Case Studies

Scenario 1: AI-Assisted Screening for a Rare Condition Predominantly Affecting African Populations

Background: An AI tool is being developed to assist doctors in screening for a rare but serious condition, which mostly affects people of African descent and which can be fatal unless treated promptly. Screening currently relies upon a team of experts reviewing reports of symptoms as well as follow-up MR imaging data. Using this approach about 10% of positive cases are missed resulting in significant mortality. There is an expensive but reliable blood test that can be used for subjects flagged up by the screening process.

Development and Training: The AI tool will be trained using the symptoms and MR imaging as input data and blood test results as output labels. Training and evaluation will be based upon data acquired from a network of 5 hospitals across the USA. It is intended that the tool will be made available globally to improve screening for the rare condition.

Questions to consider:

- What dangers/risks of the use of AI for this problem can you identify at this stage?
- How would you go about addressing these?
- What fairness metric(s) do you think might be appropriate when assessing the AI tool for potential bias?

REMINDER - DEFINITIONS OF FAIRNESS

- *False Negative Rate (FNR): the rate at which positive cases are missed by the classifier*
- *Demographic parity - equal chance of being classified positive for each protected group*
- *Equalised odds - equal true positive rate (TPR) & false positive rate (FPR) for each protected group*
- *Equal opportunity - only equalise either FPR or FNR, not both*

Suggested answers/discussion points:

- What dangers/risks of the use of AI for this problem can you identify at this stage?
    - The training data will presumably be mostly Caucasian (from the USA), so how will the tool perform on African races where the prevalence is higher?
    - Perhaps MR scanners used in some parts of the world will be older models, so how well will the tool generalize to data from these scanners?
    - Will demographic information be recorded, e.g. race, sex and age? This will be important for assessing the fairness of the tool.
- How would you go about addressing these?
    - Try to acquire data from a broader range of sources, especially from Africa.
    - Try to record demographic information, e.g. race, sex and age
    - Thoroughly test performance of the tool on different demographic groups.
    - Investigate whether the fairness metrics employed generalize to other domains, e.g. data from other held-out test hospitals where there may be a compound domain shift.
- What fairness metric(s) do you think might be appropriate when assessing the AI tool for potential bias?
    - Demographic parity would probably not be appropriate because of the higher prevalence in African races.
    - False negatives are probably more important to control than false positives – FPs will lead to unnecessary expense for the blood test but FNs could lead to mortality – so perhaps Equal Opportunity based on equalizing FNs could be appropriate?

Conclusion: This case study highlights the critical importance of addressing data imbalance and ensuring fairness in AI-assisted medical screenings, especially for conditions predominantly affecting specific demographic groups. By diversifying the training dataset, collaborating with international hospitals, applying data augmentation, and using fairness metrics such as demographic parity, equal opportunity, and false negative rate, the AI tool can be developed to provide equitable and accurate screening for the rare condition. This approach will help improve health outcomes globally, reduce disparities in medical care, and ensure that all populations benefit from advancements in AI technology.