

# SIDDHI VINAYAK DASH

DOB: 01 April 2005    Bhubaneswar, Odisha

☎ (+91) 7684931821    ✉ [siddhudash21@gmail.com](mailto:siddhudash21@gmail.com)    [in siddhivdash](#)    [siddhivdash](#)

## Education

### SRM Institute of Science and Technology

*Bachelor of Technology in Computer Science and Engineering (CGPA - 8.08)*

September 2021 – Present

Chennai, Tamil Nadu

### Vikash Residential School, Bargarh

*AISCCE – CBSE, Science (PCM) (Percentage - 81.2%)*

May 2023

Bargarh, Odisha

### Vikash Residential School, Bargarh

*AISSE – CBSE (Percentage - 92%)*

May 2021

Bargarh, Odisha

## Relevant Coursework

- Machine Learning
- Deep Learning
- Artificial Intelligence
- Probability and Statistics
- Linear Algebra and Calculus
- Data Structures and Algorithms
- Design and Analysis of Algorithms

## Work Experience

### Mahanadi Coalfields Limited (MCL)

*Data Science Intern [Certificate]*

June 2025 – July 2025

Sambalpur, Odisha

- Built an intelligent HR chatbot **MCLBuddy** using **NLP** techniques to automate employee query handling.
- Trained models with **ANN**, **RNN**, **Bi-LSTM**, and **transformers** for accurate, context-aware responses.
- Created and cleaned a domain-specific dataset from internal HR documents and FAQs.
- Used **FAISS** for fast semantic search and applied **text preprocessing** and **embedding** techniques.

## Projects

### Diamond Price Prediction

*Python, Scikit-learn, Flask, Data Pipelines*

June 2025

*GitHub — Deployed Application*

- Developed an end-to-end **regression model** to predict diamond prices using attributes like **carat**, **cut**, **color**, **clarity**, and **dimensions**.
- Performed **EDA** and **feature engineering** using libraries such as **Pandas**, **NumPy**, **Seaborn**, and **Matplotlib**.
- Trained and evaluated multiple models including **Linear Regression**, **Ridge**, **Lasso**, **ElasticNet**, and **Random Forest**.
- Used evaluation metrics like **RMSE**, **MAE**, and **R<sup>2</sup> Score** to identify the best-performing model.
- Deployed the selected model using a **Flask API** for real-time prediction through a web interface.

### Diabetes Prediction

*Python, Scikit-learn, Pandas*

May 2025

*GitHub*

- Built and compared models like **Logistic Regression**, **KNN**, **SVM**, **Decision Tree**, and **Random Forest** on the Pima Indians Diabetes Dataset.
- Performed **EDA**, **data preprocessing**, and **feature scaling** using **Pandas** and **NumPy**.
- Evaluated models using **accuracy**, **F1-score**, with **Random Forest** achieving the best performance.

## Technical Skills

- **Languages:** Python, SQL, C
- **ML/DL Libraries:** Scikit-learn, TensorFlow, Keras, XGBoost
- **Data Analysis/Visualization:** NumPy, Pandas, Matplotlib, Seaborn
- **Concepts:** Regression, Classification, Clustering, CNN, RNN, Model Evaluation
- **Tools:** Jupyter Notebook, Google Colab, Git, GitHub, Anaconda
- **Deployment:** Flask, REST APIs
- **Databases:** MySQL, MongoDB

## Extracurricular Activities

- Contributed to **Aaruush**, a national-level technical fest, as a graphic designer, creating visually appealing posters and promotional materials using **Adobe Photoshop**.
- Earned the **LeetCode 50 Days Badge 2025** for solving problems consistently for over 50 days, showcasing strong commitment to algorithmic problem solving.