

# ABUTALHA SHAIKH

Email: abutalha11503@gmail.com | Phone no: +91-9766065167

GitHub: [github.com/38832](https://github.com/38832)

## Objective

Data Science professional with expertise in machine learning, data analysis, and statistical modeling. Proficient in Python, C++, and SQL, with hands-on project experience and a passion for solving data-driven challenges.

## Education

JSPM's Rajarshi Shahu College of Engineering, Pune

2021 - 2025

• Branch – Computer Science | Final Year

(expected)

## Technical Skills and Interests

- **Programming Languages & Frameworks:** C++, Python, SQL, NumPy, Pandas, Scikit-learn
- **Concepts & Technologies:** Object-Oriented Programming (OOP), Statistics, Machine Learning, Deep Learning, Data Visualization, Natural Language Processing (NLP), Large Language Models (LLM), AWS
- **Development Tools:** Power BI, PyCharm, Jupyter Notebook, VSCode, Google Colab, Flask, MLFlow, DagsHub, DVC
- **Database:** MySQL
- **Version Control & Collaboration:** Git, GitHub
- **Soft Skills:** Leadership, Adaptability, Analytical Thinking, Team Collaboration, Problem-Solving, Quick Learning

## Projects

### Physiotopia – AI-Powered Rehabilitation Web App [\[Link\]](#)

Designed and developed a health-focused web application offering physiotherapy support, featuring a custom-built machine learning model for injury prediction with fine tuned LLM for detailed rehab plan.

- Created an intuitive and responsive interface for users to input symptoms and receive instant rehab suggestions.
- Integrated decision tree model achieving **96% accuracy** to predict injury categories and suggest possible care paths.
- Tech Stack: Flask, FastAPI, Scikit-learn, Pandas, Frontend Frameworks, Colab, LLM

### Fake Job Posting Prediction [\[Link\]](#)

Built and deployed a deep learning model within a Flask web app to identify fraudulent job postings based on job listing content.

- Pre-processed job listing data including Title, Description, Company Profile, and Benefits.
- Utilized LSTM model architecture to handle text data, achieving **98% accuracy** on test data.
- Applied SMOTE for handling class imbalance and deployed the model for real-time use.

### Student Performance Analysis [\[Link\]](#)

Developed a machine learning solution to evaluate and predict student performance based on academic behaviour and history.

- Collected and analyzed academic data including exam scores, attendance records, and study hours from a MySQL database.
- Built and trained a Linear Regression model achieving **88% accuracy** in predicting academic outcomes.
- Used DVC and MLFlow for experiment tracking and model versioning; integrated into a Flask-based dashboard.

## Certifications

Machine Learning using Python | freeCodeCamp.org [\[Link\]](#)

Data Visualization | Great Learning [\[Link\]](#)

## Responsibilities and Extracurricular Achievements

- **Public Relations Officer, Rotaract Club (2022 – 2024):** Increased social media engagement by 40% and event participation by 25%; developed effective communication strategies.
- **Event Coordinator, Entrepreneurship Club (2021 – 2023):** Organized startup pitch sessions for 200+ participants with a 10-member team.
- Participated in KPIT Sparkle Hackathon, secured 2nd Position in University-level volleyball, and organized annual Techfest sports events.