

Siddharth Hiraou

sidhiraou@gmail.com | 7167509361 | <https://www.linkedin.com/in/siddharth-hiraou> | Buffalo, NY

EDUCATION

Masters Science in Data Science | University at Buffalo, State University of New York

2024 – 2026

Bachelor's in Information Technology | International Institute of Information Technology, Pune

2020 – 2024

PROFESSIONAL EXPERIENCE

Data Engineering Intern | Hanker Analytics | California, US

May 2023 - July 2023

- Developed a **Data Quality Management (DQM) model** in Python, leading to a **36.9%** improvement in analysis report accuracy.
- Created a **Power BI dashboard** to showcase client applications, resulting in a **30%** increase in client engagement.
- Implemented an **ML model** to predict causes of disease, with an accuracy of **86.1%**.
- Designed **statistical visualization charts** to track the company's economic progress, enhancing decision-making processes by **25%**.

Data Analyst Intern | InternLabs | Pune, India

December 2022 - May 2023

- Developed a **Real Estate Data Analysis application**, integrating data from multiple informal sources including Facebook, Telegram channels and Whatsapp communities.
- Consolidated data from diverse sources into a unified **data frame structure**, processing over **12,000** real estate listings in Pune.
- Developed an automated data pipeline that removed manual data entry, saving approximately **20** hours of work per week.

Research Assistant | IsquareIT | Pune, India

June 2022 - May 2023

- Developed a real-time poultry monitoring system for Venky's Chicken farm using AI/ML algorithms, achieving model accuracy of **93.1%**.
- Utilized Python, OpenCV, and CNN-based pattern recognition to monitor flock behavior, analyzing **over 10,000 images** for health trend identification and risk mitigation.
- Went to IPMT(Institute of poultry Management and Technology) for on ground research pattern for chicken herd pattern monitoring.
- Implemented an **attendance monitoring system** for farm workers, improving workforce accountability and operational efficiency.
- Co-authored the following **research paper published** in the *International Journal of Technological Exploration and Learning (IJTE)*

Software Intern (Python) | Orgpedia | Pune, India

February 2023 - April 2023

- Worked on debugging python code for document analysis model of company, Increased accuracy of the model by **22.6 percent**.
- Collaborated with the data science team to identify and resolve critical issues in the existing codebase and **mentored 2** juniors regarding future work.

LANGUAGES, SKILLS AND CERTIFICATES

- Languages:** Python, Java, Javascript, C/C++, R, Java, SQL.
- Databases:** MySQL, MongoDB.
- Framework/Libraries:** Plotly, Seaborn, Power BI, Numpy, Pandas, Matplotlib, Tensorflow, Keras, Pytorch, BeautifulSoup.
- Software and Tools:** Git, Advanced Excel, Tableau, Power BI, Matlab, AWS, GCP, Azure DevOps.
- Certifications:** Google Data Analytics Professional Certification, Google Cloud Practitioner.

PROJECTS

Smart Poultry Farm Management with AI and Cloud Integration | Python, OpenCV, Roboflow, CNN, Git

- Developed a real-time poultry monitoring system using AI/ML algorithms for early detection of diseases to enhance productivity and reduce losses for a renowned farm called Venky's Chicken.
- Utilized Python, OpenCV and CNN integrated group pattern recognition to monitor flock behavior and health trends.
- Implemented temperature control mechanisms to optimize farm conditions and also added attendance monitoring system for workers.

Predictive Analytics for Credit Risk Evaluation | Python, Scikit-learn, XGBoost, SHAP/ LIME, PyTorch, MySQL

- Developed a predictive model using Python and Scikit-learn to evaluate the likelihood of loan defaults based on customer demographics, financial data, and credit history.
- Built and evaluated various machine learning models including Logistic Regression, Random Forest, and XGBoost.
- Optimized model performance through hyperparameter tuning using GridSearchCV, resulting in improved accuracy and predictive power.
- Deployed the final model using Flask API for real-time predictions on PyTorch.