VISVESVARAYATECHNOLOGICALUNIVERSITY

BELGAVI,KARNATAKA-590018

An Internship Report on

Criminal Face Recognition Using Python"

Submitted in partial fulfillment for the award of the degree of

Bachelor of Engineering in

Computer Science& Engineering

Submitted By

DRUPAD S **4MN20CS015**

Under the Guidance of

INTERNAL GUIDE

Prof. Suhasini

Assistant Professor

Dept. of CS&E

MIT Thandavapura

EXTERNAL GUIDE

Mr. Shaikh Faisal

Data Science Trainer

Inventeron

DEPARTMENT OF COMPUTER SCIENCE & ENGINEERING

MAHARAJAINSTITUTEOFTECHNOLOGYTHANDAVAPURA

(APPROVED BY THE AICTE & AFFILIATED TO VISVESVARAYA TECHNOLOGICAL UNIVERSITY,BELAGAVI)

(ACCREDITED BYISO9001:2015 & ISO21001:2018)

MYSURU-571302

2023-2024

DEPARTMENT OF COMPUTER SCIENCE & ENGINEERING

MAHARAJA INSTITUTE OF TECHNOLOGY THANDAVAPURA

CERTIFICATE

Certified that the internship work entitled "**CRIMINAL FACE RECOGNITAION USING PYTHON**" is a bonafide work carried out by **DRUPAD S (4MN20CS015)**, student of Maharaja Institute of Technology Thandavapura in Computer Science & Engineering as prescribed by Visvesvaraya Technological University, Belagavi during academic year 2023-2024. It is certified that all corrections/suggestions indicated have been incorporated inthe report. The report has been approved as it satisfies the course requirements in respect of Internship prescribed for the Bachelor of Engineering Degree.

……..	……..
Signature of Co-Ordinator	**Signature of HOD**
Prof. Suhasini	Dr. Ranjit K N
Assistant Professor	Assistant Professor and Head
Dept. of CS&E	Dept. of CS&E
MIT Thandavapura	MIT Thandavapura

External viva

Name of the Examiners	**Signature with date**

1)…….

2)…….

Reg. No : AAB9565

inventeron
nothing but smart..

INVENTERON TECHNOLOGIES AND BUSINESS SOLUTIONS LLP

CERTIFICATE OF INTERNSHIP

This is to certify that Mr/~~Ms.~~ Drupad S

has completed internship on Artificial Intelligence And Machine Learning

from 14-AUG-2023 to 14-SEP-2023 successfully.

We wish this intern all the best for future endeavours.

For Inventeron Technologies And Business Solutions LLP

Managing Director

Managing Director

Skill India
कौशल भारत-कुशल भारत

N·S·D·C National Skill Development Corporation
Transforming the skill landscape

IoT

Raspberry Pi

Java

AICTE

www.inventeron.com

ACKNOWLEDGEMENT

I am grateful to the **Maharaja Education Trust® Mysuru,** for providing an opportunity to get myBachelor's degree in this college.

I would like to extend my thanks to **Dr. Y T KRISHNE GOWDA,** Principal, Maharaja Institute of Technology Thandavapura, for his co-operation throughout academics, which have helped me in satisfactory completion of Internship.

I would like to extend my thanks to **Dr. H K Chethan,** Professor and Mentor, CS&E, MIT Thandavapura and **Prof.RANJIT K N,** Assistant Professor and HOD, Department of Computer Science & Engineering, for providing me invaluable support throughout the period of Internship.

I express my sincere thanks to **Mr.Shaikh Faisal, Inveteron Technologies Ptv Ltd**, BANGALORE for their constant support and appreciation during Internship Program.

I express my deepest sense of gratitude to my guide **Prof. Suhasini,** Assistant Professor, Department of Computer Science & Engineering, for her valuable guidance, suggestions and cheerful encouragement during the entire period of my internship program.

I also take an opportunity to thank all the teaching and non teaching staff members of CS&E department.

DRUPAD S

4MN20CS015

DECLARATION

I, **DRUPAD S** [4MN20CS105], student of 8th semester Computer Science and Engineering, Maharaja Institute of Technology Thandavapura, hereby declare that the internship entitled "**Criminal Face Recognition Basis using Python**" submitted to the Visvesvaraya Technological University, Belagavi during the academic year 2023-24, is a record of an original work done by me underthe guidance of my Internal guide Mrs. Suhasini, Assistant Professor, Department of Computer Science and Engineering, Maharaja Institute of Technology, Thandavapura and my external guide **Mr.Shaikh Faisal, Inveteron Technologies Ptv Ltd.**This internship dissertation report is submitted in partial fulfillment for the award of Bachelor's degree in Computer Science and Engineering. The results embodied in this report have not been submitted to any other University or Institute for the award of any degree.

ABSTRACT

Criminal Face Detection project aims to build a automated Criminal Face Detection system by levering the human ability to recall minute facial details. Identification of criminals at the scene of a crime can be achieved in many ways like fingerprinting, DNA matching or eye witness accounts. Out of these methods eye witness accounts are preferred because it stands scrutiny in court and it is a cost effective method. It is possible that witnesses to a crime have seen the criminal though in most cases it may not be possible to completely see the face of the perpetrator. The Criminal Face Detection System will be built of an existing criminal database. Input would be provided in the form of an images and matched again the existing databse and results would be provided. To identify any Criminal we need some identification regarding person, which are given by eyewitness. In most cases the quality and resolution of the recorded image segments is poor and hard to identify a face. To overcome this sort of problem we are developing software.

