David Oniani

+1 (646) 565 1224 • onianidavid@gmail.com • www.davidoniani.com • oniani • davidoniani

For my research papers, projects, and full CV, please visit https://www.davidoniani.com.

Education

08/2017 – 05/2021 **B.A. Computer Science, Mathematics**, *magna cum laude*, *Luther College*, Decorah, IA.

Relevant courses: Machine Learning, Applied Statistics (I and II), Data Modeling and Querying, Linear Algebra, Advanced Algorithms and Data Structures, Programming Languages, Operating Systems and Architecture, Object-oriented programming with C++, Computational Models, Topology, Real Analysis, Game Theory.

Technical Skills

Languages: Python, Modern C++, Rust, Shellscript, Haskell, Lua, C, R, SQL, Java, Javascript, HTML, CSS.

Tools and Libraries: Linux, macOs, kitty, Neovim, RStudio, JupyterLab, numpy, PyTorch, TensorFlow, scikit-learn, spaCy, pandas, TensorRT, SQLite, Flask, Zola, Git, GitHub, GitHub Actions.

Other Skills: LaTeX, Markdown, Technical Writing.

Work Experience

02/2022 – Present **Staff Machine Learning Research Scientist**, *University of Pittsburgh*, Palo Alto, CA (remote work).

- First-authored and co-authored several papers to be submitted to leading AI/NLP journals.
- State-of-the-art research in machine learning (ML) and natural language processing (NLP).
- Lead architect of the NLP system aimed at facilitating state-of-the-art research in the US.
- Researching few-shot learning for sentence classification of the clinical text.
- Collaborating with other researchers and scientists to write, draft, and review research papers.
- Responsible for supporting various projects within the lab.

04/2021 – 01/2022 **Machine Learning Engineer**, *DawnLight*, Palo Alto, CA

Sequoia Capital Backed AI and Edge Computing Startup (Sold).

- Architected and trained highly performant CNN-based AI model for cough detection.
- Developed and maintained an AI sensing inference runtime (C++ codebase).
- Implemented wrapper APIs for interacting with annotation platforms such as Dataloop.
- Ported the bounding box SSD model to PyTorch Lightning.
- Wrote a set of tools for automating the pipeline of data engineering and model training.
- Implemented Voice Activity Detection (VAD) system using both AI and WebRTC inference.
- Together with the Principal Scientist, wrote a library of signal processing algorithms (libdsp).
- Coordinated efforts with the Firmware Team for low-level systems and hardware support.
- Continuously improved the documentation quality of the ML and Engineering team docs.

02/2020 – 09/2020 **AI Intern**, *Mayo Clinic, Kern Center*, Rochester, MN.

- First-authored several papers that got accepted at world's major AI and biomedical informatics conferences/journals (ACM-BCB, JAMIA, etc). Publications and code available on my website.
- Created automated COVID-19 screening tool based on decision trees to assist nurses and physicians.
- Built chatbot as extension of GPT-2 model by applying BERT, BioBERT, USE, and tf-idf.
- Utilized node2vec for generating COVID-19 network embeddings and built co-occurrence network.
- Wrote set of programs for extracting information from thousands of EHRs (Electronic Health Record).
- Set up Elasticsearch and indexed millions of documents for use in NLP algorithm.
- Worked directly under Dr. Feichen Shen and Dr. Yanshan Wang in the division of Dr. Hongfang Liu.

Research Experience

Fall 2019 **Directed research in programming languages with Dr. Alan K. Zaring.**

- Extended CCL language, wrote type-checking rules, made key contributions to type system.

Summer 2019 **Collaborative research on visual persuasion with professor Richard K. Merritt.**

- Trained convolutional neural networks for image recognition using PyTorch.
- Wrote Python scripts for PDF image/text extraction and data cleanup.
- Performed extensive set of both statistical and textual analyses using state-of-the-art algorithms.

Summer 2018 **Collaborative research on unit testing with Dr. Roman Yasinovskyy.**

- Automated feedback generation for C++ programming course.
- Redesigned and significantly improved SQL and relational algebra solution checker.
- Designed testable practice problems for algorithms and data structures course.

Honors and Awards

2018, 2019	Recipient of Two Luther College Dean's Office Summer Research Awards, Luther College.
All Semesters	Dean's List Recipient, Luther College.
2017	Selected Start-up (UnleashAR), TOP 200, Wolves Summit (Largest Startup Conference in EU).
2016	Gold Medal for Academic Excellence, Ministry of Education and Science of Georgia.
2011 – 2016	Five-time Finalist (TOP 20/3,000 back-to-back five times), National Mathematics Olympiad of Georgia.