

ABOUT ME	<p>AI researcher and engineer, with</p> <ul style="list-style-type: none"> ▷ first-author papers in NeurIPS and AISTATS, ▷ extensive open-source software contributions, and ▷ experience training neural networks at fast-paced teams. 	<p>Boston, MA</p> <p>mccarter.calvin@gmail.com</p> <p>(616) 272-0909</p>
EDUCATION	<p>Carnegie Mellon University, Pittsburgh, PA</p> <p><i>Ph.D. in Machine Learning</i> August 2013 - May 2019</p> <ul style="list-style-type: none"> ▷ Advisor: Seyoung Kim GPA: 3.80 ▷ Selected Courses: Probabilistic Graphical Models, Convex Optimization, Statistical Machine Learning, Foundations of Machine Learning Theory <p>University of Michigan, Ann Arbor, MI</p> <p><i>Bachelor of Science in Engineering</i> August 2009 - May 2013</p> <ul style="list-style-type: none"> ▷ Major: Computer Science, Minor: Mathematics GPA: 3.98 ▷ Selected Courses: Operating Systems, Computer Architecture, Database Systems, Numerical Methods, Linear Algebra, Theoretical Statistics 	
PUBLICATIONS CompArch CompBio ML	<p>J Michuda, et int., C McCarter, et int., T Taxter, “Validation of a transcriptome-based assay for classifying cancers of unknown primary origin”, <i>Mol Diagn Ther</i>, 2023.</p> <p>C McCarter and N Dronen, “Look-ups are not (yet) all you need for deep learning inference.” <i>Sparsity in Neural Networks Workshop</i>, 2022.</p> <p>R Hanson, D Martin, C McCarter, J Paulson, “If Loud Aliens Explain Human Earliness, Quiet Aliens Are Also Rare.” <i>The Astrophysical Journal (APJ)</i>, 2021.</p> <p>LE Fernandes, et int., C McCarter, et al., “Real-world Evidence of Diagnostic Testing and Treatment Patterns in US Breast Cancer Patients with Implications for Treatment Biomarkers from RNA-sequencing Data.” <i>Clinical Breast Cancer</i>, 2020.</p> <p>C McCarter, J Howrylak, S Kim, “Learning Gene Networks Underlying Clinical Phenotypes Using SNP Perturbations”, <i>PLOS Computational Biology</i>, 2020.</p> <p>C McCarter and S Kim, “Large-Scale Optimization Algorithms for Sparse Conditional Gaussian Graphical Models”, <i>AISTATS</i>, 2016.</p> <p>C McCarter and S Kim, “On Sparse Gaussian Chain Graph Models”, <i>NeurIPS</i>, 2014.</p> <p>S Moon, C McCarter, YH Kuo, “Active learning with partially featured data”, <i>Proceedings of the 23rd International Conference on World Wide Web (WWW)</i>, 2014.</p> <p>C McCarter, et int., B Haab, “Prediction of Glycan Motifs Using Quantitative Analysis of Multi-lectin Binding”, <i>Proteomics Clinical Applications</i>, 7: 9-10, 2013.</p> <p>D Chatterjee, C McCarter, V Bertacco, “Simulation-based Signal Selection for State Restoration in Silicon Debug”, <i>ICCAD</i>, 2011.</p>	
PREPRINTS	<p>C McCarter, “Towards Backwards-Compatible Data with Confounded Domain Adaptation”, arXiv preprint arXiv:2203.12720, <i>Under Review</i>, 2022.</p> <p>A Basumallik, D Bunandar, N Dronen, L Levkova, C McCarter, L Nair, D Walter, D Widemann, “Adaptive Block Floating-Point for Analog Deep Learning Hardware”, arXiv preprint arXiv:2205.06287, <i>Under Review</i>, 2022.</p>	
PATENTS	<p>D Bunandar, C McCarter, A Basumallik, “Improving the accuracy of analog linear processor.” US Patent App. 18/077,177 (2022).</p> <p>J Michuda, ..., C McCarter, et al., “Systems and methods for multilabel cancer classification.” US Patent App. 17/150,992 (2021).</p>	

