INFO 4100/5101 Learning Analytics Syllabus Spring 2024



Spring 2024

CORNELL FUTURE OF LEARNING LAB

Cornell Bower College of Computing and Information Science

# INFO 4100 / 5101: Learning Analytics

**Lecture** (75 minutes)
Mondays 2:55PM – 4:10 PM Olin Hall 155

| **Sections** (75 minutes) | | | Hollister Hall |
|---|---|---|---|
| 201 | Tuesday | 2:55PM – 4:10PM | 314 |
| 205 | Wednesday | 8:40AM – 9:55AM | 314 |
| 206 | Thursday | 8:40AM – 9:55AM | 314 |
| 203 | Thursday | 2:55PM – 4:10PM | 314 |
| 202 | Friday | 11:40AM – 12:55PM | ~~206~~ moved to 401 |
| 208 | Friday | 11:40AM – 12:55PM | 401 |
| 207 | Friday | 1:25PM – 2:40PM | 401 |
| 204 | Friday | 2:55PM – 4:10PM | 401 |

**Student Question Sessions**

# INFO 4100/5101 Learning Analytics Syllabus Spring 2024



~~Slack Zoom~~

**Teaching Team**

Prof. René Kizilcec (he/him; kizilcec@cornell.edu)
Scott Allen (he/his; PhD candidate in Physics)
Jinsook Lee (she/her; PhD student in Information Science)
Yan Tao (she/her; PhD student in Information Science)
Lincy Chen (she/her; Senior majoring in ILR)
Lucia Zacek (she/her; Senior majoring in CS / Statistics)
Mitra Farzami (she/her; Senior majoring in IS)
Dave Tirtariyadi (he/him; Senior in Chemical Engineering / Data Science)
Salma Hazimeh (she/her; Senior majoring in IS)
Gemma Petrulla (she/her; Senior majoring in ISST)

For questions, send us a message on **Slack** in #help or a direct message.

**Course LMS** (we use EdX instead of Canvas in this course)
https://edge.edx.org/courses/course-v1:CornellX+INFO4100+SP24/about

**Slack Group** (we use Slack for all communication in this course)
info4100-sp24.slack.com - join the group using the link on EdX

**Course Description**

Technology has transformed how people teach and learn today. It also offers unprecedented insight into the mechanics of learning by collecting detailed interaction and performance data, such as in online courses and learning

learning. This course blends learning theories and methodologies covering a wide range of topics with weekly hands-on activities and group projects using real-world educational datasets. You will learn how learning works, major theories in the learning sciences, and data science methods. Students collect and analyze their own learning trace data as part of the course. Learning outcomes: Students will learn to articulate key ideas in the learning sciences; articulate the potential benefits and dangers of learning analytics for students, teachers, and institutions; choose and apply appropriate methods for analyzing different kinds of educational data and be able to articulate why; and interpret the results of basic learning analytics.

**Prerequisites**

This course is for undergraduate juniors, seniors, and graduate students interested in learning, education technology, educational data mining, and the broader implications of technology and data in education. Prior knowledge of probability and statistics (random variables, probability distributions, statistical tests, p values), data mining techniques (regression, clustering, prediction models), and fundamentals of programming is strongly recommended. Prior experience with the statistical programming language *R* is also recommended, as you will analyze data sets in R throughout this course.

**Learning Objectives**

analytics to analyze different kinds of educational data.
3. Evaluate the results of different methods for different applications.
4. Compare the strengths and weaknesses of methods for different applications.
5. Identify potential benefits and risks of learning analytics for students, teachers, and institutions.

**Course Structure and Grading**

Each week of the course covers a different topic and you have one lecture meeting, one lab section, readings and homework on that topic. There is **no midterm or final exam**. The key to success in this course is to stay on top of the material, put in consistent effort each week, and ask for help early.

*Individual/group work:* You will work both individually and in study groups during the course. All students will be assigned into **persistent study groups** at the start of the course based on their section, preferences, background, and interests. Students are encouraged to work together and help each other within their assigned groups but not between groups. Every student needs to submit their own version of each assignment. It is not permitted to copy someone else's code, except for group projects where you can split up tasks and everyone submits the same final product (see Academic Integrity section).

*Slack for communication:* Instead of email or Canvas, we will use **Slack** for all course communication

questions about homework, course logistics, etc., and a #reading-reflections channel to post your reflection. The teaching team monitors Slack daily and will respond in a timely fashion.

