

Vaibhav Semwal

Dehradun, UK | vaibhavsemwal15@gmail.com | [+91 9027182209](tel:+919027182209) | [linkedin.com/username](https://www.linkedin.com/username) | [Github profile](#)

EDUCATION

Indian Institute of Technology Roorkee

Aug 2022 – May 2024

Master of Science (M.Sc.) in Physics

JEST RANK: 72

IIT-JAM Rank: 295

CGPA: 7.2

- Relevant Coursework: Deep learning, Data Structures and Algorithm, Probability and Statistics

BCAS, Delhi University

Aug 2018 – July 2021

Bachelor of Science(H) (B.Sc.) in Physics

CGPA: 8.4

- Relevant Coursework: Calculus, Statistics, Computational Physics

EXPERIENCE

Data Analyst Intern | MAS

- Conducted in-depth **market surveys** of leading platforms with **Excel, SQL, Pandas, and NumPy**, and leveraging Seaborn and Matplotlib to visualize market trends and user preferences for informed product development.
- Conducted market research using analytical techniques to strategize a well-informed market entry plan.
- Impact: Conducted thorough statistical analysis to confirm the product's necessity and devised a well-structured project plan to create a Course marketplace within a **20-lakh budget**.

Data Science Trainer | NullClass

- Developed and delivered high-quality student projects with comprehensive video tutorials and dedicated support.
- CNN: An emotion detection system using **OpenCV** for facial recognition and pre-trained **TensorFlow Keras** models for emotion classification. Featuring a user-friendly **Tkinter GUI** for image input, and emotion display.

PROJECT

[Github portfolio page](#)

Credit Card Fraud Detection:

Tools: KNN, SVM, Heroku

- Performed EDA and pre-processed a **highly imbalanced dataset** of over 1 million records.
- Utilized **Logistical Regression, support Vector Machine, and K-Nearest Neighbours** to model the problem.
- Applied **grid search** for **hyperparameter tuning** compared performance of models based on **F1 scores** and **confusion Matrix** and **ROC curves**. Achieved classification **recall up to 93%** for logistic regression.

Real-time Object detection with YOLOv8

Tools: OpenCV, OCR, YOLO, streamlit

- Cleaned the dataset containing images, converting XML annotations into YOLOv8.
- Implemented a **YOLOv8** model and **pytesseract**, achieving **87% accuracy** in detecting and recognizing license plates within images and reading numbers from plate.
- Deployed the model into a user-friendly web application using **streamlit**.

Fine-Tuning Llama-2-7b LLM for Question Answering

Tools: PEFT, QLoRA, Llama

- Fine-tuned the Llama-2 model for answering questions with task specific instruction on Guanaco-Llama dataset.
- Implemented Parameter-efficient-fine-tuning (**PEFT**) and employed Quantization and Low-rank approximation (**QLoRA**) which reduced training time and memory usage by **90%**.

Speech Emotion Recognition System:

Tools: Tensorflow, librosa, Flask

- Collected audio samples from **TESS, RAVDESS, Crema-D, and Savee** datasets.
- Employed **Librosa** for feature extraction and used **CNN** based model to achieve a **97% accuracy**.
- Created a **web app** using **Flask** to upload and detect emotions in audio.

TECHNICAL SKILLS

Programming Language: Python | **Databases:** MySQL | **Data Science:** Machine Learning, Deep Learning, Natural Language Processing, Computer Vision

Frameworks: Scikit-Learn, Tensorflow, PyTorch, HuggingFace | **Model Deployment:** Streamlit, Flask, FastAPI, Docker

Additional Skills: MatLab, Web Scraping, MySQL, Tableau, MLOPs, RAG

Research Experience

Master's Thesis: Computation Analysis of Photonic Spin Hall Effect

- Developed and executed simulations for Transfer Matrix method using MATLAB and Python for Tamm Plasmon Polariton. Modelled light matter interaction and revealed a 18 micro-meter spin dependent transverse shift.

Certificates

- Business Statistics and Advanced Excel : Coursera
- IBM Data Analyst Certificate
- Machine learning and Deep learning : IIT Roorkee
- Advance Python : Udemy
- Python for Data Science : Udemy
- MySQL Basics : Great Learning