Naïve Bayes as a Default in Human Category Learning

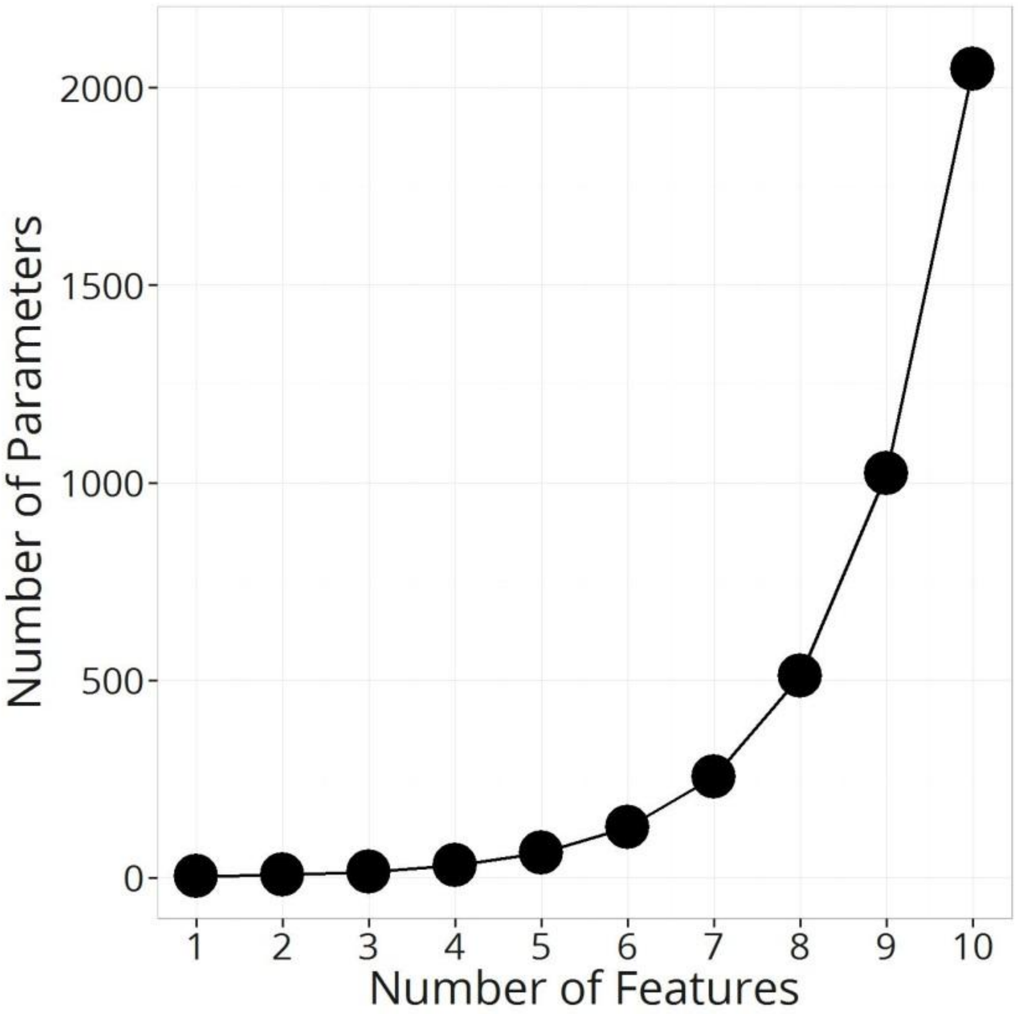
Jana B. Jarecki | Björn Meder | Jonathan D. Nelson
Center for Adaptive Behavior and Cognition (ABC)
Max Planck Institute for Human Development | Berlin

TeaP 2015, Hildesheim, Germany

The challenge

Taking a probabilistic viewpoint on categorization poses a computational challenge.

