The University of Hong Kong

Automated Analysis of Student Contributions and Behaviors

Xiao Hu

Faculty of Education

2015-09-23

Agenda

- Learning analytics
- Study 1: Automated analysis of student comments on Wikis of group writing projects
  - Refine coding scheme based on interviews with in-service teachers
  - Discover relationships among comment categories
  - Predict comment categories with automated means
- Study 2: Predicting student performance based on system logs in Learning Management Systems (LMS)
  - Build prediction models with feature selection
  - Apply models across years
  - Link LMS activities, assessment and learning outcomes

Learning Analytics

- “Uses intelligent data, learner-produced data, and analysis models to predict and advise on learning” (Siemens, 2010)
- “[LA] measures, collects, analyzes and reports data about learners and their contexts for the purposes of understanding and optimizing learning and learning environments (Ferguson, 2012).
- My work: **using automated means to analyze learning: the outcomes, the process, and the context**
  - Text mining of student discussions/reflections/writings
  - Predicting student performances based on behavior data
  - Building tools for monitoring and optimizing learning process

Study 1: Automated Analysis of Student Comments on Wikis of Group Writing Projects

Collaborators: Dr. Sam Chu; Ms. Christy Cheong

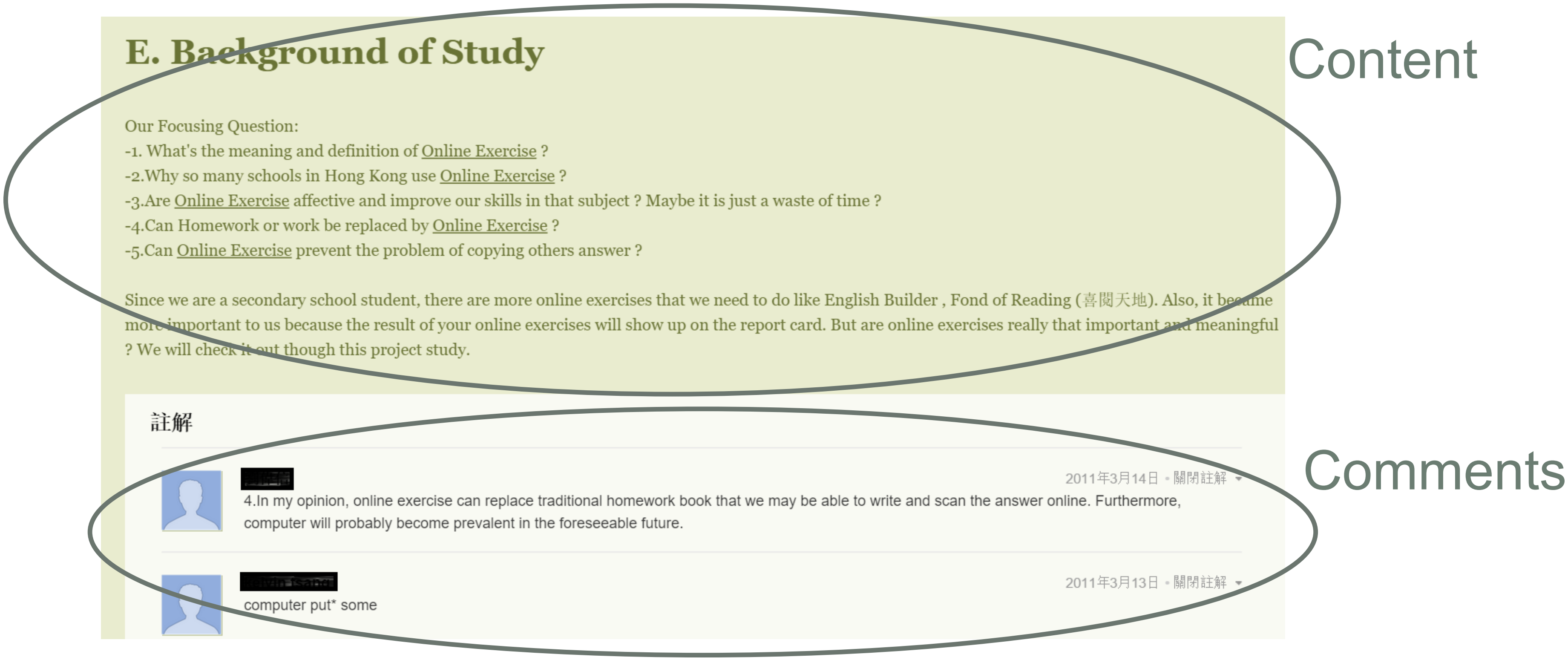
Study 1: Automated analysis of student comments on Wikis of group writing projects

