JOSHUA T. ABBOTT

joshua.t.abbott@gmail.com | (484) 554-8383
joshabbott.github.io

Computational cognitive scientist specializing in using psychological theory to evaluate and improve machine learning models.

EXPERIENCE

Language and Cognition Lab 2022 - current
Research Scientist in Computational Cognitive Science *Berkeley, CA*

- Investigating how commonly used Large Language Models (LLMs) depart from theories of psychological semantics. This involves exploring why semantic fluency and typicality are not captured well from models trained on large text-corpora only.

University of Melbourne 2018 - 2021
Postdoctoral Research Fellow in Cognitive Science *Melbourne, Australia*

- Developed and led data-science research projects investigating semantic variation in word meanings across world cultures: curating datasets from sources in cognitive anthropology, ethnobiology, and ornithology, and using ML methods for analysis.
- Joint-led an international collaboration investigating how cognitive models of human generalization and few-shot learning can improve ML methods of recommendation and evaluation (e.g., in books, music, etc.).

Max Planck Institute for Human Development 2017 - 2018
Postdoctoral Fellow in Adaptive Rationality *Berlin, Germany*

- Developed and led research projects (with graduate students, postdocs, and senior PIs) investigating how the topology of semantic embedding spaces affects forecasting models of human judgment and decision making.
- Investigated systemic biases in conference abstract submissions by utilizing various NLP methods (e.g., topic models and semantic embeddings).

University of California, Berkeley 2010 - 2017
Research Scientist in Artificial Intelligence and Graduate Student Researcher *Berkeley, CA*

- As a Research Scientist in the Berkeley AI Research (BAIR) Lab, I led projects investigating how to transform deep neural network (CNN) representations closer to human semantic space representations (recovered from similarity judgments).
- As a Graduate Student Researcher, I led over 20 projects on behavioral modeling in categorization and language usage, exploring the effects of different semantic embeddings: utilizing theories and large-scale experiments from psychological sciences to build better ML models for recommendation and computer vision systems that behave more like people do.

RELEVANT SKILLS

Languages: Python, C/C++/C#, MATLAB
ML Tools: pytorch, tensorflow; (NLP) huggingface, spaCy; (Computer Vision) opencv; (Data Science) scikit-learn, pandas
Misc: linux, git, latex, mysql, postgres, php, AWS/mturk experiments

EDUCATION

University of California, Berkeley 2016
Ph.D. in *Cognitive Science,* Dissertation: "*Statistical models of learning and using semantic representations*"

University of Cambridge 2010
M.Phil in *Computer Science,* Thesis: "*Relevance feedback and novelty detection under the Bayesian Sets framework*"

New College of Florida 2009
B.A. *(Honors)* in *Computer Science,* Thesis: "*Temporal sequence analysis of Bottlenose dolphin vocalizations*"

SELECTED PAPERS

J.C. Peterson, J.T. Abbott, and T.L. Griffiths. (2018). Evaluating (and improving) the correspondence between deep neural networks and human representations. *Cognitive Science.* 42(8), 2648-2669. (Computational Modeling Prize in Perception and Action).

- Transforming CNN feature spaces towards more human-like semantic spaces based on human similarity judgments, providing a method to better capture how human experts conceptually represent novel domains from pixel space.

D.D. Bourgin, J.T. Abbott, and T.L. Griffiths. (2017). Towards More Human-Like Recommendations. In *Proceedings of the NIPS 2017 Workshop on Cognitively Informed Artificial Intelligence: Insights from Natural Intelligence*. (Spotlight Presentation).

- Evaluates Bayesian models of generalization from cognitive sciences on user playlist history as a novel method of recommendation and collaborative filtering.