TRI NGUYEN

nguyetr9@oregonstate.edu — github.com/ductri — linkedin.com/in/tri-nguyen-cs — 541-360-9255

EDUCATION

Oregon State University — Corvallis, OR, USA   2020 - Expected 2025
Ph.D. in Computer Science (GPA:3.85)
Relevant courses: Matrix Analysis, NLP with Deep Learning, Estimation, Machine Learning, Convex Optimization.

Ho Chi Minh City University of Technology — Vietnam   2012 - 2017
Bachelor of Science, Computer Science, (GPA: 8.72/10, Gifted program)

RELATED EXPERIENCE

Graduate Research Assistant   Mar 2020 - Present
Prof. Xiao Fu, Oregon State University,   *Corvallis, OR*
- Performed research projects on topics including deep learning, identifiability of matrix/tensor decomposition, convex optimization, and statistical learning that has resulted in 3 publications in top-tier conferences and journals.
- Presented one topic every quarter in internal reading group meetings, including: diffusion model, large language model, minimax analysis to expand and keep up-to-date knowledge in machine learning.

AI Engineer   Feb 2017 - Jan 2020
YouNet Group   *Ho Chi Minh City, Vietnam*

Sentiment Analysis
- Lowered manual workload of social media research team from 65% to 40% by training LSTM-based models with a continuous integration pipeline on an increasingly large dataset to perform text-based sentiment classification.
- Built on-demand scalable sentiment classification APIs using Docker, Kubernetes, and Google Cloud.
- Improved the reliability of crawling system to consume up to 100s GBs textual data/day with an error-tolerant system with proper logging, exception handling, and comprehensive unittests.

Customized Data Retrieval and Aggregation
- Reduced 100 support requests/week by building a web-based retrieval application supporting highly customizable search syntax and offering advanced operators such as 'not', 'and', 'or', '*'.

Sample Size Estimation
- Formulated equations to determine a reliable sample size essential for the business analysis team's work.

PUBLICATIONS

- **[ICML, 2023]** **T. Nguyen**, S. Ibrahim, and X. Fu. *Deep Clustering with Incomplete Noisy Pairwise Annotations: A Geometric Regularization Approach* .
- **[ICLR, 2023]** S. Ibrahim, **T. Nguyen**, and X. Fu. *Deep Learning From Crowdsourced Labels: Coupled Cross-Entropy Minimization, Identifiability, and Regularization.*
- **[TSP, 2022]** **T. Nguyen**, X. Fu, and R. Wu. *Memory-Efficient Convex Optimization for Self-Dictionary Separable Nonnegative Matrix Factorization: A Frank-Wolfe Approach.*

PROJECTS

- **Self-dictionary Learning**: Reduced the runtime of a Matlab version of a Frank-Wolfe algorithm by 10x, while using the same amount of memory by implementing a Mex function.
- **Sentiment Shifting**: Trained a Transformer model in an unsupervised, unpaired textual data setting as in machine translation to translate positive reviews to negative ones and vice verse.
- **Community Detection**: Demonstrated superior of Tucker decomposition over CDP decomposition in the context of rich structural community detection setting.
- **Nonlinear Least Square**: Derived a formulation for a matrix decomposition problem and implemented a Gauss-Newton algorithm to optimize a nonlinear least square objective function.
- **Vietnamese Accent Restoration**: Predict accents given accent-striped sentences in Vietnamese using LSTM.

SKILLS

Python, Matlab, C++; PyTorch, TensorFlow, Ray, Weights&Biases(wandb); Google Cloud, Docker, Kubernetes, Hive/Impala, Kafka; SQL, Pandas, NumPy; Git, Tmux, Vim, LaTeX.