KØBENHAVNS UNIVERSITET

Cognitive Science 2 Exam Project

Multimodality Research in IT & Cognition

Multimodal Automatic Personality Recognition

Supervisors: Costanza Navarretta & Patrizia Paggio

Department of Scandinavian Studies and Linguistics. June 3, 2021

Group members

Navn	KUid
Lasse Goul Jensen	kvh190
Niels Krog	lkt259

Abstract

The following paper investigates the possibilities of predicting personality traits from multimodal features, namely physiological measurements and facial landmarks using the ASCERTAIN dataset. We implement and compare the performance of three different machine learning classification models, logistic regression, random forest and multilayer perceptron. We compare the results between unimodal and multimodal features between each model and with state of the art implementations. We found that multilayer perceptron and random forest outperformed the state of the art with a maximum F1 score of 0.87 and 0.82, respectively.

Division of Labour

Both authors have taken equal part in this project both in the implementation and the report. The authors have had equal influence on all decisions taken in the workflow. Below is a table of where we have contributed the most.

Section	Written by
1.0.1	Lasse
1.0.2	Niels
1.1 - 1.4	Lasse & Niels
2	Niels
3.1 - 3.2	Lasse
3.3	Niels
4	Lasse & Niels
5	Lasse & Niels
6	Lasse & Niels

Keystrokes: 19164

Contents

1 Introduction

Personality has a large impact on how people live their lives, and thus also how they interact with everything from other people, to computer interfaces and the world in general. It is relevant to all areas of computing that somehow deal with human behaviour [19] (e.g. make better user experience), and because of this, the automatic prediction of personality traits is becoming more and more in demand, and prediction using multimodal data has specifically gained a lot of traction in recent years [13]. In this report, we will explore the possibility and accuracy of predicting personality using data from the ASCERTAIN dataset [18], which includes physiological response data and facial movement data. We will compare a full multimodal approach, employing all of the features, to unimodal approaches using both modalities separately, to test our hypothesis that a multimodal approach will outperform either of the unimodal approaches.

1.0.1 Personality: Personality types or traits have been discussed at least as far back as The Roman Empire, where Galen proposed his system of the four human temperaments [16]. However, most of our modern understanding of personality types stems from Carl Jung's theory of universal types of human psychology [7] in which he describes the different modifiers of personality, among them the now well-known concepts of introversion versus extraversion. From here, a range of different modern personality models sprung, among them the Meyer-Briggs Type indicator [14], which builds on top of Jung's theory, the DISC model which is instead based on the work of William Moulton Marston [12] and the Big-Five model by Goldberg, building upon the works of Allport and Odberg [4]. The Big-Five model is the most used model in personality prediction by far [13, p. 2314], and is also the one which our dataset has used. The model builds upon the idea of five psychological traits that humans posses to a larger or smaller extent: openness, conscientiousness, extraversion, agreeableness and neuroticism. These traits are often measured by way of a Likert scale questionnaire, a good example of which has been made available online by openpsychometrics [1].

Since personality is such an abstract, psychological phenomenon, it might be hard to imagine why we can predict it using different kinds of observable data. To understand why, we take a look at the Brunswik Lens Model [3] which states that people *externalize* their personality by way of *distal cues* (e.g. frowning or timbre of voice) which others can then perceive. It is these cues that automatic personality recognition targets, and uses them to infer the personality of the person in question [19, p. 4-5].

