Personality Prediction Through CV

A COURSE PROJECT REPORT
[18CSC305J- Artificial Intelligence]

Submitted by

P VIKRAMA RAYALU[RA2011030010205]
VIGHNESAH PRADHAN[RA2011030010180]
MOHAMMED AIMEN KHAN[RA2011030010210]

Under the Guidance of

Dr.K.A VARUN KUMAR

(Associate Professor, Department of Networking and Communications)

in partial satisfaction of the requirements for the degree of

BACHELOR OF TECHNOLOGY
in
COMPUTER SCIENCE ENGINEERING
with specialization in Cyber Security

DEPARTMENT OF NETWORKING AND COMMUNICATIONS
COLLEGE OF ENGINEERING AND TECHNOLOGY SRM
INSTITUTE OF SCIENCE AND TECHNOLOGY
KATTANKULATHUR- 603 203
MAY 2023

BONAFIED CERTIFICATE

Certified that this project report **"Personality Prediction Through CV"** is the bonafide work of **"P VIKRAMA RAYALU (RA2011030010205), VIGHNESH PRADHAN (RA2011030010180) and MOHAMMED AIMAN KHAN (RA2011030010210)"** of III Year/VI Sem B.Tech(CSE) who carried out the mini project work under my supervision for the course **18CSC305J- Artificial Intelligence** in SRM Institute of Science and Technology during the academic year 2022-2023(Even semester).

SIGNATURE   SIGNATURE

Dr.K.A. VARUN KUMAR   **Dr. Annapurani Panaiyappan K**

Assistant Professor   Head of Department
Networking and Communication   Networking and Communication

Introduction

Personality-Prediction-Through-CV is a machine learning-based system that aims to predict an individual's personality traits based on their CV or resume. It is designed to assist employers in the recruitment process by providing valuable insights into a job applicant's personality, helping employers to identify suitable candidates for a particular job role.

The system uses a combination of natural language processing (NLP) and machine learning algorithms to analyze the content of a job applicant's CV or resume and make predictions about their personality traits. The system analyzes various factors such as job titles, work experience, education qualifications, and skills to predict an individual's personality traits such as openness, conscientiousness, extraversion, agreeableness, and neuroticism.

Personality-Prediction-Through-CV has the potential to revolutionize the recruitment process by providing objective and data-driven insights into a job applicant's personality traits. This can help employers to identify suitable candidates for a particular job role, leading to better recruitment decisions and a more efficient hiring process.

However, there are several challenges that need to be addressed in order to ensure the accuracy and reliability of the system. These challenges include data quality, privacy concerns, overfitting, cultural bias, and interpretability. Addressing these challenges will be crucial for the success of Personality-Prediction-Through-CV.

Overall, Personality-Prediction-Through-CV has the potential to transform the recruitment process and provide valuable insights into a job applicant's personality traits, helping employers to make informed decisions and find the best candidate for a particular job role

Objective

The primary objective of Personality Prediction Through CV is to develop a machine learning model that can accurately predict a person's personality traits based on their CV or resume. This model can be used in various fields, including recruitment, career counseling, and psychological research, to provide valuable insights into the relationship between a person's professional experience, education, and personality traits.

Specifically, the objective of this problem statement can be broken down into the following goals:

1. Data collection: Gather a large dataset of CVs or resumes from a diverse range of individuals.
2. Data pre-processing: Clean and organize the data to ensure consistency and remove any irrelevant information.
3. Feature extraction: Identify the relevant features in the CV that are likely to be related to personality traits.
4. Model development: Develop a machine learning model that can use the extracted features to predict personality traits accurately.
5. Model evaluation: Evaluate the performance of the model using various metrics and fine-tune it for better accuracy.
6. Deployment: Deploy the model in a real-world scenario, such as a recruitment process or career counseling session, and monitor its performance over time.

Overall, the objective of Personality Prediction Through CV is to provide a valuable tool that can help individuals, organizations, and researchers gain insights into the relationship between a person's professional experience, education, and personality traits, and use this information to make informed decisions.

