

RUTIKA AVINASH KADAM

Stony Brook, NY 11790 · +1 934-949-8653 · rutikakadam1005@gmail.com · LinkedIn · Portfolio · GitHub

EDUCATION

Stony Brook University

Master of Science in Data Science

Coursework: Statistics, Data Analysis, Machine Learning, Deep Learning, Natural Language Processing, Big Data Analysis.

Stony Brook, USA

August 2024 – May 2026

Savitribai Phule Pune University

Bachelor of Engineering in Information Technology

Coursework: Discrete Structures, Data Structures & Algorithms, Distribution Systems, Cloud Computing.

Pune, India

August 2016 – May 2020

SKILLS

Programming Languages: Python, R, SQL (MySQL, PostgreSQL, PL/SQL), MongoDB, C, C++, HTML, CSS, JavaScript

Data Science & Machine Learning Libraries: NumPy, Pandas, Matplotlib, Seaborn, scikit-learn, TensorFlow, Keras,

PyTorch, Hugging Face

Data Engineering & Big Data Tools: Azure Data Factory, Azure Data Lake, Databricks, SQL Server Management Studio

MLOps & Deployment Tools: Flask, Streamlit, Gradio, GitHub, Docker, Hugging Face Hub, Hugging Face Inference API

Tools & Platforms: Visual Studio Code, Jupyter Notebook, SAS, RStudio, Postman, Azure, AWS, Power BI, Tableau, Microsoft

Excel, ServiceNow, Jira, Microsoft Endpoint Configuration Manager, Qualys, Aternity

PROFESSIONAL EXPERIENCE

Stony Brook Medicine

Research Assistant

Stony Brook, USA

July 2025 – Present

- Built missing data imputation pipelines in R & SAS for a cohort of women aged 65+ in the Study of Osteoporotic Fractures.
- Developed & validated machine learning models (logistic regression, random forest, boosting algorithms) to predict physical function decline & fracture risk; applied cross-validation & hyperparameter optimization across multiple imputed datasets to ensure robust & reproducible results.

Tata Consultancy Services

System Analyst

Pune, India

August 2020 – April 2024

- Collaborated with the Vulnerability Management team to perform Risk Analysis on vulnerability datasets from Qualys VMDR across 55K+ Windows assets & 1M+ vulnerabilities, uncovering trends, anomalies, & threat vectors using Python & MySQL.
- Implemented predictive models for vulnerability prioritization & patch management timelines, leveraging engineered features like CVSS-weighted risk scores & patch urgency indices, leading to a 15% reduction in security risks.
- Developed supervised ML models (logistic regression, boosting algorithms, ANN) to predict the likelihood of exploitability using CVE metadata, asset attributes, & historical remediation data, improving prioritization efficiency by 25%.
- Employed Data Analysis Expressions (DAX)-enhanced KPIs like CVSS score, threat intelligence, exploitability, risk levels, remediation timelines, deployment status, compliance rates within Power BI driven Vulnerability Analysis Dashboard.
- Designed & deployed feasible technical solutions using MS Endpoint Configuration Manager to remediate Windows & application vulnerabilities with 99% compliance; optimized configurations & automated tasks using PowerShell, boosting productivity by 25%.

Zensar Technologies

Data Engineer Intern

Pune, India

May 2019 – July 2019

- Built ETL pipelines using Azure Data Factory to process 30GB of transactional & web traffic data stored in Azure Data Lake Gen2; leveraged this data to perform Funnel Analysis for India's leading e-commerce food ordering platform, Swiggy, identifying key drop-off points & reducing cart abandonment by 11%.
- Built interactive Power BI dashboards to visualize conversion rates, traffic sources & user journey patterns; collaborated with marketing to optimize campaigns, increasing high-intent traffic by 15%.
- Tracked pipeline enhancements & analysis results in JIRA, ensuring reproducibility & alignment with business KPIs.

PROJECT EXPERIENCE

AskYourDocument | Retrieval-Augmented Generation(RAG), Natural Language Processing

- Designed RAG application combining FAISS vector search with Google Generative AI, enabling intelligent question answering over ingested documents & web content.
- Implemented an ingestion pipeline for PDFs, DOCX, TXT, & URLs with SBERT-based semantic chunking, integrated into a FastAPI backend & Streamlit frontend to deliver semantic search & LLM-powered contextual responses.

SmartApply | Python, Gradio, Google Gemini API, Hugging Face Spaces

- Built SmartApply, a Gradio-based web app using Google Gemini API and PyPDF2 to extract resume text, analyze against job descriptions, and generate ATS match %, profile summaries, missing keywords, and skill improvement suggestions.
- Deployed the application on Hugging Face Spaces for public access, integrating prompt engineering and modular ML workflows to deliver real-time resume analysis with structured & an interactive UI.

ScoreCast: Exam Performance Prediction | Machine Learning, Python, Numpy, Pandas, Scikit-learn, HTML, CSS, Flask

- Designed a Flask web app with interactive UI allowing users to input data & predict math scores, building preprocessing pipelines & training multiple ML models including Linear Regression, bagging & boosting algorithms.
- Optimized models using GridSearchCV, achieving 86.8% accuracy (R^2), & ensured robust evaluation through cross-validation for reliable predictions.