*EdX:* Course materials are released weekly on the EdX platform. The type of content released and schedule will be consistent for most weeks.

**Sunday morning, 8am ET:** readings, videos, homework released

**Saturday night, 11:59pm ET:** reading comprehension/ reflection, homework due

*Attendance* is expected in lecture meetings and lab sections and counts for 10% of the total grade. You are responsible for submitting your weekly attendance on edX by Saturday 11.59pm ET under the Attendance section. We will perform random checks and deduct grade points if attendance is marked incorrectly.

*Readings* are posted in the form of a PDF chapter from the *Handbook of Learning Analytics*, academic articles and reports, or video presentations.

*Reading comprehension* questions ask about the reading or video; they are auto-graded for correctness (no partial credit) and count for 10% of the total grade. In the last 2-3 weeks of the course, we will be using a new AI-based tool to support reading comprehension, which will be substituted for the comprehension quiz.

confirm that you completed the assignment; they are graded in terms of completion (0/1) and count for 10% of the total grade. We will perform random checks and deduct grade points if a reading reflection is marked as complete but not posted.

*Homework* typically involves analysis of a dataset using methods introduced in the same week. The dataset and starter code will be posted and you upload your 'knitted' response report as a Word Doc(x) on edX. Homework counts for 55% of the total grade. Solutions are released after the deadline has passed. No late submissions are accepted once the solution is released.

*Group projects* are more complex homework assignments that students collaborate on; they count for 15% of the total grade.

Students enrolled in graduate-level INFO 5101 will be held to a higher standard in grading homework and group project submissions.

The grade computation is performed automatically in edX (see breakdown below).

| Lecture and Section Attendance | Reading Comprehension | Reading Reflections | Homework | Group Projects |
|---|---|---|---|---|
| 10% | 10% | 10% | 55% | 15% |

**How to Get Help**

Students can get help in several ways. The fastest and most reliable way to get help is to ask your question on the Slack #help

message on Slack to the instructor or Head TA. Besides Slack, you can talk to the instructor or TAs after class/section, and come to one of the student-question sessions.

Please note that questions asked on Slack late in the evening are likely not going to be answered until the next day. To help you 24/7 with any questions, this semester we are trying out a course-specific learning and homework support chatbot. More information to come.

**Academic Integrity**

Academic honesty is a serious matter and will be treated as such. Academic dishonesty harms trust within our community and penalties can therefore be severe. Each student in this course is expected to abide by the Cornell University Code of Academic Integrity at all times. Any work submitted by a student in this course for academic credit must be that student's own work. All outside assistance should be noted and the work of others should be properly cited. For more information, check out Cornell's academic integrity resources. If you have questions about what is allowed or noticed academic dishonesty, you can always reach out to the instructor and teaching team on Slack.

This course is participating in Accepting Responsibility (AR), which is a pilot supplement to the Cornell Code of Academic Integrity (AI). For details about the AR process and how it supplements the AI Code, see the AR website.

**You are free to use generative artificial intelligence (GenAI) tools**

help with writing or debugging code. What matters most is that you use these tools in combination with your critical thinking skills. Don't simply accept the answer– probe it, question it, ask for more explanation if it is still unclear.

**Accommodations and Counseling Resources**

Let me know as soon as possible if you require special accommodations due to learning disabilities, religious practices, physical requirements, medical needs, etc. We are committed to creating an inclusive learning environment for all students. It is not uncommon for students to experience stressful events at some point during college. Students sometimes experience anxiety, depression, financial strain, the loss of loved ones, family stress, and other stressors. It is perfectly normal for students to seek the service of mental health professionals to provide them with support and skills to cope with these experiences. Cornell Health is offering a variety of services to students on and off campus: https://health.cornell.edu/services/mental-health-care. You can also call Cornell Health at (607) 255-5155.

**General Readings**

Assigned reading introduces the topic of the week and the methods for analysis.

- Lang, C., Siemens, G., Wise, A., & Gašević, D. (2017). The Handbook of Learning Analytics – First Edition. SoLAR. https://www.solaresearch.org/hla-17/

- Ambrose, S. A., Bridges, M. W., DiPietro, M., Lovett, M. C., & Norman, M. K. (2010). *How learning works: Seven research-based principles for smart teaching*. John Wiley & Sons.


**Detailed Class Schedule** (subject to changes)

## WEEK 1: INTRODUCTION

22 January: Overview & Getting Set Up

***Lecture***
Meet the teaching team; why learning analytics matters; course outline, logistics, and systems. Advice from previous students.

