Junyu Luo

luoshuiti@outlook.com • soap117.github.io/junyu.github.io/

Education & Awards

Academic Qualifications

- **Pennsylvania State University** — **Ph.D.**
  Informatics, GPA: 3.91/4.00 — *2020.08–2024.05*
- **Sichuan University** — **Bachelor**
  Computer Science, GPA: 3.86/4.00, Major GPA: 3.924/4.00 — *2020.06*

Honors & Awards

- **The Award of Excellence, Microsoft Asia Internship Program** — **2020**
- **National Scholarship** — **(THREE TIMES) 2015-2018**
- **The First Prize Scholarship of Sichuan University** — **(THREE TIMES) 2015-2018**
- **The First Prize in Sichuan Province Lanqiao Programming Contest** — **2017**
- **The First Prize in Sichuan University Mathematics Competition** — **2016**

Skills

- Proficient in all kinds of deep learning frameworks, including Transformers, LLMs, diffusion models, GAN, graph neural networks, information retrieval frameworks, and object detection frameworks.
- Proficient in Natural Language Processing and Computer Vision related topics.
- Proficient in Python PyTorch, TensorFlow, and familiar with Keras, C#, C++, Java, and JavaScript.

Work Experience

- **Machine Learning Scientist - Content Recommendation**
  TikTok, USA — *May 2024–Present*
  - **Model Structure Enhancement:** Developing advanced Mixture of Experts (MoME), long sequence modeling, and task fusion frameworks to improve the recommendation model, validating the scaling law in recommendation systems.
  - **Generative Recommendation with VQ:** Develpoing a refined vector quantization (VQ) method for video content, integrating it into generative recommendation frameworks for enhanced video recommendations.

  Skills used: Deep Learning, Machine Learning, Recommendation Systems, Online Learning, VQ, MMoE
- **Research Intern on Natural Language Processing** — **Dr. Danica Xiao**
  Relativity, USA — *June 2023–Aug 2023*
  - **Designing Algorithm for Preventing Hallucination for Large Language Models (LLMs):**
    Paper: Zero-Resource Hallucination Prevention for Large Language Models
    Summary: Used prompt engineering to perform self-evaluation under the zero-resource setting to test the understanding of LLMs in following instructions.
    Skills used: Neural Network Pipeline, NLP, Large Language Models, Constrained Beam Search, Prompt Engineering
- **Research Intern on Machine Learning** — **Dr. Cheng Qian**
  IQVIA, USA — *May 2022–Dec 2022*
  - **Designing Clinical Trial Retrieval Algorithm Based on Trial Protocols:**
    Paper: Clinical Trial Retrieval via Multi-grained Group-based Similarity Learning
    Summary: Designed a hierarchical matching model for trial protocols with a novel group-based training loss and 2D word matching.
    Skills used: NLP, Transformers, Convolutional Networks, Group Loss, Hierarchical Attention, Information Retrieval
  - **Designing Personalized Drug Risk Prediction Model:**
    Paper: pADR: Towards Personalized Adverse Drug Reaction Prediction by Modeling Multi-sourced Data. (CIKM)
    Summary: Incorporated patient's EHR modality with drug molecular information to predict potential adverse reactions.
    Skills used: Pre-trained Language Models, Transformers, Multi-modality, SMILES Chemical Representation, Sequential Modeling, EHR, ICD Codes, Adverse Event Prediction, Spark
- **Research Intern on Knowledge Computing** — **Dr. Jinpeng Wang**
  Microsoft Research Lab - Asia (MSRA), Beijing, China — *Mar 2019–Jan 2020*

- **Automatic Pattern Recognition from PowerPoint Design:**
  Summary: Transformed the pattern matching problem into a sequential matching problem to discover potential design patterns.
  Skills used: Sequential Matching
- **Object Detection for Special Chart Images:**
  Paper: ChartOCR: Data Extraction from Chart Images via a Deep Hybrid Framework (WACV)
  Summary: Designed a high-precision point-based object detection model for chart objects.
  Skills used: Computer Vision, Object Detection, Point Detection, Web Server, Azure

Research Experience

Research Projects on Machine Learning — **Dr. Fenglong Ma**
Pennstate University IST, Pennsylvania, USA — *Feb 2020–May 2025*

- **Multi-modality Pre-training of EHR Data**
  Paper: Hierarchical Pretraining on Multimodal Electronic Health Records. (EMNLP)
  Summary: *Developing a novel, multi-modal, and unified pretraining framework MEDHMP for multi-modality health data.*
  Used Skills: Multi-modality, Pre-training, Self-supervised Learning, Representation Learning, EHR, Spark
- **Automatic ICD Coding based on Diagnosis Text**
  Paper: Fusion: Towards Automated ICD Coding via Feature Compression. (ACL)
  Summary: *Using information compression to reduce the clinical note noise and improve the speed of automatic ICD coding.*
  Used Skills: Transformers, NLP, Text Classification
  Paper: CoRelation: Boosting Automatic ICD Coding Through Contextualized Code Relation Learning.
  Summary: *Improving ICD coding performance through modeling contextualized code relations through graph network.*
  Used Skills: Bi-LSTM, Graph Attention Network, Synonym Fusion, Text Classification
- **Medical Text Simplification**
  Paper: Benchmarking Automated Clinical Language Simplification: Dataset, Algorithm, and Evaluation. (COLING)
  Summary: *Designing a controllable medical term simplification pipeline for using external medical dictionary knowledge.*
  Used Skills: Neural Network Pipeline, NLP, Text Generation, Question Answering, Constrained Generation, Knowledge Injection
- **Electric Health Record Mining**
  Paper: HiTANet: Hierarchical Time-Aware Attention Networks for Risk Prediction on Electronic Health Records. (KDD)
  Summary: *Using two-level transformers to model the complex EHR code sequential data to predict future diseases.*
  Used Skills: Transformers, Time-aware Attention, Sequential Modeling, Disease Prediction

