Yuren Sun

yurensun@stanford.edu | (608)-338-4124 | Stanford, CA 94305 | https://github.com/YurenSUN/ | https://yurensun.github.io

EDUCATION

Stanford University September 2022 - June 2024
Master of Science in Computer Science, Information Management and Analytics Specialization, 4.04/4.00
Related Courses: Embedded Systems, Bioinformatics, Natural Language Processing, Mining Massive Datasets
Teaching: Course Assistant for Computer Graphics (Blender and Python) and Mathematical Foundations of Computing

University of Wisconsin – Madison July 2018 - December 2021
Bachelor of Science in Computer Sciences, Economics, and Mathematics with Comprehensive Honors, 3.97/4.00
Related Courses: Algorithm, Database, Linear Optimization, Operating Systems, Stochastic Processes, UX
Related Awards: Holstrom Environmental Research Fellowship (2021), CS Golden Brick Award (2019 & 2021)
Teaching: Peer Mentor for Operating Systems (C and Linux) and UX development(React and JavaScript)

TECH SKILLS: Python, C/C++, Rust, Java, SQL, JavaScript, React, HTML, CSS, Swift, Linux, Git, Tableau, Stata

WORK EXPERIENCE

Software Development Engineer Intern, Sisu Data, San Francisco, CA June 2023 - September 2023
- Automated the data staging region detection based on the customer warehouse region to reduce data staging time with Rust
- Developed internal tools to detect customer warehouse regions with Rust and back-filled the warehouse regions with SQL
- Migrated YAML configuration storage from JSON string to Kubernetes ConfigMaps for better readability and robustness
- Retrieved and logged the query execution plan during data staging to understand the causes of the slow queries with Java
- Developed Rust and Python endpoints and internal tools to retrieve the staging query execution plan for provided workflows

Software Development Engineer Intern, Amazon Web Services, East Palo Alto, CA June 2021 - September 2021
- Designed and implemented a local reproduction tool on non-data dependent issues to improve debug abilities for Redshift
- Developed catalog functions in C to extract names of tables and views from queries and trace down dependencies of views
- Developed the pipeline to retrieve the data definition languages from query texts with automatic dependency tracking
- Conducted comprehensive testing of catalog functions to ensure the feasibility of functionalities and coverage of edge cases

Innovations Intern, American Family Insurance, Madison, WI May 2020 - August 2020
- Developed reproducible webpages for prototypes and minimum viable products using HTML, JavaScript, and CSS
- Designed, implemented, and refined the user interface based on customer interviews to ensure a seamless user experience
- Set up database and deployed the serverless web applications for user data collection with Amazon Web Services (AWS)
- Managed AWS resources with Terraform to streamline workflow and automated the code delivery with CI/CD pipelines

PUBLICATION

Sun, Y., et al. Classification of animal sounds in a hyperdiverse rainforest using convolutional neural networks with data augmentation. *Ecological Indicators, vol. 145,* 2022, p. 109621., https://doi.org/10.1016/j.ecolind.2022.109621.

SELECTED PROJECTS

Baynana Resume Helper January 2023 - Present
- Design and develop the Baynana resume helper website with React to help 500+ users write and improve resumes with AI
- Collect user feedback from interviews and provide agile updates on features including interactive UI and streamed chat
- Set up backend data storage with Supabase and incorporate it into the website to store, update, and retrieve user data
- Incorporate open-source tools and develop Latex templates to generate and preview PDF and Latex versions of resumes

Animal Audio Classification and Detection January 2020 - December 2021, September 2022 - Present
- Collect and process sample data to generate the training dataset and develop a testing pipeline to measure model performance
- Use convolutional neural networks to classify the animals based on the spectrogram of audios with Python and TensorFlow
- Improve the model accuracy on small training dataset to over 90% with transfer learning and data augmentation
- Research model performances on small datasets and minimum number of samples required for desired model performance
- Develop a pipeline to streamline audio processing to detect and classify frogs in the soundscape recordings across years
- Test the detection results with stratified sampling and research the frog mating pattern versus weather with detection results