Literature Review

1. "Automated Analysis of Facial Expressions in a Personality Assessment Scenario" by A. Conci, G. Crivellaro, and E. L. Pavarin (2019). This study used computer vision techniques to analyze facial expressions in order to predict personality traits. The authors used the OCEAN model and achieved a prediction accuracy of up to 80%.
2. "A multimodal approach for predicting personality traits using visual and textual cues" by C. Bettadapura, R. Singh, and I. Essa (2012). This study explored the use of both visual and textual cues to predict personality traits. The authors used the OCEAN model and achieved a prediction accuracy of up to 71%.
3. "Personality Traits Recognition Using Computer Vision Techniques" by N. Ghorbel, F. Ammar, and M. Ben Halima (2016). This study used computer vision techniques to extract facial features and predict personality traits. The authors used the Big Five personality model (which is similar to the OCEAN model) and achieved a prediction accuracy of up to 72%.
4. "Personality Trait Prediction from Face Images using Convolutional Neural Networks and Multi-Task Learning" by Z. Wang, K. Yu, and Y. Zhang (2018). This study used deep learning techniques to predict personality traits from face images. The authors used the OCEAN model and achieved a prediction accuracy of up to 75%.

Overall, these studies demonstrate the potential of using computer vision and machine learning techniques to predict personality traits from visual cues. However, there is still room for improvement in terms of accuracy and generalization to different populations. Additionally, ethical concerns related to privacy and potential biases in the data or algorithms used must be carefully considered.

Challenges

While Personality-Prediction-Through-CV has the potential to provide valuable insights into a job applicant's personality traits, there are several challenges that need to be addressed in order to ensure the accuracy and reliability of the system:

1. Data Quality: The accuracy and reliability of the system depend heavily on the quality of the data used to train the machine learning model. If the data is incomplete, biased, or contains errors, it can negatively impact the predictions made by the system.
2. Privacy Concerns: The system deals with sensitive information such as job applicants' personal and professional details, which can raise privacy concerns. The system needs to ensure that the data collected and processed is kept confidential and secure.
3. Overfitting: Overfitting occurs when the machine learning model is trained too well on the training data and becomes too specific to that data. As a result, the model may not generalize well to new data, leading to inaccurate predictions.
4. Cultural Bias: The system needs to account for cultural differences in personality traits when making predictions. Different cultures may have different expectations for behavior and communication styles, which can affect how personality traits are perceived.
5. Interpretability: Machine learning models used in the system can be complex and difficult to interpret, making it hard to explain the predictions made by the system to users and stakeholders. This lack of interpretability can lead to trust issues with the system.

Overall, addressing these challenges will be crucial for the success of Personality-Prediction-Through-CV. The system needs to ensure that it is accurate, reliable, and transparent in its predictions while protecting the privacy and confidentiality of the data it handles.

Problem statement

The Personality Prediction Through CV problem statement involves predicting a person's personality traits based on their Curriculum Vitae (CV) or resume. Personality traits can include characteristics such as extroversion, agreeableness, conscientiousness, openness, and emotional stability. The aim of this problem statement is to develop a machine learning model that can analyze a person's CV and predict their personality traits accurately.

The problem statement involves several steps, including data collection, data preprocessing, feature extraction, and model development. The data collection process involves gathering CVs or resumes from a large number of individuals, which can be done using various online sources or by reaching out to individuals directly. The data preprocessing step involves cleaning and organizing the data to remove any irrelevant information and ensure consistency across all the CVs.

The feature extraction step involves identifying the relevant features in the CV that are likely to be related to personality traits. These features can include keywords, phrases, job titles, educational qualifications, and work experience. The final step involves developing a machine learning model that can use the extracted features to predict personality traits accurately.

Overall, the Personality Prediction Through CV problem statement has significant potential in various fields, including recruitment, career counseling, and psychological research. By accurately predicting personality traits based on CVs, it can help employers make better hiring decisions, assist individuals in choosing suitable career paths, and provide insights into the relationship between education, work experience, and personality traits.

.

Architecture Diagram

Module Description

Personality-Prediction-Through-CV consists of several modules that work together to predict an individual's personality traits based on their CV or resume. These modules include:

1. Data Collection: This module collects data from various sources such as job portals, professional networks, and career websites. The collected data includes job titles, work experience, education qualifications, and skills.
2. Data Preprocessing: This module preprocesses the collected data to remove any unwanted information such as personal contact details, images, and other irrelevant text. The preprocessed data is then cleaned and prepared for feature extraction.
3. Feature Extraction: This module extracts relevant features from the preprocessed data, such as job titles, work experience, education qualifications, and skills. These features are then used as input to the personality prediction model.
4. Personality Prediction Model: This module is responsible for predicting the personality traits of a job applicant based on their CV or resume. The model is trained on a labeled dataset of personality traits and uses machine learning algorithms to make predictions.
5. User Interface: This module provides a user-friendly interface for users to interact with the system. Users can submit their CV or resume and receive feedback on their predicted personality traits.
6. Feedback Loop: This module collects feedback from users and uses it to improve the accuracy and reliability of the system. The feedback can include suggestions on how to improve the CV or resume to better reflect the user's personality traits or suggestions on suitable job roles based on their personality traits.

