Ihor Neporozhnii

Toronto, Canada ihor.neporozhnii@mail.utoronto.ca
Website linkedin.com/in/ihor-neporozhnii github.com/ineporozhnii

Education

University of Toronto **Sep. 2021 – Present**
PhD – Machine Learning for Chemistry *Toronto, Canada*

Taras Shevchenko National University of Kyiv **Sep. 2017 – Jun. 2021**
Bachelor of Science – Physics *Kyiv, Ukraine*

Technical Skills

Programming Languages: Python, R, SQL, Julia, C++
Machine Learning: PyTorch, TensorFlow, CUDA, Scikit-learn, Flux ML
High-performance computing: remote computing on GPU and CPU clusters
Version control (Git, GitHub), Linux, MacOs, Windows

Work Experience

Research Assistant **September 2021 – Present**
University of Toronto *Toronto, Canada*

- I develop computational methods to accelerate the discovery of materials and medicines. I have been extensively using programming languages **(Python, Julia)**, machine learning (ML) libraries **(PyTorch, Tensorflow)**, and quantum chemistry codes for molecular simulations **(DFT, xTB)**.
- My responsibilities include working with chemical and biological datasets ($>$ 1,000,000 entries), designing and training ML models (using **remote clusters** with multiple GPUs), and deploying ML models.
- Developed **ProDosMate** - a Machine Learning framework that reduced the computational cost of electronic structure analysis for materials by $\times\mathbf{10,000}$ times, allowing researchers to find new materials for sustainable energy faster.
- Earned **Data Science certificate** (University of Toronto, SciNet)

Machine Learning Research Intern **April 2024 – December 2024**
Valence Labs (Powered by Recursion Pharmaceuticals) *Montreal/Toronto, Canada*

- Developed **Inference Set Design** - an active learning method for efficient biological and chemical data acquisition.

Bootcamp Instructor **April 2023 – October 2023**
Alliance For AI-Accelerated Materials Discovery (A3MD) at UofT *Toronto, Canada*

- Conducted 5-day Bootcamps on **machine learning** for 25 scientists at **LG** and **Total Energies**. Lectures and tutorials covered data processing with Python and developing ML models with **PyTorch**, **TensorFlow**, **Scikit-learn**

Research Intern **March 2021 – April 2021**
CNRS, Université Paris-Saclay *Paris, France*

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

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.

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.

Scholarships and Awards

Climate Positive Energy Graduate Student Scholarship **2023 – 2024**
University of Toronto, Climate Positive Energy *Total value: $15,000*

Connaught International Scholarship for Doctoral Students **2021 – 2024**
University of Toronto *Total value: $30,000*

Publications

Efficient Biological Data Acquisition through Inference Set Design **2025**
I. Neporozhnii, *J. Roy, E. Bengio, J. Hartford* *ICLR*
doi: https://doi.org/10.48550/arXiv.2410.19631

Developed **Inference Set Design (ISD)** *- an active learning method for acquiring biological and chemical data. Deploying ISD* **reduces experimental costs by 60%** *while preserving the high accuracy of the acquired data.*

Navigating Materials Space with ML-Generated Electronic Fingerprints **2025**
I. Neporozhnii, *Z. Wang, R. Bajpai, C. Gomez, N. Chakraborty, I. Tamblyn, O. Voznyy* *Preprint*
doi: https://doi.org/10.26434/chemrxiv-2023-j1szt

I developed a Graph Neural Network (GNN) to predict the Density of States of materials that decreased the computational cost of electronic structure analysis by **4 orders of magnitude**, *providing a way for researchers to discover new materials for clean energy applications faster.*

Machine learning models for the discovery of direct band gap materials for light emission and photovoltaics **2023**
Computational Materials Science
F. Dinic, **I. Neporozhnii**, *O. Voznyy*
doi: https://doi.org/10.1016/j.commatsci.2023.112580

Developed materials data processing pipeline for machine learning model.

Strain data augmentation enables machine learning of inorganic crystal geometry optimization **2023**
F. Dinic, Z. Wang, **I. Neporozhnii**, *U. Bin Salim, R. Bajpai, N. Rajiv, V. Chavda,* *Patterns*
V. Radhakrishnan, and O. Voznyy. doi: https://doi.org/10.1016/j.patter.2022.100663

I developed a machine learning (ML) model that enables accurate prediction of the formation energy for non-equilibrium structures which previously required computationally expensive DFT calculations.

Insertion of MXene-Based Materials into Cu–Pd 3D Aerogels for Electroreduction of CO2 to Formate **2023**
Advanced Energy Materials
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.-B. Kraatz, T. Burdyny. doi: https://doi.org/10.1002/aenm.202300402

I conducted Density Functional Theory (DFT) calculations using VASP software.

Mesoscopic self-ordering in oxygen doped Ce films adsorbed on Mo(112) **2021**
T. Afanasieva, A. Fedorus, A. Goriachko, A. Naumovets, **I. Neporozhnii**, *and D. Rumiantsev.* *Surface Science*
doi: https://doi.org/10.1016/j.susc.2020.121766

I conducted experiments with nanomaterials in ultra-high vacuum conditions.

Conference Presentations

Navigating Material Space with ML-Generated Electronic Fingerprints **March 2024**
Materials for Sustainable Development Conference (MATSUS24) (Poster, presenter) *Barcelona, Spain*
Navigating Material Space with ML-Generated Electronic Fingerprints **August 2023**
Accelerate Conference 2023 (Poster, presenter) *Toronto, Canada*
Accelerated discovery of battery materials using ML-predicted Density of States **August 2023**
Climate Positive Energy Research Day (Talk, invited speaker) *Toronto, Canada*
Navigating Material Space with ML-Generated Electronic Fingerprints **June 2023**
Canadian Chemistry Conference and Exhibition 2023 (Talk, presenter, received presentation award) *Vancouver, Canada*
Machine learning methods for predicting density of states **December 2022**
MRS Fall Meeting & Exhibit 2022 (Talk, presenter) *Boston, United States*
Machine learning methods for predicting density of states **August 2022**
Accelerate Conference 2022 (Poster, presenter) *Toronto, Canada*
Machine learning methods for predicting density of states **June 2022**
The Canadian Symposium on Theoretical and Computational Chemistry (Poster, presenter) *Kelowna, Canada*
Machine learning methods for predicting density of states **June 2022**
Canadian Chemistry Conference and Exhibition 2022 (Talk, presenter) *Calgary, Canada*
Spatio-temporal correlation between Gamma-ray bursts and High-energy neutrino **September 2020**
WDS 2020 (Talk, presenter) *Prague, Czech Republic*