Deniz Bölöni-Turgut

db823@cornell.edu (407)-480-6687 denizbt.github.io linkedin.com/in/deniz-bt github.com/denizbt

EDUCATION

B.S. Computer Science (GPA: 3.93) *Aug 2022 – May 2026*
Cornell University, College of Engineering *Ithaca, NY*

Selected Coursework: Object-Oriented Programming & Data Structures, Functional Programming, Honors Discrete Math, Analysis of Algorithms, Natural Language Processing, Deep Learning, Probability & Statistics, Linear Algebra, Stochastic Processes

EXPERIENCE

NLP & ML Researcher *Jan 2024 - present*
Cornell University, Computer Science Department *Ithaca, NY*

- Finetuned T5 LLM for 91.7% increase in ROUGE score for lay summarization of biomedical articles with HuggingFace transformers. Improved readability scores by 31% with prompt engineering and OpenAI GPT API. Evaluated self-bias and length bias in LLM responses. Contributed to open-source project commit0, a framework for training and running LLM agents to write Python packages from scratch given only text specification and unit test suite. Advised by Dr. Claire Cardie.

Software Development Engineer Intern *Jun 2025 - present*
Amazon.com *New York, NY*

- Intern for Amazon.com, working on Amazon Live Creator App.

Software Engineer & Technical Project Lead *Feb 2023 - present*
Cornell Data Science (Engineering Project Team) *Ithaca, NY*

- Collaborated with peers to design, test, and deploy machine learning software projects with real-world applications.
- **Education Chair (Executive Board Position)**. Developed course materials and programming assignments for introductory machine learning class of 50-100 students. Delivered 50 minute weekly lectures. Graded coding assignments.

Research Assistant *May 2023 – May 2024*
Carnegie Mellon University, Software and Societal Systems Department *Pittsburgh, PA*

- Compiled a 800GB+ dataset of obfuscated binary code, starting from open-source C-language repositories. Designed the data pipeline to automate GitHub scraping, obfuscation, and compilation with g++ and clang. Repurposed and refactored a large-scale Python codebase. Published paper in security conference DIMVA 2025. Advised by Dr. Claire Le Goues.

PROJECTS

MathSearch | *Python, AWS Lambda, AWS Sagemaker, Docker, YOLOv5* *Sep 2023 - May 2024*

- Developed a full-stack ML web application, MathSearch, to solve problem of identifying math equations in PDFs. Led an Agile team of 15 developers through design and implementation of product with an end of semester deadline. Implemented a CI/CD ML pipeline with YoloV8, MathPix API, and Levenshtein String similarity. Designed scalable and low-latency backend using message queuing service, AWS SQS. Deployed with Docker and cloud computing services AWS Lambda and Sagemaker.

OScrabble Command Line Game | *OCaml, Scrum* *Sep 2023 - Dec 2023*

- Created 1200 line game OScrabble in functional OCaml, inspired by the titular board game. Implemented single and multi-player, easy/hard mode, automatic scoring, and board UI. Collaborated with two developers using Scrum methodology and test-driven development. Achieved 95% line coverage by designing 100 unit test cases for bug-free gameplay.

FormulaOne Winner Predictor | *Python, PyTorch, NumPy, BS4, Pandas* *Feb 2023 - May 2023*

- Trained a model to predict the winners of the 2023 Formula One races, achieving a 61% accuracy. Used multilayer perceptron, Monte-Carlo simulation and custom loss functions implemented in PyTorch. Collected training data by scraping historical betting odds from multiple sources using BeatifulSoup4 and Selenium. Applied data mining techniques to enhance models.

TECHNICAL SKILLS

Languages: Java, Python, MySQL, OCaml, C/C++, HTML/CSS/Javascript, MATLAB, Bash, Linux, LaTeX, Eclipse, VS Code, IntelliJ

Machine Learning: PyTorch, NumPy, Pandas, HuggingFace Transformers, Sentence Transformers, Scikit-learn, SciPy, Matplotlib, OpenCV, Pillow, OpenAI/GPT API, Meta Llama 3, Beautiful Soup, Selenium, Wandb, Streamlit

Other: AWS Services (Lambda, Sagemaker, SQS, S3, EC2, CloudWatch), Docker, Git, Slurm, Figma