

# Jixin Li

617-230-7140 | li.jix@northeastern.edu | jixinli.info | github.com/rexli999 | linkedin.com/in/jixin-li-432b7a123

## Summary

Data scientist and user experience researcher with an interdisciplinary background in psychology, statistics, and wearable sensing techniques. Core competencies grounded in statistical analysis of user data, human-computer interaction research, and innovative health system design. Harness the power of wearable devices and apply advanced analytical and artificial intelligence techniques to improve health outcomes.

## Education

### Northeastern University

Ph.D., Personal Health Informatics, Khoury College of Computer Sciences

Boston, MA

Expected May 2025

Relevant Courses: Machine Learning, Time Series Analysis, Human-Computer Interaction, Leadership and Project Management

### Columbia University

M.A., Statistics

New York City, NY

Dec 2017

### University of Michigan

B.A., Major in Psychology, Minor in Statistics | University Honor

Ann Arbor, MI

Apr 2014

## Skills

**Tools and Software** Python, R, Java, Android Mobile App Development, Git, AWS, SQL, SPSS, Linux, Shell (Bash), LaTeX, Microsoft Office

**Technical Skills** Experimental design and statistical testing, usability research, data engineering and analytics, machine/deep learning

## Research Experience

### Northeastern University, mHealth Research Group

Boston, MA

*"Automated Semantic Enrichment of Trajectory with OpenStreetMap"*

Sep 2021 - Oct 2022

- Created an open-source repository that automates the process of organizing OpenStreetMap semantic data into hierarchical taxonomy and assigning multi-level semantic labels to location data (Clustering Algorithms; Android, Java; Python with GeoPandas)
- Evaluated semantic enrichment of trajectory with geodatabases on a longitudinal smartphone location dataset (50 million coordinates); identified future research directions in combining self-report and automatic methods in semantic enrichment (Python with Matplotlib and Seaborn)

*"Forecasting Mental State Time Series in Ecological Momentary Assessment (EMA) Studies"*

Jan 2022 - May 2022

- Implemented traditional time series models (ARIMA) and machine/deep learning models (GBRT, LSTM) to forecast one-day-ahead emotion answers from a longitudinal EMA dataset of 50867 days (Python with Scikit-learn and PyTorch)
- Performed time series forecasts to reduce user burden by skipping 30% of prompts in EMA studies while controlling information loss at an acceptable level (6%); confirmed the feasibility of leveraging time series prediction to determine the likelihood of delivering EMA prompts

*"Literature Review on Reinforcement Learning (RL) in Just-In-Time Adaptive Interventions"*

Sep 2021 - Jan 2022

- Conducted a comprehensive literature review on the pros and cons of the state-of-the-art algorithms behind current JITAI applications
- Presented the compatibility and challenges of applying RL in EMA applications to maximize information gain and minimize user burden

*"Contextual Biases in Microinteraction Ecological Momentary Assessment Non-response"*

May 2020 - Mar 2022

- Created a data pipeline of processing system logs and sensor data of smartphones and smartwatches from 200 participants in a 1-year study (70,000 days) using parallel computing on the Discovery computing cluster (Android; Cluster Computing; Python with Numpy and Pandas)
- Built multi-level models to identify 10 contextual factors that impact the compliance of answering EMA surveys on smartphones and smartwatches, including temporal factors, device usage, and human activities (wrist motion and location) (Wearable Sensor; R with lme4)

*"MixWILD: GUI-based Desktop Application for Intensive Longitudinal Data Analysis"*

Sep 2019 - Present

- Developed new user-friendly GUI features to facilitate health researchers in conducting novel longitudinal data analysis (Swing Java)
- Conducted in-depth usability testing with health researchers using combined qualitative and quantitative methods of observation, questionnaires and interviews to inform iterative improvements for user experience (Project Management; Qualtrics)

## Work Experience

### Learnable, Inc.

Boston, MA

Data Analyst

Oct 2017 - Feb 2019

- Empowered real-time pricing for transportation delay insurance by scraping weather data and predicting delay risk using ML models (AWS)
- Implemented multi-level ML models to automate the categorization of exercise texts digitized from mathematics and physics books (NLP)

### AsiaEAP Consulting Co.

Shanghai, China

Survey Data Analyst

Jun 2015 - Aug 2016

- Designed mental health surveys and conducted group interviews with labor-intensive companies, including Honeywell, Volkswagen, Qunar.com
- Wrote organizational mental health reports and recommended mental healthcare services for clients to enhance organizational productivity