CONTENTS

SL.No	Index	Page No.
1	**COMPANY PROFILE**	**1-4**
1.1	About the Company	1
1.2	About the Founder	1
1.3	Services Offered by the Company	2
1.4	Working Process of the Company	2
1.5	Design Capabilities	3
1.6	People Working in the Organization	4
2	**TRAINING PROGRAMME**	**5-6**
2.1	Introduction	5
2.2	Week wise details	6
3	**LEARNING EXPERIENCE**	**7-17**
3.1	Knowledge Acquired	7
3.2	Skills Learned	16
3.3	Observed attitudes and gained values	17
3.4	The most challenging task performed	17
3.5	Project Introduction	18
3.5.2	Haar cascade classifier	18
3.5.3	Algorithm	19
4	**OUTPUT OF THE INTERNSHIP**	**19**
	CONCLUSION	21

LIST OF FIGURES

Figure No.	Figure Title	Page No.
Fig 4.1:	Setting up Pycharm Environment	19
Fig 4.2:	Recording face by running face_record.py	19
Fig 4.3:	Enocing each face by running encode_face.py	20
Fig 4.4:	Face recognized as criminal	20

Chapter 1

COMPANY PROFILE

1.1 About Company

Fig.1.1.1 Company logo

Inventeron was founded in 2013 by MISBAH MOHAMMED and KADGARJE MOHAMMED ATHA HUSSAIN.Inventeron Technologies and Business Solutions LLP, or ITABS, is an Indian based engineering and electronics company headquartered in Bangalore, Karnataka, India. It is both product and services oriented software company having its products in wireless communication technology and providing quality service to its valuable clients in its domain. Inventeron's core products are embedded components (including ICs, Control Boards, Controllers, Microprocessors, Fuel systems, Water Level Controllers, Security systems, Biometric Logic systems, Wireless Devices, etc.,). Industrial Products (Network Tower Management Systems) and Apps (E-commerce Apps, Website Design and Development).The founder is known for his teaching techniques which help the students in easy understandability of the most complicated concepts. With their teaching experience of 14 years and training thousands of students now, he has decided to take up the challenge to deliver the most in demand skill set currently in the industry, which is the future.

1.2 About the Founder

Founder MISBAH MOHAMMED and KADGARJE MOHAMMED ATHA HUSSAIN has started Inventron to give the opportunity to students that require upskilling and reskilling by also providing them an experience with a real working environment.

1.3 Services offered by the Company

1. It provides the following courses:
   - Full Stack Web Development.
   - Machine Learning using Numpy, pandas.
   - Embedded Systems.
2. They would work on 25 real time projects during the course period.
3. Course is exposed to In-house industry-based company projects.
4. Course is exposed to In-house industry-based company projects.

1.2 Working Process of the Company

They do Innovations in Embedded Electronics and software,Providing Quality service and varieties of Electronic Products.

The company has divided the work into different departments like:

1. Production: Production is the functional area responsible for turning inputs into finished outputs through a series of production processes. The Production Manager handles making sure that the materials needed are available at the time of developing the product. The Production manager must make sure the work is conducted smoothly and must supervise procedures for making work more efficient. A product is anything that can be offered to a market that might satisfy a want or need.
2. Marketing: These are the main section of the market departments:
3. Sales: This department manages the sales and distribution of the products tothe different regions.
4. Research & Department: It is responsible for market research and testing new products to make sure that they are suitable to be sold.
5. Promotion department: It decides on the type of promotion method for the products, arranges advertisements and the advertising media used.
6. Distribution department: It distributes the products across the industries.
7. Embedded System and Internet of Things (IOT) department.
8. Machine learning and web development department.

1.3 Design Capabilities

The main Designing procedures involved in the company are:

• Data Science

• Machine learning

• Artificial intelligence

Data Science:

Data science is an interdisciplinary field that uses scientific methods, processes, algorithms and systems to extract knowledge and insights from noisy, structured and unstructured data, and apply knowledge from data across a broad range of application domains. Data science is related to data mining, machine learning and big data.Data science is a "concept to unify statistics, data analysis, informatics, and their related methods" in order to "understand and analyse actual phenomena" with data. It uses techniques and theories drawn from many fields within the context of mathematics, statistics, computer science, information science, and domain knowledge. However, data science is different from computer science and information science. Turing Award winner Jim Gray imagined data science as a "fourth paradigm" of science (empirical, theoretical, computational, and now data-driven) and asserted that "everything about science is changing because of the impact of information technology" and the data deluge. A data scientist is someone who creates programming code and combines it with statistical knowledge to create insights from data.

Machine Learning:

Machine learning (ML) is the scientific study of algorithms and statistical models that computer systems use to perform a specific task without using explicit instructions, relying on patterns and inference instead. It is seen as a subset of artificial intelligence. Machine learning algorithms build a mathematical model base on sample data, known as "training data", to make predictions or decisions without being explicitly programmed to perform the task.Machine learning algorithms are used in a wide variety of applications, such as email filtering and computer vision, where it is difficult or infeasible to develop a conventional algorithm for effectively performing the task. Machine learning is closely related to computational statistics, which focuses on making predictions using computers.

The study of mathematical optimization delivers methods, theory and application domains to the field of machine learning. Data mining is a field of study within machine learning and focuses on exploratory data analysis through unsupervised learning. In its application across business problems, machine learning is also referred to as predictive analytics.

Artificial Intelligence:

Artificial intelligence sometimes called machine intelligence, is intelligence proved by machines, in contrast to the natural intelligence displayed by humans and animals. Leading AI textbooks define the field as the study of "intelligent agents": any device that perceives its environment and takes actions that maximize its chance of successfully achieving its goals. Colloquially, the term "artificial intelligence is often used to describe machines (or computers) that mimic "cognitive" functions that humans associate with the human mind, such as "learning" and "problem solving".As machines become increasingly capable, tasks considered to require "intelligence" areoften removed from the definition of AI, a phenomenon known as the AI effect. A quip in Tesler'sTheorem says, "AI is whatever hasn't been done yet." For instance, optical character recognitionis often excluded from things considered to be AI, having become a routine technology. Modern machine capabilities generally classified as AI include successfully understanding human speech, competing at the highest level in strategic game systems (such as chess and go),autonomously operating cars, intelligent routing in content delivery networks, and military simulations.

