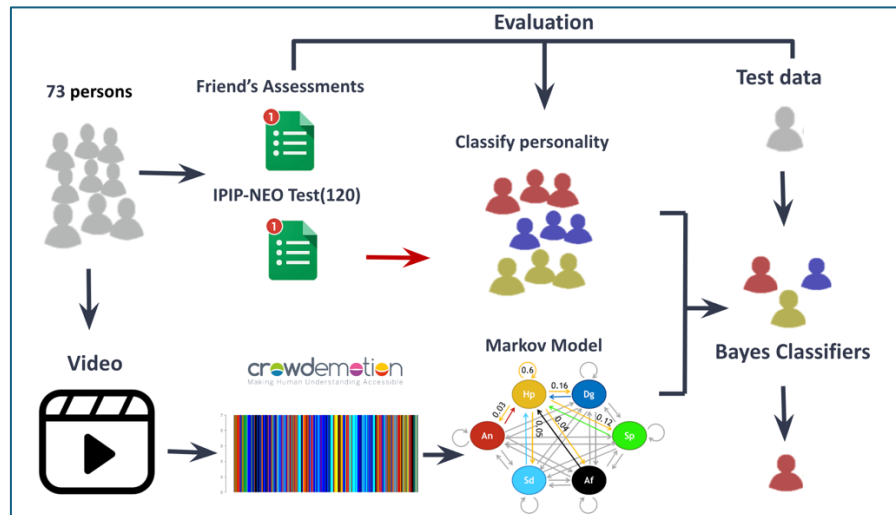


Personality Prediction from Dynamic Facial Expressions



Overview

This project explores the potential of using dynamic facial expressions to predict personality traits. Utilizing machine learning algorithms and data analysis, we aimed to establish correlations between facial expressions and the 'Big Five' personality traits.

Data Collection

The data collection phase involved two primary data sources:

- **Personality Test Data:** Participants completed the IPIP-NEO personality test, a well-regarded psychological assessment tool that quantifies the 'Big Five' personality traits: Openness, Conscientiousness, Extraversion, Agreeableness, and Neuroticism. The data collection was facilitated via a translated version of the test, administered through a Google Forms interface.
- **Facial Expression Data:** The researchers captured facial expressions using a desktop camera during controlled conversations with 73 participants. The average age was 20 (± 1.6 years old). This facial data was analyzed for emotional expression at each frame using the Crowdemotion API, which classified

expressions into six categories: Happiness, Sadness, Disgust, Anger, Surprise, and Fear.

Data Analysis

In the subsequent data analysis phase:

- The study utilized the R statistical software, equipped with the httr package, to automate the uploading of test results to the IPIP-NEO website, fetching back a report comprising 35 percentage scores of the personality traits for each participant and their friends' assessment.
- Emotional data, encoded as time-sequenced expressions, were downloaded for further analysis.

Algorithms

Three primary algorithms were applied in the study:

- **Hierarchical Clustering:** This method was used to group participants based on their personality test results. The approach involved computing similarity measures, such as correlation coefficients and Euclidean distances, between participants. The hierarchical clustering was performed using three different agglomerative techniques: complete-linkage, average-linkage, and single-linkage clustering. These techniques systematically merge the data points (or clusters) that are closest together until all points have been merged into a single cluster or into the desired number of clusters.
- **K-means Clustering:** Separately from hierarchical clustering, K-means clustering was applied to the 73 participants as a method of partitioning them into k distinct clusters based on their personality test scores. K-means clustering aims to minimize the variance within clusters and is particularly useful for identifying distinct groupings when the number of clusters (k) is specified in advance. This method iteratively assigns each data point to the nearest cluster center (mean) and recalculates the new means until the assignment no longer changes.
- **Markov Model:** A Markov Model was constructed to describe the transitions

between different emotional states, assuming that the present emotional state depends solely on the previous state and not on any earlier states. This model quantified the probability of transitions between the six defined emotional states.

- **Naïve Bayes Classifiers:** The Naïve Bayes Classifier was employed to predict the personality types based on the dynamic emotional data. The classifier worked by averaging the emotion transition probabilities for participants sharing similar personality profiles and applying these to predict the personality classification of an unseen participant based on their facial expressions.

Results

The system's predictive performance was rigorously evaluated:

- **Predictive Performance:** The Naïve Bayes Classifier yielded predictions with an average accuracy of 0.431, which was superior to the accuracy of friends' assessments, which stood at 0.287. The predictions were also found to be comparable to those made by close friends.
- **Validation Techniques:** Validation of the predictive model was conducted using the Leave One Out and Random Subsampling techniques. It was noted that while the Leave One Out method was subject to the availability of substantial training data, Random Subsampling provided a more robust assessment over repeated random selections of test data.

Conclusion

The study concluded that dynamic facial expression analysis could be employed reliably to predict personality traits within a short interaction time. The level of accuracy approached that of an acquaintance and showed considerable promise as compared to the detailed knowledge of a close friend. However, the researchers also acknowledged the need for a more extensive and diverse dataset to enhance prediction stability and generalizability.

Practical Applications

- Digital Interfaces: Tailoring user experiences based on inferred personality traits from facial expressions.
- Mental Health Services: Assisting professionals in assessing client personalities more swiftly and objectively.
- HR and Team Building: Improving recruitment processes and team dynamics with deeper personality insights.
- Educational Technology: Creating adaptive learning environments responsive to students' emotional and cognitive states.
- Customer Service Bots: Empowering bots with the ability to understand clients' needs and emotional responses more effectively.
- Gaming Industry: Adapting game narratives and difficulty in response to players' emotional cues.
- Marketing and Advertising: Tailoring campaigns to different personality profiles to increase engagement and conversion rates.

The implications of this research are vast, suggesting a future in which technology interacts more intuitively with human emotions and personalities.