# Avdhoot Patil

+1 (631) 406-0675 | avdhootdattatr.patil@stonybrook.edu | linkedin.com/in/avdhoot-patil | github.com/avdhoot0303

## Education

**Stony Brook University** — Stony Brook, NY
*Master of Science in Data Science* — *Aug 2023 – May 2025*
*Relevant Coursework*: Data analysis, Probability, Data Science fundamentals, Data management, Statistical Computing, Big data systems, Big data analysis, Deep Learning, Machine Learning, Smart energy in the Information Age.

**Xavier Institute of Engineering, University of Mumbai** — Mumbai, India
*Bachelor of Engineering in Information Technology* — *Jul 2018 – Jun 2021*

## Skills

**Languages and Frameworks**: Python, R, Java, GO, Javascript, Tensorflow, SQL, Terraform, Shell Scripting, scikit-learn (for Python), caret (for R), Flask, Django, pytest, unittest, Jenkins, Travis CI, GitLab CI, Scrum, Kanban, Apache Hadoop, Apache Spark, NLTK, SpaCy

**Tools and Platforms**: GCP AWS, Azure, Sagemaker, Kubernetes, Docker, Jira, GitHub, GitLab, Elastic Search, Kibana, Power BI, Excel, Prometheus, Grafana, REST, Redis, Pub-Sub, Nginx, MySQL, MongoDB, PostGreSQL, Linux

**Certification**: Google Cloud Platform Certified Associate Cloud Engineer [Credential link (GCP ACE)]

## Experience

**Software Engineer** | *NeoSoft Technologies, Mumbai, India* — Jun 2021 - Jun 2023

Project 1: Internal Talent Acquisition Analytics platform

- Architected and developed a scalable Talent Acquisition Analytics Platform using Docker and Kubernetes. This reduced the time-to-fill for open positions by 25%, significantly enhancing recruitment efficiency and agility.
- Leveraged microservices for analytics using Python and Elasticsearch, improving candidate quality by 20%, boosting retention rates.

Project 2: LabCorp Annotation Product

- Engineered a **model training and deployment framework in Python**, enabling seamless model training, deployment, and live servicing on **AWS Lambda and API Gateway**, contributing to **30% faster model deployment**
- Led a codebase refactoring, yielding a 20% speed boost and 30% complexity reduction. Introduced configuration files, resulting in 25% fewer bugs and streamlined development.

Project 3: Monibag

- Managed migration to **GCP**, integrating Java Microservices with **GCP Cloud Run**. Hosted microservices using **Cloud Storage**, **API Gateway**, and **load balancers**, improving performance and reducing deployment time by **30%**.
- Implemented infrastructure standardization with **Terraform**, **reducing** provisioning time by 40% and simplifying **microservice scalability**.

**Software Engineering intern** | *Tech Mahindra, Mumbai, India* — Jan 2021 - Apr 2021

- Played a key role in software development project, encompassing the design and implementation of backend system utilizing Java, SQL, PL/SQL, Hibernate, MVC, Spring Boot, and Microservices. Additionally, spearheaded the creation of a comprehensive database for a banking application through meticulous PL/SQL scripting.

## Mentorship

- **Led training programs** for new team members at Neosoft, providing in-depth **guidance on GCP, DevOps practices, Git**, and various workflows. This streamlined training approach contributed to a 30% reduction in onboarding time, fostering a more agile and efficient team.
- Presented engaging guest lectures on **DevOps methodologies** to an audience of 70+ students at Xavier Institute of Engineering. Provided personalized guidance for final year projects, resulting in a 25% increase in project quality.

## Academic Projects

**Malware Detection of Portable Executable files**

- Researched, curated, and evaluated datasets for malware detection, selecting the most relevant sources.
- Restructured and **fine-tuned a high-performing CNN model** using **Tensorflow and Keras**, achieving an impressive **95% accuracy** in **multi-class malware detection**. Mitigated class imbalances through strategic class weighting, yielding a 20% improvement in multi-class detection of the malwares and with 30% higher accuracy.

**Exploring Credit Fraud Detection Strategies using R**

- Led a project addressing challenges of unbalanced datasets, particularly in fraud credit card transactions.
- Implemented effective oversampling techniques in R, significantly improving model performance. Achieved a top score of 0.977 with XGBOOST, while Random Forest and Logistic Regression also performed well.