

# Rohan Singh

LinkedIn

GitHub

Kaggle

GfGeeks

955212382

rohan4u.official@gmail.com

## Summary

Certified in core AI and Data Science domains including **Supervised/Unsupervised Learning**, **Exploratory Data Analysis**, and networking fundamentals (Coursera). Complemented by training in Agile, software engineering, and Linux (Wingspan). **Kaggle** contributor. Actively working on NLP tasks, **RAG-based agent systems**, and **large language models (LLMs)**, bridging theoretical knowledge with real-world AI applications.

## Experience

### ML Intern, Cloudcredits

May 2025

- Worked on task-specific Small Language Models (SLMs).
- Implemented Byte Pair Encoding (BPE) for optimized tokenization and vocabulary control.
- Built and evaluated lightweight classification models for domain-specific tasks.

## Skills

- **Programming Languages:** — Python, C++, SQL
- **Databases:** — SQLite, Neo4j (Knowledge Graphs)
- **AI & Machine Learning:** — Supervised, Unsupervised, Reinforcement, Deep Learning (CNN, LSTM, Transformers), NLP (BPE Tokenizer, Embeddings), Computer Vision (CV2, Mediapipe), GenAI, PINNs, AutoML
- **Hyperparameter Experiment Tracking:** — Optuna, Weights & Biases (WB)
- **Data Science & Analytics:** — Data Preprocessing, Feature Engineering, Statistical Analysis, Hypothesis Testing
- **Frameworks & Libraries:** — TensorFlow, PyTorch, Keras, Scikit-Learn, Hugging Face, OpenAI API
- **Cloud:** — AWS-S3, Cloudinary
- **Deployment & APIs:** — Flask, FastAPI, Render

## Projects

- **MLOps Pipeline Distribution using DVC** — Developed scalable pipelines for ML model versioning and deployment using DVC.
- **Livestock Disease Prediction Model (Lumpy and Rinderpest)** — Validation Accuracy: ~95.3%, Test Accuracy: 96.56%
- **ConvoLense** — It is a cloud-native service that transforms long-form audio like debates or meetings into structured, searchable insights using LLMs, vector search, and graph reasoning.
- **Determining Leg Deformities using MobileNet** — Built a lightweight CNN model using MobileNet to detect and classify leg deformities from images, enabling early intervention and care.
- **Bubble Sheet Marks Evaluator using OpenCV and MediaPipe** — Developed an automated bubble sheet grader using contour detection and hand landmark tracking for accurate scoring and visualization.
- **AeroPINN** — Used Physics-Informed Neural Networks (PINN) to simulate airflow patterns around airfoils in real time.
- **Advanced Regression on House Price** — Built ML models for house price prediction using advanced regression techniques.

## Certifications

Supervised Learning — Coursera  
Network Fundamentals — Coursera  
Linux Basics — Wingspan

Unsupervised Learning — Coursera  
Agile Testing — Wingspan

EDA — Coursera  
Software Engineering — Wingspan

## Education

### Bennett University

2022 – 2026

Bachelor of Technology (B.Tech), Computer Science Engineering with specialization in AI

CGPA: 9.05 / 10

## Hobbies and Interests

Playing badminton, reading tech blogs