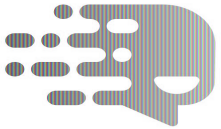
Explaining Probabilistic Reasoning

Adarsa Sivaprasad
2nd year Phd, School of Computing Sciences

University of Aberdeen

Supervisors: Dr. Ehud Reiter, Dr. Nava Tintarev, Dr. Nir Oren

1 4 9 5
UNIVERSITY OF ABERDEEN



Interactive Natur
Technology for e
Artificial Intellige
NL4X

Risk communication in healthcare

Risk score: Quantifiable risk with extensive data and understanding of controlled environment.

End user and trust:

- Effectiveness of communication is receiver dependent.
- Choice of explanation form, design considerations : graphical, textual, narrative (counterfactual, contrastive, example based)
- The individual difference in perceiving probability.
- Evaluation of explanation usefulness.

Existing work in explaining AI models for healthcare risk prediction:

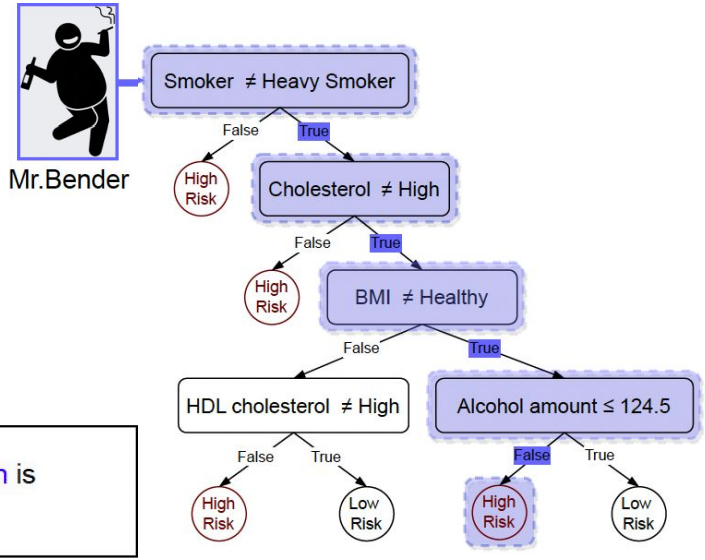
- Model fact tables.
- Intrinsically Interpretable models.
- Model agnostic explanation methods: LIME, SHAP, Anchor
- Post-hoc explanation for black box models: Model distillation, mechanistic interpretability.

These are not necessarily useful for non expert end-users.

ClinRisk QRISK®3-2018 PIS University of Aberdeen QCovid® Risk assessment QCancer®

Some patient facing probabilistic models in healthcare.

Local Explanation : Explain the decision for a particular patient input.



Local narrative explanation of a decision tree:

Mr.Bender has High risk of CHD since (BMI is not Healthy) and (daily alcohol consumption is greater than 124.5ml) even if is (not Heavy Smoker) and (Cholesterol is not High).

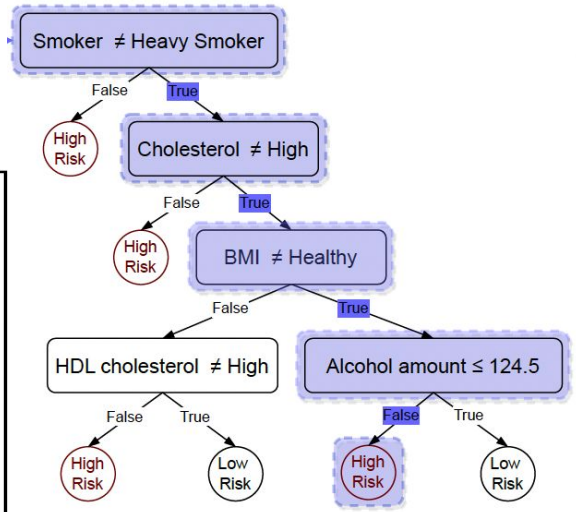
Global Explanation : Explain all possible model decisions based on the training patient records.

