**CS USDA-Auburn Meeting 4/17/25**

**Meeting Minutes**

Mehrezat:  
 GUI Updates:  
 The interface looks mostly the same, but we've added a few features:

* Savable settings
* Duration controls
* Scrollbar
* Baseline visibility

Dr. Backus:  
 If someone wants to more finely divide waveforms, how difficult would that be?

Mehrezat:  
 If you haven’t run the model yet—just opened a dataset—you can add colors. We can also add labels if none exist.

Lillian:  
 We’re working on delivering a single file that you can run directly on your computer. There are packages that help with that.

Dr. Cooper:  
 Will this need to run on personal computers?

Zach:  
 No, you’ll just double-click the file and it opens—very simple.

### 

### ML Updates

Zach:  
 Dr. Cooper noticed F1 scores weren’t calculated correctly. I’ve fixed that.

* Switched from averaging F1 across folds to calculating using raw data from all folds.
* Some numbers went down slightly, but no major changes.
* Trying another round of training with augmented data—currently in progress.

Milo:  
 Post-processing:  
 Now that models are working, we hope to finalize things before year-end.

* The semi-Markov model lets us factor in time spent in each state.
* My test took an hour, so I’m exploring faster alternatives.
* We noticed transitions in the document that don’t appear in our data or in the referenced paper.

Dr. Cooper:  
 We’d have to run experiments to validate those transitions.

Milo:  
 We’re considering two approaches:

1. Downweight unlikely transitions.
2. Trust the model entirely, even when the data differs.

We might add sliders so you can adjust trust levels dynamically.

* Good models take longer to run.
* The state diagram is mostly linear; only one cycle (M ↔ N).
* We could use a basic algorithm to find the best path from L/M to W. Is that important?

Dr. Cooper:  
 It’s hard to say. We’ll likely encounter new transitions.  
 We’re in a chicken-and-egg scenario—manual scoring may be needed every time something new appears.  
 We’d need those transitions in the data for the model to learn from them accurately.

Milo:  
 We’re considering:

1. A trust slider.
2. Setting a small, non-zero probability for all transitions.  
    Manual scoring helps you know what to expect.  
    The model struggles when data differs from its training distribution. It doesn’t truly “understand” waveforms.

Dr. Reif:  
 Having zero-probability transitions is problematic. Adding artificial transitions can help future model training, especially for new species. It should be mentioned in the methods section.

Zach:  
 The current models don’t require those artificial transitions.

Dr. Cooper:  
 This isn’t about the paper content.

Milo:  
 Right. This is more about making the model practically useful.

Dr. Cooper:  
 What about retraining the model?

Milo:  
 This tool wraps around the model—it’s separate from retraining. Retraining is a different process.

Zach:  
 We’re aiming to make retraining easy—even for non-ML folks.

Milo:  
 That’s our goal. Feed in the data, rerun the model. It takes time, but future students could do it quickly. We can wrap it in a script for easier retraining.

Zach:  
 Plan for rest of the semester:

1. Train the model and save settings.
2. Build post-processing tools into the GUI.  
    Also, we have a poster session coming up—we’ll draft a version soon.  
    We want to make sure all proper attributions are included.

### Paper Discussion

Dr. Backus:  
 Only remaining questions are for Prof. Hope. Can you stay on the call?

Dr. Cooper:  
 I have manuscript-related questions.  
 About probe splitting—used to prevent data leakage:  
 Does the GUI still split probes when running the models?

Zach:  
 Yes. When labeling a new file, there are two buttons to let users adjust probe handling. Each probe is labeled individually.

Dr. Cooper:  
 In the paper, we mention removing NP-level data for probe splitting. Do we need to discuss the GUI’s handling?

Zach:  
 We should clarify that the NP-level is removed—make it explicit.

Dr. Cooper:  
 What’s the role of precision and recall in interpreting results? Any input?

Zach:  
 Relative to what?

Dr. Cooper:  
 In the results section, I’m focusing on F1 score. How are things mislabeled in confusion matrices? How should I discuss precision/recall?

Milo:  
 Look at extremely low/high values—L is a good example.

Dr. Cooper:  
 You used “spectral analysis” in your writeup. What did you mean?

Zach:  
 I meant estimating a signal’s spectrum from its time-domain representation (e.g., frequency analysis). Probably not the best term for this paper.

Dr. Backus:  
 I can show you how to use FFT in Windaq to visualize that—helps move from time to frequency domain. We should consider adding FFT/autocad analysis to future versions of the program.

### 

### Methods Clarifications

Dr. Cooper:  
 You mention converting recordings to densely labeled CSVs. What was used?

Zach:  
 Same approach as Perry 2024. Used pandas in Python. It’s one way of doing it, but not essential to mention.

Dr. Cooper:  
 It’s been difficult writing methods. I need exact details, but papers in this field are often inconsistent. Please help me fill in the gaps—even if it means “hurting my feelings.”

Dr. Reif:  
 The paper needs to speak to CS-minded reviewers.  
 Make the workflow clear to non-CS biology collaborators too.

Zach:  
 We’ll summarize everything in our final report.

Dr. Hope:  
 We can help build a pipeline and include it in the paper. A process diagram would be helpful. Including specific tools (e.g., “used pandas”) is useful. More detail is better.

Dr. Cooper:  
 Should I send you what I’ve written?

Zach:  
 Yes, that'll give us a starting point and let us fill in the details.

Dr. Cooper:  
 Great. Final question—how do you standardize metrics with varying time window sizes? I’ll leave that section blank for now and send a draft soon. It’s coming along, but it's been challenging.

Milo:  
 CS papers often lack reproducibility. That’s part of the issue.

Dr. Backus:  
 Intimidating to think I’ll be working with a whole new group this summer. Lots of challenges ahead.

Dr. Cooper:  
 But we’re in good shape.

Zach:  
 Thanks, everyone!

