

Sonia Bae

Charlottesville, VA, 22903 | <https://soniabae.com> | 434-466-7628 | sb5ce@virginia.edu | [linkedin.com/in/sonia-bae/](https://www.linkedin.com/in/sonia-bae/)

EDUCATION

-
- University of Virginia** | PhD, System and Information Engineering (GPA: 3.94/4) Expected May 2022
- *Relevant coursework:* Statistical Modeling, Text Mining, Reinforcement Learning, Stochastic Modeling, Mobile Sensing
- Amirkabir University of Technology** | M.S. Computer Science (GPA: 18.26/20 – Distinction) Sep. 2011 – Oct. 2013
- *Relevant coursework:* Combinatorial Optimization, Advanced Artificial Intelligence, Logic Programming, Decision Support
- Amirkabir University of Technology** | B.S. Computer Science (GPA: 17.94/20 – First Class) Aug. 2007 – Sep. 2011
- *Relevant coursework:* Algorithms and Data Structures, Principles of Operating Systems, Graph Theory

SKILLS

Pytorch, Keras, Scikit-Learn, Pandas, Numpy, ggplot, Tableau, Python, Java, R, C/C++, MATLAB, SQL, HTML, AngularJs, Javascript, CSS, PHP, D3, Sequential Models, Reinforcement Learning, (un-)Supervised Learning, A/B testing, Hypothesis testing

EXPERIENCE

Machine Learning Engineer and Data Scientist Intern June 2021 – Present
Deliberate AI | New York, NY

- Creating different *deep learning models* to predict mental health disorders.
- Developed a framework to collect multi-modal data from subjects.
- Developed an application to get ground truth labels with *JavaScript, SQL, AWS, PHP, HTML*, and *Gitlab*.

Co-Instructor – Applied Statistical Modeling August 2021 – Present
University of Virginia | Charlottesville, VA

- Created syllabus and class material (data visualization, logistic regression, and bootstrapping in R).
- Developed class website and creative class engagement activity with *HTML, SQL, JavaScript*.

Graduate Researcher July 2017 – Present
University of Virginia | Charlottesville, VA

Project Lead, Predicting Engagement and Emotional Support from Social Messaging Text Data and Metadata

- Implemented a network of topics for patients with HIV (i.e., Text Mining, and Sentimental Analysis).
- Developed a data pipeline for cleaning and processing large Facebook text messaging datasets in *Python*.
- Created visualizations to enable regular communications with clinical collaborators by *JavaScript, Tableau*, and *Python*.

Project Lead, Predicting Attrition in Digital Mental Health Interventions

- Developing a conversational agent to deliver a micro-intervention for high-risk users.
- Developed a *python* machine learning framework for data processing, model development (e.g., Logistic Regression, Gradient Boosting, Random Forest, RGF, Decision Tree, Extreme Boosting, MLP, SVM), and feature engineering.

Project Co-lead, Evaluating Multi-session Online Interpretation Bias Training For Anxiety

- Created a pipeline for cleaning, processing, model development, *A/B testing*, and validation in *SQL, Python*, and R.
- Co-developed *Java* backend with *SQL* integration and optimize existing *AngularJS* of a multi-session internet-delivered cognitive bias modification program, <https://mindtrails.virginia.edu>.

Project Lead, Predicting and Learning Situational Awareness of Drivers in semi-autonomous vehicles

- Developed a novel modular framework, MEDIRL, to learn (near-)optimal policies of drivers in accident-prone situations with *deep inverse reinforcement learning* in *Python*.
- Designed experimental studies, pulled, cleaned and processed large multi-modal data streams with *Python*.
- Results indicate that MEDIRL outperforms existing models and achieves state-of-the-art performance.
- Implemented different machine learning algorithms and *deep learning models* (i.e., Logistic Regression, Gradient Boosting, Random Forest, Regularized Greedy Forest (RGF), Naive Bayes, AdaBoost, and Deep Neural Network).

SELECTED PUBLICATIONS

-
- **S. Bae**, E. Pakdamanian, I. Kim, L. Feng, V. Ordonez, L. Barnes. *MEDIRL: Predicting the Visual Attention of Drivers via Maximum Entropy Deep Inverse Reinforcement Learning*, IEEE International Conference on Computer Vision (ICCV 2021).
 - J. L. Ji, **S. Bae**, D. Zhang, C.P. Clichó-Mamani, M. J. Meyer, D. Funk, S. Portnow, L. Barnes, B. Teachman, *Multi-session online interpretation bias training for anxiety in a community sample*. Behaviour Research and Therapy (BRAT 2021).
 - S. Mendu, A. Baglione, **S. Bae**, C. Wu, B. Ng, A. Shaked, G. Clore, M. Boukhechba, L. Barnes, *A framework for understanding the relationship between social media discourse and mental health*. Proceedings ACM on Human-Computer Interaction (CSCW 2020).
- E. Pakdamanian, S. Sheng, **S. Bae**, S. Heo, S. Kraus, L. Feng, *Deeptake: Prediction of driver takeover behavior using multimodal data*. Conference on Human Factors in Computing Systems (CHI 2021).

EXTRACURRICULAR

Discussion Leader, SIGDIAL	July 2020
Vice President, Data Science and Analytic Club at UVA	2020-2021
Webmaster, Graduate Society of Women in Engineering at UVA	2018-2019
Engineering Recruiter (SWE, NSBE) at UVA	2018
Core Mentor, Research Experiences for Undergraduates at UVA	2018-2021