

Piyush R. Maharana

Linkedin Github

HuggingFace

Email : piyushmaharana15@gmail.com

EDUCATION

- **Central University of Punjab** Punjab, India
• **Masters in Computational Physics: Gold medallist CGPA: 8.36/10.00** Nov 2021 - Jun 2023
Courses: Numerical Methods, Electronic Structure Theory, Molecular Dynamics, Density Functional Theory, Atomic and Molecular Spectroscopy
- **University of Delhi** New Delhi, India
• **Bachelors in Physics: CGPA: 8.48/10.0** Jun 2018- Jun 2021
Courses: Classical Mechanics, Quantum Mechanics, Electromagnetic Theory, Statistical Mechanics, Nuclear Physics, Mathematical Physics

SKILLS SUMMARY

- **Languages:** Python, FORTRAN, C++
- **Softwares:** QUANTICS, ORCA, VASP, LAMMPS
- **Interests:** Large Language Models, Natural Language Processing, Neural Network Potentials, Mechanistic Interpretability

EXPERIENCE

- **High Entropy Alloys for hydrogen storage at room temperature** CSIR-National Chemical Laboratory, India
Advisor - Dr. Kavita Joshi Aug 2023- Present
 - **Project Assistant -I:** Using Natural Language Processing and Large Language Models for scientific data extraction from literature to create databases for training machine learning models for hydrogen storage applications. website

ACADEMIC PROJECTS/PUBLICATIONS

- **Retrieval Augmented Generation for extraction of data from scientific literature** NCL, Pune
ChemRxiv Preprint Nov 2024
 - **Abstract:** Using large language models paired with retrieval augmented generation, a dataset of metal hydrides for solid-state hydrogen storage was created. The accuracy obtained was 93 % in the cases tested. Thus, we demonstrate a pipeline to create datasets from scientific literature at minimal computational cost and high accuracy.
Maharana, P.R., & Joshi, K. (2024). Retrieval augmented generation for building datasets from scientific literature.
[Paper Link]
- **Reflections from the 2024 LLM Hackathon for Applications in Materials Science and Chemistry** NCL, Pune
arXiv Preprint Nov 2024
 - **Abstract:** With Context-Enhanced Material Property Prediction (CEMP) we sought to improve composition-based property prediction models for novel materials by providing natural language input alongside the chemical composition of interest. In doing so, we leverage the ability of large language models (LLMs) to capture the broader context relevant to the composition. We demonstrate that providing context embeddings generated from automatically extracted experimental conditions and structure-property relationships reduce MAE by 3–10% and 10–15% respectively.
[Paper Link]
- **HyStor: An Experimental Database of Hydrogen Storage Properties** NCL, Pune
Co-author May 2024
 - **Abstract:** In this work, we introduce the HyStor database, consisting of 1280 metal alloys along with their hydrogen storage capacities ($H_2wt\%$) as a function of absorption temperature. Given the lack of updates in the existing open access HydPark database since 2002, we sourced compositions from recent research articles and various patent documents, resulting in a total of 468 compositions.
Wilson, N., Verma, A., Maharana, P.R., Sahoo, A.B., & Joshi, K. (2024). HyStor: An experimental database of hydrogen storage properties for various metal alloy classes. International Journal of Hydrogen Energy. [Paper Link]
- **Quantum Dynamics of H_2 formation on interstellar dust grains** Central University of Punjab, India
Advisor - Dr. Kousik Giri Feb - Jun 2023
 - **Master's Thesis:** The details of H_2 formation in the interstellar medium are still unclear. Using MCTDH the dynamics of the formed H_2 molecule on a carbonaceous surface was determined to understand the partitioning of the H_2 recombination energy in the collinear Eley-Rideal mechanism. Our results show the introduction of the atoms neighbouring the carbon over which H is chemisorbed into the PES of Coronene $C_{24}H_{12}$, gives the correct energy partitioning with the nascent H_2 taking $\sim 85\%$ of the recombination energy.
- **Structure and Kinematics of the Milky Way Galaxy using Atomic Hydrogen**
Conference Paper Jun 2021

