SEN YANG

+1 (848) 565–5991

Mountain View, CA — sy358@rutgers.edu

ABOUT ME

I am a Senior Machine Learning Engineer and Researcher with over 10 years of experience in machine learning, representation learning, and large-scale ML systems. I currently work at Google on the Search Intelligence team, where I lead efforts in building personalized retrieval systems and end-to-end retrieval optimization under production constraints. At Waymo, I led learning-based evaluation of autonomous driving systems, transitioning driver log analysis from rule-based pipelines to representation-driven models built on Waymo's foundation models. Earlier at LinkedIn, I built and scaled personalization and home feed ranking models, translating research ideas into measurable engagement and revenue impact. I obtained my Ph.D. from Rutgers University, where my research focused on learning-based workflow modeling, evaluation, and decision support systems, and I have published in top-tier ML and data mining venues.

[Homepage] [LinkedIn] [Google Scholar]

EXPERIENCE

Google Mountain View, CA
Senior Machine Learning Engineer, Search Intelligence Jun 2025 – Present

- Lead the design and development of large-scale personalized retrieval systems, focusing on representation learning and end-to-end retrieval quality under strict latency and reliability constraints.

Waymo Mountain View, CA
Senior Machine Learning Engineer, Simulation 2022 – 2025

- Led learning-based evaluation of Waymo Driver models, leveraging large-scale representation learning and foundation-model techniques to analyze driving logs and improve robustness and generalization beyond rule-based metrics.
- Built embedding-driven scenario search and few-shot analysis pipelines to support scalable evaluation, data discovery, and targeted debugging of driving behaviors.
- Set technical direction across multiple ML initiatives, balancing model complexity, data availability, evaluation fidelity, and deployment constraints in production systems.

LinkedIn Mountain View, CA
Senior Machine Learning Engineer, Feed Ranking 2018 – 2022

- Led development of large-scale personalization and feed ranking models, jointly training user and content representations over tens of millions of entities.
- Defined true-north metrics, explored new model architectures, and validated improvements through online A/B experiments with sustained impact on engagement and revenue.
- Earlier work focused on NLP-driven data standardization and knowledge graph construction, delivering significant gains in member growth and monetization.

Bell Labs (Internship) Murray Hill, NJ
Research Intern, Representation Learning for Time Series Summers 2017, 2018

- Conducted research on deep representation learning for time-series data using RNN-based autoencoders, analyzing the gap between reconstruction quality and downstream task performance.
- Identified failure modes of autoencoder-based representations and proposed hybrid deep and statistical feature learning approaches, improving robustness under oscillatory patterns.
- Mentors: Dr. Jin Cao, Dr. Hüseyin Uzunalioglu

Huawei Technologies Shenzhen, China
System Engineer, Wireless 2012 – 2013

EDUCATION

Rutgers University — New Brunswick, NJ
Ph.D. in Electrical and Computer Engineering — 2013 – 2018
Research Focus: Data Mining
Advisors: Dr. Ivan Marsic, Dr. Hui Xiong
GPA: 3.9 / 4.0

Nanjing University of Posts and Telecommunications — Nanjing, China
B.A. in Communication Engineering — 2008 – 2012

PUBLICATIONS

1. Keyi Li, Mary S. Kim, Wenjin Zhang, **Sen Yang**, Genevieve J. Sippel, Aleksandra Sarcevic, Randall S. Burd, Ivan Marsic.
Human Intention Recognition for Trauma Resuscitation: An Interpretable Deep Learning Approach for Medical Process Data. Journal of Biomedical Informatics, 161:104767, 2025.

2. Keyi Li, **Sen Yang**, Travis M. Sullivan, Randall S. Burd, Ivan Marsic.
ProcessGAN: Generating Privacy-Preserving Time-Aware Process Data with Conditional Generative Adversarial Nets. ACM Transactions on Knowledge Discovery from Data, 18(9):1–31, 2024.

3. Wenjin Zhang, Keyi Li, **Sen Yang**, Sifan Yuan, Ivan Marsic, Genevieve J. Sippel, Mary S. Kim, Randall S. Burd.
Focusing on What Matters: Fine-grained Medical Activity Recognition for Trauma Resuscitation via Actor Tracking. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), pp. 4950–4958, 2024.