***Section***
Form study groups; get to know each other; help with R set up and test script; complete the course survey.

***Homework*** *(due Saturday at 11:59pm)*
1. Access Edx Edge course site and review content
2. Join Slack group and post a self-introduction
3. Set up R and RStudio, run test script, answer self-assessment, complete R tutorials

## WEEK 2: OVERVIEW OF LEARNING ANALYTICS

29 January: Learning analytics and educational data mining; data types and overview of methods

***Lecture***
Systems, stakeholders, and data in Learning Analytics and Educational

look like? What is unique about educational data? What methods are commonly used to analyze data in education? What is the difference between descriptive and inferential analysis?

***Section***
Review R basics; go through a few examples; work on R homework.

***Reading***
1. Watch Society of Learning Analytics video
2. Fisher et al., Mining Big Data in Education: Affordances and Challenges

***Reading Comprehension Quiz and Reflection Post*** *(due Saturday at 11:59pm)*
*Reflection Prompt:* What are 3 things that you learned from the readings that you would tell someone who has not read them? Comment on someone else's reflection post to highlight an interesting take-away that you had not previously thought of.

***Homework*** *(due Saturday at 11:59pm)*
1. Load a dataset into R and create a report with basic descriptive statistics using starter code posted online
2. Explore the posted dataset and answer basic questions about it (make plots, compute differences, compute correlations, engineer features)
3. Self-assess your progress

WEEK 3: ETHIC**AL CONSIDERATIONS**

5 February: Ethical considerations of learning analytics, including algorithmic fairness.

***Lecture***

***Section***

Discuss readings. Example question to consider: If Cornell engaged in learning and curriculum analytics specifically to improve its academic programs and courses, is there any data that would make you feel uncomfortable if it were used (socio-demographic information, grades, Canvas clickstream, id card swipes with location, web traffic)? If so, why and what question would you ask the (hypothetical) Cornell Learning Analytics team about how your data is used?

***Readings***

1. Watch keynote address at LAK 2018 conference by Neil Selwyn
2. Selwyn, N. (2020). Re-imagining 'Learning Analytics' … a case for starting again? The Internet and Higher Education, 46(February), 100745.
3. Handbook of Learning Analytics 1st Ed, Ch.4, Ethics and Learning Analytics: Charting the (Un)Charted
4. Kizilcec, R. F., & Lee, H. (2020). Algorithmic Fairness in Education.

***Reading Comprehension Quiz and Reflection Post*** *(due Saturday at 11:59pm)*

*Reflection Prompt:* What are 3 things that you learned from the readings that you would tell someone who has not read them? Comment on someone else's reflection post to highlight an interesting take-away that you had not previously thought of.

***First Group Project***

In your study group, pick a chapter of How Learning Works for the group project, start reading it, and create slides.

What do we know about how learning works? Insights from Learning Science research.

***Section***
Finish presentation slides in your group and record a 10-min video presentation.

***First Group Project continued*** *(due Saturday at 11:59pm)*
Record a 10-min video presentation about your chosen chapter of How Learning Works; upload it on YouTube and share it on Slack.

WEEK 5: CAUSAL INFERENCE

19 February: Experiments, randomization, and causal inference: theory and practice.

***Lecture***
Counterfactual thinking, policy evaluation, and how to perform robust random assignment. Randomized experiments are the most effective way to estimate the effects of educational interventions. We will review why they work.

***Section***
How to design your own experiment and analyze the results.

***Reading***
1. J Pearl, The Book of Why, Ch.1
2. Handbook of Learning Analytics 1st Ed., Ch.18, Diverse Big Data and Randomized Field Experiments in MOOCs

***Homework*** *related to Week 4 (due Friday at 11:59pm)*
1. Watch a group presentation video for each Principle of How Learning Works
2. *Video Reflection Prompt:* Translate these broad principles of how learning works into concrete

learning or meet some need, (2) how you would tailor it to the specific context and/or adjust the course to accommodate this change, and (3) how would you assess if it is working as expected. Submit your response on Slack and leave a constructive/encouraging comment on another student's response.

**WEEK 6: CAUSAL INFERENCE CONT.**

*NOTE:* No lecture due to February break but sections

***Reading Comprehension Quiz and Reflection Post*** *related to Week 5 (due Saturday at 11:59pm)*

*Reflection Prompt:* What are 3 things that you learned from the readings that you would tell someone who has not read them? Comment on someone else's reflection post to highlight an interesting take-away that you had not previously thought of.