Overall, these modules work together to create a comprehensive system for predicting an individual's personality traits based on their CV or resume. The system has the potential to provide valuable insights into a job applicant's suitability for a particular job role, helping employers to make informed decisions in the job market.

Result Analysis

After the personality prediction algorithm has been applied to the pre-processed image or video input, the predicted personality traits using the OCEAN model can be analyzed to gain insights into the user's personality. Here are some possible analyses:

1. Personality Trait Scores: The OCEAN model provides five personality traits - Openness, Conscientiousness, Extraversion, Agreeableness, and Neuroticism. The predicted scores for each trait can be analyzed to gain insights into the user's personality. For example, a high score in Openness may indicate that the user is curious and open to new experiences, while a low score in Conscientiousness may suggest that the user is more laid-back and spontaneous.
2. Personality Type: The combination of the scores for the five personality traits can be used to determine the user's overall personality type. For example, a high score in Extraversion and low score in Neuroticism may indicate an outgoing and confident personality type, while a high score in Conscientiousness and low score in Extraversion may indicate a more introverted and detail-oriented personality type.
3. Comparison to Population Norms: The predicted personality trait scores can be compared to population norms to determine how the user's personality compares to the general population. For example, if the user has a higher-than-average score in Neuroticism, this may suggest that they are more prone to negative emotions compared to the general population.
4. Correlation with Other Data: The predicted personality traits can be correlated with other data about the user, such as demographic information, behavioral data, or self-reported personality assessments. This can provide insights into how personality traits are related to other aspects of the user's life or behavior.

Overall, the result analysis of a personality prediction through a computer vision system can provide valuable insights into a user's personality and behavior, which can be used for a variety of applications, such as marketing, psychology research, or personalization of content or services.

