Course Schedule
EDUC6191: Core Methods in Educational Data Mining
Fall 2023
Professor Ryan Baker

Class 1: Introduction

August 31 & September 5
Readings

- Baker, R.S. (2023) Big Data and Education. Ch. 1, V1.
- Baker, R.S.J.d., Yacef, K. (2009) The State of Educational Data Mining in 2009: A Review and Future Visions. Journal of Educational Data Mining, 1 (1), 3-17. [pdf]
- Baker, R., Siemens, G. (2022) Educational data mining and learning analytics. In Sawyer, K. (Ed.) Cambridge Handbook of the Learning Sciences: 3rd Edition. [pdf]

Slides: [pptx]
Assignments Due: NONE

Class 2: Prediction Modeling and Algorithmic Bias

September 7 & 12
Readings

- Baker, R.S. (2023) Big Data and Education. Ch. 1, V2, V3, V4, V6, V7. Ch. 2, V7.
- Deho, O. B., Joksimovic, S., Li, J., Zhan, C., Liu, J., & Liu, L. (2022). Should Learning Analytics Models Include Sensitive Attributes? Explaining the Why. IEEE Transactions on Learning Technologies. [pdf]
- Baker, R.S., Esbenshade, L., Vitale, J.M., Karumbaiah, S. (in press) Using Demographic Data as Predictor Variables: a Questionable Choice. To appear in Journal of Educational Data Mining. [pdf]
- Baker, R.S., Hawn, M.A. (2022) Algorithmic Bias in Education. International Journal of Artificial Intelligence and Education, 32, 1052-1092. [pdf]

Slides: [pptx]

Assignments Due: NONE

Class 3: Behavior and Affect Detection

September 14 & 19
Readings

- Baker, R.S. (2023) Big Data and Education. Ch.1, V5. Ch. 2, V5. Ch. 3, V1, V2, V7.
- Botelho, A.F., Baker, R., Heffernan, N. (2017) Improving Sensor-Free Affect Detection Using Deep Learning. Proceedings of the 18th International Conference on Artificial Intelligence in Education, 40-51. [pdf]
- Hutt, S., Grafsgaard, J. F., & D'Mello, S. K. (2019). Time to scale: Generalizable affect detection for tens of thousands of students across an entire school year. In Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems (pp. 1-14).[pdf]

- Zhang, J., Andres, J.M.A.L., Hutt, S., Baker, R.S., Ocumpaugh, J., Mills, C., Brooks, J., Sethuraman, S., Young, T. (2022) Detecting SMART Model Cognitive Operations in Mathematical Problem-Solving Process. Proceedings of the International Conference on Educational Data Mining. [pdf]

Slides: [pptx]

Assignments Due: Basic: Classifier

Class 4: Diagnostic Metrics

September 21 & 26
Readings

- Baker, R.S. (2023) Big Data and Education. Ch. 2, V1, V2, V3, V4.
- Jeni, L. A., Cohn, J. F., & De La Torre, F. (2013). Facing Imbalanced Data--Recommendations for the Use of Performance Metrics. Proceedings of Affective Computing and Intelligent Interaction (ACII), 245-251.[pdf]
- Kitto, K., Shum, S. B., & Gibson, A. (2018). Embracing imperfection in learning analytics. In Proceedings of the 8th International Conference on Learning Analytics and Knowledge (pp. 451-460). ACM.[pdf]

Slides: [pptx]
Assignments Due: Creative: Behavior Detection

Class 5: Feature Engineering and Tweaking Towards Optimality

September 28 & October 3
Readings

- Baker, R.S. (2023) Big Data and Education. Ch. 3, V3, V4, V5, V6.
- Sao Pedro, M., Baker, R.S.J.d., Gobert, J. (2012) Improving Construct Validity Yields Better Models of Systematic Inquiry, Even with Less Information. Proceedings of the 20th International Conference on User Modeling, Adaptation and Personalization (UMAP 2012),249-260. [pdf]
- Slater, S., Baker, R. S., & Wang, Y. (2020). Iterative Feature Engineering through Text Replays of Model Errors. Proceedings of the International Conference on Educational Data Mining. [pdf]
- Sha, L., Gasevic, D., & Chen, G. (2023). Lessons from debiasing data for fair and accurate predictive modeling in education. Expert Systems with Applications. [pdf]

