Kevin Lee

Columbus, OH | (380) 997-8856 | kevindclee@gmail.com | kevindclee.github.io

EDUCATION

Purdue University, College of Science Aug. 2020 - May 2023
Master of Science in Computer Science West Lafayette, IN
Plan of Study GPA 3.2

Ohio State University, College of Engineering Aug. 2014 - May 2020
Bachelor of Science in Computer Science and Engineering Columbus, OH
Major GPA 3.71
Green Engineering Scholars Program

PROFESSIONAL EXPERIENCE

Ailys Frontier Oct. 2024 - Present
Data Preprocessing and Labeling Specialist Seoul, South Korea

- Preprocess approximately 15,000 data entries for a large language model (LLM) training dataset designed to generate automated responses to customer complaints, ensuring high standards of accuracy and consistency.
- Engage in quality assurance and collaborative workflows to advance data readiness and streamline preprocessing for AI model development.

Purdue University Aug. 2020 - May 2023
Department of Computer Science West Lafayette, IN
Graduate Teaching Assistant | Professors: Zhiyuan Li, PhD & Jeffrey A. Turkstra, PhD

- Assisted the professor in delivering a Software Development in C course for over 600 students, developing and testing assignment codes using C, Linux Shell Script, and Git.
- Led lab sessions and instructed undergraduates on core C concepts including data structures, memory management, file I/O, and pointers.
- Conducted over 20 hours of weekly support through office hours and online Q&A platform, answering more than 1,500 questions each semester to enhance students' comprehension.
- Graded and evaluated exams, quizzes, and assignments, ensuring fairness and consistency.
- Received consistent positive feedback from the students on end-of-semester course evaluations for exceptional support, guidance, and effective communication.

JLK, Inc. June 2021 - August 2021
Computer Vision Data Annotation Specialist Seoul, South Korea

- Curated and tagged over 1,000 images for a computer vision project in autonomous vehicle technology, achieving optimal dataset quality for model training.
- Participated in cross-functional quality reviews and collaborated on data management strategies, improving data integrity and process efficiency.

Superb AI, Data Research Lab May 2020 - July 2020
Data Annotation Specialist Seoul, South Korea

- Labeled and processed over 20,000 text data entries for a natural language processing (NLP) project supporting an AI-driven chat application.
- Developed and integrated an automated preprocessing pipeline with Python scripts and Excel formulas, drastically accelerating the labeling process for frequently encountered tags.
- Contributed to QA efforts and team-based data curation, optimizing the dataset's reliability and readiness for machine learning applications.

Ohio State University, Wexner Medical Center May 2019 - May 2020
Department of Obstetrics and Gynecology Columbus, OH
Research Assistant | Advisor: Douglas Danforth, PhD

- Implemented and tested the Virtual Patient software application's speech-to-text (STT) and text-to-speech (TTS) components using Unity, C#, and Java, resulting in a 33% improvement in STT/TTS performance.
- Analyzed automated speech recognition accuracy from the data from 620 medical students who provided over 12,000 audio files containing questions for the Virtual Patient. The results were contributed to the team's research paper publication.
- Conducted extensive research on cognitive AI services (Google Cloud, Azure, IBM Watson) and analyzed results using NLP techniques to determine the best-performing cloud service.
- Collaborated with a cross-functional team of 11 experts from Computer Science, Linguistics, and Medical fields, and participated in weekly meetings, presenting data analysis results to the team and addressing feedback for continuous improvement.
- Received positive feedback from medical students for advancing their clinical skills and enhancing patient care training through the improved Virtual Patient application.

Ohio State University Jan. 2019 - May 2019
Department of Computer Science and Engineering Columbus, OH
Undergraduate Teaching Assistant | Instructor: Larry King

- Supported the professor in delivering a Software Development in Java course for over 40 students, providing assistance during office hours and labs.
- Graded projects, quizzes, and homework, ensuring fair evaluations, and provided timely and constructive feedback.
- Managed administrative tasks and facilitated course logistics, helping the professor achieve course objectives.
- Clarified complex Java concepts, contributing to improved student understanding and performance.
- Maintained high student satisfaction and balanced teaching duties with academic workload through strong organization and communication skills.

