

Atharva Deshmukh

ad2546@rit.edu | 5859106910 | New York, United States | [linkedin.com/in/atharva-deshmukh/](https://www.linkedin.com/in/atharva-deshmukh/) | ad2546.github.io/atharva_portfolio/ | github.com/ad2546

Skills

- **Data Engineering & ML:** Data Warehousing (Star/Snowflake), Apache Airflow, Kafka, PySpark, XGBoost, Keras, TensorFlow, PyTorch, Deep Learning, LLMs, Generative AI (Ollama, Hugging Face), RAG, FAISS, LangChain
- **Programming & Scripting:** Python, SQL, R, Java, JavaScript, C, Git, Bash, Pandas, NumPy, Scikit-learn, Flask, Scipy, NoSQL
- **Big Data, ETL & Cloud:** PostgreSQL, MongoDB, Hadoop, Spark, Kafka, Snowflake, SAS, Kettle, BigQuery, dbt, CI/CD (GitLab, Azure DevOps), AWS (SageMaker, IAM, S3, Glue), Google Cloud Platform (GCP), Databricks, Docker
- **Visualization & Analytics:** Tableau, Power BI, Hypothesis Testing, Regression, Time Series, Predictive Modeling, ANOVA, Outlier Detection, Data Validation, Data aggregation, Performance Tuning

Work Experience

Graduate Teaching Assistant

Rochester Institute of Technology

Aug 2024 - May 2025

Rochester, NY

- Delivered 15 interactive lectures on Ethics in Computing, incorporating student-led case studies on data privacy, bias in AI, and ensuring accessibility; achieved a 4.8/5 average student rating.
- Advised 10+ students weekly through one-on-one consultations, providing personalized feedback on project development, resulting in improved project outcomes and student satisfaction, as indicated by course evaluations.
- Facilitated hands-on workshops and collaborative dialogue sessions on emerging trends, empowering students to troubleshoot data challenges and explore innovative applications in generative AI.

Dark Horse Digital Solutions

Software Consultant (Open Text Cordys)

Aug 2022 - Jul 2023

Mumbai, India

- Spearheaded the migration of an on-premise OpenText Cordys BPM system to AWS by re-integrating external services and deploying Docker containers, reducing operational overhead by 25%.
- Designed and maintained a centralized data warehouse in Oracle SQL, for insurance data by collaborating with stakeholders and implementing automated ETL processes with Apache Airflow, enhancing data accessibility for analytics by 30%.
- Collaborated with non-technical stakeholders and agile teams to automate data extraction and reporting through complex SQL queries and containerized stored procedures, extracting over 100,000 records daily and bolstering data-driven decision-making.

Dark Horse Digital Solutions

Software Consultant Jr (IBM BAW)

Jun 2021 – Jul 2022

Mumbai, India

- Led cross-functional teams to resolve database performance bottlenecks in an IBM BAW environment for a health insurance platform by utilizing Docker containers and optimizing SQL queries, ensuring 80% of claims processed in under 10 seconds.
- Engineered a scalable ETL pipeline using Apache Airflow which integrated into IBM BAW, processing over 10,000 insurance claims daily, and improving claims processing times by 20% within the quarter.
- Developed a scalable data warehouse to centralize insurance data by designing robust ETL pipelines and containerized deployments, improving reporting efficiency and analytics by 30%.

VESIT Renaissance

Product Development Intern

May 2020 – Jul 2020

Mumbai, India

- Developed a performance tracking application for new college students by collecting stakeholder requirements and academic data, leveraging Python and PostgreSQL to build an interactive backend and dashboard that improved academic feedback by 30%.
- Engineered an automated ETL pipeline post data cleansing, validation and preparation using Apache Spark to consolidate diverse student performance data, processing over 10,000 records daily and boosting data accuracy by 20%.
- Orchestrated containerized deployments with Docker to standardize development environments, reducing configuration time by 50% and accelerating feature rollouts.

Education

Rochester Institute of Technology (GPA: 3.7)

Master of Science in Information Technology and Analytics

May 2025

Rochester, NY

- Relevant Coursework: Visual Analytics, Data Science and Analytics (ML), Non-Relational Database Management, Information Retrieval, Data Mining, Information Assurance Fundamentals, Data Warehousing and Data Lakes, Natural Language Processing.

University of Mumbai (GPA: 3.0)

Bachelor of Engineering in Computer Engineering

May 2021

Mumbai, India

- Relevant Coursework: Artificial Intelligence, Machine Learning, Analysis of Algorithms, Data Structures, Statistics, Data Modelling, Pentaho, Software Engineering (including Agile development practices).