Weekly meeting 16/10/2020

Prep:

* Survey
* Background
* Where is dataset, I have logged in switch
* What is balanced accuracy:  Balanced accuracy is calculated as the average of the proportion corrects of each class individually. In this example, both the overall and balanced calculations produce the same accuracy (0.85), as will always happen when the test set has the same number of examples in each class. This does NOT RELIABLY COMPENSATE for disparities in training dataset within classes (more of one class than another), but it’s better than nothing.

Need to do:

* Add in what I will add to interface: where will I add what, what is the text going in there (explanations + potentially images)
  + Uncertainty measures
  + SHAP possibly
  + Instance level explanations
  + LIME
  + ALE
  + Mean / variance / median of data, and of the possible categories within the data if they exist
* Look up some rigorous Ml model comparison methodology
* Finish background
* Something wrong with the datasets / alge:
  + if column has all missing values, it is not well dealt with I believe
  + after running model, trouble in viewing ROC curve
  + will ask Antoine and jonathan
* on the survey: be more specific + add what specific measures I was thinking about
  + for uncertainty: show a “pimped” graph of the malaria triage score
* on the project:
  + start coding:
    - dummy “boxed” version of: uncertainty measures, LIME, measures of X/Y/Z so the base code is ready at least
    - when access to database: then start thinking about a good model
      * talk to (cant remember name) person doing distributed Ml
* on the organization
  + start overleaf
    - do you have a special format or can I do my own / whatever
  + start bibliography
    - I mean start using it in overleaf else it’s annoying to do all of it at the end
  + Start commenting code when start coding
  + Change the whole organization of the folders and note taking
    - 1 word document for ML papers
    - 1 word document for infectious disease papers
    - 1 word document with ALL paper names and where I found some of them
    - 1 word document PER week containing recap of:
      * What I worked on
      * What is left to do
      * Weekly meeting recap
      * Future steps
      * Anything worth investigating
      * Check jonathan’s organization
    - Maybe trello would be useful
  + Start uploading to github fairly soon
* Check the reference papers for the 2 theses I was given