

Dilochan Karki

dilochan@colostate.edu | +1 (970) 632-8491 | linkedin.com/in/dilochan-karki | github.com/karkidilochan

EDUCATION

Master of Science in Computer Science
Colorado State University, Fort Collins, Colorado

Expected May 2025
GPA: 3.84/4.0

Bachelor of Engineering in Computer Engineering
Tribhuvan University, Nepal

2021

EXPERIENCE

Machine Learning Engineer (Research Assistant), Colorado State University
Fort Collins, Colorado

August 2023 - Present

- **Model Development:** Developed a sequential LSTM model to characterize time-series network traffic with 75% accuracy and analyze patterns for TensorFlow distributed training networks.
- **Data Analysis:** Implemented a data pipeline to analyze 500GB+ of traffic data using Spark to extract and analyze traffic patterns across distributed networks with 10 nodes.

Data Engineer, Naamche Inc.
Dublin, Ohio

November 2021 - July 2023

- **Data-driven decision-making:** Increased investment ROI by 30% by developing a regression algorithm to score 300K+ properties in the US based on short-term rental market heuristics.
- **Trend Analysis:** Collected & analyzed prospects of 5M+ properties in the short-term rental market in the US by designing and deploying a data pipeline using Airflow.
- **Product Innovation:** Built and deployed a conversational AI real estate agent that ranked 20 on ProductHunt with 100+ daily users, using LLMs and custom prompt optimization.
- **Performance Optimization:** Reduced infrastructure costs by 25% with a serverless ETL architecture using AWS Lambda and Redis queues, processing terabytes of data in data lakes.

Associate Data Engineer, Logic Information Systems
Bloomington, Minnesota

April 2021 - November 2021

- **Analytics Data Warehouse:** Led migration of retail analytics data warehouse of two Fortune 500 retailers to Snowflake, reducing query processing time by 40%.
- **Cross-functional Collaboration:** Helped business analysts reduce retail analysis report development time by 20% with automated analytics workflows to generate 1000 metrics and KPIs.

Data Science Intern, Wiseyak Inc.
Kathmandu, Nepal

November 2019 - July 2020

- **Automating Processes:** Improved archival and retrieval of 5,000 medical reports by developing and integrating an OCR system with a telemedicine platform.
-

SKILLS

Languages: Python, JavaScript/TypeScript, Java, Bash

Databases: SQL, PostgreSQL, Snowflake, NoSQL, MongoDB, DynamoDB, RDS

Frameworks & Libraries: TensorFlow, PyTorch, Apache Spark, Scikit, Numpy, Pandas, Django, FastAPI

Cloud & Infrastructure: Git, Docker, AWS(EC2, S3, Lambda, Redshift), GCP, Terraform, GitHub Actions

Reporting & Monitoring: Tableau, Looker, Postman, Grafana, CI/CD

PROJECTS

Optimizing language models for low-resource languages

- Improved named entity recognition in Nepali by 15% with language-specific BERT models, using Apache Spark to extract entity annotations from text data.

Privacy-Preserving Image Classification

- Achieved >85% accuracy while mitigating the impact of noise in gradients added by differential privacy, using in-domain pre-trained models in PyTorch, while maintaining privacy audits.

Brain-Computer Interface (BCI) Signal Classification

- Automated data validation by developing an SVM classification model trained on energy parameters in different frequency spectrums, integrated in collaboration with biomedical engineers.