- Why Wikis? Why Comments?
- Goals and Research Questions
- Study Context
- Coding Scheme Development
- Content Analysis
- Relationships Between Comment Categories
- Automated Categorization of Comments
- Discussions and Conclusions

Why Wikis?

- Wikis can facilitate project-based learning (PjBL) activities
  - PjBL requires continuous assessment and monitoring of student interaction and performance
  - Wikis support collaborative writing, record page revision history
- Wikis may significantly increase teachers' workload
  → likely to discourage the adoption (Kear, Donelan & Williams, 2014).
- Need ways to facilitate assessment of student learning in Wiki context

Why Comments?



- Comments:
  - A means of communication among students during learning
  - Reflect student dynamics
  - Under studied compared to content in Wikis

Goal and Research Questions

- Goals
  - Understanding student comments on Wikis of group writing
  - Exploring the feasibility of automating the analysis of student comments on Wikis
- Research Questions
  - What categories of comments did the student post on Wikis during group writing?
  - Are there any relationships among the categories of student comments?
  - To what extent can automated methods be used to categorize students' comments on Wikis?

Study Context

- A local secondary school
- Students collaborated in groups of 4 – 5 on an inquiry-based project for their Liberal Studies course over a five-month period
- Each group wrote project report on Google Sites
- Comments posted on each Wiki page are analyzed

	Overall	Form 1	Form 2
Total number of students	238	148	90
Total number of groups	48	30	18
Number of groups studied	**40**	**23**	**17**
Number of comments	962	621	341
Total number of units	1,528	1,062	466
Number of units analyzed	**1,482**	**1,056**	**426**

Coding Scheme Development

- Started from the literature of student online asynchronous discussions in three aspects
  - Social Interaction (SI) (Bales, 1950; Tirado, Aguaded, & Hernando, 2011)
    - Give suggestions; ask for opinions; ask for help; agree; disagree; encouragement; etc.
  - Thinking Purpose (TP) (Pena-Shaff & Nicholls, 2004)
    - Clarification; consensus building; judgment; reflection; support; etc;
  - Thinking Development (TD) (Bloom et al., 1956)
    - Knowledge; comprehension; application; analysis; synthesis, evaluation
- Refined by interviews with seven in-service teachers
  - What they need to know about student comments on Wikis?
  - How they would like to modify the scheme based on practical needs?

Themes from the interviewees

- Automation is helpful but it is necessary to have fewer categories
- Some categories (e.g., reflection) were rare in comments
- It was necessary to combine code categories similar in nature to make the scheme more practically feasible.
  - ask for information/opinions/suggestions/help => ask
  - clarification; interpretation; assertion; judgement => arguments
  - knowledge; comprehension; application => low cognitive development
  - analysis; synthesis, evaluation => high cognitive development
- Knowing the overall picture of students' performance would be sufficient in view of teachers' heavy workload

