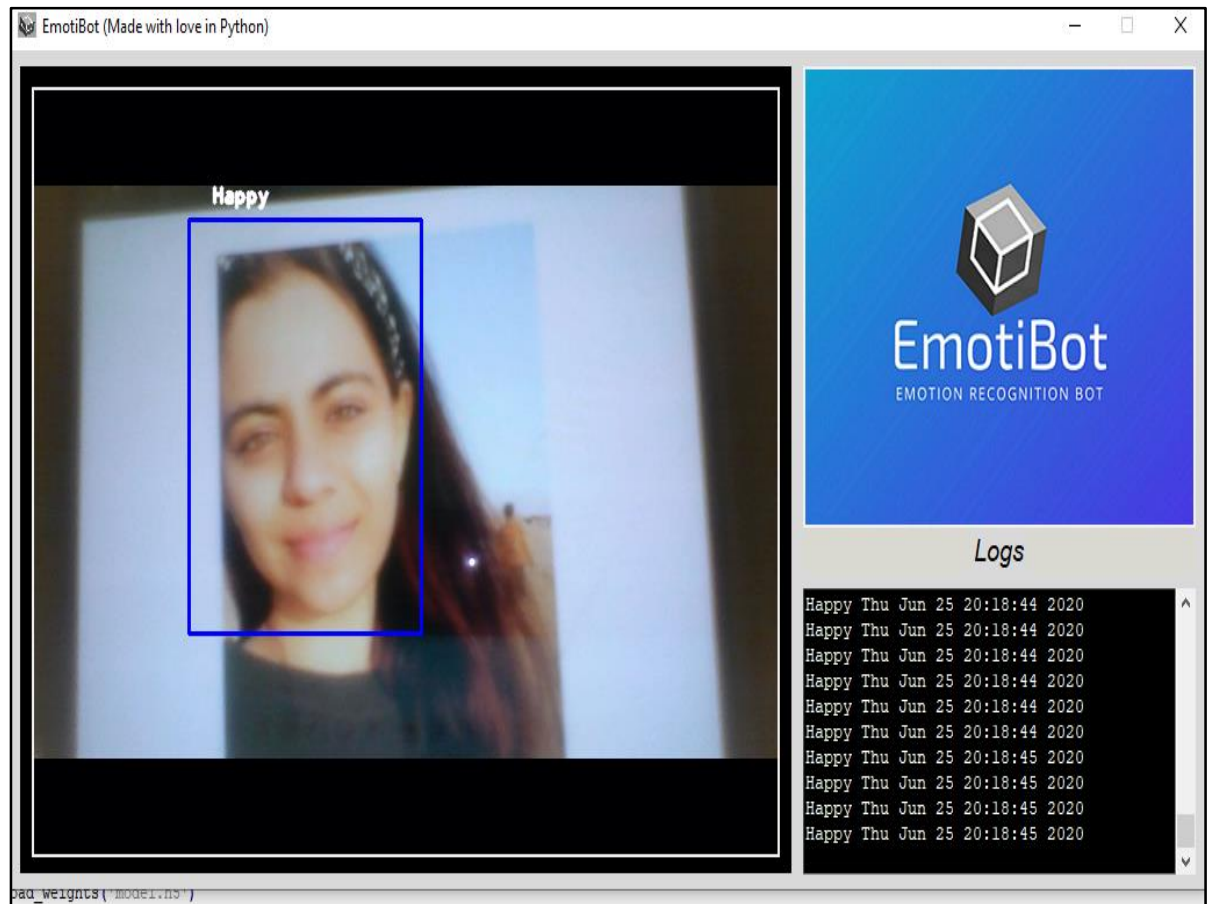
#Capture video frames
lmain = Label(frame1)
lmain.grid(row=0, column=0)

#frame2 for logo
frame2 = Frame(win, bg='#D9D9D9', width=300, height=5,padx=10,
pady=10,highlightbackground="grey", highlightcolor="black")
frame2.pack(side = TOP,fill=X,padx=10)
photo = PhotoImage(file='LogosDesign\\logom.png')
l = Label(frame2, image=photo,padx=10,pady=10)
l.pack()
l1 = Label(frame2, text="logs",font=("Helvetica", 16),bg="#D9D9D9",
fg="#0E7A3F")
l1.pack(side=BOTTOM,fill=Y,padx=2)
```