Combinatorial explosion



This challenge…

….affects cognitive categorization models

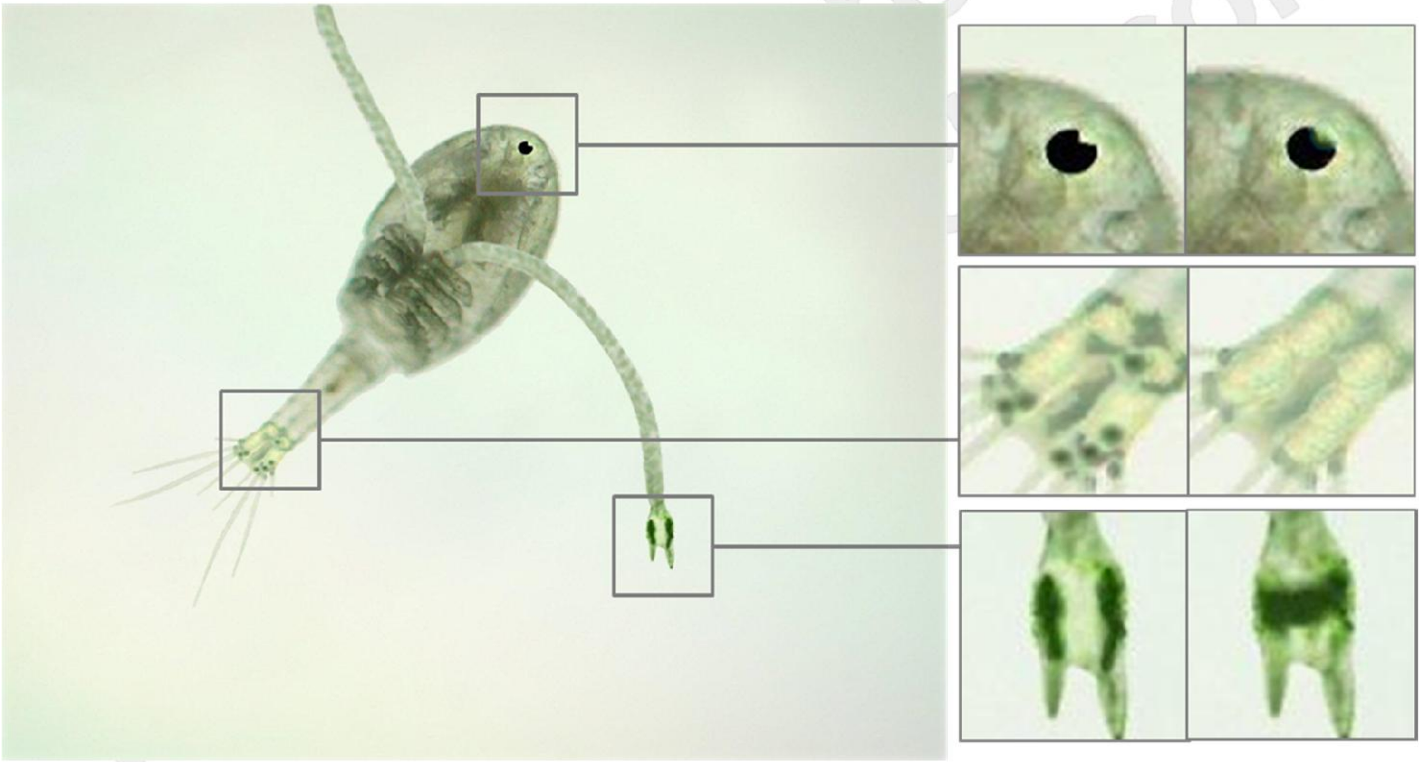
- Exemplar models (Medin & Schaffer, 1978; Nosofsky, 1986)
- Rule-based models (Ashby et al. 1998)

The challenge

This computational challenge is researched in Machine Learning.

Naïve Bayes

Species A or Species B?

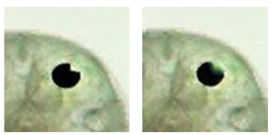


Naïve Bayes

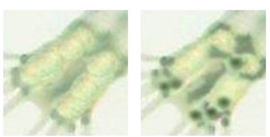
Classification algorithm assuming

class-conditional independence

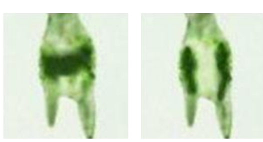
Species A



independent of



independent of

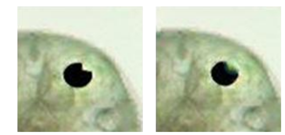


Species A

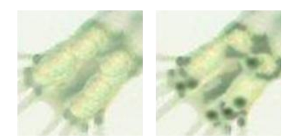
$$P(\boldsymbol{f} \mid c) = \prod_{d=1}^{D} P(f_d \mid c)$$