***Homework*** *related to Week 5 (due Saturday at 11:59pm)*
Analyze data from an experiment

**WEEK 7: MULTIMEDIA LEARNING AND VIDEO ANALYTICS**

4 March: Multimedia learning; Video analytics

***Lecture***
Multimedia learning theory is about how people learn with different content and how content should therefore be designed. Video in

video analysis and homework support.

***Reading***

1. R C Clark & R E Mayer, e-Learning and the Science of Instruction, Ch.2 and 4
2. Handbook of Learning Analytics 1st Ed., Ch.22, Video Analytics
3. Guo, P. J., Kim, J., & Rubin, R. (2014). How video production affects student engagement: An empirical study of MOOC videos. In Proceedings of the first ACM conference on Learning@ scale conference (pp. 41-50).
4. Juho Kim's lecture on Learnersourcing

***Reading Comprehension Quiz and Reflection Post*** *(due Saturday at 11:59pm)*

*Reflection Prompt:* What are 3 things that you learned from the readings that you would tell someone who has not read them? Comment on someone else's reflection post to highlight an interesting take-away that you had not previously thought of.

***Homework*** *(due Saturday at 11:59pm)*
Analyze video analytics data.

WEEK 8: PREDICTIVE MODELING

11 March: Overview of predictive modeling uses and approaches; Prediction in R

***Lecture***
Predicting learner behavior and learning outcomes is a cornerstone of learning analytics. There are various methods to choose from depending on the prediction task and available data. What teaching/learning techniques can support learning? How can

***Reading***

1. Gašević, D., Dawson, S., & Siemens, G. (2015). Let's not forget: Learning analytics are about learning. TechTrends, 59(1), 64-71.
2. Handbook of Learning Analytics, Ch.3, Predictive Modelling in Teaching and Learning

***Reading Comprehension Quiz and Reflection Post*** *(due Saturday at 11:59pm)*

*Reflection Prompt:* What are 3 things that you learned from the readings that you would tell someone who has not read them? Comment on someone else's reflection post to highlight an interesting take-away that you had not previously thought of.

***Homework*** *(due Saturday at 11:59pm)*

Compare the performance of several predictive models on two datasets.

***Second Group Project***

Review instructions and communicate within your team to set a shared timeline with clear expectations

WEEK 9: Unsupervised Learning

18 March: Guest Lecture TBD. Learn about clustering and complexity reduction to identify groups of similar students

***Lecture***

Unsupervised learning: findings patterns in data using unsupervised methods such as cluster analysis is commonly used to understand how learning behaviors and performance differs across groups of students.

~~Reading (subject to change based~~ *on guest lecture)*

1. kMeans clustering introduction video
2. Khalil & Ebner (2017), Clustering patterns of engagement in Massive Open Online Courses (MOOCs): the use of learning analytics to reveal student categories
3. Ferguson & Clow (2015), Examining Engagement: Analysing Learner Subpopulations in Massive Open Online Courses
4. Hierarchical Clustering introduction

***Reading Comprehension Quiz and Reflection Post*** *(due Saturday at 11:59pm)*

*Reflection Prompt:* What are 3 things that you learned from the readings that you would tell someone who has not read them? Comment on someone else's reflection post to highlight an interesting take-away that you had not previously thought of.

***Second Group Project continued*** *(due Saturday at 11:59pm)*
Complete project and submit project report.

WEEK 10: SELF-REGULATED LEARNING

25 March: Self-regulated learning (SRL); Measuring and supporting SRL strategies

***Lecture***
We will review the literature on self-regulated learning (SRL), the phases involved, and specific self-regulation strategies. We will also learn about how to measure and detect SRL.

***Section***

Gunderson, B. (2017). Strategic resource use for learning: A self-administered intervention that guides self-reflection on effective resource use enhances academic performance. *Psychological Science, 28*(6).

3. Watch video interview with Prof. Patricia Chen about the reading
4. Watch a video interview with Dr. Heeryung Choi about analyzing self-regulated learning

***Reading Comprehension Quiz and Reflection Post*** *(due Friday at 11:59pm)*

*Reflection Prompt:* What are 3 things that you learned from the readings that you would tell someone who has not read them? Comment on someone else's reflection post to highlight an interesting take-away that you had not previously thought of.

