Project Report
On

Sentimental Analysis

Submitted during seventh semester in partial fulfillment of the requirements for the award degree of

Bachelor of Technology

in

Information Technology

by

Keshav Kumar

(19001011031)

Under supervision of

Dr. Komal Kumar Bhatia

Department of Computer Engineering

YMCA UNIVERSITY OF SCIENCE & TECHNOLOGY

FARIDABAD-121006

2022

ACKNOWLEDGEMENT

It is with deep sense of gratitude and reverence that I express my sincere thanks to my supervisor Dr. Komal Kumar Bhatia for her guidance, encouragement, help and useful suggestions throughout. Her untiring and painstaking efforts, methodical approach and individual help made it possible for me to complete this work in time. I consider myself very fortunate for having been associated with the scholar like her. Her affection, guidance and scientific approach served a veritable incentive for completion of this work.

I shall ever remain indebted to the faculty members of J. C. Bose University of Science and Technology, YMCA, Faridabad and all my classmates for their cooperation, kindness and general help extended to me during the completion of this work.

Although it is not possible to name individually, I cannot forget my well wishers at J. C. Bose University of Science and Technology, YMCA, Faridabad and outsiders for their persistent support and cooperation.

This acknowledgement will remain incomplete if I fail to express my deep sense of obligation to my parents and God for their consistent blessings and encouragement.

(Keshav Kumar)

CANDIDATE'S DECLARATION

I hereby certify that the work which is being carried out in this Minor Project titled **"Sentimental Analysis"** in fulfillments of the requirement for the degree of Bachelor of Technology in Computer Engineering and submitted to **"J. C. Bose University of Science and Technology, YMCA, Faridabad",** is an authentic record of my own work carried out under the supervision of **Dr. Komal Kumar Bhatia**

The work contained in this thesis has not been submitted to any other University or Institute for the award of any other degree or diploma by me.

Keshav Kumar
(2019-2023)

CERTIFICATE

This is to certify that the work carried out in this project titled **"Sentimental Analysis"** submitted by Rohit Kumar to **"J. C. Bose University of Science and Technology, YMCA, Faridabad"** for the award of the degree of Bachelor of Technology in Computer Engineering is a record of Bonafide work carried out by her under my supervision. In my opinion, the submitted report has reached the standards of fulfilling the requirements of the regulations to the degree.

Mrs. Poonam

(Mentor)

Dr. Lalit Mohan Goyal

(Supervisor)

TABLE OF CONTENTS

CHAPTER 1:
INTRODUCTION

1. INTRODUCTION

●Human emotion detection is implemented in many areas requiring additional security or information about the person. It can be seen as a second step to face detection where we may be required to set up a second layer of security, where along with the face, the emotion is also detected.

●In this we will add both face expression and audio for the detection of emotions for better results.

●Human emotions can be classified as: fear, contempt, disgust, anger, surprise, sad, happy, and neutral.

2. BASIC TERMINOLOGY

This section describes all the basic terminology used in the project report. These terms form the core of the project. Before moving forward to project details, these terms are to be well understood.

1. Data cleaning and preprocessing

Data cleaning is one of the important parts of machine learning. It plays a significant part in building a model. It surely isn't the fanciest part of machine

learning and at the same time, there aren't any hidden tricks or secrets to uncover. However, the success or failure of a project relies on proper data cleaning. Professional data scientists usually invest a very large portion of their time in this step because of the belief that "Better data beats fancier algorithms".

If we have a well-cleaned dataset, there are chances that we can get achieve good results with simple algorithms also, which can prove very beneficial at times especially in terms of computation when the dataset size is large.

Obviously, different types of data will require different types of cleaning. However, this systematic approach can always serve as a good starting point.

Pre-processing refers to the transformations applied to our data before feeding it to the algorithm. Data Preprocessing is a technique that is used to convert the raw data into a clean data set. In other words, whenever the data is gathered from different sources it is collected in raw format which is not feasible for the analysis.

Need of Data Preprocessing:

- For achieving better results from the applied model in Machine Learning projects the format of the data has to be in a proper manner. Some specified Machine Learning model needs information in a specified format, for example, Random Forest algorithm does not support null values, therefore to execute random forest algorithm null values have to be managed from the original raw data set.
- Another aspect is that the data set should be formatted in such a way that more than one Machine Learning and Deep Learning

algorithm are executed in one data set, and best out of them is chosen.

2. MACHINE LEARNING MODEL

A machine learning model is a file that has been trained to recognize certain types of patterns. You train a model over a set of data, providing it an algorithm that it can use to reason over and learn from those data.

Once you have trained the model, you can use it to reason over data that it hasn't seen before, and make predictions about those data. For example, let's say you want to build an application that can recognize a user's emotions based on their facial expressions. You can train a model by providing it with images of faces that are each tagged with a certain emotion, and then you can use that model in an application that can recognise any user's emotion.

3. CONVOLUTION NEURAL NETWORK

- Convolutional neural networks are multi-layer neural networks that are really good at getting the features out of data. They work well with images and they don't need a lot of pre-processing.