Research Projects on Information Retrieval — **Dr. Min Yang**
Shenzhen Institutes of Advanced Technology(SIAT), Shenzhen, China — *Sep 2017–Jul 2018*

- **Developing methods to generate semantic embedding for long sentences and cross-modal searching**
  Paper: Cross-modal Image-Text Retrieval with Multitask Learning. (CIKM)
  Summary: *Using back-encoding to ensure the cross-modality relation between learned text and image embeddings.*
  Used Skills: Cross-modality, AutoEncoder, Representation Learning, Information Retrieval

Publications (Selected)

- Xiaochen Wang, **Junyu Luo**, Jiaqi Wang, Yuan Zhong, Xiaokun Zhang, Yaqing Wang, Parminder Bhatia, Cao Xiao, Fenglong Ma. 2023. *Unity in Diversity: Collaborative Pre-training Across Multimodal Medical Sources*. The 62nd Annual Meeting of the Association for Computational Linguistics (ACL 2024), August 11–16, 2024, Bangkok, Thailand.
- **Junyu Luo**, Cheng Qian, Lucas Glass and Fenglong Ma. *Clinical Trial Retrieval via Multi-grained Group-based Similarity Learning*. Proceedings of the SIGIR Symposium on IR in Practice (SIGIR-IRIP 2024), July 14-18, 2024, Washington D.C., USA.
- **Junyu Luo**, Xiaochen Wang, Jiaqi Wang, Aofei Chang, Yaqing Wang, and Fenglong Ma. *CoRelation: Boosting Automatic ICD Coding Through Contextualized Code Relation Learning*. The 2024 Joint International Conference on Computational Linguistics, Language Resources and Evaluation (LREC-COLING 2024). 20-25 May, 2024, Italia.
- Xiaochen Wang, **Junyu Luo**, Jiaqi Wang, Ziyi Yin, Suhan Cui, Yuan Zhong, Yaqing Wang and Fenglong Ma. 2023. *Hierarchical Pretraining on Multimodal Electronic Health Records*. Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing (EMNLP 2023), December 6-10, 2023, Singapore.
- **Junyu Luo**, Cheng Qian, Xiaochen Wang, Lucas Glass, and Fenglong Ma. 2023. *pADR: Towards Personalized Adverse Drug Reaction Prediction by Modeling Multi-sourced Data*. In Proceedings of the 32nd ACM International Conference on Information and Knowledge Management (CIKM 2023), October 21–25, 2023, Birmingham, United Kingdom.
- **Junyu Luo**, Zhi Qiao, Lucas Glass, Cao Xiao, and Fenglong Ma. 2023. *ClinicalRisk: A New Therapy-related Clinical Trial Dataset for Predicting Trial Status and Failure Reasons*. In Proceedings of the 32nd ACM International Conference on Information and Knowledge Management (CIKM 2023), October 21–25, 2023, Birmingham, United Kingdom.
- **Junyu Luo**, Junxian Lin, Chi Lin, Cao Xiao, Xinning Gui and Fenglong Ma. *Benchmarking Automated Clinical Language Simplification: Dataset, Algorithm, and Evaluation*. Proceedings of the 29th International Conference on Computational Linguistics (COLING 2022), OCTOBER 12-17, 2022, GYEONGJU, REPUBLIC OF KOREA.
- **Junyu Luo**, Cao Xiao, Lucas Glass, Jimeng Sun and Fenglong Ma. *Fusion: Towards Automated ICD Coding via Feature Compression*. Findings of the 59th Annual Meetingof the Association for Computational Linguistics (Findings of ACL), 2021.
- **Junyu Luo**, Zekun Li, Jinpeng Wang, Chin-Yew Lin: *ChartOCR: Data Extraction from Charts Images via a Deep Hybrid Framework*. Proceedings of the 2021 Winter Conference on Applications of Computer Vision (WACV 2021), 2021.
- **Junyu Luo**, Muchao Ye, Cao Xiao, Fenglong Ma. *HiTANet: Hierarchical Time-Aware Attention Networks for Risk Prediction on Electronic Health Records*. Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD 2020), 2020.
- **Junyu Luo**, Ying Shen, Xiang Ao, Zhou Zhao, Min Yang. *Cross-modal Image-Text Retrieval with Multitask Learning*. Proceedings of the 28th ACM International Conference on Information and Knowledge Management (CIKM 2019), 2019.