XI CHEN

melongone4@gmail.com ◇ https://melongone.github.io/

EDUCATION

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

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

EXPERIENCE

University of Massachusetts Amherst — **Amherst, MA**
Research Assistant — 2019.9 - present
- Refined a transformer model to analyze global news graph with trillions of edges.
- Quantified country effects on news similarity with greedy factor selection, event clustering, and regression.
- Developed a large dataset of multilingual news article pairs with human-in-the-loop active learning framework.

Zhipu AI — **Beijing, China**
Research Intern — 2018.3 - 2018.9
- Disambiguated scholars with heterogeneous random forest (deployed on prestigious AI platform Aminer).
- Predicted social links by probabilistic graph time series with latent variable.

Briup Technology corporation — **Jiangsu, China**
Software Engineer Intern — 2017.6 - 2017.8
- Developed a server-client smart traffic system on sand table as to route optimization and obstacle avoidance.

PROJECT

Large-scale Multilingual News Graph Learning
- Extract cross-lingual name entities with Spacy, Polyglot, and Wikipedia concepts.
- Recall, pre-rank, and rank similar news article pairs based on rules and active learning classifiers.
- Predict global news network for trillions of pairs with refined multilingual transformer and bi-directional index.
- Built regression models to quantify country factors with greedy feature selection and event clustering.

Personalized Opinion Graph Time Series Learning
- 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 competitive performance.

Symbolic Math Expression Learning
- Parse and regularize math formula into tree embedding with special identifiers.
- Built a sequence-to-sequence model to compute partial differential of formula.

SELECTED PUBLICATION

A Multilingual News Similarity Dataset for Media Network and Bias Research
Xi Chen, Scott Hale, David Jurgens, Mattia Samory, Przemyslaw Grabowicz. To appeal, 2023.

International News Synchronization during Covid-19 Pandemic
Xi Chen, Scott Hale, David Jurgens, Mattia Samory, Przemyslaw Grabowicz. International Conference 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 at NAACL (SEMEVAL), 2022.

Other 2 first author papers and 1 second author paper for network science and optimization
Published on peer-reviewed computer science academic journals.

SKILL

PhD Courses: Machine Learning, Optimization, Natural Language Processing, Database, Algorithm
Programming Language and framework: Python, C/C++, R, Pytorch, Spark, Java, SQL, Linux, Matlab

HONOR

Oral presentations at top-tier computer science conferences (NAACL 2022 and IC2S2 2022).
Nomination for Chinese Exceptional Student (1/6500 ISchool undergrads and grads), 2018.
Ranked global Top50 of Autochess players (50/millions).