

Nikesh Gyawali

1601 Roof Dr. Manhattan, Kansas, 66502 | (785) 477-7430 | gnikesh@ksu.edu | [LinkedIn](#) | [Portfolio](#)

Technical Skills

- Programming Languages: Python, Bash, JavaScript, SQL
- Machine learning/Data Frameworks: PyTorch, TensorFlow, Transformers, Scikit-learn, Apache Spark, Hugging Face
- Specializations: Machine/Deep Learning, Natural Language Processing (NLP), Large Language Models (LLMs)
- Data Visualization: Matplotlib, Seaborn, Plotly, PowerBI
- Others: AWS/Cloud Platform, ETL Pipeline, Prompt Engineering, RAG, RLHF, LangChain, Ollama

Education

Ph.D., Computer Science, Kansas State University, Manhattan, KS

December 2025

Bachelor's Degree, Computer Engineering, Tribhuvan University, Nepal

October 2016

Work Experiences

Fellow Post-Doc (*September 2025 – present*)

Department of Plant Pathology, Kansas State University, Manhattan, KS

- Pre-train foundational LLM model up to 1.3B parameters for fungal genome using large-scale genomic sequence data
- Build scalable HPC/ACCESS training pipelines for genome-scale data processing and model optimization
- Collaborate with computer scientists and plant pathologists to integrate biological insights from model development

Graduate Research Assistant (*January 2023 – August 2025*)

Machine Learning & Data Science (MLDS) Lab, Kansas State University, Manhattan, KS

- Develop advanced Transformers/LLMs models for NLP tasks like stance detection for multiple domains such as tweets, and financial text data from quarterly call reports ([link to paper](#))
- Leveraged Retrieval-Augmented Generation (RAG) and Prompt Engineering techniques to extract valuable insights from Large Language Models (LLMs)
- Optimized Large Language Models (LLMs) using Reinforcement Learning from Human Feedback (RLHF) to enhance response quality and alignment
- Implement state-of-the-art AI research from paper to code, adapting it for custom research requirements ([link to GitHub](#))
- Build and optimize large-scale Machine Learning models using Python and modern ML frameworks – utilize LLM to improve semi-supervised baseline accuracy on average by 10% on a challenging dataset ([link to paper](#))
- Collaborate on multiple ML/NLP projects and publish research papers in top-tier conferences ([Google Scholar](#))

Graduate Teaching Assistant (*August 2019 – December 2022*)

Kansas State University, Manhattan, KS

- Created Quizzes, graded assignments, instructed and provided academic support to students throughout the course and term projects for AI, Machine Learning, and Deep Learning courses
- Developed and evaluated technical assignments, mentored students on programming, database projects for Programming Languages, Database System, and Computer and Information Security courses

University Research Internship (*May 2021 – August 2021 and May 2022 – August 2022*)

LiuLab, Department of Plant Pathology, Manhattan, KS

- Applied Recurrent Neural Networks (RNNs), specifically Bidirectional Long Short-Term Memory (Bi-LSTM) models, to classify DNA sequences as core- or mini-chromosomes in *Magnaporthe oryzae* strains
- Developed and optimized a deep learning pipeline for genomic sequence classification, achieving high accuracy in predicting mini-chromosome presence
- Collaborated with an interdisciplinary team of computer scientists and plant pathologists to publish findings in a peer-reviewed journal (Publication: <https://doi.org/10.1093/nargab/lqae108>)

Security Analytics Engineer (*August 2017 - July 2019*)

LogPoint, Kathmandu, Nepal

- Developed Machine Learning models for User and Entity Behavior Analytics (UEBA)
- Developed algorithms for security log analysis and pattern detection, and maintained production-grade software systems

- Developed, maintained, and optimized plugins and code to efficiently parse thousands of logs using Python's Regex library, improving log parsing speed by 10%
- Took complete ownership of the development lifecycle for the product developed

Software Developer Intern (December 2016 - February 2017)

Leapfrog Technology, Kathmandu, Nepal

- Created video to gif converter using Machine Learning and Neural Networks ([link to project](#))
- Created various website templates using HTML5, CSS, and JavaScript
- Learned to work in a professional, corporate environment by collaborating with other team members (like Back-end Developers, UI/UX Designers, and Managers) to meet client's requirements

Selected Projects

Optimizing Financial Report Analysis with LLMs and Semi-supervised Learning

- Leveraged LLMs as data annotators using Prompt Engineering and RAG system in semi-supervised learning to improve label quality and reduce reliance on manually labeled data for analyzing quarterly earnings call reports for a private company
- Fine-tuned BERT-based models for stance detection, utilizing LLM-generated pseudo-labels to enhance the analysis of financial sentiment in earnings call transcripts, achieving average model's accuracy over 85%

AI-Driven Social Media Analysis for Vaccine Sentiment (Research paper under review for publication)

- Analyzed 10 years (2012-2022) of social media data using LLMs and NLP techniques, including lexicon-based analysis and topic modeling, to examine sentiment/stance and trends related to vaccines
- Implemented large-scale data crawling pipelines and conducted temporal, geospatial, and statistical analysis

FunBERT: Pre-training and Fine-Tuning Transformers for Genomic Insights ([Link to Hugging Face](#))

- Developed a custom foundation model for Fungal genomes, based on Transformer architecture, enabling genomic sequence analysis and classification in Fungal DNA
- Pre-trained the model from scratch and fine-tuned it further for downstream tasks while ensuring scalability for large-scale genomic data processing
- Achieved over 97% accuracy in DNA sequence classification, demonstrating the model's effectiveness in genomic analysis

Commercial Motor Vehicle crash severity prediction ([link to published paper](#))

- Developed and optimized Machine Learning models (Random Forest, CatBoost, XGBoost, Gradient Boost, and Transformer-based models) for crash injury severity classification
- Designed and implemented a scalable ML pipeline for data preprocessing, feature engineering, and model training on large-scale tabular datasets
- Enhanced model interpretability and explainability using the SHAP framework, providing insights into feature importance and decision-making

Time Series Anomaly Detection in Medical Break-the-Glass ([link to published paper](#))

- Implemented anomaly detection models (Gaussian Mixture Model) to identify anomalous user sessions during medical emergency access (Break-the-Glass events)
- Developed a two-tiered prediction system: a real-time fast analysis engine and a post-hoc detailed statistical evaluation model to analyze and flag abnormal access behavior
- Applied statistical modeling techniques to calculate likelihood scores for session anomalies, improving post-event auditing and security enforcement

Selected Publications

- "Predicting commercial motor vehicle crash severity in Kansas using explainable machine learning" **JTSS 2025** ([link](#))
- "GunStance: Understanding Gun Regulation Stance/Debates with ML/DL and LLMs" **ACL 2024** ([link](#))
- "Using Recurrent Neural Networks/Bi-LSTM for DNA sequence classification." **NAR Genomics 2024** ([link](#))
- "Disaster Tweet Classification using Deep Learning, Zero and Few-Shot Large Language Models." **DMTA 2023** ([link](#))
- "Time series anomaly detection using probabilistic model in medical break-the-glass" **HoTSoS 2020** ([link](#))