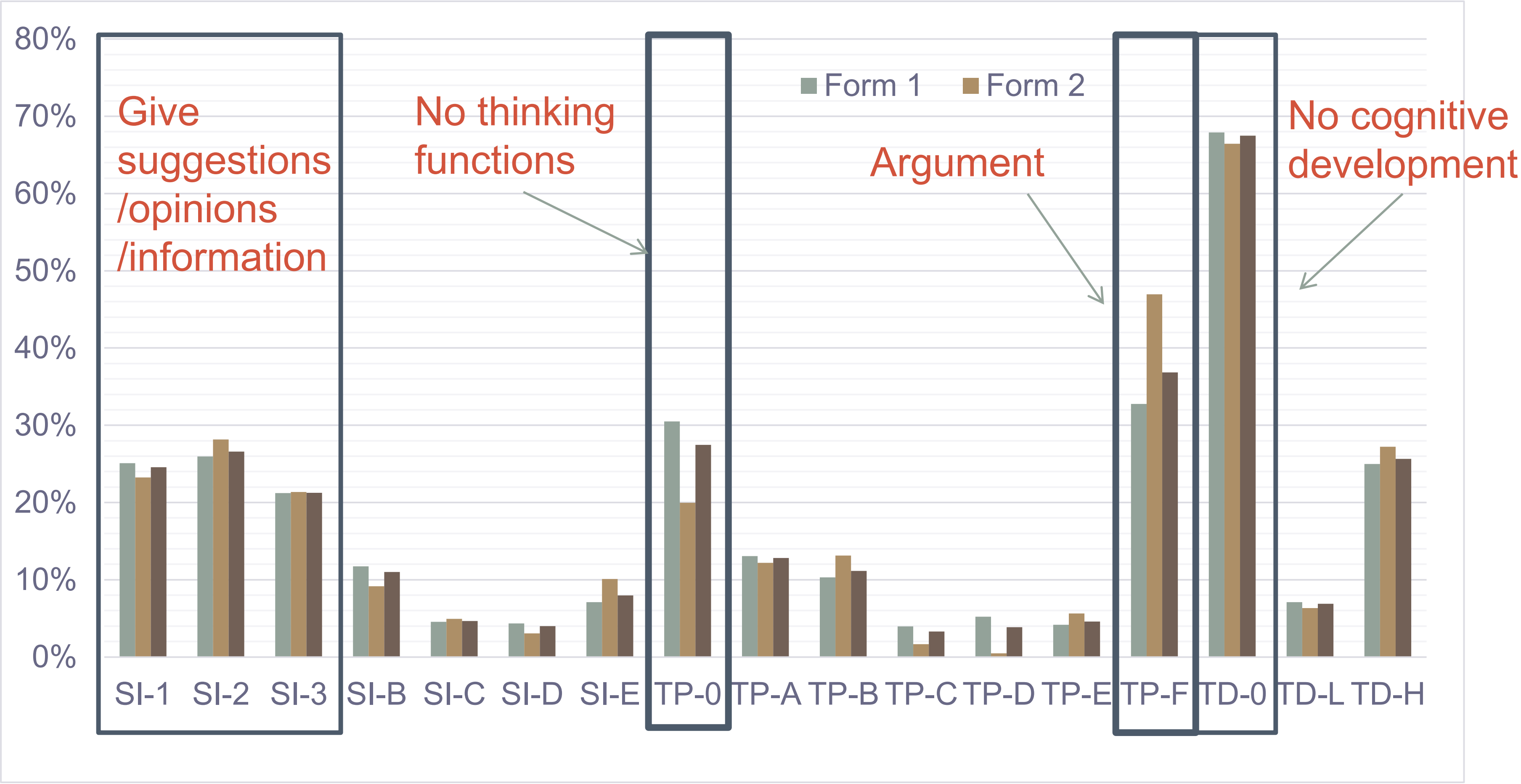
Content Analysis

- Final Scheme:

Module	No. of cat.	Categories
Social Interaction (SI)	7	N/A; give suggestions; give opinions; give information; ask; agree; disagree; others
Thinking Purpose (TP)	7	N/A; questions; simple replies; conflict; results; argument; others
Thinking Development (TD)	3	N/A; low cognitive development; high cognitive development

