Ashish Agarwal

(203) 626-2315 : agarwal.ashish.singhal@gmail.com **||** https://CRLannister.github.io/
https://www.linkedin.com/in/agarwalashishsinghal/ **||** https://github.com/CRLannister
Flexible on Location within the U.S.

SUMMARY

Data science professional with 4+ years of industry expertise in building scalable data pipelines, deploying machine learning models in production, and managing cloud-based infrastructure. Proficient in Python, AWS, Docker, Kubernetes, and deep learning frameworks. Experienced in ETL/ELT processes, MLOps practices, and leveraging LLMs for data-driven solutions. Strong background in statistical analysis, A/B testing, causal inference, time series analysis, machine learning, and cloud-native technologies with a passion for optimizing workflows and enhancing operational efficiency.

EDUCATION

Master of Science, Data Science	(**University of New Haven**)	GPA: 3.94	[West Haven, CT :- Aug 2023 – Dec 2024]
Bachelor's in computer engineering	(**Institute of Engineering**)	%: 79.2%	[Pulchowk, Lalitpur, Nepal :- Nov 2016 – Sep 2021]

RELATED EXPERIENCE

Data Engineer Intern (**North East Scientific**) [Waterbury, CT :- June 2024 – Aug 2024]

- Led the development of ETL pipelines using MasterControl and Netsuite APIs, automating data extraction, transformation, and storage in MySQL, reducing external dependencies by 30%.
- Developed and implemented KPIs, metrics, dashboards, and reports for Inventory, Sales, and Production insights using Qlik Sense Cloud, enabling real-time data-driven decision-making from raw ETL-processed data.
- Designed a Retrieval-Augmented Generation (RAG) system for document processing using **LLama3-8B** model, Open WebUI and FastAPI, improving knowledge retrieval efficiency and streamlining employee training.
- Streamlined data validation and automation processes using cron jobs and GitHub Actions, automating tasks like timely file downloads, data validation, historical data logging, and reconciliation. Delivered reports to stakeholders with automated notifications, achieving 100% accuracy and saving 95% of the time spent on manual processes, significantly boosting efficiency.
- Quickly learned and addressed a **time-sensitive** label printing issue unrelated to my role, using ZPL to reprogram the scanner and printer, delivering the solution within the required timeline.

Graduate Research Assistant (**Secure and Assured Intelligent Learning Lab**) [West Haven, CT :- March 2024 – Dec 2024]

- Engineered multi-tenant GPU environments as a **SaaS**, similar to **Google Colab**, using Docker, Helm, and Kubernetes, optimizing computational resource allocation by 40% and reducing manual resource management time and effort by 99%.
- Applied advanced statistical techniques (wavelet transforms, ICA, SVM) for EEG signal processing and classification, enhancing model accuracy by 12%.
- Automated workflows for data management, web development, and deployment with Datalore and WordPress.
- Guided students in developing and refining their capstone projects, focusing on innovative applications of machine learning and data science.

Data Engineer (**University of New Haven**) [West Haven, CT :- Oct 2023 – Dec 2024]

- Developed and optimized ETL pipelines on AWS (Lambda, EC2, Redshift), improving data integration performance by 66%.
- Implemented OCR for feature extraction from image data and fine-tuned deep learning models using Hugging Face transformers, reducing model inference time by 15%.
- Collaborated with stakeholders to develop Power BI dashboards, delivering actionable insights to senior management.

Freelance Sr. Data Engineer (**Upwork**) [Lalitpur, Nepal :- May 2022 – July 2023]

- Led migration of a legacy healthcare ERP system to AWS (S3, Redshift, Glue) and Snowflake, designing PySpark ETL pipelines and Kafka-based event streaming, reducing data latency by 35% and cutting annual cloud costs by $28K.
- Built CI/CD pipelines using Jenkins and Terraform to automate deployment of a microservices-based SaaS platform, reducing release cycles from 2 weeks to 4 hours and improving system uptime by 25%.
- Engineered a Kinesis and Scala streaming pipeline for a fintech client, processing 2M+ daily transactions and integrating with ML models to reduce fraud detection time by 20%.
- Developed REST APIs (Django) to sync data between Salesforce and client-owned NoSQL databases (MongoDB, Cassandra), eliminating 50+ hours/month of manual data entry.

Software Engineer/ BI Developer (**LIS Yomari**) [Lalitpur, Nepal :- April 2021 – May 2022]

- Migrated on-premises data warehouses to AWS, implementing secure data transfer and scalable ETL pipelines using Apache Spark, Kafka, and AWS services (Batch, Lambda, Kinesis).

