Arun 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*

EXPERIENCE

- **Deep Forest Sciences** Fremont, CA
  Machine Learning Engineer (remote) *Mar '22 - Present*
  - Designed and built a maching learning platform for computational drug discovery applications on AWS Batch service.
  - Built RestAPIs using FastAPI for submitting jobs to the platform and developed backend services.

- **Indian Institute of Technology, Madras** Chennai, India
  Project Assistant *July '21 - Feb '22*
  - Built a malware detection framework and implemented predictive modelling algorithms for detecting malware from network traffic data.
  - Built an open-source packet logging tool which can capture and parse network packets at high speed (tested upto 50 MB/sec) and detect duplicate packets using a BloomFilter(code).

- **Tata Consultancy Services** Chennai, India
  Research and Development Intern *May '19 - Nov '19*
  - Analyzed time series sales data and built time-series forecasting models for prediction future sales.
  - Studied dynamic pricing of interdependent and perishable products using reinforcement learning techniques and proposed a Q-Learning model which achieved a yield of 16% higher than passive pricing techniques.

- **in-d.ai** Chennai, India
  Machine Learning Enginnering Intern *June '20 - July '20*
  - Developed an AI engine for classifying documents and used NLP techniques for named-entity extraction. The solution was deployed for real-time use using a Flask based web framework.
  - Used IAM Handwriting dataset and augmented it to build a robust convolutional netural network for recognizing handwritten words to recognize handwritten characters (accuracy achieved: 64%).

PROJECTS

- Presented a poster on Potential Biases in Using Machine Learning for Healthcare Applications in RBCDSAI Deployable AI - 2022 conference (poster).
- Contributed machine learning model architectures to the open scientific ML library DeepChem and improved testing infrastructure by contributing to CI/CD pipeline and increasing test coverage (pull requests).
- Performed a data driven analysis of Food Corporation of India's operations by collection open data and suggested policy recommendations to efficiently utilize foodgrains (report).
- Implemented the paper Time Series Anomaly Detection using Generative Adversarial Networks for performing anomalye detection in time series data (code).
- Awarded Achievement Award for outstanding curricular, co-curricular and extra-curricular achievements in MSc Data Science class of 2021.
- Conducted *Cricket and Statistics* program at Mango Education for kids aged between 11 - 15.

PROGRAMMING SKILLS

- **Languages**: Python, C, C++
- **Technologies and Frameworks**: MongoDb, MySQL, PyTorch, Git, Github Actions, Docker, AWS EC2, S3, AWS Batch