1.4 People Working in the Organization

Presently the company have around 20 employees including all the departments like Embedded, Java, Python,IOT and so on. The company is working with many Industrial projects in different domains and working for its own products. They deliver the excellence and consequently improve what is essential for human progress by mastering Science and Technology.

Chapter 2

TRAINING PROGRAM

2.1 Introduction:

An internship is a valuable opportunity for individuals to gain work experience within an organization for a predetermined period of time. Although it was traditionally reserved for medical graduates, internships are now available in a wide range of fields, including businesses, non-profit organizations, and government agencies. Typically, students and graduates participate in internships to develop relevant skills and gain valuable experience in their chosen field. Employers find great value in internships as they often identify and recruit their top interns as full-time employees, saving time and money in the long run. Third-party organizations are often responsible for recruiting interns on behalf of industry groups, but it's important to note that rules and regulations vary by country regarding intern classification as employees. Unfortunately, some unscrupulous employers may exploit the internship system..

Typically, an internship consists of an exchange of services for experience between the intern and the organization. Internships are used to determine if the intern still has an interest in that field after the real-life experience. In addition, an internship can be used to create a professional network that can assist with letters of recommendation or lead to future employment opportunities. The benefit of bringing an intern into full-time employment is that they are already familiar with the company, their position, and they typically need little to no training. Internships provide current college students the ability to participate in a field of their choice to receive hands on learning about a particular future career, preparing them for full-time workfollowing graduation.

Internships for professional careers are similar in some ways, but not as rigorous as apprenticeships for professions, trade, and vocational jobs. The lack of standardization and oversight leaves the term "internship" open to broad interpretation. Interns may be high school students, college and university students, or post-graduate adults. These positions may be paid or unpaid and are temporary.

2.2 Week wise details:

Week 1:

- Introduction to the Company
- Introduction to Python Working with:
    - ✓ Variables
    - ✓ Expressions
    - ✓ Strings
    - ✓ Lists
    - ✓ Dictionaries
    - ✓ Tuples
    - ✓ Conditional statements
    - ✓ Looping statements
    - ✓ Files

Week 2:

- Functions with python
- Object oriented concepts

Week 3:

- Supervised machine learning: Linear regression, logistic or sigmoid function, logistic regression, naive bayes, KNN, decision tree, random forest and various model evaluation methods.
- Unsupervised machine learning: Clustering & its use cases, k means clustering.
- Introduction to Machine Learning working with python packages
    - ✓ numpy
    - ✓ pandas
    - ✓ matplotlib
    - ✓ VaderSentiment

Week 4:

Project work-Criminal face recognition

Chapter 3

LEARNING EXPERIENCES

3.1 Knowledge Acquired

What is Python?

Python is a programming language created by Guido van Rossum in 1991..

It is used for:

- web development (server-side),
- software development,
- mathematics,
- system scripting.

What can Python do?

- Python can be used on a server to create web applications.
- Python can be used alongside software to create workflows.
- Python can connect to database systems. It can also read and modify files.
- Python can be used to handle big data and perform complex mathematics.
- Python can be used for rapid prototyping, or for production-ready software development.

Why Python?

- Python works on different platforms (Windows, Mac, Linux, Raspberry Pi, etc).
- Python has a simple syntax similar to the English language.
- Python has syntax that allows developers to write programs with fewer lines than someother programming languages.
- Python runs on an interpreter system, meaning that code can be executed as soon as itis written. This means that prototyping can be very quick.
- Python can be treated in a procedural way, an object-orientated way or a functionalway.

Python Syntax compared to other programming languages

- Python was designed for readability, and has some similarities to the English language with influence from mathematics.
- Python uses new lines to complete a command, as opposed to other programming languages which often use semicolons or parentheses.
- Python relies on indentation, using whitespace, to define scope; such as the scope of loops, functions and classes. Other programming languages often use curly-brackets for this purpose.

Python Indentation

Indentation refers to the spaces at the beginning of a code line. Where in other programming languages the indentation in code is for readability only, theindentation in Python is very important.Python uses indentation to indicate a block of code.

Comments

Python has commenting capability for the purpose of in-code documentation. Comments start with a #, and Python will make the rest of the line as a comment.

Creating Variables

Variables are containers for storing data values. Unlike other programming languages, Python has no command for declaring a variable.A variable is created the moment you first assign a value to it.

Variable Names

A variable can have a short name (like x and y) or a more descriptive name (age, carname, total_volume). Rules for Python variables:

- A variable name must start with a letter or the underscore character
- A variable name cannot start with a number
- A variable name can only contain alpha-numeric characters and underscores (A-z, 0-9, and _)
- Variable names are case-sensitive (age, Age and AGE are three different variables)

Built-in Data Types

Variables can store data of different types, and different types can do different things. Python has the following data types built-in by default, in these categories:

- Text Type: `str`
- Numeric Types: `int`, `float`, `complex`
- Sequence Types: `list`, `tuple`, `range`
- Mapping Type: `dict`
- Set Types: `set`, `frozenset`
- Boolean Type: `Bool`

Python Strings

String literals in python are surrounded by either single quotation marks, or double quotation marks. `'hello'` is the same as `"hello"`.

You can display a string literal with the `print()` function

List

A list is a collection which is ordered and changeable. In Python lists are written with square brackets.

- **Access Items**

You access the list items by referring to the index number

- **Negative Indexing**

Negative indexing means beginning from the end, `-1` refers to the last item, `-2` refers to the second last item etc.