```
#frame3 for logs
frame3 = Frame(win, bg='black', width=300, height=150,padx=5,
pady=5,highlightbackground="#707173", highlightcolor="black")
frame3.pack(fill=X,expand=True,padx=5, pady=5)
frame3.pack_propagate(0) #stops frame from shrinking

#frame4 for buttons
frame4 = Frame(win, bg='grey', width=300, height=60,padx=5,
pady=5,highlightbackground="grey", highlightcolor="black")
frame4.pack(side=BOTTOM, fill=BOTH,expand=1,padx=5,pady=10)
b = Button(frame4, text='ON', padx = 40)
b.pack(side=LEFT,pady = 20, padx = 20)
c = Button(frame4, text='OFF', padx=40)
c.pack(side=RIGHT,pady = 20, padx = 20)
show_frame() #Display 2
win.mainloop()
```

screenshot:





## 6.Integration of Graphical User Interface with OpenCV:

To integrate GUI with openCV firstly required python packages were imported.Tkinter gives acces to GUI functionality along with both the Image and ImageTk classes from PIL/PILLOW to display the image in the GUI.We import CV2 to start the webcam and fit the webcam feed into the frame created by Tkinter.

### Code snippet:

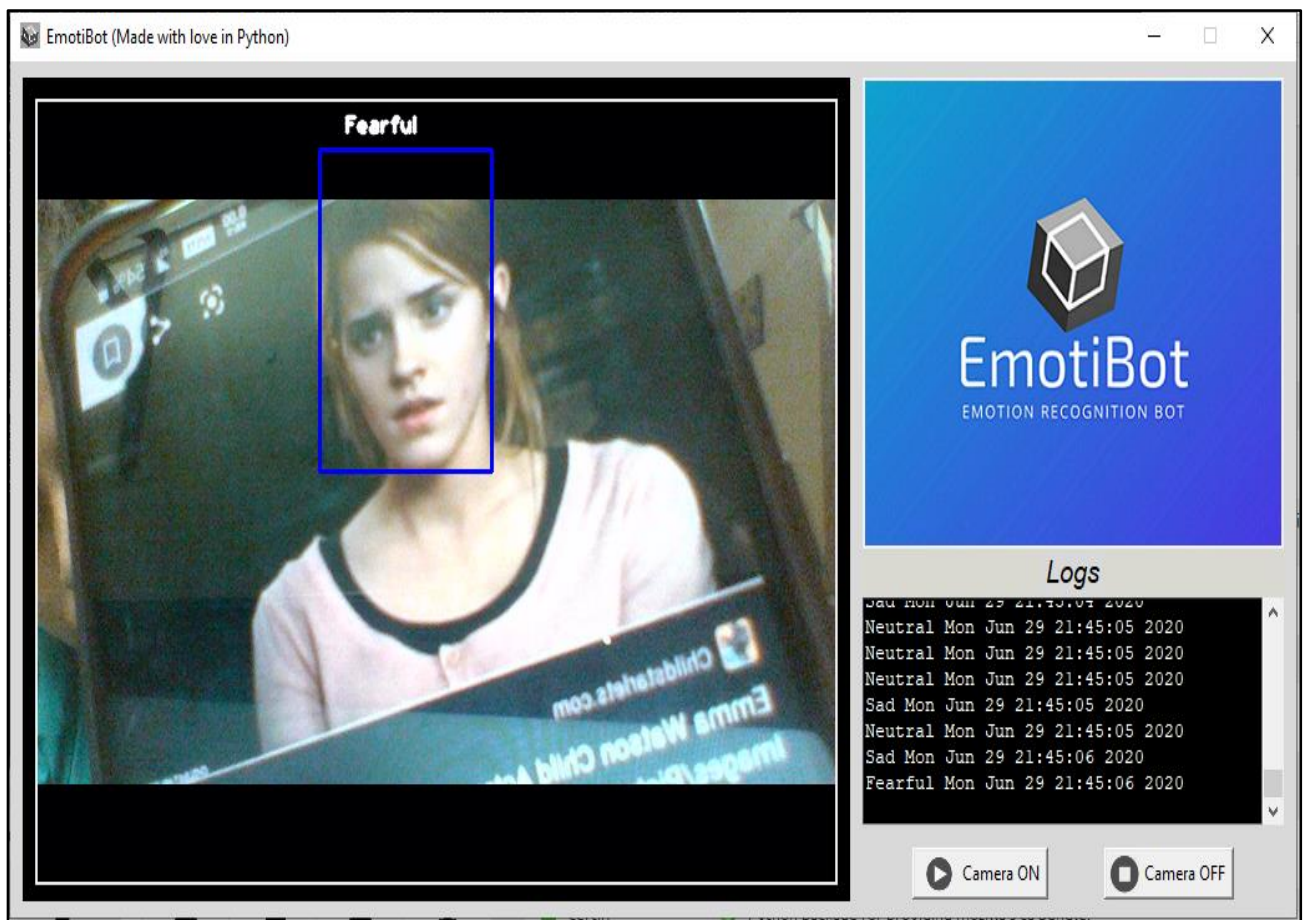
```
#import required modules
import cv2
from tkinter import Frame, Tk, PhotoImage, Button, Label, Scrollbar,
Text
from tkinter.constants import RIGHT, LEFT, TOP, BOTTOM, X, Y, BOTH,
NORMAL, END, DISABLED
from PIL import Image, ImageTk
import time
import mysql.connector

