https://andrewnc.github.io **Andrew Carr** andrewcarr06@gmail.com

Relevant Experience

Gretel.ai - Senior Applied Research Scientist Jan 2022 - Present

- Developed, improved, and launched novel GAN model that used 90% less memory while maintaining generative quality
- Optimized metrics and reporting API which lead to a 3.1x speed up in report generation
- Lead creation of new synthetic data generative model that resulted in a 6x error reduction
- Researched and designed generative architecture that trained and generated data 400x faster than previous iterations
- Contributed broadly across the organization in content creation, customer success, and marketing efforts resulting in multiple successful customer engagements

OpenAI - Member Of Technical Staff (Fellow) May - Nov 2021

- Run evaluations and benchmark performance on Codex program synthesis models included in "Evaluating Large Language Models Trained on Code" and powering GitHub copilot
- Designed, researched, and created novel datasets in Pyspark resulting in a 2x model performance improvement. These datasets replaced old benchmarks and became the standard training sets used across the entire team
- Improved throughput of model inference by 66x and scaled to hundreds of GPUs using MPI, internal tools, and distributed Pytorch training
- Developed demos, novel visualizations, and internally showcased emergent model behavior using web technologies and the Python data stack
- Led proprietary research project on deep learning theory from ideation to completion, presenting findings to CTO and chief scientist, result adopted by and inspired multiple subsequent internal research projects

Google Brain - Research Science Intern May - Nov 2020

- Contributed key mathematical and algorithmic insights into a new self-supervised pretraining method that leverages recent advances in differentiable sorting for representation learning
- Defined new state of the art performance for unsupervised audio and vision based tasks
- Streamlined and standardized several scattered experiments across notebooks and code bases. This drastically increased team productivity and we were able to launch multiple large scale experiments daily
- Designed and developed two novel metrics to measure experimental success now used by the team to communicate our findings
- Orchestrated foundational experiments across thousands of GPUs and decreased data loading time from 5 minutes to 300ms
- Explored mathematical relationship between entanglement and optimal transport distance, presented work to team
- Explored the literature and successfully reproduced results from the field which increased our ability to iterate and improve upon existing research
- Published findings in IEEE Journal for signal processing

Lyft, Level 5, Autonomous Vehicles - Software Engineering Intern June - Aug 2019

- Developed A/B testing platform in high performant C++ to compare prediction models locally and in the cloud greatly increasing my team's development velocity
- Identified predictive features and developed real-time feature extraction system for use in machine learning pipeline
- Explored statistical and neural models for dynamical vehicle motion prediction leading to a 22.5% performance improvement
- Lead 3 engineers in exploratory 20% project for semantic code search
- Presented research to members of my team, explaining relevant topics and mathematics to apply to our technology stack

Qualtrics - Software Engineering Intern May - Aug 2018

- Achieved ∼96% accuracy with a .005% false positive rate, matching state of the art on phishing detection by researching and implementing system using sophisticated NLP feature engineering and machine learning
- Increased speed of system 3x resulting in a 63% reduction in hardware costs while handling 3 million daily requests by engineering asynchronous API using parallel processing and high performance computing techniques
- Identified, explored, and implemented state of the art emerging topic tracking system which allowed my team to reach their stretch goals for the quarter and led to a **patent**
- Built question similarity tool using sentence embeddings after collecting and curating a dataset of ∼130,000 questions. Improved f1 score from .3 to ∼.7 built using both structured and unstructured datasets
- The final estimated impact of my internship is $300k - 500k in yearly savings

Amazon Alexa Prize: Team Eve - Machine Learning Research Engineer Jan - Apr 2018

- Member of team Eve for the Alexa prize challenge. One of eight teams selected out of hundreds to research and build a social chatbot system to hold arbitrary conversation for 20 minutes on any topic
- Designed and built an offensive speech filtering system using probabilistic methods, which performed ∼3% better than current industry standards

- Researched and designed a complex sentiment analysis tool that classified sentences as having complex sentiment used for noteworthy knowledge retrieval

Perception, Control, and Cognition Lab - Deep Learning Researcher Dec 2016 - May 2020

- Lead multiple projects from inception to completion while mentoring students with a variety of skill levels which resulted in a number of novel contributions and publications
- 1^{st} place Student Research Conference presentation
- Explored intersection of probabilistic programming and parametric learning
- Developed a system to improve MRI quality using a denoising auto encoder
- Designed deep architecture to improve hearing aid quality resulting in signal to noise ratio increase of 197%

Private Capital Group - Software Engineer, Intern May - Oct 2016

- Developed web solutions to significantly increase back-office employee effectiveness by creating automated systems that resulted in yearly savings of over $200,000
- Collected, cleaned, and analyzed internal and external data which was built into reporting dashboards that tracked key business insights and allowed partners to make informed decisions
- Decreased product downtime by 47% by implemented full testing suite and fixing critical bugs

Domo Inc - App Assurance Intern Feb 2016 - Apr 2016

- Combined technical and business knowledge to ensure that new business related apps were useful to consumers
- Authored and created several Domo business apps, designed to answer user business questions

BYU Math Department - PDE Research Assistant Sept 2015 - Apr 2016

- Discovered proper boundary condition equations to more accurately model pressure waves using numerical methods, resulting in a method of approximation that was 3x faster than previous methods