- Interrater Reliability
  - Double coded comments in 1/4 of the sample groups = 14 groups
  - Cohen's kappa
    - SI: κ =.69, TP: κ =.63, TD: κ =.64,
    - Total κ =.76
    - a satisfactory/excellent level of agreement (Altman,1991, Cicchetti, 1994)

Code Distribution



Mann-Whitney's tests between two forms: no significant difference on all categories

Relationship between Categories

- Association Rule Mining
  - discover relationships among multiple variables at the same time
  - "if – then" rules: "*antecedent -> consequence*"
  - *"buying diapers -> buying beers"*
  - FP-Growth algorithm (Han, 2010)

Group	Association rules	Support	Confidence	Lift	Cozine
SI-1, TP-0, TD-0	TD-0, SI-1 -> TP-0	0.13	0.61	2.21	0.54
SI-3, TP-0, TD-0	**SI-3 -> TD-0, TP-0**	0.13	0.61	2.24	0.54
SI-2, TP-F, TD-H	**TP-F -> SI-2**	0.15	0.56	3.01	0.67
	SI-2 -> TP-F, TD-H	0.15	0.58	2.65	0.63
	TD-H -> TP-F, SI-2	0.22	0.60	2.25	0.71
	TP-F, SI-2 -> TD-H	0.15	0.68	2.65	0.63

Automated Categorization

- Categorization models
  - Logistic Regression
  - Support Vector Machines (SVM)
  - Naïve Bayesian
- Text analytic features (Mayfield, Adamson, & Rosé, 2014)
  - Basic features: bag-of-words counts, part-of-speech, line length, punctuations
  - Stretchy patterns: n-gram with gaps
    - *"this [GAP] reasonable"*
  - Context patterns:
    - position in discussion: *first*, *last*, *middle, only*
    - similarity to other comments in the page
    - Reference to other information (links, quotations): *yes, no*
- 10-fold cross-validation on 1,424 units

Algorithms	Features	Social Interactions (SI)		Thinking Purpose(TP)		Thinking Development (TD)	
		Accuracy	Kappa	Accuracy	Kappa	Accuracy	Kappa
Logistic Regression	Basic	0.668	0.473	0.660	0.517	0.742	0.323
	Basic + Stretchy patterns	0.719	0.579	**0.712**	**0.607**	**0.810**	0.564
	Basic + Context features	0.665	0.471	0.653	0.509	0.750	0.357
	Basic + Stretchy patterns + Context features	**0.725**	**0.589**	0.709	0.603	[illegible]	[illegible]54
Naïve Bayes	Basic	0.554	0.419	0.534	0.424	[illegible]	[illegible]38
	Basic + Stretchy patterns	0.565	0.431	0.546	0.436	0.704	0.446
	Basic + Context features	0.555	0.419	0.533	0.423	0.698	0.440
	Basic + Stretchy patterns + Context features	0.565	0.431	0.545	0.435	0.706	0.448
SVM	Basic	0.664	0.507	0.626	0.499	0.761	0.480
	Basic + Stretchy patterns	0.697	0.560	0.666	0.554	0.804	**0.569**
	Basic + Context features	0.664	0.507	0.629	0.503	0.759	0.478
	Basic + Stretchy patterns + Context features	0.698	0.698	0.663	0.663	0.803	0.803

Human coder Agreement: κ =.63

Human agreement: SI: κ =.69, TP: κ =.63, TD: κ =.64

Confusing Categories

Prediction Label	TP-0	TP-A	TP-B	TP-C	TP-D	TP-E	TP-F
TP-0	291	3	13	3	1	2	80
TP-A	11	163	0	1	1	0	11
TP-B	36	1	90	0	0	2	30
TP-C	6	3	1	5	0	1	29
TP-D	11	1	0	1	11	0	32
TP-E	15	2	6	0	0	39	5
TP-F	70	8	12	7	4	1	415
Total	440	181	122	17	17	45	602

