RAMDULAR YADAV

Data Scientist

+977 9819936338
rammey115@gmail.com

Computer Engineering
IOE Purwanchal Campus, Dharan, Nepal

I'm a Data Scientist who loves finding meaningful patterns in complex data. My expertise in Machine Learning and Deep Learning helps me build solutions that make a real difference. I'm comfortable with Python coding and have a solid background in math and statistics that lets me approach problems from different angles. What really drives me is turning raw data into insights that people can actually use. I'm always learning new techniques and technologies because I believe staying curious is how we push the boundaries of what AI and machine learning can do.

Technical Skills

Expertise: Machine Learning, Deep Learning, NLP, OCR, Computer Vision, Image Processing, Generative AI (GANs), Retrieval-Augmented Generation (RAG), Prompt Engineering.

Programming Languages: Python, C++, SQL, Bash, HTML/CSS.

Libraries and Frameworks: TensorFlow, PyTorch,Django, Scikit-learn, OpenCV, Hugging Face, Pandas, NumPy, Matplotlib, Seaborn.

Cloud and Deployment: Microsoft Azure, Docker, Kubernetes, Streamlit.

Big Data Tools: Hadoop, Spark, SQL.

Developer Tools: Jupyter Notebook, VS Code, Git, Anaconda, Colab, GPT/LLM API integrations.

Soft Skills: Leadership, Management, Teamwork, Communication, Creative Thinking, Analytical & Problem-Solving.

Projects

AI-Based Online Exam Proctoring System GitHub Link

Bachelor's Final Year Project | Fusemachine Project

- Developed a real-time, AI-powered online proctoring system to ensure secure and fair online examinations.
- Fine-tuned **YOLOv8** to detect banned objects like mobile phones and books from webcam input.
- Used **OpenCV** and **MediaPipe** to track face, eye gaze, and head direction to detect suspicious behavior.
- Implemented audio monitoring to detect unusual sounds or multiple voices using external microphone input.
- Detected tab switching and full-screen exits to prevent cheating during exams.
- Generated automated reports summarizing candidate behavior and possible malpractice cases every 5 minutes.
- Built the entire system with **Django** as the backend framework and achieved **90% accuracy** in detecting cheating behaviors.

Resume Parser Using NLP and OCR GitHub Link

- Developed an intelligent **OCR-based tool** for extracting key details from resumes using **Natural Language Processing (NLP)**.
- Automated extraction of essential information like names, skills, education, and work experience, helping recruiters save time and make better hiring decisions.
- Used **Tesseract OCR** and **pdf2image** for text extraction from scanned resumes and PDFs.
- Leveraged **SpaCy's Named Entity Recognition (NER)** to identify and categorize key entities such as names, dates, skills, and locations.
- Built an intuitive UI with **Streamlit** to allow users to upload resumes and instantly view the extracted data.
- Achieved **90% accuracy** in extracting structured information from resumes.

Image Caption Generator GitHub Link

Bachelor's Minor Year Project

- Implemented an **Image Caption Generator** using a combination of the **Flickr8k dataset** and a personally curated collection of 150 real-world photos, each with corresponding captions.
- Initially leveraged the **VGG16 model** for image feature extraction and **LSTM (Long Short-Term Memory)** networks for generating descriptive captions, achieving about **40% accuracy**.
- Evaluated the model using BLEU scores:
  - **BLEU-1: 0.534**
  - **BLEU-2: 0.286**
  - **BLEU-3: 0.189**
  - **BLEU-4: 0.083**
- Switched to **Vision Transformer (ViT)** and **GPT-2**, significantly improving the quality and relevance of the generated captions.

Student Performance Indicator GitHub Link

- Conducted thorough **Exploratory Data Analysis (EDA)** to analyze student performance data.
- Compared the performance of various algorithms including **KNeighborsRegressor**, **DecisionTreeRegressor**, **RandomForestRegressor**, and **AdaBoostRegressor**, with **Linear Regression** emerging as the optimal choice.
- Trained and evaluated the model, achieving high accuracy on the test data.
- Developed an end-to-end machine learning pipeline, from data preprocessing to model deployment.

Face Recognition Attendance System GitHub Link

- Developed an end-to-end **face recognition attendance system**, deployed with a user-friendly interface using **Tkinter**.
- Harnessed deep learning techniques to accurately identify and record individuals, automating attendance tracking.
- Integrated **OpenCV** for real-time face detection and recognition, ensuring high accuracy and efficiency.
- Streamlined attendance management workflows by eliminating manual data entry and ensuring error-free records.
- Deployed the system in real-world scenarios for automated attendance tracking in classrooms and offices.

Education

- **Himalayan White House International College** *2019-2021*
  +2 Science **CGPA: 3.35**
- **IOE Purwanchal Campus, Dharan** *2021-2025*
  Bachelor of Computer Engineering

Work Experience / Extracurricular

- **Fusemachine**
  AI Fellowship April 2024 - November 2024
  - Successfully completed a 6-month AI Fellowship, covering advanced topics in **AI/ML**, **data science**, **regression and classification**, **clustering**, **neural networks**, **image processing**, **NLP**, and **reinforcement learning**.
  - Gained hands-on experience through projects involving **predictive modeling**, **web scraping**, **model deployment**, and **reinforcement learning**, enabling the implementation of real-world AI applications.
  - Explored advanced AI concepts like **Transformers**, **Large Language Models (LLMs)**, and **MLOps**, and gained expertise in deploying scalable and efficient AI models.
  - Collaborated on group projects, enhancing teamwork, problem-solving, and communication skills while applying AI technologies to solve complex problems.
  - Deployed machine learning models in production environments using **MLOps** practices, ensuring that models were scalable, maintainable, and efficient.
- **CodeforNepal**
  Data Fellow january 2021- April 2022
  - Completed Data Camp as part of the Fellowship, honing skills in Python, machine learning algorithms, and data analysis and manipulation.
- **Karma Technology**
  Data Analyst November 2023 - May 2024
  - Responsible for creating datasets crucial to the success of the team's data-driven projects. Played a pivotal role in ensuring the availability of high-quality information, contributing significantly to the company's effective decision-making processes.

Certifications

- Python
- Data Science Course - Mastering the Fundamentals
- Introduction to Machine Learning
- Intro to ChatGPT and Generative AI
- Microdegree in Artificial Intelligence - Fusemachine AI Fellowship