YUBO LI

Email: yubol@andrew.cmu.edu | Mobile: (858) 666-7737

EDUCATION

Carnegie Mellon University - *Ph.D. in Information Systems* 2022 – Present
Carnegie Mellon University - *M.S. in Information Systems & Management* 2020 – 2021
University of California, San Diego - *B.S. in Applied Mathematics | Minor in Business* 2015 – 2019
Leadership: Founder of the UCSD Triton Actuarial Society | GPA:3.8/4.0

PUBLICATIONS

- **Yubo L** et al. *Beyond Single-Turn: A Survey on LLM Multi-Turn Interactions*.
- **Yubo L** et al. *Firm or Fickle? Evaluating LLM Consistency Capability in Multi-Turn Interactions*. ACL 2025.
- **Yubo L** et al. *No Black Box Anymore: Demystifying Clinical Predictive Modeling with Temporal-Feature Cross Attention Mechanism*. AMIA 2025 Annual Symposium (Finalist).
- **Yubo L,** Rema P. *Enhancing End Stage Renal Disease Outcome Prediction: A Multi-Sourced Data-Driven Approach*. Journal of the American Medical Informatics Association, 2025.
- **Yubo L** et al. Time-To-Inconsistency: A Survival Analysis of Large Language Model Robustness to Adversarial Attacks. ICLR 2026 (under review).
- **Yubo L,** Weiyi S. Co-Alignment: Rethinking Alignment as Bidirectional Human-AI Cognitive Adaptation. Agents4Science 2025.
- **Yubo L**, Saba S, Rema P. *Towards Interpretable End-Stage Renal Disease (ESRD) Prediction: Utilizing Administrative Claims Data with Explainable AI Techniques*. AMIA 2024 Annual Symposium.

FELLOWSHIPS

- 2025 Presidential Fellowship, Tata Consultancy Services (TCS)
- 2024 Fellowship in Generative AI in Healthcare, Center for Machine Learning and Health, Carnegie Mellon University
- 2023 Fellowship in Digital Health Innovation, Center for Machine Learning and Health, Carnegie Mellon University

RESEARCH EXPERIENCE

- **Q&A Systems for Trustworthy Organ Transplantation Guidance** **Carnegie Mellon University**
  Supervisor: Ramayya Krishnan, Dean of the Heinz College | Rema Padman, CMU Professor 2025.01 - Present
  - **Q&A System:** Developed Q&A systems for organ transplantation patient care by integrating retrieval-augmented generation (RAG) with adaptive frameworks that synthesize patient-specific medical history, clinical guidelines, and dynamic risk factors to deliver tailored guidance and support.
  - **RAG/Retriever:** Engineered a hybrid retrieval pipeline combining both sparse and dense retrievers (DPR, ANCE) to accurately extract and verify medical content from over 100 transplant-specific handbooks.
  - **LLM Reasoning Framework:** Developed an incentivized reasoning mechanism that optimizes LLM performance by rewarding multi-step chain-of-question sequences, enabling the system to reason freely and proactively request critical diagnostic information.

- **LLM Alignment & Multi-Turn Consistency** **Carnegie Mellon University**
  Supervisor: Ramayya Krishnan, Dean of the Heinz College | Rema Padman, CMU Professor 2024.05 - Present
  - **Distributed Training:** Architected and optimized distributed training infrastructure for LLMs (Mistral, LLaMA, Deepseek) across 4 nodes with 16 NVIDIA GH200 GPUs, implementing tensor parallelism, pipeline parallelism, and ZeRO optimizer techniques to efficiently scale model training and reduce computational overhead.
  - **Fine-tuning:** Performed fine-tuning of state-of-the-art LLMs both locally (on-premises clusters) and via OpenAI's official GPT-4o API, successfully improving model consistency, confidence-aware reasoning, and achieving state-of-the-art performance.
  - **Evaluation:** Proposed a metric to quantitatively assess multi-turn response stability, prioritizing early interaction accuracy and swift recovery from initial errors. Released a comprehensive benchmark to test the model's ability to sustain coherent and contextually accurate responses during dynamic, multi-turn conversations.

- **Chronic Kidney Disease Progression Analysis with DL & XAI** **Carnegie Mellon University**
  Supervisor: Rema Padman, CMU Professor | UPMC & Highmark 2022.07 - Present
  - **Deep Learning:** Developed Temporal-Feature Cross Attention Mechanism (TFCAM), an attention-based deep learning framework that explicitly captures temporal and feature-level interactions to enhance clinical predictive modeling for Chronic Kidney Disease (CKD) progression.

- **XAI:** Provided multi-level explainability including temporal insights, feature importance ranking, and cross-temporal feature interactions, enabling clinicians to interpret model predictions transparently and identify actionable clinical insights.
- **Enhanced Representation & Delivery:** Designed an interactive, user-friendly dashboard GUI accessible to diverse user groups; integrated LLMs to automatically generate clear, insightful summaries of predictions and model interpretations, enhancing usability and clinical decision-making support.

WORKING EXPERIENCE

- **Amazon –** *Applied Scientist Intern* 2025.06 – 2025.08
  - **Generative Recommendation Systems:** Developed User-LLM pipeline for personalized recommendation by pre-training custom user encoder on 172k customer interaction sequences, then fine-tuning through progressive alignment strategies (adapter-only, adapter + LLM, full fine-tuning) and pooling mechanisms.
  - Achieved superior predictive performance (60.1% Hit Ratio@3, 67.5% NDCG@3) compared to direct prompting baselines while delivering **93.4% token reduction** and **64.8% cost savings**.
  - **RAG/Search**: Architected production-ready RAG service infrastructure using FAISS HNSW indexing for 60k user embeddings to enhance recommendation contest at scale.
  - **Distributed Training**: Engineered distributed training across multi-GPUs using FSDP with mixed-precision BF16.
- **HireBeat –** *Data Scientist* 2019 – 2020
  - Leveraged machine learning and NLP techniques to develop a data analysis pipeline that optimized content recommendation algorithms, significantly improving personalization accuracy and user engagement.
- **R-eat (Self-employed) -** *Founder* 2012 – 2014
  - **Entrepreneurship**: Founded and grew an online food delivery business by enhancing customer experience through UI optimization and personalized marketing, resulting in a 60% sales increase and a loyal customer base of over 2,000.

COURSES

- **Teaching Assistant**
  - 11-711 Advanced NLP with Prof. Graham Neubig.
  - 10-680/10-681 Mathematical & Computational Foundation for Machine Learning with Prof. Henry Chai.
  - 95-891 Introduction to Artificial Intelligence with Prof. David Steier.
  - 95-888 Data Focused Python with Prof. Michael Simko.
  - 95-865 Unstructured Data Analysis with Prof. George Chen.
  - 95-760 Decision Making under Uncertainty with Prof. David Choi.
- **Courses**
  - 10-714 Deep Learning Systems
  - 10-715 Advanced Introduction to Machine Learning
  - 10-716 Advanced Machine Learning: Theory and Methods
  - 10-703 Deep Reinforcement Learning
  - 10-725 Convex Optimization
  - 10-805 Machine Learning with Big Data
  - 11-711 Advanced NLP
  - 11-785 Introduction to Deep Learning
  - 11-777 Multi-Model Machine Learning
  - 15-513 Introduction to Computer System
  - 15-688 Practical Data Science
  - 36-707 Regression Analysis
  - 36-731/36-732 Causal Inference