

Identifying Student Behaviors Early in the Term for Improving Online Course Performance

1

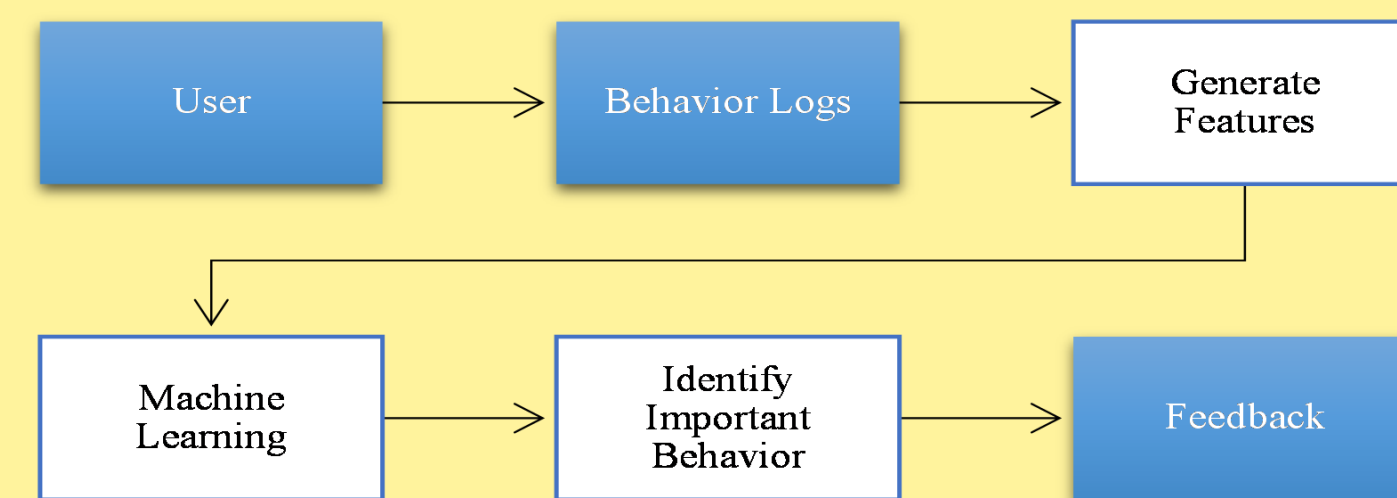
Makoto Mori & Philip Chan

Florida Institute of Technology

Goals and System Design

2

- Identify behaviors early in the course
- High-level behaviors (not just how much ,but how early/regular...)
- Behavior combinations



High-level Features

3

Feature	Description	Type
activity_coverage	Coverage of all types of activities	Syllabus
regular_day	Regular study schedule	Experience
study_days	Days studied for a test	Syllabus
test_submit_before_due	Days submitted test before due	Syllabus
(total/avg)days_after_unit_release	Days attempted units after unit released	Syllabus
(total/avg)video_after_unit_release	Days videos watched after unit released	Syllabus
(total/avg)video_before_asr	Times videos watched before ASRs	Syllabus
(total/avg)review_times	Correct materials reviewed before test	Syllabus
done_1st_form_test	Attempted 1st form of test	Syllabus
done_both_form_test	Attempted both forms of test	Syllabus

High-level Features from Sequential Patterns

4

Feature	Description
[assignment,assignment]	Two assignments occur sequentially
[test,unit,unit]	Access two units materials after test
[video,unit]	Access a unit material after video
[video,unit,unit]	Access two unit materials after video
[ASR,unit]	Access a unit material after ASR
[drill_submit,drill_review,unit]	Drill_review occurs after drill_submit, then unit occurs
[drill_review,unit]	Access a unit material after drill_review
[study_guide,unit]	Access a unit material after study_guide
[unit,study_guide,unit]	Access study_guide between two unit materials

Machine Learning – Random Forests

5

- Multiple decision trees
- Consider a random subsets of features at each node
- Cross validation to determine
 - Number of trees
 - Size of subsets