EXPERIENCE	<p>Amazon Alexa AI, Applied Scientist June 2022 - Present Researching and developing neural retrieval and ranking methods for entity resolution. Training and deploying to newly-onboarding Alexa use-cases.</p> <p>Independent Project February 2022 - June 2022 Developed domain adaptation method which conditions on confounders, and explored founding a startup to ease deployment of vector embeddings across biology.</p> <p>Lightmatter, ML Scientist January 2021 - February 2022 Researched ways to accelerate deep learning inference on photonic AI accelerator. Explored finetuning strategies to ensure model accuracy despite hardware noise and quantization. Helped invent next generation of hardware architecture to improve noise robustness.</p> <p>Tempus Labs, ML Scientist June 2019 - January 2021 Created and validated a new batch effect correction method, which was deployed on the Tempus RNA-seq pipeline as the source-of-truth for all clinical AI models and pharma data deliveries. Developed a new topic model for gene expression deconvolution in metastatic cancers. Explored network learning methods and graph neural nets for gene expression networks and chromosomal rearrangement graphs.</p> <p>Carnegie Mellon University, PhD Student August 2013 - May 2019 Developed novel sparse graphical models and scalable optimization algorithms for disease systems biology. Used statistical learning to discover the gene regulatory networks which explain the effect of genetic variation on clinical traits.</p> <p>Van Andel Research Institute, Research Intern Summer 2013 Worked under the supervision of Brian Haab to apply feature selection method in pancreatic cancer biomarker discovery and to validate method on proteomics database.</p> <p>Google, Software Engineering Intern Summer 2012 Worked on server backend for Google Flight Search, developing functionality for international results for live Flight Search queries.</p> <p>University of Michigan, Research Assistant 2011 Worked under the supervision of Valeria Bertacco and Debapriya Chatterjee to develop post-silicon validation method. Designed and implemented parallel algorithm in CUDA.</p>
SELECTED OPEN-SOURCE CONTRIBUTIONS	<p>ConDo https://github.com/calvinmccarter/condo-adapter [author] Toolbox for Confounded Domain Adaptation.</p> <p>onnx2pytorch https://github.com/ToriML/onnx2pytorch [main contributor] Converts ONNX models to PyTorch.</p> <p>PerturbNet https://github.com/SeyoungKimLab/PerturbNet [author] Learns multi-omic gene regulatory networks.</p> <p>MLPerf Inference https://github.com/mlcommons/inference Deep learning benchmark. [memory-efficient pyramidal encoder for RNN-Transducer]</p> <p>matrix-completion https://github.com/tonyduan/matrix-completion [incremental singular-vector thresholding] Classical matrix completion.</p> <p>PyTorch https://github.com/pytorch/pytorch [added LazyInstanceNorm] Deep learning framework.</p> <p>nanopq https://github.com/matsui528/nanopq Product quantization (PQ) and optimized PQ. [eigenvalue allocation initialization]</p>
LANGUAGES	Python (PyTorch, TensorFlow, pandas, NumPy, Numba), Matlab, C++, C, CUDA

PRESENTATIONS	<i>Transcriptome background tissue correction in metastatic cancers using a correlated composition admixture model.</i>	
	American Association for Cancer Research (AACR), Annual Meeting 2020.	
	<i>An efficient algorithm for learning a gene network underlying clinical phenotypes under SNP perturbations.</i>	
	Genome Informatics meeting at Cold Spring Harbor Labs, November 2017.	
	<i>Penalised Additive Least Squares Models for High Dimensional Nonparametric Regression and Function Selection.</i>	
	Large-Scale Kernel Learning Workshop @ ICML 2015.	
TEACHING	<i>Probabilistic Graphical Models</i> (Teaching Assistant, CMU)	Spring 2016
	<i>Introduction to Machine Learning</i> (Teaching Assistant, CMU)	Fall 2015
ACTIVITIES AND PROFESSIONAL SERVICE	<i>Elicitation of latent knowledge (ELK) award contest</i>	February 2022
	AI Alignment Research Center (research proposal received honorable mention).	
	<i>Paper reviewing</i>	June 2016 - Present
	Reviewer for <i>NeurIPS</i> , <i>AISTATS</i> , <i>IEEE Internet of Things</i> , and <i>SciPy</i> .	
	<i>University of Pittsburgh Biomedical Informatics Training Program</i>	2017
	Mentor to undergraduate research intern through iBRIC program.	
	<i>Middle school science fair judging</i>	2015-2020
	Science fair judge for PA Junior Academy of Science and Chicago Public Schools.	
	<i>Machine Learning Department Student Research Symposium</i>	2014
	Member of organizing committee. Created website and helped plan symposium.	
	<i>English Language Institute Conversation Circle Program</i>	2011 - 2013
	Group leader of conversation circle for ESL students at University of Michigan.	
	<i>University of Michigan Robocup (Robot Soccer) Team</i>	2009 - 2012
	Member and team leader (2010-2011). Developed computer vision subsystem.	