# Prachi Rahurkar

rahurkap@oregonstate.edu
Web: prachirahurkar.github.io
LinkedIn: prachi-rahurkar
Mobile: (+1) 669-204-1088

## Education

**Oregon State University** — Corvallis, OR
Master of Science - Computer Science, GPA: 3.71/4 *(including Fall '20)* — 2018–2020

*Coursework: Natural Language Processing, Networks in Computational Biology, Intelligent Agents and Decision-Making, Computer Vision- I and II, Deep Learning, Machine Learning*

**University of Mumbai** — Mumbai, India
Bachelor of Engineering - Computer Science, GPA: 8.26/10 — 2014–2018

*Coursework: Analysis of Algorithms, Data Structures, Artificial Intelligence, Machine Learning, Soft Computing, Data Mining, Databases, Parallel and Distributed Systems, Theoretical Computer Science*

## Research Experience

- **Question Answering** — Sep 2019 – present
  *MS Thesis (Advisor: Dr. Prasad Tadepalli)* — *Oregon State University*
  - I am working on automating the generation of text adversarial passages and building reading-comprehension question-answering models robust to the adversarial passages.
  - Performed categorization of the successful adversarial text attacks that misdirect model predictions, based on the data collected in a user study that I conducted. (Published at NeurIPS workshop)
- **Brain Network Analysis** — Sep 2020 – Dec 2020
  *Research Project (Advisor: Dr. Stephen Ramsey)* — *Oregon State University*
  - Performed structural analysis on human connectome and chimpanzee connectome data.
  - Implemented graphical measures such as centralities, motifs, clustering coefficients, max-flow, analysis of unique edges, etc. among others and showcased differences between the two brain networks.
- **NaYaNa for Universal Literacy** — Jul 2017 – Apr 2018
  *Research Intern (Supervisor: Dr. Nagarjuna G.)* — *Homi Bhabha Center, TIFR*
  - Implemented training data generation algorithm given NaYaNa script (of alphabets and numbers)
  - Developed a handwritten character recognition system for the script NaYaNa using feature extraction techniques based on character geometry, with the help of stacked neural networks in implementation.
- **NLP for Regional Languages** — May 2018 – Aug 2018
  *Independent Study (Advisor: Dr. Jayant Gadge)* — *TSEC, Mumbai*
  - Implemented English-to-Marathi translation and worked on understanding the effects of back-translations.
  - Implemented poetry generation system using SeqGAN and worked on enhancing its architecture.

## Professional Experience

- **Machine Learning Engineer (Part-time)** — Aug 2020 - Dec 2020
  *Memorial Sloan Kettering Cancer Center* — *New York City, NY*
  - RadiologyBERT: Built a radiology language model by unsupervised pretraining on radiology text data extracted from medical CT scan reports and fine-tuned the LMs for downstream tasks.

- **NLP Data Scientist (Intern)** May 2020 - Aug 2020
  *Memorial Sloan Kettering Cancer Center* *New York City, NY*
  - Developed complete NLP pipeline from scratch using public Longformer and BERT for classification on medical radiology reports, and scaled it for all metastatic disease tasks, with little training data.
- **Algorithm Engineer - Computer Vision (Intern)** Jun 2019 - Sep 2019
  *KLA–Tencor Corp.* *Milpitas, CA*
  - Proposed and implemented novel image quality estimation techniques for wafer images.

## Publications

- **Prachi Rahurkar**, Matt Olson, 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
- **Prachi Rahurkar**, Lior Gazit, Huy Nguyen
  *RadioBERT: Detecting Cancer Types in Radiology Reports*
  (Under Submission at NAACL 2021)
- Richard K. Do, Kaelan Lupton, Pamela Causa, Anisha Luthra, Michio Taya, Karen Batch, Huy Nguyen, **Prachi Rahurkar***, Lior Gazit*, Kevin Nicholas, Christopher J. Fong, Natalie Gangai, Nikolaus Shultz, Farhana Zulkernine, Varadan Sevilimedu, Krishna Juluru, Amber Simpson, Hedvig Hricak
  *Patterns of Metastatic Disease in Cancer Patients Derived from Natural Language Processing of Structured CT Radiology Reports*
  (Under Submission) (* responsible for NLP and deep learning work)

## Awards and Honors

- 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
- Bronze Medal at State-level Fencing Competition held in Nagpur, MH, India 2012

## Teaching

- **Teaching Assistant** at Oregon State University Fall 2020
  *Software Engineering (CS 361)*
- **Graduate Teaching Assistant** at Oregon State University Winter 2019 - Spring 2020
  *Senior Capstone Project (CS 467)*

## Projects

- **Generate Melody for your Poetry** Winter 2019
  *Developed a neural model using stacked LSTM networks which takes poetry text as input and generates a suitable melody for it as the output. Implemented in Python (PyTorch)*
- **Translate Katakana to English** Fall 2019
  *Performed translation of Japanese Katakana to English with the help of phonemes in WFSAs and WFSTs and an extension of the Viterbi algorithm. Implemented in Python.*
- **Pattern Studio** Spring 2019
  *Built a system that generates floral and geometric design patterns using self-attention GANs, and applies the generated pattern on the garment selected in the image of a user. Implemented in Python (PyTorch)*

## Skills

- **Languages:** Python, Java, R, Matlab
- **Libraries:** PyTorch, TensorFlow, Transformers
- **Databases:** SQL, MongoDB
- **Misc.:** Bash, Git, Flask, Node.js

## Languages

- **English:** Native/Bilingual
- **Marathi:** Native/Bilingual
- **Hindi:** Fluent/Proficient
- **Sanskrit:** Limited
- **Spanish:** Beginner