Alvin Pan

1 River Ct, Apt 1709, Jersey City, NJ 07310 | (412)708-5115 | qp2134@columbia.edu

EDUCATION

Columbia University — New York, NY
M.S. Computer Science — Expected 5/2023
Research Advisors: Krzysztof Choromanski, Matei Ciocarlie, Shuran Song
Coursework: Computer Vision, Natural Language Processing, C++, Advanced Algorithms, Causal Inference, Database

Carnegie Mellon University — Pittsburgh, PA
B.S. in Machine Learning & Statistics, Math — 6/2021
Advisors: Katerina Fragkiadaki, Peter Freeman, Rebecca Nugent
Coursework: Machine Learning, Probabilistic Graphical Models, Deep Reinforcement Learning, Optimization, Linear Algebra, Calculus, Analysis, Statistics, (Sequential & Parallel)Algorithms & Data Structures, Python, C, SML

EXPERIENCE

Machine Learning Scientist — New York, NY
NYU Langone — 8/2023-Current
- Building GPT backbone for ALL Langone physicians(medical note generation)
- Code up videoGPT & VQGAN from scratch for video based RL
- Extend dreamerV3 from image to video data
- Construct multi-head policy model on dreamerV3 in MineDojo
- Implement dreamerv3 PyTorch modules on ray/rllib package

Machine Learning Engineer Intern — Los Angeles, CA
Elementary Robotics — 1/2023-5/2023
- Built end-to-end pipeline for anomaly detection with >60% sample efficiency of automatic defect product screening for shampoos & cosmetic creams by ResNet, vision transformers, segmentation & self-distillation
- Used AWS, Amazon Sagemaker, EC2, S3 & mlops-prod-cld for large-scale data transfer, preprocessing & labelling
- Investigated and implemented SOTA deep learning models and embeddings such as DINO & SAM

Amazon Project: Confidence-Aware Reinforcement Learning for Human-in-the-Loop Decision Making — New York, NY
Columbia ROAM Lab — 4/2022-3/2023
- Designed novel deep Q learning algorithm to optimize the number of expert calls in robot motion

Machine Learning Engineer Intern — Burlingame, CA
Phantom AI — 6/2022-8/2022
- First person to implement Hierarchical Object Detection with DRL in pytorch with distributed parallel from scratch to enhance the tightness of bounding boxes
- Implemented Faster R-CNN for vehicle detections with exhaustive augmentations by opencv & albumentations

Data Science Team Associate — New York, NY
Sportsbiz — 1/2022-5/2022
- NLP project: Language data cleaning, sentimental analysis of tweets, facebook & Instagram chats & posts with >70% accuracy by pytorch, transformers & nltk
- Computer vision project: Multiple objects detection for sports logos and equipment via tensorflow & scipy

Software Engineer Intern — Dallas, TX
5Miles LLC — 5/2020-8/2020
- First to create rust apps via web assembly & node.js & docker for ML classification & visualizations in a faster and more secure manner
- Modified rust libraries for AI/ML for more readable and user-friendly web assembly applications

Teaching Assistant (300 Students) in Deep Reinforcement Learning & Control PhD — Pittsburgh, PA &
Carnegie Mellon University & Columbia University — New York, NY
- Coded & trained RL algorithms in pytorch & tensorflow(Imitation Learning, Actor Critic, Deep Q network, DDPG, PPO, MBPO etc.) — 9/2020 – 1/2022
- Led lectures/reciations & office hours, supervise RL projects

SKILLS

Programming	Python, C++, C, Linux, R, SQL, Rust, SML, Matlab, Javascript, Git, Node.js
Packages	Pytorch, Lightning, Tensorflow, Numpy, Scipy, Opencv, Pandas, Dask, Sklearn, Matplotlib, Boto3, Gym, Json
Toolkits	Cuda, AWS, GCP, Anaconda, Enroot, Ubuntu

Other	Leadership, Communication, Latex

PUBLICATIONS

- **Machine Learning (Columbia University)** Krzysztof Choromanski, Arijit Sehanobish, Han Lin, Yunfan Zhao, Eli Berger, Tetiana Parshakova, **Alvin Pan**, David Watkins, Tianyi Zhang, Valerii Likhosherstov, Somnath Basu Roy Chowdhury, Avinava Dubey, Deepali Jain, Tamas Sarlos, Snigdha Chaturvedi, Adrian Weller
  "Efficient Graph Field Integrators Meet Point Clouds": https://proceedings.mlr.press/v202/choromanski23b/choromanski23b.pdf, ICML(ACCEPTED)
- **Robotics (Columbia University)** Siddharth Singi, Zhanpeng He, **Alvin Pan**, Sandip Patel, Gunnar A Sigurdsson, Robinson Piramuthu, Shuran Song, Matei Ciocarlie
  "Decision Making for Human-in-the-loop Robotic Agents via Uncertainty-Aware Reinforcement Learning": https://arxiv.org/abs/2303.06710, ICRA(ACCEPTED)
- **Machine Learning (Columbia University)** Yunfan Zhao*, **Alvin Qingkai Pan***, Krzysztof Choromanski*, Deepali Jain, Vikas Sindhwai
  "Implicit Two-Tower Policies": https://arxiv.org/abs/2208.01191, ICLR 2024 (ACCEPTED workshop)
- **Machine Learning (University of Pittsburgh)** Jason Xiaotain Dou, **Alvin Qingkai Pan**, Runxue Bao, Harry Haiyi Mao, Lei Luo
  "Sampling Through the Lens of Sequential Decision Making": https://arxiv.org/abs/2208.08056, ICML(submitted)
- **Health Data Analysis (Carnegie Mellon University & University of Pittsburgh)**
  "Variations in Non-Pharmaceutical Interventions by State Correlate with COVID-19 Disease Outcomes": https://www.medrxiv.org/content/10.1101/2021.07.28.21261286v1

TEACHING

Columbia University Fall 2021
Reinforcement Learning(ELEN E6885), Course Assistant, Graduate

Carnegie Mellon University Fall 2020, Spring 2021
Deep Reinforcement Learning & Control(10-703), Teaching Assistant, Graduate

VOLUNTEER

ADRA Australia, Nepal 12/2015 – 1/2016
A member of a kindergarten reconstruction project in Nepal after earthquake

ADRA Australia, Cambodia 12/2016 – 1/2017
Helped building a new local church in Cambodia

REFERENCES

Eric K. Oermann AI scientist/Neurosurgeon, NYU Center for Data Science/Langone Hospital	Eric.Oermann@nyulangone.org
Krzysztof Choromanski Staff Research Scientist/Adjunct Professor, Google Brain/Columbia University	choromanski1@gmail.com
Krishna Gopalakrishnan Vice President of Machine Learning, Elementary Robotics	krishna@elementaryrobotics.com
Dat Do Director of Machine Learning, Elementary Robotics	datdo@rocketmail.com
Katerina Fragkiadaki Assistant Professor, Carnegie Mellon University MLD	katef@cs.cmu.edu
Peter Freeman Co-director of Undergraduate Statistics & Data Science, Carnegie Mellon University	pfreeman@cmu.edu