- Using convolutions and pooling to reduce an image to its basic features, you can identify images correctly.

- It's easier to train CNN models with fewer initial parameters than with other kinds of neural networks. You won't need a huge number of hidden layers because the convolutions will be able to handle a lot of the hidden layer discovery for you.

- One of the cool things about CNNs is the number of complex problems they can be applied to. From self-driving cars to detecting diabetes, CNNs can process this kind of data and provide accurate predictions.

4. OPENCV

OpenCV is the huge open-source library for the computer vision, machine learning, and image processing and now it plays a major role in real-time operation which is very important in today's systems. By using it, one can process images and videos to identify objects, faces, or even handwriting of a human. When it integrated with various libraries, such as NumPy, python is capable of processing the OpenCV array structure for analysis. To Identify image pattern and its various features we use vector space and perform mathematical operations on these features.

3. MOTIVATION & PROBLEM STATEMENT

- F.E.R. can help to prevent violence improves the overall security of a place.
- As a employee we think to provide the best service to customer. By the use of F.E.R. we can do that in a better way as the customer's emotions can be compared before and after going inside the center to determine how satisfied they are with the service they've received. And if there is a low score, the system can advise the employees to improve the service quality.
- F.E.R. can also help to improve the different sectors like:

- ➢ Audience engagement
- ➢ Video gaming test
- ➢ Differently abled children
- ➢ Health care

.

4. OBJECTIVE AND SCOPE

The sole intention behind the consideration of this Project is:

- The primary objective of our project is to detect emotions of any person during video call or simple interaction.
- Using the face fact that our facial features undergo significant changes with emotions.

CHAPTER 2:
MODULES

1. NUMPY

- NumPy is a Python library used for working with arrays.
- It also has functions for working in domain of linear algebra, fourier transform, and matrices.
- NumPy was created in 2005 by Travis Oliphant. It is an open source project and you can use it freely.
- NumPy stands for Numerical Python.

- NumPy is a Python library used for working with arrays.
- It also has functions for working in domain of linear algebra, fourier transform, and matrices.
- NumPy was created in 2005 by Travis Oliphant. It is an open source project and you can use it freely.
- NumPy stands for Numerical Python.

2. PANDAS

- Pandas is a Python library for data analysis. Started by Wes McKinney in 2008 out of a need for a powerful and flexible quantitative analysis tool, pandas has grown into one of the most popular Python libraries. It has an extremely active community of contributors.
- Pandas is built on top of two core Python libraries—matplotlib for data visualization and NumPy for mathematical operations. Pandas acts as a wrapper over these libraries, allowing you to access many of matplotlib's and NumPy's methods with less code. For instance, pandas' .plot() combines multiple matplotlib methods into a single method, enabling you to plot a chart in a few lines.

2.3 MATPLOTLIB

- Matplotlib is a low level graph plotting library in python that serves as a visualization utility.
- Matplotlib was created by John D. Hunter.
- Matplotlib is open source and we can use it freely.
- Matplotlib is mostly written in python, a few segments are written in C, Objective-C and Javascript for Platform compatibility.

2.4 SEABORN

- Seaborn is an amazing visualization library for statistical graphics plotting in Python. It provides beautiful default styles and color palettes to make statistical plots more attractive. It is built on the top of matplotlib library and also closely integrated to the data structures from pandas.
- Seaborn aims to make visualization the central part of exploring and understanding data. It provides dataset-oriented APIs, so that we can switch between different visual representations for same variables for better understanding of dataset.

2.5 Scikit-learn

scikit-learn is an open-source Python library that implements a range of machine learning, pre-processing, cross-validation, and visualization algorithms using a unified interface.

Important features of scikit-learn:

- Simple and efficient tools for data mining and data analysis. It features various classification, regression and clustering algorithms including support vector machines, random forests, gradient boosting, k-means, etc.

- Accessible to everybody and reusable in various contexts.
- Built on the top of NumPy, SciPy, and matplotlib.
- Open source, commercially usable – BSD license.
- In this article, we are going to see how we can easily build a machine learning model using scikit-learn.

2.6 TENSORFLOW KERAS

Keras is compact, easy to learn, high-level Python library run on top of TensorFlow framework. It is made with focus of understanding deep learning techniques, such as creating layers for neural networks maintaining the concepts of shapes and mathematical details. The creation of framework can be of the following two types −

- Sequential API
- Functional API

Consider the following eight steps to create deep learning model in Keras −

- Loading the data
- Preprocess the loaded data
- Definition of model
- Compiling the model
- Fit the specified model
- Evaluate it
- Make the required predictions
- Save the model

2.7 OPENCV

OpenCV is the huge open-source library for the computer vision, machine learning, and image processing and now it plays a major role in real-time operation which is very important in today's systems. By using it, one can process images and videos to identify objects, faces, or even handwriting of a human. When it integrated with various libraries, such as NumPy, python is capable of processing the OpenCV array structure for analysis. To Identify image pattern and its various features we use vector space and perform mathematical operations on these features.