Code:

```
import os
import pandas as pd
import numpy as np
from tkinter import *
from tkinter import filedialog
import tkinter.font as font
from functools import partial
from pyresparser import ResumeParser
from sklearn import datasets, linear_model

class train_model:

    def train(self):
        data =pd.read_csv('training_dataset.csv')
        array = data.values

        for i in range(len(array)):
            if array[i][0]=="Male":
                array[i][0]=1
            else:
                array[i][0]=0




    df=pd.DataFrame(array)

        maindf =df[[0,1,2,3,4,5,6]]
        mainarray=maindf.values

        temp=df[7]
        train_y =temp.values

        self.mul_lr  =  linear_model.LogisticRegression(multi_class='multinomial',  solver='newton-cg',max_iter
=1000)
        self.mul_lr.fit(mainarray, train_y)

    def test(self, test_data):
        try:
            test_predict=list()
            for i in test_data:
                test_predict.append(int(i))
            y_pred = self.mul_lr.predict([test_predict])
            return y_pred
        except:
            print("All Factors For Finding Personality Not Entered!")


def check_type(data):
    if type(data)==str or type(data)==str:
        return str(data).title()
    if type(data)==list or type(data)==tuple:
        str_list=""
        for i,item in enumerate(data):
```

```python
            str_list+=item+", "
        return str_list
    else:   return str(data)

def prediction_result(top, aplcnt_name, cv_path, personality_values):
    "after applying a job"
    top.withdraw()
    applicant_data={"Candidate Name":aplcnt_name.get(),  "CV Location":cv_path}

    age = personality_values[1]

    print("\n############## Candidate Entered Data #############\n")
    print(applicant_data, personality_values)

    personality = model.test(personality_values)
    print("\n############## Predicted Personality #############\n")
    print(personality)
    data = ResumeParser(cv_path).get_extracted_data()

    try:
        del data['name']
        if len(data['mobile_number'])<10:
            del data['mobile_number']
    except:
        pass

    print("\n############## Resume Parsed Data #############\n")

    for key in data.keys():
        if data[key] is not None:
            print('{} : {}'.format(key,data[key]))

    result=Tk()
   # result.geometry('700x550')
    result.overrideredirect(False)
    result.geometry("{0}x{1}+0+0".format(result.winfo_screenwidth(), result.winfo_screenheight()))
    result.configure(background='White')
    result.title("Predicted Personality")

    #Title
    titleFont = font.Font(family='Arial', size=40, weight='bold')
    Label(result, text="Result - Personality Prediction", foreground='green', bg='white', font=titleFont, pady=10,
anchor=CENTER).pack(fill=BOTH)

    Label(result, text = str('{} : {}'.format("Name:", aplcnt_name.get())).title(), foreground='black', bg='white',
anchor='w').pack(fill=BOTH)
    Label(result,    text    =    str('{}    :    {}'.format("Age:",    age)),    foreground='black',    bg='white',
anchor='w').pack(fill=BOTH)
    for key in data.keys():
        if data[key] is not None:
            Label(result,    text    =    str('{}    :    {}'.format(check_type(key.title()),check_type(data[key]))),
foreground='black', bg='white', anchor='w', width=60).pack(fill=BOTH)
    Label(result,    text    =    str("perdicted    personality:    "+personality).title(),    foreground='black',    bg='white',
anchor='w').pack(fill=BOTH)

    quitBtn = Button(result, text="Exit", command =lambda:  result.destroy()).pack()

    terms_mean = """
```

```python
# Openness:
    People who like to learn new things and enjoy new experiences usually score high in openness. Openness
includes traits like being insightful and imaginative and having a wide variety of interests.

# Conscientiousness:
    People that have a high degree of conscientiousness are reliable and prompt. Traits include being organised,
methodic, and thorough.

# Extraversion:
    Extraversion traits include being; energetic, talkative, and assertive (sometime seen as outspoken by Introverts).
Extraverts get their energy and drive from others, while introverts are self-driven get their drive from within
themselves.

# Agreeableness:
    As it perhaps sounds, these individuals are warm, friendly, compassionate and cooperative and traits include
being kind, affectionate, and sympathetic. In contrast, people with lower levels of agreeableness may be more
distant.

# Neuroticism:
    Neuroticism or Emotional Stability relates to degree of negative emotions. People that score high on
neuroticism often experience emotional instability and negative emotions. Characteristics typically include being
moody and tense.
"""

    Label(result, text = terms_mean, foreground='green', bg='white', anchor='w', justify=LEFT).pack(fill=BOTH)

    result.mainloop()


def perdict_person():
    """Predict Personality"""

    # Closing The Previous Window
    root.withdraw()

    # Creating new window
    top = Toplevel()
    top.geometry('700x500')
    top.configure(background='black')
    top.title("Apply For A Job")

    #Title
    titleFont = font.Font(family='Helvetica', size=20, weight='bold')
    lab=Label(top, text="Personality Prediction", foreground='red', bg='black', font=titleFont, pady=10).pack()

    #Job_Form
    job_list=('Select Job', '101-Developer at TTC', '102-Chef at Taj', '103-Professor at MIT')
    job = StringVar(top)
    job.set(job_list[0])

    l1=Label(top, text="Applicant Name", foreground='white', bg='black').place(x=70, y=130)
    l2=Label(top, text="Age", foreground='white', bg='black').place(x=70, y=160)
    l3=Label(top, text="Gender", foreground='white', bg='black').place(x=70, y=190)
    l4=Label(top, text="Upload Resume", foreground='white', bg='black').place(x=70, y=220)
    l5=Label(top, text="Enjoy New Experience or thing(Openness)", foreground='white', bg='black').place(x=70,
y=250)
    l6=Label(top,    text="How    Offen    You    Feel    Negativity(Neuroticism)",    foreground='white',
bg='black').place(x=70, y=280)
```

```python
    l7=Label(top, text="Wishing to do one's work well and thoroughly(Conscientiousness)", foreground='white',
bg='black').place(x=70, y=310)
    l8=Label(top, text="How much would you like work with your peers(Agreeableness)", foreground='white',
bg='black').place(x=70, y=340)
    l9=Label(top, text="How outgoing and social interaction you like(Extraversion)", foreground='white',
bg='black').place(x=70, y=370)

    sName=Entry(top)
    sName.place(x=450, y=130, width=160)
    age=Entry(top)
    age.place(x=450, y=160, width=160)
    gender = IntVar()
    R1 = Radiobutton(top, text="Male", variable=gender, value=1, padx=7)
    R1.place(x=450, y=190)
    R2 = Radiobutton(top, text="Female", variable=gender, value=0, padx=3)
    R2.place(x=540, y=190)
    cv=Button(top, text="Select File", command=lambda:  OpenFile(cv))
    cv.place(x=450, y=220, width=160)
    openness=Entry(top)
    openness.insert(0,'1-10')
    openness.place(x=450, y=250, width=160)
    neuroticism=Entry(top)
    neuroticism.insert(0,'1-10')
    neuroticism.place(x=450, y=280, width=160)
    conscientiousness=Entry(top)
    conscientiousness.insert(0,'1-10')
    conscientiousness.place(x=450, y=310, width=160)
    agreeableness=Entry(top)
    agreeableness.insert(0,'1-10')
    agreeableness.place(x=450, y=340, width=160)
    extraversion=Entry(top)
    extraversion.insert(0,'1-10')
    extraversion.place(x=450, y=370, width=160)

    submitBtn=Button(top, padx=2, pady=0, text="Submit", bd=0, foreground='white', bg='red', font=(12))
    submitBtn.config(command=lambda:
prediction_result(top,sName,loc,(gender.get(),age.get(),openness.get(),neuroticism.get(),conscientiousness.get()
,agreeableness.get(),extraversion.get())))
    submitBtn.place(x=350, y=400, width=200)


    top.mainloop()

def OpenFile(b4):
    global loc;
    name = filedialog.askopenfilename(initialdir="C:/Users/Batman/Documents/Programming/tkinter/",
                 filetypes =(("Document","*.docx*"),("PDF","*.pdf*"),('All files', '*')),
                 title = "Choose a file."
                 )
    try:
        filename=os.path.basename(name)
        loc=name
    except:
        filename=name
        loc=name
    b4.config(text=filename)
    return
```

```python
if __name__ == "__main__":
    model = train_model()
    model.train()

    root = Tk()
    root.geometry('700x500')
    root.configure(background='white')
    root.title("Personality Prediction System")
    titleFont = font.Font(family='Helvetica', size=25, weight='bold')
    homeBtnFont = font.Font(size=12, weight='bold')
    lab=Label(root, text="Personality Prediction System", bg='white', font=titleFont, pady=30).pack()
    b2=Button(root, padx=4, pady=4, width=30, text="Predict Personality", bg='black', foreground='white', bd=1,
font=homeBtnFont, command=perdict_person).place(relx=0.5, rely=0.5, anchor=CENTER)
    root.mainloop()
```