#integration
# start the webcam feed
cap = cv2.VideoCapture(0)
def show_frame():
    (_, frame) = cap.read()
    frame = cv2.flip(frame, 1)
facecasc =
cv2.CascadeClassifier('haarcascade_frontalface_default.xml')
    cv2image = cv2.cvtColor(frame, cv2.COLOR_BGR2RGBA)
    gray = cv2.cvtColor(frame, cv2.COLOR_BGR2GRAY)
    faces = facecasc.detectMultiScale
    (cv2image, scaleFactor = 1.3,minNeighbors=5)

    img = Image.fromarray(cv2.cvtColor(frame,
cv2.COLOR_BGR2RGBA))
    imgtk = ImageTk.PhotoImage(image=img)
    lmain.imgtk = imgtk
    lmain.configure(image=imgtk)
    lmain.after(10, show_frame)

# Capture video frames
lmain = Label(frame1)
lmain.grid(row=0, column=0, pady=(3, 0))
show_frame() # Display 2
win.mainloop()
```

screenshot:



## Assignment 2: Data Mining:

Data Mining is the process of examining pre-existing databases in order to generate new information. It is also known as Knowledge Discovery in Databases.

On the basis of a real estate dataset simple linear regression is created. Along with it scatter plot of regression is created. R-squared value mean, intercept and co-efficient are calculated and prediction for an apartment with size 750 sq. ft. is made.

Here the dependent variable is 'price', while the independent variable is 'size'.

### Screenshots:

```
Import the relevant libraries

In [1]: import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
sns.set()

from sklearn.linear_model import LinearRegression

Load the data

In [2]: data = pd.read_csv('real_estate_price_size.csv')
data.head()

Out[2]:
```

	price	size
0	234314.144	643.09
1	228581.528	656.22
2	281626.336	487.29
3	401255.608	1504.75
4	458674.256	1275.46

```
In [3]: data.describe()

Out[3]:
```

	price	size
count	100.000000	100.000000
mean	292289.470160	853.024200
std	77051.727525	297.941951
min	154282.128000	479.750000
25%	234280.148000	643.330000
50%	280590.716000	696.405000
75%	335723.696000	1029.322500
max	500681.128000	1842.510000

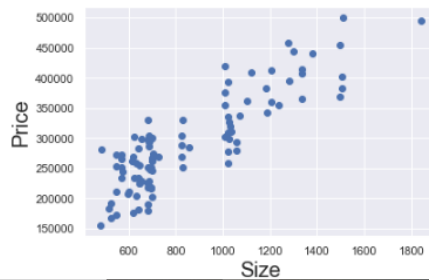
## Create the regression

### Declare the dependent and the independent variables