PUBLICATIONS

Artificial intelligence in virtual standardized patients: Combining natural language understanding and rule based dialogue management to improve conversational fidelity.

Maicher, K. R., Stiff, A., Scholl, M., White, M., Fosler-Lussier, E., Schuler, W., Lee, K., & Danforth, D. R. et al. (2022). Medical teacher, 1–7. Advance online publication. *https://doi.org/10.1080/0142159X.2022.2130216*

PROJECTS

Virtual Patient Project May 2019 - May 2020

https://accad.osu.edu/research-gallery/virtual-patient-project

- Participated in the Virtual Patient research project, developing an AI system for Virtual Standardized Patients (VSPs) using Unity 3D, Natural Language Processing (Speech to Text and Text to Speech), and Neural Network technologies (CNN, RNN).
- Analyzed data from over 12,000 audio files from 620 first-year medical students, contributing to a significant improvement in system accuracy from 75% to 90%.

Backdoor Attack on Language Model Code Generators Feb. 2023 - May 2023

https://github.com/kevindclee/Backdoor-Attack-on-Language-Model-Code-Generators

- Engaged in a project at Purdue aimed at investigating whether inserting trigger words into models could generate malicious code, potentially causing backdoor attacks on users' computers.
- Wrote web scraping code to gather training data for the GPT-2 code generator model and implemented trigger insertion code using Python to experiment with the insertion of trigger words, streamlining the research process.
- Collaborated with team members on various aspects of the project, including data collection, code implementation, and experimentation, ensuring comprehensive coverage of research objectives.

INVOLVEMENT

Google, Machine Learning Bootcamp Seoul, South Korea / Sep. 2023 - Dec. 2023

- Participated in Google's Machine Learning Bootcamp, where I honed skills in deep learning, computer vision, and natural language processing.
- Demonstrated proficiency with TensorFlow, achieving a top 7% finish in a Kaggle competition.
- Developed expertise in advanced deep neural network models including ResNets, MobileNet, YOLO, U-Net, and Transformer Networks, enhancing knowledge in text analytics, natural language processing, and anomaly detection techniques.
- Acquired hands-on experience with Python, PyTorch, scikit-learn, Numpy, Pandas, Tableau, and Google Cloud Platform, solidifying proficiency in frameworks for machine learning applications and data visualization.

En-core, Playdata Bootcamp Seoul, South Korea / May 2022 - Dec. 2022

- Involved in the En-core Playdata Bootcamp, acquiring a comprehensive skill set covering both front-end and back-end development, along with artificial intelligence applications.
- Proficiently developed and tested multiple web applications, including a food information website and a travel review website, using JavaScript, React, Next.js, Spring Boot, MySQL, GitHub, RestAPI, JSON, and Python.
- Applied agile methodologies to manage project workflow, including regular stand-ups, sprint planning, and progress tracking.
- Gained valuable leadership experience through leading teams and delivering project presentations, which enhanced my ability to collaborate effectively and communicate complex technical concepts.

QUALIFICATIONS & SKILLS

- **Front-end frameworks:** React, JavaScript, Material UI, Tailwind CSS, HTML, CSS, JSON, XML, RestAPI, Angular, GraphQL, Flask.
- **Back-end frameworks:** Node.js, Spring, SpringBoot.
- **AI/ML frameworks:** Python, PyTorch, TensorFlow, scikit-learn, Numpy, Pandas, Keras.
- **Programming languages:** Java, C, Python.
- **Databases:** Oracle, MySQL, Neo4j.
- **Cloud platforms:** Google Cloud Platform, Microsoft Azure, IBM Watson.
- **Languages:** English (fluent), Korean (fluent), Japanese (intermediate)

HONORS & SCHOLARSHIPS

Ohio State University Trustees Merit Scholarship 2014 - 2016
Ohio State University Green Engineering Scholars Program 2014 - 2016