SANCHIT SINGLA

Data Scientist

📞 +91-7973498002 @ sanchitsingla1403@gmail.com 🔗 https://github.com/sa-1-2
🔗 https://www.linkedin.com/in/sanchit-singla/ 📍 Punjab, India

SUMMARY

Certified Data Scientist and Data Analyst Professional from DataCamp, currently seeking a Data Scientist role. Proficient in Python, SQL, Tableau, and Power BI for data analysis, visualization, and machine learning. Skilled in data exploration, hypothesis testing, Large Language Model fine-tuning, deep learning model construction, and CI/CD pipeline setup for ML model deployment on cloud platforms.

SKILLS

Machine Learning: Supervised, Unsupervised Algorithms   Deep Learning: ANN, CNN, RNN   Databases: SQL

Large Language Models (llm)   Programming Language: Python, MATLAB   Natural Language Processing (NLP)

Data Visualisation: Tableau, Microsoft Power BI   Platforms: TensorFlow, Jupyter Notebook, VS code   Azure

AWS: EC2, Sagemaker, S3   Advanced Excel   Python API Framework: Flask, Streamlit   Predictive Modelling

Web Scraping: Selenium, BeautifulSoup   Dockers   MLOps   Mlflow   Git/Gihub   CI/CD Pipelines

Communication   Analytical & Problem Solving Skills   Teamwork

EDUCATION

Google Advanced Data Analytics Professional Certificate

Awarded By: Coursera

📅 10/2023 - 01/2024

- Skills: Data Science, Data Analysis, Python Programming, Machine Learning: Naive Bayes, Linear & Logistic Regression, Decision Tree & Random Forest Classification, Ensembling Techniques, K-Means, DBSCAN, Statistical Analysis, Tableau Software, Data Visualization, Predictive Modelling, Exploratory Data Analysis (EDA), Regression Models.

Certification in Data Science & Machine Learning

Awarded By: Indian Institute of Technology Mandi

📅 12/2022 - 08/2023

- Skills: Python, Machine Learning (Supervised: Regression, Classification Models; Unsupervised: Clustering Models, PCA), Data Analytics tools (Tableau, Power BI, SQL), Statistics (A/B Testing, Hypothesis Testing, ANOVA, Post hoc), Mathematics (Calculus, Gradient Descent), Data Collection, Jupyter Notebooks, Matplotlib, Seaborn, Plotly.

B.E. Mechanical Engineering

Chandigarh College of Engineering & Technology

📅 08/2017 - 07/2021

WORK EXPERIENCE

Data Science Intern

Innomatics Research Labs

📅 01/2024 - Present   📍 Remote

- Engaged in real-time projects, applying practical skills and knowledge to address contemporary challenges and achieve tangible outcomes.

Machine Learning Intern

Feynn Labs

📅 12/2023 - Present   📍 Remote

- Proposed an AI-based prototype for detecting tomato crop diseases, aiming to mitigate yield losses **(10-50%)** and annual production loss **(10-20%)** associated with plant diseases in tomatoes.
- Conducted market research and segmentation in the Indian electric vehicle (EV) industry, revealing that **28%** of consumers express dissatisfaction. This finding suggests an opportunity to penetrate the competitive EV market in India by addressing the needs of this dissatisfied segment.

WORK EXPERIENCE

Data Science Intern

ByteIQ Analytics

📅 09/2023 - 10/2023   📍 Remote

Project: Singapore Bank Loan Risk Analyzer: ML Classification + POWER BI

- Analyzed a Singaporean bank dataset (**1000** rows, **17** columns) on customer loan transactions, employing machine learning classification to identify defaulters.
- Conducted EDA and feature engineering and implemented different classification models for different test splits (**0.25, 0.3, 0.2**).
- Achieved Logistic Regression as the best model with a balanced accuracy of **73%,** an AUC score of **0.73**, and an F1 score of **0.62** with a **0.25** test split.
- Leveraged Power BI to craft a compelling narrative and develop a Data Product for clients, resulting in a significant boost in loan recovery rates from **40.22%** to **54.94%.**

Business Development Associate

BYJUS

📅 01/2022 - 09/2022   📍 Mohali, India

Achievements

- Generated Total Revenue of **INR 23,00,000** by counselling prospects, leveraging a ticket size range of **46,000** to **98,000**.
- Recognised as a Top sales performer from the General Manager.

PERSONAL PROJECTS

SQL Web-App using Google Gemini Pro LLM

Gemini Pro Large Language Model

- Recognized the need for simplifying SQL query generation for data analysis tasks.
- Developed an SQL web application to streamline query generation using **Google Gemini Pro LLM** and **Streamlit**.
- Integrated **Google Gemini Pro LLM** for natural language understanding and utilized Streamlit for interface development, enabling automatic SQL query generation from user questions.
- Achieved significant reduction in query complexity, enhancing user experience and facilitating faster data analysis with increased efficiency and accuracy.

German Traffic Sign Classification Model

Convolution Neural Network Model

- Tackled the challenge of German traffic sign classification with a dataset comprising **50,000+** images from the INI Benchmark Website, containing **40+** distinct sign classes.
- Performed exploratory data analysis and feature engineering, implemented keras tuner, and achieved **99%** training and **98%** validation accuracy.
- Validated the model performance with a **98%** confidence level in predictions, a pivotal contribution towards enhancing road safety through accurate German traffic sign recognition.

CERTIFICATION

Data Scientist Professional Certificate	**Data Analyst Professional Certificate**
Awarded By: DataCamp	Awarded By: DataCamp

LANGUAGES

English, Hindi, Punjabi