1.0.2 Multimodality: A modality in cognitive science can be defined as being either perceptive (sense of hearing, sight etc.) or communicative (speech, facial expressions etc.). When we combine one or more of these modalities for any purpose, we work with what is known as multimodality. Multimodality is interesting to us for a number of reasons. First of all, a growing body of neuroscientific and psychological literature all point towards the brain as fundamentally wired for multimodal perception and integration [17]. Because of these findings, there has been a shift towards a view of human cognition more grounded in Gestalt theory, which dictates that instead of looking at isolated parts, we should look at the holistic whole [15, p. 20]. Translating this idea to classification tasks would mean that looking at multiple modalities when trying to classify, instead of a single modality, would work better. There are also plenty of examples of this being just the case, such as Kindiroglu et al. using audio and visual modalities to detect certain personality traits [8], Zhang et al. also using audio and visual modalities got the highest accuracy in ChaLearn Challenge 2016 for perceived personality analysis [21] and An who combined text and audio modalities for personality detection [2]. It is worth mentioning that just as these three, so are the vast majority of multimodal classification models bimodal. Very few trimodal models have been attempted as of yet [13, p. 2321].

1.1 Importance of Problem

The problem of automatic personality prediction is important because of its wide range of applications. Academic discussion and research have already begun on some of these applications, such as using personality as the basis for a recommender system [20], real-time personality detection for better interactions with robots [5], and the use of personality detection as a tool in the hiring process of companies [11]. Besides these, many other areas can easily be imagined to gain from automatic personality detection. HCI is an obvious area, where personality detection might be used as a tool to customize system feedback to best suit the user. Areas where psychological analysis of people are a key part may also be areas of interest, for example health and police work.

1.2 What is Lacking in Current Knowledge

The biggest problems of automatic personality detection in general is a lack of large, diverse datasets for training models, especially data-hungry neural networks, as well as a lack of diversity in the personality models employed, with the vast majority relying on the Big-Five model [13]. For specific areas, such as for use in HCI, the problem is to find the specific features that generate the best results, and the best classification algorithms to work with said features, which is what we are investigating in this report.

1.3 Research Question and Hypotheses

We want to investigate whether a multimodal approach to automatic personality recognition, using a series of different features, works better than a unimodal approach using only the features that are related to a single modality. We therefore focus on the following research question:

> "Is bimodal automatic personality recognition using physiological response data and non-communicative gestures better at personality recognition than unimodal recognition using each modality separately?"

Our hypothesis is that a multimodal approach, using all of the available features, will outperform a unimodal approach using only the physiological data, as well as a unimodal approach using only the facial data. We base this on previous research, which implies that a multimodal approach likely outperforms unimodal ones, as mentioned in subsubsection 1.0.2.

When we speak of automatic personality recognition, we mean the ability to perform "recognition of the true personality of an individual" [19, p. 1], as separate from personality perceived by others. This means, that the personality traits were gathered from self-report done by the test subjects.

In terms of classification methods, we will apply logistic regression (LR), random forest (RF) classifier and a multilayer perceptron (MLP). We hypothesise, that RF will perform the best, as LR is a simple classifier and the MLP may be too data-hungry for our chosen dataset.

1.4 Structure of Paper

In this paper, we will first take a look at the ASCERTAIN dataset by Subramanian et al. [18] in section 2. Here we will explain what the dataset features are and how they were obtained. In section 3 we will cover the methodology of the project. In section 4 we present our results, and in section 5 we discuss and conclude upon our findings.

2 The ASCERTAIN Dataset

The ASCERTAIN dataset (databa**AS**e for impli**C**it p**ER**sonali**T**y and **A**ffect recognit**I**o**N**) is a multimodal dataset created by Subramanian et al.[18]. It contains self-reported Big-Five personality traits from 58 participants along with Electroencephalogram (EEG), Electrocardiogram (ECG), Galvanic Skin Response (GSR) and facial activity recordings (EMO) when exposed to 36 different affective video clips with lengths ranging from 51 to 127 seconds.

The physiological readings were obtained using commercially available sensors, which assures scalability and ecological validation [18, p. 1]. However, this means the readings vary in quality, especially the ECG readings.

