

Sangwon Lee

Last updated: Feb 10, 2024

Professional, Data Analysis PJT, Analytical Sciences Center, LG Chem

LG Science Park, Seoul, Republic of Korea

lsw91.main@gmail.com | (+82) 10-2291-6467 | [GitHub](#) | [Google Scholar](#)

Education

Korea Advanced Institute of Science and Technology (Daejeon, Korea)

Postdoctoral Researcher in Applied Science Research Institute (Apr. 2021)

Ph.D. in Chemical and Biomolecular Engineering (Feb. 2020)

Dissertation: Nanoporous Materials Discovery for Energy and Environmental Applications using Machine Learning
(Advisor: Jihan Kim)

University of Seoul (Seoul, Korea)

M.S. in Chemical Engineering (Feb. 2016)

Thesis: Prediction of chemical potential and phase transition of molecular crystals using Monte Carlo simulations
(Advisor: Jaeeon Chang)

B.S. in Chemical Engineering (Feb. 2014)

Research interests

- Artificial intelligence for materials
- Machine learning
- Bayesian inference
- Energy management system
- Molecular simulations
- Analytical chemistry
- Software engineering
- Numerical analysis

Publications

*The number of citations is obtained by scraping Google Scholar

First and co-first author papers

Title	Citations	Year
(Contributed Equally) Inverse design of porous materials using artificial neural networks, B Kim, S Lee, J Kim, <i>Science advances</i> 6 (1), eaax9324, 2020 <div> <div>deep learning</div> <div>molecular simulations</div> <div>materials discovery</div> </div>	248	2020
Computational screening of trillions of metal–organic frameworks for high-performance methane storage, S Lee, B Kim, H Cho et al., <i>ACS Applied Materials & Interfaces</i> 13 (20), 23647-23654, 2021 <div> <div>deep learning</div> <div>molecular simulations</div> <div>materials discovery</div> </div>	94	2021
User-friendly graphical user interface software for ideal adsorbed solution theory calculations, S Lee, JH Lee, J Kim, <i>Korean Journal of Chemical Engineering</i> 35, 214-221, 2018 <div> <div>numerical analysis</div> <div>software design</div> <div>thermodynamics</div> </div>	90	2018
Predicting performance limits of methane gas storage in zeolites with an artificial neural network, S Lee, B Kim, J Kim, <i>Journal of Materials Chemistry A</i> 7 (6), 2709-2716, 2019 <div> <div>deep learning</div> <div>molecular simulations</div> <div>materials discovery</div> </div>	43	2019
(Contributed Equally) Computational Analysis of Linker Defective Metal–Organic Frameworks for Membrane Separation Applications, H Kim, S Lee, J Kim, <i>Langmuir</i> 35 (11), 3917-3924, 2019 <div> <div>deep learning</div> <div>molecular simulations</div> <div>numerical analysis</div> </div>	10	2019
(Contributed Equally) Machine learning-based discovery of molecules, crystals, and composites: A perspective review, S Lee, H Byun, M Cheon, J Kim, JH Lee, <i>Korean Journal of Chemical Engineering</i> , 1-12, 2021 <div> <div>machine learning</div> <div>materials discovery</div> </div>	6	2021
(Contributed Equally) Performance Evaluation of Deep Learning Architectures for Load and Temperature Forecasting under Dataset Size Constraints and Seasonality, W Choi, S Lee, <i>Energy and Buildings</i> , 113027, 2023 <div> <div>deep learning</div> <div>forecasting</div> <div>energy management systems</div> </div>	5	2023
Chemical potential and solid-solid equilibrium of near-spherical Lennard-Jones dumbbell crystal, S Lee, M Kim, J Chang, <i>Korean Journal of Chemical Engineering</i> 33, 1047-1058, 2016 <div> <div>molecular simulations</div> <div>numerical analysis</div> <div>statistical mechanics</div> </div>	1	2016

Title	Citations	Year
(Contributed Equally) Interpretable deep learning model for load and temperature forecasting: Depending on encoding length, models may be cheating on wrong answers, W Choi, S Lee, <i>Energy and Buildings</i> , 113410, 2023 deep learning forecasting energy management systems		2023
Explainable kinetic modeling for aerobic biodegradation of poly(lactic acid) using Bayesian inferences, S Lee, K Park, J Lee et al., <i>Submitted</i> Bayesian inferences kinetic modeling bio-plastics		2024

