Liaison CS Meeting Agenda – Thursday 10/31/24

1. Zach: discussion of change points (15 minutes)
2. Mehrezat: demonstration of GUI (graphical user interface) design (10 minutes)
3. Milo: presentation of confusion matrices for his model(s) (15 minutes)
4. Time for liaison items (20 minutes)

Meeting minutes:

Zach: Change point detection

* Transitions associated with change in signal amplitude or mean
* Looking for these changes is similar to how people label signals
* Create divisions so that mean/standard deviation within region is the same
* Algorithm requires initial estimate for number of transitions
* Dr. Backus question: how do detected states correspond to waveforms?
* The states detected by model are regions where mean and stdev (amplitude) are similar
* Dr. Cooper: Z or W is withdraw, program may have difficulties with this (very short)
* Dr. Reif question: do we impose a limit on the length of a state?
* Model may not perform well on transitions which aren’t abrupt
* Model only identifies transitions, not the types of the states
  + Working on model to identify these
* This is a simple version of a hidden markov model, which should perform better
  + Will take into account ordering of states
* Dr. Backus: can be challenging to decide whether to split each different waveform appearance, or lump them together
  + Pathway waveforms will be challenging, as model may split every time the appearance changes (a human labeler would lump together)
* Dr. Backus/Dr. Reif: other researchers have been considering hidden markov models for this problem as well — identifying legal/illegal transitions, labeling waveform types

Mehrezat: User interface development

* Dr. Reif: okay to have separate interface from engr team if it hinders progress to integrate, just need output of one to work with the other
* Dr. Backus: CS team from last year already developed software for visualizing waveform
* Interface will let user upload and visualize signal, waveform labels, and confidence for each waveform generated by model
* Important for user to see labels overlaid on the waveform and modify them, since model will never label the waveform perfectly
* Dr. Reif question: will the model learn from modifications made by humans?
* Zach: this will depend on the model used
* Dr. Reif: will be helpful to read CS final report from last year for user interface feedback
* Dr. Reif: would be good to have feature to add comments
* Dr. Backus: will be frustrating if the user has to remind the program of the same adjustment over and over: ideally it will be able to learn — perhaps a training mode?

Milo: deep learning update

* precision and recall may provide better way to quantify accuracy
  + precision: only return data points in a class
  + recall: identify all data points in a class
* dilated CNN: each layer looks at larger and larger regions
  + advantages: scalable in terms of memory, can look at entire data set
  + cons: looks at chunks, harder to have long-term memory
  + confusion matrix never predicts W
  + Ideally, confusion matrix has 100% along the diagonal
* UNet
  + layers progressively condense information
  + advantages: whole dataset at once, long term connections
  + cons: hard to scale for long datasets, might not be able to look at full length
  + might be able to split up data
  + NP: precision < recall, overclassifying
  + J/K: precision > recall, underclassifying
  + Weighting: next steps are to upweight minority classes and downweight edge points