Ihor Neporozhnii

ihor.neporozhnii@mail.utoronto.ca | +1 (778)-697-8423

linkedin.com/in/ihor-neporozhnii | github.com/ineporozhnii | Google Scholar | My Website

Education

University of Toronto — Toronto, Canada

PhD – Computational Chemistry — *September 2021 – expected March 2027*

- Thesis focus: Machine Learning for Accelerated Discovery of Medicines and Materials

Taras Shevchenko National University of Kyiv — Kyiv, Ukraine

Bachelor of Science – Physics — *September 2017 – June 2021*

Selected Publications

- **I. Neporozhnii**, J. Roy, E. Bengio, J. Hartford, Efficient Biological Data Acquisition through Inference Set Design, In proceedings of the International Conference on Learning Representations (ICLR), 2025.
- **I. Neporozhnii**, Z. Wang, R. Bajpai, C. Gomez, N. Chakraborty, I. Tamblyn, O. Voznyy, Navigating Materials Space with ML-Generated Electronic Fingerprints, Physical Review Letters (Under review).
- F. Dinic, **I. Neporozhnii**, O. Voznyy, Machine learning models for the discovery of direct band gap materials for light emission and photovoltaics, Computational Materials Science, 2024.
- F. Dinic, Z. Wang, **I. Neporozhnii**, U. B. Salim, R. Bajpai, N. Rajiv, V. Chavda, V. Radhakrishnan, O. Voznyy, Strain data augmentation enables machine learning of inorganic crystal geometry optimization, Patterns, 2023.
- M. Abdinejad, S. Subramanian, M. K. Motlagh, M. Noroozifar, S. Duangdangchote, **I. Neporozhnii**, D. Ripepi, D. Pinto, M. Li, K. Tang, J. Middelkoop, A. Urakawa, O. Voznyy, H. Kraatz, T. Burdyny, Insertion of MXene-Based Materials into Cu–Pd 3D Aerogels for Electroreduction of CO2 to Formate, Advanced Energy Materials, 2023.

Work Experience

Machine Learning Research Engineer Intern — August 2025 – Present

Cognano — *Toronto, Canada*

- Developing data processing and prediction scoring pipelines for the VHH Epitope Prediction Challenge.

Machine Learning Research Intern — April 2024 – December 2024

Valence Labs | Recursion Pharmaceuticals — *Montreal/Toronto, Canada*

- Developed Inference Set Design - an active learning approach that **reduces costs of high-throughput library screens by up to 60%** while preserving the high accuracy of the acquired data, using **Python** and **PyTorch**.
- Trained **large transformer models** to predict the biological effects of compounds on cells.
- Utilized **protein language models** for feature engineering to improve a downstream model.
- Implemented probabilistic stopping criterion to avoid unnecessary experimentation.

Research Assistant — September 2021 – Present

University of Toronto — *Toronto, Canada*

- Developing machine learning methods to **improve the efficiency of discovering novel materials for clean energy applications and medicines**.
- Built ProDosNet - a Graph Neural Network (GNN) that decreased the computational cost of materials electronic structure analysis **by 3 orders of magnitude**, providing a way to accelerate the discovery of new materials for clean energy applications.
- Developed a data processing pipeline and a neural network using **Python** and **PyTorch** to classify materials band gap into direct and indirect. The resulting model achieved a **40% improvement** in identifying direct band gap materials compared to random search.
- Performed **density functional theory (DFT)** simulations of the CO_2 reduction reaction to support experimental results, contributing to a publication in Advanced Energy Materials.

Machine Learning Bootcamp Instructor April 2023 – October 2023

University of Toronto | Alliance for AI-Accelerated Materials Discovery (A3MD) *Toronto, Canada*

- I organized and conducted two 5-day bootcamps on **machine learning and chemistry for 25 scientists** at LG Group and Total Energies.
- Prepared lectures and tutorials that covered chemical data processing with **Python** and developing deep learning models for chemistry with **PyTorch** and **TensorFlow**.

Research Intern March 2021 – April 2021

CNRS, Université Paris-Saclay *Paris, France*

- Developed an algorithm to analyze signals from the JUNO neutrino experiment.
- Implemented the algorithm in **C++**, which resulted in a **50% improvement** in the accuracy of signal reconstruction.

Research Engineer April 2019 – August 2021

Institute of Physics, National Academy of Sciences of Ukraine *Kyiv, Ukraine*

- Conducted experiments with nanomaterials in ultra-high vacuum conditions. The results of this study are published in Surface Science.

Research Intern July – August 2020

Jagiellonian University *Krakow, Poland*

- Performed cross-match of neutrino and Gamma-ray burst datasets, conducted statistical analysis with **Python**.

Research Intern July – August 2019

Institute of Nuclear Physics, Polish Academy of Sciences *Krakow, Poland*

- Analyzed data from CERN Atlas experiment using **Machine Learning** and **Monte Carlo** methods.

