

Madan K C
madankc71@gmail.com | (616) 816-7469 | github.com/madankc71
Irving, Texas, USA

PROFILE

AI/ML-focused Software Engineer with experience in LLM/RAG-based applications, full-stack systems, and data science, supported by prior Graduate Research and Teaching Assistant roles. Interested in PhD research in ML/AI and data-driven intelligent systems at the intersection of computer science and real-world domains.

EDUCATION

M.S. in Data Science and Analytics | Grand Valley State University, Michigan, USA, 2023
Coursework: *Machine Learning, Statistical Programming, Multivariate Analysis, Statistical Modeling and Regression, Health & Bioinformatics, Data Mining, Information Visualization*

B.E. in Computer Engineering | Tribhuvan University, Nepal, 2020

Coursework: *Data Structure and Algorithms, AI, Databases, Math (Linear Algebra, Probability, Numerical Methods, Calculus), Operating Systems, Computer Network, Discrete Structure*

RESEARCH & TEACHING EXPERIENCE

Graduate Research Assistant - Grand Valley State University | Aug 2022 - April 2023

- Explored deep learning models for image classification and evaluated different architectures.
- Conducted literature reviews and designed experiments, tuned hyperparameters, and analyzed model performance using standard metrics.
- Documented findings in structured reports to support faculty research.

Graduate Teaching Assistant - Grand Valley State University | Jan 2023 - April 2023

- Supported EGR 112 (C programming) classes, assisting in lab sessions and helping undergraduate students debug code and understand core programming concepts.
- Graded programming assignments and exams and provided written feedback to help students improve their code quality and problem-solving.

INDUSTRY EXPERIENCE

S3 Technical Solutions - Software Engineer | Nov 2024 - Present

- Work at the intersection of backend development and applied ML/AI, designing, and implementing services and user-facing interfaces for internal tools.
- Led development of an internal retrieval-augmented generation (RAG) assistant that combines a large language model with a vector database of curated organizational documents to answer employees' natural-language queries.
- Implemented the document ingestion and embedding pipeline, similarity search over the vector store, and prompt/context construction, along with logging for error analysis and quality monitoring.

Calcgen Solutions - Data Analyst | Aug 2023 - Nov 2024

- Performed statistical modeling, data analysis, and visualization on real-world datasets.

- Built predictive models and evaluated them using appropriate statistical and ML metrics.
- Designed data cleaning and preprocessing workflows to improve model robustness.
- Communicated insights via reports and dashboards to support stakeholders' decisions.

SmartSites Nepal - Software Developer | Dec 2019 – Nov 2021

- Developed and maintained responsive client-facing websites and web applications.
- Collaborated with project managers and non-technical clients to translate requirements into reliable, maintainable software solutions.
- Performed performance and usability improvements (page load time, layout, and accessibility).

ACADEMIC PROJECTS

- **Coral Reef Image Classification (PyTorch, CNN):** Implemented and trained a convolutional neural network in PyTorch to classify coral reef images with normalization and augmentation, achieving 99% test accuracy and strengthening deep learning and computer vision skills.
- **MNIST Digit Classification (MLP in R):** Built a multi-layer perceptron from scratch in R for MNIST digit classification, implementing custom activation functions and He-Initialization, and trained it via backpropagation to reach 98.43% accuracy.
- **eBird Species Recommendation (KNN from Scratch):** Implemented a k-nearest neighbors-based recommendation system from scratch to suggest likely bird species based on eBird checklist data, practicing feature engineering and evaluation of distance-based models.
- **Multi-modal CITE-seq Prediction (Bioinformatics ML):** Predicted ADT values from CITE-seq data using parallel multivariate regression and gradient descent on high-dimensional biological data, demonstrating the application of ML in an interdisciplinary bioinformatics setting.
- **California Unemployment Analysis (EDA & Visualization):** Performed exploratory data analysis and cleaning on California unemployment data using pandas and built Tableau visualizations to highlight spatial and temporal patterns across counties and metro areas.

TECHNICAL SKILLS

ML & AI: Supervised and unsupervised learning, classification and regression, deep learning, NLP, large language model (LLM), retrieval-augmented generation (RAG) pipelines, vector-database, model evaluation and error analysis

Statistics, Data Science & Research Methods: Probability and statistical inference, regression modeling, multivariate analysis, hypothesis testing, sampling, experimental design, exploratory data analysis, data cleaning and preprocessing

Programming: Python, R, Java, React, SQL, C, C++, PHP

Data & Analytics Tools: PyTorch, scikit-learn, pandas, NumPy, Jupyter Notebook, Tableau, Power BI, SAS, Excel, RStudio

Web & Cloud: Spring Boot, React, REST APIs, Git, Linux, AWS (including experience deploying and integrating services in cloud), GPU-accelerated model training

Collaboration, Teaching & Communication: Graduate Teaching Assistant and Research Assistant experience, worked with non-technical stakeholders