4. Wenjin Zhang, Keyi Li, **Sen Yang**, Chenyang Gao, Wanzhao Yang, Sifan Yuan, Ivan Marsic.
MaskMatch: Boosting Semi-supervised Learning through Mask Autoencoder-driven Feature Learning. arXiv preprint arXiv:2405.06227, 2024.

5. Keyi Li, **Sen Yang**, Chenyang Gao, Ivan Marsic.
Rutgers Multimedia Image Processing Lab at SemEval-2023 Task 1: Text-Augmentation-based Approach for Visual Word Sense Disambiguation. Proceedings of the 17th International Workshop on Semantic Evaluation (SemEval), pp. 1483–1490, 2023.

6. Xin Dong, Yaxin Zhu, Yupeng Zhang, Zuohui Fu, Dongkuan Xu, **Sen Yang**, Gerard de Melo.
Leveraging Adversarial Training in Self-Learning for Cross-Lingual Text Classification. Proceedings of the 43rd ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR), pp. 1541–1544, 2020.

7. **Sen Yang**, Aleksandra Sarcevic, Richard A. Farneth, Shuhong Chen, Omar Z. Ahmed, Ivan Marsic, Randall S. Burd.
An Approach to Automatic Process Deviation Detection in a Time-Critical Clinical Process. Journal of Biomedical Informatics, 85:155–167, 2018.

8. Wangsu Hu, Zijun Yao, **Sen Yang**, Shuhong Chen, Peter J. Jin.
Discovering Urban Travel Demands through Dynamic Zone Correlation in Location-based Social Networks. European Conference on Machine Learning and Knowledge Discovery in Databases (ECML-PKDD), pp. 88–104, 2018.

9. **Sen Yang**, Weiqing Ni, Xin Dong, Shuhong Chen, Richard A. Farneth, Aleksandra Sarcevic, Ivan Marsic, Randall S. Burd.
Intention Mining in Medical Process: A Case Study in Trauma Resuscitation. IEEE International Conference on Healthcare Informatics (ICHI), pp. 36–43, 2018.

10. Ran He, **Sen Yang**, Jingyuan Yang, Jin Cao.
Automated Mining of Approximate Periodicity on Numeric Data: A Statistical Approach. ACM International Conference on Compute and Data Analysis, 2018.

11. **Sen Yang**, Moliang Zhou, Shuhong Chen, Xin Dong, Omar Ahmed, Ivan Marsic, Randall S. Burd.
Medical Workflow Modeling Using Alignment-Guided State-Splitting HMM. IEEE International

Conference on Healthcare Informatics (ICHI), 2017.

12. **Sen Yang**, Xin Dong, Leilei Sun, Yichen Zhou, Richard A. Farneth, Hui Xiong, Randall S. Burd, Ivan Marsic.
A Data-driven Process Recommender Framework. ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD), pp. 2111–2120, 2017.

13. **Sen Yang**, Xin Dong, Moliang Zhou, Xinyu Li, Shuhong Chen, Rachel Webman, Aleksandra Sarcevic, Ivan Marsic, Randall S. Burd.
VIT-PLA: Visual Interactive Tool for Process Log Analysis. KDD Workshop on Interactive Data Exploration and Analytics (IDEA), 2016.

14. **Sen Yang**, Moliang Zhou, Rachel Webman, JaeWon Yang, Aleksandra Sarcevic, Ivan Marsic, Randall S. Burd.
Duration-Aware Alignment of Process Traces. Industrial Conference on Data Mining (ICDM), pp. 379–393, 2016.

15. **Sen Yang**.
Applied Process Mining, Recommendation, and Visual Analytics. Ph.D. Thesis, Rutgers University, 2019.

ACADEMIC SERVICE

Journal Reviewer
ACM Transactions on Knowledge Discovery from Data (TKDD)
Journal of Biomedical Informatics (JBI)
ACM Transactions on Management Information Systems (TMIS)
IEEE Transactions on Big Data

Conference Program Committee
ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD)
ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR)
IEEE Conference on Computer Vision and Pattern Recognition (CVPR) Workshops
IEEE International Conference on Healthcare Informatics (ICHI)
Conference on Empirical Methods in Natural Language Processing (EMNLP)
Annual Meeting of the Association for Computational Linguistics (ACL)