Other papers

Title	Citations	Year
Applications of machine learning in metal-organic frameworks, S Chong, S Lee, B Kim, J Kim, <i>Coordination Chemistry Reviews</i> 423, 213487, 2020	116	2020
Finding hidden signals in chemical sensors using deep learning, SY Cho, Y Lee, S Lee et al., <i>Analytical chemistry</i> 92 (9), 6529-6537, 2020	37	2020
Size-Matching Ligand Insertion in MOF-74 for Enhanced CO ₂ Capture under Humid Conditions, BL Suh, S Lee, J Kim, <i>The Journal of Physical Chemistry C</i> 121 (44), 24444-24451, 2017	35	2017
Finely tuned inverse design of metal–organic frameworks with user-desired Xe/Kr selectivity, Y Lim, J Park, S Lee, J Kim, <i>Journal of Materials Chemistry A</i> 9 (37), 21175-21183, 2021	21	2021
New model for S-shaped isotherm data and its application to process modeling using IAST, S Ga, S Lee, G Park, J Kim, M Realff, JH Lee, <i>Chemical Engineering Journal</i> 420, 127580, 2021	14	2021
Computational design of metal–organic frameworks with unprecedented high hydrogen working capacity and high synthesizability, J Park, Y Lim, S Lee, J Kim, <i>Chemistry of Materials</i> 35 (1), 9-16, 2022	12	2022
Isotherm parameter library and evaluation software for CO ₂ capture adsorbents, S Ga, S Lee, J Kim, JH Lee, <i>Computers & Chemical Engineering</i> 143, 107105, 2020	10	2020
Deep learning-based initial guess for minimum energy path calculations, H Park, S Lee, J Kim, <i>Korean Journal of Chemical Engineering</i> 38, 406-410, 2021	1	2021
Real-time probabilistic backfill thermal property estimation method enabling estimation convergence judgment, W Choi, S Lee, BH Dinh, YS Kim, <i>Case Studies in Thermal Engineering</i> 48, 103108, 2023		2023

Title	Citations	Year
Automatic Object Extraction from Electronic Documents Using Deep Neural Network, H Jang, Y Chae, S Lee, J Jo, <i>KIPS Transactions on Software and Data Engineering</i> 7 (11), 411-418, 2018		2018

Projects

LG Chem

Predicting oxidation states and coordination numbers from XANES using deep learning (2021, with LG AI Research)

- Contributions: designed deep learning architecture, developed knowledge transfer method
- *Patent application submitted*

Building MLOps infrastructure for instance segmentation models (2021-2022)

- Contributions: designed infrastructure architecture, established the infrastructure

Development of a segmentation tool using segment anything model (2023)

- Contributions: worked on back-end development

Estimation of length dimension of carbon nanotubes in SEM images (2023)

- Contributions: developed a length estimation algorithm, worked on back-end development

Monocrystallinity analysis of cathodes from EBSD using machine learning (2022-2023)

- Contributions: developed the theory of monocrystallinity, designed deep learning models, worked on back-end development
- *Patent application submitted*

Tortuosity estimation of battery separators using Monte Carlo simulations (2022)

- Contributions: developed a tortuosity calculation algorithm, created a 3D visualization tool

Predicting biodegradation of polymers using reaction kinetics and Bayesian inference (2021-2023)

- Contributions: developed a reaction kinetics model for polymer biodegradation, designed a Bayesian inference model, contributed to back-end development

EBSD orientation analysis for lithium-ion path in cathodes (2022)

- Contributions: developed the theory, contributed to back-end development

Determination of the diffusion coefficient of lithium in separators (2023)

- Contributions: developed the theory for the diffusion model

Orientation analysis of glass fibers in the engineering plastics (2021)

- Contributions: developed deep learning model

Classification of crystal phases in cathodes from 4D-STEM data (2023)

- Contributions: developed a diffraction pattern classification algorithm

Deep document understanding (2020, with LG AI Research)

- Contributions: developed a SMILES image to smiles text model using ResNet and transformer

Softwares & Libraries

DeepTimeSeries (Python, PyTorch)

[GitHub Repository](#)

A deep learning library designed for time series forecasting, built on the PyTorch framework

PORMAKE (Python)

[GitHub Repository](#)

