Predicting Personality through Multimodal Signals

Omar Kassem

University of Freiburg
Prof. Dr. Hannah Bast

October 7, 2019

Outline

Motivation

Computer understanding of human behaviour:

- Enhancing human-computer interaction, e.g. conversations.
- Augmenting interactions between humans, e.g. sales.

Problem statement

Exploring the ability of deep learning approaches to predict the personality traits of people based on data provided by them.

- Text
- Audio
- Images

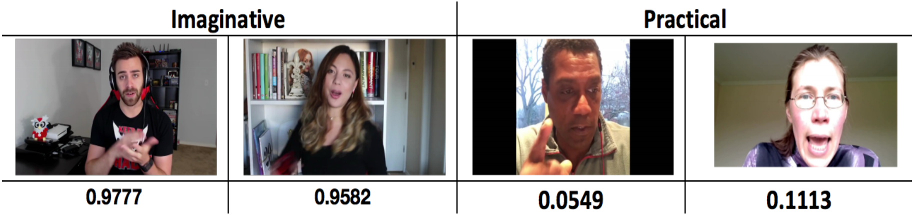
How to define personality

Big five model (OCEAN):

- Openness
- Conscientiousness
- Extraversion
- Agreeableness
- Neuroticism

Openness

Creativity and imagination versus being predictable and straight forward.



Correlation with used words (Yarkoni 2010):

- Positive: culture, films and poetry
- Negative: anniversary, diaper and hubby

ChaLearn LAP 2016: First Round Challenge on First Impressions - Dataset and Results, Victor et al.

Conscientiousness

A measure of self-discipline and dedication, a sign of long term success.

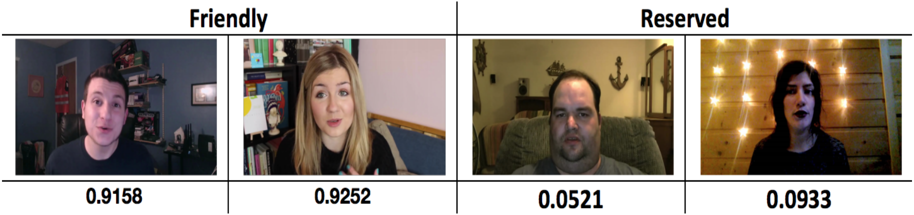
Correlation with used words (Yarkoni 2010):

- Positive: achieved, discipline and persistence
- Negative: boring, drunk and deny

ChaLearn LAP 2016: First Round Challenge on First Impressions - Dataset and Results, Victor et al.

Extraversion

Willingness to share emotions with others and enjoying their company.



Correlation with used words (Yarkoni 2010):

- Positive: bar, concert and friends
- Negative: books, computer and winter

ChaLearn LAP 2016: First Round Challenge on First Impressions - Dataset and Results, Victor et al.

Agreeableness

Understanding or helpful versus selfish and cautious with other people.

Correlation with used words (Yarkoni 2010):

- Positive: gifts, together and joy
- Negative: a**hole, harm and violence

ChaLearn LAP 2016: First Round Challenge on First Impressions - Dataset and Results, Victor et al.

Stability (Neuroticism)

Emotional stability versus vulnerability to displeasing emotions.

Correlation with used words (Yarkoni 2010):

- Positive: poem, mountain and sunset
- Negative: awful, ashamed and stress

ChaLearn LAP 2016: First Round Challenge on First Impressions - Dataset and Results, Victor et al.

Outline

First Impressions dataset

Predicting personality traits from short videos.

- 15 sec video clips
- Collected from Youtube vloggers
- 10,000 clips (total of 41 hours)
- Regression labels in $[0, 1]$

http://chalearnlap.cvc.uab.es/dataset/20/description/

First Impressions competition

Team	AVG	E	A	C	N	O
NJU	**91.3**	91.3	**91.3**	**91.7**	**91.0**	**91.2**
evolgen	91.2	**91.5**	91.2	91.2	91.0	91.2
DCC	91.1	91.1	91.0	91.4	90.9	91.1

