Winson Chen

(949)-232-4947 | winsonchen108@gmail.com | linkedin.com/in/Winson-Chen
github.com/W1nson | scholar.google.com/Winson

Education

University of California, Santa Cruz — Santa Cruz, CA
Scientific Computing and Applied Mathematics, Master of Science — *June 2023*
- GPA: 3.94

University of California, Santa Cruz — Santa Cruz, CA
Computer Science, Bachelor of Science, Cum laude — *June 2022*
- GPA: 3.86

Irvine Valley College — Irvine, CA
Computer Science, Associate in Science, Cum laude — *Aug. 2020*
- GPA: 3.58
- Honors Program

Experience

Business Analyst / Full-Stack Engineer — Mar. 2024 – Present
InfoIMAGE, Inc — *Brisbane, CA*
- Led a team of four in the design and development of an internal website, coordinating tasks, managing timelines, and ensuring successful project completion.
- Developed a new internal Core platform using Next.js and Flask, resulting in a 80% increase in data sample lookup
- Designed Core ORM by using SQLAlchemy, to increase development time by 20% and improve security
- Utilized RESTful API to facilitate front-end to back-end communication, improving data consistency by 50%
- Integrated AG-Grid to optimize the user experience resulting 15% increase in filtering and searching process

Machine Learning Engineer — Aug. 2022 – Sept. 2023
University of California, Santa Cruz / Amazon Alexa Prize — *Santa Cruz, CA*
- Reduced training time by 23% using LoRA and lightning package to fine-tune RedPajama 3B
- Enhanced system speed by 34% via AWS EC2 auto-scaling deployment of RESTful server with Red Pajamas 3B
- Increased 30% user ratings by fine-tuning RedPajamas 3B with dialog history for personalized questions
- Over three months, we achieved a 10% increase in user evaluations by implementing an APL (Alexa Presentation Language) detail template with text that auto-scrolls in sync with the voice-over
- Through the design and A/B testing of multiple variations of dialog managers and APL templates, we successfully attained a 13.5% improvement in satisfactory evaluations, striving to unveil the ultimate user experience

Data Scientist Intern — Jun. 2022 – Sept. 2022
Lawrence Livermore National Laboratory — *Remote*
- Designed web scraping pipeline using Scrapy's parallel process, improved 25% data collection
- Fine-tuned CodeBERT with DataRaceBench to adapt to parallel syntax in C/C++
- Optimized training speed by freezing hidden layers, reducing CodeBERT training time by 10%
- Set up a data pipeline to visualize SARS-CoV-2 Inhibitor for training a CNN in classification

ERIC Lab Assistant — Aug. 2021 – Sept. 2023
University of California, Santa Cruz — *Santa Cruz, CA*
- Received Amazon Alexa Prize Award to work on Alexa Prize SocialBot, TaskBot Challenge
- Collecting data with by simulating drone controller to create visual language navigation dataset
- Developed toolbox to help users to find published papers from Association for Computational Linguistics

Research Assistant — Aug. 2021 - Dec. 2021
University of California, Santa Cruz — *Santa Cruz, CA*
- Utilizing PySINDY package to fit the spatial-temporal data of drosophila genes using machine learning
- Experimenting various models to find the behavior of drosophila gap genes

Group Tutor — Jan. 2021 – May 2022
University of California, Santa Cruz — *Santa Cruz, CA*
- Helping students with programming, debugging, and data structure
- Providing computer science help include Java, C, C++, and Python

Code Release

Stock Assistant May 2024 – Present

- Developed a stock analyst agent with LangChain to summarize financial information for better decision-making
- Integrated Function Calling with yfinance API and Chromadb retrieve latest financial news in embedding
- Allow users to switch between OpenAI models, Ollama models, or Google Models for customization

Mahjong Calculator Jun. 2023 – Present

- Designing a MahJong Score Calcualtor to understand the rule of winning score
- Increased 20% labeling efficiency by setting up remote server with Label Studio for Mahjong tile annotation
- Fine-tuned using YoloV8 with customize MahJong dataset to detect the tile and its face value

