SURAJ KUMAR

Patna, Bihar

📞 +91-7617477447 ✉ surajksharma7@outlook.com 🔗 surajksharma7 ⃝ surajksharma7 ▪ surajksharma7 🌐 Portfolio

EDUCATION

University of Petroleum and Energy Studies **Aug 2024 – Present**
Masters of Computer Applications - CGPA - 7.5 *Uttarakhand, India*

Bhagwan Mahavir University **Aug 2021 – May 2024**
B.Sc-Data Science - CGPA - 8.74 *Surat, India*

Research Works

"Elevating Large Language Models' Comprehensibility with SHAP Values", in Revolutionizing Machine Learning with Large Language Models: Security, Scalability, and Sustainability (Book Chapter), Series: Advancements in Mist AI and AI-Driven Cloud Networking.

TECHNICAL SKILLS

Languages: C++, Python, MySQL
Model Development & Experimentation: Model Development & Deployment, Machine Learning, Deep Learning, GenerativeAI
Data Management & Preprocessing: Data Pipeline Management, Data Preprocessing, Feature Engineering, Web Scraping
Statistics & Probability: Hypothesis Testing, Probability Distributions, Sampling, Inferences
Libraries/Frameworks: Scikit-learn, PyTorch, Pandas, Numpy, Langchain, Matplotlib, Seaborn, Plotly, NLTK, XGBoost, Selenium

PROJECTS

Breast Cancer Diagnosis Predictor ⌐ **Mar 2025**

- Developed an interactive **Streamlit-based web app** to predict whether a breast mass is benign or malignant using cell measurement inputs.
- Processed and modeled data from the public **Breast Cancer Wisconsin (Diagnostic) Dataset**; implemented a full ML workflow for diagnosis prediction.
- Created a dynamic **radar chart** visualization for clinical feature inputs, enhancing interpretability of diagnostic factors.
- Serialized the trained ML model and scaler into reusable assets, seamlessly integrating backend inference with UI rendering.
- Styled the application with custom **CSS** for clean, intuitive UX, including color-coded diagnostic badges.
- Packaged the app with clear setup — including a 'requirements.txt' and usage instructions — for easy deployment and evaluation.

Laptop Recommendation System ⌐ **Feb 2025 – Mar 2025**

- Collected and curated **1000+ laptop specifications** from **Smartphone**, engineered features like processor, RAM, GPU, display, storage, and OS for model training.
- Implemented a **Content-Based Recommendation System** using **TF-IDF Vectorization** and **Cosine Similarity** to suggest top-3 laptops similar to user input.
- Developed an interactive **Streamlit web app** with input forms, prediction page, file upload, resource download, and contact page modules.
- Optimized preprocessing pipeline, cached similarity matrix, and improved inference time by **40%**.
- Deployed the system with modular code structure and dataset support, enabling easy scalability and maintainability.

COURSES

- **Data Science Mentorship Program (DSMP) - CampusX** **Aug 2024 – Aug 2025**
  Highlights: Machine Learning & Fundamentals of MLOps.

Achievements

- Secured **2nd Position** in College Chess Competition at Bhagwan Mahavir University.