Figure: Percentages of mean absolute accuracy by top-3 competitors

$$MAA = 1 - MAE = 1 - \frac{1}{n}\sum^{n}|Y_{pred} - Y_{true}|$$

Outline

Text model

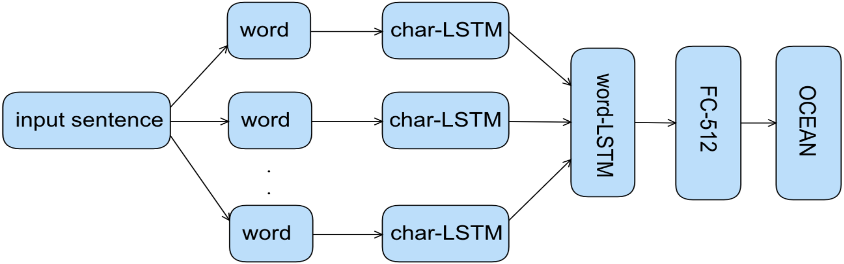


Figure: Character-based model for text modality

Audio model

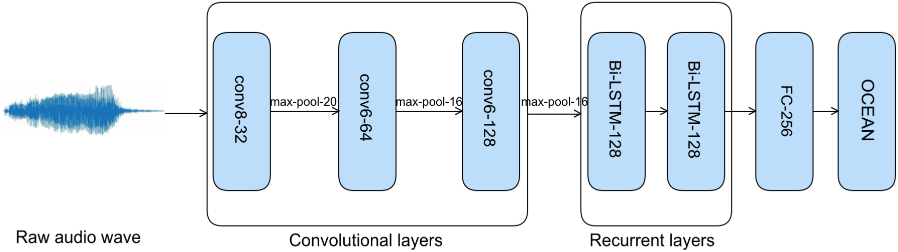


Figure: End-to-End audio model for audio modality

Image model

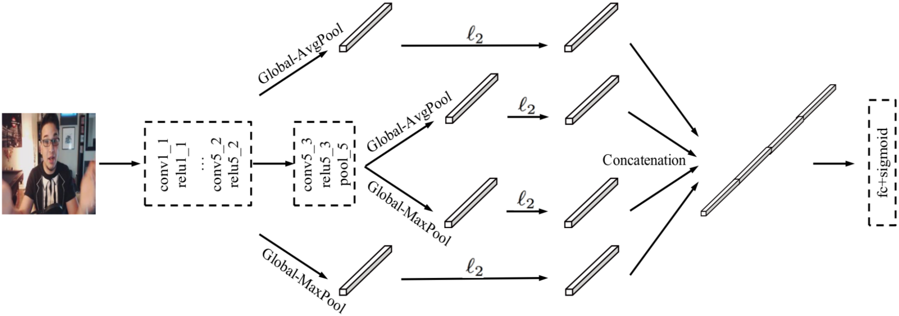


Figure: DAN+ model for video modality

Deep Bimodal Regression of Apparent Personality Traits from Short Video Sequences, Wei et al.

Image model

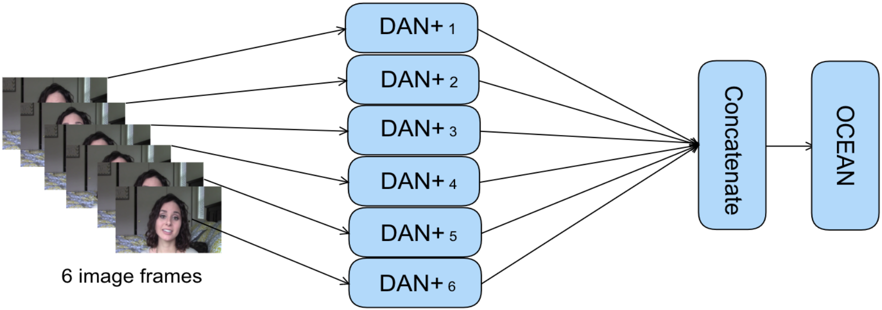


Figure: DAN+V model for video modality

Multimodal models

Fusion techniques:

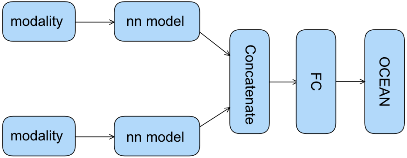


Figure: Early fusion

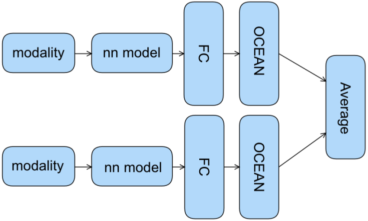


Figure: Late fusion (ensemble)

Outline

MAA results

Model	AVG	E	A	C	N	O
Baseline	88.4	88.1	89.9	87.4	87.9	88.5
Video	**91.5**	**91.7**	**91.3**	**92.2**	**91.1**	**91.5**
DAN+ (NJU)	91.1	–	–	–	–	–
ResNet (NJU)	91.0	–	–	–	–	–
Audio	**89.7**	**89.6**	**90.3**	**89.0**	**89.8**	**90.1**
EtE. (NJU)	89.5	–	–	–	–	–
Lin. Reg. (NJU)	89.0	–	–	–	–	–
Text	**88.8**	**88.3**	**89.9**	**88.4**	**88.6**	**88.8**

Table: Percentages of models' accuracies that use one modality

Multimodal MAA results

Model	AVG	E	A	C	N	O
Baseline	88.4	88.1	89.9	87.4	87.9	88.5
Audio, Text, Video(E.F.)	91.3	91.4	91.1	91.9	91.0	91.1
NJU follow-up	**92.1**	–	–	–	–	–
Audio, Video(E.F.)	91.5	**91.5**	**91.3**	**91.9**	**91.3**	**91.3**
Audio, Video(L.F.)	91.4	91.4	91.3	91.5	91.1	91.5
NJU	91.3	91.3	91.3	91.7	91.0	91.2
evolgen	91.2	91.5	91.2	91.2	91.0	91.2
Video, Text(E.F.)	91.2	91.2	90.8	91.8	90.9	91.2
Audio, Text(E.F.)	89.8	89.6	90.3	89.2	89.8	90.0

Table: Percentages of models' accuracies that use multiple modalities

Data distribution

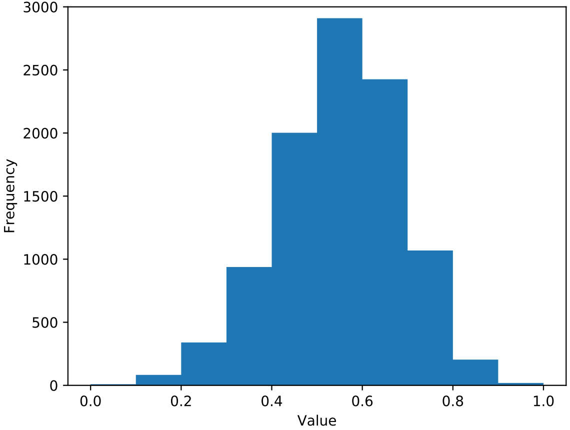


Figure: Agreeableness data distribution

Receiver Operating Characteristic curve

Area Under ROC Curve (AUC).

- Classification metric
- Plots TPR against FPR
- Distinguishing between positive and negative examples

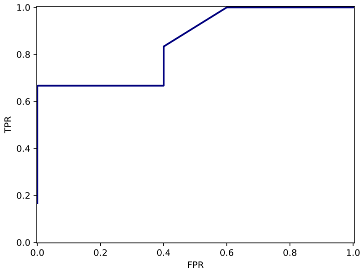


Figure: ROC curve example

AUC Results

Model	AVG	E	A	C	N	O
Baseline	51.0	49.9	51.3	51.6	50.1	51.9
Video	**82.7**	**84.2**	**76.5**	**88.0**	**82.5**	**82.4**
Audio	74.2	74.9	71.3	72.5	76.4	75.9
Text	65.3	63.0	65.3	67.5	66.9	64.0

