Ethan Ooi

Machine Learning Enthusiast

+60 19 488 0383
+44 78 3518 5692
ethanooiyien@gmail.com
linkedin.com/in/ethan-ooi

Experience

Regov Technologies, Kuala Lumpur — AI Developer

April 2021 - PRESENT

• Deliver Text/Voice chatbots on GCP Dialogflow for client acquisition. Architect and Deliver custom GCP components into our existing AWS Infrastructure to increase supported languages from 1 to >50 (Lambda, DynamoDB, AWS Connect, Aws Lex, Voximplant, Dialogflow)

• Developed CNN based Audio classifier with Spectrograms for Voicemail or Human Voice differentiation. Wrote blog post and solid docs (Tensorflow, PIL, Scipy, Dimensionality Reduction, CNN, LSTM, flask)

• Create version control system and oversee automated validation of malformations in client input data (Pandas, Python). Report directly to VP for Architecting and Orchestrating releases of our chatbot RICHA, with unit and QA Testing, and hold code quality and documentation up to standard.

• Conduct knowledge-sharing workshops on all the above for the team and VP, use this knowledge to deliver a full custom CI/CD workflow integrating components from Voximplant, GCP, and AWS infrastructure

Axiata Analytics Centre, Kuala Lumpur— Data Scientist

September 2020 - April 2021

• Used Pandas and AWS EC2, S3, and EMR to carry out cell-tower analytics and value creation through location planning for new cell towers. Delivered a robust ETL Pipeline to ingest data over multiple sources and produced development and planning insights for client EdotCo (Python and Linux)

• Spearhead R&D in ML based enrichment of current data sources (Keras)

• Lead the team in adopting AGILE Software Engineering principles such as Test-Driven-Development, Continuous Integration and Deployment, Version Control, Peer Review and Pair Programming

Atmos International, Manchester — *Associate ML Engineer*

October 2019 - March 2020 (Part Time 20h/week while completing Masters)

• Constructed a microservice architecture to ingest training data and setup training in parallel (Golang + Python)

• Adhere to the Google Protobuf communication protocol to create fault-tolerant APIs to connect microservices through RPC (Golang)

About Me

I am a creative and resourceful software engineer with an interest in the practical uses of collecting and refining data.

Passion in the creative discovery and productionization of neural networks.

I love linux and the open source movement, I run mostly only open source software, and use a unix based operating system as my daily driver

In my free time I live and breathe music, honing my skills on the instruments I play, or collaborating with members of the online music community

ML/Ops Skills

Pandas, Matplotlib, Numpy

Tensorflow and Keras

Autoencoders

Reinforcement Learning (Policy Gradient, DDPG, MDP, Q-Learning, Actor Critic)

CNN, RNN, Ensemble (Bagging/Boosting)

Computer Vision, Classification, Regression, and other Supervised tasks

Sentiment Analysis and Graph Embedding for NLP

CI/CD Linux/Git workflow

· Delivered AIGA- an automated gas advisory tool that currently operates the compressors on a gas pipeline going from Chile to Brazil- heavy use of Reinforcement Learning (DDPG, A2C, A3C in Tensorflow)

Microsoft Research Labs, Beijing — *ML Research Intern*

June 2019 - August 2019

· Produced reports comparing techniques in multiple research papers and how they can apply to our client- Agricultural Bank of China in dealing with credit-score estimation and fraud-ring detection using graph databases. Embedding, CNNs, Reinforcement Learning, and Adversarial Networks were applied in a graph database structure.

· Additional research in using LSTMs and Graph Networks to disambiguate user queries in a natural language model for an internal virtual-assistant project through querying a Knowledge Graph

Education

University of Manchester — *MEng in Computer Science*

September 2016 - May 2020 **High 2:1 (68%)**

Notable Postgrad Modules:
COMP60711 Data Engineering
COMP60532 Principles of Digital Biology
COMP62521 Agile and Test-Driven Development
COMP60542 Introduction to Health Informatics

Garden International School, Kuala Lumpur— *A levels*

September 2014 - May 2016 Maths (A), Physics (A), Computer Science (A), Music (A) (CIE)

Projects

Generative Music — *Undergrad Dissertation*

The goal of the project was to generate an endless stream of newly created "musical wallpaper".

Native JS and HTML implementations of music theory to produce a pleasing stochastic system of generative music
Mark Achieved for Project- 81%
https://gee842.github.io/genmusic

Hornstat — Personal Project
Run a team to deliver a discord bot serving more than 60,000 members of a community .

Large scale data processing of the leaderboards of the videogame "Brawlhalla". Creating meta and trend reports for the strongest characters and weapons based on data from top rated players.

Heavy use of Data ingestion pipeline (Node), Storage (Neo4j), Analytics (Scipy, Pandas, Numpy), and Display (Matplotlib, Seaborn).

Example of regional meta analysis: https://pastebin.com/GFzKCYx4

EC2, Lamda, Dialogflow

Languages
English (Native)
Mandarin (Working Prof.)

Notable Undergrad Modules

Algorithms Design and Analysis
Discrete Mathematics and Formal Logic
Graphics and Computer Vision
Machine Learning and Optimization
Distributed Computing and Systems
Compilers, Cryptography and Network Security

Programming Languages
GoLang (Infrastructure and Backend)

Python (NumPy, Pandas, Matplotlib,flask,discordPy)

Backend Javascript (NodeJS)

Java with Spring MVC and Maven

Hobbies

Self taught Jazz musician Band leader in 20+ Projects (Guitar and Drums)

Have music released on all major platforms under the artist name "ETHVN"

Top 30 competitor of eSport: "Brawlhalla"

eSports team captain for UoM at UK University Esports League "League of Legends" (Top 16)

References Available on Request