SARTHAK SAHU

+91 8920686320 | sarthaksahu776@gmail.com
LinkedIn: https://www.linkedin.com/in/sarthak-sahu-9664922a8
GitHub: https://github.com/S1aERT

SUMMARY

Computer Science undergraduate with Data Science , Artificial Intelligence and Machine Learning as core domains with strong hands-on experience in applied analytics, predictive modeling, and deep learning. Proven ability to formulate data-driven solutions using machine learning, statistical reasoning, and experimentation across healthcare and business domains. Experienced in end-to-end data science workflows including data preprocessing, model evaluation, validation, and deployment. Seeking a Machine Learning Intern role to apply analytical and algorithmic skills to real-world product and customer problems at scale.

EDUCATION

Bachelor of Technology (B.Tech) in Computer Science Engineering
Symbiosis Institute of Technology, Pune
Aug 2022 – Aug 2026

Intermediate (Class XII): 88.2% – City International School, Pune
Matriculation (Class X): 95.6% – Delhi Public School, Ghaziabad

TECHNICAL SKILLS

Programming & Tools: Python, R, Java, C, SQL (MySQL), Git
Data Science & Analytics: Data preprocessing, Exploratory Data Analysis (EDA), feature engineering, supervised learning, deep learning, model evaluation, cross-validation, performance metrics
Machine Learning & AI: Regression, classification, CNNs, RNNs, transfer learning, hyperparameter tuning
Frameworks & Libraries: TensorFlow, Keras, PyTorch (basic), Scikit-learn, Flask
Data & Visualization: Pandas, NumPy, Matplotlib

Cloud: AWS Sagemaker, Azure ML studio

Large Language Models (LLMs) and Generative AI(GenAI)

PROJECT EXPERIENCE

Healthcare Data Science & Machine Learning

Alzheimer's Disease Classification using Brain MRI Scans
• Developed and evaluated deep learning models (ResNet50, VGG19, EfficientNetB3) to classify Alzheimer's disease stages from brain MRI images.
• Performed data preprocessing, normalization, and augmentation to improve model generalization.
• Compared multiple architectures using performance metrics to identify the most effective model for prediction:Analyzed model bias and class imbalance using precision-recall trade-offs
• Interpreted results and analyzed limitations of imaging data for clinical decision support use cases.
• Deployed the trained model using a Python Flask-based web application to demonstrate real-world applicability.

Alzheimer's Disease Detection using Retinal Images
• Built an AI-based diagnostic pipeline using TranNet OCT data and UNet-based image segmentation.
• Applied deep learning techniques to extract meaningful patterns from medical images.
• Evaluated segmentation quality and classification performance to assess model reliability in healthcare scenarios.

Eye Disease Classification System
• Implemented CNN-based classification models (ResNet50, VGG19, EfficientNetB3) for multi-class eye disease detection.
• Conducted comparative analysis across models to understand performance trade-offs.
• Integrated the system with a Flask frontend to enable end-user interaction and model inference.

Applied Data Science Projects

House Price Prediction – California Dataset
• Built a supervised regression model to predict housing prices using structured tabular data.
• Performed exploratory data analysis to identify influential features and data trends.
• Applied feature engineering and model evaluation techniques to improve prediction accuracy.

AI-Based Recipe Recommendation System
• Developed an RNN-based model to generate food recipes using sequential data.
• Analyzed textual data patterns to improve sequence generation quality and relevance.

ADDITIONAL TECHNICAL PROJECTS

Secure Blood Bank Management System
• Designed a RESTful application using Spring Boot, JWT authentication and JPA.

• Used SQL(MySQL) for designing normalized relational schemas and performing efficient CRUD operations in a Spring Boot–based healthcare management system.

• Implemented role-based access control and secure data handling.

Smart Zombie Shooter Game
• Developed a 3D game using Unity and C# demonstrating problem-solving, logic design, and performance optimization.

CERTIFICATIONS

• The Ultimate Job Ready Data Science Course

• Introduction to Generative AI- Google Cloud Completion Badge

• Introduction to Large Language Models -Google Cloud Completion Badge

• Machine Learning Operations (MLOps) for Generative AI -Google Cloud Completion Badge

---

INTERESTS

Data-driven problem solving, applied machine learning, technology for social and healthcare impact