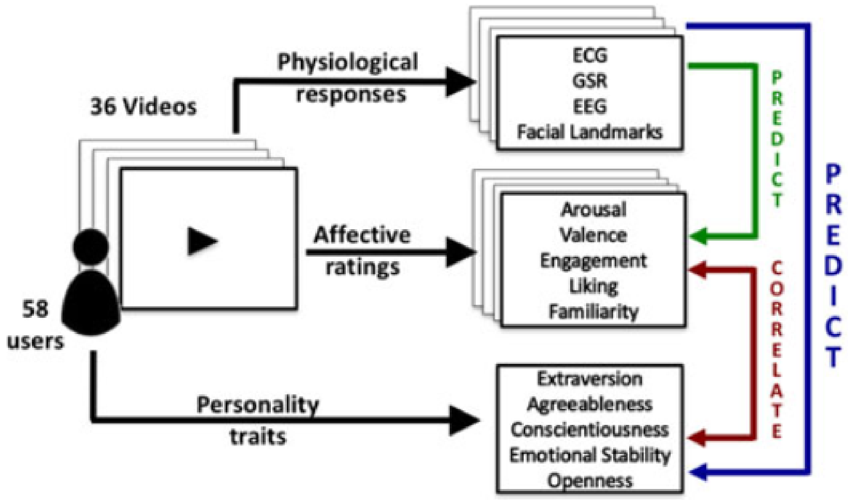

Figure 1: An overview of the contents, features and methods from the ASCERTAIN article[18]

3 Methods

In this section, we will first present the features used for predicting personality types, then the machine learning models used for solving the task and finally, how they are evaluated and trained. The results obtained from the models, along with the state of the art results for comparison will be presented in section 4.

3.1 Features

The ASCERTAIN dataset contains features for physiological responses, affective ratings and personality traits. Their connections made in [18] can be seen in figure 1, which shows how the physiological responses can predict both the affective ratings and personality traits, while the affective rating correlates with the personality traits. For the sake of the scope of this paper, we will focus on using physiological responses and facial markers to predict the Big Five personality traits in the participants. The trait, neuroticism is inverted and labeled *emotional stability* in the dataset.

We use the features extracted from the dataset, as opposed to the raw recordings. The features are presented in the tables below:

Modality	Number of Features	Feature Types
ECG	32	Low frequency (0-2.4 Hz) power spectral densities (PSDs) Very slow response (0-0.04 Hz) PSDs Inter beat intervals (IBI) Heart rate (HR) Statistics over heart rate variablility (HRV)
GSR	31	Mean skin resistance and mean of derivative Mean differential for negative values Proportion of negative derivative samples Number of local minima in GSR signal Average rising time for GSR signal Spectral power in 0-2.4 Hz band Zero crossing rate of skin conductance slow response (0-0.2 Hz) Zero crossing rate of skin conductance very slow response (0-0.08 Hz) Mean SCSR SCVSR peak magnitude
EEG	88	Average of first derivative Proportion of negative differential samples Mean number of peaks Mean derivative of the inverse channel signal Average number of peaks in the inverse signal Statistics over each of the 8 signal channels
EMO	72	Statistics over horizontal and vertical deformations of 12 motion units

Some features contain statistical measurements, which consits of six features: the mean, standard deviation, skewness, kurtosis of raw feature over time, percent of times the feature value is above the mean plus standard deviation and the percent of times the feature value is below the mean minus the standard deviation.

The features, low frequency PSDs, very slow response PSDs and IBI were removed from the ECG modality due to being corrupted. For some of the video clips, the sensors malfunctioned, resulting in missing values. The reading for such instances were dropped from the dataset.

3.2 Data Preprocessing

The task of classifying personality traits is a binary recognition, meaning the continuous values for the traits cannot be used as labels. Therefore, following the methodology in [18, p. 7], each method classifies whether a given trait is high or low. This definition is derived from being above or below the median in each trait.

Subramanian et al. used principal component analysis (PCA) to reduce their feature space when training their models. For each modality they captured 99% of the variance in the data with PCA. Though the dimensionality of the data is not too large for a consumer laptop to process, the reason behind reducing the feature space is to prevent overfitting, which is a risk when working with small amounts of data[18, p. 8]. The same preprocessing was applied to our data.