- **Abstract:** Presented this paper based on my undergrad thesis in 3rd International Conference on Recent Advances In Fundamental And Applied Sciences(RAFAS), 25-26 June 2021 *VBatra and PR Maharana 2022 J.Phys.:Conf.Ser.2267*
[Paper Link]

• **Spiral Structure and Dynamics of the Milky Way Galaxy**

University of Delhi, India

Jan -Jun 2021

Advisor - Dr. Vandana Batra

- **Undergraduate Thesis:** The hyper-fine transition of atomic hydrogen produces radiation of $\nu=1.4$ GHz which is used to detect intergalactic HI gas in the Milky Way. An observation was carried out from $20^\circ < l < 220^\circ$ and $-20^\circ < b < 20^\circ$ in steps of 5° using the remotely operated SALSA-Onsala 2.3m Radio Telescope located at Onsala Space Observatory, Sweden. The Perseus, Local and Outer arms of the Milky Way Galaxy were clearly identified. Using the Tangent Point Method the rotation curve till the solar circle was determined giving a value of $V_c(R_o) = 240.16 \pm 7.03 \text{ km s}^{-1}$. [Blog Link]

• **Spectral Index of Synchrotron emission of Crab Nebula**

University of Delhi, India

Sep 2020

Advisor - Dr. Vandana Batra

- **Abstract:** Using archival flux density data ranging from radio to optical from different observations, the spectral index of the nebula in the different electromagnetic regimes was determined. [Paper Link]

COURSES/WORKSHOPS

IIIT Hyderabad-8th summer school on AI : research training event aiming at equipping participants with the relevant advances in the critical and fast developing area of artificial intelligence. Link **July 2024**

UCL—Future of Materials Discovery Workshop : workshop on AI and machine learning in the physical sciences Link **June 2024**

LLM Hackathon for Applications in Materials and Chemistry: teamed up with the University of Liverpool to apply LLMs to improve property prediction for chemical compositions. [Our Submission] **May 2024**

IISER Pune—Natural Language Processing: Attended this semester course on NLP **Jan-May 2024**

CDAC—Quantum Accelerated Computing: Attended lectures and hands-on session on Quantum Computing. **Dec 2023**

IUCAA—Introductory Summer School in Astronomy and Astrophysics: Attended lectures on various different topics in astrophysics. **May-June 2021**

RAD @ Home: Learned how to make images using the data collected by GMRT and other telescopes and understand UV-optical-IR-radio(RGB-C) images of galaxies to classify galaxies into different subtypes. **June 2021**

EXTRACURRICULARS

Volunteering: I teach underprivileged kids and distribute surplus food to those in need as part of the NGO Robin Hood Army **Aug 2023-Present**

Astronomy Club: Searching for suspected comets visible in SOHO/LASCO satellite images by participating in the Sungrazer project and taking part in astronomical observations and workshops and lectures on astrophysics and programming at Nehru planetarium. Link **2018-2022**

International Asteroid Search Campaign: Two provisional asteroid discoveries recorded **Apr-May 2018**

Supernova Hunters: Discovered supernovae in Pan-STARRS1 data as part of a citizen science project Link **2017-2019**

REFERENCES

Dr. Kavita Joshi Principal Scientist

Physical and Materials Chemistry Division, CSIR-National Chemical Laboratory. k.joshi@ncl.res.in Link

Dr. Kousik Giri Assistant Professor

Department of Computational Sciences, Central University of Punjab kousik.giri@cup.edu.in Link

Dr. Vandana Batra Associate Professor

Department of Physics, Bhaskaracharya College of Applied Sciences, University of Delhi vandana.batra@bcas.du.ac.in Link

Dr. Felix Bast Professor and HOD

Department of Botany, Central University of Punjab felix.bast@cup.edu.in Link