# <SAAHIL KATARIA>

Kolkata,West Bengal, India

saahilkataria1@gmail.com
https://saahilkataria.com/
www.linkedin.com/in/saahil-kataria

Motivated Machine Learning Engineer with knowledge in NLP, LLMs, and Deep Learning. Familiar with DevOps tools such as Git, DVC, Docker, Jenkins, and Ray, along with pipeline monitoring using WhyLogs and EvidentlyAI. Capable of working with big data frameworks like PySpark and SparkSQL, and deploying scalable machine learning models.

## PROJECTS

### Torch Ensemble-Based Microscopic Fungi Classification

Built a PyTorch K-fold ensemble model for microscopic fungi classification using custom CNN architecture, data augmentation, and stratified cross-validation.

Deployed the model as an API using FastAPI, allowing users to upload images and receive predictions.

### Abalone Age Estimation using LGBM and PostgreSQL

Developed a web-based application for abalone age estimation using a pre-trained LightGBM model (Optuna for HP tuning) . Integrated FastAPI, PostgreSQL for storing predictions, Jinja2 templates for web pages, and Evidently.ai for data drift monitoring.

The app allows users to input features and receive predictions with stored results for future review.

### Advanced DeBERTa-Powered Language Identifier with Gradio Interface

Developed a robust language identification system using a fine-tuned pre-trained DeBERTa model. Integrated Gradio for a user-friendly web interface that allows users to input text and receive language predictions.

## EDUCATION

**BTech in Computer Science & Engineering (Data Science)**
MCKV Institute of Engineering
2020 - 2024

Higher Secondary Examination
Apeejay School Park Street
2020

Secondary Examination
Don Bosco Liluah
2018

## SKILLS & PROFICIENCIES

- Scikit-learn, XGBoost, CatBoost, LightGBM
- TensorFlow, PyTorch, , PyTorch Lightning , Keras , ARIMA
- Pandas, NumPy, Matplotlib, Seaborn, Plotly , Optuna , Albumentations
- AWS (SAAC02 Certified)
- MLflow, Weights & Biases (W&B)
- Flask, FastAPI , Gradio
- MySQL, PostgreSQL
- Git , DVC , Docker-Compose , Jenkins , Docker , Ray
- WhyLogs ,  EvidentlyAI
- PySpark , SparkSQL

# </MACHINE LEARNING ENGINEER>