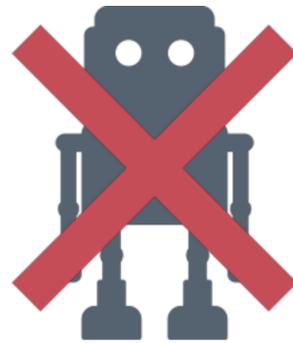
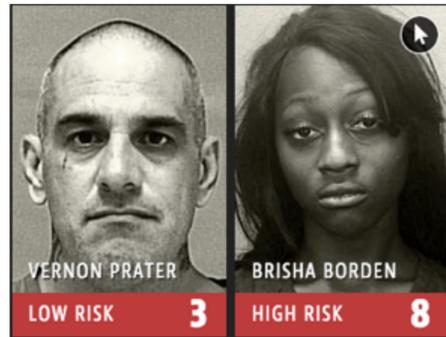


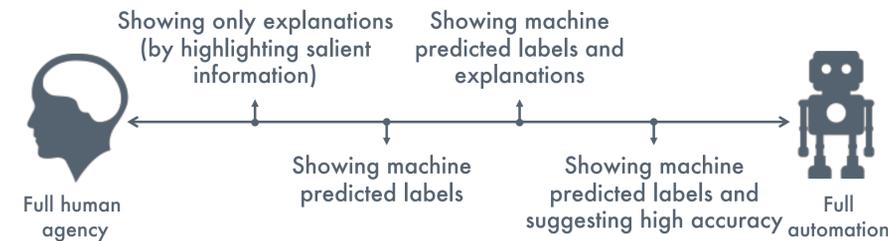
On Human Predictions with Explanations and Predictions of Machine Learning Models: A Case Study on Deception Detection

Motivation

Machine learning holds promise for addressing important societal challenges, but full automation is not desired in many tasks due to ethical and legal concerns.



Goal



- The goal of our work is to investigate best practices for integrating machine learning into human decision making.
- We propose a spectrum between full human agency, where humans make decisions entirely on their own, and full automation, where machines make decision without human intervention.

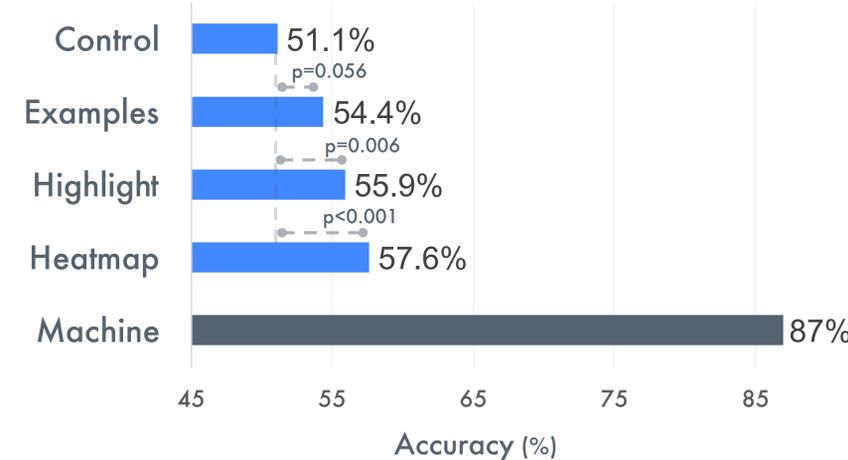
Vivian Lai and Chenhao Tan
 @vivwylai | @chenhaotan
 vivlai.github.io | chenhaot.com
 University of Colorado Boulder
deception.machineintheloop.com