Identifying Important Behaviors from Forests

6

- Individual behaviors
 - Most frequent behavior in the root nodes
- Behavior combinations
 - feature combination f_i
 - tree r
 - positive examples $P_r(f_i)$
 - total number of examples $T_r(f_i)$
 - score: $s(f_i) = \sum_{r=1}^n \frac{P_r(f_i)}{T_r(f_i)}$

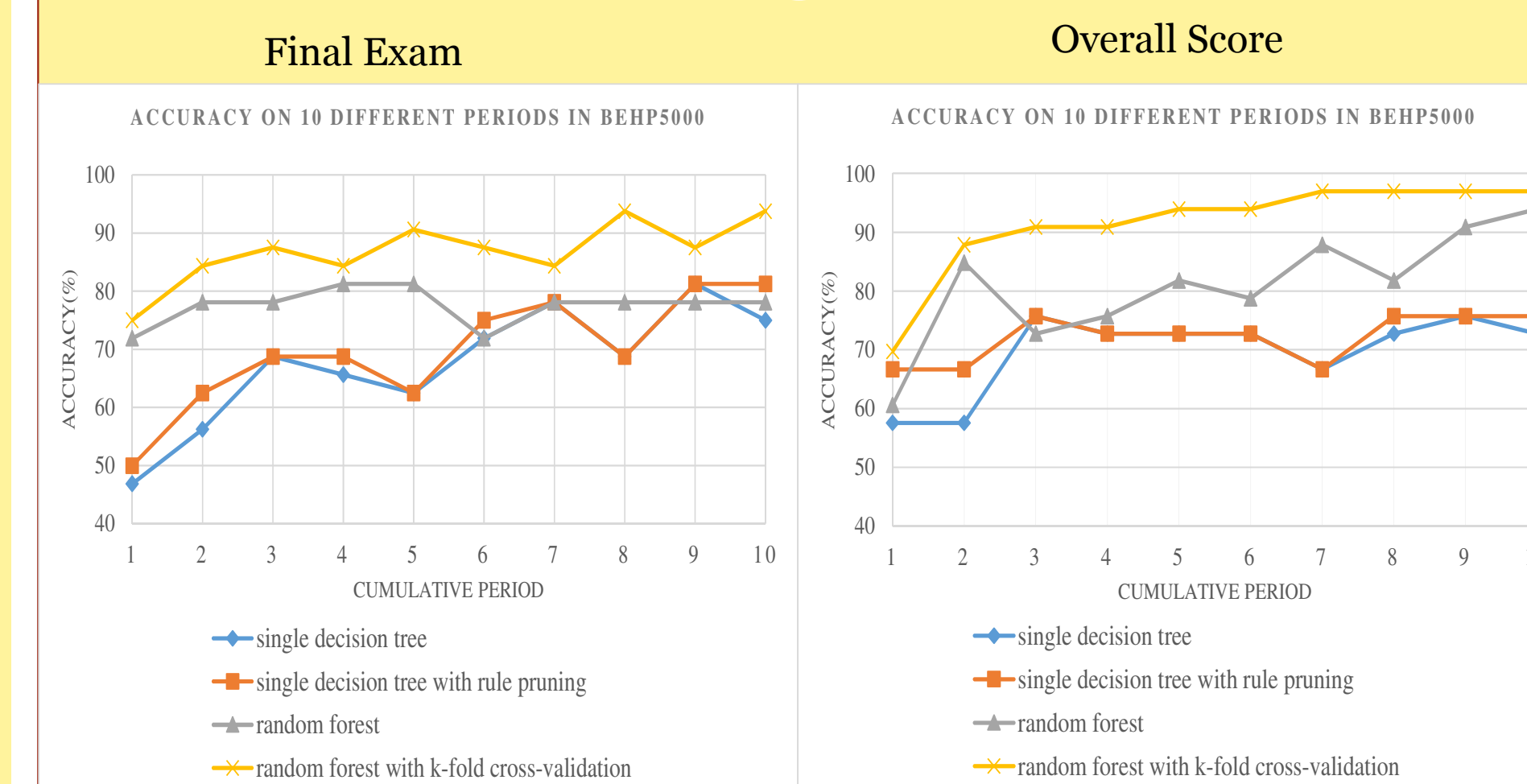
Evaluation

7

- “Concepts and Principles of Behavior Analysis”
- January to April 2013
- 7 course units
- 110 students