CHAPTER 3:
DESCRIPTION

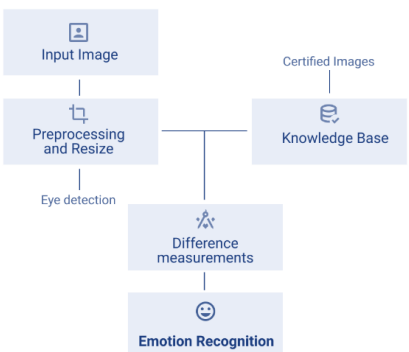
.

In this section, the overall working of the project has been described. How the project started and how the project works and how the various phases of project were carried
out.

1. BRIEF DESCRIPTION

Facial emotion detection technology is becoming more and more advanced every year. The AI used detects and studies the expressions depending on many factors to conclude what emotion the person is showing. Factors such as:

Location of the eyebrows and eyes
Position of the mouth
Distinct changes of the facial features
The study held in regarding emotion recognition summarized the system's algorithm as follows:



KNOWLEDGE BASE

This base contains images that are used for comparison and recognizing emotion variations. The images are stored in the database. Every time an input is given to the system, it finds a relevant image from its knowledge base by comparing the stored pictures and the input to come up with an output.

PREPROCESSING AND RESIZE

This step enhances the input and removes different types of noises. After that, the input image will be resized, typically with the use of the eye selection method.

DIFFERENCE MEASUREMENTS

During this step, the system will find any differences between the input image and the stored images and will finally lead to the emotion recognition step.

EMOTION RECOGNITION

This is the final step of the process. The comparison is made, and the final output is given depending on the differences found.

CHAPTER 4:
HARDWARE AND SOFTWARE REQUIREMENTS

Hardware and software requirements of any project are must to be satisfied, so that the virtual environment can be set up on any machine to run the project. So in this section the software and the hardware requirements are discussed completely.

1. HARDWARE REQUIREMENTS

The program can run on a machine with following configurations:

1) Windows 10 Operating system/ Linux/ Ubuntu
2) At least 4 GB RAM
3) At least 3 GB of disk space should be available
4) AMD Radeon Graphics Card (Recommended for faster processing)

4)2. SOFTWARE REQUIREMENTS

Following are the software requirements:

4)2.1. GOOGLE COLAB

Collab, or "Collaboratory", allows you to write and execute Python in your browser, with

- Zero configuration required
- Free access to GPUs

4)2.2. VS Code

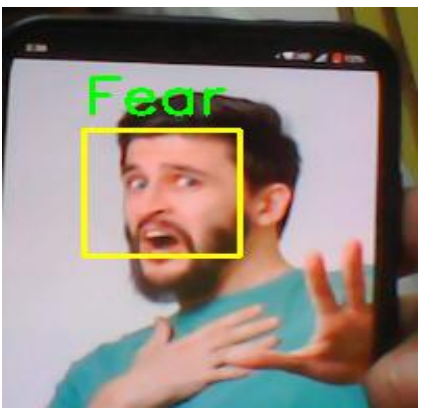
Visual Studio Code is a lightweight but powerful source code editor which runs on your desktop and is available for Windows, macOS and Linux. It comes with built-in support for JavaScript, TypeScript and Node.js and has a rich ecosystem of extensions for other languages and runtimes (such as C++, C#, Java, Python, PHP, Go, .NET).

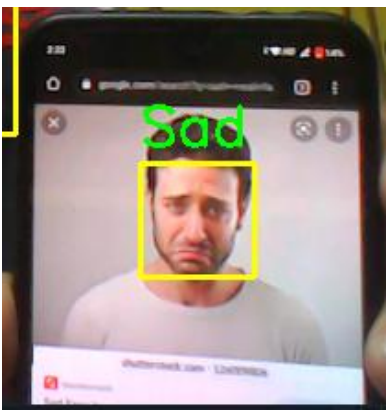
4)2.3. KAGGLE

Kaggle, a subsidiary of Google LLC, is an online community of data scientists and machine learning practitioners. Kaggle allows users to find and publish data sets, explore and build models in a web-based data-science environment, work with other data scientists and machine learning engineers, and enter competitions to solve data science challenges.

CHAPTER 5:
RESULTS

Below are the self-explanatory screenshots of my front end model with its work:





CHAPTER 6:
CONCLUSION

FUTURE ENHANCEMENT:

1. Improve the accuracy of prediction by increasing the dataset emotions like distgust, fear.

2. Intermix the result of facial and audio model which is already created for detect the with better accuracy.

AI Emotion Recognition technology is continuously being studied and improved to solve and provide solutions to arising risks and issues and to avoid any cultural and ethical problems in the society.

REFERENCES

https://youtube.com/playlist?list=PLfP3JxW-T70HvifebGl3d5d5jzI1Kp0i8

https://youtube.com/playlist?list=PLeo1K3hjS3uu7CxAacxVndI4bE_o3BDtO

https://youtu.be/WQeoO7MI0Bs

https://youtu.be/Bb4Wvl57LIk