Table: AUC percentages of models that use one modality

Multimodal AUC results

Model	AVG	E	A	C	N	O
Baseline	51.0	49.9	51.3	51.6	50.1	51.9
Video, Audio, Text(L.F.)	83.5	84.2	79.0	87.1	84.1	83.4
Video, Audio(L.F.)	**84.1**	**84.8**	**79.1**	**88.0**	**84.5**	**83.9**
Video, Audio(E.F.)	83.2	84.1	77.9	87.4	83.9	82.5
NJU	82.3	83.9	76.3	87.0	82.0	82.2
evolgen	82.1	83.8	77.7	84.9	82.6	81.4
Video, Text(L.F.)	83.2	84.4	78.0	87.7	83.4	82.7
Audio, Text(L.F.)	74.0	74.2	71.3	73.5	76.0	75.2

Table: AUC percentages of models that use multiple modalities

Results overview

Model	AUC	MAA
Baseline	51.0	88.4
Video	82.7	**91.5**
Video, Audio, Text(E.F.)	81.7	91.3
Video, Audio, Text(L.F.)	83.5	90.8
Video, Audio(E.F.)	83.2	**91.5**
Video, Audio(L.F.)	**84.1**	91.4
Video, Text(E.F.)	81.5	91.2
Video, Text(L.F.)	83.2	90.9
Audio, Text(E.F.)	74.1	89.8
Audio, Text(L.F.)	74.0	89.6

Table: A comparison between AUC and MAA performances

Conclusion

- Automatic personality prediction from different data types
- Best MAA performance for single modality models
- Best AUC performance
- Comparing different fusion techniques for the same models

Demo

Questions?