Species A

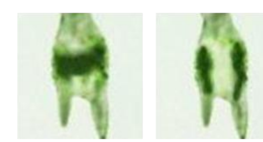
Species B



independent of

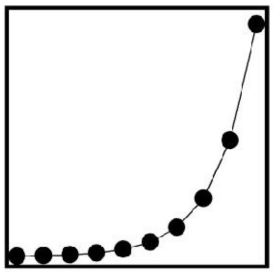


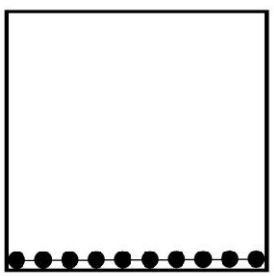
independent of



Naïve Bayes

Complexity Reduction



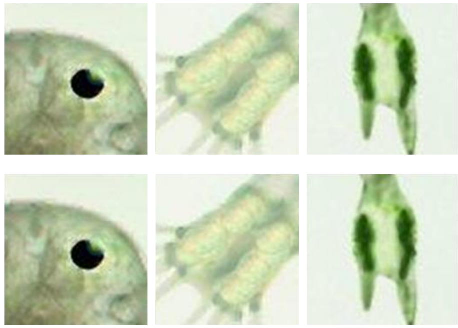


Robust

against violations of conditional independence

(Domingos & Pazzani, 1997; Rish et al., 2001)