Project Experience

Active Learning for building Maps of Biology | *Valence Labs, Recursion Pharmaceuticals* 2024

- Applied **transfer learning**, **active learning**, and **model ensembling** techniques to predict gene-compound relationships achieving **80% accuracy**, using **Python** and **PyTorch**.
- Handled and analyzed biological data, including molecular structures, cell images, model embeddings, experimental measurements, and simulation results.
- The project resulted in a publication and a poster presentation at ICLR 2025 conference.

Discovering Organic Light Emitting Molecules with ML | *University of Toronto* 2023

- Utilized **Python** and **RDKit** to preprocess and featurize molecular datasets.
- Developed machine learning model ensembles using **Python**, **PyTorch**, and **Scikit-learn** to predict emission properties of organic molecules, achieving 0.78 R^2 on experimental data.
- Ran multiple **active learning** rounds to acquire additional data, improving model predictions by 10%.
- Identified promising molecules with machine learning that were synthesized and characterized in the lab, validating model predictions.

Exploring Space of Materials with Deep Learning | *University of Toronto* 2023

- Implemented a deep learning model for predicting materials electronic densities of states, using **Python** and **PyTorch**.
- Developed robust material representations, enabling targeted exploration of chemical space and demonstrating effectiveness across 5 distinct applications, including **catalysts** and **solid electrolytes**.
- Performed **molecular dynamics simulations** to verify model predictions.
- Designed a loss function that improved training stability and reduced prediction error by 20%.
- Deployed the model and a chemical space analysis tool: ProDosMate.

Teaching Experience

Teaching Assistant | *Physics for Life Sciences at University of Toronto* September 2021 – Present

- I conduct practical sessions for undergraduate students enrolled in courses Physics I and II for Life Sciences, including lab experiments and computer simulations.

Honors and Awards

Connaught International Scholarship for Doctoral Students | *University of Toronto ($50,000)* 2021 – 2025
Climate Positive Energy Graduate Student Scholarship | *University of Toronto ($15,000)* 2024
Invited talk: Accelerated discovery of materials with Machine Learning | *Climate Positive Energy* 2024
Best presentation award: Navigating Material Space with ML | *Canadian Chemistry Conference* 2023
Scholarship for Bachelor students | *Taras Shevchenko National University of Kyiv ($4,000)* 2017 – 2021
Second place at the All-Ukrainian Physicists' Tournament | *International Physicists' Tournament* 2019

Conference Presentations

Efficient High-Throughput Compound Library Screens with Active Learning 2025
Poster presenter, Accelerate Conference, Toronto, Canada

Efficient Biological Data Acquisition through Inference Set Design 2025
Poster presenter, International Conference on Learning Representations (ICLR), Singapore

Navigating Material Space with ML-Generated Electronic Fingerprints 2024
Poster presenter, Materials for Sustainable Development Conference (MATSUS24), Barcelona, Spain

Navigating Material Space with ML-Generated Electronic Fingerprints 2023
Poster presenter, Accelerate Conference, Toronto, Canada

Accelerated discovery of battery materials using ML-predicted Density of States 2023
Talk, Climate Positive Energy Research Day, Toronto, Canada

Navigating Material Space with ML-Generated Electronic Fingerprints 2023
Talk, Canadian Chemistry Conference and Exhibition, Vancouver, Canada

Predicting Materials Density of States with Graph Neural Networks 2022
Talk, MRS Fall Meeting and Exhibit, Boston, United States

Predicting Materials Density of States with Graph Neural Networks 2022
Poster presenter, Accelerate Conference, Toronto, Canada

Machine Learning Methods for Predicting Density of States 2022
Poster presenter, Canadian Symposium on Theoretical and Computational Chemistry, Kelowna, Canada

Machine Learning Methods for Predicting Density of States 2022
Talk, Canadian Chemistry Conference and Exhibition, Calgary, Canada

Spatio-temporal correlation between Gamma-ray bursts and High-energy neutrino 2020
Talk, WDS Conference, Prague, Czech Republic

Technical Skills

- **Languages**: Python, R, Julia, C++
- **Frameworks and libraries**: PyTorch, TensorFlow, JAX, Transformers, Scikit-learn, RDKit, Pandas, Numpy.
- **Machine Learning**: Deep Learning, Active Learning, Transfer Learning, GNNs, LLMs, Machine Learning Force Fields.
- **Computational Chemistry**: Density Functional Theory (VASP, CP2K, PySCF), Semi-empirical methods (xTB), LAMMPS.
- **High-performance computing**: Training models and running molecular dynamics simulations on multi-node GPU and CPU clusters with SLURM, Hydra, and PyTorch Lightning.
- **Software development**: Version control (Git, Github), VS Code, JupyterLab.