```
In [4]: x = data['size']  
y = data['price']
```

### Explore the data

```
In [5]: plt.scatter(x,y)  
plt.xlabel('Size',fontsize=20)  
plt.ylabel('Price',fontsize=20)  
plt.show()
```



### Transform the inputs into a matrix (2D object)

```
In [13]: x_matrix = x.values.reshape(-1,1)
```

### Regression itself

```
In [14]: reg = LinearRegression()  
reg.fit(x_matrix,y)
```

```
Out[14]: LinearRegression(copy_X=True, fit_intercept=True, n_jobs=None, normalize=False)
```

### Calculate the R-squared

```
In [15]: reg.score(x_matrix,y)
```

```
Out[15]: 0.7447391865847586
```

### Find the intercept

```
In [16]: reg.intercept_
```

```
Out[16]: 101912.60180122906
```

### Find the coefficients

```
In [13]: reg.coef_
```

```
Out[13]: array([223.17874259])
```

### Making predictions

You find an apartment online with a size of 750 sq.ft.

All else equal what should be its price according to the model?

```
In [14]: reg.predict(750)
```

```
Out[14]: array([269296.65874718])
```

## **Work Division (w.r.t time):**

### **January:**

1. Data Collection

### **February:**

1. Data Collection
2. Theory and study of research papers

### **March:**

1. Data Preprocessing
2. Data Classification/Categorization

### **April:**

1. Data classification/Categorization
2. Database Modelling
3. Database Connectivity
4. Data Mining using Real Estate Dataset

### **May:**

1. Graphical User Interface
2. Integration of GUI with OpenCV

## **Challenges or difficulties faced during assignment:**

I was a python beginner when I joined as an intern. I was aware about the advance concepts but being a beginner i didn't had any practical experience hence, implementing a whole system using the advance concepts was the most difficult part to achieve. Machine learning being the concept which is the closest to reality in reality is really difficult to understand and implement.

Collecting data using the techniques used my researchers and not using the available dataset fully as the dataset was outdated and not accurate. It would not have been possible without my collaboration partner. It took us almost a month to just gather the specified raw data.

Creating GUI and integrating it with OpenCV was challenging. Data preprocessing required a lot of research regarding different methods and to implement the best suited method. Data classification took a lot of time as it was very challenging to classify data into the respective emotion folder.

Due to the covid-19 pandemic, we weren't able to physically meet the guide or visit the company which caused a lot delay in work. Due to the availability of applications like Google Meet and Microsoft Teams we finally caught up and the work started to gain its speed.

## **Correlating with the Academics:**

During the entire academic program of BCA we were taught different courses like OOP, Data Structures, Computer Networks, software engineering which came in handy whenever required. Also the extra courses like management and communication skills were the great supporters during the entire internship Duration.

Our teachers gave us all the knowledge required and many practical demonstrations to deal with different situations which were the source of all the assignments I was able to complete during internship.

## **CORPORATE CULTURE**



## **Working Atmosphere:**

Rishabh Software is a good company with a bright future. It has an excellent platform to learn and grow in professional career. It has an open culture to work in a collaborative atmosphere. Colleagues are friendly, and management is cooperative. All the employees in the organisation are very helpful and friendly. All of them are experts in their field of work.

My guides at the company Dr.Falguni Ranadive and Mr.Gagan Dubey were very calm,poised and composed. They were very patient and eager to help anytime. Their dedication inspired me a lot to push myself to achieve the tasks.

## **Culture within the Organization:**

### **Cultural norms followed at Rishabh Software:**

Employees/Interns at Rishabh Software have to follow certain norms. Every employee receives a mail regarding the norms at Rishabh Software before joining so that they are well aware beforehand. Following details were described in the mail I received:

As an intern with RISE, you are now a representative of the company and hence what you display, physically or through your behavior and etiquettes matter a lot. We have listed some for your reference and action, however no charter/ rules /policy can be written to cover all situations. Each individual must exercise his or her high ethical standards.

### **Access control card system:**

1. The access control cards will be issued to every individual by the HR department after they sign the ID register maintained for this purpose.
2. The card once issued, binds the use for the particular individual and cannot be used or accessed by any other person than the one who owns the card.
3. It is a must to flash the card to the IN card reader during entry and to the OUT card reader during exit, even if the door is found open. PLEASE FLASH YOUR CARD AT TIME OF EACH ENTRY/ EXIT.
4. Do not Misuse the access authority given to you. What constitutes Misuse of your card?
  - Allowing any other person to gain entry or exit.
  - Being in possession of more than one Access card.
  - Allowing another person to make use of your card for any purpose.
5. Lost and stolen cards must be reported promptly to the HR department. Use of Company Resources:

1. Use of company resources, assets and consumable items, including time, materials, and equipment shall be only for business purposes and training benefit.
2. Any personal use, without proper permission, tantamount to misuse of Company property.
3. Use of training rooms, desk, projector or any other fixed asset is to be done with care.
4. Wi-Fi will be enabled on your personal systems to assist you with your training. Misusing of this facility is a misconduct.
5. Viewing or accessing unwanted sites on the internet is strictly prohibited. This may result in disciplinary action, up to and including termination.
6. Make sure to switch off lights, fans, Air-conditioners, computers, UPS when not in use/before leaving for the day

**Workplace Cleanliness:**

1. All employees are expected to keep their desks and surroundings neat and tidy.
2. They should keep office premises clean e.g. throwing waste in waste bins, wiping shoes on doormats, etc.
3. In cafeteria, only vegetarian food would be entertained.
4. When celebrating an event, do not dirty the floor or tables unnecessarily.
5. Consumption of chewing gums and tobacco is grossly discouraged.

**Mobile Use:**

1. The use of Mobile phones in office premises should be kept minimum.
2. Usage of mobile phones while in a session is strictly prohibited.

3. To attend to important/emergency calls, you need to come out of their work areas and take/place the call in the open areas.
4. Mobile phones are expected to be kept on vibration or silent mode by all employees. This is to ensure that we don't disturb others in the office premises.

**Physical Safety:**

1. You are expected to treat fellow employees, directors, customers and vendors with courtesy and resolve any difference in a professional, non-abusive and non-threatening manner.
2. Rishabh's policy is committed to provide to its employees a work environment free from sexual or racial harassment.
3. Rishabh does not permit employees, contractors or visitors to bring weapons onto any part of Rishabh's property, including parking areas.
4. Discussions pertaining to sensitive subjects like, religion, politics, current affairs etc. is best avoided. Dress Code:

**Workdays Dress Code:** Need to wear formal clothes throughout the week days as indicated Below: (Monday to Thursday)

- Men: Shirt, formal trousers, Shoes and socks.
- Women: Salwar-Kurta, Saree, Shirt, Pant, Formal Skirt, Shoes or Sandals. Flip-flops (shower type shoes) and thong-type shoes are not permitted.

**Weekend Dress Code:** Employees can wear casual clothes during the last working day of the week:

- Men: T-shirt, Jeans, and Sandals
- Women: T-shirt, Jeans, Skirt (no shorter than the knee length, sitting or standing), normal footwear. Understand what you wear.

## **Cultural norms I picked up:**

As an intern while following the norms I picked up quite a few of those:

- i. Being an Intern I always used to reach office on time
- ii. Amenities provided by the company were never misused by me to save myself from work.
- iii. I always made sure that I'm dressed according to the guidelines provided to me.

## **Insights gained about corporate culture that I will take into consideration when searching for full time employment:**

Internship led me to apply all the theoretical as well as the practical knowledge, which I had gained during my entire academic period but this time the application was real time.