The data was shuffled and standardized to zero mean and one standard deviation before PCA being applied. The models was then trained using stratified 10-fold cross validation. Each model was trained 30 times, once for each combination between modality and personality trait.

In the paper by Subramanian et al. [18], each physiological measurement is considered a modality. They classify using each modality alone and a combination of them all. Like them, we will classify using this approach for comparison. However, we consider all physiological measurements as belonging to a single modality. Therefore, we will also test a combination of all the physiological features as one modality, and the facial landmark data as another.

For the combined physiological features (PHY), ECG and EMO combined (explained in section 4) and fusion of all features (ALL), we used what is known as early fusion [13, p. 2320] which is a fusion method where the chosen features extracted from different modalities are used together in classification. In the experiment performed by Subramanian et al. [18], they use something called decision fusion, described by Koelstra & Patras in their text *Fusion of Facial Expressions and EEG for Implicit Affective Tagging* [9]. However, as we wanted to use other models than SVM, this was not an option for us.

3.3 Machine Learning Models

Subramanian et al. used a Naïve Bayes classifier (NB), linear support vector machine (SVM) and radial basis function (RBF) SVM to predict the personality traits of their participants. We will compare their results to three different classifiers: logistic regression, random forest and a multilayer perceptron. All models will be implemented in Python 3 with the Scikit-Learn library and the metrics used for accuracy will be the F1 score to compare the results with [18]. All training times were achieved using a 2.3 GHz Intel core i5 processor. All training times and overall performance metrics of the models can be seen in appendix Table 3

3.3.1 Baseline: To know how well our solution can predict the personality traits of the participants better than chance, it will be compared to the baseline established in [18, p. 10], a class-ratio based voting, resulting in an F1 score of 0.5 for all methods on all traits. Our goal is to have a higher F1 score than 0.5 with our classifiers.

3.3.2 Logisitc Regression: LR is a simple classifier and we do not expect it to perform well, but has been included to compare different methods. The only parameter changed from the default is the maximum iterations, which was set to 2500 for all results to converge.

3.3.3 Random Forest: RF is an ensemble classifier, consisting of a series of tree based classifiers. It usually provides reasonable results with a short training time compared to neural networks. RF implemented for this purpose had 500 estimators.

3.3.4 Multilayer Perceptron: The MLP is a type of feed-forward neural network. It can outperform other classifiers, but as mentioned in subsection 1.2, they require a large quantity of training data, which is not the case for the ASCERTAIN dataset. The key hyperparameters for the classifier are maximum 2000 iterations and a architecture of one layer with 200 perceptrons.

4 Results

In this section, we present the performance of the classifiers on each trait/modality combination. The results from [18] are included for comparison. Furthermore, visualisations of the models will be showcased.