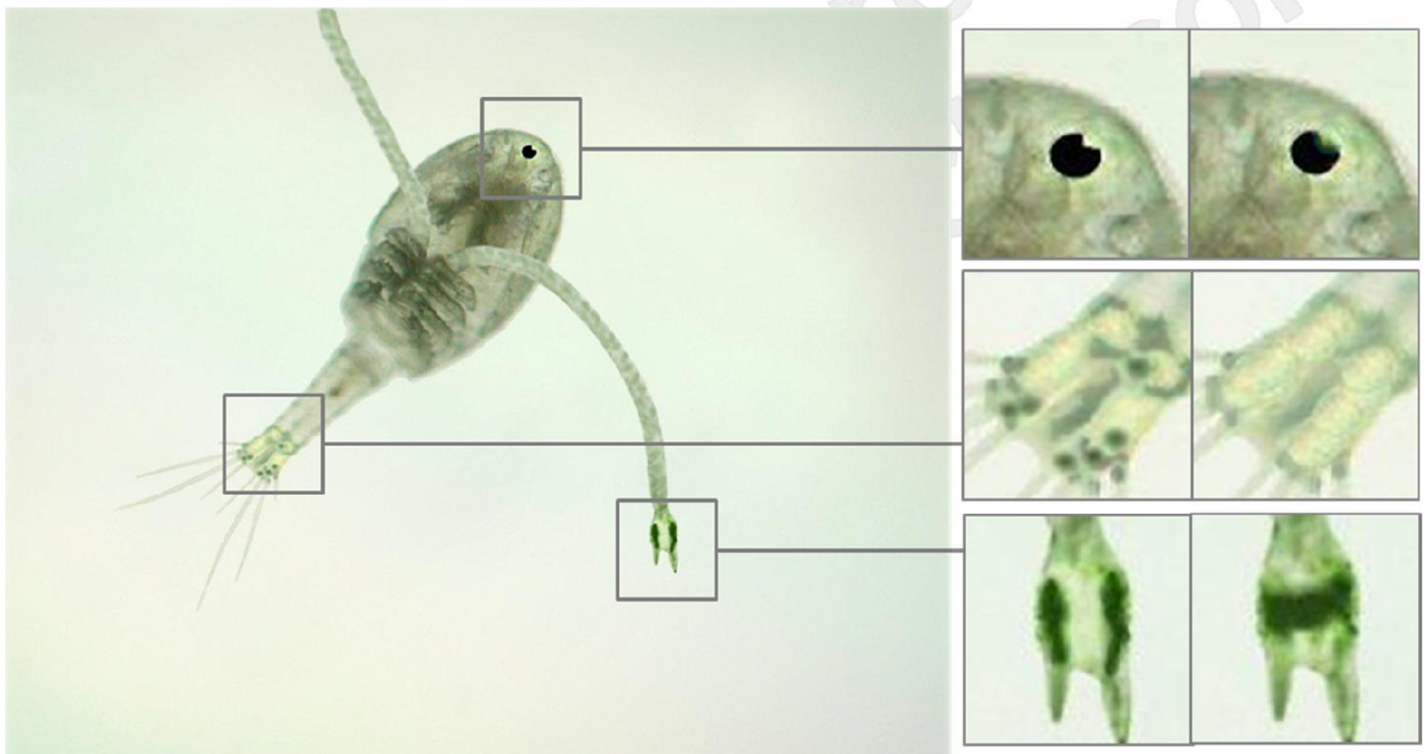
Extrapolate



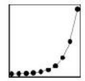
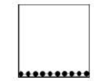
Naïve Bayes
as a Default in Human Category Learning