- **Range of Indexes**

You can specify a range of indexes by specifying where to start and where to end the range. When specifying a range, the return value will be a new list with the specified items

Tuple

A tuple is a collection which is ordered and **unchangeable**. In Python tuples are written with round brackets.

- **Access Tuple Items**

You can access tuple items by referring to the index number, inside square brackets:

- **Negative Indexing**

Negative indexing means beginning from the end, `-1` refers to the last item, `-2` refers to the second last item etc.

- **Range of Indexes**

You can specify a range of indexes by specifying where to start and where to end the range. When specifying a range, the return value will be a new tuple with the specified items.

Dictionary

A dictionary is a collection which is unordered, changeable and indexed. In Python dictionaries are written with curly brackets, and they have keys and values.

- **Accessing Items**

You can access the items of a dictionary by referring to its key name, inside square brackets:

- **Change Values**

You can change the value of a specific item by referring to its key name.

- **Loop through a Dictionary**

You can loop through a dictionary by using a `for` loop. When looping through a dictionary, the return value are the *keys* of the dictionary, but there aremethods to return the *values* as well.

- **Adding Items**

Adding an item to the dictionary is done by using a new index key and assigning a value to it:

- **Removing Items**

There are several methods to remove items from a dictionary

Python Conditions and If statements

Python supports the usual logical conditions from mathematics:

- Equals: `a == b`
- Not Equals: `a != b`
- Less than: `a < b`
- Less than or equal to: `a <= b`
- Greater than: `a > b`
- Greater than or equal to: `a >= b`

These conditions can be used in several ways, most commonly in "if statements" and loops.
An "if statement" is written by using the `if` keyword.

- **Indentation**

Python relies on indentation (whitespace at the beginning of a line) to define scope in the code. Other programming languages often use curly-brackets for this purpose.

- **Elif**

The `elif` keyword is python way of saying "if the previous conditions were not true, then try this condition".

- **Else**

The `else` keyword catches anything which isn't caught by the preceding conditions.You can also have an `else` without the `elif`:

- **And**

The `and` keyword is a logical operator, and is used to combine conditional statements:

- **Or**

The `or` keyword is a logical operator, and is used to combine conditional statements:

- **Nested If**

You can have `if` statements inside `if` statements, this is called *nested* `if` statements.

- **The pass Statement**

`if` statements cannot be empty, but if you for some reason have an `if` statement with no content, put in the `pass` statement to avoid getting an error.

Python Loops

Python has two primitive loop commands:

- while loops
- for loops

The while Loop

With the while loop we can execute a set of statements as long as a condition is true.

The while loop requires relevant variables to be ready, in this example we need to define an indexing variable, i, which we set to 1.

- **The break Statement**

With the break statement we can stop the loop even if the while condition is true:

- **The continue Statement**

With the continue statement we can stop the current iteration, and continue with the next:

- **The else Statement**

With the else statement we can run a block of code once when the condition no longer is true

Python For Loops

A for loop is used for iterating over a sequence (that is either a list, a tuple, a dictionary, a set, or a string).

This is less like the for keyword in other programming languages, and works more like an iterator method as found in other object-orientated programming languages.

With the for loop we can execute a set of statements, once for each item in a list, tuple, set etc.

The for loop does not require an indexing variable to set beforehand.

- Looping Through a String

Even strings are iterable objects, they contain a sequence of characters:

- The break Statement

With the break statement we can stop the loop before it has looped through all the items:

- The continue Statement

With the continue statement we can stop the current iteration of the loop, and continue with the next:

- The range() Function

To loop through a set of code a specified number of times, we can use the range() function,

The range() function returns a sequence of numbers, starting from 0 by default, and increments by 1 (by default), and ends at a specified number

Python Functions

A function is a block of code which only runs when it is called.You can pass data, known as parameters, into a function. A function can return data as a result.

- Creating a Function

In Python a function is defined using the def keyword:

- Calling a Function

To call a function, use the function name followed by parenthesis:

- Arguments

Information can be passed into functions as arguments. Arguments are specified after the function name, inside the parentheses. You can add as manyarguments as you want, just separate them with a comma.

The following example has a function with one argument (fname). When the function is called, we pass along a first name, which is used inside the function to print the full name

Python Classes/Objects

Python is an object oriented programming language. Almost everything in Python is an object, with its properties and methods.A Class is like an object constructor, or a "blueprint" for creating objects.

- Create a Class

To create a class, use the keyword class:

- Create Object

Now we can use the class named MyClass to create objects:

- The__init_() Function

The examples above are classes and objects in their simplest form, and are not really useful in

real life applications.

To understand the meaning of classes we have to understand the built-in __init_() function.

All classes have a function called___init_(), which is always executed when the class is being initiated.

Use the___init_() function to assign values to object properties, or other operations that are necessary to do when the object is being created:

- Object Methods

Objects can also contain methods. Methods in objects are functions that belong to the object.

Let us create a method in the Person class:

- The self Parameter

The self parameter is a reference to the current instance of the class, and is used to access variables that belongs to the class.

It does not have to be named self , you can call it whatever you like, but it has to be the first parameter of any function in the class

File Handling in Python

Python too supports file handling and allows users to handle files i.e., to read and write files, along with many other file handling options, to operate on files. The concept of file handling has stretched over various other languages, but the implementation is either complicated or lengthy, but alike other concepts of Python, this concept here is also easy and short. Python treats file differently as text or binary and this is important. Each line of code includes a sequence of characters and they form text file. Each line of a file is terminated with a special character, called the EOL or End of Line characters like comma {,} or newline character. It ends the current line and tells the interpreter a new one has begun. Let's start with Reading and Writing files**.**

- **Working of open() function**

We use open () function in Python to open a file in read or write mode. As explained above, open ( ) will return a file object. To return a file object we use open() function along with two arguments, that accepts file name and the mode, whether to read or write. So, the syntax being: open(filename, mode)**.** There are three kinds of mode, that Python provides and how files can be opened:

- **"** r **",** for reading.
- **"** w **",** for writing.
- " **a** ", for appending.
- " **r+** ", for both reading and writing

One must keep in mind that the mode argument is not mandatory. If not passed, then Python will assume it to be " **r** " by default. Let's look at this program and try to analyze how the read mode works:

The open command will open the file in the read mode and the for loop will print each line present in the file.

