

Ask the Right Questions

Dynamic Perception Testing with Active Machine Learning

Ash Hanzl · Institute of Phonetics, Faculty of Arts, Charles University · prokophanzlo3@gmail.com



INSTITUTE
OF PHONETICS
Faculty of Arts
Charles University

THE PROBLEM

Perception experiments can get tiring for participants if too many stimuli are presented.

THE GOAL

Reduce the number of stimuli presented without compromising on results.

THE SOLUTION

Model the participant's behavior live & let the model take over for obvious answers.

How It Works

First stimuli are chosen at random

Class diversity is ensured before moving on

The initial model is built

Predictions are made for all remaining stimuli

The most informative* stimulus is chosen

*Stimulus with the lowest-certainty model prediction

The model is rebuilt with the new information

Are all predictions good enough?

*Remaining prediction certainties exceed the preset value

If not, go back

The perception experiment ends

Each stimulus has either been classified by the participant, or the model's prediction is high-certainty

Example Experiment

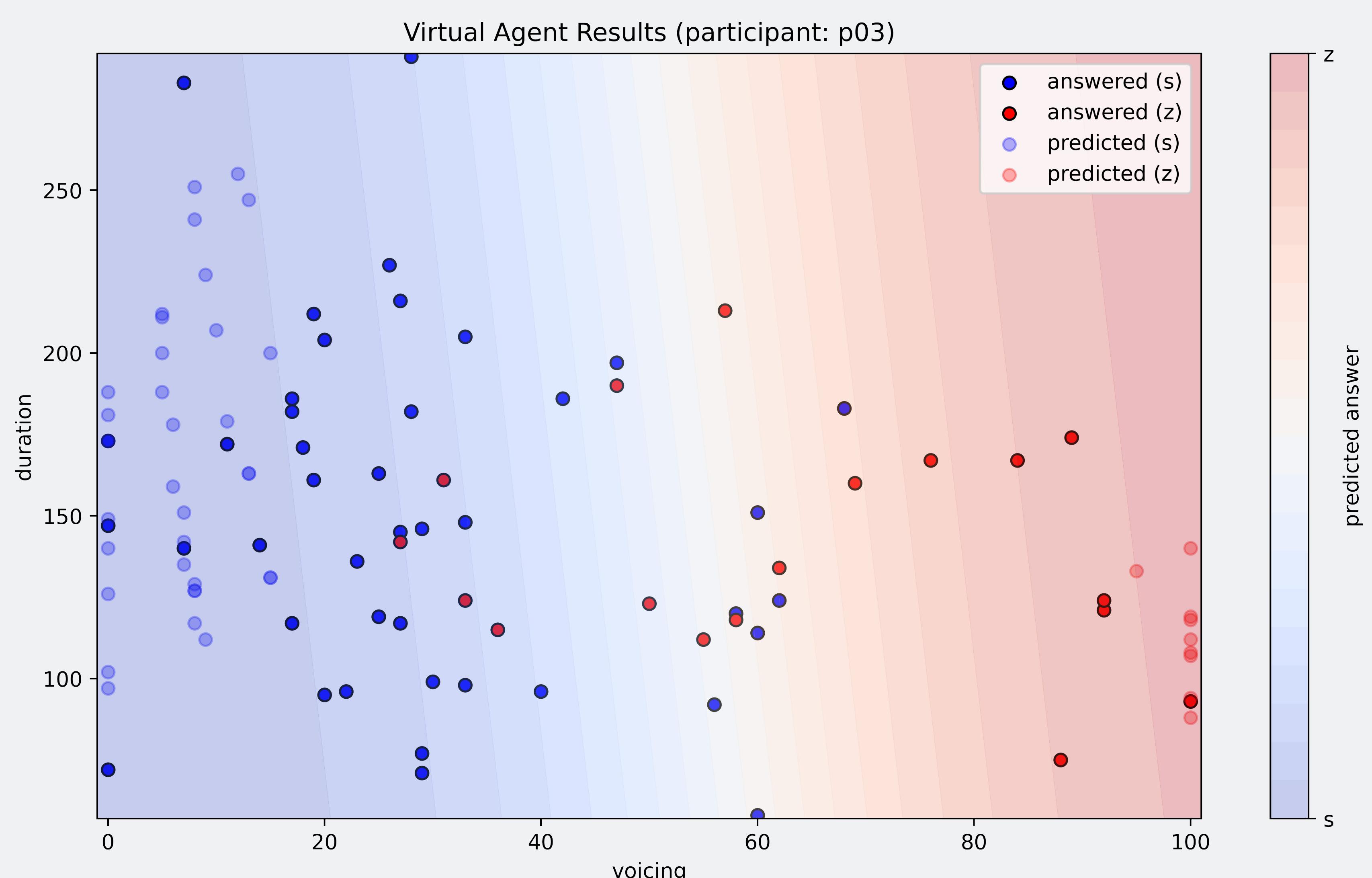
To test the framework (AsTRiQue – ASk The RIght QUEstions), data from a perception experiment investigating how Czech speakers categorize the sibilants /s/ and /z/ was used. Each stimulus was parametrized by two acoustic features: voicing (percentage of the segment with vocal fold vibration) and duration (in ms).

In the original experiment, 31 human participants each manually labeled all 104 stimuli.

For testing, these real human responses were used in a virtual agent setup (each participant was simulated as an oracle via a lookup table).

Key outcomes:

On average, the virtual agent only had to classify ~83 out of 104 stimuli (~20% reduction in participant workload). For the remaining stimuli, the model's prediction accuracy exceeded 95%; the effective overall classification reliability (participant + model) thus exceeded 99%.



What's Next?

AsTRiQue is in the early stages of development. As of now, the framework supports only binary classification tasks with two numeric predictors, using a logistic regression model.

Additionally, stratified sampling will be implemented in the initialization stage to account for data diversity.

Currently, the framework has the option to include a cleanser (high-certainty) stimulus every n iterations to reduce fatigue. This will be expanded with more distractor options.

In its finished form, AsTRiQue will be a low-code solution for researchers to administer perception experiments while maintaining granular control over the technical configuration of the system (initialization configuration, minimum iterations, cleanser stimuli/distractors, model certainty cutoff, model type used for training), and it will be deployable online with a GUI.

poster PDF, reference & more info on GitHub:

