# Arun Pa Thiagarajan

https://www.arunppsg.in/

Email : arunppsg@gmail.com

Github: arunppsg


## Education

- **PSG College of Technology** — Coimbatore, India
  *Integrated Masters in Data Science; GPA: 3.63 (9.08/10.0)* — *July 2016 – May 2021*

## Publications, Posters and Patent

- Advika Vidhyadhiraja, Arun Pa Thiagarajan et. al, *Open Source Infrastructure for Differentiable Density Functional Theory* (workshop paper) In: SynS & ML Workshop @ ICML 2023.
- Bharath Ramsundar and Arun Pa Thiagarajan, *Chiron: A Cloud Scientific Machine Learning Programming Environment*, 2022 (patent), United States Provisional Patent Application.
- Arun Pa Thiagarajan, *Potential Biases in Using Machine Learning for Healthcare Applications* (poster). In: RBCDSAI-FCAI Conference on Deployable AI, 2022
- Arun Pa Thiagarajan et. al *Data-Driven Analysis of Food Corporation of India's Operations and Policy Recommendations*, 2021 (unpublished work, report).

## Experience

- **Deep Forest Sciences** — Fremont, CA
  *Machine Learning Engineer (remote)* — *Mar '22 - Present*
  - Developed and benchmarked graph neural networks and other deep learning algorithms for molecular property prediction tasks.
  - Designed and constructed a machine learning platform for computational drug discovery applications utilizing AWS services.
  - The platform incorporates a web interface built using the React framework, a backend utilizing REST APIs to process user requests, a MySQL database for storing user data, and AWS Batch Service for managing machine learning tasks like model training and inference.
- **Indian Institute of Technology, Madras** — Chennai, India
  *Project Assistant* — *July '21 - Feb '22*
  - Developed a malware detection framework for detecting malware from network traffic data. The framework integrated machine learning algorithms for classifying network flows, algorithms for detecting suspicious domain names based on DNS patterns, and TLS fingerprinting techniques to identify malicious servers.
  - Created an open-source packet logging tool capable of capturing and parsing incoming network packets at high speeds (tested up to 50 MB/sec) utilizing ring buffer and memory map (code).
- **Tata Consultancy Services** — Chennai, India
  *Research and Development Intern* — *May '19 - Nov '19*
  - Analyzed time-series sales data and constructed time-series forecasting models to predict future sales.
  - Investigated dynamic pricing strategies in e-commerce sector using reinforcement learning techniques and developed a Q-Learning model that outperformed static pricing techniques by 11% in terms of revenue for the seller.

## Open Source Contributions

- Contributed machine learning model architectures to the open scientific machine learning library DeepChem and enhanced the testing infrastructure by contributing to CI/CD pipeline and expanding test coverage (pull requests).
- Implemented the paper Time Series Anomaly Detection using Generative Adversarial Networks for performing anomaly detection in time series data (code).
- Contributed code patches to PyTorch (contributions), a widely used deep learning framework, and PyTorch-Geometric, a deep learning framework designed for implementing graph neural networks (contributions).

## Talks

- Gave a lightning talk at Pycon India 2021 about DeepChem and a 20-min talk at December 2022 monthly FOSS United Bangalore meetup on the same.
- Gave a talk on Technical Documentation in Regional Languages at FOSS Goa, 2023 meetup

## Others

- Earned an A+ grade in the Deep Generative Models online course offered by the Center of Continuing Education at the Indian Institute of Science, Bangalore, in May 2023.
- Awarded Achievement Award for outstanding curricular, co-curricular and extra-curricular achievements in MSc Data Science class of 2021.
- Attended ACM Winter School on Cybersecurity held in Dec 2019 at NISER, Bhubaneshwar
- Conducted *Cricket and Statistics* program at Mango Education for kids aged between 11 - 15.
- Volunteer at FOSS United - a non-profit organisation which promotes open source software ecosystem in India.

## Programming Skills

- **Programming Languages**: Python, C, C++
- **Technologies and Frameworks**: MongoDb, MySQL, PyTorch, Git, Github Actions, Docker, AWS Services

Last updated: November 28, 2023