

Darshan Kholakiya

darshankholakiya12@gmail.com | +91 7710020979 | <https://github.com/Darshan660> | Mumbai, India

Summary

A highly motivated and detail-oriented individual with a strong passion for Data Science. Graduated in BSc Data Science, seeking a challenging opportunity to apply my knowledge and skills in a practical setting. Equipped with a strong foundation in Data Analysis, Statistical Modeling, NLP, Generative AI, and Machine Learning techniques, I am eager to contribute to a dynamic organization where I can utilize my expertise to its maximum level and contribute towards organization success.

Education

KES Shroff College | BSc Data Science | Mumbai, India | Jul 2020 - May 2023

Shri T.P. Bhatia College | H.S.C. (Science) | Mumbai, India | Jun 2018 - Mar 2020

Rustomjee International School | S.S.C. | Mumbai, India | Jun 2008 - Apr 2018

Experience

BugendaiTech | Data Science Intern

Mumbai, India (Remote) | Jun 2023 - Dec 2023

- Developed and lead the implementation of the "Project Documentation Generator" using Large Language Models (**LLMs**).
- Conceptualized and designed an innovative system to **automate** and personalize project documentation or summarization of meetings notes, revolutionizing the project management process, which increased team **efficiency** by **50%**.
- Precisely **fine-tuned** a pre-trained model, ensuring data quality aligned with the organization's specific requirements.
- Prioritized **data privacy** as a core objective of the project while concurrently achieving **increase** in team productivity and a remarkable **reduction** of making project documentation time by **60%**.
- Pioneered the design of the User Interface on Gradio, enabling real-time **streaming** of the output.

DoWell Research India | Data Analyst Intern

Mumbai, India (Remote) | Aug 2022 - Feb 2023

- Utilized Google Spreadsheet to **collect, clean, and standardize** global indicators data for **49** countries.
- Orchestrated the generation of various visualizations, including filled maps, pie charts, funnel and many more charts using **Looker Studio**; created **20+** reports for stakeholders across departments.
- Collaborated with a team of data analysts to conduct in-depth data analysis and provide valuable insights to support strategic **business decisions**.
- Actively participated in cross-functional collaboration, fostering **effective communication** and information exchange to identify data trends, patterns, and meet project requirements.
- Played a key role in developing and implementing **data quality** control measures to ensure the accuracy and integrity of data, facilitating reliable data analysis.
- Conducted statistical analysis and transformed findings into visually appealing representations, offering **meaningful insights** to stakeholders.
- Maintained comprehensive data records for multiple teams, ensuring **accuracy** and **consistency** in report generation for all 49 countries.

Skills

Programming Languages: Python, HTML, SQL, C++

Platform & Proficiency: Machine Learning, Deep Learning, NLP, Generative AI, Vector Database, LLM, RAG, Langchain, Chainlit, Open AI, Hugging Face, Flask, Gradio, Streamlit, Streamlit Cloud, Exploratory Data Analysis, Looker Studio, Tableau, Power BI, GitHub, Git, Google Sheets, Google Slides, MS Office

Databases: MySQL, RDBMS

Projects

WhatsApp Chat Analyzer

Data Analysis & Preprocessing, Data Visualization, Sentiment Analysis

This tool will help users understand their WhatsApp chat's patterns, frequency of messages exchanged with different contacts, and more

Laptop Price Predictor

Machine Learning, Web Scraping, HTML, Streamlit

The Laptop Price Predictor is a machine learning application designed to predict the prices of laptops and provide 3 recommendations based on various features and specifications

Asian Landmark Detection

TensorFlow Hub, Image Processing, Geo Spatial Visualization

The Asian Landmark Detection application takes an image as an input and predicts the name of the Asian Landmark and also provide you the geolocation of the landmark

Auto Tech News

Web Scraping, HTML, CSS, Flask

This application web scrapes the latest technology news and presents the top 6 news articles in a visually designed template, making it easy for users to access and download the latest news

PM Modi Speech Comparison (2020 vs 2021)

NLP, Python, Data Presentation, Jupyter Notebook

Analyze and gain insights from PM Modi's Speech of 15th August 2020 vs 2021. Utilize NLP techniques such as Stop Word Removal, Lemmatization, Sentiment Analysis, Name Entity Recognition, Topic Modeling

Certifications

Data Scientist (TryCatch Classes), Introduction to Generative AI (Google), Data Visualisation: Empowering Business with Effective Insights (Tata), Data Analytics and Visualization Virtual Experience (Accenture), Data Analytics Consulting Virtual Internship (KPMG), Python Basics (HackerRank), SQL Basics (HackerRank), Data Analysis with Python (IBM Developer Skills Network), Data Science 101 (IBM Developer Skills Network)