Nicholas Hanoian

Data Scientist • Denver, CO
nicholashanoian.com • nicholashanoian@gmail.com • (401) 935–9522

Experience

Milliman | Data Scientist — Sep. 2020–Present, San Francisco, CA

Flood Risk Consulting

- Performed review of systems and models used to price national rating plans for flood insurance, identifying potential areas of improvement. Led efforts to overhaul data pipelines and refine modeling strategies.
- These enhancements have massively increased the efficiency of the modeling process, yielding up to 50x reductions in analyst time and 10x reductions in computational runtime, all while providing improvements in model performance and transparency.

Telematics Research

- Produced minimal viable product for a novel type of telematics model designed to predict risk based on environmental factors as opposed to factors related to individual drivers.
- Developed efficient systems using PySpark and Databricks to clean and analyze large road networks containing up to one billion observations and hundreds of variables.
- Designed and analyzed XGBoost models to predict risk of specific routes, validating results against existing telematics products.

Milliman | Data Scientist Intern — May 2020–Aug. 2020, San Francisco, CA

- Researched and helped build telematics risk score used to price automobile insurance via XGBoost models.
- These models have since been approved by regulators for use in most of the U.S. and are sold by Milliman through an API.

University of California, Los Angeles | Deep Learning Student Researcher — June 2019–Aug. 2019, Los Angeles, CA

- Worked on a team of nine undergraduate researchers, led by a professor, to develop a topic-aware chatbot utilizing an attention-based structure and non-negative matrix factorization for topic modeling.
- arXiv preprint: https://arxiv.org/abs/1912.00315

University of Vermont | Machine Learning Final Project — Dec. 2019, Burlington, VT

- Designed a neural network-based model which predicts whether communities within Reddit are at risk of violating Reddit's content policy.
- Achieved over 85 percent recall on now-banned communities, which would result in an estimated 11x decrease in the number of communities which Reddit would need to manually moderate.

Education

University of Vermont | Bachelor of Science in Data Science and Mathematics — Sep. 2017–Dec. 2020, Burlington, VT

GPA: 4.0

Selected Undergraduate Coursework

Machine Learning, Data Structures and Algorithms, Data Science I & II, Probability Theory, Combinatorial Graph Theory, Database Systems, Data Ethics, Data Privacy, Experimental Design, Applied Multivariate Analysis, Survival and Logistic Regression.

Technical Skills

Languages

- **Proficient**: Python, R, SQL
- **Some Experience**: C++, Java, SAS, PHP, JavaScript, HTML, CSS

Libraries and Technologies

- **Data analysis and visualization**: Databricks, PySpark, Azure, NumPy, scikit-learn, pandas, Matplotlib, dplyr, ggplot2, Keras, Microsoft Power BI, Microsoft Excel
- **Web Development**: Django, React JS, Bootstrap

Tools | Git, UNIX, Make, CMake, Emacs, VS Code, LaTeX