

NGU DANG

SUMMARY

I am a fifth-year Ph.D. candidate in Computer Science. I am interested in Complexity Theory and Algorithm Designs, particularly the hardness and lower bounds of natural computational problems. During my undergraduate studies, I did some research on Computer Vision. I occasionally did individual projects related to Machine Learning, Data Science, and Natural Language Processing for personal enrichment.

EDUCATION

Department of Computer Science, Boston University

Boston, MA

*Ph.D. in Computer Science*2020 - 2026 (*expected*)

- Advisor: Prof. Steven Homer.
- Research area: Algorithms Design, Circuit Complexity, and The Minimum Circuit Size Problem (MCSP).
- GPA: 3.93/4.00.

Department of Computer Science, Clark University

Worcester, MA

B.A. in Computer Science

2018 - 2020

- Minors: Data Science and Mathematics.
- GPA: 3.93/4.00 — Graduated with Summa Cum Laude and High Honors.
- First Honors Dean's List in 2018, 2019, and 2020.

PUBLICATIONS & MANUSCRIPTS

1. Marco Carmosino, Ngu Dang, Tim Jackman. **Finding Circuit Extensions For XOR in Polynomial Time.** 2024. *Computational Complexity Conference 2025 (CCC' 25)*. Under Submission.
2. Marco Carmosino, Ngu Dang, Tim Jackman. 2023. **Minimal XOR Circuits: The One True Shape is a Binary Tree.** Unpublished Manuscript. This work was integrated with *Finding Circuit Extensions For XOR in Polynomial Time* above. A set of slides presenting this work can be found [here](#). A manuscript on this work can be found [here](#).
3. Mariah Papy, Duncan Calder, Ngu Dang, Aidan McLaughlin, Breanna Desrochers, and John Magee. 2019. **Simulation of Motor Impairment with “Reversed Angle Mouse” in Head-Controlled Pointer Fitts’s Law Task.** In *Proceedings of the 21st International ACM SIGACCESS Conference on Computers and Accessibility (ASSETS '19)*; ACM, Pittsburgh, PA, USA. DOI.

TEACHING EXPERIENCE

Teaching Fellow | Boston University

2021 - present

- CS131: Combinatorics Structures — Summer 2022, 2023.
- CS132: Geometric Algorithms — Summer 2022.
- CS235: Algebraic Algorithms — Spring 2021.
- CS237: Probability in Computing — Summer 2024.
- CS332: Theory of Computation — Spring 2023, Fall 2023, 2024.
- CS630: Advanced Algorithms — Fall 2021.

Grader | Boston University

2023 - present

- CS535: Complexity Theory — Fall 2023.

Undergraduate Teaching Assistant | Clark University

2018 - 2019

- CS120: Introduction to Computer Science — Fall 2018.
- CS121: Data Structures — Spring 2019.
- CS180: Automata Theory — Fall 2019.

PROJECTS

Churn Predictor for Subscription Service

. *Coursera's Challenge* — *Github Link*

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, log transforms, and behavioral buckets) on top of 20 given features to boost signal quality and improve model discrimination.
- The model achieved a ROC AUC score of 0.75 (91st percentile).

Edible Mushroom Classifier

. *Kaggle's Challenge* — *Github Link*

07.2024 - 08.2024

- Implemented a Random Forest model classifying edible mushrooms from toxic ones in Python based on their physical properties and characteristics such as the shape, size, and color of the mushroom's cap and stem, etc. which yields a good mix of numerical and categorical features for the data.
- The dataset used in this project (train and test) was generated from a deep learning model trained on the UCI Mushroom dataset. The training set contains 3116945 data points; the test set contains 2077964 data points, with 22 features.
- The model achieved an accuracy score of 0.987 on the hidden test set.

Disaster Tweets Classifier

. *Kaggle's Challenge* — *Github Link*

04.2024 - 05.2024

- Implemented a model classifying disastrous Tweets from regular ones in Python using DistilBERT by HuggingFace, which was trained on over 7000 tweets. The tweets data were scraped directly from X.com (formerly known as Twitter) where the contents were original and unprocessed (i.e. with many emojis, special characters, hyperlinks, typos, etc).
- The model achieved an accuracy score of 0.818 on the hidden test set.

Digit Recognizer

. *Kaggle's Challenge* — *Github Link*

10.2023 - 11.2023

- Implemented a Digit Recognizer model in Python using a Convolutional Neural Network (CNN), which was trained on the MNIST dataset consisting of handwritten digits (0-9).
- The model achieved an accuracy score of 0.988 on the hidden test set.

House Price Predictor

. *Kaggle's Challenge* — *Github Link*

07.2022 - 08.2022

- Implemented a House Price predictor model using CatBoost Regression where the data contains 2919 entries, each with 79 explanatory features describing most aspects of residential homes in Ames, Iowa, such as number of bathrooms and bedrooms, utilities, locations, front-yard and backyard condition, etc. which yields a good mix of numerical and categorical features for the data.
- The model achieved an RMSE score of 0.13 on the hidden test set.

| | |
|-------------------------|--|
| PAST EXPERIENCE | Undergraduate Research Assistant Worcester, MA 05.2019 - 05.2020 |
| | <ul style="list-style-type: none"> Contributed to computer vision and computational geometry research projects for the Computer Science Department. Implemented experiments, statistical analysis, visualization, and geometrical simulations in Python and Java. |
| | CMS Assistant Worcester, MA 04.2018 - 08.2018 |
| | <ul style="list-style-type: none"> Participated in building Clark University's new website on WordPress with the University's Marketing Department. Fixed 300 broken links as they were encountered and edited contents as needed. Handled tickets from other departments in the university that resolved their problems with accessing new website features. |
| SKILLS | <p>Programming: Python, Java, C, C++, MySQL.</p> <p>Libraries: Pandas, Numpy, Tensorflow, PyTorch, Natural Language Toolkit (NLTK), Keras, Scikit-Learn, Seaborn.</p> <p>Tools: Git, Jupyter, Google Colab, Visual Studio, Microsoft Office Suite.</p> <p>Scripting: LaTeX, HTML, CSS.</p> <p>OS: Windows, Linux.</p> <p>Languages: English (fluent), Vietnamese (native).</p> |
| CERTIFICATES | <ul style="list-style-type: none"> IBM Data Science by IBM on Coursera. Certificate earned at 08.31.2023. Neural Networks and Deep Learning by DeepLearning.AI on Coursera. Certificate earned at 12.31.2024. |
| AWARDS AND HONORS | <ul style="list-style-type: none"> Outstanding Academic Achievements, awarded by the Department of Computer Science at Clark University. Inducted to Phi Beta Kappa, Lambda of Massachusetts at Clark University on 05.24.2020. |
| ACADEMIC SERVICES | <p>Reviewer for: <i>Journal of Computer and System Science (JCSS)</i>.</p> <p>Organizer for: <i>Boston University Computer Science's Theory Seminar (Spring 2021)</i>.</p> <p>Vice President for: <i>Clark University Computer Science's Competitive Programming Club</i>.</p> |