Backup slides: AUC validation

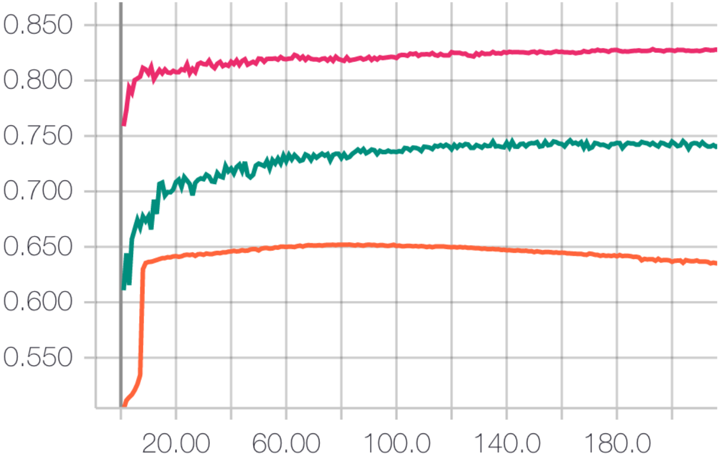


Figure: Validation AUC performance through training time

Backup slides: MAA train

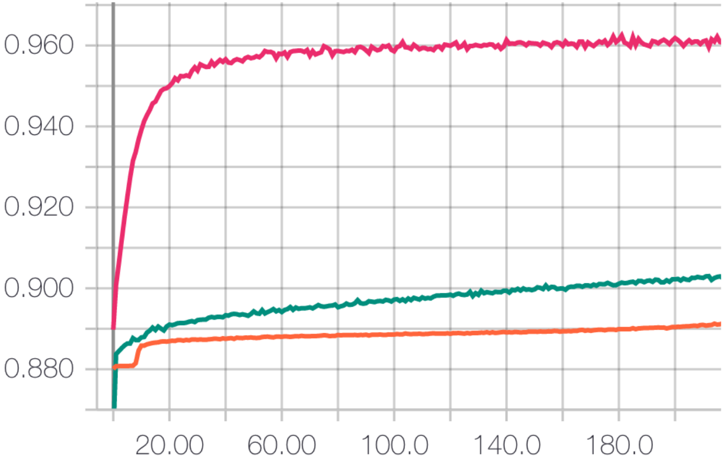


Figure: Train MAA performance through training time

Backup slides: MAA validation

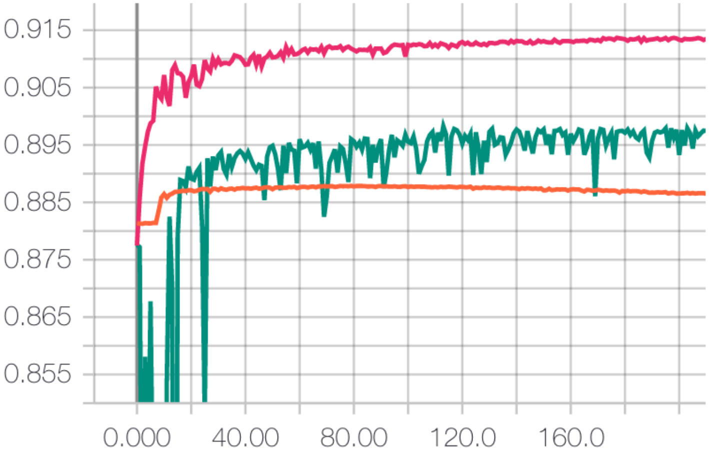


Figure: Validation MAA performance through training time

Backup slides: Heat maps

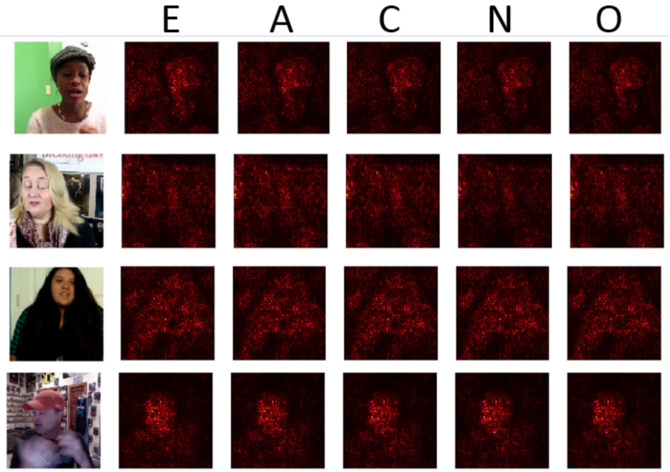


Figure: Saliency heat maps of a ResNet that predict OCEAN

Prediction of Personality First Impressions With Deep Bimodal LSTM. Yang et al.

Backup slides: Heat maps

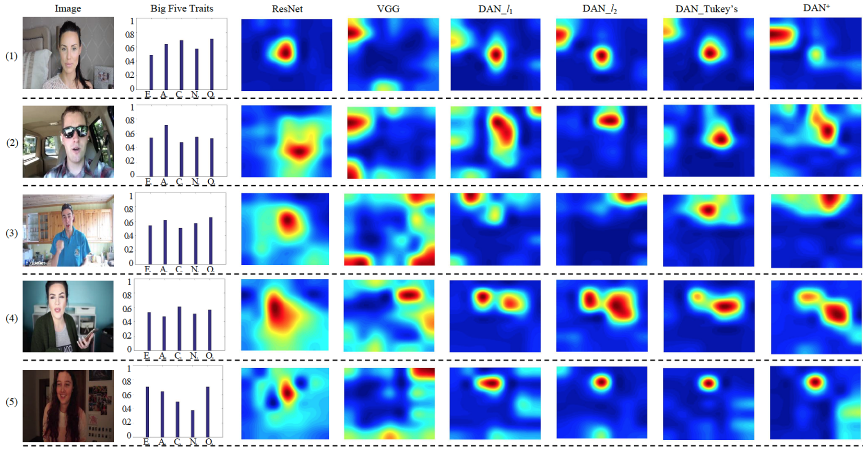


Figure: Saliency heat maps of different CNNs that predict OCEAN

Deep Bimodal Regression of Apparent Personality Traits from Short Video Sequences. Wei et al. 2017

Backup slides: Personality questionnaire

Instruction: How well do the following statements describe your personality?

