Tulga-Erdene **Sodjargal**

tulgaerdene.sodjargal@gmail.com | tulga-rdn | tulga-erdene | Google Scholar

Education

KAIST (Korea Advanced Institute of Science and Technology) — *Daejeon, South Korea*

B.S. in Bio and Brain Engineering; Double Major in Chemistry & AI-Specialized Program — *Aug. 2021 – Present*

- Cumulative GPA: 93.5/100 (as of Aug. 2025).
- Relevant Coursework: Big Data and Machine Learning in Biotechnology, Bio-Data Engineering, Bio-Information Processing, Bio-Data Structures, Statistical ML, Statistical Methods with Computing, AI Chemistry, ML for Molecules and Materials (graduate), Methods in Drug Development (graduate).
- Exchange Program: École Polytechnique Fédérale de Lausanne (EPFL), Switzerland (Sep. 2024 – Jun. 2025).

Skills

Programming	Python, MATLAB, Bash
ML& Data Science	PyTorch, TensorFlow, scikit-learn, pandas
Computational Modeling	RFDiffusion, ProteinMPNN, AlphaFold, MDAnalysis, AutoDock Vina, NAMD
Software Engineering	Git, CI/CD, pytest
Wet-Lab Techniques	Air-sensitive reactions, flash chromatography, $^{1}H/^{13}C$ NMR (1D, 2D)
Languages	English (proficient; TOEFL iBT: 116/120), Russian (bilingual), Mongolian (native)

Experience

Wellman Center for Photomedicine, Harvard Medical School (Prof. Mei X. Wu) — *Boston, USA*

Research Intern — *Jun. 2025 – Present*

- Solely responsible for establishing the lab's computational research capabilities, as the only dry-lab researcher in a wet-lab–focused group.
- Independently designed and executed molecular docking simulations to support and interpret experimental results.
- Acted as the bridge between computation and experimentation, ensuring results were directly applicable to ongoing biological studies.
- Tools: AutoDock Vina, AlphaFold3, NAMD, Gaussian, Bash

Institute of Materials, EPFL (Prof. Michele Ceriotti) — *Lausanne, Switzerland*

Research Intern — *Sep. 2024 – Present*

- Advanced atomistic ML models for molecular property prediction and molecular dynamics by integrating long-range interactions.
- Contributed to open-source educational codebases in Jupyter Notebooks, improving accessibility for students and researchers.
- Gained expertise in PyTorch internals and software engineering practices (CI/CD, Git), optimizing open-source ML tools.
- Results presented as a poster at the 2025 German Physical Society Spring Meeting; contributed to 2 preprints.
- Tools: PyTorch, ASE, LAMMPS, ipywidgets, Git, Bash

SpiderCore Inc. — *Daejeon, South Korea*

Research Intern — *Jun. 2024 – Aug. 2024*

- Implemented graph neural networks for gene therapy design, providing domain-specific chemical expertise.
- Developed a chemically-inspired self-supervised learning task, improving ML performance to state-of-the-art levels.
- Tools: TensorFlow, RDKit

University of Illinois Urbana-Champaign — *Remote*

Research Intern — *Feb. 2024 – Present*

- Conducted comparative analysis of cellular membranes with differing compositions for joint drug design projects.
- Refactored and optimized analysis methods, reducing runtime by >30× and enabling novel structural insights.
- Co-authored a manuscript currently in preparation.
- Tools: MDAnalysis, scikit-learn

KAIST — *Daejeon, South Korea*

Undergraduate Research Intern — *Mar. 2022 – Nov. 2023*

- Designed potentially therapeutic antibody variants (Prof. Byung-Ha Oh, Jun 2023 – Nov. 2023).
- Developed novel chemical reactions using combinatorial methods (Prof. Yoonsu Park, Mar. 2022 – Jun. 2023).
- Tools: RFDiffusion, AlphaFold2, ProteinMPNN, Bash

Publications

Learning Long-Range Representations with Equivariant Messages

E Rumiantsev, MF Langer, TE Sodjargal, M Ceriotti, P Loche

arXiv:2507.19382

scicode-widgets: Bringing Computational Experiments to the Classroom with Jupyter Widgets

A Goscinski, TJ Baird, D Du, J Prado, D Suman, TE Sodjargal, S Bonella, G Pizzi, M Ceriotti

arXiv:2507.05734

Projects

Predicting Demand for Electronic Parts

3rd POSTECH–UNIST–KAIST Data Science Competition

- Developed forecasting models with cost-aware optimization to predict demand for electronic parts.
- Achieved 5th place (Silver Award) among 20+ teams from leading Korean universities.
- Tools: Nixtla

Analyzing Workplace Discrimination in Korea

Final Project, *Statistical Methods with Computer*

- Conducted statistical analysis to uncover national trends in workplace discrimination.
- Applied EDA, hypothesis testing, and clustering to identify key patterns.
- Tools: scikit-learn, pandas

Biomedical Information Systems for Future Healthcare

Final Project, *Bio-Information Processing*

- Designed a biomedical and pharmacokinetics database system integrating genomics and pharmacokinetics datasets.
- Implemented an intuitive command-line interface simulating healthcare provider use cases.
- Tools: PostgreSQL, psycopg2

Housing Price Prediction

Final Project, *Statistical ML*

- Built predictive models for housing prices using open-source datasets.
- Performed EDA, feature engineering, preprocessing, and hyperparameter tuning.
- Ranked 7th out of 60 projects in class competition.
- Tools: scikit-learn, pandas, Optuna

Prediction of pK_{BHX} on a Small Dataset

Final Project, *AI Chemistry*

- Developed a graph convolutional neural network (GCNN) to predict hydrogen bond basicity (pK_{BHX}) on a dataset of ~350 molecules.
- Applied preprocessing, feature engineering, regularization, and hyperparameter tuning.
- Ranked 4th out of 13 (1st among undergraduates).
- Tools: PyTorch, RDKit

SNP Analysis for COVID-19 Delta Variant Surge

Final Project, *Bio-Data Structures*

- Implemented a heuristic global sequence alignment algorithm in pure Python to detect SNPs.
- Analyzed SNP patterns associated with the Delta variant surge in England.
- Performed literature review to interpret biological roles of identified SNPs.