Slides: [pptx]
Extra Materials for Class:[docx]
Assignments Due: Basic: Diagnostic Metrics

Class 6: Network Analysis

October 5 & 10
Readings

- Baker, R.S. (2023) Big Data and Education. Ch. 5, V5, V6. Ch. 8, V2.
- Dawson, S. (2008) A study of the relationship between student social networks and sense of community. Educational Technology & Society, 11(3), 224-238.[pdf]
- Gasevic, D., Zouaq, A., & Janzen, R. (2013). "Choose Your Classmates, Your GPA Is at Stake!": The Association of Cross-Class Social Ties and Academic Performance. American Behavioral Scientist

[pdf]
- Pantic, N., Galey, S., Florian, L., Joksimovic, S., Viry, G., Gasevic, D., ... & Kyritsi, K. (2021). Making sense of teacher agency for change with social and epistemic network analysis. Journal of Educational Change, 1-33. [pdf]

Slides: [pptx]
Assignments Due: Creative: Feature Engineering

Class 7: Bayesian Knowledge Tracing

October 12 & 17
Readings

- Baker, R.S. (2023) Big Data and Education. Ch. 4, V1, V2, V5; Ch. 6, V8.
- Corbett, A.T., Anderson, J.R. (1995) Knowledge Tracing: Modeling the Acquisition of Procedural Knowledge. User Modeling and User-Adapted Interaction, 4, 253-278. [pdf]
- Kang, J., Baker, R., Feng, Z., Na, C., Granville, P., Feldon, D.F. (2022) Detecting threshold concepts through Bayesian knowledge tracing: examining research skill development in biological sciences at the doctoral level. Instructional Science. [pdf]

Slides: [pptx]
Assignments Due: Basic: SNA

Class 8: Logistic Knowledge Tracing

October 19 & 24
Readings

- Baker, R.S. (2023) Big Data and Education. Ch. 4, V3, V7.
- Pavlik, P.I., Cen, H., Koedinger, K.R. (2009) Performance Factors Analysis -- A New Alternative to Knowledge Tracing. Proceedings of AIED2009.[pdf]
- Choffin, B., Popineau, F., Bourda, Y., & Vie, J. J. (2019). DAS3H: Modeling Student Learning and Forgetting for Optimally Scheduling Distributed Practice of Skills. Proceedings of the International Conference on Educational Data Mining (EDM 2019). [pdf]
- Pavlik, P. I., Eglington, L. G., & Harrell-Williams, L. M. (2021). Logistic Knowledge Tracing: A Constrained Framework for Learner Modeling. IEEE Transactions on Learning Technologies, 14(5), 624-639. [pdf]

Slides: [pptx]
Assignments Due: Basic: BKT

Class 9: Deep Knowledge Tracing

October 26 & 31
Readings

- Baker, R.S. (2023) Big Data and Education. Ch. 4, V6.
- Yeung, C. K., & Yeung, D. Y. (2018). Addressing two problems in deep knowledge tracing via prediction-consistent regularization. In Proceedings of the Fifth Annual ACM Conference on Learning at Scale (p. 5-14). ACM.[pdf]
- Pandey, S., & Karypis, G. (2019). A Self-Attentive Model for Knowledge Tracing. Proceedings of the International Conference on Educational Data Mining. [pdf]

- Gervet, T., Koedinger, K., Schneider, J., & Mitchell, T. (2020). When is deep learning the best approach to knowledge tracing?. Journal of Educational Data Mining, 12(3), 31-54. [pdf]

Slides: [pptx] **Assignments Due:** Basic: PFA

Class 10: Association Rule Mining and Sequential Pattern Mining

November 2 & 7

Readings

- Baker, R.S. (2023) Big Data and Education. Ch. 5, V3, V4.
- Merceron, A., Yacef, K. (2008) Interestingness Measures for Association Rules in Educational Data. Proceedings of the 1st International Conference on Educational Data Mining,57-66. [pdf]
- Kinnebrew, J. S., Loretz, K. M., & Biswas, G. (2013). A contextualized, differential sequence mining method to derive students' learning behavior patterns. JEDM-Journal of Educational Data Mining, 5(1), 190-219.[pdf]
- Cukurova, M., Khan-Galaria, M., MillÃ¡n, E., & Luckin, R. (2022). A Learning Analytics Approach to Monitoring the Quality of Online One-to-one Tutoring. Journal of Learning Analytics, 1-16. [pdf]

