Meeting's report for ELLEN WANG

Honour's student

09-09-2021

1. Agenda for today's meeting

- Present findings on readings I've completed regarding my understanding of the Random Forest (10 min)
- Discuss any problems with my understanding.
- Next Steps for the Data
- Advice for writing the literature review

2. Work already completed

- Set up meeting with Audrey (until they won't be needed as frequently)
- Completed some literature readings
- Completed reading of Melinda's thesis
- Set up overleaf
- Add previous week's meeting notes to github
- Add resources to github
- Share github and all links to the team
- Complete the Data Access Form
- Reach out to Pierre regarding Honour's Courses

3. Meeting Minutes from previous week

- Jiyang: Lacune and medical expert
- Pierre: Supervisor, RF expert and has a Masters in Neuroscience
- Audrey: RF expert
- Method: Random forests
- Explore this method and all its different iterations and variations
- Project: Predict location of lacunes in MRI scans, supervised learning process
- Data: will need access to this after filling out a form. This data has already been manually labelled and will be easier to predict.
- A lot of the work in Honours year is to address interesting research questions, and this may change over the year.
- Melinda used Deep Learning to do this task, try out Random Forests for the same purpose and try and improve the prediction
- Random Forests are easier to interpret, can be modified to better fit the specific dataset
- The medical images data is very difficult to interpret, can depend on the orientation of the brain, may need to slice 3D images to see 2D versions.
- Lots of recent literature online, try and reproduce their results first before further modifying and creating your own.
- Aim to have a consistent meeting schedule once every week or two weeks (30 minutes to 1 hour) to jointly discuss findings and results.

- Share github and link to overleaf and Trello
- Before each meeting, send out an agenda 2 days before the meeting, write down minutes and discussions, action tasks and what things to do before the next meeting.
- There should be one file for each meeting
- You may create a google drive if you want
- Confidential data, will need to fill out data application form, talk to Jiyang about this
- Will need access to Katana, a computing cluster
- Github should be private for now
- Reach out to Jiyang for journals to read and explore
- Pierre to provide advice on literature surveying, and existing data results
- Aim to create some meaningful features to feed into the model, some findings will need to be fed back to Jiyang to figure out their medical importance.
- Question from Pierre and Audrey: Will the model just be images or will it include additional information such as the age of the patient and their medical history?
- Answer: At the moment, we will start with images as lacunes are defined in their image and not so much the patient
- Localisation of the candidates, lacunes are a sudden change in grey scale?
- The aim is to reproduce some existing results in our own data before using more recent developments in the algorithm to improve.
- This project will be driven by Jiyang's questions and we will aim to aim to address these.
- Random forest for MRI will be built on understanding the classical random forest. Remember to consider things such as hyperparameters!

4. Planned work after this meeting

- Set up Jabref and input articles you have read through
- Set up a system to keep a record of the main points, proofs, agreements, disagreements in each article you read
- Use google scholar to do your research
- Revise random forests using the decision trees to bagging to random forest approach
- Revise the history of random forests: read up about it, gini impurity and entropy
- List parameters involved in the random forest and what they control
- Program the basic version of the random forest.
- Apply this basic version of your decision tree on a simulated dataset
- Go through good articles and tie in their method with yours. Can it help you?
- Audrey: bump on equations that are relevant for bagging
- Fix Latek format on Overleaf
- Organise meeting with Jiyang