

MD SUHAIL AHMED

<https://www.linkedin.com/in/md-suhail-ahmed-4578a125b/>

 mdsohailahmed1843@gmail.com



<https://github.com/Sohail1173>



7411884484

Profile: I am passionate about Data Science & AI, having good object oriented programming skills, well versed knowledge of making full end-to-end Data Science pipeline from Data Gathering to Scalable deploying applications in various cloud platforms, having good analytical and problem solving skills.

Work Experience

1. AI Engineer Intern

- Algorithmia digitech pvt ltd: Apr 2023-July- 2023
- Collaborated with a team to develop the **SmartBasket.ai** mobile application. Contributed to the design and implementation of key features. During my internship at **Algorithmia**, I utilized web scraping techniques to collect data for over **2000 companies financial data** from websites to develop a **Large Language Model (LLM)**, named as **FinGPT**. Preprocessed and cleaned the collected data to ensure high-quality inputs for the model. Trained the LLM on a diverse dataset of company information, allowing it to generate relevant and contextually accurate text. Fine-tuned the model using **custom langchain agents**, optimizing its performance for specific business applications and use cases. Participated in regular team meetings and provided updates on progress.
- Technology Stack:** Langchain, Elasticsearch, Pinecone, Chromadb, BeautifulSoup, Huggingface, Transformers, whisper ai.

2. Data Science

HappyMonk.ai: Aug-2023- Present

- Collaborated with the **HappyMonk** team to collect and curate a diverse dataset of images for computer vision projects.
- Conducted meticulous image labeling, ensuring data quality and accuracy to support machine learning model training. Applied image enhancement techniques and effects to improve the dataset's quality and diversity.
- Played a crucial role in video collection efforts, focusing on capturing diverse video footage for **YOLO** model testing.
- Actively participated in model testing and evaluation using **YOLOv8**, contributing to the development of cutting-edge computer vision solutions.
- Training the object detection, classification, and segmentation models. Building models using NVIDIA TAO toolkits**
- Analyzed and cleaned data for a chatbot to build a scalable Retrieval-Augmented Generation (RAG) system capable of answering user queries effectively.
- Worked on a video and image analysis tool to summarize content related to school children. Implemented validation checks to detect duplicate or expired videos.
- Designed and developed an ETL pipeline to extract data from BigQuery and visualize chatbot performance in Looker Studio.
- Built and optimized search APIs using semantic search techniques to retrieve user ideas and survey responses.
- LangChain, MongoDB, ChromaDB, OpenAI (LLM), Looker Studio, Python, Hugging Face

Core Skills

Languages

Python, SQL, Object Oriented Programming

Deployment

Docker, AWS, Container Registry, Amazon S3, Amazon EC2, CI/CD pipelines, MLflow, DagsHub, GCS, BigQuery

Basic Data Structure

Linear Search, Binary Search, Bubble Sort, Quick Sort, Insertion Sort, Selection Sort, Merge Sort, DataFrames, Series, Recursion, Hashing.

Mathematics for ML & DL

Algebra, Probability, Statistics, Calculus, Matrices

Education

M.sc Mathematics

6.9sgpa

BCU, Bangalore
2020 - 2022

B.sc Mathematics

72%

Gulbarga University
2017 - 2020

12th Board

Crescent P.U College.
2015 - 2017

78%

Language

English
French
Hindi
Urdu
Kannada

Interest

Cricket
Photography
Cooking

Frameworks & Libraries

Tensorflow, Keras, Pytorch, Transformers, Simple Transformers, Pandas, NumPy, SpaCy, NLTK, Matplotlib, Seaborn, Langchain, Llama Index, OpenCV, Ultralytics, Huggingface, LLM, Git, GitHub, Flask, Fastapi

ML & DL Algorithms

Random Forest	ANN, CNN, RNN
KNN	Vgg, ResNet
Decision Tree	AlexNet, LeNet
K-Means	InceptionNet
DB-Scan	EfficientNet
	SSD, Yolo, R-CNN, Mask-RCNN, MTCNN

Tools

LabelStudio, Labelimg, Lableme, Linux, Ubuntu, Deepstream, Onnx, TensorRT, Postman