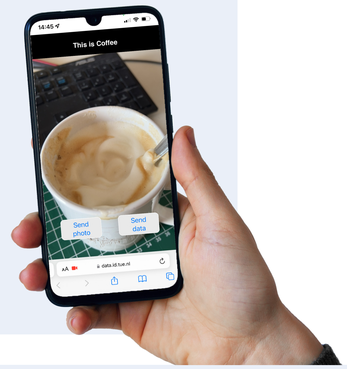
“Correct Me If I Am Wrong” Exploring How AI Outputs Affect User Perception and Trust

Jun Li Jeung
Eindhoven University of Technology
J.l.Jeung@hotmail.com

Janet Huang
Eindhoven University of Technology
Y.c.huang@tue.nl

Intro

Unexpected AI results may lead to frustration, disappointment, and technology abandonment, causing mistrust or over-trust [1]. To seamlessly integrate AI into our daily life, this work aim to investigate how AI outputs influence people's perceptions and experiences in their everyday practice, particularly when they are given the opportunity to correct AI mistake.



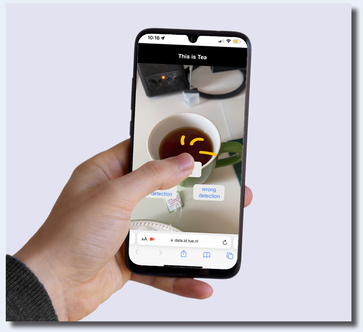
TrinkTracker

In this work, we present TrinkTracker, an AI-powered system for monitory daily drinking practices. The TrinkTracker is capable of recognizing six types of beverages based on a photo and generating a data report summarizing people's drinking behaviors.

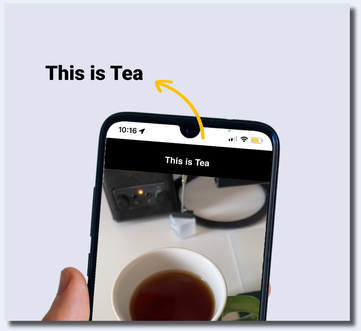
SCAN ME

Test the prototype

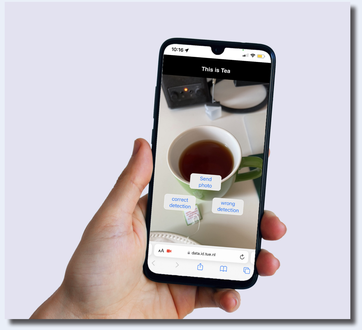
1. User takes a picture



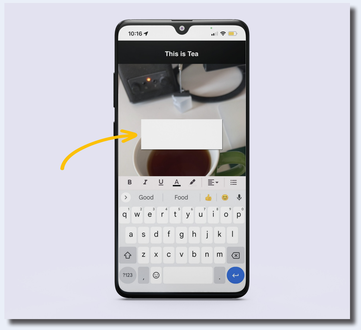
2. AI makes a prediction



3. User send the data

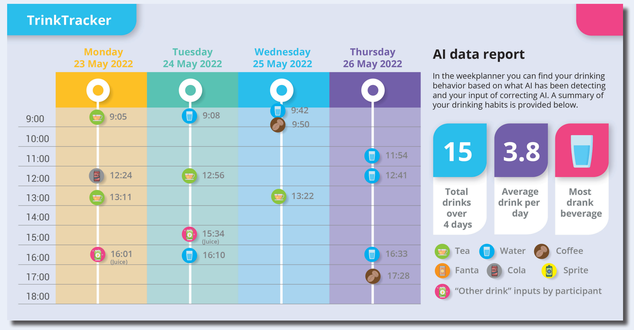


4. User modify the prediction



(The baseline system did not include this step)

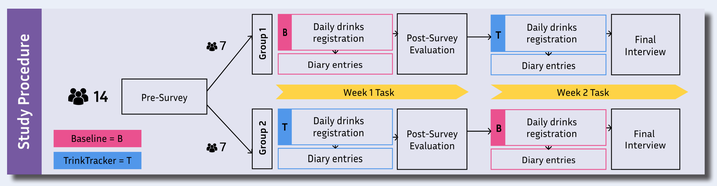
5. Generating a data report



Weekly overview of users drinking behavior

Field Study

We conducted a two-week field study with 14 participants to understand user perception, experience and trust on AI predictions by comparing two conditions: TrinkTracker and a baseline system.



Key Highlights

(1) Trust
Higher trust when giving opportunity to correct AI errors ($p<0.5$)

(2) Correct detections
Higher satisfaction when seeing AI made correct detections ($p<0.5$)

(3) Wrong detections
Wrong detections sparked people's curiosity, leading them to speculate on the underlying reasons behind the AI model

(4) Drinking awareness
Data report helps increase people's awareness and then facilitate them to reflect on their drinking behaviors

Future and Discussion

- Users usually hold expectations for AI to perform tasks accurately.
- Unexpected AI mistakes did not result in frustration but rather sparked curiosity among users.
- Adviced to provide explanations for mistakes to support people in understanding the AI capabilities and limitations [2].
- A follow-up study to see how end-users behave and respond to an AI model that learns from human input over time.

Reference

[1] Kocielnik et al., (2019). Will You Accept an Imperfect AI? Exploring Designs for Adjusting End-User Expectations of AI Systems.(CHI '19). https://doi.org/10.1145/3290605.3300641
[2] Maltbie et al., (2021). XAI Tools in the Public Sector: A Case Study on Predicting Combined Sewer Overflows. (ESEC/FSE 2021). https://doi.org/10.1145/3468264.3468547