HCML HEURUSTICS

Principles for Human-Machine Learning Design

by Corinne Schillizzi

1 Data Translation

Description

Interactions should always translate data into human interpretable visualisations or explanations.
This translation process is fundamental for users to comprehend the system's results effectively, thereby enabling them to make informed decisions.

Guiding questions

- *Does the interaction provide easy to understand visualizations or explainations of the underlaying data used?*
- *Does the interaction asks for feedback in an easy to understand visualizations or explainations of the underlaying data asked?*

Examples

Thumbs up or down feedback

A simple interaction like rating content with a thumbs up or down translates the user's preference into data for the machine to learn from and refine recommendations.

Ranked e-commerce items

In e-commerce or streaming platforms, recommendation systems present users with lists of items ranked by relevance. These rankings reflect the underlying data processed by the system, making it easier for users to see what's most likely to interest them.

2 Convey Uncertainty

Description

The system should always convey the intrinsic uncertainty and susceptibility to errors transparently. Users should be able to gauge the confidence level of ML outputs, fostering a deeper understanding and enabling users to make decisions accordingly.

Guiding questions

- *Does the interaction clearly communicate uncertainty in the data or predictions?*
- *Are confidence levels or error margins presented in a way that users can understand?*

Examples

Showing scores

Displaying a probability percentage (e.g., 80% confidence) alongside a medical diagnosis prediction helps users understand the reliability of the output.

Showing errors

Displaying error bars on a graph for a sales forecasting model. The error bars represent the uncertainty around the predictions, helping users understand the range of potential sales outcomes.

3 Convey and Limit Impact

Description

The system should communicate to the user how their input affects the learning of the system. Impacting not only their experience but also that of others. Sometimes, it is prudent to limit the impact of feedback on the system deliberately, or through user decision.

Guiding questions

- *Is the effect of user input on the system's learning process visible to users?*
- *Can users limit their feedback's influence on the system when necessary?*

Examples

Personalized contents

Users may receive a notification stating how their ratings on a movie are improving future movie recommendations for themselves and others.

Private mode

In a fitness app, users may turn on "private mode," preventing their workout data from affecting the community-wide leaderboard.

4 User Augmentation

Description

ML should primarily serve as a tool for human empowerment by providing adjustable actions and parameters to align the technology with diverse user needs and preferences, ultimately aiding users in accomplishing more than they could independently

Guiding questions

- *Does the system offer customizable parameters that enable users to adapt ML features to their specific needs?*
- *Does the interaction enhance user capabilities, enabling them to achieve better outcomes than they would alone?*

Examples

An AI-powered image enhancement tool that allows users to tweak brightness, contrast, or saturation after the automatic adjustments.

A word processor suggests sentence completions, which users can accept, edit, or ignore, saving time while ensuring the text matches their style.

5 Balanced Trust and Distrust

Description

The system should balance earning trust and distrust among users by acknowledging the strengths and fallibility of ML. This balance ensures that while users can rely on the system, they remain critical of the outputs and do not accept results without scrutiny.

Guiding questions

- *Does the system provide transparency around its limitations to prevent over-reliance?*
- *Are users encouraged to review or verify critical outputs?*

Examples

Bias warning

In hiring tools the system shows a disclaimer that it may be subject to bias and recommends human review for critical decisions.

Double-check

After an AI-suggested route is displayed, the app prompts users to confirm that the path looks reasonable before starting.

6 Balanced Autonomy and Control

Description

The system should strike a balance between machine autonomy and human control when necessary to obtain a quality outcome and overcome errors. Ensuring that final decisions rest with the users is imperative.

Guiding questions

- *Does the system provide users with control over important decisions while automating less critical tasks?*
- *Can users intervene or override system actions when necessary?*

Examples

User intervention

A self-driving car can take over for highway driving but requires the user to make final decisions in tricky city environments.

Automated routine tasks

In financial portfolios users can let the system handle routine investments but maintain control to approve or reject major portfolio changes.

CONTENT FROM