# Neil Yeung

Data Engineer with ML & software engineering expertise, interested in scaling & shipping data-driven products & systems.
+1(650)-656-6380 | neil.y.yeung@gmail.com | Website | GitHub | LinkedIn | Google Scholar

## Education

**University of Rochester** — Rochester, NY
*B.S. Computer Science, Minor in Mathematics* | ***GPA: 3.7*** — *Expected May 2023*

## Skills and Relevant Coursework

- **Programming Languages**: Python, Scala, R, SQL, Java, C/C++, Javascript, Go, Kotlin, Assembly
- **Tools & Technologies**: Spark, AWS, Docker, Databricks, MLFlow, PyTorch, TensorFlow, Pandas, GraphQL, Git
- **Coursework**: Computer Vision, Data Mining, Data Structures & Algorithms II, Honors Linear Algebra, Statistics

## Experience

**Angelo Gordon** — May 2022 – Aug 2022
*Data Engineering Intern* — *New York City, NY*
- Built a distributed big data tool that processes 1 billion+ data points for hedge fund with 50 B AUM, shipped tool to production, and presented tool to CTO and other executives.
- Improved data pipeline for tool using multi-dimensional clustering and bin-packing, reducing query latency by 80%.

**Intel** — May 2021 – Dec 2021
*Machine Learning Engineering Intern* — *Santa Clara, CA*
- Developed a Vision Transformer model for segmentation, resulting in a filed patent.
- Created a containerized data pipeline for segmentation and object detection that ingests 1 million+ images.

**Intel** — June 2020 – Oct 2020
*Software Engineering Intern* — *Folsom, CA*
- Developed 5 novel statistics to measure performance for graphics performance simulator API.
- Created cloud-based dashboard to measure performance changes, using Elasticsearch and GraphQL to query simulator API, viewed 50+ times monthly.

**VISTA Lab** — Jan 2020 – Nov 2021
*Machine Learning Researcher under Prof. Jiebo Luo* — *Rochester, NY*
- Proposed generative Variational Autoencoder (VAE) video prediction model for medical images, collaborated with Amazon researcher, resulting in a co-first author paper accepted to *BME Frontiers 2022*.
- Conducted data mining research on time-series Twitter data, resulting in a first author paper accepted to 2020 *IEEE Big Data* (17% accept rate) conference and an interview on a popular data science podcast, Data Skeptic.

**CIeNet Technologies** — Summer 2016, 2017
*Software Engineering Intern* — *Santa Clara, CA*
- Developed white noise model to augment car speech recognition engine, accuracy increased by 50%.
- Developed Android music player app for infotainment system and shipped app to a major American car company.

## Papers & Patent

- Co-author of filed computer vision segmentation patent application: "Tools and Methods for Dual Attention Vision Transformers for Interactive Image".
- First co-author computer vision video prediction paper in journal *BME Frontiers 2022*: "Breast Cancer Induced Bone Osteolysis Prediction Using Temporal Variational Auto-Encoders".
- First author data mining paper in *IEEE Big Data 2020*: "Face Off: Polarized Public Opinions on Personal Face Mask Usage during the COVID-19 Pandemic".

## Activities & Awards

- **Academic**: Rush Rhees Scholarship (50% tuition), National Merit Scholar, Research & Innovation Grant ($4000)
- **Competitive Programming**: Passed North American Qualifiers ICPC 2021, Top 30 in ICPC Northeast Regionals
- **Other**: Amino Capital Fellowship, Won 3rd in Forbes Entrepreneurial Competition, Accepted to HackMIT 2022