Decisions



How these predictions will affect the end-user? What are the implications of these decisions to the end-user when the system's predictions are wrong?

Al Assessment Task

Type of assessment, Group variables,

Output to explain, Worst-case analysis.



Safety propositions



What are we trying to do for the enduser(s) of the system? What objectives are we serving? How our decisions will affect the end-user(s) prospects?.

Data Sources

using? Are we infringing

individual/company rights of privacy

and agency in using these data

sources? (internal and external)



Collecting Data



Are the subject aware of the data collection? What are the access/control of the subjective over its data? The data being collected are actually being used during the modelling process, or are we just storing it?

Making Assessments

When do we make on new

the system's decisions?

predictions/decisions? How long do we

have to wait to see the implications of

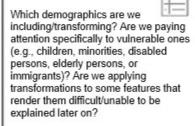


Offline assessment



Methods and metrics to evaluate the system before deployment, such as on Fairness, Explainability and Robustness.

Features



Building Models



What are the limitations of the model being created/updated (Robustness. Explainability and Fairness)? What are the measures put in place to mitigate or eliminate these limitations are being addressed?

Live Assessment and Monitoring

