Sungjin Nam

namsungjinn@gmail.com

Fort Collins, CO 80524 LinkedIn | Google Scholar | GitHub

Applied Research Scientist

Understand and Support Users from Linguistic and Behavioral Data

Experienced applied scientist in educational technology, especially in evaluation, language learning, user engagement, and content improvements. Expertise in applied natural language processing (NLP) and machine learning (ML). Fluently delivers the message with quantitative analysis, visualizations, and prototyping in cross-functional team settings.

Professional Experience

ACT, Inc, Iowa City, IA July 2021 – Present

Senior Research Scientist

- Researching and developing CRASE+ automated essay scoring engine. CRASE+ is actively used in multiple high-stakes writing assessments, including ACT Writing, scoring 200k student responses per year (FY24). Contributing to multiple ML module developments, such as scoring engine, disturbing content classifier, and off-topic detector.
- Investigating large language model applications, including open-ended survey response analysis, solution step generations, reading passage identification, and essay writing tutoring, for content developers and students.
- Developing ML solutions for operation automatization, including form assembly and sales territory recommendation.

PEARSON, Centennial, CO June 2020 – May 2021

Data Scientist, Intelligent Systems Research and Development (Aida Calculus product)

- Created neural-network models to represent mathematical equations and find similar items or matching templates.
- Developed an IRT-based Bayesian network learner model to estimate students' skill proficiency.
- Implemented RESTful APIs for the app's popular features, such as practice, diagnostic dashboard, and hint selection.

Technical Skills

Programming Skills

- Python: PyTorch, scikit-learn, pandas, NumPy, Matplotlib/seaborn, NetworkX, Gensim, TensorFlow, Flask.
- R: LME4, tm, topicmodels, bnlearn, JAGS, ggplot, Shiny.
- Others: SQL, Git, Databricks, Docker, Java, AWS, Google Cloud, Bash shell.

Research Methods

- ML: Deep neural networks, classification, mixed-effect regression, clustering, active learning, curriculum learning.
- NLP: Contrastive learning, semantic representation, language modeling, topic modeling.
- Others: Visualization, crowdsourcing, A/B hypothesis testing, eye-tracking, mixed research methods.

Education

- **Ph.D.**, Information, University of Michigan, Ann Arbor, MI (Advisor: Prof. Kevyn Collins-Thompson) 2020
  - Thesis: Understanding Language Learning via Machine Learning Methods on Behavioral and Linguistic Data
- **M.Sc.**, Information, University of Michigan, Ann Arbor, MI 2014
  - Specialization in Information Analysis & Retrieval and Human-Computer Interaction

- **B.A.**, Psychology and Information Science & Culture, Seoul National University, Seoul, South Korea 2012

Publications

Peer-Reviewed Proceedings & Journals

- Yang, K. B., Nam, S., Huang, Y., & Wood, S. (2024, September). Rhetor: Providing LLM-Based Feedback for Students' Argumentative Essays. In *European Conference on Technology Enhanced Learning* (pp. 201-205). Cham: Springer Nature Switzerland.
- Nam, S., Collins-Thompson, K., Jurgens, D., & Tong, X. (2024, May). Finding Educationally Supportive Contexts for Vocabulary Learning with Attention-Based Models. In *Proceedings of the 2024 Joint International Conference on Computational Linguistics, Language Resources and Evaluation (LREC-COLING 2024)* (pp. 7286-7295).
- Yang, B., Nam, S., & Huang, Y. (2023). "Why My Essay Received a 4?": A Natural Language Processing Based Argumentative Essay Structure Analysis. In *International Conference on Artificial Intelligence in Education* (pp. 279-290). Cham: Springer Nature Switzerland. (Acceptance rate: 21.11%)
- Nam, S., Bylinskii, Z., Tensmeyer, C., Jain, R., Wigington, C., & Sun, T. (2020). Why are you reading this? Predicting reading goal and familiarity from people's mobile interaction behaviors. *Journal of Vision, 20*(11), 951-951. (Abstract)
- Nam, S., Samson, P. (2019). Integrating Students' Behavioral Signals and Academic Profiles in Early Warning System. In *International Conference on Artificial Intelligence in Education*, 345-357. (Acceptance rate: 25%)
- Nam, S., Frishkoff, G. A., & Collins-Thompson, K. (2018). Predicting Students' Disengaged Behaviors in an Online Meaning-Generation Task. In *IEEE Transactions on Learning Technologies 11* (3), 362-375.
- Nam, S., Collins-Thompson, K., & Frishkoff, G. A. (2017). Predicting Short- and Long-Term Student Learning with a Vocabulary Tutoring System via Semantic Features of Partial Word Knowledge. In *Educational Data Mining*, 80-87. (Acceptance rate: 25%)
- Waddington, R. J., Nam, S., Lonn, S., Teasley, S. D. (2016). Improving Early Warning Systems with Categorized Course Resource Usage. In *Journal of Learning Analytics 3* (3), 263-290.
- Nam, S. (2016). Predicting Off-task Behaviors in an Adaptive Vocabulary Learning System. In *Educational Data Mining*, 672-674. (Doctoral Consortium)
- Oh, J., Nam, S., & Lee, J. (2014). Generating highlights automatically from text-reading behaviors on mobile devices. In *CHI'14 Extended Abstracts on Human Factors in Computing Systems* (pp. 2317-2322).
- Nam, S., Lonn, S., Brown, T., Davis, C. S., & Koch, D. (2014). Customized course advising investigating engineering student success with incoming profiles and patterns of concurrent course enrollment. In *Proceedings of the Fourth International Conference on Learning Analytics and Knowledge* (pp. 16-25). ACM, 16-25. (Acceptance rate: 27%)
- Waddington, R. J., & Nam, S. (2014). Practice exams make perfect: incorporating course resource use into an early warning system. In *Proceedings of the Fourth International Conference on Learning Analytics and Knowledge* (pp. 188-192). ACM, 188-192. (Acceptance rate: 27%)

Book Chapters

- Nam, S., Collins-Thompson, K., & Frishkoff, G. A. (2017). Modeling Off-task Behaviors and Learning Outcomes on a Meaning-Generation Task. In *Analytics for Learning White Paper: Measurement in digital environments*.

