

UTKARSH MATHUR

(425) 633-5220 • utkarsh.mathur@gmail.com • [LinkedIn](#) • [GitHub](#) • [Portfolio Website](#) • Greater Seattle Area, WA

- Data Science and Software Engineering professional with 2+ years of experience in developing end-to-end solutions with Front-end and Back-end systems, Machine Learning & Deep Learning models, LLMs, Data Pipelines (ETL/ELT), and Data Analysis.
- Proficient in Python, C++, JavaScript (MERN stack), SQL, & cloud technologies, with expertise in AI, CI/CD, & Web Development.

EDUCATION

University at Buffalo, State University of New York (UB) | New York, USA January 2023 – June 2024

Master of Science in Data Science | *Capstone Project*: “Graph Neural Networks and Large Language Models: A Literature Review”

Coursework: Design & Analysis of Algorithm, Statistical Learning, Data Model Query Language, Computer Vision, Machine Learning

Indian Institute of Technology Roorkee (IIT Roorkee) | Uttarakhand, India

July 2018 – July 2022

Bachelor of Technology in Polymer Science (Chemical Engineering) | *Capstone Project*: “Production of Sustainable Aviation Fuel”

Coursework: Artificial Neural Networks, Mathematics (Calculus, Algebra, Statistics & Numerical Methods), Data Structure & Algorithms

EXPERIENCE

Data Scientist | Atriano | Remote, CA September 2024 – Present

- Developing educational story generation app (*EdTech*) for kids using *Large Language Models (LLMs)* and *recommendation systems*.

Data Engineer, Imaging | ImagoAI | Milpitas, CA

October 2024 – January 2025

- Deployed Galaxy (mycotoxin test – *Food Quality Assessment*) for 10 clients by gathering data and fine-tuning ML models (*Python*).
- Automated (*Python* & *Shell*) the process of Hyperspectral Image camera calibration *increasing the process efficiency by 80%*, *reducing the manual efforts by 95%*, and *saving \$30000*.

Data Scientist | Quinbay | Bengaluru, India

May 2022 – October 2022

- Optimized order return process by deploying *Object Detection model* reducing the *approval time from 6 hours to 15 seconds*.
- Developed a pipeline to flag counterfeit products by leveraging *Computer Vision*, *Natural Language Processing*, and *Data Analysis* for image recognition and product screening. This feature *reduced manual efforts by 30%* and *accelerated the process by 70%*.
- Spearheaded deployments (*E-commerce*) using *Python*, *PyTorch*, *Node*, *React*, & *SQL* and analytics with *GCP*, *Tableau*, & *PowerBI*.

Machine Learning Engineer | Hono | Gurugram, India

July 2021 – April 2022

- Deployed (*Django* & *JavaScript*) Predictive Analytics pipelines to automate production of client specific services (*HR Tech*) using ETL/ELT pipelines (*Python*, *AWS*, *SQL*), Machine Learning (*Python*), & A/B Testing (*VWO*), *generating \$120000+ revenue in 2022*.
- Developed Employee Attrition Forecasting, Promotion Suggestion, Hike Estimation, and Leave Prediction features for *12 clients*.

Data Scientist- Intern | ImagoAI | Gurugram, India

April 2021 – May 2021

- Designed mycotoxin tests (*Food Quality Assessment*) for *3 products* (Corn, DDGS & Wheat) using ML Algorithms (*Python* & *AWS*).

Research Intern | Department of Physics, IIT Roorkee | Roorkee, India

September 2020 – June 2021

- Experimented on Deep Learning models (*TensorFlow*) for semantic segmentation to estimate ocular tumor growth and trends.
- Developed *MATLAB* software for image data annotation of tumor-infected rat eyes *accelerating the process by 85 - 90%*.

SKILLS

Programming Languages: Python, C, C++, SQL, JavaScript, Java, C#, HTML, CSS, R, MATLAB, Perl

Data Science Skills: PyTorch, TensorFlow, Keras, Tableau, PowerBI, PySpark, Databricks, Qlik, Snowflake, Hadoop

Software Skills: Node, React, MongoDB, Django, Kafka, Jenkins, Selenium, Git, Kubernetes, Docker, AWS, GCP, Jira, Confluence

PROJECTS

Metaheuristic Optimization vs Backpropagation | Computer Vision, University at Buffalo

January 2024 – May 2024

- Developed (*PyTorch*) Metaheuristic optimizers using Particle Swarm Optimizer and Grey Wolf Optimizer for Convolutional Neural Networks training and contrasted their performance with backpropagation (Stochastic Gradient Descent and Adam Optimizer).

Top Spotify Tracks Database | Data Models Query Language, University at Buffalo

August 2023 – December 2023

- Built a web application (*Flask app*) to showcase an intuitive dashboard for “Top 50” Spotify playlists across 67 countries.
- Engineered a pipeline (*Python*) to aggregate real-time data from Spotify Web API in a normalized relational database (*SQL*).