I see myself as someone who …	Disagree strongly	Disagree a little	Neither agree nor disagree	Agree a little	Agree strongly
… is reserved	(1)	(2)	(3)	(4)	(5)
… is generally trusting	(1)	(2)	(3)	(4)	(5)
… tends to be lazy	(1)	(2)	(3)	(4)	(5)
… is relaxed, handles stress well	(1)	(2)	(3)	(4)	(5)
… has few artistic interests	(1)	(2)	(3)	(4)	(5)
… is outgoing, sociable	(1)	(2)	(3)	(4)	(5)
… tends to find fault with others	(1)	(2)	(3)	(4)	(5)
… does a thorough job	(1)	(2)	(3)	(4)	(5)
… gets nervous easily	(1)	(2)	(3)	(4)	(5)
… has an active imagination	(1)	(2)	(3)	(4)	(5)

Figure: BFI-10

Measuring personality in one minute or less. Rammstedt et al.(2007)

Backup slides: Image model

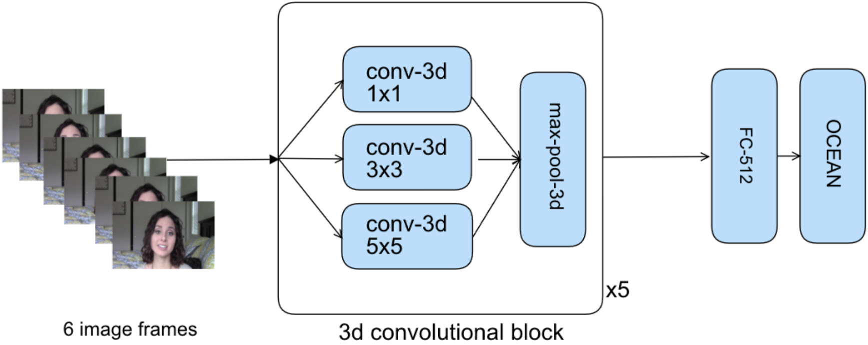


Figure: Conv-3d model

Backup slides: Text model

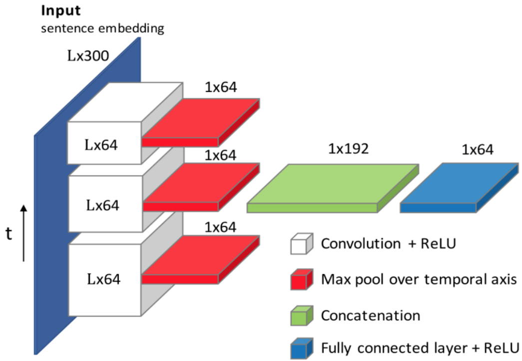


Figure: Word based model

Investigating Audio, Video, and Text Fusion Methods for End-to-End Automatic Personality Prediction. Kampman et al.

