SHUBHANSHU SINGH

 +91-9335208375 shubhanshusingh8081@gmail.com   LinkedIn   GitHub   GeeksforGeeks

EDUCATION

Noida Institute of Engineering and Technology   **2022 – 2026**
B.Tech in Computer Science Engineering (AIML), CGPA: 7.82   *Greater Noida, India*
Divine Sainik School   **2021 – 2022**
12th,PCM, Percentage: 75.2%   *Varanasi, India*
Divine Sainik School   **2019 – 2020**
10th, Percentage: 84.8%   *Varanasi, India*

Experience

Machine Learning Intern (Remote) —*(Edunet Foundation)*   Sep 2023 - Nov 2023

- Engineered a machine learning pipeline to diagnose diseases from patient symptoms, achieving a peak prediction accuracy of **95% with the Random Forest model**.
- Implemented and compared four algorithms (**Decision Tree, Random Forest, KNN, Naive Bayes**) for robust prediction.
- Engineered SQLite-based data pipeline to efficiently store and retrieve over 1,000 patient records and ML model outputs, enabling actionable data insights and real-time analytics.
- Orchestrated end-to-end system integration between Streamlit frontend and backend APIs (ML models + SQLite), ensuring 100% functional data flow and user interactivity across platforms.
- **Tech Stack:** Python, scikit-learn, Pandas, SQLite, Streamlit, Data Engineering, Predictive Analytics, Machine Learning.

Technical Skills

Languages:	Python, Java, SQL, HTML, CSS, JavaScript
Frameworks/Libraries:	Tkinter, Django, Streamlit, Scikit-learn, Pandas, NumPy, keras, Matplotlib, Seaborn
Tools:	VS Code, Google Colab, Jupyter Notebook, Git, GitHub
Concepts:	Artificial Intelligence, Machine Learning, Deep Learning, DSA, DBMS, Operating System

Projects

Gen-AI Assistant for Research Summarization (GitHub Link — Live App)

- Built an AI-powered summarization assistant leveraging Google Gemini Pro + Streamlit to automate extraction and NLP-driven summarization of academic PDFs/TXT files, reducing manual effort by 90%.
- Architected scalable backend microservices for file parsing and NLP-based text extraction, enabling robust API integration with Gemini Pro for real-time generative AI workflows.
- Develop a Challenge Mode that formulates 3 subjective questions and evaluates user input through contextual LLM scoring.
- Deployed the system on Streamlit Cloud using a stateless API backend, cutting local setup time to zero and resolving 100% of dependency issues across platforms.

Tech Stack: Python, Streamlit, Google Gemini Pro, API Integration, Natural Language Processing (NLP), Document Management.

Human Disease Prediction Using Symptoms (GitHub Link — Live App)

- Developed a predictive ML-based healthcare tool capable of diagnosing over 40 diseases using 130+ symptoms, achieving up to 85% accuracy.
- Evaluated performance across four classification models and selected the most reliable predictor, improving diagnostic trustworthiness.
- Designed and deployed a responsive, user-centric web UI using Streamlit, improving accessibility and patient interaction by 40%.
- Implemented persistent SQLite logging system to capture user inputs and prediction outputs, supporting traceable diagnostics and longitudinal data analysis.

Technologies: Python, Streamlit, Pandas, Scikit-learn, ML Algo(Decision Tree, Random Forest, K-NN, Naïve Bayes), Matplotlib, SQLite

Certifications & Achievements

- **AWS Academy Certified:** Cloud Foundations, Machine Learning Foundations
- **Cisco Certified:** Data Analytics Essentials ,Cyber Threat Management
- **300+ DSA Problems Solved on GeeksForGeeks.** (Profile Link)
- **Campus Ambassador at GeeksforGeeks** (Certificate)