# .Jixin l i

📱 617-230-7140 | 🗷 li.jix@northeastern.edu | 🎢 jixinli.info | 🖬 linkedin.com/in/jixin-li-432b7a123 | 🕿 google scholar

# Summary

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

# Education

#### **Northeastern University** Ph.D., Personal Health Informatics, Khoury College of Computer Sciences

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

### **Columbia University**

M.A., Statistics

### **University of Michigan**

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

# Skills

Tools and Software Python, Pytorch, R, Java, Android Mobile App Development, Git, AWS, SQL, SPSS, Shell (Bash), EFX, 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

"Improving the efficiency of self-reported data collection using machine learning methods"

- Designed and implemented a ML-enhanced system for adaptive-length real-time surveys on in-situ human behavior and mental states
- Evaluated the system on multiple open-source datasets, reducing average survey length by over 50% while minimizing information loss

### "Automated Semantic Enrichment of Trajectory with OpenStreetMap"

- Created an open-source repository that automates the process of organizing OpenStreetMap semantic data into hierarchical taxonomy and annotating location data with multi-level semantic labels Clustering Algorithms; Android, Java; Python with GeoPandas
- Evaluated automatic annotations from commonly used geodatabases on a longitudinal smartphone location dataset (50 million coordinates); proposed a location-labeling EMA system that combines automatic annotations from geodatabases Python with Matplotlib and Seaborn

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

- 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

### "Contextual Biases in Microinteraction Ecological Momentary Assessment Non-response" (Distinguished Paper Award)

- 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 Ime4

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

- 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.

Data Analyst

Boston, MA Oct 2017 - Feb 2019

Shanghai, China

Jun 2015 - Aug 2016

- 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.

Survey Data Analyst

- 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

Expected May 2025 New York City, NY Dec 2017 Ann Arbor, MI

Apr 2014

Boston, MA

Boston, MA

Sep 2021 - Oct 2022

Jan 2022 - May 2022

Feb 2023 - Feb 2024

Mar 2022

Sep 2019 - Present