Feng CHEN

fengc9@uw.edu | (781)266-7238| 3031 184th ST SW, Apt C104, Lynnwood, WA 98037

EDUCATION

University of Washington Seattle, WA Sep 20223 – now
PhD candidate in Biomedical and Health Informatics

Harvard Medical School Longwood, MA Sep 2021 – Dec 2022
Master of Biomedical Informatics **Cumulative GPA:** 3.96

Brandeis University Waltham, MA Jan 2018 – Dec 2020
Double major: Biochemistry & Computer Science **Cumulative GPA:** 3.59 (Dean's list) **GPA (Math&CS):** 3.78
Award: Jacques Cohen Award in Computer Science (awarded for excellent researchers)
Main Coursework: Deep Learning for Natural Language Processing, Data Structure and Algorithms, Statstical Machine Learning, Deep Learning for Biomedical Data, Clinical Informatics, Data visualization, Precision Medicine, Data Science for Medical Decision Making

SKILLS

Programming: Python (proficient), Java (proficient), SQL, R, Scheme, Matlab, HTML
Software and tools: AWS, Github, Photoshop, Git, Microsoft Office, Linux, Cortellis, CLI
Research skills: Machine learning; deep learning; natural language processing; large language models; clinical data processing; automatic speech recognition; bioinformatics analysis; data cleaning, data analysis and visualization; proposal and manuscript writing; paper reviewing, qualitative coding.

RESEARCH EXPERIENCES

University of Washington | Advisors: Prof. Trevor Cohen Seattle, WA
Research assistant

Natural Language Processing for Social Signals Detection in Patient-Provider Clinical Dialogs Sep 2023 - now
- Utilized ASR (Whisper) and speaker diarization (Pyannote) to transcribe and analyze patient-clinician conversations.
- Applied multiple NLP and LLM models to train and classify the social signals scores for patients and clinicians.
- Investigated implicit racial bias in clinical dialogs and optimized predictive performance for detecting social signals.

Comparative Analysis of LLMs in Detecting PTSD in Clinical Interviews Jan 2025 – now
- Developed and compared transformer-based and embedding-based NLP models (e.g., Mental-RoBERTa, SentenceBERT, LLaMA) for PTSD classification from clinical transcripts.
- Implemented prompt-based PTSD classification using LLaMA with DSM-5-informed prompts.
- Conducted error analysis on symptom severity and depression comorbidity, revealing model sensitivity to clinical subtleties.

Developing Data-driven Clinical Signatures for People with Hallucinations Mar 2024 - now
- Building automatic pipeline to derive multimodal data-driven clinical signatures from mobile behavioral tasks to predict individual differences in severe negative outcomes among people experiencing hallucinations.
- Developed and improved automatic sentence coherence measurement package to evaluate verbal recall tasks audio recordings.
- Identifying and mitigating bias in automatic speech recognition and predictive models across demographic groups.

LLM-based Chatbot Design for the Elderly Community Mar 2024 - now
- Designed and developed an LLM-powered chatbot to improve life for residents at a retirement community.
- Conducted user interviews and training sessions to refine features like adjustable font sizes, high-contrast modes, and voice-to-text support.
- Engineered prompts to ensure clear responses and implement interface to enhance accessibility and user satisfaction.

Harvard Medical School | Advisors: Dr Li Zhou & Prof. Pengyu Hong Longwood, MA
Research assistant/Project core member

Bias and Missingness Handling in the Task of Mortality Prediction for ARDS Feb 2022 – Dec 2022

- Preprocessed MIMIC-IV data and retrieved data for ARDS patients with ICU ventilation data. Then, Calculated and visualized the demographics distribution and missing rate for all features used for mortality prediction.
- Built a workflow to impute missing values based on different types of missingness with four imputation methods.
- Used propensity score matching to detect and narrow the bias in the data and compared the machine learning model performance for mortality prediction.
- Helped to apply for grants for the project for the lab in the project of handling bias in EHR data.

Paper Review for Bias in Healthcare AI Applications using EHR data Dec 2022 – Dec 2023

Project leader/ First author | Advisors: Dr Li Zhou & Dr Liqin Wang

- Used scripts and key-word search to extract paper related to bias handling application healthcare AI and followed the flow of PRISMA workflow to manual review and filter over 200 papers that satisfied our pre-set requirements.
- Reviewed the full text of the paper and form recorded the evaluation metrics, detailed methods and classified types of bias and methods to handle bias.
- Published a systematic review paper that summarizes the results of bias handling in EHR data to JAMIA.

