– Seminar –
Cognitive Reasoning Seminar

Final Meeting

apl. Prof. Dr. Dr. Marco Ragni
Nicolas Riesterer, Daniel Brand

December 20th, 2019

Cognitive Computation Lab,
Department of Computer Science,
University of Freiburg

Task

Given known real data of human syllogistic reasoning and artificial datasets, label unknown datasets as real or artificial.

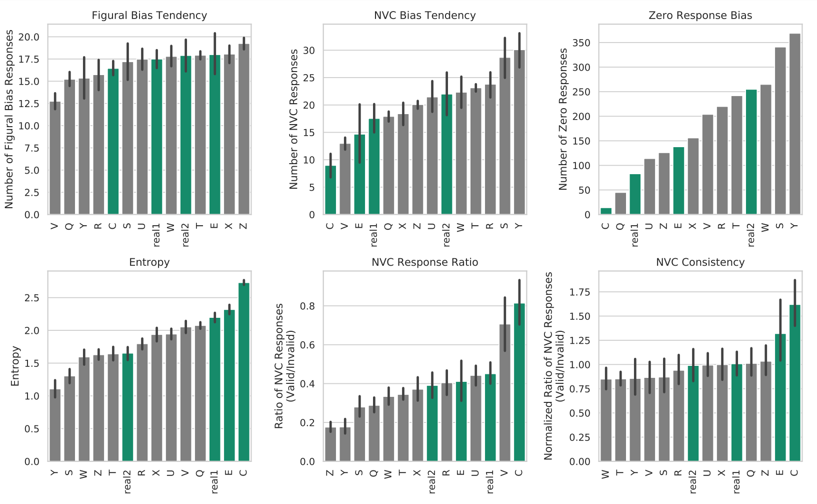
New Datasets

Dataset	Real	Description
Q	No	Cognitive model TransSet
R	No	mReasoner and PHM with noise
S	Yes	20 Students
T	No	Sampling from distribution (real2)
U	Yes	Combination of real1, real2 and E
V	No	Same entropy as real1
W	No	Voting like F, based on real2
X	No	Voting like F, mixed with old F
Y	Yes	Re-test of S, same students
Z	No	Most frequent answer with noise

Aggregate Analysis Methods

Method	Description
Figural Bias Tendency	Number of fig. bias responses (1: ac, 2: ca)
NVC Bias Tendency	Number of NVC responses
Zero Response Bias	Number of responses not given by participants
Entropy	Average syllogism-based entropy
NVC Response Ratio	Ratio of NVC responses (valid/invalid)
NVC Consistency	Normalized NVC Response Ratio

Aggregate Analysis



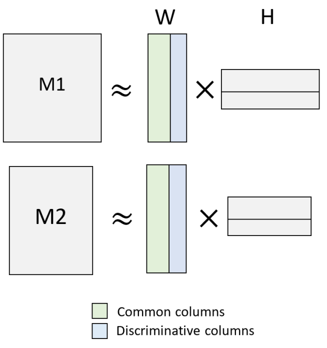
Aggregate Results

Method	Q	R	S	T	U	V	W	X	Y	Z
Figural Bias Tendency						f				f
NVC Bias Tendency			f						f	
Zero Response Bias			f						f	
Entropy			f						f	
NVC Response Ratio						?			f	f
NVC Consistency										
Prediction	r	r	f	r	r	?	r	r	f	f
Ground Truth	f	f	r	f	r	f	f	f	r	f

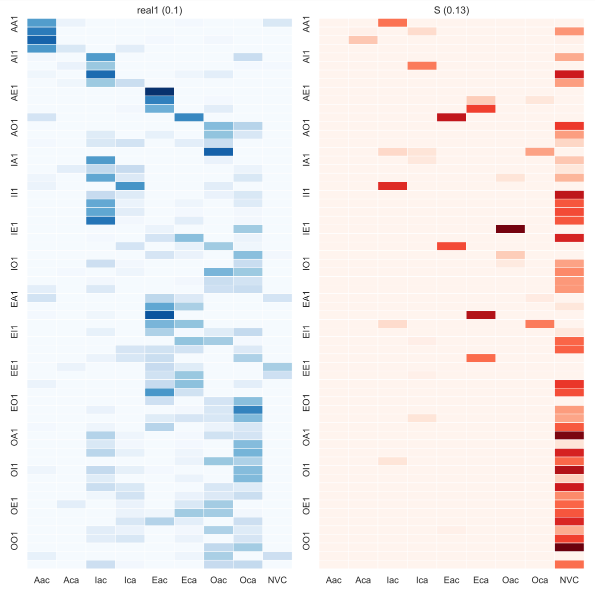
General aggregate analysis fails for these datasets.

Analysis based on NMF

- Real data, if sum of H-matrices about 0.1
- Note: S, V and Y are very small
- Decision based on maximum total H-values and the maximum H-values of the real datasets (to account for small datasets)
- We compared to real1, real2 and E



Example: real1 and S



NMF predictions

Method	Q	R	S	T	U	V	W	X	Y	Z
max	0.16	0.17	0.16	0.17	0.1	0.3	0.13	0.12	0.2	0.51
max_{real}	0.16	0.17	0.13	0.11	0.07	0.29	0.11	0.1	0.12	0.16
Prediction	f	f			r	f	r	r		f
Truth	f	f	r	f	r	f	f	f	r	f

Problems with NMF

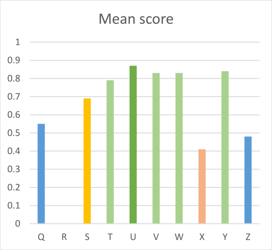
- Some results are hard to interpret
- Small datasets hurt the approach
- We know, that voting/distribution based data can't be detected

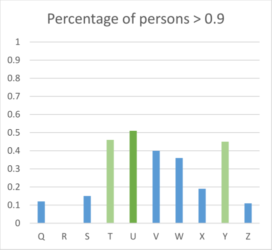
Aggregate analyses don't really work.
We need to shift the focus to individuals.

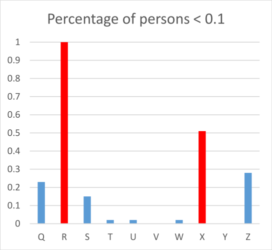
Classification with Neural Networks

- We have labels for artificial and real datasets
- Train a classifier!
- Neural network trained on single persons (576 vector)
- Additional random users are added to have more training data for artificial data
- To rate a dataset, each person is classified
- Final decision based on mean results

Results with Classifier







Results with Classifier

Method	Q	R	S	T	U	V	W	X	Y	Z
Mean Score		f		r	r	r	r	f	r	
Percentage of persons > 0.9		f		r	r				r	
Percentage of persons < 0.1		f				r		f	r	
Prediction		f		r	r	r		f	r	
Ground Truth	f	f	r	f	r	f	f	f	r	f

Concluding Observations

1. Different methods focus on different aspects of the data leading to different results
2. Identifying fakes is difficult if features of the data are highly independent
3. Algorithmic process-based fakes are easier to detect than stochastic fakes

Take-Home Message

- Quality of data-driven ML methods usually demonstrated on feature-rich domains
  - Large databases
  - Highly dependent inputs (e.g., pixel data)
- However, in practical applications often expected to perform similarly in less accessible domains
- Often unclear if learned patterns actually represent the real dependencies of the processes underlying the data
- If properties of the predictive approach are known compromising results is incredibly easy by injecting artificially generated data points.