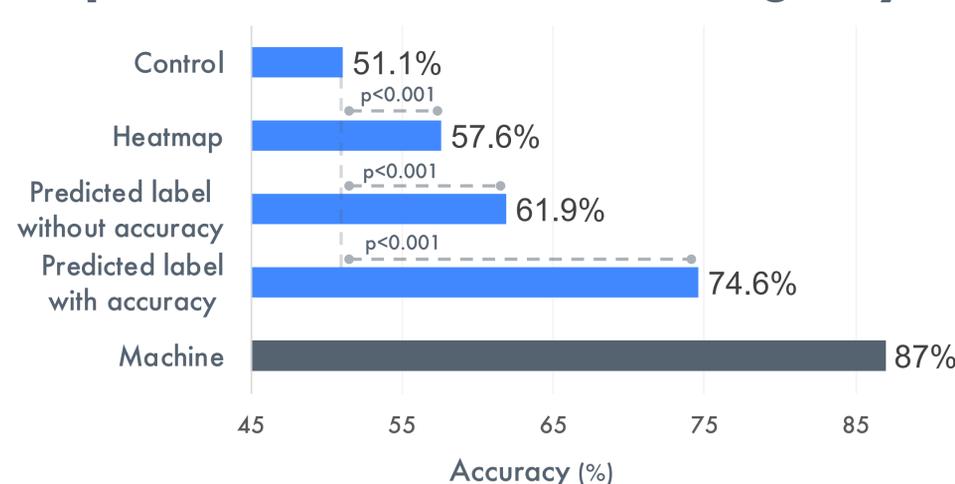
Deception Detection as Testbed

I must say, this hotel is absolutely beautiful. My wife and I were blown away by the penthouse fitness center and pool all the way up on the 42nd floor. Very well designed rooms with great linens and pillows. The bathrooms were a nice change of pace from the boring hotel bathrooms we usually see. Overall an excellent experience here, great views and close to everything you could want to do in Chicago.

Explanations alone slightly improve human performance

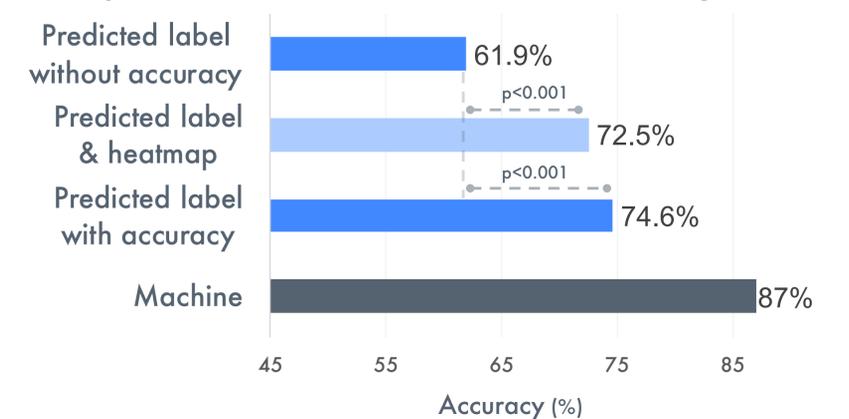


Tradeoff between human performance and human agency

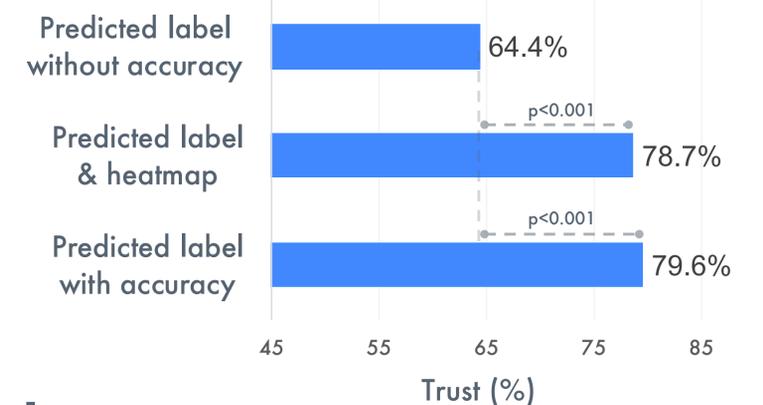


Explanations can moderate the tradeoff

Predicted labels coupled with explanations induce a similar effect as compared to predicted labels and explicit statement of machine accuracy.



Explanations help increase humans' trust



Takeaway