Trait	Classifier	ECG	EEG	GSR	EMO	PHY	ALL
	NB	0.56	**0.63**	0.35	0.45	-	0.65
	SVM	0.06	0.52	0.31	0.00	-	0.52
Extraversion	RBF SVM	0.53	0.48	0.45	0.35	-	0.57
	LR	0.63	**0.63**	0.57	0.66	0.66	0.74
	RF	0.72	0.62	**0.59**	0.76	0.72	0.77
	MLP	**0.74**	0.55	**0.59**	**0.82**	**0.78**	**0.86**
	NB	0.55	0.52	0.40	0.39	-	0.58
	SVM	0.45	0.12	0.35	0.34	-	0.46
	RBF SVM	0.32	**0.54**	0.42	0.27	-	0.53
Agreeable	LR	0.52	0.42	0.49	0.66	0.53	0.71
	RF	0.72	0.44	**0.59**	0.76	0.61	0.66
	MLP	**0.73**	0.51	0.53	**0.81**	**0.73**	**0.83**
	NB	0.60	0.35	0.39	0.57	-	0.65
	SVM	0.51	0.35	0.36	0.35	-	0.59
Conscient	RBF SVM	0.55	0.31	0.34	0.54	-	0.67
	LR	0.56	0.48	0.49	0.62	0.58	0.67
	RF	**0.72**	0.44	0.54	0.71	0.59	0.46
	MLP	**0.72**	**0.52**	**0.58**	**0.76**	**0.69**	**0.77**
	NB	0.53	0.26	0.44	0.49	-	0.59
	SVM	0.60	0.46	0.36	0.56	-	0.64
Em. Stab.	RBF SVM	0.58	0.51	0.47	0.73	-	0.77
	LR	0.62	**0.64**	0.59	0.56	0.68	0.72
	RF	0.72	0.41	0.58	0.69	0.75	0.78
	MLP	**0.73**	0.62	**0.60**	**0.78**	**0.78**	**0.82**
	NB	0.48	0.34	0.50	0.28	-	0.56
	SVM	0.35	0.36	0.36	0.36	-	0.42
Open	RBF SVM	0.49	0.37	0.26	0.53	-	0.53
	LR	0.48	0.43	0.25	0.58	0.57	0.67
	RF	**0.72**	**0.64**	**0.61**	0.43	0.53	0.56
	MLP	0.67	0.45	0.48	**0.78**	**0.69**	**0.80**

Table 1: F1 scores for each modality/feature combination. The feature *PHY* is all physiological measurements combined, which is not included in [18] and are therefore not available. The modality, *ALL* is the multimodal fusion of physiological measurements and facial landmarks. The higest F1 score for each combination of trait and modality is highlighted with bold.

From Table 1, we see that our models outperform the models created in [18], with the MLP performing best in the majority of cases. As hypothesized, multimodal features provide better predictions than unimodal, however, this is not always the case for RF.

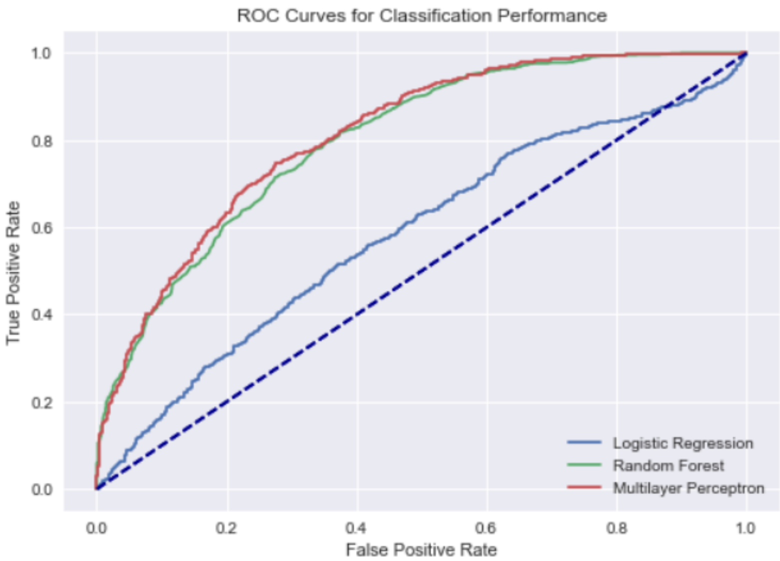

Figure 2: A ROC curve showing the overall classification performance of the different models.

The classification ability in form of a receiver operating characteristic figure (ROC) curve seen in Figure 2, shows how logisitc regression performs worse than the two other models, as expected. RF and MLP have similar performance, with the latter being slightly better at classifying.

From Table 1 we noticed, that the modalities ECG and EMO consistently have the highest F1 score, giving reason for creating a bimodal feature by concatenating these two and see if it would perform better than the combination of all features. For RF and MLP, the results were better on all traits except emotional stability. LR did not improve. The results can be seen in Table 2.

Classifier	Extra	Agree	Cons	Stab	Open
LR	0.71	0.68	0.66	0.60	0.61
RF	**0.82**	**0.75**	**0.73**	0.73	**0.72**
MLP	**0.87**	**0.84**	**0.81**	0.82	**0.81**

Table 2: The F1 score of each model when classifying traits using ECG and EMO combined. F1 scores higher than the *ALL* feature from Table 1 for each model is marked with bold.

5 Discussion

In this section, we discuss the findings and models established. We will look into the results and performance of the different models with respect to the initial hypothesis. Finally, we will provide suggestions for future research.

5.1 Interpretation of Results

State of the art methods of personality recognition vary depending on the area, e.g. in the visual modality CNNs generally yield the best results, but in bimodal approaches, deep bimodal regression currently gives the best results by utilizing Deep Residual Networks [13]. Examples of this approach is presented by Güclütürk et al. who achieved an accuracy score of 0.9109 [6], and Li et al. who achieved an accuracy score of 0.9188 [10].

Almost all of our full multimodal models outperformed the baseline at 0.5, as did all of our ECG/EMO combination methods. Several of the single physiological measurements scored below.

Unlike our expectations, our models outperformed the models from [18]. One explanaition for this may be because Subramanian et al. focused on creating the dataset and simply showcased that prediction above chance is possible. Had they yearned for high F1 scores, they would likely have used deep learning.

In line with our expectations, our model did not outperform the state of the art models, With a mean accuracy of 0.85 as our best result (Combination of ECG and EMO features using MLP). This is within expectations, as most state of the art uses deep networks, which we did not.

The results for ECG and EMO were consistently higher for all models than EEG and GSR. One reason for this may be because GSR and EEG have more noisy results with random spikes in the readings. EMO is directly controllable by the test subject and might be a better indicator for personality than GSR for a subject who tends to perspire a lot.

5.2 Strengths and Limitations of Approach

When comparing our three models, LR is outperformed by a large margin by RF and MLP. However, the performance of the two latter models are within the same range, making RF the optimal choice, when taking training time into account. Were the dataset larger, the MLP would likely outperform RF by a larger margin.

The architecture of the MLP is puzzling. It has one layer with 200 perceptrons. One would think that having hidden layers would yield better results, but it was not the case. A possible explanation for this is the lack of data, as a more complex architecture would require more data to be trained optimally.

Classifying personality using multimodality works significantly better than a unimodal approach, which is fitting the research mentioned in subsubsection 1.0.2. One can guess, that even though the classifiers have access to physiological readings (i.e. EEG), which humans usually do not, it is still limited how much this single modality can tell about a subject's personality. Like humans, our classification models benefit from having access to multiple modalities, which makes sense, given that human personality is expressed in more than one modality as discussed in subsubsection 1.0.1.

5.3 Future Research

As we have shown in this report, the combination of features and modalities to use in a personality recognition model, as well as what model is chosen, is just as important as what features to extract. As we have only looked at different combinations of modalities, more insight can almost certainly be gained by analysing the individual features themselves.

While we have in this report worked with a dataset were personality traits were defined by being above or below median [18, p. 7], this might not actually be all that meaningful from a psychological point of view [19, p. 16]. Vinciarelli and Mohammadi instead argues that ranking people in an experiment based on personality might have more merit [19, p. 16].

Lastly, the ASCERTAIN dataset is sparse in data. With only 58 test subjects, it is not enough data for deep learning, which could probably outperform all our models. Perhaps data augmentation is a possibility to expand on the quantity of the training data.

6 Conclusion

As we hypothesised, the multimodal approach utilizing all of the features, did outperform the two unimodal approaches, as well as any single single physiological measurement in almost all cases. However, we did find that a bimodal combination of face movement (EMO) and ECG data actually outperformed the fusion approach in many cases, with a top F1 score of 0.87 for extraversion using MLP (avg. score for MLP = 0.83). We assume that the reason for this, is due to GSR and EEG not contribution to the classification, but instead introduce noise to the data, resulting in worse performance. More research into the individual features of each combination might yield even better results, as may other classification models.