Carnegie Mellon University - IT Lab Research Fellow June - Aug 2015

- Performed secondary research on police effectiveness in the presence of body cameras. We found a 70% decrease in violence on both sides when using body cameras
- Analyzed data from user studies and developed a custom web game to help local refugees learn English

Full-Stack Web Developer - BYU Studies Jan 2014 - Mar 2015

- Led a web team of 3 in maintaining a VB/ASP.NET website with thousands of unique visitors, increasing traffic and profitability by over 38%
- Managed large SQL databases while analyzing customer information to improve overall business increasing customer retention by 11%

Courses Taught or Assisted

BYU CS Department - Intro to Data Science TA Jan 2020 - Apr 2020

- Assisted in designing a course targeted to teach computer science students about the basics of data science, statistics, ML, and Python
- Held office hours, recitations, gave lectures, and assisted in grading
- Designed labs, curriculum, assignments, and tests

BYU Math Department - Optimal Control Theory TA Jan 2018 - Apr 2018

- Designed final project and labs for class while simultaneously enrolled to assist multidisciplinary professor

BYU Math Department - Competitive Coding Instructor Aug 2017 - Apr 2018

- Designed a course targeted to teach applied math students about technical problem solving while also teaching interview strategy, and various programming languages
- Resulted in 12 out of our 14 teams placed in the top 20 of the annual university coding competition
- Received a course rating of 4.8/5.0 which is 0.5 points higher than the department average

BYU CS Department - Graduate Deep Learning TA Jan 2017 - Apr 2020

- Assisted, over multiple semesters, in teaching a course targeted to help computer science students understand open problems in deep learning including methods and solutions
- Held office hours, recitations, gave lectures, and assisted in grading
- Designed labs, curriculum, assignments, and tests

Publications and Patents

M Chen., et al (2021) Evaluating Large Language Models Trained on Code. arXiv preprint arXiv:2107.03374

Carr, A., (2021). Everyday Data Science. Self Published Book - reached top 50 in Computer & Technology Education

Carr, A., Berthet, Q., Blondel, M., Teboul, O., Zeghidour., N (2020). Self-Supervised Learning of Audio Representations from Permutations with Differentiable Ranking. IEEE SPL

Carr, A. N. (2020). Geometric Extensions of Neural Processes. Master's Thesis

Carr, A., Nielson, J., & Wingate, D. (2019). Wasserstein Neural Processes. OTML Workshop NeurIPS arXiv:1910.00668.

Carr, A., & Wingate, D. (2019). Graph Neural Processes: Towards Bayesian Graph Neural Networks. arXiv preprint arXiv:1902.10042.

Text Analytic Notifications (2019) US 20055.538

Fulda, N., Etchart, T., Myers, W., Ricks, D., Brown, Z., Szendre, J., ... Carr, A., & Wingate, D. (2018). Byu-eve: Mixed initiative dialog via structured knowledge graph traversal and conversational scaffolding. Proceedings of the 2018 Amazon Alexa Prize.

Talks and Awards

Poster ProbAI Summer School 2022: Presented our work on effect of sampling methods on synthetic data quality score

#100 of 100 Most Influential People in AI 2021: Based on my written and spoken contributions over the year 2020

Workshop Poster NeurIPS 2019: Presentation of our work on Wasserstein Neural Processes

1^{st} **place Student Research Conference 2019**: Presentation of our work on Graph Neural Processes to peers and professors

President's Leadership Council Presentation: Selected by faculty and staff to represent my college's 4,000+ students by presenting my research to BYU's $1 million+ donors and top administration

Judge and Presenter 2019, 2020: Mentored hackathon participants, presented technology relevant to multiple projects, served on the judging panel at the Global Legal Hackathon

No Code Presentation 2019: Presented novel visual coding platform to group of students at the Global Legal Hackathon

Burton Scholarship 2017 - 2018: Full tuition academic merit scholarship

Excellence in Mathematics 2016: Nominated by research advisor for research contributions

Multiple internal speaking opportunities across companies in industry: OpenAI, Google, Lyft

Education

M.S. Computer Science; 4.0 2020
Brigham Young University *Provo, UT*

B.S. Applied and Computational Mathematics; 3.81 2018
Brigham Young University *Provo, UT*

Other Experience

1^{st} **place BYU AI Club Hackathon 2020**: Built a computer vision controlled robotic hand

2^{nd} **place BYU ACM Hackathon 2019**: Built a computer vision pong game that is controlled with hand detection

2^{nd} **place BI Wolff Hackathon 2018**: Built prescriptive ML solution to predict individual risk of becoming homeless

1^{st} **place BYU ACM Hackathon 2017**: Created Auto Dino program to perfectly play the chrome dino no wifi game

1^{st} **place BYU ACM Hackathon 2016**: Created *Mathify* app using polynomial interpolation to display text as math

1^{st} **place BYU ACM Summer Coding Competition 2018, 2019**

2^{nd} **place Global Legal Hackathon Utah 2018**: Made a chrome extension using NLP to summarize terms and conditions which I turned into a product, grew to 2000 active users, and sold

Python 3.8 Open Source: Fix small doc bug in cpython pull #11683

pyprobml Open Source: A primary contributor for Machine Learning a Probabilistic Perspective v2 Python code with Dr Kevin Murphy

Data Science Blog: 300+ monthly readers. Data science problems solved with esoteric programming languages

Ranked 8th in world: Tetris in spring of 2011