

SHOBHIT KUMAR SINGH

+91-9580727944

shobhit22iit@gmail.com

linkedin.com/in/shobhit-singhh

github.com/Shobhit-Singhh

EDUCATION

- | | | |
|---|---------------------|-----------|
| • Indian Institute Of Technology Guwahati, <i>M.Tech in Medical Devices & Diagnostics</i> | GPA: 9.44/10 | May 2024 |
| • Kamla Nehru Institute Of Technology, <i>B.Tech in Mechanical Engineering</i> U.P., India | GPA: 7.75/10 | Aug 2021 |
| • Divine Public School, <i>Mathematics + Computer Science</i> Gorakhpur (U.P.), India | GPA: 7.53/10 | July 2017 |

EXPERIENCE

- **Senior Data Scientist**, (*Primary Healthtech Private Limited* | Delhi, India) (March'24 - Present)
 - Developed **Python and C++ scripts** to streamline data extraction and analysis processes, **resulting in a 25% reduction** in device testing time. Utilized **hypothesis testing** to **eliminate redundant testing levels 2 & 4**, while **automating procedures for levels form 1 to 7** using **pySerial** and **numpy**, thereby enhancing overall efficiency of manufacturing process.
 - Developed a **manufacturing dashboard with Tableau** for monitoring multiple stages of the manufacturing process and automated manual data analysis using **Scikit-learn, Streamlit and Tkinter**, **reducing processing time by 90 mins/batch**.
- **Data Science Intern**, (*Primary Healthtech Private Limited* | Delhi, India) (May'23 - Feb'24)
 - Utilized clinical trial data for process optimisation and **improve accuracy by 30% and precision by 25%** of the diagnostic device and optimising reaction parameters **reduces the device's testing time (incubation + analysing time) by 50%**.
 - **Steps: Implemented data validation, data cleansing, and data transformation, Cross Validation and statistical analysis.**
 - Processed raw clinical trial data using **scikit-learn and Pandas**, ensuring accuracy and integrity for thorough analysis. Organized and prepared the data to **uncover deeper insights in PowerBI** and improve diagnostic device performance.

DATA SCIENCE PROJECTS

- **Advance Data Visualisation And Exploration Platform**, (*All-In-One Data Science App*) **Deployed, GitHub**
 - Created a **No Code platform** focused on exploratory data analysis and hyper-parameter tuning and model training.
 - **Drag and Drop Tool kit** that simplifies model training and makes machine learning accessible to a broader audience.
 - **Tech Stack:** Python, Numpy, Pandas, Matplotlib, Plotly, Scikit-learn, Pygwalker, Machine learning, Sklearn, Streamlit.
- **Automated Inventory Control And Churn prediction**, (*AI-Based Inventory Management Tool*) **GitHub**
 - Implemented an **inventory control system, reducing stockouts possibility by 80%** and improved pipeline management.
 - Utilized data trends for **inventory forecasting and churn prediction**, enhancing decision-making and resource allocation.
 - Designed a dashboard to visualize **inventory metrics and churn probabilities**, make data-driven decisions in real-time.
 - **Tech Stack:** Python, Numpy, Pandas, Matplotlib, Plotly, Scikit-learn, Pygwalker, Machine learning, Sklearn, Streamlit.

AI/ML PROJECTS

- **AI-based Question Generation**, (*AI Assistant for EdTech*) **GitHub**
 - Developed an NLP model for generating **MCQ, True/False, Fill in the Blanks, Match the Following** questions.
 - **Reducing time by 70-80%** in question framing, allowing teachers to provide valuable feedback to students.
 - **Tech Stack:** Python, Numpy, Matplotlib, Plotly, Scikit-learn, NLTK, T5 transformer, Word2vec, Sent2vec, SpaCy, BERT.
- **Web Scraping for Keyword Extraction and Sentiment Analysis**, (*Feedback Evaluation Tool*) **Github**
 - Created a web scraper to **extract insights from web-based textual data** for tracking feedback and market trends.
 - **Performed sentiment analysis**, empowering businesses to optimize their marketing strategies and resource allocation.
 - **Tech Stack:** Python, Numpy, Pandas, Matplotlib, Plotly, Scikit-learn, Beautiful Soup, NLTK, SpaCy, Google sheets.
- **Querying SQL Database text-to-SQL Query model**, (*SQL Assistant*) **GitHub**
 - Developed a user-friendly platform that allows easy uploading and interaction with databases in a common language.
 - Enhanced database query interpretation by employing **Prompt Engineering** and **LLM fine-tuning** on queries dataset.
 - **Tech Stack:** Python, Numpy, Pandas, Plotly, Scikit-learn, Hugging Face's Transformer, TensorFlow, NLTK, SpaCy.

TECHNICAL SKILLS

- | | |
|--|---|
| • Programming languages: C/C++, Python, MATLAB*. | • Deployment: Streamlit, Tkinter, FastAPIs, LangChain. |
| • Development: ML, ANN, CNN, RNN, NLP, LLM. | • Databases: SQL, MongoDB. |
| • Libraries: Pandas, Numpy, TensorFlow, Sklearn, SpaCy. | • MLOps: DVC, MLflow, MLOps, Docker, Kubernetes. |
| • Dashboards: Tableau, Power BI, Excel Dashboard. | • Miscellaneous: MacOS, Windows, LaTeX*, Ubuntu. |

COURSES AND CERTIFICATIONS

- **IITG course:** Data structure and algorithm, Database management system, SQL, Mathematical Modelling & Simulation, Image Processing with Machine learning, Supervised and Unsupervised Machine Learning, Python Programming languages.
- **Certification: Stanford online:** Advanced Learning Algorithms, Supervised Machine Learning, **Google:** Foundations of Data Science. Google Advanced Data Analytics Certification, Regression Analysis: Simplify Complex Data Relationships.

POSITION OF RESPONSIBILITY AND EXTRA-CURRICULAR

- **DPR at IITG:** Department Placement Representative for the 2023-2024 batch at the Indian Institute of Technology Guwahati.
- **Coding Milestone:** Solved more than **500 DSA coding questions**, demonstrating proficiency in my problem-solving skills.
- **LeetCode:** Achieved a **rating of 1429 on LeetCode** in weekly contests and have developed a proficiency in problem-solving.
- **Codeforces Pupil:** Attained the title of "**Pupil**" on **Codeforces**, demonstrating a commendable rating ranges from 1200-1399.
- **Intrapersonal Skills and Interests:** Strength Training, Sketching, Swimming, High-Intensity Interval Training and Gym.
- **Sharing through lens:** Won the **first prize in the 'Shutter UP'** photography contest for my creative visual storytelling.