SAHIL SAINI

+91 9653335663 | saini.sahil6766@gmail.com | Linkedin | Github |Thane, India

EDUCATION

B Tech in CSE   **CGPA: 8.06**

Mumbai University   Aug 2020 - May 2024

M Tech in AI/ML   **CGPA: 7.00**

Symbiosis Institute of Technology   Aug 2024 - May 2026

SKILLS

- **Programming:** Java, Python, SQL
- **Data Engineering:** ETL, EDA, Machine Learning, MySQL, PL/SQL, MongoDB, Data Visualise
- **Data Analyst:** Power BI, Matplotlib, Seaborn
- **Software Development:** OOP, Data Structures, Algorithms, Operating Systems
- **Technologies, Cloud & Tools:** JavaScript, React, Hibernate, JSP, Spring, Azure, PyTorch, TensorFlow, Machine Vision, GAN, Gen AI, NLP, NLTK, LLM, Git
- **Soft Skills:** Leadership, Ethical Problem-Solving, Communication, curiosity, adaptability

PROJECTS

1. Stock Performance & Insights Analysis(Power BI, SQL, Financial Analysis, Data Visualization)

Role: Data Visualization Specialist & Financial Analyst
Responsibilities:

- Developed an interactive Power BI dashboard to analyze stock performance of major companies like Walmart, Apple, and Amazon.
- Conducted financial and operational analysis, integrating free cash flow, net profit margins, CEO compensation, and workforce trends.
- Implemented historical data analysis and trend mapping to provide real-time insights for investors and analysts.
- Designed dynamic visualizations for easy comparison of stock growth indicators.

2. Retail Business Intelligence for Supermart (Power BI, DAX, SQL, ETL, Data Visualization)

Role: Data Analyst
Responsibilities:

- Designed and developed an interactive Power BI dashboard for Super Mart, enabling real-time sales forecasting and performance monitoring .
- Implemented ETL processes to ensure data accuracy and consistency.
- Utilized DAX functions to calculate key performance metrics, including average order-to-ship time, profit trends, and customer segmentation analysis.
- Integrated slicers and interactive visualizations to enhance user experience and decision-making.

3. Movie Recommendation Chatbot(Python, Azure AI, NLP, Machine Learning, Flask)

Role: Machine Learning Engineer & NLP Developer
Responsibilities:

- Developed an AI-powered chatbot for personalized movie recommendations based on actors, genres, and release years.
- Utilized Azure's semantic analysis API to improve recommendation accuracy by 85%, enhancing user experience.
- Designed and implemented collaborative filtering & content-based filtering models to refine recommendations.
- Conducted A/B testing to validate the chatbot's effectiveness, leading to a 30% increase in user engagement.

4. E-commerce Shopping Website(Java, Spring Boot, MySQL, React, JavaScript, Hibernate)

Role: Full-Stack Developer
Responsibilities:

- Engineered a scalable e-commerce platform with separate modules for customers and administrators.
- Managed a database of over 500 products, reducing inventory errors by 20% through optimized SQL queries.
- Developed a user-friendly shopping cart and checkout system, increasing customer retention by 15%.

- Optimized backend queries to improve page load speed by 30%, enhancing website performance.

5. Pill and Text Summarization (YOLO v5, PyTorch, Tesseract OCR, NLP, Deep Learning, Wikipedia API)

Role: Computer Vision & NLP Engineer

Responsibilities:

- Developed a computer vision system using YOLO v5 and PyTorch, achieving 90% accuracy in text recognition on pill packages.
- Integrated Tesseract OCR for real-time text extraction and summarization, improving information retrieval efficiency by 50%.
- Designed an automated summarization pipeline, reducing content length by 70% while maintaining key information.

6. Building a Music App for Composers (Python, Librosa, Random Forest, MLP, KMeans Clustering)
(Spotify Dataset, HuggingFace Text2Music)

Role: AI & Machine Learning Developer

Responsibilities:

- Designed an AI-driven music app that assists composers in song recommendations and musical note generation.
- Leveraged IRMAS and Spotify Million Song Dataset for feature extraction using Librosa and MFCC scaling.
- Developed an instrument classification model using SVC, Random Forest, and MLP, achieving a macro recall of 44.7%.
- Proposed a hybrid clustering algorithm for efficient temporal data handling in music composition.

7. Disease Diagnosis Prediction using ML & EDA (Python, Spacy, Pandas, Matplotlib, Seaborn, Scikit-learn, Logistic Regression, Random Forest, XGBoost)

Role: Data Scientist & Healthcare Analyst

Responsibilities:

- Developed a machine learning pipeline to predict the likelihood of various diseases based on patient medical attributes.
- Performed Exploratory Data Analysis (EDA) using Seaborn and Matplotlib to uncover trends, correlations, and anomalies in healthcare datasets.
- Cleaned and preprocessed data using Pandas, handling missing values and encoding categorical features for ML model readiness.
- Trained and compared multiple models, including Logistic Regression, Random Forest, and XGBoost, achieving high accuracy and precision for disease classification.
- Delivered insights to support data-driven decisions in preventive healthcare and early diagnosis.

CERTIFICATIONS

- **Completed Java with MySQL** Link   Jan 2023
- **Completed Microsoft Power BI** Link   Dec 2024

Achievements

- Gold Badge in Java & Silver Badge in MySQL &Python (HackerRank)

- **Participated in Code Cubicle 2.0 hackathon (Link)**

  1. Educate consumers on the hidden realities of product labeling, highlighting marketing tactics that emphasize benefits (e.g., "Source of Vitamin C") while downplaying less favorable aspects (e.g., high added sugar content).
  2. Empower consumers to make informed, healthier choices by offering a multilingual application that uncovers and explains these labeling discrepancies.