Uthaipon “Tao” Tantipongpipat

Machine Learning Engineer / AI Researcher

[uthaipon@gmail.com](mailto:uthaipon@gmail.com) <https://www.linkedin.com/in/uthaipon/>

# SUMMARY

Machine learning engineer and researcher with 4 years of experience after a PhD. Led a team on ML development by planning and executing quarterly milestones in ranking and recommendation systems. Experienced in responsible AI / AI ethics at Twitter. Has driven successful cross-functional projects (resulting in 3+ billions press click-read) and impacted engineers by applying research to set company-wide metrics. Track record (10+) of top-tier peer-reviewed publications in ML, algorithms, and statistics, with emphasis on ML theory, ML fairness, and optimization, with 1st place award from the US government on differential privacy.

# EXPERIENCE

## Agoda, Bangkok, Thailand *- Lead Data Scientist* Jan 2023 – Now

* Implemented and optimized TensorFlow Decision Forest model, and transformed and optimized RNN with attention dataset, improving offline validation metrics by 1-2%.
* Initiated and led a research project to develop a user-specific parameter in the ranking model, integrating customer-loyalty components, resulting in 50% improvement in MSE for predicting customer loyalty and future profit.
* Engineered a CTR (click-through-rate) prediction model, reducing RMSE by 75%. Provided strategic insights to the advertising team for optimizing ad revenue from clicks.
* Designed, implemented, and maintained Scala Spark jobs and SQL alerts for dataset monitoring for models, including anomaly detection and dashboards for trends on user behaviors and devices.
* Managed and strategized with managers, data scientists, and ML engineers on deployment design, decisions, and timelines of the project for the team’s milestones.

## Twitter, remote US *- Machine Learning Researcher* Jun 2020 - Jan 2023

* Led Twitter’s image cropping algorithmic bias audit resulting in a published academic paper and $1.5M press ad equivalency and 3B readership from 500 news articles in 49 countries. Led to another follow-up work by team members resulted in additional $1.4M, 2.7B reads, and 800 articles from 47 additional countries, and contributed to the decision to remove the algorithm in production.
* Proposed a 13-18% precision-recall video classification model improvement with no additional cost to partnering team to fix offensive misclassifications on Tweet topic annotations and discovered correlation bias with demographics despite a lack of private individual data.
* Established a data-driven guideline for company-wide engineers to adopt an inequality metric in A/B statistical testing and got business approval from leadership to finally deploy the metric.
* Provided statistical analysis to customer teams to evaluate and quantify bias in ML models; redesigned common ML statistical significance tests required for bias measurement.
* Published two papers in social computing conference and one in data science journal.

## Microsoft, Redmond WA *- Research Intern* May 2019 - July 2019

* Implemented privacy guarantee on large-scale natural language processing models (RNNs and LSTMs) to protect against personal deidentification due to model usage.
* Researched private correlation clustering algorithm, private submodular optimization, and surveyed literature for private stochastic gradient descent for training deep models.

# EDUCATION

## Georgia Institute of Technology, Atlanta GA Aug 2016 - May 2020 *PhD in Algorithms, Combinatorics, and Optimization (ACO). GPA 4.00/4.00*

## University of Richmond, Richmond VA Aug 2012 - May 2016 *BS in Mathematics, with Thesis (Algebraic Combinatorics). GPA 3.97/4.00*

## University of Oxford, Oxford UK Oct 2014 - Jun 2015 *Study Abroad Program in Mathematics and Computer Science. First Class.*

# 

# SELECTED PUBLICATIONS

* Measuring Disparate Outcomes of Content Recommendation Algorithms with Distributional Inequality Metrics Patterns Journal 2022
* Volume Sampling and Approximation Algorithms for A-Optimal Design MathOR 2022; SODA 2019
* Image Cropping on Twitter: Fairness Metrics, their Limitations, and the Importance of Representation, Design, and Agency CSCW 2021
* Fast and Memory Efficient Differentially Private-SGD via JL Projections NeurIPS 2021
* Differentially Private Mixed-Type Data Generation for Unsupervised Learning IISA 2021
* Combinatorial Algorithms for Optimal Design COLT 2019
* The Price of Fair PCA: One Extra Dimension NeurIPS 2018

See Google Scholar <https://scholar.google.co.th/citations?hl=th&user=nzO_5FMAAAAJ> for the full list.

# SELECTED AWARDS

* **Impact Recognition** Award, CSCW (the social computing conference) 2021
* **Best Reviewer** of NeurIPS (top-tier machine learning conference) 2019
* **1st Prize** and **People’s Choice** Awards, The Unlinkable Data Challenge, National Institute of Standards and Technology (NIST), US Department of Commerce 2018
* **Honorable Mention** (top 2.5%), William Lowell Putnam Mathematical Competition 2015
* **Bronze Medal** and **Honorable Mention**, Asia-Pacific Mathematics Olympiad (APMO) 2010, 2011
* **Gold** and **Bronze Medals**, IWYMIC International Mathematics Competition 2008, 2009

# SKILLS

**Technical:** Python (pandas, numpy, scipy, sklearn), Scala, PySpark, Scala Spark, SQL (BigQuery, Impala), Java, C++, Tensorflow, PyTorch, Kubeflow, Hadoop, Git, GCP, Superset, Grafana, Mathematica, LaTeX

**Languages:** Thai (native); English (full proficiency)