

KIEN TRAN

Atlanta, GA 30332 • (+1) 470 923 7528 • ktran332@gatech.edu • master-data.science • github.com/trantrikien239

EDUCATION

GEORGIA INSTITUTE OF TECHNOLOGY

Aug 2022 - Dec 2023

Master of Science in Analytics – Computational Analytics track – GPA: 4.0/4.0

- Coursework: Machine Learning, Deep Learning, Web Search and Text Mining, Regression Analysis, Operations Research for Supply Chain Engineering I, High Performance Computing, ML with Limited Supervision, Deep Learning for Text Data

FOREIGN TRADE UNIVERSITY

Sep 2013 - Dec 2017

Bachelor's degrees – Double majors in Economics and Business Administration – GPA: 3.62/4.0

SKILLS

Programming: Python (pandas, scikit-learn, Pytorch), SQL, R, C++

Visualization: Pyplot & Seaborn; Microsoft Excel & PowerPoint

Deployment: MySQL, BigQuery, PySpark, Google Cloud Platform, AWS, Kubeflow, Docker, Git

Machine Learning: Gradient Boosting and Bagging trees; Linear models; Deep Learning including Multi-layer Perceptron, RNNs (GRU, LSTM), Transformers, Graph Neural Networks (GraphSAGE, Graph Attention), Convolutional Neural Network

EXPERIENCE

ZILLOW GROUP

Summer 2023

Applied Science Intern

- Applied boosted tree & transformer models to enhance renter personalization and gain insights into users and properties
- Enhanced overall performance by 3% (roc_auc & recall), with an 8% boost on a key user segment over a strong baseline

GEORGIA INSTITUTE OF TECHNOLOGY

Jan 2023 - May 2023

Graduate Research Assistant – Machine Learning

- Researched ranking models for eldercare equipment recommendation, improved performance by 20%

ONE MOUNT GROUP – Service provider in supply chain, real estate, and retails

Aug 2019 - Jul 2022

Lead Data Scientist (promoted from Data Scientist)

- Led creation, design, and development of customer data products, which enable personalized services for 11M end users
- Oversaw and mentored a data science team (task requirements, code review, model performance monitoring, etc.)
- Built an Auto-ML engine, using Catboost and mutual information, which would iteratively find and target potential customers to increase click-through rate ~2.5x for multiple marketing campaigns
- Developed a system of ML models, based on boosted tree, sequence modeling, and multi-task learning, to infer 70 common customer attributes which led to 2x increase in ML models' reusability
- Established a promotion budget allocation system using uplift modeling and integer programming to increase ~90% transaction count and save \$35K through a 75% reduction in excessive promotion cost

BOSTON CONSULTING GROUP

Jul 2018 - Jul 2019

Consulting Business Analyst

- Performed data analysis of 15 million customers for a large Asian bank, breaking revenue down to individual customers then reconstructing overall revenue streams, received approval to pilot the recommendations in 5 major branches
- Analyzed survey results to recognize clusters of various customer personas and build appropriate customer journeys

PROJECTS

- **Cetasearch:** A conversational search engine utilizing gpt-3.5 API and Sentence Transformer for sourced, chat-like results
- **Recommender system using GNN:** Implemented 2 link prediction models using GraphSAGE and GAT for the Instacart dataset, achieving 2% increase in roc_auc over a strong NCF baseline (Neural Collaborative Filtering)
- **Generative inpainting survey:** Researched and evaluated Diffusion (pre-trained Stable Diffusion) against DCGAN and Contextual Attention methods on the Places205 dataset. Diffusion had the best FID and the most natural completion.
- **Group recommender system:** Used Python, SQLite, and JavaScript to create a matrix factorization model (NRMSE = 0.14) and a web UI to recommend novel and relevant anime for groups of friends
- **Automated essay scoring model:** Utilized BERT, RNN, and traditional ML techniques to predict six analytic measures of essays. Implemented transfer learning and unsupervised clustering algorithms for improved performance and insights.
- **Sequence embedding:** Applied CBOW and GRU to sequences of grocery products to obtain item and user embeddings