DANH PHAN

Data Scientist

Full working rights in Australia

+61 416 978 987
danh.phan.mq@gmail.com
https://danhphan.net
/danhpt
/danhpt
/danhphan

Hi, my name is Danh Phan. Applying Machine learning and Big data analysis to derive actionable insights is my passion. I am a Data Scientist at the Pacific service team, Johnson Controls Australia. I am also a researcher working on various machine learning methods for intelligent transport systems at Monash University. I have contributed to open-source projects like PyMC, a popular probabilistic programming framework in Python.

EXPERIENCE

Data Scientist — Apr 2022 - Present
Data Analyst — Oct 2021 - Mar 2022
Johnson Controls Australia — ***Melbourne, Australia***

Develop Dynamic Pricing Models; Margin and Conversion rate optimisation; Customer churn modelling; Perform data analysis and build various KPIs and BI dashboards; Improve data quality, data governance and data security; Build and deploy End to End ML applications on Microsoft Azure.

Open-source software developer — Jan 2022 - Present
PyMC contributor — ***PyMC Github***

Develop Python functions, classes, and docker images for open-source projects that focus on Bayesian Modelling and Probabilistic Machine Learning, including packages like PyMC, Aesara, and PyMC-Examples.

Data Science Associate Instructor — Jul 2021 - Present
Data Fluency, Monash University — ***Melbourne, Australia***

Prepare technical materials; Teach hands-on workshops on machine learning, data analysis with Python, and data visualisation (Power BI) for professional staff and researchers at Monash University; Lead a group of four instructors.

Data Analyst Intern — Aug 2018 - Nov 2018
Data Department, Domain Group — ***Sydney, Australia***

Developed customer clustering models for marketing and sales teams; Performed data analysis using SQL, ETL, Python; Building various dashboards with Tableau.

Casual Teaching Academic — Feb 2018 - Nov 2018
Computing Department, Macquarie University — ***Sydney, Australia***

Tutored and consulted students on System Integration, and IT Systems & Projects.

IT Specialist — Sep 2013 - Oct 2016
Dept. of Information and Communications of NgheAn — ***Vietnam***

Developed and maintained web applications and SQL databases.

RECENT PROJECTS (please check other projects at https://danhphan.net/projects)

[1] Multi-output Gaussian Processes in PyMC (github.pymc.gp_experiments)

This project aims to add support for multi-output Gaussian Processes in PyMC.

[2] Aussie Social Sentiment Analysis (https://aussie-sentiments.herokuapp.com)

This project collects data from Twitter's APIs and build sentiment analysis models.

[3] Melbournian Daily Activities (https://melbourn-city.herokuapp.com)

This project visualises the daily activities of Melbournians in different areas.

EDUCATION

PhD Candidate on Intelligent Transport Systems Feb 2020 - Present
Institute of Transport Studies*, *Monash University ***Melbourne, Australia***
Supervisors: Prof. Hai Vu, Prof. Graham Currie
Topic: Machine learning for transport systems; Activity-based Agent-based models.

Master of Information Technology Feb 2017 - Dec 2018
Macquarie University ***Sydney, Australia***
Final internship: Data Analyst at Domain Group
***Core units*:** Machine Learning, Data Mining, Information Systems, Security Management
GPA 6.733/7 (4.0/4.0**); Top one** postgraduate student; **Academic Excellence Award**

Bachelor of Information Technology Aug 2005 - Jul 2010
Hanoi University of Science and Technology (HUST) ***Hanoi, Vietnam***
***Core units*:** Probability and Statistics, Algebra and Analytic, Numerical methods, Discrete Mathematics, Artificial Intelligence, Data structures and algorithms, Software Engineering.
GPA 8.3/10; High Distinction; Faculty rank 1/39

PUBLICATIONS

[**1**] Danh T. Phan, Hai L. Vu, Graham Currie. **AttentionChoice: Discrete choice modelling supported by a deep learning attention mechanism**. *The Transportation Research Board (TRB), 2022, under review.*

[**2**] Danh T. Phan, Hai L. Vu, Eric J. Miller. **A New Approach to Improve Destination Choice by Ranking Personal Preferences**. *Transportation Research Part C: Emerging Technologies, 2022.* https://www.sciencedirect.com/science/article/pii/S0968090X22002406

[**3**] Danh T. Phan, Hai L. Vu. **Activity pattern generation incorporating deep learning for transport demand models**. *2021*. https://arxiv.org/abs/2104.02278

[**4**] Danh T. Phan. **Smart Mobility Improvement: Classifying Commuter Satisfaction in Sydney, Australia**. *In Proceedings of the 3rd International Conference on Machine Learning and Soft Computing (ICMLSC 2019).* https://dl.acm.org/citation.cfm?id=3311021

[**5**] Danh T. Phan. **Australia Housing Price Prediction using Machine Learning Algorithms: The Case of Melbourne City, Australia**. *The 2018 International Conference on Machine Learning and Data Engineering*. https://ieeexplore.ieee.org/document/8614000

CORE SKILLS

Technical: Python, R, Java, C++, C#, .NET, SQL, ETL, Data Factory, Data Warehouse

Data analysis: Numpy, Jax, Pandas, GeoPandas, Scipy, NLTK, Dask, PySpark

Machine learning: Scikit-learn, Pytorch, Spacy, Statsmodels, PyMC, Aesara

Other tools: MySql, Sql Server, Docker, CI/CD, DevOps, Tableau, Power BI

Environments: Linux, AWS, Microsoft Azure, Google Cloud, Salesforce, ServiceMax

CERTIFICATES

- Microsoft Certified: Azure Data Scientist Associate - Microsoft Sep 2022
- Bayesian Data Analysis in Python - DataCamp Nov 2021
- Deep Learning Specialization - Coursera Sep 2021
- Scalable Machine Learning with Apache Spark - Databricks Jul 2021
- Discrete Choice Analysis: Predicting Market Demand - MIT Jun 2021