Jay Zern Ng

https://jayzern.github.io

155 Claremont Avenue, New York, NY 10027
jn2717@columbia.edu ⋄ +16463014870

Education

Columbia University — New York, NY
M.S. in Data Science | Software Systems & Machine Learning | GPA: 3.67/4.00 — *Aug 2019 – Dec 2020*

- **Systems:** Operating Systems, Programming Languages & Translators, Computer Networks, Introduction to Databases
- **ML:** 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: 80.2% — *Oct 2016 – Jul 2019*

- Bayesian Forecasting, Bayesian Statistics, Stochastic Procceses, Multivariate Statistics, Neural Networks & Reinforcement Learning, Machine Learning, Mathematical Analysis I, II & III, Algorithms & Data Structures, Mathematical Statistics A & B

Programming Skills

Languages: C/C++, Java, Python, SQL, R, Javascript, OCaml
Web: Node.js, Angular, React.js, Flask, Django, PostgreSQL, MongoDB
DevOps: AWS, GCP, Docker, Kubernetes, Jenkins
Data: Scikit-learn, TensorFlow, PyTorch, Numpy, Pandas, Spark, D3.js, ggplot2
Others: Advanced UNIX, LLVM, Git

Work Experience

Boost — Kuala Lumpur, MY
Software Engineer Intern | Frontend · Backend · DevOps — *May 2020 – Sep 2020*

- Boost is a lifestyle e-wallet app with 10 million+ users. Built a registration system and RESTful API for 10k+ merchants using **Angular, Node.js, RxJS**, to accelerate digital adoption during Covid
- Deployed the web app and CI/CD using **Kubernetes, Docker, AWS, Jenkins**. Maintained SDLC lifecycle using **JIRA, Confluence**

Columbia Law School — New York, NY
Data Engineer (Research Assistant) | NLP · SQL — *Sep 2019 – Apr 2020*

- Researched the Harvard Caselaw Access Project. Wrote a data engineering pipeline for 150GB+ JSON/XML data using **Python, PostgreSQL** to study the origins of anglo-saxon law
- Implemented Named Entity Recognition models for NLP using **spaCy, Regex** for automatic detection of legal citations

Barclays — London, UK
Cybersecurity Intern | Checkmarx · SQL — *Jun 2018 – Aug 2018*

- Worked on application security (AppSec) dashboards using **Checkmarx** and researched security methodologies based on **OWASP**
- Restructured the AppSec database schema using **SQL** normalization and built a prototype chatbot using **Django, DialogFlow**

VLT Labs — Kuala Lumpur, MY
Software Engineer Intern | Frontend · Backend · NoSQL — *Jul 2017 – Sep 2017*

- 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. Implemented a frontend framework for a retailer using object oriented CSS

Selected Projects

TI-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

Automatic event understanding using Deep Learning [TensorFlow, Google Cloud Platform, 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 attention based, non-sequential (wide-and-deep, ResNet) and sequential (LSTM, RNN) architectures.

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