Brandeis University | Advisors: Prof. Hongfu Liu & Prof. Pengyu Hong Waltham, MA

Research assistant/Project core member/Co-first author

Virus-host Interactions of SARS-CoV-2 Predicted by a Network-based Cross-species Model Mar 2020 – Aug 2020

- Researched the virus-host interactions of other published coronaviruses and compared SARS-CoV-2 with them after multiple sequence alignment in terms of domains, functions and mutations.
- Built graphs presenting sequence similarities in viral and host proteins and networks of the protein-protein interaction.
- Performed graph embedding and used deep learning to train a network-based cross-species model; used it to predict the highly potential PPI and pathways of SARS-CoV-2 among animals.

The Prediction of SARS-CoV-2 Mutation Using Deep Learning and Evolutionary Algorithms Apr 2020 – Sep 2021

- Collected the data of the identified SARS-CoV-2 mutations along the time and location of their occurrences and analyzed the datasets to calculate the mutation rates and identify the hot zone.
- Built a model with evolutionary algorithms and GAN to predict the mutation of SARS-CoV-2 based on its mutation history.
- Studied the mutation influence on the 3-D structure of the virus proteins.

Massachusetts Institute of Technology | Advisor: Prof. Jean-Francois Hamel Cambridge, MA & Online

Testing Simulators for Bioreactors and Optimizing Parameters for Vaccine Productions Jan 2020 – Aug 2020

Research intern

- Learned and used bioreactor simulators to cultivate cells and study the bioprocesses for both mammalian and yeast cells.
- Researched experimental settings for vaccine production, performed tests using AAV on the simulators and compared with experimental results; tried to optimize the parameterization of vaccine production for COIVD-19 to maximize the yield.
- Wrote weekly reports and a summary report about the optimal environment for the potential expanded-scale production.

INTERNSHIP EXPERIENCES

Amazon Software Development Engineer Intern | Bellevue, WA May 2022 – Aug 2022

- Developed and optimized the algorithm for the shortest path calculation for the Supply Chain Optimization group.
- Used clustering and machine learning methods to find the optimal partition the logistic network for transportation.
- Visualized the network before and after the optimization as well as the statistics the algorithm improved.

Gordian Ventures Investment Intern | Shanghai, China Mar 2023 – Aug 2023

- New technologies research including t-RNA, Aptamers, Brain-computer interfaces and write publish information book.
- Collected and mapped Biotech companies early funded by several large ventures. Summarized their key technology and pipelines and clustered and visualized the recent trends of drug discovery.
- Collected and calculated the comprehensive market size and the potential growth rate for peptide drug market and supported the valuation of our newly incubated companies using DCF model.

MANUSCRIPT

Feng Chen*, Hangyu Du*, Hongfu Liu, Pengyu Hong, *Network-based Virus-Host Interaction Prediction with Application to SARS-CoV-2,* published in Cell Patterns (* co-first authors) DOI: 10.1016/j.patter.2021.100242

Feng Chen, Liqin Wang, Julie Hong, Jiaqi Jiang, Li Zhou, *Unmasking Bias and Inequities: A Systematic Review of Bias Detection and Mitigation in Healthcare Artificial Intelligence Using Electronic Health Records*, published in JAMIA. DOI: 10.1093/jamia/ocae060

Feng Chen, Manas Satish Bedmutha, Ray-Yuan Chung, Janice Sabin, Wanda Pratt, Brian R. Wood, Nadir Weibel, Andrea L. Hartzler, Trevor Cohen, *Toward Automated Detection of Biased Social Signals from the Content of Clinical Conversations*, oral presented as full paper in AMIA 2024 Annual Symposium.

Feng Chen, Dror Ben-Zeev, Gillian Sparks, Arya Kadakia, Trevor Cohen, Detecting PTSD in Clinical Interviews: A Comparative Analysis of NLP Methods and Large Language Models, submitted to AMIA 2025 Annual Symposium.

Feng Chen, Ray-yuan Chung, Yein Jeon, Oleg Zaslavsky, *Improve Technology and eHealth Literacy in a Retirement Community via The Use of LLM-Powered Chatbo*t, presented in GSA 2024 as a poster. DOI: 10.1093/geroni/igae098.4005

Feng Chen, Luna Xingyu Li, Ray-yuan Chung, Yein Jeon, Wenyu Zeng, Oleg Zaslavsky, Learning from Elders: Making an LLM powered Chatbot for Retirement Communities more Accessible through User-centered Design, submitted to CALD-AI @ ASIS&T 2025.

Xiruo Ding, Zhecheng Sheng, Brian Hur, **Feng Chen**, Serguei VS Pakhomov, Trevor Cohen, *Enhancing Robustness of Foundation Model Representations under Provenance-related Distribution Shifts*, presented in Workshop on Distribution Shifts, 37th Conference on Neural Information Processing Systems (NeurIPS 2023).