- **Working of read() mode**

There is more than one way to read a file in Python. If you need to extract a string that contains all characters in the file then we can use file.read()**.** The full code would work like this: Another way to read a file is to call a certain number of characters like in the following code the interpreter will read the first five characters of stored data and return it as a string:

- **Creating a file using write() mode**

To manipulate the file, write the following in your Python environment:

The close() command terminates all the resources in use and frees the system of this particular program.

- **Working of append() mode**

It is designed to provide much cleaner syntax and exceptions handling when you are working with code. That explains why it's good practice to use them with a statement where applicable. This is helpful because using this method any files opened will be closed automatically after one is done, so auto-cleanup.

Using write along with with() function We can also use write function along with with() function:

- **split() using file handling**

We can also split lines using file handling in Python. This splits the variable when space is encountered. You can also split using any characters as we wish. Here is the code.

There are also various other functions that help to manipulate the files and its contents. One can explore various other functions in Python Docs.

3.2 Skills Learned

- **Machine learning**:

Machine learning (ML) is a scientific study that involves algorithms and statistical models used by computer systems to perform tasks without explicit instructions. Instead, it relies on patterns and inference. It is a subset of artificial intelligence. The machine learning algorithms use sample data, called "training data," to build a mathematical model that enables them to make predictions or decisions without being explicitly programmed to perform the task. Machine learning algorithms are used in various applications, such as email filtering and computer vision, where developing a conventional algorithm for effectively performing the task is difficult or not feasible.Machine learning is closely related to computational statistics, which focuses on making predictions using computers. The study of mathematical optimization delivers methods, theory and application domains to the field of machine learning. Data mining is a field of study withinmachine learning, and focuses on exploratory data analysis through unsupervised learning.[3][4] In its application across business problems, machine learning is also referred to as predictive analytics.

- **TASKS**

1. Data acquisition and cleaning
2. Data visualization
3. Data modelling
4. Testing
5. Comparison and measurement

1. **DATA ACQUISITION AND CLEANING:** Buying and merging the data from all the proper sources. Data Exploration and Pre-processing: Cleaning and preprocessing the data to create homogeneity, performing exploratory data analysis and statistical analysis to understand the relationships between the variables

2. **DATA VISUALIZATION** Data visualization is the graphical representation of information and data. By using visual elements like charts, graphs, and maps, data visualization tools supply an accessible way to see and understand trends, outliers.

3. **DATA MODELLING** Data modeling is the process of producing a descriptive diagram of relationships between several types of information that are to be stored in a database. One of the goals of data modeling is to create the most efficient method of storing information while still supporting complete access and reporting.

4. **COMPARISON AND MEASUREMENT** Comparison of data points is probably the most common and easy-to-understand method for data analysis. As the name suggests, we use comparison to evaluate and compare values between two or more data points. With comparison you can also easily find the lowest and highest values in the chart.

5. **TESTING** We are keeping 20% of our dataset to treat it as unseen data and be able and evaluate the performance of our models. We are splitting our dataset in a way such that all the news is represented proportionally equally in both training and testing dataset. Other than that, the choiceis being done randomly with uniform distribution.

3.3 Observed attitudes and gained values:

They are enough enthusiastic and dedicated towards their work and they are ready towork anytime regardless of time. They were professionals and training us in a professional manner. The art of working with dedication and under pressure are some of the values gained from them.