***Homework*** *(due Friday at 11:59pm)*

1. Keep a diary of your own SRL activities
2. Analyze a dataset to identify evidence of self-regulation

**- SPRING BREAK -**

WEEK 11: ASSESSMENT

8 April: Measurement and validation; Modeling learning with Bayesian knowledge tracing, deep knowledge tracing

***Lecture***

Assessment in the context of learning analytics. What exactly are the data produced by learning environments measuring? How to measure learning? Developing robust assessments. BKT and IRT.

Analytics

2. Wang, J., & Bao, L. (2010). Analyzing force concept inventory with item response theory. American Journal of Physics, 78(10), 1064-1070.
3. Video interview with Prof. Neil Heffernan about BKT and ASSISTments
4. Optional reading on BKT mentioned in the interview: Pardos, Z. A., & Heffernan, N. T. (2010). Modeling individualization in a bayesian networks implementation of knowledge tracing. In *International Conference on User Modeling*, Adaptation, and Personalization (pp. 255-266).

***Reading Comprehension Quiz and Reflection Post*** *(due Saturday at 11:59pm)*

*Reflection Prompt:* What are 3 things that you learned from the readings that you would tell someone who has not read them? Comment on someone else's reflection post to highlight an interesting take-away that you had not previously thought of.

***Homework*** *(due Saturday at 11:59pm)*
Evaluate the psychometric properties of a standardized assessment.

## WEEK 12: EMOTIONAL LEARNING ANALYTICS

### 15 April: Emotion and affect in learning

***Lecture***
Emotional learning analytics; what is affect, why does it matter, how can it be used?

***Section***

Ch.12, Emotional Learning Analytics
3. Baker et al., Towards Sensor-Free Affect Detection in Cognitive Tutor Algebra
4. Hutt et al., Gaze-based Detection of Mind Wandering during Lecture Viewing

***Reading Comprehension Quiz and Reflection Post*** *(due Saturday at 11:59pm)*

*Reflection Prompt:* What are 3 things that you learned from the readings that you would tell someone who has not read them? Comment on someone else's reflection post to highlight an interesting take-away that you had not previously thought of.

***Homework*** *(due Saturday at 11:59pm)*
Train an affect detector using the dataset provided.

WEEK 13: LA DASHBOARDS

22 April: Communicating learning analytics to different stakeholders with visualizations using a dashboard

***Lecture***
Dashboards are commonly used to communicate your learning analytics insights to different stakeholders like students and course instructors. What makes a good dashboard? How do you create one? Students work in groups to create student- and instructor-facing dashboards. Using visualizations to communicate data effectively. Transparency, trust, actionability in dashboards.

***Reading***

Learning Analytics Driving Teacher's Attention
4. Watch R Shiny and Plotting with ggplot2 tutorials

***Section***
Work on the group project with support.

***Reading Comprehension Quiz and Reflection Post*** *(due Saturday at 11:59pm)*

*Reflection Prompt:* What are 3 things that you learned from the readings that you would tell someone who has not read them? Comment on someone else's reflection post to highlight an interesting take-away that you had not previously thought of.

***Third Group Project***
Review project instructions and communicate within your team to set a shared timeline with clear expectations As a team, create a dashboard for clicker data using R shiny. Create a plan and a sketch to get feedback from teaching staff during the section.

WEEK 14: **CURRICULUM ANALYTICS**

29 April: Analytics in higher education, modeling college pathways; machine learning to solve logistical problems in higher education.

***Lecture***
College students make choices about which courses to choose. What information informs these choices and how can macro-level learning analytics in higher education shape college pathways.

Marketplace (15min)
2. Kizilcec et al. (2023), "From Pipelines to Pathways in the Study of Academic Progress
3. Chaturapruek et al. (2021), "Examining Undergraduate Course Consideration at Scale" AERA Open.

***Reading Comprehension Quiz and Reflection Post*** *(due Saturday at 11:59pm)*

*Reflection Prompt:* What are 3 things that you learned from the readings that you would tell someone who has not read them? Comment on someone else's reflection post to highlight an interesting take-away that you had not previously thought of.

***Third Group Project continued*** (due *Saturday at 11:59pm)*
Submit a write-up about your design choices with screenshots and your code.

**WEEK 15: WRAP-UP**

6 May: Course review and feedback.

***Lecture***
Looking back and looking forward, what you learned this semester, and what you can now accomplish with your new knowledge and skills.

***Homework*** *related to Week 14 (due Saturday at 11:59pm)*
Analyze course sequence data to visualize pathways and/or identify toxic course pairings.

**End Of Course -- No Final Exam**

INFO 4100/5101 Learning Analytics Syllabus Spring 2024