Backup slides: AUC visualization

Visual demo of AUC: `http://www.navan.name/roc/`

$$TPR = \frac{TP}{TP + FN}$$

$$FPR = \frac{FP}{FP + TN}$$

Backup slides: FI data distribution

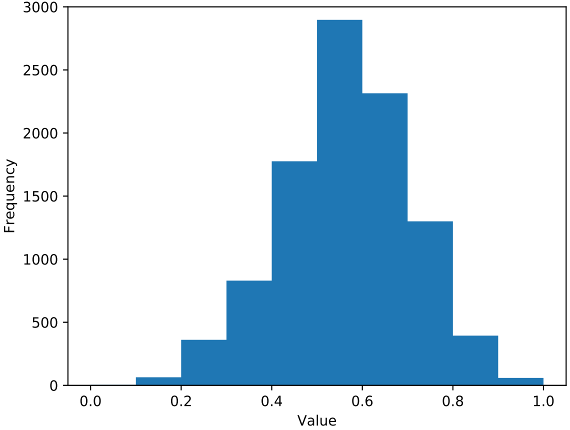


Figure: Openness data distribution of First Impressions dataset

Backup slides: FI data distribution

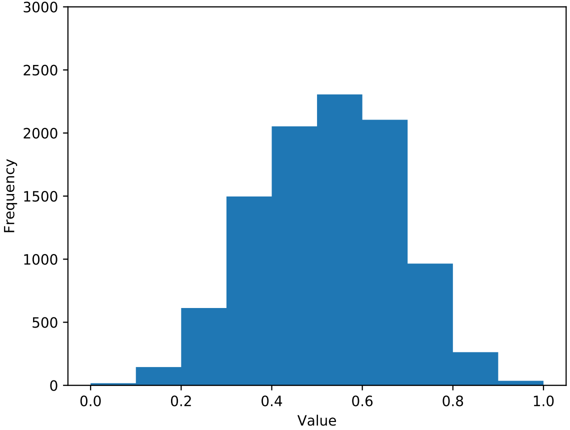


Figure: Conscientiousness data distribution of First Impressions dataset

Backup slides: FI data distribution

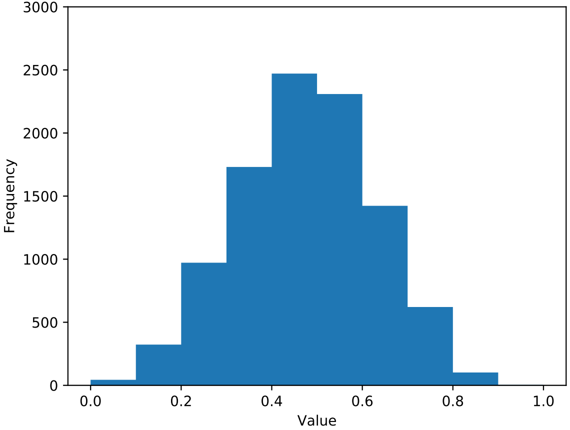


Figure: Extraversion data distribution of First Impressions dataset

Backup slides: FI data distribution

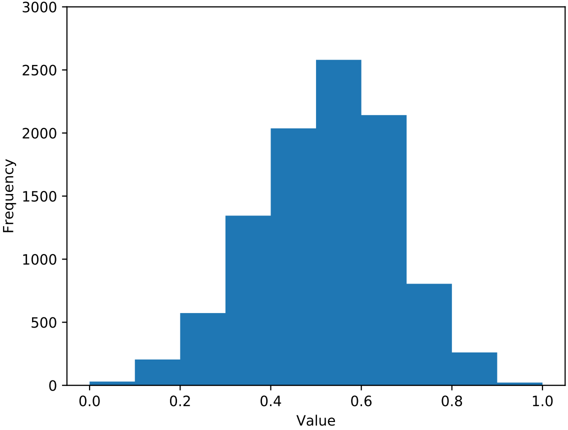


Figure: Neuroticism data distribution of First Impressions dataset

Backup slides: Labels construction

First Impressions labels:

- Collected by Amazon Mechanical Turk (AMT)
- Uses pairwise comparisons
- The labels are output of a Bradley Terry Luce (BTL) model
- Fits a maximum likelihood
- Sigmoid output layer

Backup slides: Pairwise comparisons

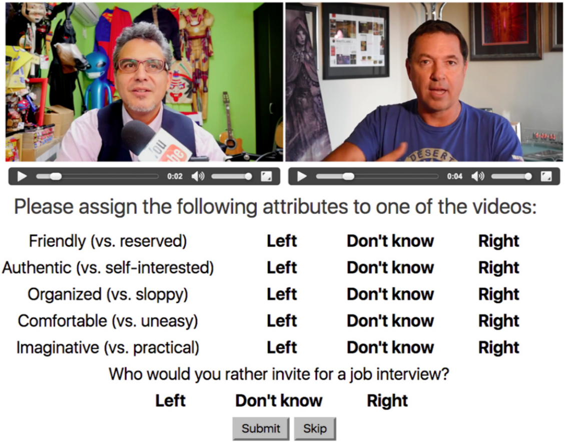


Figure: Data labeling

ChaLearn LAP 2016: First Round Challenge on First Impressions - Dataset and Results, Victor et al.

Backup slides: PAN dataset

PAN dataset:

- Text dataset from twitter.
- A tweet is maximum 140 characters.
- The dataset has more than 28,000 tweet.

