

Joe Davison – curriculum vitae

Email joeddav@utah.cs.edu
Website joeddav.github.io
Google Scholar WJATKB8AAAAAJ
GitHub GITHUB.COM/JOEDDAV

Education

2023-Present **Ph.D., Computing (*emphasis: Artificial Intelligence*)**
University of Utah
advisor: Vivek Srikumar

2018-2020 **M.S., Data Science**
Harvard University

2015-2018 **B.S., Computer Science**
Brigham Young University

Experience

Jul. 2021 – Dec. 2022 **Enveda Biosciences**
Senior Machine Learning Scientist / Manager
Developed machine learning system for crucial component of drug discovery platform, providing industry-leading internal tools for molecular structure elucidation.

Feb. 2020 – Jul. 2021 **Hugging Face**
Research Engineer
Implemented state-of-the-art zero-shot classification tools, enabling text classification via the open-source TRANSFORMERS library and Hugging Face API endpoint in 100 languages without supervised training.

Jun. 2019 – Sep. 2019 **IBM Research**
Research Intern
Joe Davison, Kristen Severson, and Soumya Ghosh. Cross-population variational autoencoders. In *4th workshop on Bayesian Deep Learning (NeurIPS)*, 2019b

May 2018 – Aug. 2018 **Pluralsight**
Data Science Intern
Improved search and recommendation systems by developing universal model for embedding 5 educational content formats into common vector space. Created dashboard for comparison of models including DOC2VEC, FASTTEXT, LDA, and TF-IDF.

Sep. 2017 – Apr. 2018 **Zeff**
Machine Learning Engineer (*Part-time*)
Developed image embedding model trained with distributed MXNET framework on over 40 million images.

May 2017 – Aug. 2017 **Microsoft**
Software Engineer Intern
Expanded CNTK deep learning framework by creating R bindings to existing Python interface.

- Aug. 2016–
May 2017 **Qualtrics**
Software Engineer (*Part-time*)
Developed REDIS microservice for logging and maintenance of internal export utility.
- May 2016–
Aug. 2016 **Instructure**
Software Engineer Intern
Converted Pages module in Canvas app from outdated OBJECTIVE C to more modern SWIFT, updating with modern styling.
- Aug. 2015–
Apr. 2016 **Rollins Center for Entrepreneurship & Technology — BYU**
Web Developer / IBMC Student Co-Lead (*Part-time*)
Designed new website and led student team for 2016 *International Business Model Competition* held at Microsoft campus in Seattle.
- Jan. 2015
Aug. 2015 **Dept. of Microbiology & Molecular Biology — BYU**
Web Designer (*Part-time*)
Designed interactive website to increase student interest in department’s programs.

Publications

First Author

1. Joe Davison, Joshua Feldman, and Alexander Rush. Commonsense knowledge mining from pretrained models. In *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)*, November 2019a
EMNLP Oral Presentation: youtube.com/watch?v=NBYL7s8cVfw
2. Joe Davison, Kristen Severson, and Soumya Ghosh. Cross-population variational autoencoders. In *4th workshop on Bayesian Deep Learning (NeurIPS)*, 2019b
Workshop Poster: [dropbox.com/.../CPVAE Poster.pdf](https://dropbox.com/.../CPVAE%20Poster.pdf)

Contributing Author

3. Thomas Butler et al. MS2Mol: A transformer model for illuminating dark chemical space from mass spectra. *ChemRxiv preprint*, September 2023
4. Gennady Voronov, Rose Lightheart, Joe Davison, Christoph A Kretzler, David Healey, and Thomas Butler. Multi-scale sinusoidal embeddings enable learning on high resolution mass spectrometry data. *arXiv preprint arXiv:2207.02980*, 2022
5. Quentin et al. Datasets: A community library for natural language processing. In *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing: System Demonstrations*, November 2021, (**Best Demo Award**)
6. Thomas Wolf et al. Transformers: State-of-the-art natural language processing. In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing: System Demonstrations*, October 2020, (**Best Demo Award**)
7. Mark Hamilton et al. Flexible and scalable deep learning with mmlspark. In Claire Hardgrove, Louis Dorard, and Keiran Thompson, editors, *Proceedings of The 4th International Conference on Predictive Applications and APIs*, volume 82 of *Proceedings of Machine Learning Research*, PMLR, 24–25 Oct 2018

Highlighted Coursework

Harvard	Machine Learning for Natural Language Topics in Machine Learning: Deep Bayesian Models Machine Learning: Advances in Uncertainty Quantification, Structured Prediction, and Large-Scale Models	taught by Alexander Rush Finale Doshi-Velez Jasper Snoek et al.
BYU	Probabilistic Machine Learning Deep Learning: Theory & Practice Tools for Machine Learning Intro to Artificial Intelligence	David Wingate David Wingate Tony Martinez Jacob Crandall

Open Source Projects

2017	joeddav/devol Proof-of-concept repository enabling genetic search of ConvNet architectures represented by fixed-length vectors.
2017	microsoft/CNTK-R Proof-of-concept repository enabling genetic search of ConvNet architectures represented by fixed-length vectors. Note: Developed as 2017 Microsoft internship project
2016	joeddav/get_smarties Standalone tool for dummy variable generation in Python with full sklearn fit/transform compatibility.

Personal

Books	THE WAY OF KINGS SHŌGUN BORN A CRIME EDUCATED	by Brandon Sanderson by James Clavell by Trevor Noah by Tara Westover
Theater	HADESTOWN (2019) CAMELOT (1960) BRIGHT STAR (2014)	by Anaïs Mitchell by Alan Jay Lerner by Steve Martin and Edie Brickell
Languages	ENGLISH RUSSIAN	native conversational
Hobbies	piano, guitar, ukulele, racquetball, photography (nature & wildlife), hiking, skiing, reading	