**Date: 24 September, 2023**

**PROJ-CS781: Mid-term Review**

|  |  |
| --- | --- |
| **Project Title:** Personality Traits Analysis | **Group: 14** |
| **Team (Roll / Name):**   1. 13000120063 / Sanjeev Pratap 2. 13000120078 / Sudakshina Majumdar 3. 13000120080 / Suman Raj 4. 13000120081 / Sumana Sen | **Guide: Prof. Nairanjana Chowdhury** |

**Questions:**

1. Justify the uniqueness of your project.

**Solution:** The uniqueness of the project on personality traits analysis lies in its ability to leverage images as a primary data source for predicting user private traits. Unlike traditional studies that rely on text-based information, this project taps into the growing popularity of visual content, particularly among younger generations, and the shift towards image-centric social networks like Facebook and Twitter. By conducting content analysis solely on images, the project offers a distinct approach to assessing personality dimensions. This innovative methodology not only broadens the scope of information used for prediction but also reflects the evolving landscape of social media and user behavior, making it highly relevant and timely for understanding and predicting private traits in today's digital age.

2. What are the expected benefits from your project?

**Solution:** This project holds immense usefulness across diverse fields:

* Personal Growth and Development: Individuals can gain self-awareness, identify strengths, and work on areas for improvement, leading to personal growth and better life decisions.
* Interpersonal Relationships: Better understanding of one's own and others traits enhances communication, conflict resolution, and relationship-building skills.
* Organizational Productivity: Businesses can optimize team dynamics, job-role fit, and leadership development for enhanced productivity and employee satisfaction.
* Education: Educational institutions can adapt teaching methods to students learning styles, leading to improved engagement and academic outcomes.

New Ideas / Innovations for the Project:

* **Real-time Analysis**: Develop tools that provide real-time personality traits analysis, adapting to individuals changing behaviours and life circumstances. For example, observations from their social media profiles where they like, share, comment and post will help in analysing their personality.

Potential Market Demand:

* **Corporate Sector**: Businesses are increasingly recognizing the value of personality traits analysis for optimizing team dynamics, leadership development, and employee satisfaction.
* **Education Industry**: Schools, colleges, and educational platforms can use personality traits analysis to personalize learning experiences and improve student outcomes.
* **Mental Health Services**: Therapists and counsellors can gain useful insights from personality traits analysis for more effective treatment planning and emotional support.
* **Recruitment and HR**: Companies can use personality traits analysis in recruitment processes to match candidates with job roles, leading to improved employee retention.
* **Marketing and Consumer Insights**: Businesses can tap into personality traits analysis for targeted marketing strategies and product development.

3. a) State the findings from your analysis till date.

**Solution:** Successful data preprocessing has significantly improved data quality, leading to better feature extraction results. Feature extraction techniques, both for text and image data, have shown promising initial results in capturing relevant personality traits. The choice of machine learning models for text-based analysis has yielded competitive accuracy and performance. Image-based personality prediction has demonstrated the potential to complement text-based predictions.

b) What is the % completion progress of analysis? Explain the calculation logic.

**Solution:** About 79% analysis completed. The % completion progress of analysis can be calculated by assessing the tasks completed relative to the total tasks planned for the analysis phase. For instance, if the analysis phase was planned to have 10 key tasks, and 7 of them have been completed, the progress would be 70%. The calculation logic is: (Tasks Completed / Total Planned Tasks) \* 100 = % Completion Progress.

4. a) State the findings from your design till date.

Solution: The user interface design has been successfully developed, focusing on user-friendliness and efficient data input handling. The integration of both text and image data in the multimodal fusion technique has shown potential for enhanced personality prediction. Fine-tuning strategies for machine learning and deep learning models have resulted in improved model performance.

b) What is the % completion progress of design? Explain the calculation logic.

Solution: The % completion progress of design can be calculated similarly to the analysis phase. For example, if there were 8 design tasks planned, and 5 have been completed, the progress would be 62.5%. The calculation logic is: (Tasks Completed / Total Planned Tasks) \* 100 = % Completion Progress.

5. State the tools you are using for analysis and design (refer Software Engineering).

Solution: These are the following tools and Technologies which we are using for analysis and designing the project:

* Programming Language: Python
* Development Environment: Jupyter Notebooks, IDEs (Visual Studio Code, PyCharm, Spyder)
* Libraries and Frameworks: Keras, Pandas and Numpy, TensorFlow, Scikit-Learn
* Github
* NLTK (Natural Language Toolkit), spaCy, or Hugging Face Transformer for text processing and sentiment analysis.

6. Are you foreseeing any risk in completing your project?

**Solution:** Some foreseeing risk in completing our project are:

* It might not be 100% accurate prediction. For Some people it may be exact but for some people we can’t analyze their personality by their social media activities.
* Gathering and analyzing personal images from social media platforms may raise privacy concerns.
* Analyzing users' private traits, even from publicly available data, raises ethical questions. Ensuring responsible and ethical use of the collected data is essential to maintain public trust and avoid backlash.
* The quality of image data can vary significantly. Low-resolution images or images with noise can hinder the accuracy of personality trait predictions.
* Building accurate personality prediction models, especially from images, can be computationally intensive and complex. Managing the computational resources and ensuring model scalability is a potential risk.
* Personality traits inferred from images may not be as straightforward to interpret as text-based traits. Developing clear and reliable methods for interpreting visual cues is a challenge.
* Social media platforms and user behaviors are constantly evolving. So, our project may face challenges in keeping up with these changes and adapting its methods accordingly.

7. State the study references you have used in this semester.

**Solution:** References which we have taken in this semester are:

1. Golbeck, J., Robles, C., & Turner, K. (2011, May). Predicting personality with social media. In CHI'11 Extended Abstracts on Human Factors in Computing Systems (pp. 253-262). ACM

2. R. Bin Tareaf, P. Berger, P. Hennig, C. Meinel,” Personality Exploration System for Online Social Networks: Facebook Brands as a Use Case.” In 2018 IEEE/WIC/ACM International Conference on Web Intelligence (WI), pp. 301-309. IEEE, 2018.

3. Q. Daniele, M. Kosinski, D. Stillwell, J. Crowcroft.” Our twitter profiles, our selves: Predicting personality with twitter.” In 2011 IEEE third international conference on privacy, security, risk and trust and 2011 IEEE third international conference on social computing, pp. 180-185. IEEE, 2011.

4. https://www.ijert.org/personality-analysis-using-social-media

8. Submission of RM (excel file) and PP (Microsoft Project Plan) as separate attachments following the given templates

Solution: Submitted.

9. What is the % completion progress of the prototype? Explain the calculation logic.

Solution: It is about 70% completed.

10. Additionally, Guides should ask questions (what’s, how-to’s) on understanding of the target system and expected functions.