

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 working at the Pacific Service team, Johnson Controls. I have actively contributed to open-source projects like [PyMC](#), a popular Probabilistic Machine Learning framework.

## EXPERIENCE

**Data Scientist - Data Analyst**, [Johnson Controls](#), Australia Oct 2021 - Present

- Develop Dynamic Pricing models and Customer Attrition models for millions of dollars projects.
- Perform A/B testing, Build and deploy End to End models on Microsoft Azure.
- Perform data analysis and build various KPIs reports using SQL, Python and Power BI.

**Open-source Software Contributor**, [PyMC](#), Github Jan 2022 - Present

- **Google Summer of Code 2022**: Add supports for [Multi-output Gaussian Processes](#) in PyMC
- Develop features, fix issues and docker images for open-source projects that focus on Probabilistic Machine Learning, including packages like [PyMC](#), [Aesara](#), and [PyMC-Examples](#).

**Data Associate Instructor**, [Data Fluency](#), Monash University, Australia Jul 2021 - Present

- Prepare technical materials; Teach data-related [hands-on workshops](#) on Python, Git, Bash, High Performance Computing, and Power BI for researchers at Monash University.

**Data Analyst Intern**, [Domain Group](#), Australia Aug 2018 - Nov 2018

- Developed [a chat bot](#) and customer clustering models; Performed data analysis using SQL and Python on AWS RedShift and Snowflake; Building various dashboards with Tableau.

## CERTIFICATES

- [Microsoft Certified: Azure Data Scientist Associate](#) - Microsoft Sep 2022
- [Deep Learning Specialization](#) - Coursera Sep 2021
- [Scalable Machine Learning with Apache Spark](#) - Databricks Jul 2021

**RECENT PROJECTS** (For further information, please check <https://danhphan.net/projects>)

- Multi-output Gaussian Processes in PyMC: [danhphan.net/multi-output-gps](https://danhphan.net/multi-output-gps)
- Aussie Social Sentiment Analysis: <https://aussie-sentiments.herokuapp.com>
- Melbournian Daily Activities: <https://melbourn-city.herokuapp.com>

## EDUCATION

**PhD researcher**, Monash University, Australia Feb 2020 – Feb 2023 (Expected)

- Machine learning for intelligent transport systems. Received Monash Graduate Full Scholarship.

**Ms. of Information Technology**, Macquarie University, Australia Feb 2017 - Dec 2018

- Core units: Machine Learning, Data Mining. GPA **4.0/4.0**. Received Academic Excellence Award.

**RECENT PUBLICATIONS** (Please check my [Google Scholar profile](#))

- Danh Phan, Hai Vu, Graham Currie. AttentionChoice: Discrete choice modelling supported by a deep learning attention mechanism. *The 102<sup>nd</sup> Transportation Research Board*, 2023.
- Danh Phan, Hai Vu, Eric Miller. A New Approach to Improve Destination Choice by Ranking Personal Preferences. *Transportation Research Part C: Emerging Technologies*, 2022.
- Danh Phan. Australia Housing Price Prediction using Machine Learning Algorithms: The Case of Melbourne City, Australia. *The 2018 IEEE International Conf. on Machine Learning*, 2018.

## SKILLS

- **Technical**: Python, SQL, Bash, R, Java, C++, ETL, Data Warehouse, Data Lake
- **Data analysis**: Numpy, Jax, Numba, Pandas, Dask, Seaborn, Plotly, Tableau, Power BI
- **Machine learning**: Scikit-learn, PyMC, Pytorch, XGboost, Statsmodels, Transformers
- **ML Ops & Data Ops**: Docker, CI/CD, Github Actions, MLFlow, Airflow, PySpark, Dbt
- **Environments**: Linux, Microsoft Azure, AWS SageMaker, RedShift, Snowflake, Google Cloud