TP-C: conflict; TP-D: results (consensus building, reflection); TP-F: argument

Cross-form Categorization

	Features	SI		TP		TD	
		Acc.	*κ*	*Acc.*	*κ*	*Acc.*	*κ*
Form 1, → Form 2	Basic + stretchy patterns	0.614	0.426	**0.705**	**0.570**	**0.727**	**0.368**
	Basic + stretchy patterns + context features	**0.634**	**0.462**	0.697	0.559	0.725	0.365
Form 2, → Form 1	Basic + context features	0.618	0.395	0.592	0.441	0.717	0.270
	Basic + stretchy patterns + context features	0.618	0.395	0.593	0.446	0.714	0.262

Models trained on Form 1 data can be applied to Form 2 data with moderate agreement with human coder.

Discussions

- Dominance of TP-0 (no thinking purpose) and TD-0 (no cognitive development)
  - Improved design of commenting functions on Wiki platforms
  - Early intervention from teachers
- Relationship between categories
  - May further improve prediction
- Misclassifiied/confusing categories
  - Data sparseness (“conflict” had only 45 instances)
  - Too fine granularity?
- Generalizability across Forms
  - Practical use in handling cases with insufficient training data

Conclusions and Future Work

- Wiki comments are worth teachers' attention but could equally consume their efforts
- Data mining methods can help analyze and manage student comments, providing teachers with practical implications in an economical fashion to devise timely and appropriate learning support
- Future work direction 1: to verify results in larger datasets
- Future work direction 2: to automate analysis of comments written in Chinese

Study 2: Predicting Student Performance Based on System Logs in Learning Management Systems (LMS)

Collaborators: Dr. Leon C. U. Lei; Dr. Gaowei Chen; Prof. Nancy Law; Prof. Ricky Kwok; Ms. Peggy Chi; Ms. Jessica Wong; Mr. Chen Qiao

Study 2: Predicting Student Performance Based on System Logs in Learning Management Systems (LMS)

- Goals and Research Questions
- Study Context
- Prediction paradigm 1: Assessment tasks
- Prediction paradigm 2: Moodle “activities”
- Discussions

Goals and Research Questions

- Goals:
  - Derive a scientific and efficient method to monitor student learning behaviors on LMS (Moodle)
  - Develop a Moodle Plugin to help instructors and students monitoring learning progress
- Research Questions
  - Can event logs on Moodle be used to predict student performance?
  - Can Moodle logs be used to estimate student learning progress towards learning outcomes?

Study Context

- One Common Core course in HKU, Two years' Moodle logs

Year	2013	2014
No. students	104	152
No. of log events	94K	151K

User	Time	Module	Action	URL	Info
10115	2013.9.27 9:30	course	view	?id=1234	CCST1234
10109	2013.9.29 19:15	forum	post	?id=203	Dis. forum
10101	2013.10.12 12:10	wiki	edit	?id=229	Group wiki

