Evaluation of WIT using the [Guidelines for human-AI interaction design](#)

From Ali Bektas and Christopher Kilian

Issue 1: What-If prediction unresponsive when deleting a feature from a datapoint

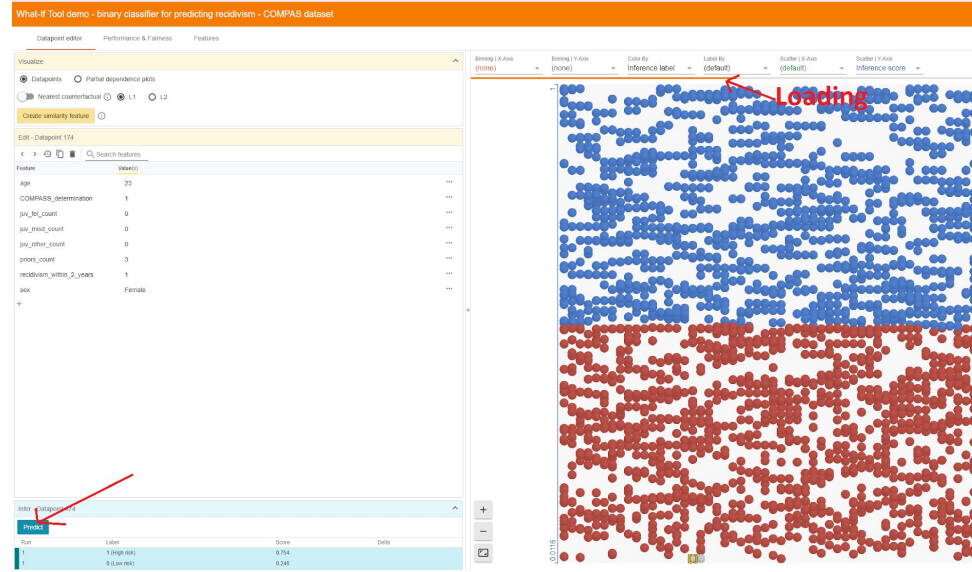| |
|---|
| **Heuristic:** |
| G1 - Make clear what the system can do. |
| **Description:** |
| When selecting a datapoint in the Datapoint Editor View, deleting a feature (e.g. race) and trying to make a What-If prediction, the view looks like it is processing something but nothing happens. Aborting the action is only possible by refreshing the page. |
| **Severity***: 3 (between 2 and 3) |
| **Recommendation**: |
| Grey out the Predict button and describe in which scenarios this is possible and in which not. |

Issue 2: Provide information on how good or bad the used distance metric is when using Nearest Counterfactual

Heuristic:
G2 - Make clear how well the system can do what it can do
Description:
Depending on the number of features a datapoint has, the used distance metric (e.g. euclidean distance) can perform different (e.g. it can perform worse with high dimensional features). Especially for people without data science background it would be good to add such information.
Severity*: 3 The user is less efficient.
Recommendation:
Add information about the distance metric to the info box or documentation

pair-code.github.io/what-if-tool/demos/compas.html

What-If Tool demo - binary classifier for predicting recidivism - COMPAS dataset

Datapoint editor   Performance & Fairness   Features

Visualize

◉ Datapoints   ○ Partial dependence plots

◯ Nearest counterfactual ⓘ ◉ L1   ○ L2

Create similarity feature ⓘ

Edit - Datapoint 174

< > ↺ ▢ 🗑 🔍 Search features

Feature	Value(s)
age	23
COMPASS_determination	1
juv_fel_count	0
juv_misd_count	0
juv_other_count	0