Local+Global explanation of a decision tree:

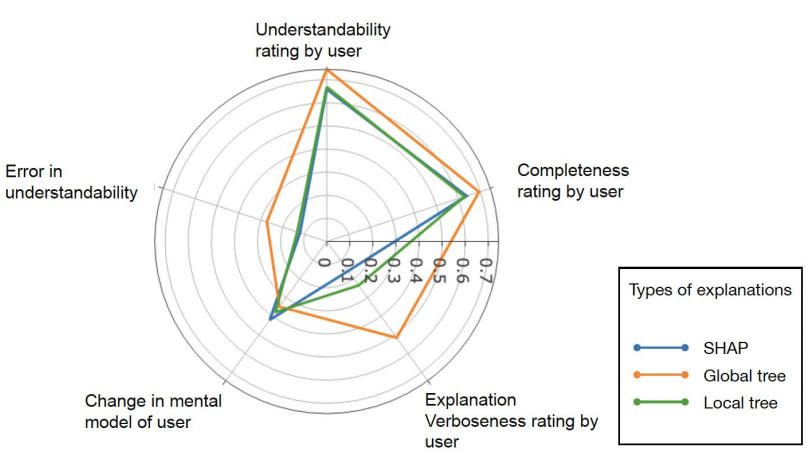
A patient has High risk of CHD if:

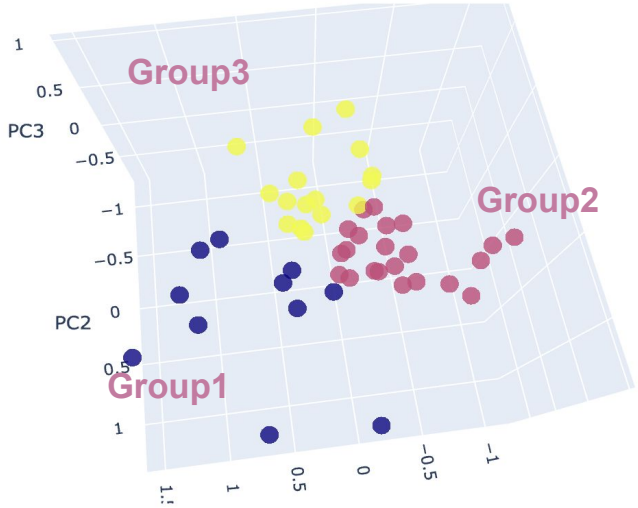
1. (Is Heavy Smoker).
2. (Cholesterol is High) even if (not Heavy Smoker).
3. (HDL cholesterol is High) even if is (not Heavy Smoker) and (Cholesterol is not High) and (BMI is Healthy).
4. (BMI is not Healthy) and (daily alcohol consumption is greater. than 124.5ml) even if (not Heavy Smoker) and (Cholesterol is not High).

Mr.Bender has High risk of CHD since he follows Rule 4.



Shortcoming of XAI methods for non-expert users - Coronary Heart Disease prediction





1. Certain people strongly prefer specific type of explanation. This preference does not necessarily translate to understandability.
2. The explanation understanding is strongly dependent on the complexity of the feature interaction being explained.

Sivaprasad, A., Reiter, E., Tintarev, N., Oren, N. (2024). Evaluation of Human-Understandability of Global Model Explanations Using Decision Tree. In: Nowaczyk, S., et al. Artificial Intelligence. ECAI 2023 International Workshops. ECAI 2023.