Backup slides: PAN data distribution

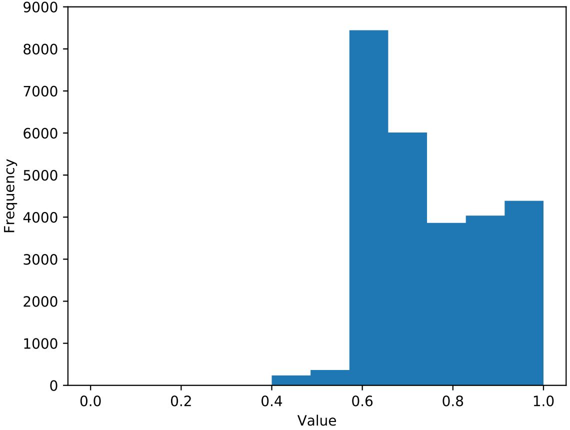


Figure: Openness data distribution of the PAN dataset

Backup slides: PAN data distribution

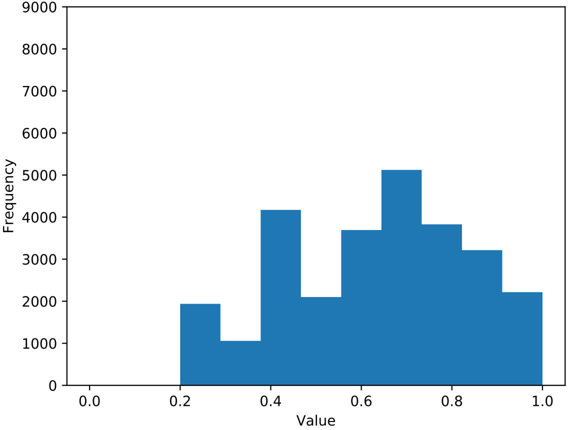


Figure: Conscientiousness data distribution of the PAN dataset

Backup slides: PAN data distribution

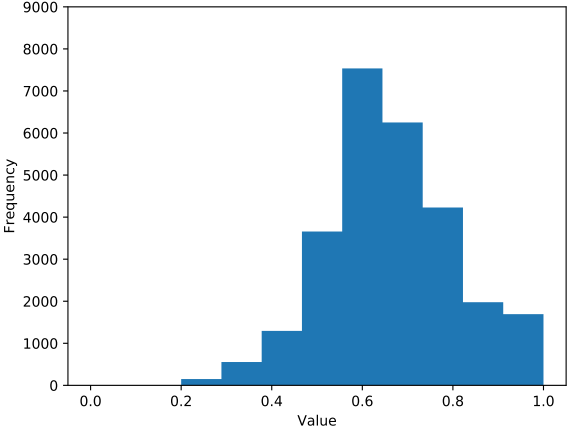


Figure: Extraversion data distribution of the PAN dataset

Backup slides: PAN data distribution

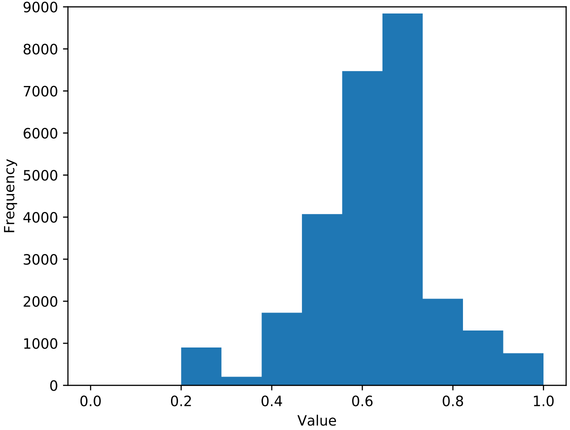


Figure: Agreeableness data distribution of the PAN dataset

Backup slides: PAN data distribution

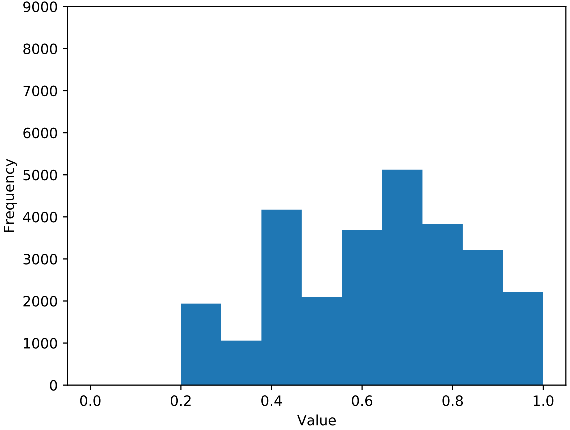


Figure: Neuroticism data distribution of the PAN dataset