YongGeon Lee

Austin, TX 78723, USA

yg910524@gmail.com | LinkedIn

SUMMARY

Master's student developing a research focus in materials informatics and scientific machine learning.
Interested in integrating ML with first-principles models to study structure-property relationships in materials.

EDUCATION

The University of Texas at Austin, Austin, TX — Expected May, 2026
M.S. in Information Studies (Data Science track)
GPA: 4.0/4.0

Chungnam National University, Daejeon, Korea — August, 2017
B.S. in Fine Chemical Engineering

RESEARCH EXPERIENCE

Computational Materials Science, UT Austin — September, 2025 - Present

- Built a GNN pipeline for cathode thermodynamic stability prediction (E_hull) using data from Materials Project and OQMD, with pymatgen and ASE for structure-to-graph construction.
- Applied interpretable ML to identify structural factors that govern stability and thermodynamically accessible configurations in high-Ni NCM cathodes.
- Developing a DFT+U workflow in Quantum ESPRESSO for layered oxide cathode materials to validate ML-based stability screening with first-principles energetics.

Prediction of Stroke Rigidity, UT Austin — January, 2025 - August, 2025

- Developed GNN models to analyze structured medical datasets and interpret multi-relational patterns, achieving 76% accuracy and 80% ROC-AUC.
- Built a LangChain-based pipeline that used an LLM to label clinical relevance and enforce guardrails for high-quality, low-manual GNN knowledge graph enrichment.
- Applied GNN Explainer to identify key predictive features, delivering transparent, explainable AI outputs validated by clinical domain experts.

Materials Knowledge Retrieval System, UT Austin — January, 2025 - April, 2025

- Built a multi-agent RAG system for retrieving materials-related information from structured databases and vector indexes, using PostgreSQL and FAISS.
- Developed specialized agents with guardrails to reduce hallucinations and ensure citation-backed responses.
- Achieved 90%+ retrieval accuracy and reduced latency and compute load with SQL-first, on-demand strategy.

Energy Conversion Laboratory, CNU — January, 2017 - July, 2017

- Conducted adsorption experiments on activated carbon to examine uptake behavior and surface characteristics.
- Performed Zn^{2+} adsorption experiments and analyzed uptake using UV-Vis, BET, and FTIR.

WORK EXPERIENCE

The University of Texas at Austin
Quality Assurance Analyst — June, 2025 - Present

- Analyze software quality through QA testing, utilizing Power BI to deliver data-driven insights to stakeholders.
- Prototyped an LLM-assisted QA tool for automated detection of recurring error patterns and defect causes.

Circle Platform
Data Team Lead April, 2022 - June, 2024

- Developed 100+ APIs, optimized data collection pipelines, and improved data processing speed by 50%.
- Engineered and refined keyword search platforms, improving large-scale data management efficiency by 36%.
- Applied Scikit-learn for sales forecasting, boosting the user base from 1,100 to 2,300 by enhancing satisfaction.

Umicore
R&D and Quality Engineer August, 2018 - October, 2021

- Analyzed coin-cell performance, XRD profiles, and process datasets to identify degradation patterns in high-Ni NCM cathodes and guide additive/coating trial strategies.
- Investigated moisture-driven Li_2CO_3-LiOH imbalance and supported corrective actions that improved cathode stability and product reliability.
- Produced 200+ technical reports using Six Sigma and statistical methods, collaborating with R&D teams to resolve issues and enhance process consistency.
- Built a foundational understanding of structure-property relationships that inform current materials research.

SKILLS

- **Battery Materials:** DOE (Minitab), Six Sigma (DMAIC), SPC, LIMS (Electrochemical, XRD, Coin-cell)
- **Materials Informatics:** Pymatgen, ASE, Materials Project API, DFT+U workflows (Quantum ESPRESSO)
- **Machine Learning:** Python (Pandas, NumPy, SQL), PyTorch, PyG, Scikit-learn, NLP, GNNs, XAI, CNNs
- **Visualization & Analytics:** MATLAB, OriginPro, Tableau, Power BI, Matplotlib

PUBLICATIONS & PRESENTATIONS

- Xu, J. ǂ, **Lee, Y. ǂ,** et al.**,** "Beyond Feature Importance: Feature Interactions in Predicting Post-Stroke Rigidity with Graph Explainable AI" Preprint. (2025)
- Yi, S., **Lee, Y.**, et al., "Thematic Analysis of Unstructured Clinical Transcripts with Large Language Models" *GenAI4Health Workshop @ NeurIPS*. (2025)
- **Poster:** "Predicting Post-Stroke Rigidity with Graph XAI" presented at the 2025 UT System AI Symposium in Healthcare, Houston, TX
- **Presentation:** "ML-based study on post-stroke rigidity prediction", CHITA 2025, Austin, TX

HONORS & AWARDS

- Excellent Thesis Award, Korean Society of Industrial and Engineering Chemistry (2017)
  Undergraduate Thesis: "Characteristics of Zinc Ion Adsorption in an Aqueous Solution Using Modified Carbon"
- Merit-Based Scholarship (4 times), Chungnam National University, Daejeon, Korea
  Awarded for outstanding academic achievements, 2010 - 2017

TEACHING EXPERIENCES

- **AI in Health** (Fall 2025): Mentored students on end-to-end ML workflows with emphasis on reproducibility.
- **Database Management** (Fall 2025): Led SQL and data modeling labs for managing large-scale databases.
- **Explainable AI** (Spring 2025): Guided students in applying XAI techniques relevant to interpretable ML.
- **Fine-tuning LLMs** (Spring 2025): Taught transformer fine-tuning for scientific text-to-knowledge pipelines.

LEADERSHIP & ACTIVITIES

- **Vice President,** UT Korean Student Association (2025 - 2026)
- **Mentor**, UT Austin Mentor–Mentee Program (2025)
- **Data Team Leader**, Circle Platform (2022 - 2024)
- **Volunteer**, Austin Annual Book Fair (2024)