SAHIL KOCHAR

sahilkochar08@gmail.com 9352994726 Bengaluru

https://www.linkedin.com/in/sahilkochar08 https://github.com/sahil-008-ops

SUMMARY

Data Scientist with strong foundations in Statistics, Probability, and Machine Learning, experienced in analyzing structured and unstructured datasets. Proficient in Python, Pandas, NumPy, Scikit-learn, and SQL for data cleaning, feature engineering, model building, and validation. Skilled in regression, classification, clustering, EDA, and predictive analytics. Capable of translating business problems into scalable data science solutions and creating dashboards and analytical reports to support decision-making.

WORK EXPERIENCE

Operations and Data Intern
Scaler
- Analyzed large structured datasets from AI training pipelines to ensure data quality and consistency.
- Performed data validation, anomaly detection, and root cause analysis to improve operational KPIs.
- Designed performance dashboards and SLA tracking reports to support data-driven decisions.
- Standardized review frameworks reducing rework rates and improving efficiency.
- Collaborated with QA and engineering teams to translate operational challenges into analytical solutions.

SKILLS

Technical
Programming: Python
Libraries: Pandas, NumPy, Scikit-learn, Matplotlib, Seaborn, TensorFlow
Database: SQL, MySQL
Machine Learning: Regression, Classification, Clustering, Supervised & Unsupervised Learning, Feature Engineering, Model Evaluation, Cross-Validation, Hyperparameter Tuning
Statistics: Probability, Hypothesis Testing, Descriptive & Inferential Statistics
Data Analysis: EDA, Data Cleaning, Data Validation, KPI Analysis
Visualization: Power BI, Advanced Excel (Pivot Tables, XLOOKUP)
Tools: Jupyter Notebook, VS Code, Git, Flask
Concepts: Model Deployment (Basic), Data Bias, Data Leakage, Model Drift

EDUCATION

Master of Computer Application in AI/ML 08/2024 – Present | Bangalore
Jain (Deemed-to-be University)

Bachelor of Computer Application 08/2021 – 05/2024 | Bikaner
Maharaja Ganga Singh University

PROJECTS AND RESEARCH PAPERS

PulmoScan – EfficientNetB0 Lung Cancer Classification
Python, TensorFlow, Flask, Grad-CAM
- Built a multi-class lung cancer detection system using EfficientNetB0 trained on 5,000+ CT scans, achieving 90–93% accuracy. Improved image preprocessing quality by ~40% and integrated Grad-CAM visual explanations, reducing clinician review time by 20% while improving model transparency and trust.

WildEye - Wildlife Prediction & Semantic Embedding Model
Python, Scikit-learn, Semantic Embeddings, FAISS, RAG, Pandas, NumPy.
- Built a predictive model leveraging **semantic embeddings** and **RAG-style retrieval** for data-driven decision-making.
- Reached **89.3% classification accuracy** and **0.91 F1-score** on a curated wildlife dataset.
- Applied feature engineering and ML pipelines to highlight actionable insights for stakeholders.

AI in Supply Chain Management: Enhancing Efficiency
(Research Paper Number - ISSN: 2320-2882)
- Analyzed operational data across 6+ units to identify inefficiencies, improving forecast accuracy by 18% and reducing delays by 22%. Translated insights into actionable recommendations supported by KPI dashboards.