

Atharva Dumbre

☎ (862)381-9846 | ✉ atharva.dumbre1@gmail.com | [in /in/atharvadumbre](https://in.linkedin.com/in/atharvadumbre) | github.com/atharvadumbre | atharvadumbre.github.io

Education

New Jersey Institute of Technology | **M.S. in Computer Science** | GPA: 3.78 / 4.00

Sep2022 - May 2024

University of Mumbai | **B.E. in Computer Engineering** | GPA: 7.88 / 10.00

Aug 2018 - May 2022

Skills

Languages & Databases: Python, SQL, Java, JavaScript, HTML, CSS, MySQL, Redis, MongoDB, FAISS Vector DB

Tools & Frameworks: Excel, Power BI, Tableau, Git, VS Code, Jupyter Notebook, React, Express.js, Node.js, Flask, OpenAI, Scikit-learn, TensorFlow, Keras, Pytorch, LangChain, Hugging Face, Pandas, NumPy, Matplotlib, Seaborn, XGBoost, NLTK, FastAPI, Streamlit, CUDA

Cloud Technologies: AWS, Azure, GCP, Docker, Apache Airflow, AWS Services (S3, Glue, Lambda, Athena)

Management Skills: Communication, Teamwork, Problem-solving, Analytical reasoning, Attention to Detail, Time management, Detail-oriented

Other: Machine Learning, Deep Learning, Data visualization, Data Science, Regression, Classification, Clustering, NLP (Natural Language Processing), LLM (Large Language Models), RAG (Retrieval Augmented Generation), Data Analytics, Time-series analysis, Testing, Debugging

Work Experience

NJIT - Ying Wu College of Computing

Newark, NJ

Research Assistant – Data Scientist

Jan 2024 - May 2024

- Created a reinforcement learning algorithm with Python & the PyTorch framework for a traffic control system that led to a 9% reduction in the average waiting time of vehicles at 5 simulated intersections.
- Guided a research team to read through 20+ research papers, prepared documentation standards for them, & set standard of comparison.
- Strove the team to analyze the data using Power BI & communicated the outcomes at regular meetings, which expanded their insights, as well as pointing the project in the right direction.
- Achieved 99% accuracy in annotating over 300 images of simulated traffic intersections on Roboflow, in collaboration with other researchers, improving the dataset quality for reinforcement learning models.
- Utilized supervisor feedback to increase system readiness by 67% by installing SUMO & CARLA on 10 systems, ensuring optimal functionality.

AICTE - EduSkills Foundation

Remote

AI-ML AWS Intern

Oct 2021 - Dec 2021

- Administered 75% of common student queries autonomously using an NLP chatbot implemented with AWS Lex and AWS Lambda, integrating the chatbot into WhatsApp using the Twilio programmable messaging API.
- Slashed student inquiry response time by 80% via integration with university support services, measured by average response time metrics.
- Grew student satisfaction with support services by 30% as measured by post-interaction survey scores, by improving chatbot interface.

Projects

AI Logo Generation Platform (creative-dalle.com) – MERN (Mongo, Express.js, React, Node.js) | OpenAI | Hostinger | Cloudinary

- Launched a custom logo generation platform utilizing OpenAI's DALL-E 3 model to generate logos with prompts tuned to industry standard designs, featuring "Surprise Me," social sharing, & keyword search; ensured 99.9% uptime with Render.com & Hostinger.com cloud hosting & served images to users via the Cloudinary platform. [Link](#)

PDF Chat with RAG System – Python | Streamlit | LangChain | HuggingFace | FAISS

- Developed a chatbot using Streamlit to convert PDFs into a searchable knowledge base with the FAISS vector database, LangChain, & Google's T5 LLM with average query retrieval in 4 seconds. [Link](#)

YouTube Trending Videos Data Engineering & Analysis – AWS | Power BI | SQL | DAX

- Implemented an ETL pipeline using AWS (S3, Glue, Lambda, Athena, CloudWatch) to process & analyze over 70,000 YouTube Trending Videos records for 3 countries (USA, Canada, Great Britain), converting JSON to Parquet & created a Power BI dashboard with 5 KPIs & trends. [Link](#)

Fintech - Revenue Analysis: Hospitality Domain – Power BI | SQL | DAX

- Designed an intuitive Power BI dashboard for the Codebasics Resume Challenge, using DAX & SQL queries to preprocess hospitality data & create 26 key performance indicators (KPIs). [Link](#)

Fintech - Company Bankruptcy Predictor – Python | Pandas | Numpy | Matplotlib | Sklearn | Pickle

- Built a predictive ML model to predict the bankruptcy of a company using Python for data cleaning, analysis, & visualization. Attained 96.18% accuracy with the help of the Random Forest Classifier algorithm & presented the findings in the form of a presentation. [Link](#)

Research Publication

- Classification Of Indian Sign Language Characters Utilizing Convolutional Neural Networks & Transfer Learning Models With Different Image Processing Techniques (IEEE) | [Link](#) Electronic ISBN:978-1-6654-7988-2

Certificates & Achievements

- Oracle Generative AI Professional | AWS Cloud Practitioner | TensorFlow Developer Certificate | LLaMa for Developers (LinkedIn) | Neural Networks & Deep Learning (Coursera) | HTML, CSS, & Javascript for Web Developers (Coursera) | Building Web Applications in PHP (Coursera) | Fundamentals of Deep Learning for Computer Vision (NVIDIA) | Introduction to AI in the Data Center (NVIDIA) | CCNA: Introduction to Networks | Cisco Cybersecurity Essentials | **Foundations of Machine Learning & Applications Bootcamp** IIT Mandi
- Secured 1st place for "Most Creative Use of Redis Cloud" at HackNJIT Oct 2023 | **Kaggle Expert (3647 Highest Rank)**.