Howard Yen

hyen@princeton.edu | 347-446-2283 | github.com/howard-yen | howard-yen.github.io | linkedin.com/in/howard-yen/

Education

Princeton University, Princeton, NJ May 2023
GPA: 3.985 out of 4.0, Departmental GPA: 4.0 out of 4.0
Intended in B.S.E. Computer Science, certificates in Robotics and Intelligent System and Statistics and Machine Learning
Coursework: Natural Language Processing, Advanced Computer Vision, Embodied Language Understanding, Operating Systems
Awards: Shapiro Prize for Academic Excellence 2021, ICPC 2021 North America Finalist, 2nd Place at Citadel Terminal Live 2020

Skills

Languages: Native in Chinese (Mandarin) and Fluent in English
Technical Skills: Advanced: Python, Java, C, C++, React JS, PostgreSQL; Proficient: OCaml, R, Figma

Work Experience

Software Engineering Intern, **Meta Reality Labs**, Seattle/Redmond, Washington May 2022 - August 2022
- Implemented data augmentation interface in the internal Automatic Speech Recognition (ASR) & Natural Language Understanding (NLU) library using PyTorch for augmentation in both the audio and the text dimension.
- Experimented with different augmentation techniques like inserting pauses, duplicating words, and replacing phrases with other phrases of the same slot type on the STOP dataset. Achieve an 1% and 0.25% improvements in exact match on synthetic and natural speech training data, respectively.
- Investigated the effectiveness of data augmentation in the audio vs. text dimension and on low-resource training sets.

Teaching Assistant, **Princeton University Computer Science Department**, Princeton, New Jersey February 2022 - May 2022
- Graded for the class COS 484: Natural Language Processing, which had approximately 100 students, provided feedback to students' theory and coding assignments on correctness and efficiency on topics like LSTM networks and Markov Chain.
- Answered students' questions on class forum about general NLP subjects like generative vs. discriminative networks.

Software Engineering Intern, **Facebook AI Applied Research**, Menlo Park, California (Remote) May 2021 - August 2021
- Interned on the Privacy ML team at the Responsible AI org. Researched and implemented sketching based gradient approximation algorithm called FetchSGD in PyTorch to optimize client-server communication in Federated Learning.
- Achieved more than 80% gradient compression rate with less than 5% drop in performance on computer vision and natural language datasets like CIFAR10, CelebA, and Sent140 that are competitive with state-of-the-art FL training algorithms
- Worked closely with other engineers to refactored the code base and design new interfaces for the FL library to be more generic to accommodate new algorithms like FetchSGD

Projects

Why is the Sky Blue? A Few Words Are Not Enough to Answer Open-Domain Questions Sept 2021 - June 2022
- Advised by Professor Danqi Chen and Jinhyuk Lee. Submitted to EMNLP 2022 (Short, under review)
- Introduced a new task **M**ulti-type **O**pen-domain **Q**uestion **A**nswering (MOQA) that requires building open-domain question answering (ODQA) systems that can provide short, medium, long, and yes/no answers.
- Established strong baselines for MOQA using state of the art ODQA systems DPR and DensePhrases and modified the two existing systems to enable them to output different forms of answers, and found that current systems fail at MOQA.

TigerResearch February 2021 - May 2021
- Designed an interface to browse research mentors and mentees and match them together for Princeton's Office of Undergraduate Research in Figma by collaborating with administrators and teammates and interviewing users
- Developed frontend in React JS and backend in Flask to handle real-time searching and automatic matching features

Leadership and Extracurriculars

Vice Chair, **Association for Computing Machinery**, Princeton University September 2019 - Present
- Host the annual Princeton Computer Science Contest by working out logistics, contacting sponsors, and writing problems
- Practice problem solving skills and algorithm and data structure optimization through competitions like ICPC

Design and Development Team Member, **RESearch INnovation DEsign**, Princeton University September 2020 - Present
- Conduct user interviews on the course selection process for college students to find ways to improve students' course selection experiences and develop a web app that integrates calendar and course reviews in Python and React

Princeton NLP Group, Badminton Club, Swimming, Fencing Club, Club Climbing