Cara Van Uden

website: caravanuden.com
email: cara@cs.stanford.edu
github: @caravanuden

Education

Stanford University
MS in Computer Science
| Sept 2021 - June 2023 (expected)

- **GPA**: 4.12/4.30
- **Coursework**: Practical Machine Learning, Mining Massive Datasets
- **Teaching Assistant**: Data Management and Data Systems (1 term)

Dartmouth College
BA in Computer Science &
BA in Cognitive Science
| Sept 2015 - June 2019

- **GPA**: 3.87/4.00, summa cum laude
- **Coursework**: Data Structures, Algorithms, Databases, Computer Systems, Machine Learning
- **Teaching Assistant**: Introduction to Computer Science (6 terms), Foundations of Applied Computer Science (1 term)

Papers

- **Van Uden, C. E.**, Nastase, S. A., Connolly, A. C., Feilong, M., Hansen, I., Gobbini, M. I., & Haxby, J. V. (2018). "Modeling semantic encoding in a common neural representational space." *Frontiers in Neuroscience.*

Awards

- Phi Beta Kappa (2019)
- High Honors Thesis in Computer Science (2019)
- Neukom Award for Outstanding Undergraduate Research in Computational Science (2019)
- Academic Award in Cognitive Science (2019)

Skills

- **Languages**: Python, SQL, Scala, C, R
- **Libraries/Frameworks/Platforms**: PyTorch, PySpark, Spark MLlib, Keras, scikit-learn, Docker, Airflow, Google Cloud Platform

Work Experience

Wayfair
Software Engineer II - Machine Learning Platform | Feb 2020 - Aug 2021

- Built and deployed ETL and machine learning training/inference pipelines for computer vision, NLP, and multimodal product matching at scale.
- Redesigned and moved pipelines from on-prem to GCP.
- Was sole technical contributor for first three months; onboarded rest of the team. Collaborated with data scientists and other engineering teams.

Data Scientist I - Visual Similarity | Aug 2019 - Feb 2020

- Incorporated noisy environmental imagery into the "visually similar" product recommendation pipeline. Reduced product coverage gap by 50% and decreased runtime by 20%.
- Led "visual clustering" workstream - developed POCs for use cases of clustering visual embeddings across Wayfair. Example partner teams included marketing and search.

Celgene
Data Science Intern - Translational Data Science | Summer 2016

- Built an exploratory data analytics and visualization tool for analyzing gene expression and drug response data.
- Used site-wide by scientists for exploratory target deconvolution/validation in translational drug development for blood and bone marrow cancers.

Research Experience

Stanford Machine Learning Group (Ng), Center for AI in Medicine and Imaging (Langlotz), Biomedical Informatics Lab (Shah)
Research Intern | Oct 2021 - present

- Various deep learning, representation learning, and few-shot learning projects using patient images (X-ray and CT) and text (electronic health record).
- Deploying one such model in a hospital setting for pneumonia screening.

Dartmouth Computational Cognitive Neuroscience Lab (Haxby)
Research Intern | Jan 2018 - June 2019 (full-time Winter 2018)

- Used fMRI data alignment and forward encoding models to predict neural responses to naturalistic video stimuli across people (**paper**).
- Compared the representations learned by different CNN architectures to those of the human brain's ventral visual stream. Achieved state-of-the-art neural response prediction performance in late-stage visual areas (**thesis**).

Dartmouth Biomedical Data Science Lab (Hassanpour)
Research Intern | Jan 2016 - June 2017

- Built LSTM pipelines that extracted word and sentence embeddings from social media text data for substance abuse risk stratification.
- Combined CNN/LSTM ensemble estimated risk of alcohol abuse, and found social media data characteristics associated with high-risk alcohol use.

Projects

- Predicting sustainable development indices from geolocated text (graduate course project in 2021, received A+)
- Data quality validation and cleaning with Deequ (Wayfair Hackathon 2021, finalist)
- Visual similarity product clustering for browse (Wayfair Hackathon 2019, winner)
- Decoding neural representations of emotion in resting-state fMRI data (full-time internship at CMU in Summer 2018)