David Oniani

+1 (646) 565-1224 ○ oniani.ai ○ onianidavid@gmail.com ○ oniani ○ davidoniani ○ @davidoniani

Visit oniani.ai for my papers, software, and more! I also make YouTube videos about AI and Software Engineering

Education

- **Luther College** — Decorah, IA
  B.A. Computer Science, Mathematics (magna cum laude) — Aug. 2017 – May. 2021

Technical Skills

- Languages: Python, Rust, Shell Script, Modern C++, PostgreSQL, Lua, Haskell, JavaScript, HTML, CSS
- Tools and Libraries: Linux, Neovim, Git, Docker, PyTorch, Transformers, scikit-learn, matplotlib, numpy, pandas, Flask, Zola
- Theory: Applied Statistics, Deep Learning, Information Theory, Linear Algebra, Multivariable Calculus, Real Analysis, Topology
- Other Skills: LaTeX, Markdown, Technical Writing

Work Experience

- **University of Pittsburgh** — Work Remotely from Brooklyn, NY
  Machine Learning Research Scientist — Feb. 2022 – Present
  - First-authored papers accepted at major AI/ML and health informatics conferences and journals
  - Led engineering and system design efforts for CDM-compliant research data warehouse (ReDWINE)
  - Automated Extract, Transform, Load (ETL) pipelines manipulating hundreds of GBs of data
  - Researched Few-Shot Learning and Transformers for accurate and efficient text classification
- **DawnLight** — Palo Alto, CA
  Research Engineer — Apr. 2021 – Jan. 2022
  - Developed and maintained AI sensing inference runtime for audio and radar (C++ codebase)
  - Architected and trained highly performant CNN-based DL model for human activity detection
  - Implemented Python APIs and data engineering pipelines for interacting with data platforms
  - Together with Principal Scientist, wrote C++ library of signal processing algorithms (libdsp)
  - DawnLight was Sequoia Capital Backed Ambient Intelligence and Edge Computing Startup
- **Mayo Clinic** — Rochester, MN
  Artificial Intelligence Researcher — Feb. 2020 – Sept. 2020
  - First-authored papers accepted at leading AI and health informatics conferences and journals
  - 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
  - Worked directly under Dr. Feichen Shen and Dr. Yanshan Wang in division of Dr. Hongfang Liu

Select Publications

- A Qualitative Evaluation of Language Models on Automatic Question-Answering for COVID-19
  Proceedings of the 11th ACM International Conference on Bioinformatics, Computational Biology and Health Informatics (ACM-BCB), 2020
  First Author, 30 citations
- Constructing Co-occurrence Network Embeddings to Assist Association Extraction for COVID-19 and Other Coronavirus Infectious Diseases
  Journal of the American Medical Informatics Association (JAMIA), 2020
  First Author, 19 citations
- Toward Improving Health Literacy in Patient Education Materials with Neural Machine Translation Models
  American Medical Informatics Association (AMIA) Informatics Summit, 2023
  First Author
- Social and Behavioral Determinants of Health in the Era of Artificial Intelligence with Electronic Health Records: A Scoping Review
  Health Data Science (HDS), 2021
  Co-Author

Honors and Awards

- National Interest Waiver (NIW) Green Card Immigrant Petition Approval for Exceptional Ability, USCIS (2022)
- Dean's List Recipient, Luther College (2017 – 2021)
- Recipient of Multiple Highly Competitive Scholarships, Luther College (2017 – 2020)
- Recipient of 2 Dean's Office Summer Research Awards, Luther College (2018, 2019)
- Gold Medal for Academic Excellence, Ministry of Education and Science of Georgia (2016)
- Multiple-Time Finalist, National Mathematics Olympiad of Georgia (2011 – 2016)
- 7th Place (TOP 10), National Mathematics Olympiad of Georgia (2014)
- IMO & IPhO Nominee, Ministry of Education and Science of Georgia, (2014)