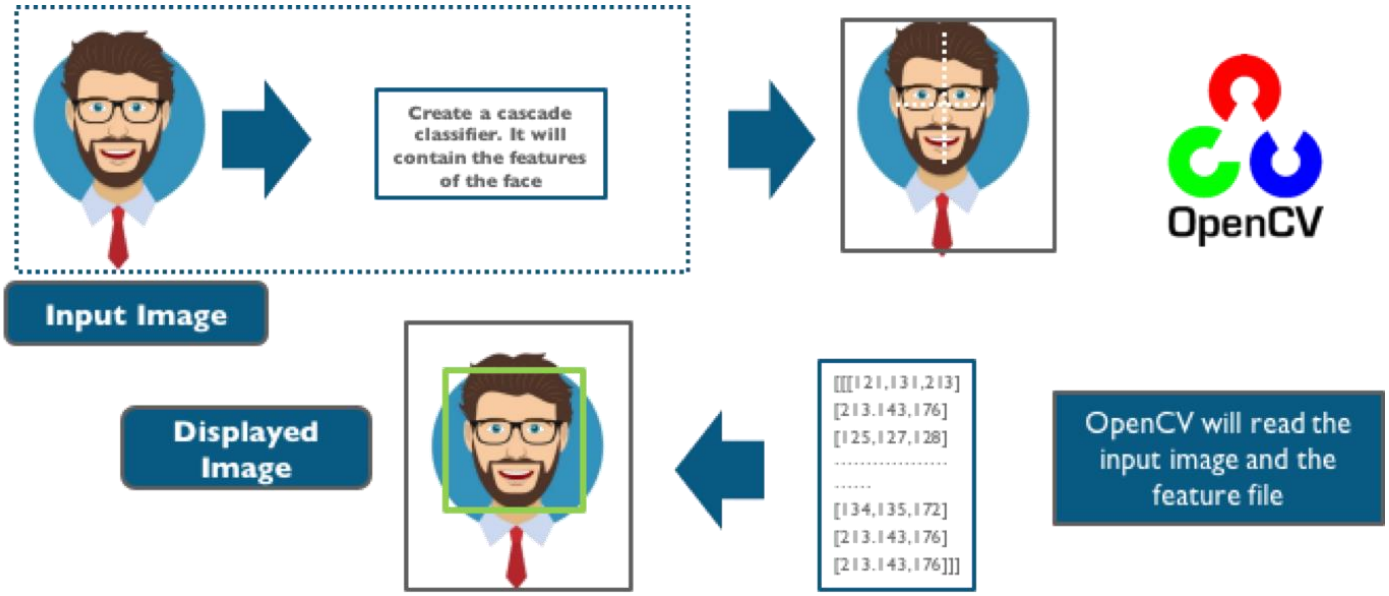
3.4 The most challenging task performed

Understanding the mind-set and working methodology of the corporate working people and start working with them.

3.5 Project Introduction

Face detection and recognition have become essential tools in criminal identification and investigation. Face detection involves identifying individuals by comparing their facial features to a database of known faces. In criminal identification, face detection and recognition are used to identify suspects and match them to surveillance footage or images from crime scenes. It has the potential to significantly improve the accuracy and speed of criminal investigations and to increase public safety by enabling law enforcement to identify and apprehend dangerous criminals more quickly. However, the use of face detection and recognition technology also raises ethical and privacy concerns, as it can potentially be used to track individuals without their consent or to discriminate against certain groups. It is therefore important to use these tools responsibly and transparently and to ensure that appropriate safeguards are in place to protect individual rights and freedoms.

3.5.2 Haar cascade classifier



Haar cascade classifiers are a machine-based learning method that use Haar features to identify objects in images. Viola and Jones first introduced Haar cascades in 2001 as a method for identifying artifacts and real-time face detection .Haar cascade classifiers work by learning a cascade function from both positive and negative photos, and then using that function to classify items in other pictures. The algorithm uses Haar-like features to traverse the image in window-sized sections to compute and match features.

