Vineet Vinayak Pasupulety

vvinayak@gatech.edu | vineetvinayak | VVinayak | 971-801-0410

EXPERIENCE

QUALCOMM | Senior Machine Learning Engineer | August 2021 - Present | San Diego, CA

- Work in Corporate Engineering's Advanced ML Solutions team
- **Thermite**: Enhancing battery and device material lifetime by 1.5 years for 7 popular mobile models
  - Prototyped space-efficient Reinforcement Learning solution to control CPU/GPU frequencies based on sequential change in device surface temperature (performance-temperature tradeoff)
  - Modelled simulator to mimic real device physics using just 8 performance features and 10k temperature samples trained with a neural network
  - Devised a novel policy training and adaptive reward mechanism that allowed faster training convergence with minimal future training
  - Currently productionizing for Samsung Galaxy devices as on-device solution, aiming to increase device lifetime by 1.5 years and reducing material wastage/device disposal by 24%
- **Xerocole**: Predicting device surface temperature on-edge using minimal internal heat sensors
  - Processed 36 sensor datasets from 7 devices with 1000s of performance & temperature features
  - Analyzed benchmark tests on devices to measure external device metrics produced over 50 hours
  - Prototyped several prediction models to predict external device temperature (upto 84% accuracy) using the least number of internal heat sensors (determined by pareto-optimality)
  - Allowed easy tuning of sensors programmatically using ML regression without need of manual engineer measurements and updates
  - Worked with external stakeholders to produce on-device model in 6 months, thus reducing $100M annually in sensor expenses and cutting engineer effort from 16 hours to 30 minutes/week
  - Guided interns in improving speed of feature processing from 2 hours to 12 minutes with Recursive Feature Elimination enhanced GridSearch

AMAZON | Research Scientist | August 2019 - June 2021 | Seattle, WA

- Worked in the Communications AI team under Alexa AI on multiple projects
- **FIRM**: Protecting 100 million Alexa users from 175k monthly unintended/false invocations
  - Analyzed 250k customer utterances in **Amazon SQLite** and root caused 6 data quality issues affecting 1st generation supervised models, thus improving annotator efforts and reducing annotation time from 10 hours/1000 utterances to 6 hours
  - Wrote **Bash** workstreams to get high quality offline training/validation/testing data using **AWS Kinesis, S3 and CloudWatch** within 2 hours (previously 3 days)
  - Performed extensive ASR and NLU feature search in Alexa's data stores that resulted in adding 16 new features and discarding 7 older ones; produced documentation on future search efforts, engineering owners etc. to streamline engineering effort to make such data available at runtime
  - Trained and analyzed 2 **Anomaly Detection** models in **Scikit-Learn** over **SageMaker CLI** and designed 27 experiments over 2 years that lifted model precision from 35% to 76%
  - Produced robust production code for data parsing, featurization, training and inference with extensive unit testing for fast and reproducible packaging of training artifacts into production pipelines/infrastructure setup without scientist intervention
- **Affinity**: Making large scale recommendations on Alexa app/devices for increased skill discovery
  - Produced 5 point-of-view data analyses from customer usage, profile & clickstream data
  - Designed 3 novel True Success Rate metrics to define model success for adoption, discovery and engagement; helped partner teams design A/B tests on custom audiences and production logic
  - Built proof-of-concept **Collaborative Filtering** model in **PySpark** on **EMR** with documentation
  - Built production model that increased card recommendations (over rule based systems) by 12.75M, resulting in 1.8M customers trying atleast 1 new Comms skill monthly
  - Onboarded new scientists to workflows and documented **Contextual Multi-Armed Bandits**

strategy for future models across international locales

- **WAY**: Forecasting the best devices to reach customers amongst several devices
  - Performed data analysis to root cause customer frustrations due to miscommunications to wrong devices in 750k monthly utterances
  - Prototyped a time series forecasting model based on **GRU-ODE-Bayes** in **PyTorch** that reduced reported frustrations by 80%
- Led team initiatives to increase team morale and efficiency
  - Mentored a team of 4 engineers in ML science by providing support for onboarding to existing ML team projects, literature reviews, 1-1 guidance, MLScience/Ops talks, presentations etc.
  - Built templates to make experiment documentation quick to write+relevant for future science, thus increasing experiments by 24%, and reduced documentation time from 2 days to 3 hours.
  - Built ticket mechanisms to help scientists focus on ambiguous science work, thus increasing commit-to-deliver ratio from 33% to 54% in every sprint

AMAZON | Research Intern | May - August 2018 | Boston, MA

- Worked in the Conversational AI team in Alexa AI Research under **Dr. Alborz Geramifard**
- Designed deep NLP models to capture sentiment in unique dialogue turns in a context-sensitive setting; provided this as a reward signal to improve Alexa's turn-level dialogue policy
- Curated datasets of 500 dialogues and 7000 user utterances to capture live utterance characteristics
- Developed with **PyTorch** as a deep learning classifier for an Alexa skill called **MovieBot**.

ADOBE | Research Intern | May - August 2017 | Bangalore, India

- Designed deep learning models to capture complex edits Photoshopped onto natural images;
- Wrote web crawlers to create curated datasets of 9000 images, 1300 Photoshop tools & 178 tutorials
- Performed Dependency Parsing/Vector Space Analysis/Topic Modelling to map Photoshop tool-effect-region tuples into meaningful tutorial steps for novices
- Developed with **TensorFlow/NLTK/OpenCV and D3** for plugin for Photoshop CS6. Patent Submitted.
- Awarded **"Most Creative Research Project 2017"** amongst 23 teams at Adobe Big Data Labs

RESEARCH

DOMAIN RANDOMIZATION FOR EMBODIEDQA

- Working with **Dr. Devi Parikh**, **Dr. Stefan Lee** and **Dr. Peter Anderson** on using Domain Randomization to improve reinforcement learning agents abilities on **EmbodiedQA**
- Built using **PyTorch**. Engineered training framework to permit dynamic domain randomization

INTRINSICALLY MOTIVATED CONTEXTUAL BANDITS

- Worked with **Dr.Swati Gupta** and developed in **Tensorflow**
- Posed exploration in reinforcement learning as a contextual bandits Q-function ensemble; each bandit used intrinsic motivation functions (curiosity) to discover improved trajectories
- Increased the levels reached (over basic UCB) of a Mario agent by 24% on OpenAI Gym

CAPTIONING-VQA | Code | Video

- Worked with **Dr.Devi Parikh** and developed in **PyTorch**
- Created word, image embeddings from pre-trained Image Captioning models to act as context attention for a deep neural Visual Question Answering encoder-decoder framework

INTERPOLO | Code | Poster

- Worked with **Dr.Polo Chau** and developed in **Flask, AJAX, D3.JS and HTML/CSS**
- Interactive Visualisations and Analytics on-demand of worldwide Terrorism between 1970-2016

EDUCATION

GEORGIA INSTITUTE OF TECHNOLOGY | MS in Computational Science & Engineering
Specialization: Computer Vision, Deep Reinforcement Learning, Natural Language Processing
August 2017 - May 2019 | College of Computing, Atlanta, GA