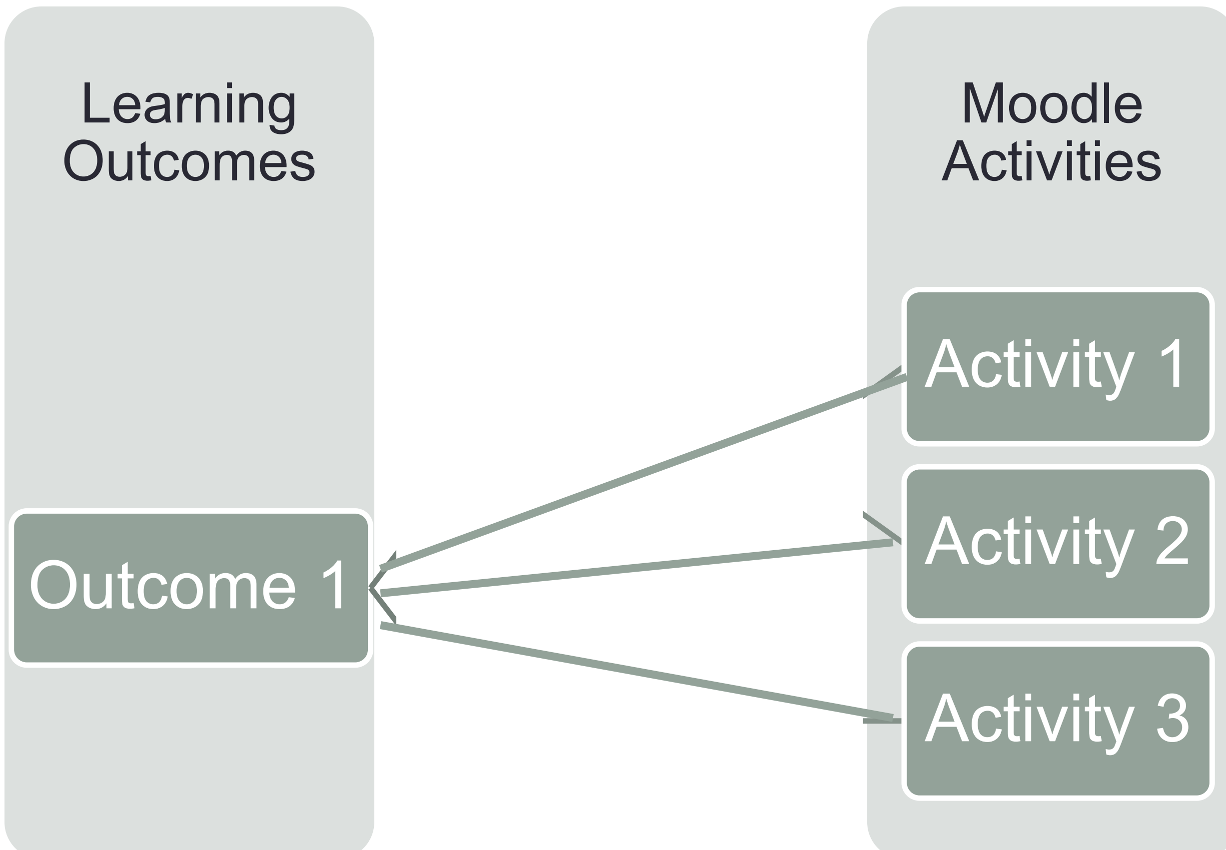
- 17 different modules: forum, quiz, wiki,…
- 53 types of actions: wiki view, add, update, post, edit,…

Predicting Performances of Assessment Tasks

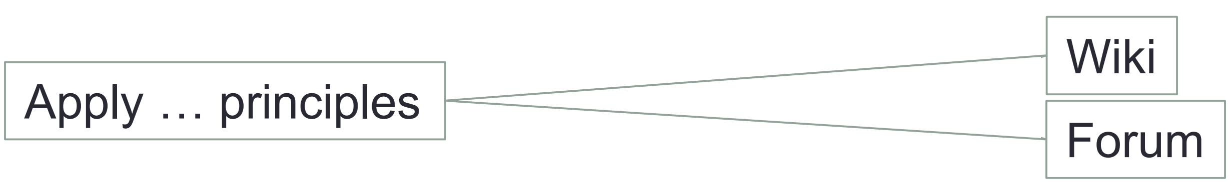
- Predict student overall performances and performance of various assessment tasks
  - Homework; Quiz; Tutorial,…
- 90 module-action features/variables
  - Course-view; quiz-attempt, questionnaire-submit, ….
- Linear regression with stepwise backwards feature selection

		Over all	Home work	Quiz	Tuto rial	Group Wiki	Group Pres.	Indiv. Essay	Indiv. Pres.
2013 N = 104	No. of feat.	7	8	1	9	9	8	5	5
	R^2	0.43	0.40	0.37	0.15	0.25	0.21	0.09	0.12
2014 N= 152	No. of feat.	14	13	11	12	15	11	10	13
	R^2	0.66	0.54	0.59	0.33	0.40	0.36	0.23	0.24
13->14	R^2	0.27	0.20	0.51	0.05	0.02	0.02	0.00	0.04