7 Appendix

The implementation of the models and the features used in this project are available on GitHub. The seed for the random number generation is set to 42 for reproducibility.
URL: `https://github.com/kroglkt/CogSci2`

7.1 Training Time and Metrics

Model	Training time in minutes	Mean F1	Std F1	Mean Accuracy	Std Accuracy
Logistic Regression	3	0.58	0.10	0.64	0.06
Random Forest	13	0.62	0.10	0.69	0.06
Multilayer Perceptron	82	0.69	0.11	0.71	0.10

Table 3: The training time and performance metrics for the three models.

7.2 PCA Analysis

Investigating the trait/modality combinations, some are generally lower than others, independently of the model. To gain insight in this, we performed PCA with two components to create the following figure.

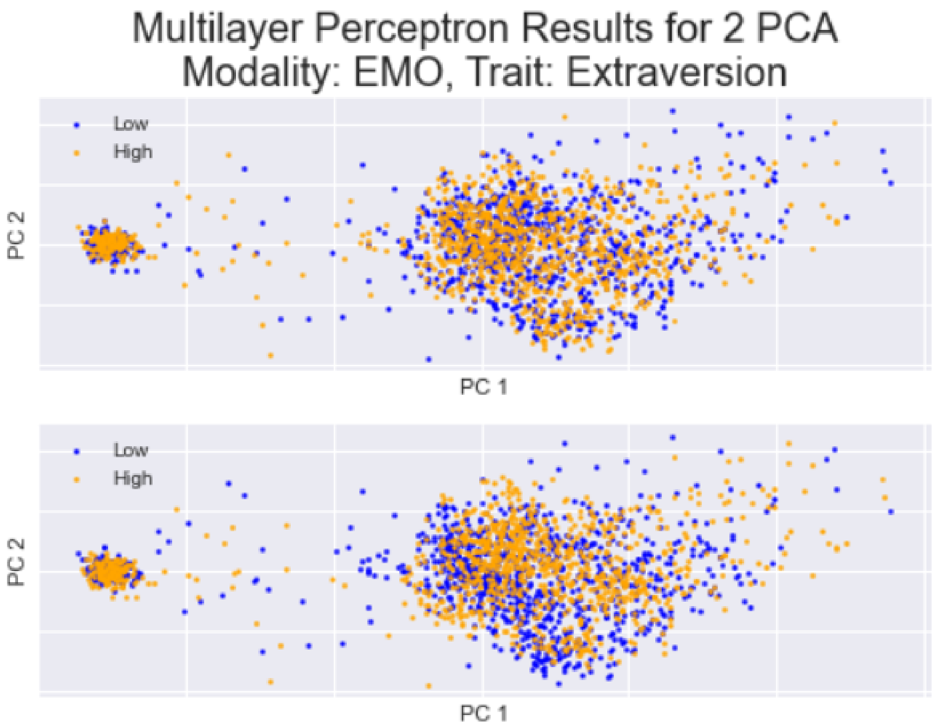

Figure 3: Two dimensional PCA scatter plots showing the actual (top) and predicted (bottom) classifications from the MLP on the extraversion trait.

Figure 3 shows the first two principal components with colours indicating high and low trait. The top subplot shows the target values, while the bottom subplot shows the predicted values for the model. Multiple figures for different combinations were made, but did not provide great insight in the data, as there would be no clear grouping of high and low traits.

References

[1] Big five personality test. openpsychometrics.org, Date accesed: 2021-05-28. https://openpsychometrics.org/tests/IPIP-BFFM/.

[2] G. An and R. Levitan. Lexical and acoustic deep learning model for personality recognition. *Proceedings of interspeech*, page 1761–1765, 2018.