SamBasketballTW Mar. 2023 – Jul. 2023

- Designed the **scalable** static website which presents the various basketball leagues in Taiwan
- Reduced time cost by 50% to streamline **data collection** from various websites
- Leveraged Bootstrap 5 and incorporated custom functions to elevate the user experience and **optimize performance**

ACLTool | *Python, Flask, Pandas, HTML, Javascript, BeautifulSoup* Nov. 2021 – Jan. 2022

- Users can find all the papers that published from ACL Rolling Review for easy access.
- Uses natural language processing to find recent trending topics in Machine Learning.
- **Github Repository**

Publications

Active Listening: Personalized Question Generation in Open-Domain Social Conversation with User Model Based Prompting
EMNLP'24
Kevin Bowden, Fan Yue, **Winson Chen**, Wen Cui, Davan Harrison, Eric Xin Wang, Marilyn Walker

- We hypothesize that users of conversational systems want a more personalized experience, and existing work shows that users are highly receptive to personalized questions (PQs).
- Populate prompt templates with these 400 interests and use an LLM to generate PQs customized to user interests.

Early Experience with Transformer-Based Similarity Analysis for DataRaceBench
Correctness Workshop @ SC'22
Winson Chen, Tristan Vanderbruggen, Pei-Hung Lin, Chunhua Liao, Murali Emani

- Discovered the strengths and limitations of the Transformer-based approach and pointed out future research direction
- Performed **data analysis** on new source codes with CodeBERT to understand the clone detection
- Experimented on CodeBERT by **fine-tuning** with multiple datasets to adapt different programming style

Aerial Vision-and-Dialog Navigation
ACL'23
Yue Fan, **Winson Chen**, Tongzhou Jiang, Chun Zhou, Yi Zhang, Eric Xin Wang

- Performed **data collection** which collected over 1,000 recorded navigation trajectories with asynchronous human-human dialogues with simulator we built
- Increased 20% efficiencies in **data visualization** by generated graphs and charts using wordcloud, matplotlib

Athena 3.0: Personalized multimodal chatbot with neuro-symbolic dialogue generators
Alexa Prize SocialBot Grand Challenge 5 Proceedings
Yue Fan, Kevin K Bowden, Wen Cui, **Winson Chen**, Vrindavan Harrison, Angela Ramirez, Saaket Agashe, Xinyue Gabby Liu,

- Received **$50,000 for Science and Innovation Award**
- **Multimodal** interactivity with the Alexa screen devices by implementing APL Manager to generate one of three unique user interfaces with relevant content
- **Fine-tuned RedPajama 3B** with different datasets using **LoRA** to achieve 20% speedup in training
- **Annotated** 37,235 responses to train a GPT-2 style decoder-only Transformer to increase the rating by 0.1

Talks

Early Experience with Transformer-Based Similarity Analysis for DataRaceBench
Dallas, Texas 11/18/2022
In-person talk at Correctness workshop at SC22

Teaching

Graduate Teaching Assistant Sept. 2022 - June 2023
University of California, Santa Cruz *Santa Cruz, CA*

- MATH 19A (Calculus for Science, Engineering, and Mathematics)
- AM 10 (Mathematical Methods for Engineers I)
- AM 20 (Mathematical Methods for Engineers II)
- AM 214 (Applied Dynamical System)
- AM 250 (Intro to High-Performance Computing)

Honors and Awards

- Dean's Honor: Fall 2020, Spring 2021, Fall 2021, Winter 2022, Spring 2022
- Next Gen. SAM Scholar (Funded by National Science Foundation)

Technical Skills

Languages: Python, C, C++, CUDA/HIP(C++), SQL, JavaScript, HTML, CSS, Java
Frameworks: Flask, Django, PostgreSQL, React Native, jQuery, Bootstrap, Next.js
Developer Tools: Docker, Postman, Git, Jira
Libraries: PyTorch, TensorFlow, Pandas, NumPy, Matplotlib, Seaborn, Lightning, Gradio, Keras, BeautifulSoup, Scrapy, Selenium, OpenMP, MPI, OpenCV, Langchain, Streamlit, ChromaDB
Services : RESTful API, SLURM, AWS: SageMaker, S3, EC2