Slides: [pptx]
Assignments Due: NONE

Class 11: Clustering

November 9 & 14

Readings

- Baker, R.S. (2023) Big Data and Education. Ch. 6, V1, V2, V3, V4, V5.
- Bowers, A.J. (2010) Analyzing the Longitudinal K-12 Grading Histories of Entire Cohorts of Students: Grades, Data Driven Decision Making, Dropping Out and Hierarchical Cluster Analysis. Practical Assessment, Research & Evaluation (PARE), 15(7), 1-18. [pdf]
- Lee, J., Recker, M., Bowers, A.J., Yuan, M. (2016). Hierarchical Cluster Analysis Heatmaps and Pattern Analysis: An Approach for Visualizing Learning Management System Interaction Data. Poster presented at the annual International Conference on Educational Data Mining (EDM). [pdf]

Slides: [pptx]
Assignments Due: Basic: Sequential Pattern Mining

Class 12: Knowledge Structure Discovery

November 16 & 21
Readings

- Baker, R.S. (2023) Big Data and Education. Ch. 6, V6, V7.
- Desmarais, M.C., Meshkinfam, P., Gagnon, M. (2006) Learned Student Models with Item to Item Knowledge Structures. User Modeling and User-Adapted Interaction, 16, 5, 403-434.[pdf]
- Desmarais, M. C., & Naceur, R. (2013). A matrix factorization method for mapping items to skills and for enhancing expert-based Q-Matrices. Proceedings of the International Conference on Artificial Intelligence in Education, 441-450. [pdf]
- Koedinger, K.R., McLaughlin, E.A., Stamper, J.C. (2012) Automated Student Modeling Improvement. Proceedings of the 5th International Conference on Educational Data Mining, 17-24.[pdf]

- Chen, P., Lu, Y., Zheng, V. W., Chen, X., & Yang, B. (2018). KnowEdu: a system to construct knowledge graph for education. IEEE Access, 6, 31553-31563.[pdf]

Slides: [pptx]

Assignments Due: Basic: Clustering

No Class Thanksgiving Break

November 23

Note that Online Class is now before In-Person Class

Class 13: Correlation Mining

November 28 & 30
Readings

- Baker, R.S. (2023) Big Data and Education. Ch. 5, V1.
- Slater, S., Ocumpaugh, J., Baker, R., Scupelli, P., Inventado, P.S., Heffernan, N. (2016) Semantic Features of Math Problems: Relationships to Student Learning and Engagement. Proceedings of the 9th International Conference on Educational Data Mining, 223-230.[pdf]
- Matayoshi, J., & Karumbaiah, S. (2021). Investigating the Validity of Methods Used to Adjust for Multiple Comparisons in Educational Data Mining. Proceedings of the International Conference on Educational Data Mining. [pdf]

Slides: [pptx]
Assignments Due: Creative: Knowledge Structure

Class 14: Prompt Engineering

December 5 & 7

Readings

- Baker, R.S. (2023) Big Data and Education. Ch. 7, V1, V2.
- DAIR.AI (2023) Prompt Engineering Guide. [html]
- Nardo, C. (2023) The Waluigi Effect. [html]
- Grogan, J.C. (2023) JamesGPT (Just Accurate Markets Estimation System) [html]
- Moraco, S. (2023) [pdf]

Assignments Due: Basic: Correlation Mining

Class 15: Reinforcement Learning

December 12 & 14

Readings

- Baker, R.S. (2023) Big Data and Education. Ch. 8, V3, V4.
- Rafferty, A., Ying, H., & Williams, J. (2019). Statistical consequences of using multi-armed bandits to conduct adaptive educational experiments. Journal of Educational Data Mining, 11(1), 47-79. [pdf]

- Ju, S., Zhou, G., Barnes, T., & Chi, M. (2020). Pick the Moment: Identifying Critical Pedagogical Decisions Using Long-Short Term Rewards. Proceedings of the International Conference on Educational Data Mining. [pdf]
- Bassen, J., Balaji, B., Schaarschmidt, M., Thille, C., Painter, J., Zimmaro, D., ... & Mitchell, J. C. (2020, April). Reinforcement learning for the adaptive scheduling of educational activities. Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems (pp. 1-12). [pdf]

Slides: [pptx]
Assignments Due: Creative: Prompt Engineering