- Frishkoff, G. A., Collins-Thompson, K., Nam, S., Hodges, L., & Crossley, S. (2016). Dynamic Support of Contextual Vocabulary Acquisition for Reading (DSCoVAR): An intelligent tutor for contextual word learning. In Crossley, S. A., & McNamara, D. S. (Eds.), *Adaptive Educational Technologies for Literacy Instruction*. N.Y., NY: Routledge.

Preprints

- Nam, S., Jurgens, D., & Collins-Thompson, K. (2022). An Attention-Based Model for Predicting Contextual Informativeness and Curriculum Learning Applications. *arXiv preprint arXiv:2204.09885.*
- Nam, S., Bylinskii, Z., Tensmeyer, C., Wigington, C., Jain, R., & Sun, T. (2020). Using Behavioral Interactions from a Mobile Device to Classify the Reader's Prior Familiarity and Goal Conditions. *arXiv preprint arXiv:2004.12016.*

Community Service

- Reviewer Service:
  - NLP: EMNLP, ACL, BEA
  - HCI: ACM CHI, ACII, UMUAI
  - Educational Technology: AIED, ACM LAK, ACM L@S, BJET, JATT, EM: I.P.
- Student Organizer of Academic Innovation at Michigan Analytics Talk Series (Jan 2016 – May 2019)
- Local Arrangement Chair for ACM SIGIR (2018)

Invited Talks

- Machine Learning Research in Educational Technology (October 2023) at the University of Wisconsin, Computer Science Department.
- Learning from Implicit Signals in Learning (August 2018) at Seoul National University, Department of Psychology.
- Understanding Online Behavior and Performance Signals in Educational Systems (April 2018) at the University of Oregon, Computer and Information Science Department.
- Predicting Short- and Long-Term Student Learning with a Vocabulary Tutoring System via Semantic Features of Partial Word Knowledge (June 2017) at Seoul National University, Department of Communication.
- Learning Analytics with Mass Data: Analyzing Higher Education's Academic Performance and Decision Making with Data (June 2014) at Seoul National University, Department of Psychology, and Department of Communication.

Public Presentations

- Nam, S., & Samson, P. (2018). Mining Students' In- and Out-of-class Behaviors to Create Earlier Warning System. In Michigan Institute for Data Science Annual Data Science Symposium.
- Nam, S., & Collins-Thompson, K. (2018). Toward Understanding Contextual Information in Text. In University of Michigan Data Science Learning Analytics Challenge Symposium.
- Nam, S., Collins-Thompson, K., & Frishkoff, G. A. (2017). Predicting Short- and Long-Term Student Learning with a Vocabulary Tutoring System via Semantic Features of Definition Responses. In University of Michigan Data Science Research Forum.
- Nam, S., & Collins-Thompson, K. (2017). Understanding Questions with Object and Action Terms: Bi-modal topic modeling of student questions from Piazza. Presented at Michigan Institute for Data Science Annual Symposium.

- Nam, S., Collins-Thompson, K., Frishkoff, G. A., Bhide, A., Muth, K. B., & Perfetti, C. (2016). Modeling Real-time Performance on a Meaning-Generation Task. Presented in AERA'16 Annual Meeting; Michigan Institute for Computational Discovery & Engineering Symposium; and Twenty-Second Annual Meeting of Society for the Scientific Study of Reading.
- Oh, J., Nam, S., & Lee, J. (2013). Enhancing Effectiveness of Skim Reading with Highlighting Text. Presented in HCI Korea 2013.

Research Internship/Assistant Experience

ADOBE, San Jose, CA June – Nov 2019

Machine Learning Intern, Document Intelligence Lab (Mentor: Dr. Tong Sun)

- Designed an experiment to predict mobile device readers' prior knowledge levels and goal conditions from scrolling behaviors. The experiment included a crowdsourcing task with a web application, analyzing eye-tracking signals, user interviews, and developing a mixed-effect classifier.

ECHO 360, Ann Arbor, MI May – Aug 2018

Graduate Research Intern (Mentor: Dr. Perry Samson)

- Developed early warning prediction model from behavioral logs and academic profiles and suggested customized learning strategies for different student groups.
- Supervised undergraduate interns in developing software for transcription, text search, and text clustering.

UNIVERSITY OF MICHIGAN, Ann Arbor, MI Sep 2012 – May 2020

Graduate Research Assistant

- Developed a deep learning model to identify more instructive sentences for a vocabulary learning system and examined how sentences improved efficiency in training NLP models like word2vec or FastText.
- Created an interpretable semantic scale for text by projecting biases in semantic differential with word2vec.
- Implemented a prediction model for students' disengaged behaviors and intervention moments.
- Expanded a topic model with grammar dependency on online forum texts for richer representation.
- Analyzed students' successful learning behaviors, like exploring course-taking patterns from >10 years of enrollment data or behavioral logs from multiple STEM courses' learning management systems.