My goal is to be a part of an organization where I can apply all the knowledge I have acquired for the betterment of the organization as well for my growth.

Some of the learnings to grow in corporate culture:

- i. Relate to people who have the capacity to mentor me and widen my scope of encouragement.
- ii. Continue to educate to update my knowledge and skill for my position, and position I desire to hold in future.
- iii. Manage my tasks and actions to contribute to the performance outcomes set for my department and company.
- iv. Build relationship outside the company that can be leveraged for its benefit.
- v. Exhibiting knowledge, I have been imparted and skills that make me a uniquely valuable employee.
- vi. Communicate and use all opportunities to strengthen my network throughout the organization.

## **COMMUNICATION**

## **Ways in which I communicated within and outside the organization:**

There was no restriction to interact to other employees I used to communicate verbally with each other.

Other modes of communication like video calls, phone calls and emails were used for communication outside the organization. Professionally emails are considered as the best mode of communication.

Due to the covid-19 pandemic, the whole communication cycle was dependent on emails and conference calls as employees/interns weren't allowed to visit the campus.

We created a WhatsApp group for discussions regarding the project.



## **Examples of good communication skills utilized by the employees:**

There are many examples of good communication skills utilized by the employees, and I would like to mention a few.

- i. **Listening:** The employees were active listeners. They used to involve in paying close attention to what the other person is saying, asking and clarifying questions, and rephrasing what the person says to ensure understanding.
- ii. **Respect:** Employees always used to respect each other and their ideas. Many of them used person's name, and make an eye contact while interacting. Many employees make the other person feel appreciated at the organisation.
- iii. **Confidence:** They were confident while interacting with one another.
- iv. **Nonverbal Communication:** They were also good at expressing through body language, eye contact and hand gestures.

## **Methods to improve the effectiveness of communication within the organization:**

Effective ways to learn precise, professional oral and written communication skills include:

- i. People in the organisation, such as a supervisor, should provide feedback on strengths and weaknesses as a communicator.
- ii. Alternatively, the employees should analyse the strengths and techniques of excellent communicators. They should even imitate strong communicators.
- iii. When sending a message, employees must think of the target audience, being sure to use terms and phrases that readers or listeners will understand.

## **CAREER REFLECTION**

## **Internship related to field of study:**

### **Internship Domain: Machine Learning & Data Mining**

Machine learning is a method of data analysis that automates analytical model building. It is a branch of artificial intelligence based on the idea that systems can learn from data, identify patterns and make decisions with minimal human intervention. Machine learning focuses on the development of computer programs that can access data and use it learn for themselves. data mining is a process used to extract usable data from a larger set of any raw data.

### **The technologies and subjects learned during the internship and domain in which I'm going to start my career:**

Some of the technologies which I learnt during the Internship:

- i. Advance Python and its packages and modules:
  - a. Numpy
  - b. Pandas
  - c. Matplotlib
  - d. Scikit Learn
  - e. Tkinter
  - f. OpenCV
- iii. Advance Statistics:
  - a. mean, median, mode,
  - b. frequency distribution
  - c. standard deviation
  - d. variance
  - e. skewness,
  - f. covariance and correlation

- v. TensorFlow 2
- vi. Data Science
- vii. Data Mining
- viii. Machine Learning

## **Conclusion: Experience and Learning**

Developing a whole system that is still the part of research work in this emerging era of machine learning from reading numerous research papers to planning and implementing was an experience in itself.

