Christina Baek

kbaek@cs.cmu.edu
kebaek.github.io

Education

Carnegie Mellon University, Advised by Zico Kolter — Aug. 2021– Present
Ph.D. in Machine Learning

University of California, Berkeley, Advised by Yi Ma — Aug. 2020 – Jun. 2021
5th Year M.S. in Electrical Engineering and Computer Science; Mathematics Breadth

University of California, Berkeley — Aug. 2016 – Jun. 2020
B.S. in Electrical Engineering and Computer Science; Minor in Bioengineering

Experience

Carnegie Mellon University, Department of Machine Learning Advised by Zico Kolter — Aug. 2021 - Present
Research Assistant

- Focus: out-of-distribution generalization, continual learning
- Worked on showing that the disagreement rate between a pair of models trained on different seeds is often exactly equal to their average test error, even for out of distribution test data. Published in ICLR 2022 (Spotlight).
- Current working on better understanding the calibration of ensembles.

UC Berkeley Artificial Intelligence Research Lab, Advised by Yi Ma — Jan. 2020 - Dec. 2021
Research Assistant

- Focus: continual learning, dictionary learning, reinforcement learning
- Reformulated the Maximal Coding Rate Reduction loss such that the number of log determinants required does not grow linearly with the number of classes. Saw 10x faster training on Tiny-ImageNet. Currently in submission to CVPR 2022.
- Worked on a journal publication theoretically analyzing the improvements in computational complexity we observe in practice when the agent is given intermediate rewards in reinforcement learning tasks. Published in JAIR 2022.
- Showed that ReduNet, a network constructed by forward propagation, performs significantly better on class incremental learning tasks than deep networks trained by backpropagation. Worked on global convergence proofs of loss functions over the Steifel manifold. Published in CVPR 2021.

UC Berkeley Automation Lab, Advised by Ken Goldberg — Jan. 2014 - Oct. 2016
Research Assistant

- Designed a particle filter algorithm to tackle the mechanical search problem of grasping a target object in a cluttered bin.

UC Berkeley Molecular Cell Biomechanics Lab, Advised by Mohammad Mofrad — Jan. 2019 - Jun. 2019
Research Assistant

- Designed convolutional neural networks to predict the punctual stress during unfolding in molecular dynamics simulations of double globule tethered proteins. Discovered patterns between punctual stress and a protein's secondary structure during protein unfolding.

Harvard Medical School, Department of Biomedical Informatics, Advised by Chirag Patel — Jun. 2017 - Aug. 2017
Intern for Summer Institute of Bioinformatics

- Created a database of annotated microbiome studies that use whole-genome sequencing https://microbial-genes.bio .
- Built a pipeline in R that conducts a metagenome-wide association study of microbiome data and outputs significant genetic/functional markers.

UC San Diego, Department of Medicine, Advised by John Chang — Jun. 2016 – Jan. 2017
Research Assistant

- Studied the role of TGFβ pathway in cancer. Showed that modulating USP11 expression altered the stability of TGFβ receptor type 2 (TGFBR2) and TGFβ downstream signaling in human breast cancer cells.

Publications

⋆ denotes equal contribution

[1] **Assessing Generalization of SGD via Disagreement** [arxiv]
Yiding Jiang⋆, Vaishnavh Nagarajan⋆, **Christina Baek**, J. Zico Kolter
International Conference in Machine Learning (ICML) Workshop on Overparameterization: Pitfalls & Opportunities, 2021 + International Conference on Learning Representations (ICLR), 2022 (Spotlight)

[2] **Computational Benefits of Intermediate Rewards for Hierarchical Planning** [arxiv]
Yuexiang Zhai, **Christina Baek**, Zhengyuan Zhou, Jiantao Jiao, Yi Ma
Journal of Artificial Intelligence Research (JAIR), 2022

[3] **Incremental Learning via Rate Reduction** [arxiv]
Ziyang Wu⋆, **Christina Baek⋆**, Chong You, Yi Ma
Conference of Computer Vision and Pattern Recognition (CVPR), 2021

[4] **The Landscape of Genetic Content in the Gut and Oral Human Microbiome** [pubmed]
Braden Tierney, Zhen Yang, Jacob Luber, Marc Beaudin, Marsha Wibowo, **Christina Baek**, Chirag Patel, Aleksandar Kostic
Cell Host and Microbe, 2019

[5] **Ubiquitin specific peptidase 11 (USP11) enhances TGFβ-induced epithelial-mesenchymal plasticity and human breast cancer metastasis** [pubmed]
Daniel Garcia, **Christina Baek**, M Valeria Estrada, Tiffani Tysl, Eric Bennett, Jing Yang, John Chang.
Molecular Cancer Research, 2018

[6] **Inhibition of Spontaneous and Experimental Lung Metastasis of Soft-Tissue Sarcoma by Tumor-Targeting Salmonella typhimurium A1-R** [pubmed]
Shinji Miwa, Yong Zhang, **Kyung-Eun Baek**, Fuminari Uehara, Shuya Yano, Mako Yamamoto, Yukihiko Hiroshima, Yasunori Matsumoto, Hiroaki Kimura, Katsuhiro Hayashi, Norio Yamamoto, Michael Bouvet, Hiroyuki Tsuchiya, Robert Hoffman, Ming Zhao.
Oncotarget, 2014

Editor

[1] **High-Dimensional Data Analysis with Low-Dimensional Models: Principles, Computation, and Applications**
Yi Ma, John Wright
Cambridge University Press.

Teaching

CS189/289A Introduction to Machine Learning Summer 2019 – Spring 2021
Content TA, Spring 2021: Designed exam questions, held discussion sections. Instructor: Jonathan Shewchuk.
Project-Lead TA, Fall 2020: Designed the final project, studying a machine-learning perspective of the night sky and the evolution of our understanding of it across cultures and time. Instructor: Anant Sahai.
Head TA, Spring 2020: Lead course staff, wrote supplementary material. Instructor: Jonathan Shewchuk.
Content TA, Summer 2019: Designed exam questions, held discussion sections. Instructor: Jonathan Shewchuk.

CS170 Efficient Algorithms and Intractable Problems Fall 2019
Reader: Held weekly office hours and designed homework rubrics. Instructor: Satish Rao.

CS70 Discrete Mathematics and Probability Theory Spring 2018
Mentor: Held mini-discussion sections for a group of 4 students. Prepared students for exams. This was a part of UC Berkeley's Computer Science Mentors club.

Honors & Scholarships

CMU Presidential Fellowship in Machine Learning 2021
Awarded to 1 student per graduate school application cycle.

Outstanding GSI Award 2021
Awarded by UC Berkeley for outstanding work in teaching on campus.

Koret Research Scholarship 2020
Received $4000 *from UC Berkeley to conduct my proposed research with Professor Yi Ma over Summer 2020*

Thermo Fisher Scientific Scholarship 2016-2020
Received $\$20,000$ *for scholastic excellence.*

Eta Kappa Nu Honors Society 2018
National Electrical Engineering and Computer Science Honors Society.

Tau Beta Pi Engineering Honors Society 2017
National Engineering Honors Society.

Regents' and Chancellor's Scholarship 2016
Awarded to $< 2\%$ *of entering class for creativity and leadership.*

Relevant Coursework

STAT 240: Robust Statistics
STAT 210: Theoretical Statistics
EE 229: Information Theory
EE 227C: Convex Optimization
CS 285: Deep Reinforcement Learning
CS 288: Natural Language Processing
CS 270: Combinatorial Algorithms
MATH 140: Differential Geometry
MATH 104: Intro to Real Analysis
BIOE 145: Intro to Machine Learning in Computational Biology