Patient facing risk prediction model - In-vitro fertilisation outcome prediction

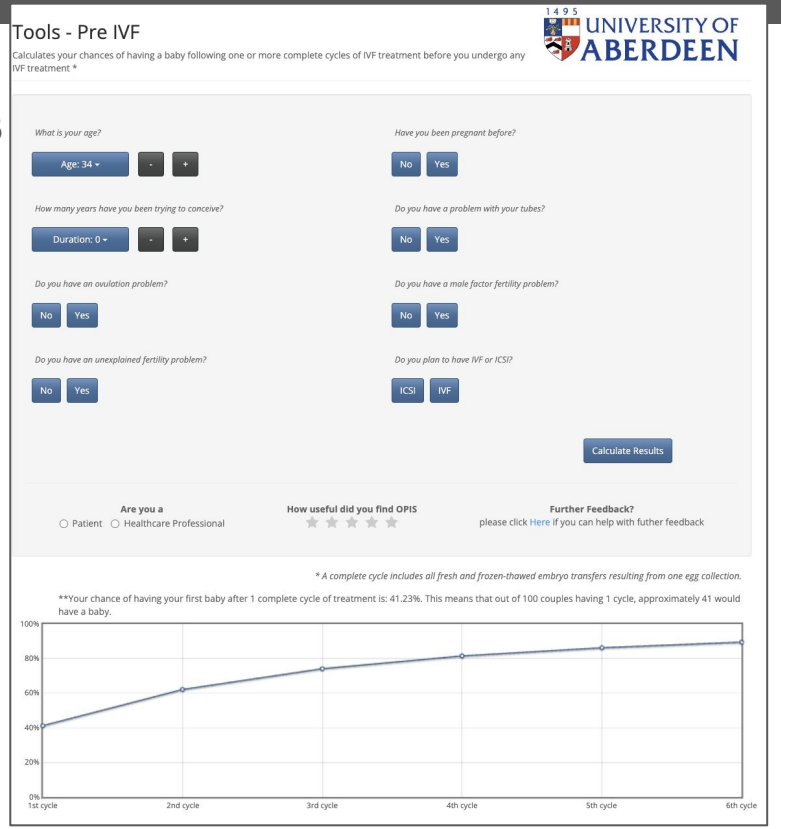
McLernon model : Model rained on HFEA data from 113873 women who underwent IVF in United kingdom from 1999-2008. (cross-validated on more recent data)

Estimated lifetime prevalence of infertility globally: 17.5%

Estimated percentage of population currently experiencing infertility: 12.6%

The tool is used by ~3000 users in a month.

https://w3.abdn.ac.uk/clsm/opis/tool/ivf1



McLernon DJ, Steyerberg EW, Te Velde ER, Lee AJ, Bhattacharya S. Predicting the chances of a live birth after one or more complete cycles of in vitro fertilisation: population based study of linked cycle data from 113 873 women. BMJ. 2016;355:i5735.

OPIS: In-vitro Fertilization Outcome Prediction – User Feedback

All the healthcare professionals rated OPIS was user-friendly.

24% of patients said they did not understand what the results meant.

13% of patients found the OPIS presentation Not user friendly.

Most recurring patient feedback:

- “I have a specific condition of **PCOS** that affects the chance of pregnancy, but this model does not consider it. “
- “My specific case of using (**no transfer in first cycle**) is not included as an option to input”
- Unable to interpret the probability.

Sivaprasad, A., Reiter, E.,. Linguistically Communicating Uncertainty in Patient-Facing Risk Prediction Models. In Proceedings of the 1st Workshop on Uncertainty-Aware NLP, EACL 2024.

Uncertainty Communication

Model confidence

- Considerations for communicating predicted probability as cumulative graph and percentages.
- Users question of model confidence under factors not considered in modelling.

(*I do not understand the graph.*)

Model Reasoning

- Consideration of difference in model reasoning and mental model of the patient for expectation management.
- Required to achieve a balance between complexity and the need for a global model explanation.

(*I do not understand why the probability reduces when I add male factor infertility.*)

Confidence in individual prediction

- For an informed user, required to communicate the confidence of individual prediction under personal medical history.

(*Has the model been trained on patients who has similar demography as me?*)

Unknown Knowns

- The current model lacks information on factors such as patient BMI and smoking status, which are recognized as influencers of infertility.

(*This model does not consider my endometriosis. I do not trust this model.*)

Ongoing Work

User profiling: Specific to the context of infertility treatment, what is the influence of patient demographic, and personal characteristics in the understandability of a particular kind of explanation.

Uncertainty communication:

- What are effective way to communicate the performance metrics and confidence of a particular prediction to the user?
- Elicit the user expectations from an explanation-based UI.

Generating narrative explanations:

- Formal framework for communicating uncertainty through a dialogue interface.

Thankyou.

Connect at: a.sivaprasad.22@abdn.ac.uk

LinkedIn