Predicting Performances of Moodle “Activities”



Activities:
- Forum
- Quiz
- Wiki
- Assignment
- Resources
- Feedback
- …



Predicting Performances of Moodle "Activities"

- Features
  - Module-action counts
  - Time-based: e.g., quiz attempt duration; lag time of first view, etc.
- Linear regression with stepwise backwards feature selection

		Multi-attempt Quiz	Single-attempt Quiz	Wiki	Assignment
2013 N = 104	No. of feat.	8	3	1	1
	R^2	0.21	0.10	0.08	0.06

Discussion

- Prediction of assessment task performances
  - Worked well
    - Some tasks can be predicted across years
  - May not be generalizable across courses, or different designs of the same course
- Prediction of Moodle activity performances
  - Potentially generalizable across courses
    - All course Moodle consists of activities
  - May be challenging to obtain accurate models

Summary

- Two studies on automated analysis of student learning based on evidences they left in e-learning platforms
- Goals: to design methods to obtain reliable indicators of learning progress in an efficient manner
- Techniques: association rule mining; categorization, prediction, etc.
- Connect to learning and teaching practice
- Some encouraging results, still more challenges
- Future directions
  - Evaluate actual impacts on learning and teaching
  - Contribute to fundamental questions in science of learning

THANK YOU !
APPRECIATE COMMENTS & SUGGESTIONS

xiaoxhu@hku.hk

References

- Altman, D.G. (1991). Practical statistics for medical research. London: Chapman and Hall.
- Aydin, Z., & Yildiz, S. (2014). Using wikis to promote collaborative EFL writing. *Language Learning & Technology*, *18*(1), 160-180.
- Bales, R. F. (1950). A set of categories for the analysis of small group interaction. *American Sociological Review*, *15*(2), 257-263.
- Berry, K. J., & Mielke, P. W. JR. (1988). A generalisation of Cohen's kappa agreement measure to interval measurement and multiple raters. *Educational and Psychological Measurement*, *48*(4), 921-933.
- Bloom, B. S., Engelhart, M. D., Furst, E. J., Hill, W. H., & Krathwohl, D. R. (Eds.). (1956). *Taxonomy of educational objectives: The classification of educational goals. Handbook 1: Cognitive domain*. London, WI: Longmans, Green & Co. Ltd.
- Castaneda, D. A., & Cho, M. H. (2013). The role of wiki writing in learning Spanish grammar. *Computer Assisted Language Learning*, *26*(4), 334-349.
- Cicchetti, D. V. (1994). Guidelines, criteria, and rules of thumb for evaluating normal and standardized assessment instruments in psychology, *Psychological Assessment*, Vol. 6, pp. 284-290.
- Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, *20*(1), 37-46.
- Cullen, R., Kullman, J., & Wild, C. (2013). Online collaborative learning on an ESL teacher education programme. *ELT Journal*, *67*(4), 425-434.
- Kear, K., Donelan, H., & Williams, J. (2014). Using wikis for online group projects: Student and tutor perspectives. *International Review of Research in Open and Distance Learning*, *15*(4), 70-90.
- LAK. (2011). The first International Conference on Learning Analytics and Knowledge 2011. Available at https://tekri.athabascau.ca/analytics.
- Li, X., Chu, S. K. W., & Ki, W. W. (2014). The effects of a wiki-based collaborative process writing pedagogy on writing ability and attitudes among upper primary school students in Mainland China. *Computers & Education*, *77*, 151-169.
- Lo, H. C. (2013). Design of online report writing based on constructive and cooperative learning for a course on traditional general physics experiments. *Educational Technology & Society*, *16*(1), 380-391.
- Mayfield, E., Adamson, D., & Rosé, C. P. (2014). LightSide Researchers' Workbench: User manual. Retrieved 12 November 2015 from http://ankara.lti.cs.cmu.edu/side/LightSide_Researchers_Manual.pdf.
- Merceron, A., & Yacef, K. (2008). Interestingness measures for associations rules in educational data. *Educational Data Mining*, 8, 57-66.
- Pena-Shaff, J. B., & Nicholls, C. (2004). Analyzing student interactions and meaning construction in computer bulletin board discussions. *Computers & Education*, *42*(2004), 243-265.
- Tirado, R., Aguaded, I., & Hernando, A. (2011). Collaborative learning processes in an asynchronous environment: An analysis through discourse and social networks. *Journal of Latin American Communication Research*, *2*(1), 115-146.
- Vach, W. (2005). The dependence of Cohen's kappa on the prevalence does not matter. *Journal of Clinical Epidemiology*, *58*(7), 655-661.

