DOEUN LEE

+1(817) 736-5263 ◇ lee.11501@osu.edu ◇ Columbus, OH

EDUCATION

The Ohio State University
M.S. in Computer Science and Engineering — Aug 2024 - Current
GPA: 4.00/4.00
Advised by Dr. Sachin Kumar

The University of Texas at Austin
B.S. in Computer Science — June 2020 - May 2024
GPA: 3.87/4.00
Certificate: Applied Statistical Modeling
Advised by Dr. Eunsol Choi and Dr. Amy Pavel

PUBLICATIONS

Sanchaita Hazra*, **Doeun Lee***, Bodhisattwa Prasad Majumder, Sachin Kumar, "Accepted with Minor Revisions: Value of AI-Assisted Scientific Writing". Conditionally accepted in the Annual ACM Conference on Intelligent User Interfaces (IUI 2026)

Doeun Lee, Yi Yu, Ashish Manne, Oluwatoba Moninuola, Andrew Srisuwananukorn, Frank Wen, Ping Zhang, Sachin Kumar, "When Case Gets Rare: A Retrieval Benchmark for Off-Guideline Medical Question Answering". Under review in the Fourteenth International Conference on Learning Representations (ICLR 2026)

EXPERIENCE

Graduate Research Assistant, *The Ohio State University* — October 2024 - Current

Advised by *Dr. Sachin Kumar*. Researching the capabilities of large language models in scientific writing and domain-specific question answering.

- Co-Leading a study on the ability of text generation models in scientific writing and corresponding user behavior.
- Leading the creation of a benchmark dataset in the medical domain to assess the capability of LLMs in the domain-specific long-form question answering (LFQA) task.

Undergraduate Research Assistant, *UT Austin* — June 2023 - May 2024

Advised by *Dr. Eunsol Choi* and *Dr. Amy Pavel*. Researched user behavior and perception in LLM-based and traditional search interfaces for question answering.

- Analyzed user information-seeking patterns across Google search and LLM-based interfaces.
- Led user study sessions and curated a dataset of 200 classified questions for behavioral analysis.
- Evaluated the similarity between the source and the user response using BERTScore and analyzed behavioral trends.

QA Test Engineer Intern, *SparkCognition* — May 2022 - Aug 2022

Involved in AI-driven analytics and solutions for various industries to enhance operational efficiency and data security.

- Built Dredd tests for backend API payload validation and contributed to debugging and QA workflows.
- Wrote both unit test and integration test scripts for verifying feature management APIs and fabric endpoint responses using Pytest.
- Collaborated in a scrum team of 7 as QA engineer for optimization of AI fabric hyperopt and auto-nbm service.

AWARDS AND SCHOLARSHIPS

KAAGA Next Generation Leadership Forum 2023
Won 2nd place among 100 participants and was awarded a $500 scholarship

Grace Hopper Celebration - UTCS Department Scholarship 2022
Selected based on academic excellence and leadership qualities as women in technology.

UT Austin University Honors 2021 - 2024
Academic excellence acknowledged by the College of Natural Science.

UT Austin Dean's List 2020 - 2021
Academic excellence acknowledged by the School of Undergraduate Studies.

PROJECTS

Radiology And Data Augmented Classifier for X-ray Paper

Explored the impact of anatomical segmentation masks on chest X-ray classification by fine-tuning BiomedCLIP under varied training to enhance model interpretability and diagnostic focus.

- Applied anatomical segmentation masks to chest X-rays to guide model attention toward clinically relevant regions during fine-tuning and inference.
- Designed and implemented two training strategies: (1) linear classifier fine-tuning with a frozen vision encoder, and (2) full model fine-tuning including the vision encoder.
- Demonstrated that segmentation masks serve as lightweight spatial priors, improving interpretability without requiring retraining on localized features.

Adversarial Attacks for LLM and Defense Techniques Paper

Examined the susceptibility of three Transformer-based models against adversarial examples and implemented two mitigation techniques for different real-world conditions.

- Investigated the robustness of BERT, RoBERTa, and ELECTRA to adversarial examples generated via TextFooler.
- Developed augmented training by combining original and adversarial examples, increasing adversarial test accuracy by 36%.
- Designed a weighted voting ensemble across three models to improve robustness under unknown adversarial conditions, achieving 76% accuracy without adversarial training.

Analyzing and Mitigating Dataset Artifacts Paper

Analyzed dataset artifacts in SQuAD using adversarial attacks, and mitigated them through augmentation of adversarial data and the original training data.

- Identified spurious correlations in the SQuAD dataset using adversarial attacks on ELECTRA-small.
- Enhanced the evaluation accuracy by 30.5% through combined training on adversarial and original data.
- Conducted manual evaluation of 50 sampled adversarial examples and categorized them by question types.

Pancreatic Cancer Survival Prediction Google Colab

Using dataset with patient information relevant to pancreatic cancer and its treatment, used diverse models and feature engineering after initial model run to predict survival of the patient.

- Preprocessed 3,171 patient records and applied feature engineering for survival prediction modeling.
- Compared models including Naïve Bayes, Random Forest, SVM, and Neural Networks, achieving 68% baseline accuracy.

- Improved accuracy to 72% by removing outliers using SVM then applying SMOTE for class balancing.

ACTIVITIES

UT Austin Computer Science Transfers' Society, Officer of Outreach 2023 - 2024
Helped potential transfers prepare for CS transfer admission cycle.

UT Austin Undergraduate Korean Association, Vice President of Informatics 2022 - 2023
Created and managed website for UKA and led a group of directors for yearlong publication project.

UT Austin Korean Engineering Student Association, Vice President and Project Leader 2021 - 2022
Led weekly meetings with a group to teach web application development.

SKILLS

Programming	Python, Java, C, R
Frameworks and Libraries	PyTorch, Tensorflow, AWS
Language	Korean (Native), English (Professional)