3.5.3 Algorithm

1. Importing Libraries (Numpy, pandas etc)

2. Importing Dataset (pd.read_csv, etc)

3. Data processing (Using Numpy & Pandas)

4. Data visualisation (Matplotlib, seaborn etc)

5. Importing Machine learning algorithms (From sklearn, keras, tensorflow etc)

6. Training data into algorithm (Train test split & use Train data)

7. Testing algorithm for accuracy (Use Test data)

Chapter 4

OUTPUT OF THE INTERNSHIP

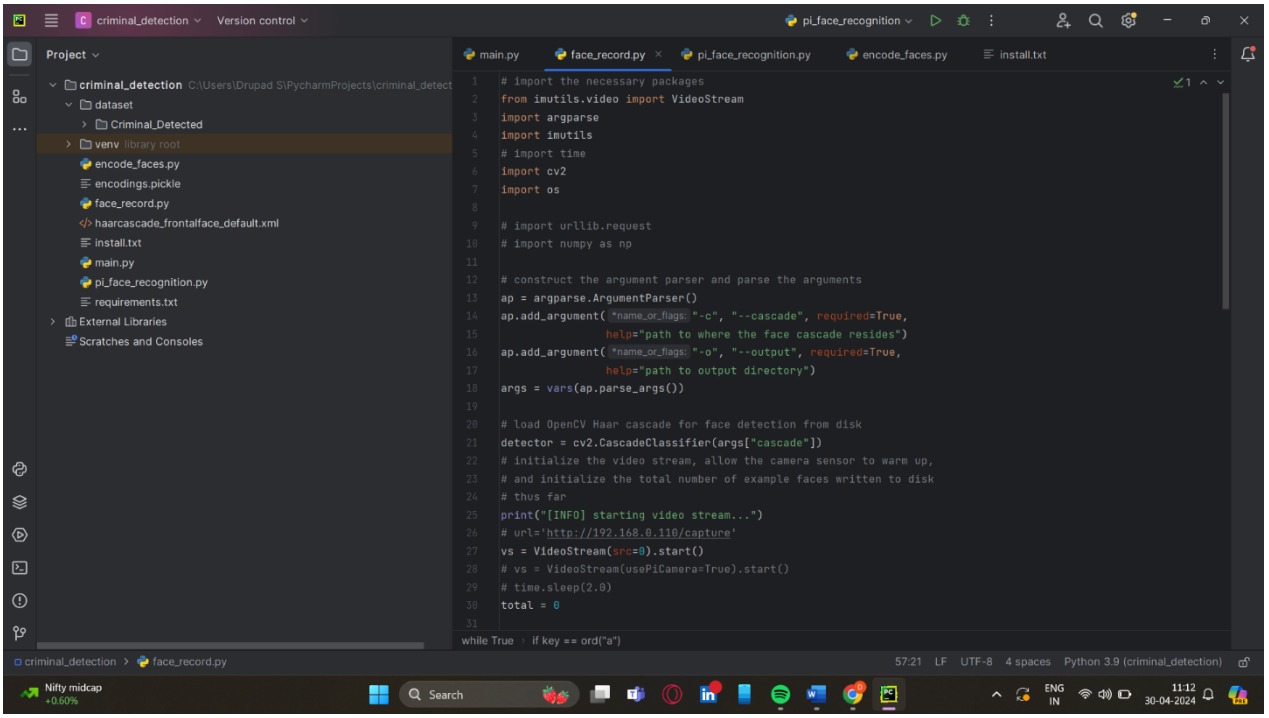


Fig 4.1: Setting up Pycharm Environment

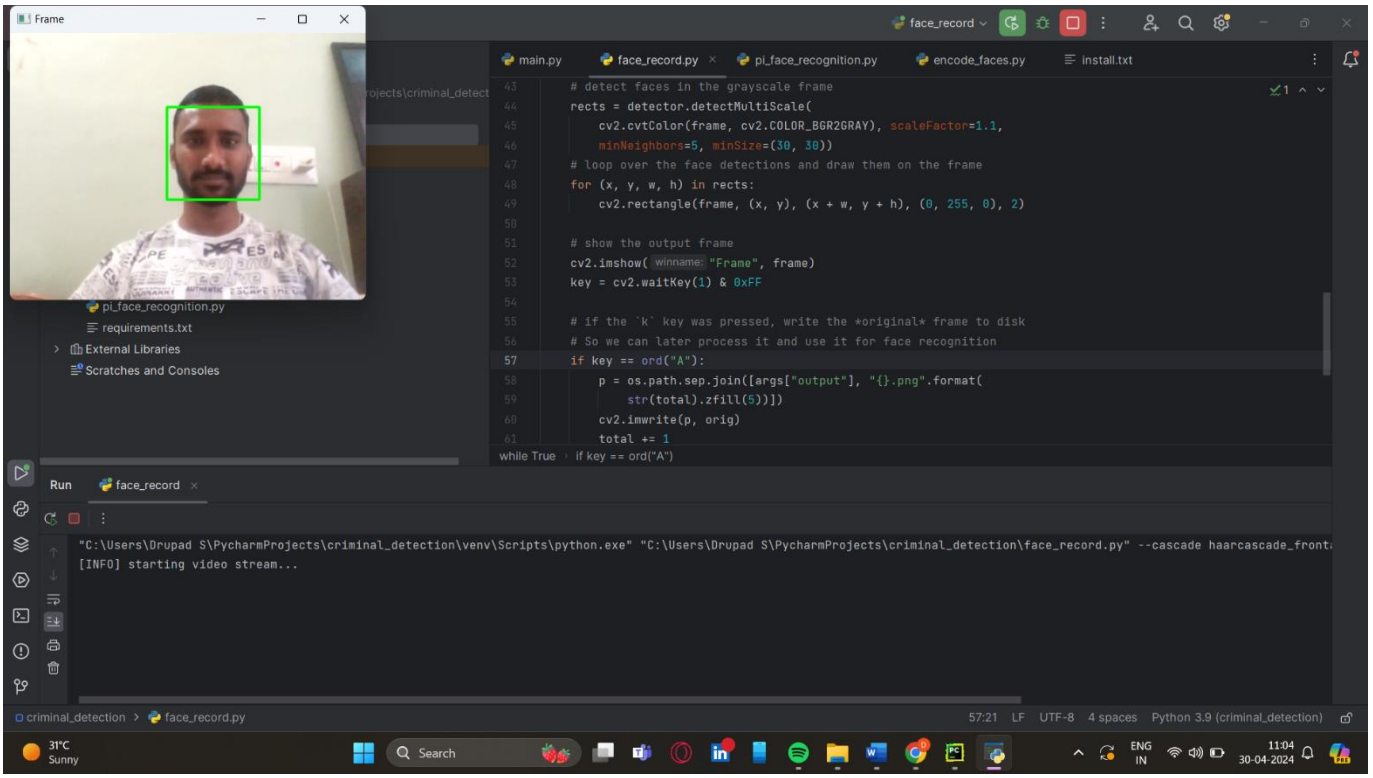


Fig4.2: Recording face by running face_record.py

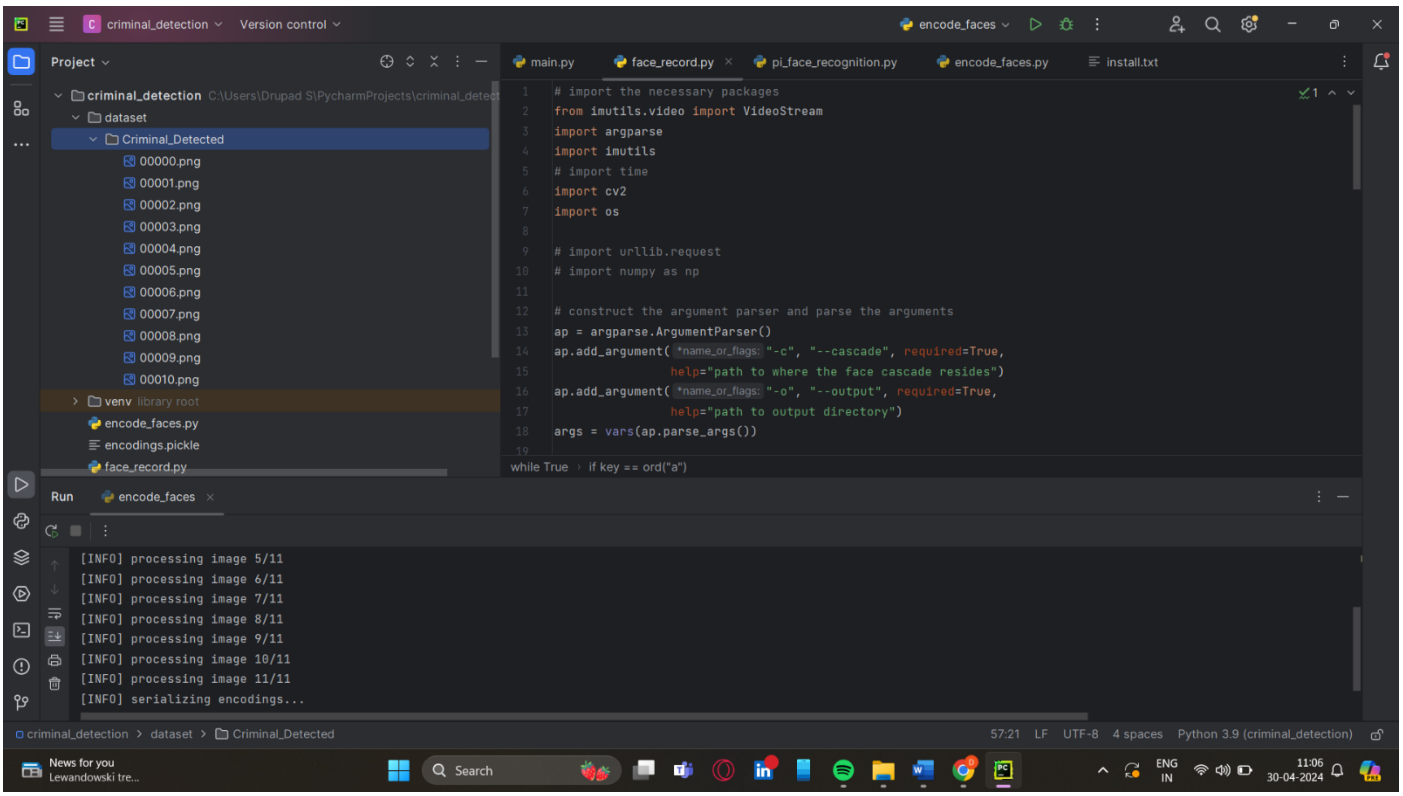


Fig 4.3: Enocing each face by running encode_face.py

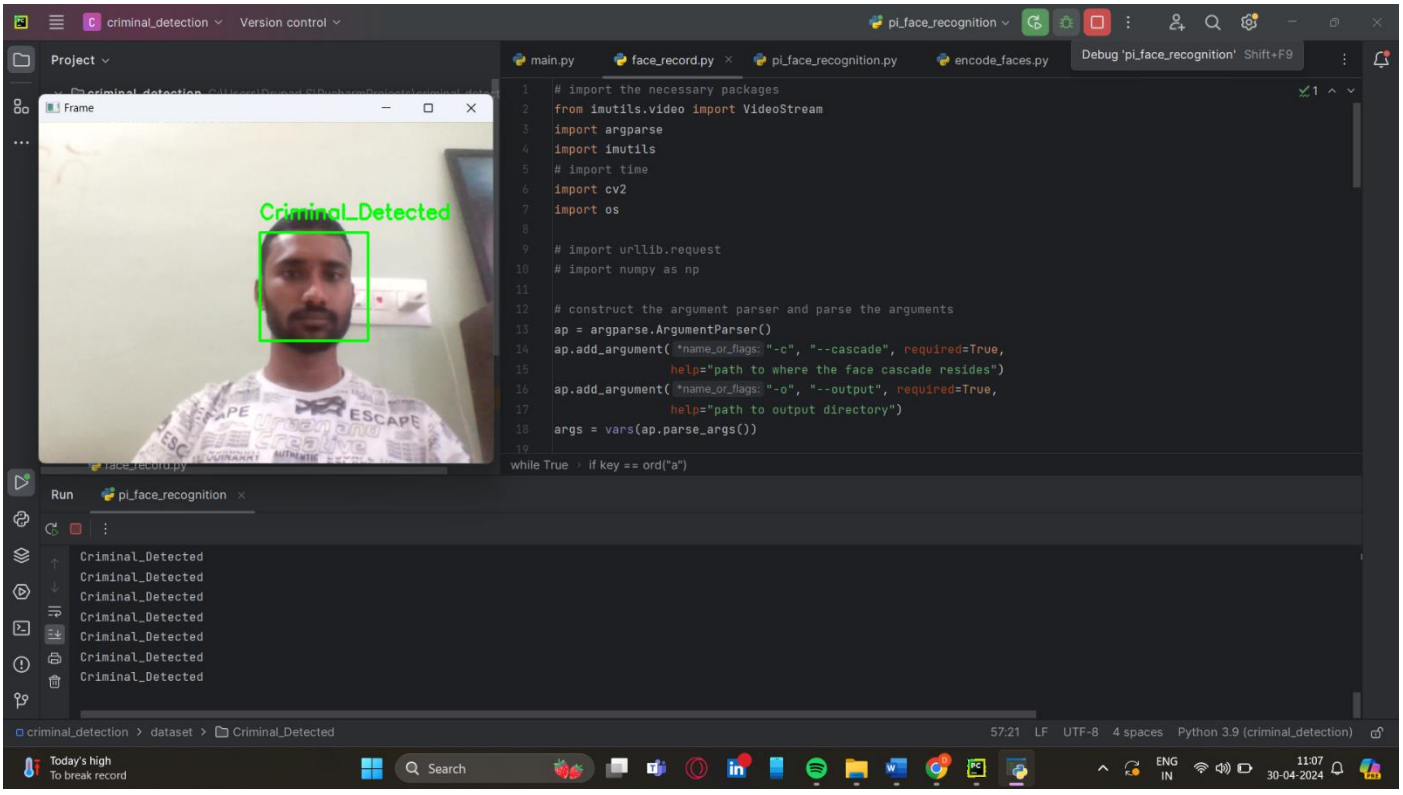


Fig 4.4: Face recognized as criminal

CONCLUSION

I recently completed a month-long internship to gain a deeper understanding of the industry and the mindset of corporate professionals. This experience proved to be both fulfilling and informative. Throughout the internship, I had the opportunity to meet and network with numerous individuals who I am confident will help me in the future. Additionally, I was able to develop important skills such as verbal and non-verbal communication, problem-solving, time management, observation, and self-motivation. The senior staff in the office provided me with valuable encouragement and motivation. I also gained knowledge of machine learning algorithms, including supervised and unsupervised learning, and their strengths and weaknesses. Overall, I have a solid understanding of the fundamental issues related to this field.