Predicting Performance

8



Important Behaviors

9

Detected important student behavior			
Period	NO.	final exam	overall score
period1	top. 1	avg(days_after_unit_release)	total(cyberarrat_times)
	top. 2	total(discussion_times)	total(discussion_times)
	top. 3	total_activities_one_log	total_activities_one_log
period2	top. 1	test_submit_before_due	total(login_times)
	top. 2	total(login_times)	min(ppt_hours)
	top. 3	avg(days_after_unit_release)	min(login_times)
period3	top. 1	total(days_after_unit_release)	total(days_after_unit_release)
	top. 2	total(asr_times)	total(unit_times)
	top. 3	test_submit_before_due	total(unit_hours)
period4	top. 1	total(discussion_times)	total(days_after_unit_release)
	top. 2	total(asr_times)	test_submit_before_due
	top. 3	total(days_after_unit_release)	min(meetinging)
period5	top. 1	total(days_after_unit_release)	total(days_after_unit_release)
	top. 2	total(asr_times)	total(login_times)
	top. 3	total(study_guide_times)	total(study_guide_times)

Important Behavior Combinations (Final Exam)

10

Period	Rank	Detected important behavior	Score	Delta
period1	top. 1	total(days_after_unit_release)>4.5^total(ask_question_times)<=3.5	291	60.2
	top. 2	total(days_after_unit_release)<=4.5^total(video_times)>3.5	231	
period2	top. 1	test_submit_before_due<=9.5^total(discussion_times)>1.0	217	62.6
	top. 2	total(days_after_unit_release)>17.0^test_submit_before_due<=9.5	155	
period3	top. 1	total(days_after_unit_release)>24.5^test_submit_before_due<=12.5	41.3	10.9
	top. 2	total(days_after_unit_release)<=24.5^total(asr_hours)>4.5	30.3	
period4	top. 1	total(discussion_times)>5.5^activity_coverage<=37.5	36.3	12.8
	top. 2	total(asr_times)<=4.5^[unit,study_guide,unit]=false	23.5	
period5	top. 1	total(days_after_unit_release)<=48.5^study_days<=17.0	59.6	15
	top. 2	total(days_after_unit_release)>48.5^total_activities_one_log>63.5	44.6	

Important Behavior Combinations (Overall Score)

11

Period	Rank	Detected important behavior	Score	Delta
period1	top. 1	total(cyberarrat_times)<=1.5^total(drill_times)<=2.5	50.6	7.25
	top. 2	total(discussion_times)<=8.5^total(study_guidehours)<=118.0	43.4	
period2	top. 1	total(login_times)>10.5^min(ppt_hours)>0.5	119	15.5
	top. 2	min(ppt_hours)>3.5^total(meeting_hours)<=110.5	104	
period3	top. 1	total(unit_times)>9.5^test_submit_before_due>7.5	26.5	2.72
	top. 2	total(days_after_unit_release)>22.5^total(discussion_times)>3.0	23.8	
period4	top. 1	total(days_after_unit_release)<=40.0^test_submit_before_due>=8.5	64.2	9.04
	top. 2	min(meetinging)<=2.5^done_both_form_test>5.0	55.2	
period5	top. 1	total(days_after_unit_release)<=48.5^	66.7	9.89
	top. 2	total(Supplemental_materials_hours)<=20.5	56.8	

Concluding Remarks

12

- ~70% accuracy in the 1st cumulative period
- ~90% accuracy in the 5th cumulative period
- total(days_after_unit_release) as the most important single behavior---also in combinations.
- For final exam, combination of
 - submitting a test closer to the due date and
 - participates more in the discussions.
- For overall score, combination of
 - login sessions and
 - spends more time on the powerpoint materials during each access.