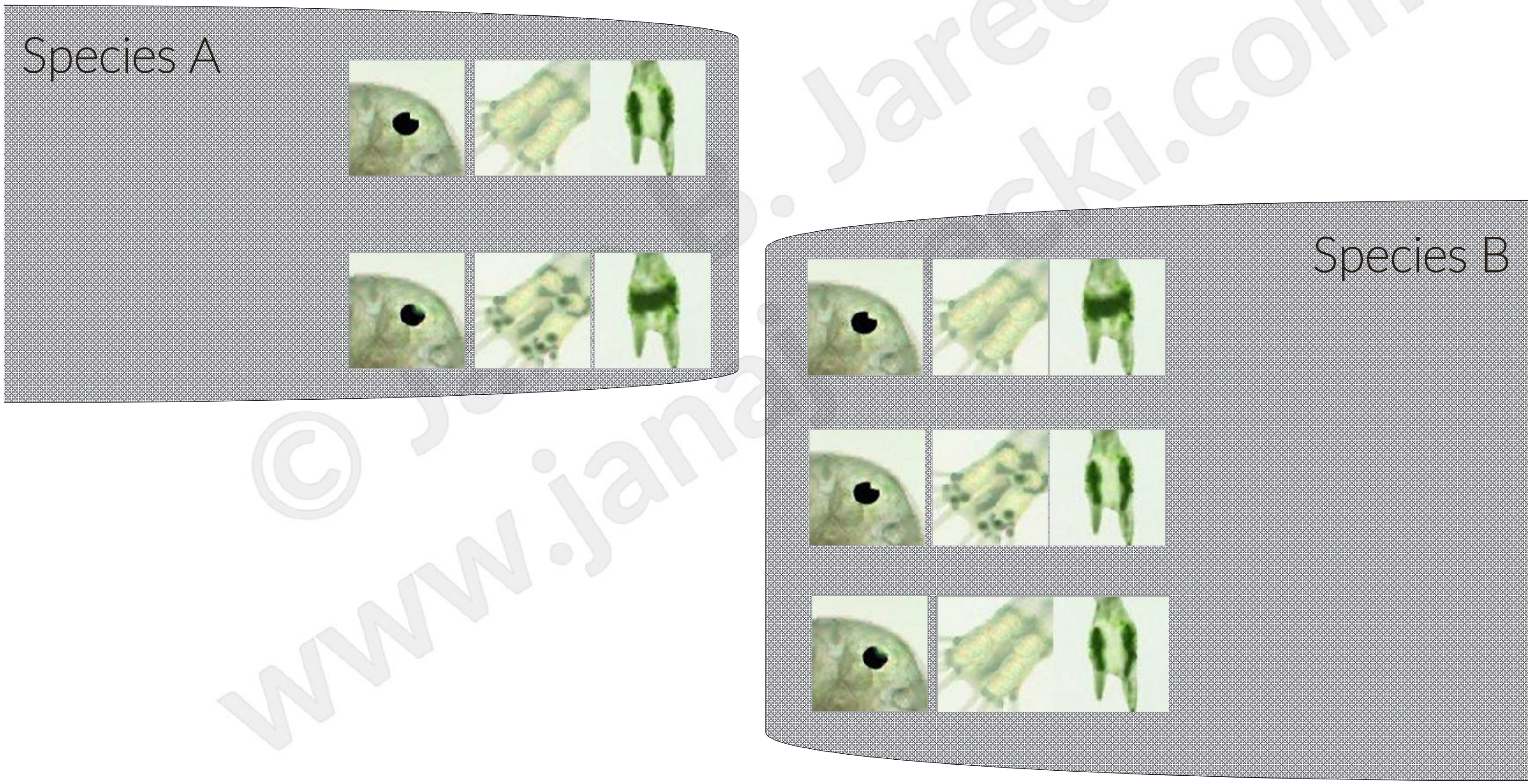
Naïve Bayes

Species A or Species B?



