XI(SEAN) CHEN

melongone4@gmail.com (citations 200+)⋄ https://melongone.github.io/

EDUCATION

University of Massachusetts Amherst — **Amherst, MA**
- PhD in Computer Science (GPA: 3.9/4) — 2019-present

Central South University — **Changsha, China**
- B.E. in Computer Science (GPA: 3.9/4) — 2014-2018

EXPERIENCE

Amazon (AGI Foundations) — **Seattle, WA**
Applied Scientist Intern — 2024.7 - 2024.10
(Incorporated into Amazon core LLM with a ICML submission, saliently beat existing methods).
- Applied parameter efficient LLM training with Quantized LoRA for scalable model adaptation.
- Devised a model merging framework by aligning activation to accelerate multi-task LLM learning.
- Implemented adaptive online learning with sample re-weighting for continuous LLM improvement.

University of Massachusetts Amherst — **Amherst, MA**
Research Assistant — 2019.9 - present
- Designed a cluster-aware LLM active learning framework for extracting events from temporal textual graph.
- Zero-shot learning on LLM with an event retrieval pipeline from global multilingual news database.
- Built a refined transformer-based global news graph with trillions of edges for event detection, recommendation.
- Developed the largest multilingual news similarity dataset, leveraging active learning for retrieval and ranking.
- Applied greedy factor selection and regression to quantify country effects on news similarity.

University of California, Los Angeles — **Los Angeles, CA**
Research Intern — 2018.7 - 2018.12
- Modeled the dynamics of opinions and personalities with probabilistic graph time series.
- Devised an extended EM algorithm for learning the graph time series and get state-of-the-art performance.

Zhipu AI (Chinese OpenAI) — **Beijing, China**
Research Intern — 2018.3 - 2018.9
- Searched and rank scholars with heterogeneous random forest (deployed on prestigious AI platform Aminer).
- Clustered and visualize temporal scientific topics with refined genetic algorithm.

PROJECT

Merge LLM with Online Activation Alignment
- Developed an activation alignment framework to merge LLMs across diverse tasks (math, coding, medical).
- Adapted activation weight for LLM online learning during parameter-efficient training and sampling.

Cluster-wise LLM Active Learning for Temporal Events from Social Polls
- Designed a cluster-aware active learning framework to detect temporal events from Twitter poll texts.
- Modeled the evolution of social discourse by graph-based topic tracking and statistical analysis.

Global Multilingual News Graph Learning with Transformers
- Devised a global news retrieval system, generating trillion-scale relations with dense embeddings and indexing.
- Built rank models for similar multilingual news using rules-based and active learning classifiers.

LLM Question Answering Generation for Identifying Disasters
- Engineered zero-shot learning prompts to optimize LLM precision and recall in news event retrieval system.
- Built regression models to quantify country factors with greedy feature selection and event clustering.

Personalized Opinion Graph Time Series Learning
- Model the dynamics of opinions and personalities with probabilistic graph time series.
- Devise an extended EM algorithm for learning the graph time series and get competitive performance.

SELECTED PUBLICATION

Adaptive Activation Alignment for Merging Large Language Models
Xi Chen, Yue Zhang, Akshaya Shanbhogue. Submitted to International Conference on Machine Learning (**ICML**), 2025.

Detecting Global Disaster Event with Question Answering on Large Language Model
Xi Chen*, Erica Cai*, Brendan O'connor, Przemyslaw Grabowicz. International AAAI Conference on Web and Social Media (**ICWSM**), 2025.

Cluster-aware Large Language Model learning for Event Identification from Social Polls
Xi Chen, Mayank Bumb, Vishal Kalakonnavar, Przemyslaw Grabowicz. To appeal, 2024.

International News Synchrony and Diversity during the Start of Covid-19
Xi Chen, Scott Hale, David Jurgens, Mattia Samory, Ethan Zuckerman, Przemyslaw Grabowicz. International World Wide Web Conference (**The Webconf**), 2024.

A Multilingual Similarity Dataset for News Article Frame
Xi Chen, Scott Hale, David Jurgens, Mattia Samory, Przemyslaw Grabowicz. International AAAI Conference on Web and Social Media (**ICWSM**), 2023.

Multilingual Document-level Similarity
Xi Chen, Ali Zeynali, Chico Camargo, Fabian Flock, Devin Gaffney, Przemyslaw Grabowicz, Scott Hale, David Jurgens, Mattia Samory. International Workshop on Semantic Evaluation (**@NAACL**), 2022.

Modeling Personalized Dynamics of Social Network and Opinion at Individual Level
Xi Chen, Jie Tang, Yizhou Sun. Preprint on Arxiv, 2019.

SKILL

Core Expertise:
Large language model: RLHF, Parameter-efficient fine-tuning, Model merging, Retrieval-augmented generation(RAG).
Search & Retrieval: News similarity search, Transformer-based retrieval models, Semantic indexing.
Recommendation Systems: Graph modeling, Active learning for ranking, Multilingual content recommendations.
Machine Learning & Optimization: Event detection, Time-series analysis, Probabilistic modeling.

Tools & Frameworks:
Deep Learning: PyTorch, TensorFlow, Hugging Face Transformers.
Big Data & Search: Spark, Elasticsearch, FAISS, SQL.
Programming: Python, C++, R, Java, Linux.

SERVICE

Program Member:

- NAACL(2023)
- ICWSM(2023, 2024)
- SEMEVAL (2022, 2023, 2024)
- IC2S2 (2022, 2023, 2024)

HONOR

NSF Student Travel Award, 2024
UMass CICS best portfolio finalist, 2023.
Research presented at top-tier computer science venues (Webconf, NAACL, ICWSM, TADA, IC2S2).
Nomination for Chinese Exceptional Student (1/6500 ISchool undergrads and grads), 2018.
Ranked global Top50 of Autochess players (50/millions, a world-class strategy game).