

PATHIK RUPWATE

pathikrds@gmail.com | 9321194480 | Mumbai, India
pathikrupwate.com | LinkedIn | GitHub

PROFESSIONAL SUMMARY

Data Scientist and Machine Learning Engineer with over 4 years of experience building and deploying end-to-end predictive models, computer vision pipelines, and NLP solutions. Proven track record of leveraging advanced statistical techniques and modern AI frameworks (TensorFlow, PyTorch, LangChain, RAG) to drive measurable business outcomes, from improving campaign ROI by 30% to increasing forecasting accuracy by 25%. Adept at bridging the gap between complex data science architecture and strategic business KPIs.

CORE COMPETENCIES & SKILLS

Machine Learning:

Classification, Regression, Clustering, Time-Series Forecasting, Natural Language Processing, Computer Vision, Causal Inference, Anomaly Detection.

Generative AI:

Large Language Models, Retrieval-Augmented Generation (RAG), Agentic Workflows (LangChain), API Integration.

Programming:

Python (PyTorch, scikit-learn, TensorFlow, Pandas, Keras), R (ggplot2, caret), SQL, C++, Java, HTML, CSS.

Infrastructure:

Docker, Flask, GCP, AWS, Linux, GitHub Copilot, MongoDB.

Analytics & Viz:

A/B Testing, Multivariate Testing, Tableau, Power BI.

PROFESSIONAL EXPERIENCE

Data Scientist / Machine Learning Engineer

Mar 2026 – Present

ABM Knowledgeware

Mumbai, India

- **Municipal AI Agent (Mainet Upgrade):** Spearheaded the development of a municipal AI agent utilizing the Model Context Protocol (MCP), Retrieval-Augmented Generation (RAG), and complex Data APIs; integrated this agent as a major feature upgrade to the flagship Mainet software, significantly improving data accessibility for government officials.
- **Automated Localization Pipeline:** Directed the end-to-end architecture of an automated, multi-lingual tutorial video generation pipeline using Python and the ElevenLabs API, completely eliminating manual voiceover work and saving over 80 hours of production time per release cycle.
- **Large-Scale Data Restructuring:** Engineered a sophisticated data parsing and recovery solution to salvage 400,000+ unstructured files recovered from an AWS incident; developed pattern-matching algorithms to accurately restore the original directory structure and metadata for 80% of the disorganized dataset, preventing massive critical data loss.
- **AI Vulnerability Tester:** Designed and implemented an AI-driven vulnerability testing suite by integrating OWASP ZAP with Gemini AI, enabling automated discovery of security flaws alongside intelligent, autonomous generation of remediation code snippets to accelerate patch deployment.
- **Odisha Government Agent (Ongoing):** Currently leading an initiative to build and deploy a specialized, highly scalable AI agent for the Odisha Government, designed to modernize regional administrative tasks and automate workflows for millions of residents.

Data Scientist

Oct 2023 – Jan 2026

Kearny Bank

Austin, TX

- Improved campaign ROI by 30% through advanced customer segmentation (K-means, hierarchical clustering, DBSCAN, PCA, Naive Bayes, Random Forest), directly optimizing ad targeting and performance.
- Increased forecast accuracy by 25% for loan default predictions by implementing isolation forests and one-class SVM architectures.
- Performed sophisticated NLP on unstructured user feedback using TF-IDF and LDA to extract actionable insights and model user sentiment at scale.
- Conducted rigorous A/B and multivariate testing to assess marketing strategies and product optimizations; collaborated closely with Product Managers to align data initiatives with business KPIs.
- Designed and maintained dynamic dashboards in Tableau and Power BI, utilizing heatmaps, time-series charts, and decision trees to visualize key metrics for executive stakeholders.

Data Science Researcher

Jan 2023 – Jun 2023

University of Chicago

Chicago, IL

- Built and deployed a real-time BMI prediction pipeline utilizing a VGG-16 CNN architecture, served via a Flask API processing live video input.
- Optimized facial recognition models for complex video analysis using TensorFlow and OpenCV, reducing latency and enhancing model performance through iterative experimentation.

Data Scientist

Jan 2022 – Dec 2022

Aiolux

Chicago, IL

- Predicted S&P 500 sector performance across three distinct time horizons (50, 100, and 200 days) using Generative Adversarial Networks (GANs), SMOTE, and an Extremely Randomized Tree Classifier, achieving an average predictive accuracy of 80%.
- Generated sophisticated synthetic data to address severe class imbalance, successfully increasing overall model robustness by 20%.
- Identified macroeconomic drivers, such as CPI, utilizing causal inference techniques, and applied findings to predictive ad strategy simulations.

Database Engineer

Aug 2020 – May 2021

Beloit College Powerhouse

Beloit, WI

- Designed, architected, and implemented a robust database and front-end website to track 5,000+ inventory items using SQL, Python, and HTML.

Software Engineer

Jan 2020 – Dec 2020

Open Energy Dashboard

Beloit, WI

- Built interactive diagrams and implemented multilingual architecture for an open-source energy dashboard utilizing Docker, Ubuntu, and React.

PROJECTS

AI-Powered Portfolio Chatbot

2026

- Architecting and developing an interactive, AI-driven chatbot integrated into a personal portfolio website.
- Utilizing modern LLM APIs, Retrieval-Augmented Generation (RAG), and LangChain for agentic workflows to dynamically retrieve and answer questions regarding professional experience, skills, and projects.

EDUCATION

University of Chicago

Jun 2023

Master of Science, Applied Data Science

Chicago, IL

Beloit College

May 2021

Bachelor of Science, Computer Science & Quantitative Economics

Beloit, WI

- **Leadership:** President, Beloit Student Government – Managed and allocated a \$250,000 annual funding budget for student initiatives and campus organizations.