Kecen Yao

University of Toronto
+86 13660739596 kecen.yao@mail.utoronto.ca linkedin.com github.com

Education

University of Toronto **September 2024 – June 2026**
Master of Science in Applied Computing *Toronto, Ontario, Canada*

New York University **September 2020 – June 2024**
Bachelor of Science in Computer Science (Honors) & Data Science (Concentration in AI) *Shanghai, China & NY, NY, United States*

- Cumulative GPA: 3.89/4.00, Major GPA: 3.99/4.00
- Selected Awards:
  - Dean's List of NYU 2021-2023
  - Latin Honors: Magna cum laude
  - NYU Shanghai Excellence Award

Relevant Coursework

- Algorithms (A)
- Special Topic: NLP (A)
- Computer Systems Org (A)
- Recommendation System (A)
- Computer Network (A)
- Applied Internet Tech (A)
- Operating Systems (A-)
- Probability and Statistics (A)
- Machine Learning (A)
- Reinforcement Learning (A)
- Database Design & Implementation (A)
- Linear and Nonlinear Optimization (A)

Technical Skills

Programming Languages: Python (proficient), C/C++, C#, JavaScript, HTML, CSS, SQL, MATLAB
Developer Tools: Linux, Git, Docker, MongoDB, Unity
Frameworks: PyTorch, TensorFlow, OpenCV, ROS, Sklearn

Experience

Machine Learning Intern @ Deepmirror **May 2024 – Aug 2024**
Supervisor: Zebing Cai *Guangzhou, China*

- Benchmarked and optimized relocalization algorithms for in-car VR applications, focusing on pose error metrics and reducing inference time by integrating transformer-based approaches.
- Conducted detailed comparisons of image retrieval and matching frameworks, including CricaVPR, Cosplace, GIM, LightGlue, SuperGlue, and IMP, across diverse datasets and tasks.
- On this framework, developed and trained a seatback classification detector using NANOdet, handling data collection, cleaning, and augmentation for deployment on edge devices in various lighting conditions.
- Worked on and helped build a unified training framework for machine learning, Uranus, utilizing Docker environments and Determined AI clusters for model training and deployment.

Machine Learning Intern @ UniDT Technology (Shanghai) Co., Ltd **August 2023 – December 2023**
Supervisor: Dr. Xiaohua Xuan & Dr. Kehuan Shi, Algorithms Department *Shanghai, China*

- Participated in building a Large Language Model (LLM) called "Wisdom-Interrogatory" specifically designed for legal queries. Performed massive legal data regularization, cleaning, and analysis for training and fine-tuned the LLM.
- Improved the LLM's website by developing the web development framework. Used Vue.js for improved user interaction on the front end and established a comprehensive database and API management on the back end.

Research Assistant & Tandon UGSRP member - AI4CE Lab @ New York University **December 2022 – Present**
Supervisor: Professor Chen Feng, assistant professor of Civil and Urban Engineering at NYU Tandon *New York, NY, United States*

- Built a webcam data downloading pipeline and down-sampled video loading pipeline independently, contributing to a new street webcam streaming video database targeting crowd detection. Benchmarked state-of-the-art object detection algorithms using the database.
- Implemented conformal inference to generate uncertainty aligned with densities map in object counting task. Developed a method to integrate pixel-wise uncertainties into the object counting framework to refine the counting model. Proved its statistical significance and utility in enhancing counting accuracy.

Research Assistant - NYU Multimedia and Visual Computing Lab @ New York University **July 2023 – September 2023**
Supervisor: Professor Yi Fang, associate professor of Electrical and Computer Engineering at NYU Abu Dhabi and NYU Tandon *New York, NY, United States*

- Integrated pre-trained language model like GPT-4, visual language model like CLIP, and task-driven model such as GR-ConvNet, to develop a zero-shot learning system that can generate gestures for robot grasping and placing tasks using RGBD images.
- Developed a robust pipeline that enabled autonomous robotic self-checking through the implementation of Visual Question Answering (VQA), empowering the model to assess and validate its actions.
- Explored robotic simulation using ROS (Robot Operating System) and Gazebo, successfully running robotic arm simulation demos.

Selected Projects

Advanced Deep Reinforcement Learning Algorithm Analysis in BipedalWalker Environment | *Python, Pytorch, RL* **September 2023 - January 2024**

- Conducted an extensive comparative study of prominent Deep Reinforcement Learning (DRL) algorithms, including Proximal Policy Optimization (PPO), Deep Deterministic Policy Gradients (DDPG), Twin Delayed DDPG (TD3), and Soft Actor-Critic (SAC), within the OpenAI Gym's BipedalWalker environment.
- Enhanced and applied new performance metrics for more accurate evaluation of algorithm efficacy, demonstrating distinct advantages and limitations of each algorithm through both quantitative and qualitative analysis.

Neutral Summarization for Framing Biased Articles | *Python, Pytorch, NLP* **January 2023 - May 2023**

- Developed an innovative approach to automatic neutral summarization using the T5 encoder and fine-tuned T5 decoder, proposing a solution for mitigating media framing bias caused by selective journalistic writing, aiming to foster social solidarity and reduce division among readers.
- Demonstrated the model excellence in Avg. Framing Bias metric (arousal) and Salient Info metrics (BLEU and ROUGE1-R) compared to the baseline model (Bert), underscoring the model's ability to produce highly balanced and informative content.
- Proposed an innovative single-input-document neutral summarization architecture, which is more applicable to the real world sector.

NYC Taxi Prediction - Machine Learning Projects | *Python, Pytorch, Sklean* **January 2022 - May 2022**

- Applied commonly-used machine learning model to forecast future taxi cab demand in New York City.
- Conducted data cleaning, data preprocessing, and applied standard PyTorch training frameworks.
- Implemented supervised-learning algorithms in Python, such as Linear Regression, Tree-based models, KNN (K-Nearest Neighbors), and Recursive Neural Network (RNN). The best-performing model RNN attained an R2 score exceeding 0.99, demonstrating exceptional predictive accuracy.

New York University Shanghai Hackathon | *Python, Pytorch* **January 2022**

- Developed a hybrid machine learning model by combining multiple algorithms for data training. Focused on diagnosing specific diseases using health indicators.
- Achieved the highest level of accuracy compared to all competing attendants, showcasing exceptional predictive performance in disease diagnosis.