Yuxing Fei

Phone: (+1)-510-325-3316
Email: yuxingfei@berkeley.edu
Website: https://yuxingfei.com

Last updated on October 7, 2024

Berkeley, CA 94720
Linkedin: @yuxing-fei
Google Scholar: Yuxing Fei
Github: @idocx

Summary

- Worked as a core member to build the world's first fully **automated laboratory for solid-state powder synthesis** (A-Lab); developed **customized robots** and **management software** for high-throughput material experimentation; applied A-Lab to **accelerate battery material research**.
- **Material experiment skills**, especially in solid-state synthesis, XRD, SEM/EDS, and battery testing.
- Experience in **AI for Science**, emphasizing graph neural network and text mining for materials science.

Education

University of California, Berkeley Berkeley, CA, USA — Aug. 2021 – May 2026 (expected)
Ph.D. in Material Science, Department of Materials Science and Engineering
GPA: 3.914/4; Advisor: Prof. Gerbrand Ceder

University of California, Berkeley Berkeley, CA, USA — Aug. 2021 – Aug 2024
M.S. in Material Science, Department of Materials Science and Engineering

Wuhan University Wuhan, Hubei, China — Sept. 2017 – June 2021
B.S. in Chemistry, College of Chemistry and Molecular Sciences & **Honor Science Program** of Hongyi Honor School
GPA: **3.96**/4.00; Rank:**1**/17

University of California, Berkeley Berkeley, CA, USA — Jan. 2020 – Aug. 2020
Visiting Undergraduate Student

Research Experience

Graduate Student Researcher @ UC Berkeley — Aug. 2021 – Present
Overview:
I am one of the core developers in a multi-disciplinary team to build an autonomous laboratory (A-Lab) to accelerate the discovery of inorganic materials with computational materials, machine learning, robotic experimentation, and text mining.

- Designed AlabOS, a general-purpose **experiment and data management platform for self-driving lab**. (Tech stacks: Python, MongoDB, RabbitMQ, Flask, ReactJS)
- **Prototyped, tested, and integrated customized robots** into the automated lab using Python, Fusion 360 and 3D printing.
- Prototyped an **automated coin-cell assembly** pipeline for high-throughput battery testing in A-Lab, based on drop-casting method.
- Deploying **active learning** algorithm in A-Lab for **rapidly determining Fluorine-dopant's solubility limit in disorder rocksalt** (DRX) cathode material for next-gen lithium-ion battery, using lattice parameters from powder X-ray diffraction.

Keywords: self-driving Lab; solid-state synthesis; active learning; battery testing; hardware development; programming; disorder rocksalt cathode.

Undergraduate Research Assistant @ UC Berkeley — Feb. 2020 – Aug. 2020
1. Self-supervised Language Model for Abbreviation Sense Disambiguation in Material Science Documents
Overview:
This research focuses on developing a novel learning-based toolkit to detect, expand, and disambiguate abbreviations that appear in material science documents more accurately, which is an important part of automatic knowledge extraction from material science literature.

- Modified the abbreviation detection algorithm by Schwartz to get better performance in material science documents.
- Designed a self-supervised task to teach deep learning model (*Bert*) to **catch latent knowledge in the scientific documents** and used this model to disambiguate material abbreviations in texts.
- Built a **search engine** to link the long forms to PubChem entries based on Elasticsearch.

Keywords: scientific text mining; material informatics; natural language processing (NLP); word disambiguation.

2. COVIDScholar: COVID-19 literature search powered by advanced NLP algorithms
Overview:
This work is aimed at facilitating researchers to retrieve COVID-19-related publications more efficiently using state-of-the-art natural language processing techniques.

- Trained a **multi-label text classification** model for categorizing COVID-19 related articles (e.g. `prevention`, `diagnosis`, `treatment`) based on *SciBert*.
- Built a super-fast autocomplete system for search queries.
- Trained a *FastText* model for search **query re-ranking** to improve search accuracy.
- Created word embedding visualization to show latent relationship between terminologies.
- Wrote **scrapers and parsers** to collect up-to-date papers from various sources.

Keywords: natural language processing; data mining; COVID-19; scientific search engine.

Work Experience

Intern @ Machine Learning Group, Microsoft Research Asia Nov. 2020 – Apr. 2021
Overview:
In this project, we designed a reinforcement learning (RL) agent to generate novel and stable crystal structures that are not included in the experiment database like ICSD.

- Built a **PPO-based RL model** using CGCNN and RLlib.
- Proposed an **ionic substitution environment** for RL-based novel crystal generation.
- Designed various rewards (step rewards & final rewards) to guide the RL agent to make good decisions and tried to find a way to aggregate them into a single reward value.

Keywords: deep reinforcement learning; graph neural network; crystal structure generation.

Publications

[1] N. J. Szymanski, B. Rendy, **Y. Fei**, R. E. Kumar, T. He, D. Milsted, M. J. McDermott, M. Gallant, E. D. Cubuk, A. Merchant, *et al.*, "An autonomous laboratory for the accelerated synthesis of novel materials," *Nature*, vol. 624, no. 7990, pp. 86–91, 2023.

[2] **Y. Fei**, B. Rendy, R. Kumar, O. Dartsi, H. P. Sahasrabuddhe, M. J. McDermott, Z. Wang, N. J. Szymanski, L. N. Walters, D. Milsted, Y. Zeng, A. Jain, and G. Ceder, "AlabOS: A Python-based reconfigurable workflow management framework for autonomous laboratories," *Digital Discovery*, 2024.

[3] S. Wang, N. J. Szymanski, **Y. Fei**, W. Dong, J. N. Christensen, Y. Zeng, M. Whittaker, and G. Ceder, "Direct Lithium Extraction from α-Spodumene through Solid-State Reactions for Sustainable Li_2CO_3 Production," *Inorganic Chemistry*, vol. 63, no. 29, pp. 13 576–13 584, 2024.

[4] Z. Wang, Y. Sun, K. Cruse, Y. Zeng, **Y. Fei**, Z. Liu, J. Shangguan, Y.-W. Byeon, K. Jun, T. He, *et al.*, "Optimal thermodynamic conditions to minimize kinetic by-products in aqueous materials synthesis," *Nature Synthesis*, vol. 3, no. 4, pp. 527–536, 2024.

[5] J. Dagdelen, A. Trewartha, H. Huo, **Y. Fei**, T. He, K. Cruse, Z. Wang, A. Subramanian, B. Justus, G. Ceder, *et al.*, "COVIDScholar: An automated COVID-19 research aggregation and analysis platform," *Plos one*, vol. 18, no. 2, e0281147, 2023.

[6] Z. Wang, K. Cruse, **Y. Fei**, A. Chia, Y. Zeng, H. Huo, T. He, B. Deng, O. Kononova, and G. Ceder, "ULSA: Unified language of synthesis actions for the representation of inorganic synthesis protocols," *Digital Discovery*, vol. 1, no. 3, pp. 313–324, 2022.

[7] Z. Wang, O. Kononova, K. Cruse, T. He, H. Huo, **Y. Fei**, Y. Zeng, Y. Sun, Z. Cai, W. Sun, *et al.*, "Dataset of solution-based inorganic materials synthesis procedures extracted from the scientific literature," *Scientific Data*, vol. 9, no. 1, p. 231, 2022.