NGUYEN VAN BINH

(+65) 83073859 | nguyen.binh@u.nus.edu | linkedin.com/in/nvbinh/ | github.com/nvbinh15 | nvbinh.com | medium.com/@nvbinh

SKILLS

Programming Languages	Python, C/C++, Java, Rust, RISC Assembly
AI/ML Domains and Tools	Computer Vision, NLP, Graph ML, MLOps | Scikit-Learn, PyTorch, TensorFlow, MLflow
Database/Data Processing	RDBMS, PostgreSQL, Hadoop MapReduce, Spark, DVC
Web Development	FastAPI, Django, Django REST Framework, HTML/CSS/JavaScript
Others	Docker, Cloud Platforms (Amazon Web Services, Azure), Computer Networks, Airflow, Linux

EDUCATION

National University of Singapore   **Aug 2020 – now (Expected: May 2024)**

Bachelor of Engineering, Computer Engineering, Minors in Management and Data Engineering

- Grade Point Average (GPA): 4.91/5.0 – Dean's List in all eligible semesters
- NUS Overseas Colleges Program, NOC Paris (2023): participated in entrepreneurship program at Université PSL

EXPERIENCE

Software Engineer, Machine Learning Intern, Vebits (Singapore)   **Jan 2024 – now**

- Train and quantize ResNet-50 and YOLOv7 with ByteTrack for image classification, object detection and tracking; build real-time video streaming system (1080p on RTSP server) integrating ML inference on edge devices with C++ and Python
- Create tools for camera calibration, image orthogonalization, image stitching, distance, and depth estimations with OpenCV, speeding up internal image processing workflow by 55%

Data Science Intern, Science Feedback (Paris, France)   **Feb 2023 – Aug 2023**

- Developed web domain credibility ranking models with graph neural networks adopting PyTorch and NetworkX, set up end-to-end ML pipeline, deployed models in production (making prediction for > 1 million domains) using FastAPI
- Established data infrastructure (data warehouse, computation, orchestration) on Azure Cloud Services, enhancing data security, quality, and integrity, improving efficiency of data and software teams, and fastening ML life cycle by 40%
- Evaluated and customised social media data scraping tools to gather insights for scientists doing fact-checking task

Teaching Assistant, School of Computing – National University of Singapore   **Aug 2021 – Dec 2022**

- Placed on the Honour List of Student Tutors for Excellence in Teaching with 4.8/5 feedback point (faculty average: 4.2)
- Conducted classes for Master-level IT5003 Data Structures and Algorithms (covered DSA concepts and implementation in Python) and CS2102 Database System (discussed Relational Algebra, ER model, SQL, and Database Normalization)

Machine Learning Engineer Intern, Neuron Mobility (Singapore)   **May 2022 – Aug 2022**

- Trained and deployed ML models for scooter parking image classification on AWS SageMaker achieving 95% accuracy and 0.98 precision score, detecting bad parking behaviours, enhancing user experience, and saving US$200,000 annually
- Proposed and coded backend for an in-house annotation platform utilizing Django REST Framework and PostgreSQL, streamlining collaboration between ML and data annotation teams, leading to time savings of 5+ hours per week

PROJECTS

Quora Duplicate Question Detection and Generation (source code | report)

- Led team of 6, built duplicate question identification models including RNN, GRU, LSTM, fastText and word2vec embeddings, and fine-tuned BERT family LLMs culminating in 90.1% test accuracy rate and 0.87 F1-score
- Constructed generative language models harnessing byte-pair encoding, n-grams, MLP, and decoder-only transformers accomplishing test 76.8 perplexity; augmented generated questions to classification task in a semi-supervised manner

Laser Tag (demo | source code – capstone project): Augmented Reality shooting game

- Collected data and created an ML model classifying actions based on IMU data mounted on players' hands attaining 96.14% test accuracy; translated model into high-level synthesis (C++) and ran inference on Avnet Ultra96-V2 board
- Designed architecture, optimized inter-component communication overhead through 20% data packet size reduction

AWARDS & HACKATHON

- Bronze Medal, European Physics Olympiad 2019
- Champions, National AWS DeepRacer League 2019: trained the best reinforcement learning model for autonomous car

EXTRA-CURRICULAR ACTIVITIES

- Community management, GoodHub SEA: connect with NGOs and social enterprises in Southeast Asia to deliver technical solution and consultation; organize quarterly community meetups having over 25 attendees
- Founder & Technical Writer, Binary Dissect – weekly newsletter: help hundreds of engineers level up by sharing intermediate and advanced programming and computer science concepts and techniques