A Python library for the construction of porous materials using topology and building blocks

IAST++ (C++)

[GitHub Repository](#)

User-friendly software for ideal adsorbed solution theory calculations

GRIDAY (C++)

[GitHub Repository](#)

An energy shape calculator for the porous materials

PC-SAFT (Fortran)

[GitHub Repository](#)

A software tool that facilitates the calculation of chemical equilibrium using perturbed chain statistical associating fluid theory

MonteCarloSimulator (Fortran)

[GitHub Repository](#)

A molecular Monte Carlo simulator employed for free energy calculations of molecules in both fluid and solid states

dartwork-mpl (Python, Matplotlib)

[GitHub Repository](#)

A Matplotlib styling and utility package crafted for the creation of publication-ready figures

LabeledImage (Python)

(private, LG Chem)

A library designed for managing labeled images, specifically tailored for instance segmentations. This library includes tools for format conversion, patch detection aggregation, and image augmentation.

Deep learning architectures

MOF-NET (Tensorflow)

[GitHub Repository](#)

[GitHub Repository](#) (Graph-network version)

A deep neural network engineered for predicting MOF (Metal-Organic Framework) properties based on topology and building blocks

ZeoGAN (Tensorflow)

[GitHub Repository](#)

A generative adversarial network (GAN) designed for the inverse design of zeolites

ESGAN (Tensorflow)

[GitHub Repository](#)

A generative adversarial network (GAN) created for generating energy shapes of porous materials

XANES-Net (PyTorch)

(private, LG Chem, with LG AI Research)

An architecture developed for predicting oxidation states and coordination numbers from XANES data, leveraging knowledge transfer from crystal structures

Skills

•: Usable with some time investment

••: Proficiently usable

•••: Usable at an expert level

Programming languages

- C/C++ •••
- Python •••
- Fortran •••
- Mathematica •••
- Javascript •
- Rust •

Machine learning

- PyTorch •••
- Tensorflow •••
- PyTorch Lightning •••
- PyTorch Geometric •
- Detectron2 ••
- MMDetection ••
- scikit-learn ••
- XGBoost •
- JAX ••
- Darts ••

Bayesian inference

- PyMC •••
- Pyro •

Data visualization

- Matplotlib •••
- Seaborn ••
- Plotly ••
- Bokeh •
- d3.js •

Version control

- Git ••
- GitHub ••

CI/CD

- Github Actions ●●

Package and project management

- Poetry & Pyenv ●●●
- Conda ●●

Unit test

- pytest ●●

Data structure & numerical analysis

- Pandas ●●●
- Xarray ●●
- NumPy ●●●
- SciPy ●●●

GUI programming

- Qt, PySide6 ●●●

Symbolic computation

- Mathematica ●●●
- SymPy ●●

Computer vision

- OpenCV ●●●
- scikit-image ●●●

3D visualization

- PyVista ●●●
- OpenGL (GLSL) ●
- Panda3D ●●

Containerization

- Docker ●●
- Docker compose ●●

Back-end & ML serving

- FastAPI ●●
- MLflow models ●●

Documentation

- Sphinx ●●
- LaTeX ●●

MLOps

- MLflow ●●●
- Tensorboard ●

Reverse proxy

- Traefik ●●

Configuration management

- Hydra ●●

Operating system

- Linux (Ubuntu, CentOS) ●●

Molecular simulations

- PyMatGen ●●●
- ASE ●●●
- LAMMPS ●●●
- VASP ●
- PySCF ●

Database

- SQLite ●

Presentations

Oral presentations

- Construction of MOF Structures from Topology and Building Blocks, 2019 KICChE Fall Meeting and International Symposium
- Evaluating Performance Limits of Methane Storage Separations of Porous Materials Using Artificial Neural Network, IUMRS-ICEM 2018

Posters

- IAST++: Software for Ideal Adsorbed Solution Theory (IAST) Calculations with User-friendly Interface, 2017 KICChE Fall Meeting and International Symposium
- Sensitivity Analysis of CO₂ Capture Materials in Post-combustion Flue Gas, ChemIndex 2016
- Order-disorder transition and free energy of the crystal of Lennard-Jones diatomics, 2014 KICChE Fall Meeting

Awards

Gold Prize (2nd), Engineering Mathematics Competition, University of Seoul, Nov. 2012