SEA RAN CLEON LIEW

linkedin.com/in/cleon-liew | srcliew@uwaterloo.ca | cleon01.github.io

EDUCATION

University of Waterloo — Waterloo, ON
Bachelor of Mathematics — Sep 2020 – Jun 2024

Major in Computational Mathematics; Minor in Computing
Cumulative GPA: 85/100; Graduated with Distinction
Relevant Coursework: Designing Functional Programs; Elementary Algorithm Design and Data Abstraction; Data Types & Structures; Computational Statistics and Data Analysis; Neural Networks; Data Visualization

WORK EXPERIENCE

Artificial Intelligence Software Development Research Intern — Waterloo, ON
University of Waterloo

FHIR Data Encoding Project under Prof. Bryan Tripp — Dec 2023 – Apr 2024

- Researched methods of encoding FHIR-formatted data to enhance transformer model performance
- Implemented methods (e.g., Spatial Semantic Pointers, Sinusoidal Encoding) to model the "importance function" of values

Metastatic Cancer Imaging Project under Prof. Subha Kalyaanamoorthy — May 2023 – Dec 2023

- Designed and developed deep learning programs using Pytorch, processing nerve cell images to predict cancer metastasis
- Applied models and techniques including Support Vector Machine, Random Forest, Vision Transformer, ResNet, Inception, Transfer Learning, Convolutional Autoencoder, and various mixed approaches
- Optimized and trained models on Compute Canada's High-Performance Computing (HPC) clusters

Deep Packet Inspection Quality Assurance Intern — Ottawa, ON
Nokia — Sep 2022 – Dec 2022

- Automated traffic detection testing for Web and Mobile platforms using Python, Selenium, and Appium
- Tested application filters to ensure proper capture of application traffic and application of Quality of Service (QoS) policies
- Inspected network packets using Wireshark to troubleshoot and verify the performance of the traffic detection system

Data Quality Assurance Intern — Calgary, AB (Remote in Toronto, ON)
Ontopical — Jan 2022 - Apr 2022

- Developed web scrapers using Scrapy, Splash, and Lua to support the Extract, Transform, Load (ETL) process for collecting and organizing multimedia content from various sources
- Tested web scrapers in AWS Batch and monitored their performance with CloudWatch
- Leveraged Git, Bitbucket, Jira, and Confluence for version control and task management within an Agile Scrum framework

Assurance Practice Intern — Hong Kong
PricewaterhouseCoopers — May 2021 - Aug 2021

- Conducted detailed testing and analysis of financial statements using Excel (VBA and Macro) and PwC's auditing software Aura, identifying trends, fluctuations, and discrepancies to inform audit conclusions

SKILLS

Languages: Python, R, SQL, C, Java, Bash, Lua, HTML, CSS
Machine Learning: scikit-learn, Keras, PyTorch, TensorFlow, Cuda, Hugging Face
Web Scraping and Automation: BeautifulSoup, Scrapy, Splash, Selenium, Appium
Collaboration Tools: Git, Bitbucket, Jira, Confluence, AWS Batch and Cloudwatch
Others: Linux, Putty, Wireshark, LaTeX, MATLAB, Tableau, Power BI, Excel (VBA, Macros), NumPy, Pandas

PUBLICATIONS

Use Of Subword Tokenization For Domain Generation Algorithm Classification
Cybersecurity 6, 49 (2023). (https://cybersecurity.springeropen.com/articles/10.1186/s42400-023-00183-8)

- Integrated scheme consisting of subword-based models for better classification of domains generated by various DGAs

Beam - An Algorithm For Detecting Phishing Link
2022 APSIPA Annual Summit and Conference, pp. 598-604. (https://ieeexplore.ieee.org/document/9979860)

- Developed an attention-based phishing detector by performing sub-word tokenization and fine-tuning the BERT model