

Nikhil Ghugare

Seattle, WA | +1 (206) 535-9528 | nikhil.ghugare20702@gmail.com | [LinkedIn](#) | [GitHub](#) | [Portfolio](#)

EDUCATION

Master of Science in Data Science, Seattle University GPA - 3.5/4

Relevant Coursework: Statistical Machine Learning, Big Data Analytics, Data Visualization & Business Intelligence

June 2026

Seattle, WA

Bachelor's in Computer Engineering, Savitribai Phule Pune University GPA- 3.4/4

Relevant Coursework: Data Science, Machine Learning, Data Structures, Operating Systems, Computer Networks, OOPS

May 2024

Pune, India

EXPERIENCE

Data Science Intern

Houston, Texas

Motherson Group – Chairman's Office, Americas

Jun 2025 – Aug 2025

- Spearheaded the development and automation of executive Power BI dashboards for the Chairman and senior leadership, integrating **7 years of historical HR and salary data from 5 countries and 80+ corporate entities**.
Result: Enabled real-time tracking of critical workforce metrics and provided instant, high-level insights for strategic oversight.
- **Revolutionized data reporting efficiency** by designing and building SQL pipelines that fully automated the data entry process for all **80+ corporate entities** into the company's internal software.
Impact: Eliminated the previous, error-prone manual process of Excel sharing, ensuring data integrity and a single source of truth.
- Designed and delivered interactive visualizations to analyze **10+ critical trends**, including **manpower, turnover, absenteeism, legal cases, salaries, and negotiations**, directly supporting the Chairman's Office team in strategic planning and high-stakes executive decision-making.
- Conducted rigorous data cleansing and transformation on multi-source datasets spanning seven years, utilizing advanced Excel to resolve numerous data inconsistencies and errors. Reported directly to the **HR Director and General Manager** while partnering with **HR teams** to standardize KPIs and maintain data confidentiality.

Data Science Intern

Mumbai, India

Prodigy InfoTech

Jan 2024 – Feb 2024

- Executed comprehensive **Exploratory Data Analysis (EDA)** on large, complex datasets using **Python (pandas, NumPy, matplotlib)**, successfully identifying critical trends and extracting actionable business insights to inform predictive modeling strategies.
- Engineered and deployed robust predictive machine learning models for business forecasting and decision support, applying algorithms like **Linear Regression and Random Forest** that resulted in improved prediction accuracy and model performance.
- Managed end-to-end data lifecycle by utilizing **SQL, MySQL, and Oracle** for efficient data extraction, schema design, and query optimization, which directly supported ERP database development (documented with ER diagrams and flowcharts).
- Partnered with cross-functional teams to translate complex analytical findings into **data-driven solutions**, delivering enhanced workflow efficiency and reporting clarity through the development of interactive visualizations and reports.

Web Development Intern

Remote

Oasis InfoByte

Feb 2023 – Mar 2023

- Developed and deployed interactive web interfaces by leveraging a comprehensive tech stack, including **HTML, CSS, and JavaScript**, while also utilizing **WordPress** for efficient content management and site building.
- Contributed directly to **UI/UX** design by creating compelling visual mock-ups and high-fidelity prototypes using **Figma**, effectively translating design concepts into intuitive, user-centric pages that enhanced engagement.
- Successfully executed foundational web development initiatives, strengthening practical skills in front-end development principles and ensuring cross-browser compatibility.
- Utilized a versatile skill set to create intuitive user pages prioritizing engagement, deepening understanding of creating seamless digital experiences.

PROJECTS & RESEARCH

Online Transaction Fraud Detection | Python, Scikit-learn, XGBoost, Pandas, NumPy, Matplotlib, MySQL

Project

- Built and fine-tuned a fraud detection model using XGBoost and Random Forest, enhancing prediction accuracy.
- Conducted EDA and feature engineering on large financial datasets to optimize model performance and reduce false positives.

Youth Drug Use Risk Modeling | Python, TensorFlow, Keras, NumPy, Librosa, CNN

Github

- Processed 55K+ behavioral health records using feature engineering, encoding, and SMOTE to improve model balance.
- Trained and optimized Random Forest, XGBoost, and SVM models, achieving high ROC-AUC and recall for early risk detection.

Bird Audio Classification Using Deep Learning | Python, TensorFlow, Keras, NumPy, Librosa, CNN

Github

- Converted bird call MP3 recordings into Mel spectrograms and trained a custom CNN for binary and multiclass species classification.
- Applied data augmentation and dropout regularization, achieving over 90% precision in noisy real-world audio environments.

TECHNICAL SKILLS

- **Programming & Analytics:** Python, R, SQL, Java, C++, Excel (Pivot Tables, VLOOKUP, Data Cleaning)
- **Machine Learning & AI:** Scikit-learn, TensorFlow, PyTorch, XGBoost, Keras, CNNs, SVM, NLP (BERT, SpaCy, VADER)
- **Visualization & Geospatial Tools:** Power BI, Tableau, ArcGIS Pro, Matplotlib, Seaborn, Plotly, Google Charts
- **Data Engineering & Cloud:** AWS (EC2, S3, Lambda, EKS), MySQL, ETL Pipelines, GCP, Azure, Hadoop, PySpark, Hive, Docker, MongoDB, Data Warehousing
- **GenAI & Professional Skills:** RAG Pipelines, LangChain, GenAI Tools (Gemini, Claude, Cursor), Leadership, Problem-Solving, Communication