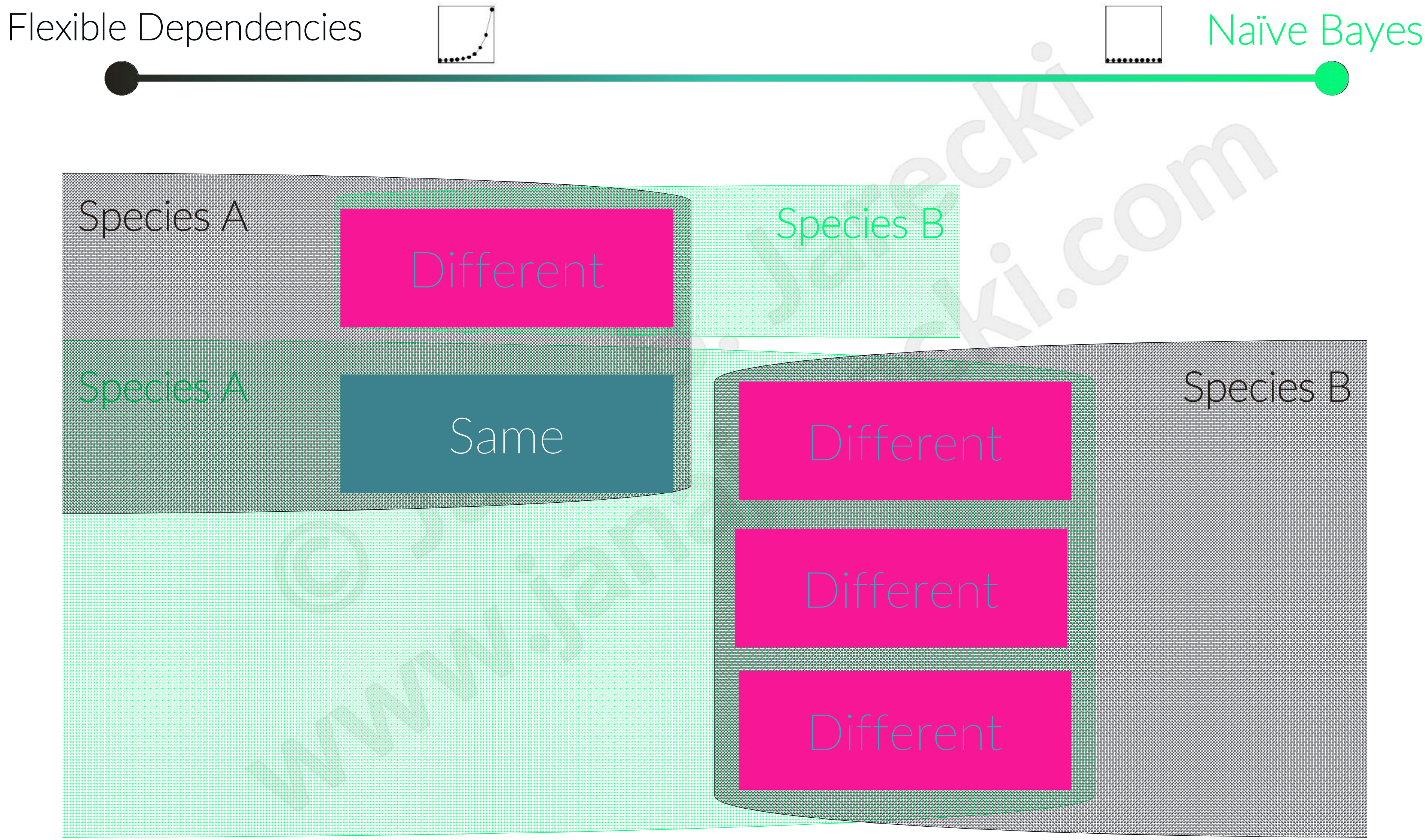
Optimal Experimental Design

Flexible Dependencies   Naïve Bayes

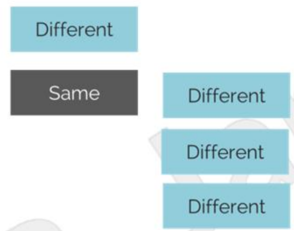


Optimal Experimental Design

Optimal Experimental Design

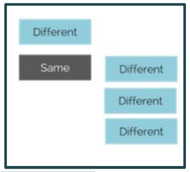


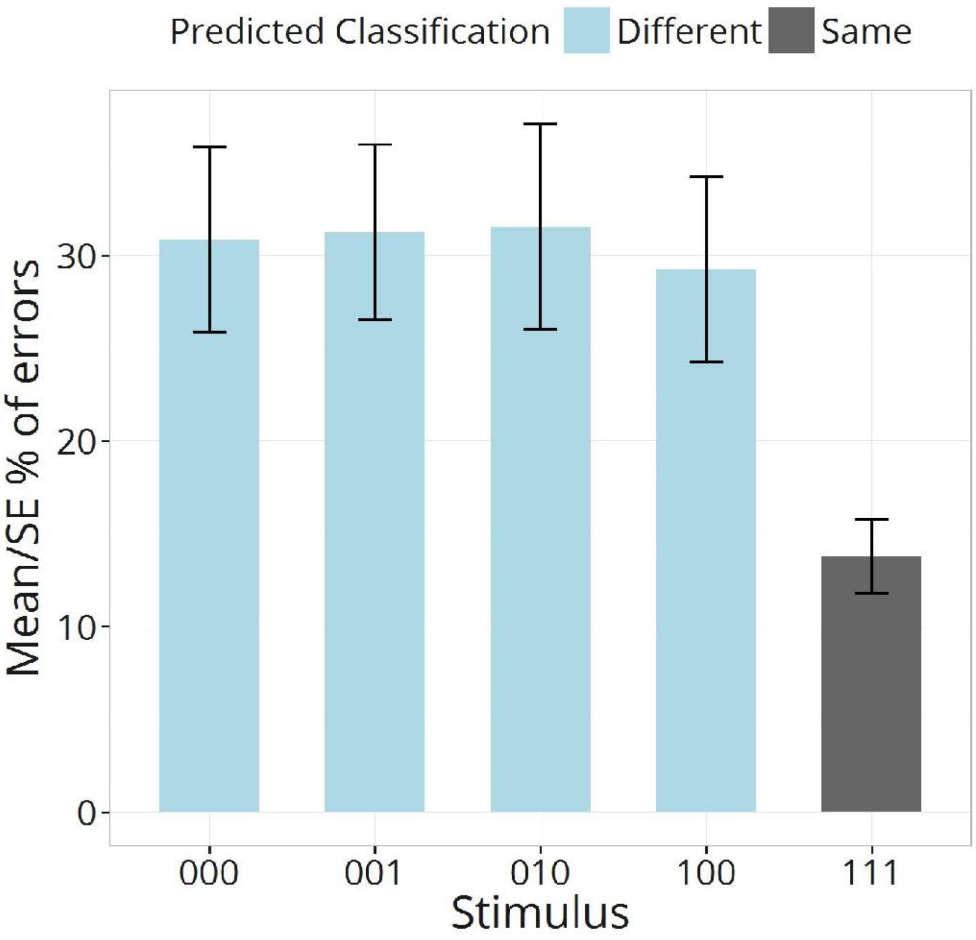
Optimal Experimental Design



For four out of five stimuli we predict different classification decisions given flexible dependencies versus Naïve Bayes.