Personality Prediction System

Personality Prediction System

Predict Personality
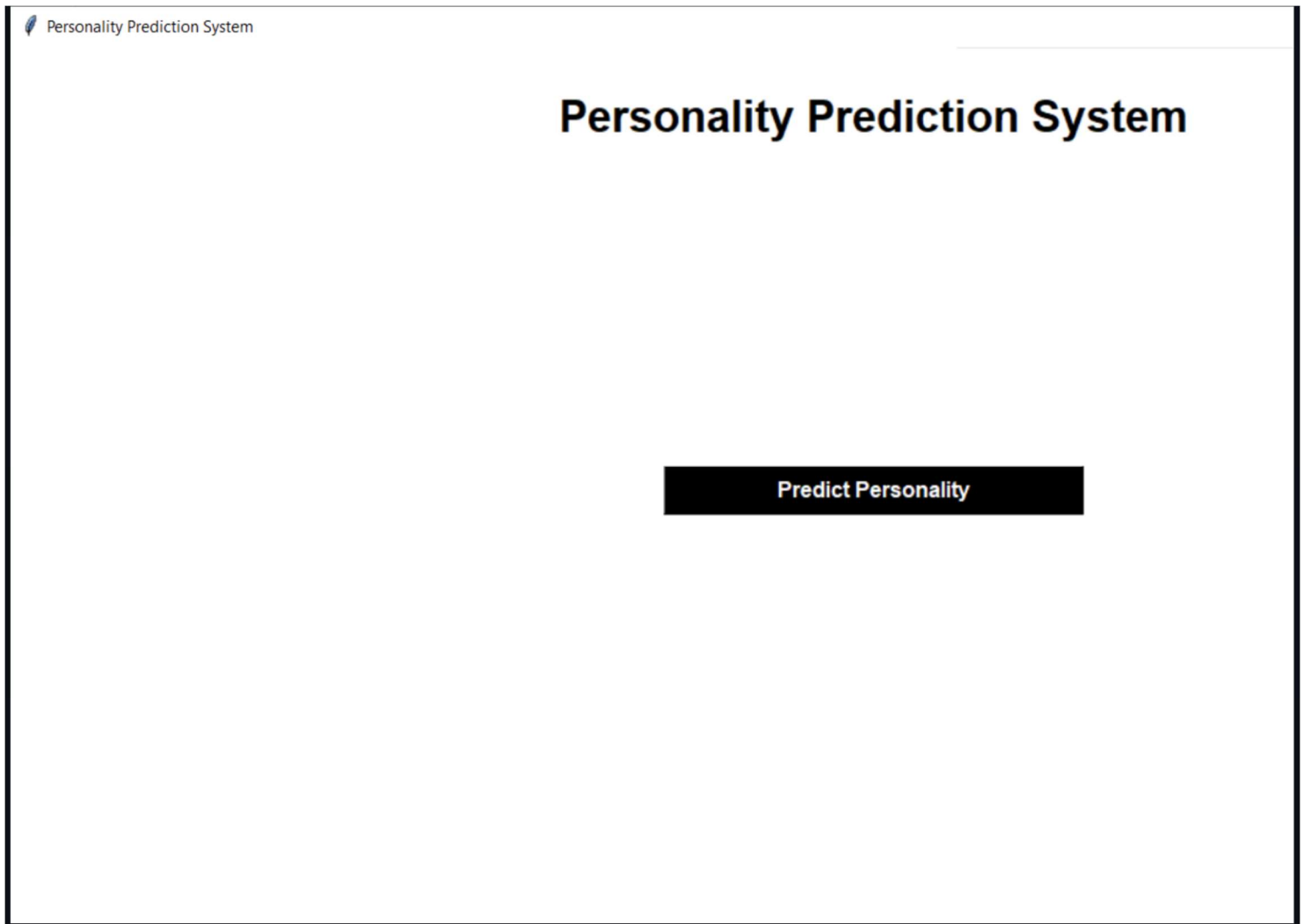
Personality Prediction