Creating Graphical User Interface using Tkinter module,Collecting data using various methods, preprocessing the data ,database modeling and atlast integrating all these pieces to form a system would not have been possible without my collaborative partner, mentors at the company and ofcourse the project guide.

Its easy to solve difficulties when you have your mentors to back you up, every time I was stuck in some problem my project partner, mentors at Rishabh software and the project guide were always there to help me out, they pushed me and encouraged me to solve the queries on my own which ended up being solved my me at last which made me feel really proud of myself.

I tried my hands on advance python while I was a beginner. Learned to create GUI in python, learned new technologies. This project really pushed me to go beyond my boundaries.

Last 5 months have been really daunting as well as a learning experience for me.I gained a lot of knowledge, learned to work in a professional environment in a team, lots of life lessons.

## References and Bibliography:

1. <https://www.udemy.com/course/the-data-science-course-complete-data-sciencebootcamp/> : Data Science Boot camp Course, where I learnt most of the concepts related to statistics, and basics of ML, DL, and Neural Networks
2. Tkinter
  - a. [https://www.tutorialspoint.com/python/python\\_gui\\_programming.htm](https://www.tutorialspoint.com/python/python_gui_programming.htm)
  - b. <https://www.geeksforgeeks.org/python-gui-tkinter/>
  - c. <https://www.javatpoint.com/python-tkinter>
  - d. <https://docs.python.org/3/library/tk.html>
3. OpenCV
  - a. <https://opencv.org/>
  - b. <https://docs.opencv.org/2.4/doc/tutorials/tutorials.html>
  - c. [https://opencv-python-tutroals.readthedocs.io/en/latest/py\\_tutorials/py\\_objdetect/py\\_face\\_detection/py\\_face\\_detection.html](https://opencv-python-tutroals.readthedocs.io/en/latest/py_tutorials/py_objdetect/py_face_detection/py_face_detection.html)
  - d. <https://www.ijsr.net/archive/v6i3/25031706.pdf>
  - e. [https://opencv-python-tutroals.readthedocs.io/en/latest/py\\_tutorials/py\\_objdetect/py\\_face\\_detection/py\\_face\\_detection.html](https://opencv-python-tutroals.readthedocs.io/en/latest/py_tutorials/py_objdetect/py_face_detection/py_face_detection.html)
3. Numpy
  - a. <https://numpy.org/doc/>
  - b. [https://numpy.org/devdocs/docs/howto\\_document.html](https://numpy.org/devdocs/docs/howto_document.html)
4. Research Papers
  - a. <https://www.sciencedirect.com/science/article/pii/S1877050917305264>
  - b. [https://ecs.utdallas.edu/research/researchlabs/msp-lab/publications/Busso\\_2004.pdf](https://ecs.utdallas.edu/research/researchlabs/msp-lab/publications/Busso_2004.pdf)
  - c. <https://www.ijrte.org/wp-content/uploads/papers/v8i4/D8275118419.pdf>
  - d. [https://www.researchgate.net/publication/328694744\\_Emotion\\_recognition\\_using\\_facial\\_expression\\_analysis](https://www.researchgate.net/publication/328694744_Emotion_recognition_using_facial_expression_analysis)
  - e. [https://scholarworks.sjsu.edu/cgi/viewcontent.cgi?article=1643&context=etd\\_projects](https://scholarworks.sjsu.edu/cgi/viewcontent.cgi?article=1643&context=etd_projects)
  - f. <https://ieeexplore.ieee.org/document/8398861>

- g. [https://curve.carleton.ca/system/files/etd/9d64f172-02e4-4172-890e-d751eee18de4/etd\\_pdf/e5ac8a1b25ad65e92ce18a84241280f8/fratesi-automatedrealtimeemotionrecognitionusingfacial.pdf](https://curve.carleton.ca/system/files/etd/9d64f172-02e4-4172-890e-d751eee18de4/etd_pdf/e5ac8a1b25ad65e92ce18a84241280f8/fratesi-automatedrealtimeemotionrecognitionusingfacial.pdf)
  - h. [http://cs231n.stanford.edu/reports/2016/pdfs/022\\_Report.pdf](http://cs231n.stanford.edu/reports/2016/pdfs/022_Report.pdf)
5. Database Connection
- a. <https://pynative.com/python-mysql-database-connection/>
  - b. [https://www.tutorialspoint.com/python/python\\_database\\_access.htm](https://www.tutorialspoint.com/python/python_database_access.htm)
  - c. <https://www.javatpoint.com/python-mysql-database-connection>
  - d. [https://www.w3schools.com/python/python\\_mysql\\_getstarted.asp](https://www.w3schools.com/python/python_mysql_getstarted.asp)
6. Pandas
- a. <https://matplotlib.org/>
7. Matplotlib
- a. <https://matplotlib.org/>