Contact

+1 (857) 398-5107 (Mobile)
n.c.chaudhari10@gmail.com

www.linkedin.com/in/niraj-
chaudhari-76934a173 (LinkedIn)
ncchaudhari10.github.io/ (Portfolio)

Top Skills

MySQL
Stored Procedures
Data Analysis

Certifications

IBM AI Engineering Specialization
Intermediate SQL for Data Scientists
MTA: Introduction to Programming
Using Python - Certified 2018
The Data Science of Experimental
Design
Machine Learning with Python

Honors-Awards

Winner - Smart India Hackathon
2020
Best Team Award
On the Spot (Team) Award
Special Achievement Award
Star of the Month Award

Publications

Deep learning–based downscaling of
summer monsoon rainfall data over
Indian region

Niraj Chaudhari

MS in Data Science at Northeastern University, Boston
Boston, Massachusetts, United States

Summary

I'm a Data Scientist with a strong interest in Generative AI and its
real-world applications, from finance to protein design. Currently
pursuing my Master's in Data Science at Northeastern University,
I bring close to four years of industry and research experience,
including roles at MFS Investment Management, Tata Consultancy
Services, and academic research labs.

My current work focuses on generative models for protein design,
using diffusion models, graph neural networks, and tools like
AlphaFold and RFdiffusion. I enjoy translating research into practical
solutions, whether it's building machine learning systems from
scratch or designing workflows that solve complex problems.

Outside of work, you'll find me exploring trails, playing guitar, or
diving into new cultures and ideas.

I'm actively looking for:
→ Full-time opportunities starting Jan 2026
in AI/ML Research, Generative AI, or Data Science roles.

Let's connect! I'm always up for a great conversation about AI,
science, or ideas worth building.

_____

Experience

Institute for Experiential AI at Northeastern University
AI Research Scientist
April 2025 - Present (4 months)
Boston, Massachusetts, United States

Khoury College of Computer Sciences
Research Apprenticeship
January 2025 - April 2025 (4 months)
Boston, Massachusetts, United States

• Developed generative AI models for enzyme design, leveraging diffusion and flow matching to generate protein structures up to 300 amino acids in length.
• Built sequence-to-structure and structure-to-sequence pipelines using RFdiffusion, AlphaFold, ProteinMPNN, and ProtT5, processing large-scale 3D protein data for structural engineering.
• Applied Graph Neural Networks (GNNs) and 3D diffusion models to design functional enzymes with desired structural properties.

MFS Investment Management
Data Scientist
June 2024 - December 2024 (7 months)
Boston, Massachusetts, United States

• Deployed Streamlit application combining NLP techniques with OpenAI API to rate the quality of meeting notes in real-time, defining 15 metrics to improve note-taking and boosting documentation efficiency by 20%.
• Conducted web scraping with BeautifulSoup to gather financial advisor data and utilized 70B LLAMA model to summarize insights.
• Built interactive Tableau dashboards to visualize data from STAR schema (over 80+ tables), facilitating business-driven storytelling.
• Implemented Informer Transformer model to predict 25-day leading redemption rates for mutual funds, achieving r^2 score of 0.58.

Tata Consultancy Services
Full Stack Data Scientist
August 2020 - August 2023 (3 years 1 month)
Thane, Maharashtra, India

• Led development and implementation of Single Sign-On (SSO) across 100+ products, enhancing security and performance
• Contributed to the development of the interactive passwordless authentication mechanisms using Spring framework to authorize and authenticate half a million TCS employees
• Worked alongside system architects to re-design applications with improved security and performance
• Utilized Redis Pub/Sub, Websockets, Caching, and JWT to improve the security and availability of products, resulting in a 25% reduction in latency
• Tested products in alpha and beta testing modes, achieved 89% code coverage, and ensured service feasibility using Selenium and JMeter
• Deployed services on VMSS Azure cloud and configured ELK for telemetry and troubleshooting, accomplishing 99.97% availability

- Integrated advanced SQL queries into ETL pipelines to collect data from different sources and created visualizations using Power BI to make data-driven decisions, leading to a 30% reduction in overall infrastructure cost
- Increased click-through rate by 15% through A/B testing of features and onboarded an analytical portal for seamless data exports
- Constructed a predictive model using LSTM to forecast CPU consumption of servers, configured thresholds, and performance metrics to monitor drift and enable proactive scaling of virtual machines, preventing 10 incidents per month
- Built a multi-class classifier to handle user tickets, cutting down resolution time by 15%
- Provided technical support and troubleshooting during incident calls to stakeholders
- Mentored new associates on project details and knowledge transfer sessions to build secure enterprise applications

Indian Institute of Tropical Meteorology
Data Scientist
June 2019 - April 2020 (11 months)
Pune, Maharashtra, India

- Developed deep learning-based downscaling methods, enhancing rainfall map resolution by 4x ($0.125°\times0.125°$ from $1°\times1°$) using deep Convolutional Neural Networks
- Utilized SRCNN and DeepSD models, trained on 17 years of Indian meteorological data on Pratush GPU supercomputer
- Validated model on 3 different data, attained a correlation of 0.86 with validation data, emerging deepSD as superior downscaling method, outperforming traditional interpolation approaches
- Improved model architecture by incorporating advanced activation functions and kernel initializers, reducing training time on Pratyush GPU supercomputer by 60%
- Developed a custom method to determine optimal patch sizes based on data resolution and conducted hyperparameter optimization, decreasing model loss by 20%
- Proficient in various libraries including Keras, PyTorch, Scikit-learn (sklearn), NumPy, pandas, X-array, netCDF4, Basemap, Matplotlib, and Seaborn

Enfini Labs Tech Consultant Pvt. Ltd.
Software Engineer Internship
December 2018 - January 2019 (2 months)

• Blueprinted and drove the construction of a security layer for existing applications, incorporating JSON Web Token and Auth0.
• Deployed applications on the Microsoft Azure Cloud platform and configured metrics for 100% availability of web applications.

Education

Northeastern University
Master of Science - MS, Data Science · (September 2023 - May 2025)

Ramrao Adik Institute of Technology
Bachelor of Engineering - BE, Computer Engineering · (August 2016 - October 2020)