Jay Zern Ng

🌐 Portfolio: https://jayzern.github.io

📍 601 W 57 St, New York, NY 10019

✉ jn2717@columbia.edu ⋄ 📞 +16463014870

EDUCATION

- **Columbia University** New York, NY
  M.S. in Data Science | GPA: 3.67/4.00 *Aug 2019 – Dec 2020*
  - **Software Systems**: Operating Systems, Programming Languages & Translators, Computer Networks, Intro to Databases
  - **Machine Learning**: Applied Deep Learning, Applied Machine Learning, Machine Learning with Probabilistic Programming, Statistical Inference & Modeling, Probability & Statistics, Exploratory Data Analysis & Visualization

- **University of Warwick** Coventry, UK
  B.S. in Data Science | First Class Honors *Oct 2016 – Jul 2019*

PROGRAMMING SKILLS

Languages: Python, C/C++, JavaScript, TypeScript, Java, OCaml, SQL **Web:** Node.js, Angular, React.js, Flask, Django
Data: Scikit-learn, TensorFlow, PyTorch, Numpy, Pandas, Spark, Kafka **DevOps:** AWS, Docker, Kubernetes, Airflow, Jenkins

EXPERIENCE

- **Nielsen** New York, NY
  Data Engineer I *March 2021 - Present*
  - **Application**: Core developer for Data Fusion team. Wrote Airflow DAGs to deploy the Fusion application, wrote unit/integration tests for CI/CD pipeline, and implemented new features for the Data Fusion product
  - **Cloud**: Migrated tools using internal cloud tools (Intelligence Studio) and Nielsen Data Lake (S3 columnar storage using Hive metastore). Implemented a Datacache to store intermediate tables, wrote a data pipeline to convert tables into parquet form for the Ingestion API, and fine-tuned Spark clusters memory to optimize performance

- **Boost** Kuala Lumpur, MY
  Software Engineer Intern *May 2020 - Sep 2020*
  - **Frontend, Backend, DevOps**: Boost is a lifestyle e-wallet app with 10 million+ users. Built a registration system using Angular, Node.js, Redux, RxJS. Deployed the web app and CI/CD using Kubernetes, Docker, AWS, Jenkins

- **Barclays** London, UK
  Summer Analyst *Jun 2018 - Aug 2018*
  - **Cybersecurity**: Worked on application security dashboards using Checkmarx and researched security methodologies based on OWASP. Restructured the application security database. Built a prototype chatbot using Django, DialogFlow

- **VLT Labs** Kuala Lumpur, MY
  Software Engineer Intern *Jul 2017 - Sep 2017*
  - **Frontend, Backend**: VLT Labs is a venture builder based in Southeast Asia. Built a chatbot using Python, Node.js, MongoDB, Angular and Facebook API to automate customer support for Uber Malaysia

PROJECTS

- **Type Inference Lisp** [OCaml, C++, LLVM] [Code] [Docs]:
  - Type Inference Lisp is a new functional programming language written in OCaml, C++ and LLVM, inspired by the Scheme language from MIT. It combines the power of a Lisp language and type checking, providing the best of dynamic and static typing

- **Bayesian Online Changepoint Detection** [Python, Numpy, GPFlow] [Code] [Paper] [Slides]:
  - My bachelor's thesis on time series. Implemented the Log Gaussian Cox Process model, with applications on weather forecasting data. Extended the model using sparse variational inference for faster inference and multitask learning for modeling covariances

- **Kitchen Operating Systems** [C, Advanced UNIX, Makefiles] [Fridge][Freezer][Pantry]:
  - For my OS class, I implemented a "Freezer" round-robin scheduler for multi-processor CPUs, a "Fridge" kernel key-value store with blocking/non-blocking syscalls, and a "Pantry" filesystem with fully featured read/write/delete operations using C, UNIX

- **Automatic event understanding using Deep Learning** [TensorFlow, GCP, Docker] [Code] [Paper] [Slides]:
  - Columbia Data Science Capstone with Bloomberg L.P., on understanding events using NLP based semantic role labelling and thematic fit. Researched architectures based on attention mechanism, wide-and-deep, ResNet, LSTMs and RNNs