Error Rates

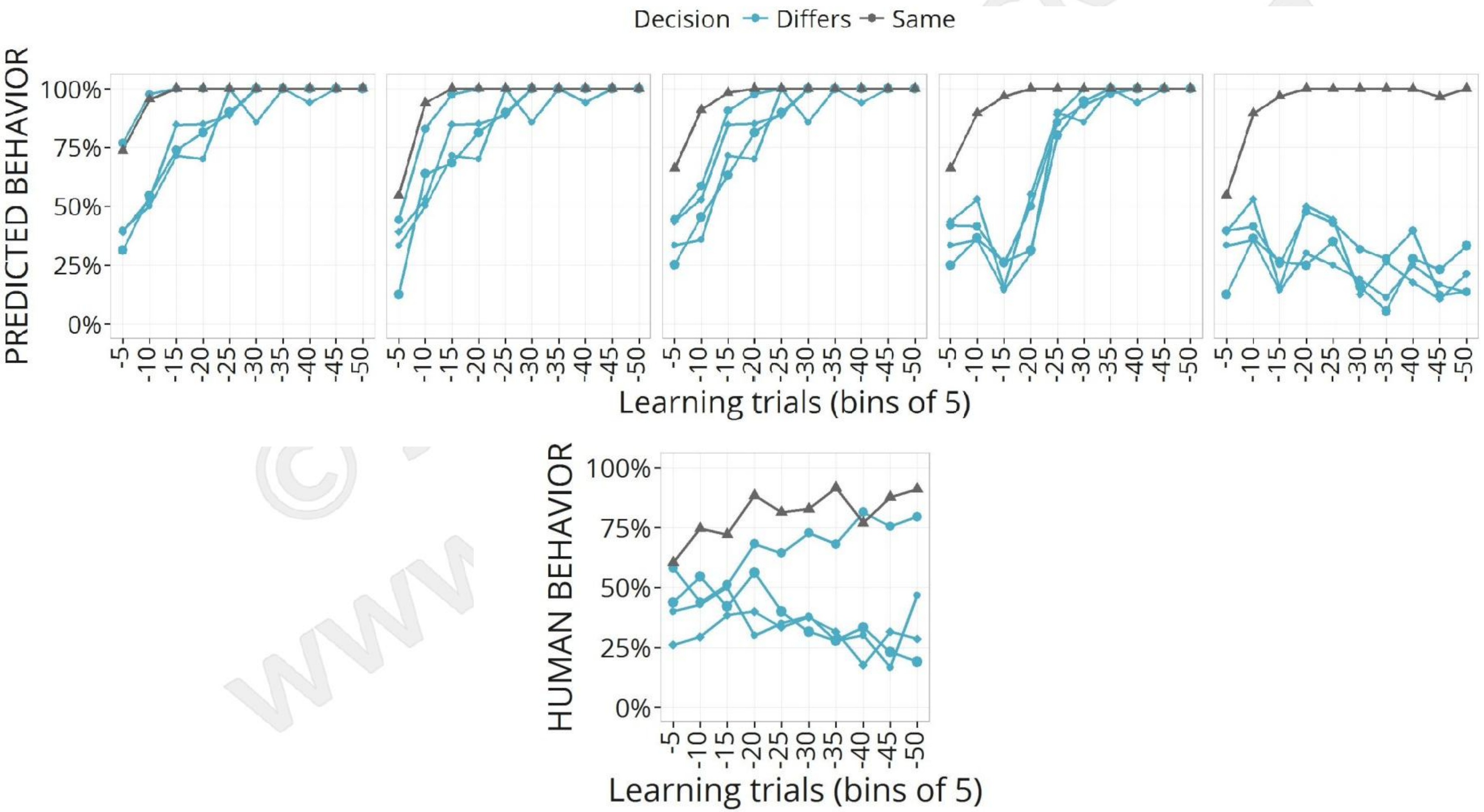

Different
Same
Different
Different
Different


Predicted Classification
Different
Same
Mean/SE % of errors
0
10
20
30
000
001
010
100
111
Stimulus

