

ANOOSHKA BAJAJ

✉ anooshkabajaj@gmail.com
in [linkedin.com/in/anooshkabajaj](https://www.linkedin.com/in/anooshkabajaj)

☎ (812) 822-7876
🔗 github.com/ab490

📍 Bloomington, IN, United States
🌐 ab490.github.io

EDUCATION

Master of Science (M.S.) in Data Science

Indiana University Bloomington

August 2024 – May 2026

- **Research Assistant (R.A.):** Working in Dr. Zoran Tiganj's lab, focusing on in-context learning mechanisms in Transformers. Investigating LLMs interpretability and correlating findings with human brain language understanding.
- **Courses:** Introduction to Statistics, Applied Machine Learning, Applied Database Technologies.

Bachelor of Technology (B.Tech.) in Bio-Engineering with Minor in Computer Science Engineering

Indian Institute of Technology (IIT) Mandi, India

August 2019 – May 2023

- **CGPA:** 8.50 / 10.00
- **Relevant Courses:** Data Science I (Data Analysis), Data Science II (Probability and Statistics), Data Science III (Machine Learning), Deep Learning and Applications, Applied Biostatistics, Data Structures and Algorithms.

EXPERIENCE

Engineer - Radisys India Ltd.

Bangalore, Karnataka, India

July 2023 – June 2024

(Intern : July 2023 – November 2023)

- Improved the **5G** Network infrastructure by modifying the **OAM** code base in **C++** and **Python**.
- Handled requests across multiple modules, including Configuration Management (**CM**) (added YANG model configurations per 3GPP specifications) and Performance Management (**PM**) (handled performance counters).
- Resolved client requests related to **CU** (Centralized Unit) and collaborated using **BitBucket**.

Data Engineer Intern – Technocolabs Softwares Inc.

Indore, Madhya Pradesh, India (Remote)

July 2022– Sep 2022

- Pre-processed dataset of over 1 million rows in **Python**, performing Exploratory Data Analysis and Feature Engineering
- Created a data warehouse with **BigQuery**. Managed ETL workflows and generated DAGs using **Cloud Composer** built on **Apache Airflow**. Analyzed the data by generating interactive dashboard on Google Data Studio.

Research Intern – IISER Bhopal (Dr. Sunando Datta)

Bhopal, Madhya Pradesh, India

June 2022– July 2022

- Assisted PhD students in analyzing 2,212 YAP target genes in the Hippo Signaling Pathway to inhibit tumor growth.
- Produced the dataset by **web scraping** using Selenium, and performed **data visualization** of target genes belonging to 28 gene groups.
- Deployed a **Python** web app for cell migration analysis from time-series coordinate files using **Flask**.

PROJECTS

Skin Cancer Classification App | Deep Learning, CNN, Flask

- Developed a **web application** to predict and classify skin cancer types from dermoscopic images uploaded by users, utilizing **Convolutional Neural Networks**.
- Trained the Deep Learning model on the HAM10000 dataset to classify skin cancer into one of 7 categories.
- Enhanced model accuracy by evaluating over different architectures and implementing three regularization techniques, resulting in a final deep learning model integrated into an efficient **Flask**-based web application.

Predictive Analytics | Machine Learning, Random Forest

- Created a predictive ML model using **Random Forest** to assess diabetes risk based on features such as age, smoking history, BMI, HbA1c level, and blood glucose levels.
- Performed data cleaning, label encoding, feature engineering, and hyperparameter tuning, employing cross-validation to optimize and evaluate model performance on medical and demographic data.

Audio Summarizer App | GenAI, NLP, LLM, Streamlit

- Developed an audio summarization **web application** using OpenAI's **Whisper** model for transcription and Hugging Face's **Transformers** (t5-small) for generating concise summaries.
- Built a **Streamlit**-based interface with PyTorch integration for efficient processing of WAV/MP3 audio files up to 200MB.

TECHNICAL SKILLS

Programming Languages: C++, Python, R, MATLAB

Development Tools & IDEs: Jupyter Notebook, Visual Studio Code, Spyder, RStudio

Databases & Data Management: SQL, MySQL, PostgreSQL, Tableau, Snowflake

Frameworks & Libraries: Git, Tensorflow, PyTorch, Flask, Streamlit

Certifications: [Intermediate Machine Learning](#), [Feature Engineering](#): Kaggle. [MATLAB Onramp](#): Mathworks.