Improvement of the Bayesian generalization model in order to handle negative examples and discontinuous hypotheses

May 16, 2022

abstract-text

1 Background

The problem our project is focusing on is Shepard's ideal generalization problem. The generalization problem focuses on how humans build hypothesis spaces for a given consequence after observing stimuli. In the paper "Generalization, similarity, and bayesian inference", they discuss how using a model of bayesian inference, we can predict the probabilty of given stimuli being included within the consequential region [**Tenenbaum**]. The model uses the equation $p(y \in C \mid x) = \sum_{h:y\in h} p(h|x)$ where h is a hypothesis from the hypothesis space $\mathcal{H}$ and $p(h|x)$ is the posterior probabilty of the hypothesis after observing x. We plan to extend this model to investigate how including negative examples within the x vector(x is the observed stimuli) affect how the model limits hypotheses. We also plan to explore how different distrubtions and models compare to the original model for generalization.

2 Question

To generalize the model by Tenenbaum et al, we want to find a way, how it can be improved, so that it can handle negative examples and discontinuous hypotheses.

3 Negative examples in the baseline model

The baseline model (Figure **??**), which was introduced by Tenenbaum et al is by itself not capable of handling negative examples.
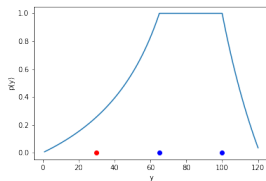
A naive approach to incorporate negative examples, is to fit a model the same way the
Evaluate how negative examples affect the predictions of the original model and use it as a baseline (2) Evaluate the predictions made by the original model for multi cluster sample data (3) Evaluate how the model could be altered so that negative examples can be incorporated to limit cluster intervals assuming a continuous consequential region (4) Evaluate how the model could be altered so that negative examples can be incorporated to limit cluster intervals assuming a discontinuous consequential region 1 (5) Evaluate how the model can be expanded to handle discontinuous regions without negative examples We want to initially solve the first three issues. Due to the added co

We start with a naive approach, where we will use the base model and fit one to positive and one to negative examples.

References

[JLK17]   Alan Jern, Christopher G. Lucas, and Charles Kemp. "People learn other people's preferences through inverse decision-making". In: *Cognition* 168 (Nov. 2017), pp. 46–64. DOI: 10.1016/j.cognition.2017.06.017. URL: https://doi.org/10.1016%2Fj.cognition.2017.06.017.