

## ACADEMIC QUALIFICATION

**Master of Science in Data Science , University of Minnesota - Twin Cities , MN , USA | 2025-2027.**

- CSCI 5521 Machine Learning fundamentals
- CSCI 5707 Principles of Database Systems
- CSCI 5523 Introduction to Data Mining

**Bachelor of Engineering in Biotechnology, RV college of Engineering, Bengaluru, India | 2021–2025.**

**Minor in Artificial Intelligence , Indian Institute of Technology, Ropar | August 2024 - August 2025.**

## TECHNICAL SKILLS

- **Programming:** Python, C, C++, R.
- **Developer Tools:** GitHub, Google Colab, Jupyter Notebook, VS Code, Google Cloud Platform , Gradio, Streamlit, Kaggle.
- **Data Analysis:** MySQL, Excel, Power BI , Tableau.
- **Libraries:** NumPy, Pandas, Matplotlib, Seaborn, Scikit-learn, TensorFlow, PyTorch, Hugging Face Transformers, OpenCV, PyPDF2, FAISS, SentenceTransformers.
- **Soft Skills:** Teamwork, Time Management, Communication, Problem Solving.

## WORK EXPERIENCE

**Octakaigon Bock Pvt Ltd.(BOCK AI), Bengaluru,India**

December 2024 – June 2025

### AI and Automation Intern

- Collected and processed domain-specific medical datasets from clinical textbooks and public health sources, using fitz (PyMuPDF) for PDF parsing and Python's json module to convert raw content into structured JSON format suitable for language model fine-tuning.
- Fine-tuned the lightweight transformer model DistilGPT-2 on the medical dataset to generate accurate diagnostic text based on patient symptoms.
- Built and deployed a scalable LLM fine-tuning pipeline on Google Cloud's Vertex AI to support continuous adaptation on medical datasets.
- Developed a prototype text-to-video generation pipeline using transformer-based prompt encoding and diffusion-based video synthesis to convert clinical instructions into medically relevant visual sequences.

**Center of Excellence for Computational Genomics, RVCE ,Bengaluru,India**

November - December 2022

### Project Intern

- Implemented an equivariant diffusion model (DDPM), a generative machine learning framework that uses noise scheduling and reverse diffusion to create molecular structures, for generating de novo drug compounds with optimal chemical properties.
- Applied conditional and inpainting models, which incorporate specific molecular constraints and fill in missing atom groups, to refine ligand-pocket interactions and improve docking accuracy.
- Performed ADMET (Absorption, Distribution, Metabolism, Excretion and Toxicity) property predictions using tools like SwissADME and protox-II, ensuring the drug-like and safety profiles of compounds.

## ACADEMIC PROJECTS

### A machine learning-based diagnosis tool for Parkinson's disease.

Developed a multimodal diagnostic system combining structured patient metadata (e.g., age, comorbidities, symptom onset) and spiral test images for early-stage Parkinson's detection.


- Implemented a Retrieval-Augmented Generation (RAG) pipeline using fine-tuned OpenChat 3.5 and domain-specific transformers like PubMedBERT and BioClinicalBERT to generate clinician-style diagnostic responses.
- Applied prompt engineering techniques to improve clinical relevance and specificity in the generated diagnostic outputs.
- Integrated a ResNet18 convolutional neural network to analyze spiral drawings and correlate motor impairment patterns .
- Prototyped the full system using Gradio to support real-time clinical inputs and generate interpretive text-based summaries from combined image and metadata sources.

### A Web-Based Platform for Storing and Accessing Genomic, Diagnostic, Proteomic, and MRI Data of Alzheimer's Disease.

Developed a Flask-based web application to centralize genomic, diagnostic, proteomic, and MRI data related to Alzheimer's disease.

- Created RESTful APIs with JSON responses and used Python's OS module for structured data retrieval and file management.
- Built a responsive and user-friendly frontend using HTML, Bootstrap, and JavaScript to enable smooth navigation and real-time data access for researchers and clinicians.

## PUBLICATIONS

- De novo Design of Anti-COVID Drugs Using Machine Learning-Based Equivariant Diffusion Model Targeting the Spike Protein (Published: MDPI, 12 May 2023) 
- Impact of Nitrogen Deficiency on Bacopa monnieri: A Computational and Experimental Study (Published: Metszet, Volume 9, Issue 9, September 2024). 