Ngu Dang

Boston, MA, 02215 — dsngu97@gmail.com — nathandang97.github.io — linkedin.com/in/ngu-dang — github.com/NathanDang97

Education

Boston University **09/2020 - 05/2026 (expected)**
Doctorate in Computer Science
- **GPA:** 3.93/4.00 – Passed the PhD Candidate Qualifying Exam. Thesis Proposal scheduled in 11/2025.
- **Relevant Coursework:** Machine Learning, Natural Language Processing, Neural Networks and Deep Learning (DeepLearning.AI Professional Certificate), Data Science (IBM Professional Certificate).

Clark University **01/2018 - 05/2020**
Bachelor of Arts in Computer Science
- **Minors:** Data Science, Mathematics
- **GPA:** 3.93/4.00 – First Honors Dean's List in 2018, 2019, and 2020.
- Graduated with *Summa Cum Laude* and obtained *Outstanding Academic Achievement Award in CS*.

Skills

Programming: Python, Java, C++, MySQL.
Libraries: Pandas, Numpy, Scipy, PyTorch, TensorFlow, Keras, Scikit-learn, Seaborn, NLTK, Z3.
Script: HTML, CSS, LaTeX.
OS/Tools: Linux, Windows, Git, Jupyter, Google Colab, Visual Studio, Microsoft Office.

Projects

Tweet Dialect Classifier **06/2025 – 07/2025**
- Built a dialect classifying pipeline in Python with BERTweet-based model that distinguishes African American Vernacular English from Standard and regular African American English and achieved 0.95, 0.99, and 0.97 for accuracy, recall, and F1 score respectively.
- Integrated the classifier into a bias-aware sentiment analysis pipeline, with statistical analysis (Kruskal-Wallis H Test) to provide insights on fairness in interpretation of social media text across different models (i.e. RoBERTa, RoBERTa-Latest, BERTweet). Github link.

Real-Time Object Detector **05/2025 – 06/2025**
- Built a real-time object detection system by training YOLOv8 on Pascal VOC (Python) and implementing C++ ONNX Runtime inference with OpenCV for webcam-based detection.
- Applied transfer learning with pretrained YOLOv8n weights and integrated ONNX Runtime C++ API to deliver fast, resource-efficient object detection with dynamic bounding box visualization and minimal latency. Github link.

Human Activity Recognition Using Deep Learning **04/2025 – 05/2025**
- Built a deep learning pipeline using Python and PyTorch to classify human activities from Wi-Fi CSI data, achieving 0.98 accuracy score with a custom CNN-LSTM model.
- Designed a complete preprocessing workflow including reshaping, normalization, smoothing, and statistical feature augmentation to improve model robustness. Github link.

Churn Predictor for Subscription Services **03/2025 – 04/2025**
- Implemented an end-to-end churn prediction pipeline in Python for a video streaming service using a real-world imbalanced subscription dataset using an ensemble of three models — a neural network, XGBoost, and Random Forest — using weighted soft voting to optimize class ranking and maximize AUC.
- Engineered advanced features (e.g., ratio metrics, interaction terms behavioral buckets, etc.) on top of 20 given features to boost signal quality and improve model discrimination and achieved achieved a ROC AUC score of 0.75. Github link.

Experience

Boston University **09/2020 – present**
Graduate Research Assistant
- Working on joint projects on computational lower bounds and upper bounds of complex algorithms and design improvements on top of current state-of-the-art results.
- Implementing our algorithms, experiments, and scripts simulating Boolean circuits in Python using Z3 SAT-solver library.

Clark University **05/2019 – 05/2020**
Undergraduate Research Assistant
- Contributed to computer vision and computational geometry research projects for the Computer Science Department.
- Implemented experiments, statistical analysis (e.g. ANOVA, Kruskal-Wallis), visualization, and geometrical simulations in Python and Java.