

SUMAN LAMSAL

3262 Clary Boulevard South Drive, Greenwood, IN | (+1) 317-699-9077 | sumanlamsal246@gmail.com
linkedin.com/in/sumanlamsal | github.com/SumanLamsal

Summary

Computer Science graduate student with 2+ years of experience in Data Science and AI/ML. Skilled in developing scalable data pipelines, APIs, and deploying machine learning models using Docker, Airflow, MLflow, and AWS. Experienced in MLOps, CI/CD, and database management with a focus on automation, reliability, and performance. Driven to build production-ready, data-centric systems that deliver measurable business impact.

Education

Master of Science in Computer Science <i>Indiana University</i>	Aug 2024 – Present <i>Indianapolis, IN — GPA: 3.9 / 4</i>
Bachelor in Electronics, Communication and Information Engineering <i>Tribhuvan University, Institute of Engineering (Pulchowk Campus)</i>	Nov 2019 – May 2023 <i>Lalitpur, Nepal — GPA: 3.88 / 4</i>

Skills

Programming Languages: Python, R, Java, C/C++, SQL, HTML, CSS, JavaScript

Frameworks & Libraries: TensorFlow, Keras, PyTorch, Scikit-learn, OpenCV, YOLO, OpenPose, NumPy, Pandas, Matplotlib, Seaborn

Tools & Platforms: Linux, Bash, Git, GitHub, GitHub Actions, Docker, AWS (S3, EC2), Airflow, MLflow, Tableau, Excel

Core Competencies: Machine Learning, Deep Learning, NLP, Computer Vision, Data Engineering, Statistical Modeling, MLOps, CI/CD, Model Deployment, Data Visualization, Feature Engineering, Agile Development, Team Collaboration

Experience

Backend Engineer <i>OCM Engineering Groups, LLC</i>	Aug 2023 – Present <i>Greenwood, IN</i>
– Developed and maintained end-to-end data ingestion and validation pipelines for Duke Energy's pole inspection platform, reducing manual data preparation time by 20% . – Employed GIS-based record-update processes and asset-management tools to streamline pole-inventory workflows, reduce data-entry errors, and improve decision-making for clients. – Pole data evaluation and inspection – Maintained strong client relationships (e.g., telecom carriers, utility companies) by consistently delivering accurate data, meeting milestones and responding to change-orders with flexible solutions.	

AI Fellowship

<i>Fuse Machines</i>	Jan 2023 – Jun 2023 <i>Kathmandu, Nepal</i>
– Implemented ML models including Logistic Regression, Random Forest, and CNNs for predictive analytics, improving model accuracy by 15% through parameter tuning and feature engineering. – Built NLP and Computer Vision prototypes using PyTorch and TensorFlow , contributing to real-world applications in text classification and image recognition. – Conducted data preprocessing, feature selection, and model evaluation pipelines using Scikit-learn , Pandas , and MLflow for reproducible experiments.	

Data Engineering & Analytics Intern

<i>Everestwalks Groups</i>	Nov 2022 – May 2023 <i>Kathmandu, Nepal</i>
– Developed and optimized ETL pipelines using Python , SQL , and Pandas , reducing data latency by 25% and improving analytics availability. – Integrated dashboards in Tableau for real-time business insights, supporting data-driven decision-making across teams. – Automated data cleaning and transformation scripts with Bash and Airflow , minimizing manual intervention and ensuring consistency.	

Projects

Offline Handwritten Signature Verification | *CNN, Computer Vision, Flask, CI/CD*

– Developed and deployed a CNN-based signature verification system using TensorFlow , OpenCV , and Flask , achieving 92% classification accuracy .
– Implemented end-to-end automation via GitHub Actions for testing, containerization, and deployment, reducing manual deployment overhead by 40% .
– Collaborated in a 3-member team to manage user authentication , data storage (PostgreSQL) , and API integration for real-time verification requests.

Lane Detection for Self-Driving Cars | *OpenCV, NumPy, Flask, Docker*

- Implemented a computer vision pipeline for lane detection and curvature estimation using **OpenCV**, improving edge detection accuracy by **18%**.
- Containerized the project with **Docker** and deployed on **AWS EC2**, enabling scalable testing and collaborative model benchmarking.
- Built a lightweight **Flask-based API** for serving processed lane detection results to client-side dashboards.

Heart Rate Prediction Model | 1D-CNN, Deep Learning, MLOps

- Developed a deep learning model with **1D-CNN** on PPG and accelerometer data achieving **MAE < 3 bpm**.
- Integrated experiment tracking using **MLflow** and automated model retraining via scheduled **Airflow** DAGs for production scalability.
- Implemented CI/CD workflow on **GitHub Actions** for continuous testing, linting, and deployment of the prediction API.

BERT for AI-Generated Text Detection | NLP, Transformers, Web Deployment

- Fine-tuned a **BERT-based classifier** achieving **89% accuracy** on 100k text samples for AI-generated content detection.
- Deployed the inference service via **Flask REST API** with frontend integration, handling over **2k concurrent user requests**.
- Configured automated deployment pipelines using **GitHub Actions** and Docker for version control and reproducibility.

Generative Adversarial Networks (GAN) | PyTorch, Deep Learning, Collaboration

- Trained and optimized GANs on the Fashion-MNIST dataset using **Wasserstein loss and gradient penalty**, improving generation quality by **22%**.
- Collaborated with peers using **Git and GitHub** for version control, establishing pull request workflows and structured code reviews.
- Automated training logs and checkpoint management using **MLflow** for reproducibility and experiment tracking.