Week 3 Video 2

Data Synchronization and Grain-Sizes
You have ground truth training labels…

- How do you connect them to your log files?
- The problem of synchronization
- Turns out to be intertwined with the question of what grain-size to use
Grain-size

- What level do you want to detect the construct at?
Orienting Example

- Let’s say that you want to detect whether a student is gaming the system, and you have field observations of gaming.

- Each observation has an entry time (e.g. when the coder noted the observation), but no start of observation time.

- The problem is similar even if you have a time for the start of each observation.
Data

Monday 8am

Monday 3pm

Friday 3pm

Gaming

Not Gaming
Data

Monday 8am

Monday 3pm

Friday 3pm

Notice the gap; maybe students were off this day… or maybe the observer couldn’t make it
Orienting Example

- What grain-size do you want to detect gaming at?
  - Student-level?
  - Day-level?
  - Lesson-level?
  - Problem-level?
  - Observation-level?
  - Action-level?
Student level

- Average across all of your observations of the student, to get the percent of observations that were gaming
Student level

Monday 8am
Gaming

Monday 3pm

Friday 3pm
Not Gaming

5 Gaming
10 Not Gaming

This student is 33.33% Gaming
Student level

Monday 8am

Monday 3pm

Friday 3pm

5 Gaming

10 Not Gaming

This student is 33.33% Gaming
Seen early in behavior detection work, when synchronization was difficult (cf. Baker et al., 2004)

Makes sense sometimes
- When you want to know how much students engage in a behavior
- To drive overall reporting to teachers, administrators
- To drive very coarse-level interventions
  - For example, if you want to select six students to receive additional tutoring over the next month
Day level

- Average across all of your observations of the student on a specific day, to get the percent of observations that were gaming
Day level

Monday 8am

Monday 3pm

Friday 3pm

Monday 40%

Tuesday 0%

Wednesday 20%

Thursday 0%

Friday 40%
Affords finer intervention than student-level

Still better for coarse-level interactions
Lesson level

- Average across all of your observations of the student within a specific level, to get the percent of observations that were gaming.
Lesson level

Monday 8am
Lesson 1: 40% gaming

Monday 3pm

Friday 3pm
Lesson 2: 30% gaming
Notes

- Can be used for end-of-lesson interventions
- Can be used for evaluating lesson quality
Problem level

- Average across all of your observations of the student within a specific problem, to get the percent of observations that were gaming
Problem level

Monday 8am

Monday 3pm

Friday 3pm
Can be used for end-of-problem or between-problem interventions

Fairly common type of intervention

Can be used for evaluating problem quality
Challenge

- Sometimes observations cut across problems
- You can assign observation to
  - problem when observation entered
  - problem which had majority of observation time
  - both problems
Observation level

- Take each observation, and try to predict it
Observation level

Monday 8am

Monday 3pm

Friday 3pm

Gaming

Not Gaming
“Most natural” mapping

Affords close-to-immediate intervention

Also supports fine-grained discovery with models analyses
Challenge

- Synchronizing observations with log files
- Need to determine time window which observation occurred in
  - Usually only an end-time for field observations; you have to guess start-time
  - Even if you have start-time, exactly where in window did desired behavior occur?
- How much do you trust your synchronization between observations and logs?
  - If you don’t trust it very much, you may want to use a wider window
Challenge

- How do you transform from action-level logs to time-window-level clips?
  - You can conduct careful feature engineering to create meaningful features out of all the actions in a clip
  - Or you can just hack counts, averages, stdev’s, min, max from the features of the actions in a clip (cf. Sao Pedro et al., 2012; Baker et al., 2012)
You could also apply your observation labels to each action in the time window. And then fit a model at the level of actions treating actions from the same clip as independent from one another. Offers the potential for truly immediate intervention.
Unfortunately, building detectors at the action-level has not worked particularly well for my group.

We’ve tried it a few times.

Maybe you’ll find a clever way to make it work great.

And then you can make fun of me in a talk at some future academic conference…
Bottom-line

- There are several grain-sizes you can build models at
- Which grain-size you use determines
  - How much work you have to put in (coarser grain-sizes are less work to set up)
  - When you can use your models (more immediate use requires finer grain-sizes)

- It also influences how good your models are, although not in a perfectly deterministic way
Next Lecture

- Feature Engineering