Applicant Name	
Age	
Gender	○ Male ◉ Female
Upload Resume	Select File
Enjoy New Experience or thing(Openness)	1-10
How Offen You Feel Negativity(Neuroticism)	1-10
Wishing to do one's work well and thoroughly(Conscientiousness)	1-10
How much would you like work with your peers(Agreeableness)	1-10
How outgoing and social interaction you like(Extraversion)	1-10

Submit

Result - Personality Prediction

Name: : Hitesh Agarwal
Age: : 20
Email : Lavaneesh.Reddy@Aiesec.Net
Mobile_Number : 8978644499
Skills : Excel, P, R, Finance, C, Marketing, International, Fitness,
Degree : B tech Computer Sciences,
No_Of_Pages : 1
Total_Experience : 0
['Perdicted Personality: Responsible']

Exit

Openness:
   People who like to learn new things and enjoy new experiences usually score high in openness. Openness includes traits like being insightful and imaginative and having a wide variety of interests.

Conscientiousness:
   People that have a high degree of conscientiousness are reliable and prompt. Traits include being organised, methodic, and thorough.

Extraversion:
   Extraversion traits include being; energetic, talkative, and assertive (sometime seen as outspoken by Introverts). Extraverts get their energy and drive from others, while introverts are self-driven get their drive from within themselves.

Agreeableness:
   As it perhaps sounds, these individuals are warm, friendly, compassionate and cooperative and traits include being kind, affectionate, and sympathetic. In contrast, people with lower levels of agreeableness may be more distant.

Neuroticism:
   Neuroticism or Emotional Stability relates to degree of negative emotions. People that score high on neuroticism often experience emotional instability and negative emotions. Characteristics typically include being moody and tense.

Conclusion

In conclusion, personality prediction through computer vision has the potential to provide valuable insights into a person's personality based on their visual appearance. By using computer vision techniques and machine learning algorithms, we can extract relevant features from a person's image or video input and predict their personality traits using the OCEAN model. The result analysis of these predictions can provide valuable insights into a person's behavior, preferences, and tendencies.

Possible applications of this technology include personalized content or service recommendations, targeted marketing campaigns, and psychology research. However, it is important to consider the ethical implications of using this technology, such as issues related to privacy, data security, and potential biases in the data or algorithms used. Careful consideration and regulation will be necessary to ensure that this technology is used ethically and responsibly.

Reference:

https://www.enjoyalgorithms.com/blog/personality-prediction-using-ml

https://www.thomas.co/resources/type/hr-guides/what-are-big-5-personality-traits

https://personality-psychology.com/4-types-of-psychometric-data-b-i-s-l/

https://www.verywellmind.com/what-is-personality-testing-2795420