Modeling Learning

Belief in Naïve Bayes structure $P(NB) = \pi$

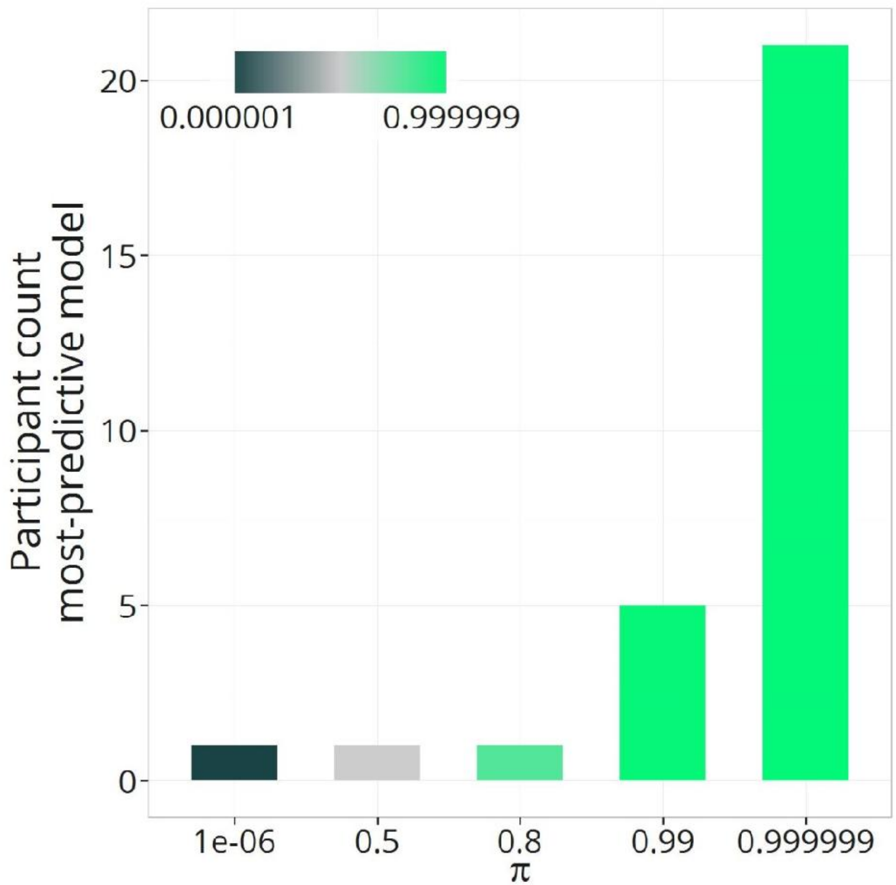
$\pi = 0$ $\pi = 1$



Most predictive model

Belief in Naïve Bayes structure $P(NB) = \pi$

$\pi = 0$ $\pi = 1$



Summary

- Given the computational complexity of classification
- We used Naïve Bayes as computationally plausible principle from Machine Learning
- As default assumption for human classification
- Classification learning is in line with Naïve Bayes
  - Proportion of errors
  - Classification learning decisions
- People are not stuck with it

References

Ashby, F. G. & Gott, R. E. (1988) Decision rules in the perception and categorization of multidimensional stimuli. *Journal of Experimental Psychology - Learning, Memory, and Cognition* 14, 33–53. doi:10.1037/0278-7393.14.1.33

Domingos, P., & Pazzani, M. J. (1997). On the optimality of the simple Bayesian classifier under zero-one loss. *Machine Learning*, *29*(2), 103–130. doi:10.1023/A:1007413511361

Katsikopoulos, K. V, & Martignon, L. (2006). Naive heuristics for paired comparisons: Some results on their relative accuracy. *Journal of Mathematical Psychology*, *50*, 488–494. doi:10.1016/j.jmp.2006.06.001

Medin, D. L. & Schaffer, M. M. (1978) Context theory of classification learning. Psychological Review 85, 207–38. doi:10.1037//0033-295X.85.3.207

Nosofsky, R. M. & Clark, S. E. (1989) Rules and exemplars in categorization, identification, and recognition. *Journal of Experimental Psychology - Learning, Memory, and Cognition* 15, 282–304. doi:10.1037/0278-7393.15.2.282

Rish, I., Hellerstein, J., & Thathachar, J. (2001). *An analysis of data characteristics that affect naive Bayes performance*.