Protocol: Midterm Presentation Bachelor Thesis "Demographic Biases in Dermatology AI"

Date: 07 April 2025, 09:00–10:00
Participants: Jürg Schelldorfer, Ludovic Amruthalingam, Nadja Stadelmann
Protocol written by: Nadja Stadelmann

Topic	**Details**
Clarifications regarding the PASSION Dataset and Model	• The PASSION dataset was collected by researchers. • The PASSION model is a plain ResNet50 architecture which was trained on the PASSION dataset. • The PASSION model predicts dermatological conditions stored in the labels *conditions_PASSION* (eczema, scabies, fungal or others) and *impetig* (presence of impetigo) based on the input picture. • The PASSION model is not yet used in practice. • Bias in the PASSION model should be reduced, so that the model can serve as a benchmark model to assess other dermatology models in regards of fairness; highlighting biases.
General Advice	• It is important to be precise e.g. regarding whether one talks about the dataset or the model -> clearly differentiate them. • It is important to be knowledgeable when talking about biases and fairness, since it is a very diverse area. • Take into consideration the technical, but also the dermatological aspects. E.g. under-representation of different ages in a dataset is only an issue, if the disease presentation differs in reality based on the age of a patient.
Bias in Models vs. Representation in Datasets	• The provided definition of bias in the context of AI is good - keep in mind that it is focusing on the model's output only. • Even if the dataset is skewed in regards of representation, the models output can still be unbiased (= fairness metrics report fairness over subgroups). That's why a dataset itself is not "biased". • On the other hand, an unbiased model does not necessarily mean that the dataset is fully inclusive and has no limitations. • For the dataset, it's a question how representative the dataset is for given subgroups.
Dataset Limitations	• It's important to know (and clearly state) the limitations of the dataset (e.g. representation issues, what data is in-, what out-of-distribution) • In practice, when provided data belongs to a out-of-distribution-case, no result is provided by the AI model. (This is especially important in health care, since now false-positive diagnoses should be provided.)
Mitigation Methods	• Oversampling is not good practice, as it can lead to misleading results. (Was intended to serve as an example only in the presentation)
Report	• The list of relevant biases and mitigation methods which will be provided to the PASSION team can be added to the appendix. In the main report, focus on top 5 to 10 items.