

Geonwoo Lee



Charlotte, NC, USA

EDUCATION

University of North Carolina at Charlotte

M.S. in Computer Science (Concentration: Artificial Intelligence)

Charlotte, NC, USA

Aug 2025 – May 2027

- GPA: 3.83/4.00

Eulji University

Bachelor of Medical IT and Marketing (Minor in Silver Care Start-ups)

Seongnam, South Korea

Feb 2016 – Feb 2022

- Relevant Coursework: Computer Algorithm, Medical Image Processing, Data Structure, Software Engineering, Practice of Data Analysis, Database, Mathematics.
- Awards: Department Highest Academic Scholarship (2021), Academic Excellence Scholarship (2021), Excellence in Club Activities Award (2021).

RESEARCH INTERESTS

Ontologies and Knowledge Representation, Neural Networks and Deep Learning, Retrieval-Augmented Generation (RAG), Human-Centered AI, and Intent-Aware AI Systems.

RESEARCH EXPERIENCE

Socratic Reasoning and Questioning Chatbot

Research Assistant (Advisor: Prof. Mohsen Dorodchi)

Charlotte, NC, USA

Sep 2026 – Present

- Developing instructor and student dashboards that provide access to multiple educational chatbots, including a Socratic-questioning chatbot designed to guide students through questions and support recall, critical thinking, and deeper understanding of course concepts.

Socratic Questioning with Retrieval-Augmented Generation

Independent Study (Advisor: Prof. Mohsen Dorodchi)

Charlotte, NC, USA

Dec 2025 – Present

- Developing a Socratic-questioning chatbot using retrieval-augmented generation to ground educational dialogue in course materials and instructions.
- Using Python-based frameworks and the OpenAI API to implement document retrieval and dialogue generation.
- Implemented a Socratic dialogue approach that responds to students' questions with brief hints and guiding questions instead of direct answers, encouraging students to reason toward solutions independently.

Privacy Risks in AI Systems / LLM-Based Applications

Graduate Researcher (Advisor: Prof. Liyue Fan)

Charlotte, NC, USA

Jan 2026 – Aug 2026

- Conducted research on privacy risks in LLMs, focusing on membership inference attacks and sensitive-information exposure in black-box settings.
- Reviewed literature on privacy, data leakage, retrieval-augmented generation, and security risks in AI-driven systems.
- Explored evaluation methods for sensitive-information handling during retrieval and response generation, including in-context probing and group-level information attacks.

A Study on Automatic Reporting of Vehicle Accident Sensors

Undergraduate Researcher (Advisor: Prof. Ki-Young Lee)

Seongnam, South Korea

Oct 2019 – Nov 2020

- Developed an MCU-based sensor integration system using temperature, humidity, GPS, and Wi-Fi data to support collision detection and automated reporting to the nearest fire station.
- Investigated automated emergency reporting to support faster intervention when drivers are incapacitated.

WORK EXPERIENCE

Electric Power Research Institute (EPRI)

Data Scientist Intern (Full-time)

Charlotte, NC, USA

May 2026 – Aug 2026

- Designed and built an end-to-end, domain-specific RAG pipeline, integrating document processing, retrieval, reranking, and answer generation.
- Designed an evaluation approach to assess chatbot answer correctness and retrieval quality, combining custom evaluation criteria with RAGAS, Recall@K, and faithfulness measures.
- Implemented query decomposition, cross-encoder reranking, and source-aware reranking to retrieve and prioritize supporting evidence for multi-hop question answering.
- Built a 100-question gold-standard retrieval benchmark with expected chunk IDs, including multi-hop questions; retrieved all expected evidence within the top 20 results for 90% of questions.

GEOTWO Co., Ltd.

Full-stack Developer (Full-time)

Gyeonggi, South Korea

Jul 2022 – May 2023

- Developed web applications integrating front-end and back-end systems, collaborating with cross-functional teams on project requirements.
- **QA Web Application (Sep 2022 – Apr 2023):** Developed front-end features using React, TypeScript, jQuery, and Semantic UI; supported back-end development with Java and Spring Boot; used SVN for version control.
- **Reference Web Tutorial (Jul 2022 – Aug 2022):** Built tutorials for O2Platform, O2Map, and O2asis, including JavaScript asynchronous operations and OpenLayers-based map functionality.

SELECTED PROJECTS

House Price Prediction with Regression Models

Individual Project

Charlotte, NC, USA

Sep 2025 – Dec 2025

- Built a house-price regression pipeline using 1,460 observations and 79 features, comparing seven models with shared preprocessing and an 80/20 train-validation split.
- Achieved the best validation results with XGBoost: $R^2 = 0.9217$, RMSE $\approx$ \$24.5K, and 71.6% of predictions within 10% of actual prices.

U.S. Electricity Generation Analysis for Data Center Site Selection

Visual Analytics Course Project / Tableau

Charlotte, NC, USA

Spring 2026

- Developed an interactive Tableau dashboard analyzing state-level electricity generation, seasonal trends, and fuel sources to inform data center site selection.

Student Retention and At-Risk Profile Analysis

Visual Analytics Course Project / Tableau

Charlotte, NC, USA

Spring 2026

- Built an interactive Tableau dashboard exploring student retention across GPA groups, academic majors, and school-event engagement to identify potentially at-risk student groups.

PROFESSIONAL DEVELOPMENT

Java Android Web App Developer Program, Ezen IT Academy

Trainee

Seoul, South Korea

Dec 2021 – Jun 2022

- Completed project-based training in Java, Android, and web development, using Vue.js, Spring Framework, MySQL, and AWS EC2.
- **Hey, Buddy:** Led a team-based collaboration-tool project using Spring, MyBatis, and Tomcat; deployed on AWS EC2 and implemented video conferencing with WebRTC and the Zoom API.
- **Today's House Admin BBS:** Built features using Spring, Bootstrap, and MySQL, including authentication, file uploads, validation, and map integration.

SKILLS

LLM & Retrieval: RAG, query decomposition, reranking, RAGAS, retrieval evaluation

Machine Learning: Supervised regression, tree-based ensembles, XGBoost, scikit-learn, model evaluation

Programming & Frameworks: Python, Java, JavaScript, React, Spring Boot

Data & Databases: pandas, NumPy, SQL, MySQL, PostgreSQL, Oracle

Tools & Visualization: Git/GitHub, AWS, Jupyter Notebook, Matplotlib, Seaborn, Tableau