Dat Nguyen

(801) 800-6538 | datnguyen@seas.upenn.edu | GitHub | Website

Experience

NVIDIA - GeForce NOW Santa Clara, CA

Data Science Intern *May 2021 – Sep 2021*

- Investigated users' feedback to identify **root cause** of game lag through different domain-specific metrics, improving customer's satisfaction score from 90 to 95, and decreasing number of complaints by 15% in three months
- Designed and productionalized four end-to-end ML pipelines to measure impact of game related issues, running 15 times faster with 1% margin of errors compared to manual methods
- Collaborated with 3 engineering teams to deliver weekly presentations to business stakeholders
- Proposed and developed personalized recommender using graph algorithm for 10 million users as self-initiated project

NBCUniversal Media - PeacockTV New York, NY

Data Science Intern *Jan 2021 – April 2021*

- Introduced and launched graph embedding model using node2vec to handle cold-start situation for live content, increasing nDCG@10 by 5.4%
- Optimized metadata enhancer model with BERT embedding, transforming movies' binary tags to weighted tags
- Initiated and improved **recommender system** based on movie synopses and click events using GCP stack

Deloitte - Transactions and Business Analytics Costa Mesa, CA

Business Intelligence (BI) Senior Consultant *Sep 2015 – Jul 2020*

- **A Major University of California** – Performed financial analysis on performance of 441 research grants
  - Revised and optimized SQL queries to extract data with 2 million rows and reduced retrieval time by 30%
  - Served as firm's liaison and successfully secured 2 follow-on deals with $500K worth of revenue
- **A Leading Alternative Investment Firm** – Led team of 8 to apply multiple valuation models for asset allocation
  - Implemented statistical analysis models to calculate the theoretical value of options using current stock prices
  - Utilized web scrapping techniques to download public data and streamlined whole process, resulting in 20% increment in research efficiency
  - Generated actionable insights from financial structured data, subsequently led to total saving of $640mm
- **A Global Digital & Printing Corporation** – Conducted data modeling to facilitate restructuring plan
  - Deployed ML models including random forest and **XGBoost** to predict churn rate, introducing 10+ new attributes and outperforming traditional model's F1 score by 4.1%
  - Created visualization dashboards in Excel and Tableau for data reporting, achieving a 25% higher attendance rate
- Mentored and provided technical training for 15 interns and junior consultants to accelerate onboarding process

Projects

Facial Expression Recognition | *Python, PyTorch, Detectron2, ResNet18*

- Implemented custom-built CNN and ResNet18 based CNN for facial expression recognition and achieved best accuracy score 35% higher than KNN model

US Car Accidents Severity Prediction | *Python, Scikit-Learn, LTSM, Apache Spark*

- Constructed **LSTM** network to pick up signals from text description and predicted severity with 91% accuracy, beating baseline model by 14%

Education

University of Pennsylvania Philadelphia, PA

Master of Science in Engineering, Data Science *Sep 2020 – December 2021*

- Honors: Dean's Master's Scholar (top 15%) — **GPA: 4.0/4.0**

Brigham Young University Provo, UT

Master of Science, Accounting; Bachelor of Science, Accounting *Aug 2010 – Apr 2015*

- Honors: Brigham Young Scholarship; Dean's List Award (top 5%) for academic excellence — GPA: 3.6/4.0

Technical Skills

Languages: Python, SQL/NoSQL, JavaScript, and R
ML & Big Data: Scikit-Learn, PyTorch, MLlib, MXNet, Spark, Hive, and Hadoop
Cloud Platform / Other: Databricks, GCP, BigQuery, Kubeflow, and AWS