[3] E. Brunswik. *Perception and the Representative Design of Psychological Experiments*. University of California Press, 1956.

[4] L. R. Goldberg. The structure of phenotypic personality traits. *American Psychologist*, (48):26–34, 1993.

[5] H. Gunes, O. Celiktutan, and E. Sariyanidi. Live human–robot interactive public demonstrations with automatic emotion and personality prediction. *Philosophical Transactions of the Royal Society B: Biological Sciences*, (374):20180026, April 2019.

[6] Y. Güçlütürk, U. Güçlü, M. A. J. van Gerven, and R. van Lier. Deep impression: Audiovisual deep residual networks for multimodal apparent personality trait recognition. *ECCVChaLearn workshop*, pages 349–358, 2016.

[7] C. G. Jung. *Psychological Types*, volume 6 of *The Collected Works of C. G. Jung*. Princeton University Press, 1971.

[8] A. A. Kindiroglu, L. Akarun, and O. Aran. Multi-domain and multi-task prediction of extraversion and leadership from meeting videos. *EURASIP Journal on Image and Video Processing*, (1):1–5, December 2017.

[9] S. Koelstra and I. Patras. Fusion of facial expressions and eeg for implicit affective tagging,. *Image and Vision Computing*, (31):164–174, February 2013.

[10] Y. Li, J. Wan, Q. Miao, S. Escalera, H. Fang, H. Chen, X. Qi, and G. Guo. Cr-net: A deep classification-regression network for multimodal apparent personality analysis. *International Journal of Computer Vision*, (128):pages 2763–2780, 2020.

[11] C. C. S. Liem, M. Langer, A. Demetriou, A. M. F. Hiemstra, A. S. Wicaksana, M. P. Born, and C. J. König. *Psychology Meets Machine Learning: Interdisciplinary Perspectives on Algorithmic Job Candidate Screening. In: Escalante H. et al. (eds) Explainable and Interpretable Models in Computer Vision and Machine Learning.* Springer, 2018.

[12] W. M. Marston. *Emotions of Normal People.* Kegan Paul, Trench, Trubner & Co. Ltd., 1928.

[13] Y. Mehta, N. Majumder, A. Gelbukh, and E. Cambria. Recent trends in deep learning based personality detection. *Artificial Intelligence Review*, (53):2313–2339, April 2020.

[14] I. B. Myers and P. B. Myers. *Gifts Differing: Understanding Personality Type.* Davies-Black Publishing, 1980.

[15] S. Oviatt, B. Schuller, P. R. Cohen, D. Sonntag, G. Potamianos, and A. Krüger. *The Handbook of Multimodal-Multisensor Interfaces: Foundations, User Modeling, and Common Modality Combinations*, volume 1. ACM, 2017.

[16] D. H. Saklofske, H. J. Eysenck, S. G. B. Eysenck, R. M. Stelmack, and W. Revelle. *Extraversion-Introversion. In: Encyclopedia of Human Behavior.* Academic Press, second edition, 2012.

[17] B. E. Stein. *The New Handbook of Multisensory Processing.* MIT Press, 2nd edition, 2012.

[18] R. Subramanian, J. Wache, M. K. Abadi, R. L. Vieriu, S. Winkler, and N. Sebe. Ascertain: Emotion and personalityrecognition using commercial sensors. *IEEE Transactions on Affective Computing*, (2):147–160, April-June 2018.

[19] A. Vinciarelli and G. Mohammadi. A survey of personality computing. *IEEE Transactions on Affective Computing*, (5):1–20, December 2014.

[20] H.-C. Yang and Z.-R. Huang. Mining personality traits from social messages for game recommender systems. *Knowledge-Based Systems*, (165):157–168, February 2019.

[21] C.-L. Zhang, H. Zhang, X.-S. Wei, and J. Wu. Deep bimodal regressionfor apparent personality analysis. *European conference on computer vision*, page 311–324, 2016.