Issue 3: Unclear choices in dropdowns

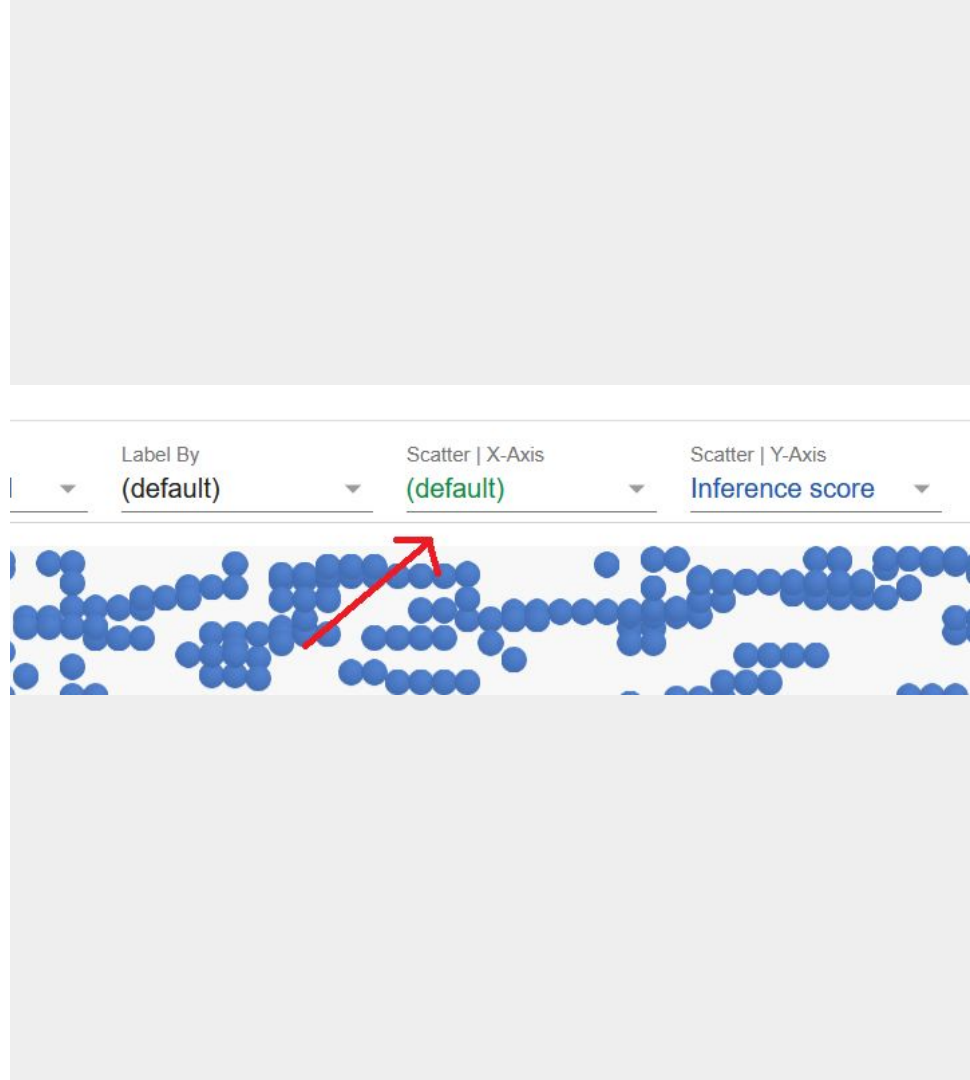| |
|---|
| **Heuristic:** |
| G1 - Make clear why the system did what it did. (not directly related to AI Guidelines) |
| **Description:** |
| It is not clear what is set as default as parameters for the visualization |
| **Severity***: 3, as it impacts the understanding/interpretation of the visualization |
| **Recommendation**: |
| Provide a description for the different choices |

Issue 4: Select features to be included in the model (Feature)

Heuristic:	
G17 - Provide global controls.	
Description:	
Provide controls for selecting the features that should be used for training the model	
Severity*: 2	
Recommendation:	
Use gear 'button' at the top right corner to select features (open a settings menu?)	

3496 datapoints loaded

Scatter | X-Axis
(default)

Scatter | Y-Axis
Inference score