**BMI meeting 2**

8 March 2021

Alex, Leire, Josh, Wenli, Pierre (GTA)

**Conclusion:**

In the next week:

1. Implement ECOC, classification tree and random forest
2. Use mean trajectory to predict (x, y) arm coordinates
3. Fit code to submission format

Next meeting: 15 March 2021

* Update each other on progress made
* Discussion of what worked and what didn’t
* Finalise choice on “best” classifier

1. **Difficulties with formatting data to input into functions**

For now, we can all use the same structured data as inputs to our classifier. Refer to the GitHub repository.

1. **Thoughts on approaching the problem as classification instead of regression?**

Good idea to use classification. It is a nice way to break down the problem instead of going straight to decoding the movements (regression). But classifiers can perform badly in the presence of minimal data so see where the error comes from. There are many improvements that can be made but think of the cost-benefit ratio of each method and the time limit (5 minutes) as well.

1. **Opinion on ensemble methods like random forests?**

Pierre doesn’t know. He likes the principle but not sure how well it will work

1. **What models/techniques are typically used?**

Pierre doesn’t do motor decoding of neural spike trains. But for other spiking data, RRNs are often used. He’s not sure if mini-batches and weighted sampling would be good, but it is not uncommon.

1. **How to improve NN/regression models?**

Focus on one angle to see where the errors come from more clearly. Then study what the link between the x and y coordinates is for that angle.

1. **Thoughts on Bayesian modelling**

Pierre is excited about Bayesian modelling but he does not know about Bayesian methods in MATLAB. But it is promising since we are working with small data, and we are given a lot of constraints/priors. We can set priors of the x and y coordinates based on the mean trajectory to constrain the values they can take on which might result in a lower error/variance.

1. **Do we have enough data for nested CV?**

Pierre is not sure. We’ll have to try it out for ourselves.

1. **Representation of spike train as series of Poisson/Gaussian processes**

Realm of information geometry. It is mostly used for inferring the characteristics of a NN. In principle, it should hold as much information as the spike trains themselves but the main limitation would be the testing data being fed in increasing lengths. This is more commonly used in “science” as opposed to “engineering”.