Biratal Raj Wagle

biratal@hotmail.com • +01 6033498536 • GitHub • LinkedIn• Website

EDUCATION

Dartmouth College, Hanover, NH — June 2025
MS Health Data Science — GPA 4.00
Related Coursework*: Biostatistics I & II, Algorithms in Biomedical Data Science, Principles of Machine Learning.*

Ashoka University, Sonipat, India — June 2023
BS Physics & Computer Science — GPA 3.68
Related Coursework: *Algorithms, Data Structures, Mathematical Physics I II & III, Statistical Mechanics*

EXPERIENCE

GenAI Engineer Intern — January 2025 - Present
River Records — Boston, MA
Developed a quality assurance pipeline for River Records, a startup out of Boston Children's Hospital, focused on developing a medical scribe assistant. This position was part of the "Data Analytics Project Lab" course at Dartmouth.

- **Software Development**: Built a robust QA pipeline using the **Azure OpenAI Service API**, reducing hallucinations from the prompting algorithm and improving the reliability of the medical scribe assistant. Automated data validation processes to catch transcription errors and ensure high-quality, HIPAA-compliant data outputs.

Data Scientist — December 2024 - Present
Geisel School of Medicine — Hanover, NH
Developed an automated data extraction pipeline in **Python** for computer posturography. Created machine learning algorithms to instantly diagnose vestibular migraines' causes.

- **Clinical Research:** Conducted data analytics on sensitive clinical data, adhering to all related policies. Utilized **pytorch**, **pandas**, and other statistical packages to create an algorithm in a self-supervised machine learning task to successfully predict outcomes.

GenAI Graduate Student Intern — December 2024 - Present
Dartmouth College — Hanover, NH
Developed a pipeline using **RouteLLM**, to optimize model selection for Dartmouth's LLM chat service to reduce the emissions and cost of every prompt. Developed starter guides (JupyterBooks) for the Dartmouth AI research community for Dartmouth's implementation of comprehensive ***LangChain***: a framework used to build AI tools.

- **Author:** Wrote an introductory guide on vector semantics, detailing advanced **NLP** related methods such as TF-IDF and PPMI. Employed dimensionality reduction techniques to visualize embeddings of queries and documents to represent their similarity.
- **Machine Learning:** Compiled an overview of **RAG** models, outlining their functionalities and implementation strategies. Supplied users with a template for training custom RAG models. (Python)

Graduate Research Assistant — August 2024 - Present
Dartmouth College — Hanover, NH
Engaged in a **research** lab focused on wearable devices, emphasizing **data exploration** and **machine learning** methodologies. Focused on other areas of digital health. Published a paper in ACM

- **Exploratory Data Analysis**: Conducted visual and basic inferential statistics on an existing dataset, and to discover that increased sleep variability is associated with worse blood glucose management. (Python) An associated qualitative study was accepted for presentation at an international conference.

- **Deep Learning:** Implemented two models (Stacked **LSTM** and WaveNet) that increased the generalizability of blood glucose prediction for Type 1 Diabetes patients, compared to traditional methods. **Under peer review.**

Research Associate September 2022 – May 2024
MitraLab Sonipat, HR
Engaged in quantitative biomedical research focusing on mitochondrial studies for oncology applications.

- **Tool development**: Led the development of MitoSinComp, a custom pipeline for analyzing structure-function relationships in mitochondria, enabling automated data processing and advanced foci detection using **machine learning**. (Python, C).
- **Data Analytics**: Conducted comprehensive analysis of mitochondrial data employing ImageJ, Python, and ParaView, creating visualizations and output files that facilitated deeper insights into the association between cancer and the structure of mitochondria and its implication for energy production. **Under peer review.**

Volunteer Researcher June 2018 – September 2018
European Council for Nuclear Research (CERN) Geneva, Switzerland

- **Computational Physics:** Did a short-term volunteer internship at CERN with Dr. Suyog Shrestha, focusing on analyzing data from hadron collisions in the Large Hadron Collider. Discovered a bug in the CERN programming language (**ROOT**), which was reported and updated in future versions.
- **Programming Languages:** Learned to apply Python and C++ code to apply Multi-Layer Perceptron Neural Networking to data and produce visualizations. Became acquainted with the intersection between computation and the natural sciences. (Python, C++)

TECHNICAL SKILLS AND INTERESTS

Python, SQL, MATLAB, R, Tableau, SAS, Kotlin, Pandas, NumPy, OpenCV, Vim, LangChain, Machine Learning, Natural Language Processing (NLP), Statistical Analysis, Statistical Modelling, Linux/Unix, Data Analysis, Microsoft Office, Power BI, Pytorch, ROOT, Tensorflow, scikit-learn

PUBLICATIONS

- Yanjun Cui, **Biratal Raj Wagle**, Shriti Raj, Enzo Plaitano, Catherine Stanger, and Temiloluwa Prioleau. 2025. *Empowering Self-Management of Diabetes through Long-Term Wearable Data and Seasonal Visualizations*. In Proceedings of the Extended Abstracts of the CHI Conference on Human Factors in Computing Systems (CHI EA '25). Association for Computing Machinery, New York, NY, USA, Article 218, 1–8. https://doi.org/10.1145/3706599.3719780
- Agarawala, S. Saini, M. **Wagle, B.** Spurlock S. Golchha, B. Parker, D. Mitra, K. *Quantitative analyses of mitochondrial structure-function reveals an early role of 'Small Mitochondrial networks' (SMNs) in carcinogenesis.* Submitted December 2024 to PNAS. https://doi.org/10.1101/2024.12.26.630414
- Lu, B. **Wagle, B.** Cui, Y. Prioleau, T. *Investigating the Reproducibility and Generalizability of Deep Learning Methods for Blood Glucose Prediction.* **In Progress**