- Enhanced data models with star/snowflake schemas, boosting query performance by 30%.
- Developed MicroStrategy dashboards, Dossiers and Documents and automated reporting.

Data Science Intern   ([Tootle](#))   [Lalitpur, Nepal :- Jan 2021 – March 2021]
- Developed customer segmentation models using LRFM algorithms, identifying high-value user segments and increasing customer retention by 4%.
- Built real-time dashboards for data visualization using MongoDB, Django, and Python, improving decision-making efficiency.

RELEVANT PROJECTS

Multi-Tenant GPU Cluster   West Haven, CT
[Secure and Assured Intelligent Learning Lab](#)   Sep 2024 – Dec 2024
- Built a Kubernetes-based GPU cluster with JupyterHub integration, enabling multi-user access and efficient GPU resource sharing.
- Configured multi-tenant resource profiles using Kubernetes and Helm, supporting customized resource allocations, which improved utilization by 40%.
- Developed a secure access framework through Kubernetes Dashboard and JupyterHub authentication, ensuring isolated and reliable user access.
- Authored detailed documentation covering setup, deployment, troubleshooting, and maintenance steps, streamlining cluster management for research and high-compute workloads.

RAG-based Knowledge Management System for Training and Information Retrieval   Waterbury, CT
[North East Scientific](#)   June 2024 – Aug 2024
- Developed a Retrieval-Augmented Generation (RAG) system by scraping and cleaning data from the company's website and specification documents, converting them into markdown format.
- Built a Pinecone and Chroma vector database incorporating LLM models such as Llama3, Phi3, and Qwen2 with quantization techniques to optimize GPU usage.
- Employed advanced prompt engineering to ensure accurate, non-hallucinated responses by reranking documents and interfacing with the database.
- Designed a user-friendly web interface, similar to ChatGPT or OpenWebUI, enabling user context tracking and delivering reliable, query-specific responses for training and operational purposes.

Retail Sales Inventory and Traffic ETL and Reporting   Lalitpur, Nepal
[LIS Yomari](#) **[Client- Ralph Lauren]**   July 2021 – March 2022
- Architected a robust ETL/ELT pipeline on AWS, ingesting data from TrueVUE APIs, S3 buckets, GCP, and FTP servers, leveraging EC2, EMR, Kinesis, Glue, Lambda, S3, Airflow, Redshift, CloudWatch, CloudFormation, and IAM.
- Orchestrated workflows using AWS (EC2, EMR, Lambda, Airflow) and implemented automated monitoring with CloudWatch, ensuring 99.9% uptime.
- Developed metadata objects, metrics, and attributes based on specifications in MicroStrategy, conducted data validation and performance testing, designed dashboards and dossiers tailored to client requirements.

TECHNICAL SKILLS

Languages & Tools: Python, SQL, NoSQL, GraphQL, R, Pytorch, TensorFlow, Scikit-Learn, Streamlit, FastAPI, Spark, Hugging Face, Docker, Kubernetes, Git, Bash, Linux, Dask, Excel, Polars
Cloud & DevOps: AWS (Lambda, EMR, S3, IAM, EC2, Redshift, Pinecone, Databricks, Kinesis, Glue, Sagemaker, ECR, SNS), Azure, GCP (BigQuery, GKE, GCE, VPC, AutoML, Vertex AI, Dataproc), Kafka, Airflow, Jenkins, GitHub Actions, Terraform
Data Engineering & MLOps: ETL/ELT Pipelines, CI/CD, Distributed Training, Spark, Data Lakes, CloudFormation, Helm, MLflow, OCR
Visualization Tools: Power BI, QLIK, Tableau, MicroStrategy, Looker, Excel, Networkx, Graphviz, Altair, Bokeh, Plotly
ML/LLMOps: Supervised Learning, Unsupervised Learning, Reinforcement Learning, Deep Learning, Hyperparameter Tuning, Model Optimization(LORA, Quantization), LLM Eval, Unsloth, Hugging Face TRL, wandb

EXTRACURRICULAR ACTIVITIES

[Helping Hands Nepal, HENN](#)   Kathmandu, Nepal
Volunteer and Technical Supervisor   May 2017 – July 2019
- Implemented Linux-based educational systems in remote schools, designing network infrastructure, subnets, and VM using open-source technologies like Mozilla and SSH, ensuring sustainable, secure access for underprivileged students.
- Developed offline educational materials in Nepali, leveraging open-source technologies to improve education access in remote areas.