# Prachi Rahurkar

San Jose, CA

rahurkap@gmail.com **|** 669-204-1088 **|** linkedin.com/in/prachi-rahurkar/ **|** https://prachirahurkar.github.io/

## WORK EXPERIENCE:

**Amazon Music**                                                                                                                                                San Francisco, CA
**Software Engineer, Machine Learning**                                                                                                         Aug 2021 - present

- Designed and implemented customer embeddings training and validation jobs utilizing AMU customer track playback history data to generate personalized track recommendations (with a runtime of only 45-47 mins) along with its AWS clusters and CDK.
- Extended the above pipeline of same-entity (track-to-track similarity) use case to cross-entity (track-to-artist similarity) use cases, which was adopted by 4 client services enhancing multiple customer listening experiences.
- Proposed and implemented the backend selection and sequencing logic to play the first 5 tracks of playback for Free-tier Amazon Music customers requesting a track, as Free-tier customers cannot play songs on-demand. This feature significantly boosted customer engagement and retention (via Voice) on Amazon Music Free by 6%.
- Proposed and implemented the model to showcase personalized playlist MyDiscoveryMix which contains 15 never-heard-before tracks based on customer's listening history, increasing the Unlimited-plan sign-ups by 3.8% in NA and EU regions.
- Crafted and implemented a thorough metrics and alarms package tailored for the recommender service, along with the development of monitoring dashboards. This initiative resulted in an impressive 80% reduction in response time to critical issues caused by the model, and a notable 50% improvement in team efficiency and usability.

**Memorial Sloan Kettering Cancer Center**                                                                                                                  New York, NY
**Machine Learning Engineer**                                                                                                                            May 2020 - Feb 2021

- Designed and implemented the complete natural language processing pipeline from scratch that predicts whether a metastatic (mets) disease of cancer is present in a given radiology report, using open-source language models BERT and Longformer.
- Orchestrated the pipeline to address a spectrum of 13 metastasis tasks (such as kidney mets, liver mets, lung mets, spleen mets, among others), resulting in an accuracy >96%. This reduced the lookup time per report from 40 mins to just 3 mins.
- Designed and built a radiology-dedicated new language model RadioBERT by unsupervised pre-training of NLP model RoBERTa on radiology text data extracted from medical CT scan reports, and then fine-tuned the language model for multiple downstream tasks. Also developed and established guidelines for effective model verification within the institution.

**Oregon State University**                                                                                                                                      Corvallis, OR
**Research Assistant**                                                                                                                                       Aug 2020 - Mar 2021

- Conducted research to generate syntactically and semantically similar passages in the Question-Answering (QA) task.
- Implemented linguistic categorization for perturbed passages that are successful in misleading language model's answer predictions, focusing on adversarial attacks in the reading-comprehension QA task.
- Automated the generation of these adversarial text passages and conducted research to enhance the robustness of reading-comprehension question-answering models against these challenges. (Published at NeurIPS 2020)

## EDUCATION:

**Master of Science** - Computer Science, Artificial Intelligence: GPA 3.7/4                                                             2018 - 2021
Oregon State University                                                                                                                                Corvallis, OR, USA

**Bachelor of Engineering** - Computer Science: GPA 8.26/10                                                                                 2014 - 2018
Thadomal Shahani Engineering College - University of Mumbai                                                                              Mumbai, India

## PUBLICATIONS:

- **Prachi Rahurkar**, Matt Olson and Prasad Tadepalli. "Human Adversarial QA: Did the Model Understand the Paragraph?"**\*** **NeurIPS** 2020 - Workshop on Human and Model in the Loop Evaluation and Training Strategies (HAMLETS).          **\*(M.S. Thesis)**
- Richard Do, **Prachi Rahurkar\***, Lior Gazit\*, et al. "Patterns of metastatic disease in patients with cancer derived from natural language processing of structured CT radiology reports over a 10-year period" Journal of **Radiology** 301 (1), 115-122                                                                          **\***(Responsible for NLP and Deep Learning work)

## AWARDS:

- 1st Prize at the State-level Project Competition DJ ASCII (amongst over 909 participants)                                                 2018
- 2nd Prize at the Project Expo of Thadomal Shahani Engineering College (amongst over 300 participants)                    2018

## PROJECTS:

- YouTube Assistant: Web-app that answers questions about any YouTube video, using OpenAI's LLM agent. (LangChain, Streamlit)
- Pets Name Generator: Web-app that suggests creative names for your pet based on the inputs, animal type and its skin color.

## TECHNICAL SKILLS:

- **Programming languages:** Python, JavaScript, Java, TypeScript, Scala, Go, C, C++, Bash, Ruby, C#, HTML, CSS, Rust, PHP
- **Machine Learning and AI:** PyTorch, Tensorflow, LLMs, NLP, Computer Vision, Deep Learning, Spark, Git, Docker, Kubernetes
- **AWS Tools:** EMR, Infrastructure CDK, EC2, Lambda, S3, Redshift, CloudWatch, Kinesis, DynamoDB, SQS and CloudFormation
- **Frameworks:** React, Redux, Node.js, Express, UNIX, Flask, GraphQL, Apollo, Angular, Spring, Vue, jQuery, JUnit, RabbitMQ, Jira, Jenkins, Django, Fast API, Apache Kafka, Espresso. **Database Technologies:** MongoDB, MySQL, Redis, PostgreSQL, Firebase