SHUBHAM MURTADAK
Data Scientist

+91-9322191338 | shubhammurtadak022@gmail.com | linkedin | github

ABOUT ME

Entry-level Data Scientist with proficiency in Python, Machine Learning ,Generative Ai, eager to make a meaningful impact through contribution. I'm passionate, energetic, and geeky individual whose desire to learn is endless.

EDUCATION

B.E. in Artificial Intelligence & Data Science **2021-2025**
PES Modern College of Engineering
CGPA: **9.09**

Senior Secondary (XII), Science **2020**
Shree Ganesh Jr College, HSC Board
Grades: **83.38%**

Secondary (X) **2018**
New English School Korhale, India
Grades: **89%**

SKILLS

Primary Skills:

- **Programming Languages:** Python, R, C++, SQL
- **Python/ML/CV Packages:** Keras, OpenCV, Pandas, Scikit-Learn, Numpy, Matplotlib, Seaborn, Plotly
- **Knowledge of Machine Learning**
- **Deep Learning:** Neural Network, ANN, CNN, DNN, Transfer Learning, Back Propagation, TensorFlow
- Text understanding, representation & classification techniques, Text clustering
- **Libraries:** BOW, TFIDF, word2vec, doc2vec, sent2vec, keyphrase extraction
- LSTM RNN, Transformers
- Generative AI, Large Language Models, LangChain Framework
- Statistics
- DSA, Problem Solving

Secondary Skills:

- **WEB TECHNOLOGIES:** HTML, CSS, JavaScript, Django, Flask, Streamlit
- **DATABASES:** MySQL, MongoDB
- **Visualization Tools:** PowerBI, Microsoft Excel
- **Web Scraping:** Scrapy, BeautifulSoup, Selenium
- **Cloud Platform Services:** AWS

EXPERIENCE

Data Science Intern May 2024 - present
DataNnovite Solutions LLP *Pune, Maharashtra (Onsite)*

- Working on Data Science technologies including Machine Learning, Deep Learning, Large Language Models, Generative AI.
- Developed a solution for analyzing and processing clients' transaction and campaign data to feed into a RAG Architecture, enabling the retrieval of answers to specific business questions via a dashboard.

PROJECTS

1. Automated Business Insights and Campaign Optimization

- Developed an LLM-powered solution using RAG architecture to analyze client transaction and campaign data, enabling data-driven decision-making through a custom dashboard.
- Segmented customers using customer lifetime value (CLV) and predicted next purchases using LSTM based on historical data.
- Targeted specific customer segments with personalized texts generated by an LLM for future campaigns.
- Containerized the entire application using Docker, simplifying deployment, ensuring consistency across environments, and enhancing scalability.
- **Tech Stack:** Python | Pandas | Numpy | LLM | LangChain | Milvus VectorDb | Flask |Docker

2. DocEase - Chat with PDF

- Developed a user-friendly web application that allows users to upload and interact with PDF documents through a chat interface, leveraging generative AI models for enhanced document processing functionalities.
- Implemented a React frontend for seamless PDF uploads, chat interactions, and content visualization, ensuring a responsive and intuitive user experience.
- Created a Flask backend to handle secure PDF storage, process user queries, and communicate with generative AI services using LangChain, optimizing the LLM integration process.
- Integrated various generative AI capabilities for question answering, summarization, and translation using the Gemini-1.5-pro model, providing users with actionable insights directly from their documents.
- Established a MongoDB database for storing user chat history and processed PDF information, enabling efficient retrieval and analysis of interactions.
- **Tech Stack:** Python | React | Flask | MongoDB | LangChain | Gemini-1.5-pro | ChromaDB

3. Anuwad - English to Marathi Translator click here

- Developed a neural machine translation model using TensorFlow, leveraging the Bahdanau Attention mechanism to translate sentences from English to Marathi.
- Implemented an encoder-decoder architecture, where the encoder processes input sequences and the decoder generates target sequences based on context.
- Integrated the Bahdanau Attention algorithm to enhance translation quality by allowing the decoder to focus on relevant parts of the input sequence during generation.
- Created a user-friendly Flask-based UI, enabling users to input English sentences and receive Marathi translations with ease.
- Demonstrated model effectiveness through various sample translations, showcasing high accuracy and contextual relevance in translations.
- **Tech Stack:** TensorFlow | Python | Flask | Bahdanau Attention

4. D-predicto click here

- Developed an end-to-end application predicting **diabetes**, **heart disease**, and **Parkinson's disease** using **XGBoost** and **RandomForest** models.
- Integrated an NLP-driven chatbot into the website to improve user interaction and access to information.
- Included a feature for users to book appointments with doctors if needed.
- **Tech Stack:** Python | SK Learn | ML | Numpy | Pandas | NLP | Flask | GitHub

5. RetainIQ click here

- Developed a system for predicting employee churn using **RandomForest** and **XGBoost** models to analyze workforce data.
- Implemented features for model training, batch prediction, and single employee prediction to determine retention likelihood.
- Built using Flask for web-based interaction
- **Tech Stack:** Python | SK Learn | ML | Numpy | Pandas | NLP | Flask | GitHub

6. GestureFlow click here

- Innovatively utilized Computer Vision and Gesture Recognition to redefine presentation interaction by creating a dynamic, hands-free platform for slide navigation and annotation.
- **Tech Stack:** Python | OpenCV | cvzone | OS Module | Numpy | Hand Tracking Module

CERTIFICATES

- Quatium Data Analytics
- Cognizant Artificial Intelligence
- Google Cloud Jam

ACHIEVEMENTS

- Winners Of Prostart 2023 -Team Matrix First Rank In College
- Academic topper throughout the graduation
- Earned Golden Badge in Python On HackerRank
- Solved over 300+ problems across diverse platforms