NICHOLAS PETOSA

(631) 759-6146 • nick.petosa@gmail.com
https://linkedin.com/in/petosa • https://github.com/petosa

EDUCATION

Georgia Institute of Technology, M.S. in Computer Science — August 2018 – December 2019
Specialization in Machine Learning — **GPA: 4.0/4.0**

Georgia Institute of Technology, B.S. in Computer Science — August 2015 – May 2018
Concentrations in Artificial Intelligence and System Architecture — **GPA: 4.0/4.0**

EXPERIENCE

Two Sigma — May 2019 – August 2019
Software Engineer Intern – Alpha Insights Team — *New York, NY*

- **Gradient-Boosted Trees.** Created a Python package that trains gradient-boosted tree classifiers to automate thousands of human decisions about data quality. Best model achieves an AUC of .70 on the noisy, limited data available. Package automates retraining with hyperparameter search on new data, so models can improve over time. Used scikit-learn and xgboost.
- **Deep Language Model Fine-tuning.** Implemented ULMFiT natural language transfer-learning to predict data quality from free-text descriptions. The prediction from this deep LSTM network is the most important feature in the gradient-boosted tree.

Microsoft — May 2018 – August 2018
Software Engineer Intern – PowerBI Advanced Analytics Team — *Redmond, WA*

- **Deep Learning.** Designed and trained a deep Keras LSTM classifier to predict English descriptions of arbitrary data visualizations with 95% accuracy. Trained on generated synthetic data and validated on real-world data.
- **Interpretable AI.** Created a web app that won 2nd place in the AI & Ethics category of Microsoft's 2018 week-long global company hackathon. The web app trains and visualizes deep PyTorch general additive models (GAMs) to build human-interpretable models.

Amazon — May 2017 – July 2017
Software Development Engineer Intern – Customer Account Protection Team — *Seattle, WA*

- **Fraud Detection.** Integrated a random forest classifier into Amazon's production sign-in systems, which handle millions of sign-ins per day. Model predicts if a customer cluster is malicious and surfaces predictions to fraud investigators. Used scikit-learn and Weka.
- **Data Visualization.** Created an account cluster d3.js visualization within dashboard used by hundreds of Amazon fraud investigators.

Cisco — May 2016 – August 2016
Software Engineer Intern – Global Support Experience Team — *Research Triangle Park, NC*

- **Data Science.** Created an R web app that visualizes partner performance. Executives estimated the tool would save $150k/quarter.
- **Web Development.** Redesigned a service wrapper with high visibility within Cisco with Java Spring.

Georgia Tech Graduate Teaching Assistant, CS 7646 Machine Learning for Trading — August 2018 – Present

- Grade assignments, answer questions, and hold regular office hours for the course.

RESEARCH

Georgia Tech Graduate Research Assistant, Quantitative Software Research Group — January 2018 – Present

- Research deep learning applied to quantitative finance under Dr. Tucker Balch, focus on timeseries classification and deep RL.
- Created a deep Q-learning trading agent with Keras. Agent was trained using walk-forward validation over historical price data.
- Implemented AlphaZero from scratch in Python, currently investigating applications to a stock market exchange simulator.

Georgia Tech Undergraduate Research Assistant, Sherrill Group — August 2016 – August 2017

- Developed a Flask app with MongoDB backend for a chemistry research package, published undergrad thesis on this research.

AWARDS

2nd Place (Home Depot Deep Learning Hackathon 2018) • **FINRA Data Analysis Prize (HackGT Hackathon 2017)**
1st Place (Coca-Cola Hackathon 2016) • **Firebase Prize (MHacks Hackathon 2016)** • **1st Place (SwampHacks Hackathon 2016)**

SKILLS

Programming Languages – Python, Java, JavaScript, C
Tools & Frameworks – PyTorch, Keras, TensorFlow, CNTK, d3.js, scikit-learn, Flask, Git
Areas of Interest – Quantitative Finance, Deep Learning, Machine Learning, Artificial Intelligence, Data Visualization