References

- Siemens, G. (2010) What are learning analytics, retrieved on September. 20, 2013 from http://www.elearnspace.org/blog/2010/08/25/what-are-learning-analytics/
- Ferguson, R. (2012). Learning analytics: drivers, developments and challenges. *International Journal of Technology Enhanced Learning*, 4 (5/6), 304
- Macfadyen, L. P., & Dawson, S. (2010). Mining LMS data to develop an “early warning system” for educators: A proof of concept. *Computers & Education*, *54*(2), 588-599.
- Romero, C., Ventura, S., & García, E. (2008). Data mining in course management systems: Moodle case study and tutorial. *Computers & Education*, *51*(1), 368-384.
- Romero, C., Espejo, P. G., Zafra, A., Romero, J. R., & Ventura, S. (2013). Web usage mining for predicting final marks of students that use Moodle courses. *Computer Applications in Engineering Education*, *21*(1), 135-146.
- Falakmasir, M. H., & Habibi, J. (2010, June). Using educational data mining methods to study the impact of virtual classroom in e-learning. In *Educational Data Mining 2010*.
- Taylor, C., Veeramachaneni, K., & O'Reilly, U. M. (2014). Likely to stop? Predicting stopout in massive open online courses. *arXiv preprint arXiv:1408.3382*.
- Porter, G. W., King, J. A., Goodkin, N. F., & Chan, C. K. (2012). Experiential learning in a common core curriculum: Student expectations, evaluations, and the way forward. *International Education Studies*, *5*(3), p24.
- Romero, C., Ventura, S., & De Bra, P. (2009). Using mobile and web-based computerized tests to evaluate university students. *Computer Applications in Engineering Education*, *17*(4), 435.
- Cole, J., & Foster, H. (2007). *Using Moodle: Teaching with the popular open source course management system*. " O'Reilly Media, Inc.".
- Kohavi, R., & John, G. H. (1997). Wrappers for feature subset selection. *Artificial intelligence*, *97*(1), 273-324.
- Kohavi, R. (1995, August). A study of cross-validation and bootstrap for accuracy estimation and model selection. In *Ijcai* (Vol. 14, No. 2, pp. 1137-1145).
- Kuhn, M. (2008). Building predictive models in R using the caret package. *Journal of Statistical Software*, *28*(5), 1-26.
- Kim, J. H. (2009). Estimating classification error rate: Repeated cross-validation, repeated hold-out and bootstrap. *Computational Statistics & Data Analysis*, *53*(11), 3735-3745.
- Ambroise, C., & McLachlan, G. J. (2002). Selection bias in gene extraction on the basis of microarray gene-expression data. *Proceedings of the national academy of sciences*, *99*(10), 6562-6566.
- Veeramachaneni, K., O'Reilly, U. M., & Taylor, C. (2014). Towards feature engineering at scale for data from massive open online courses. *arXiv preprint arXiv:1407.5238*.
- Wang. Y. (2014). Using wikis to facilitate interaction and collaboration among EFL learners: A social constructivist approach to language teaching. *System*, *42*, 383-390.
- Weltzer-Ward, L. (2011). Content analysis coding schemes for online asynchronous discussion. *Campus-Wide Information Systems*, *28*(1), 56-74.