ANUGYA KHADKA

Dallas, Texas | +1 (940) 977-6893 | anugyakhadka234@gmail.com | GitHub | LinkedIn | Website

---

PROFESSIONAL SUMMARY

Results-driven AI/ML Engineer with Master's degree in Artificial Intelligence and 2+ years of production-focused experience in deep learning, NLP, and computer vision. Proven expertise in training and fine-tuning transformer models, building end-to-end ML pipelines, and optimizing model performance at scale. Skilled in PyTorch, TensorFlow, and HuggingFace with hands-on experience deploying models in real-world applications.

EDUCATION

Master of Science in Artificial Intelligence | University of North Texas, Dallas, TX Expected May 2026

- *Relevant Coursework: Deep Learning, Natural Language Processing, Computer Vision, Reinforcement Learning, Advanced ML Systems*

Bachelor of Science in Computer Science & Information Technology (BSc CSIT) 2018–2023

Institute of Science and Technology, Tribhuvan University, Kathmandu, Nepal

---

PROFESSIONAL EXPERIENCE

Machine Learning Engineer Intern | Neelcamp | Dallas, TX **May 2025 – Aug 2025**

- **Led transformer-based model development** for text classification tasks: fine-tuned BERT and RoBERTa models achieving **93%+ F1 score** on production dataset (50K+ samples)
- **Optimized attention mechanisms** through comprehensive research on embedding strategies, reducing inference latency by **28%** while maintaining accuracy
- **Built end-to-end ML pipeline**: data preprocessing, feature engineering, hyperparameter tuning, validation, and model evaluation using PyTorch and HuggingFace Transformers
- **Experimented with ensemble approaches**: combined transformer-based models with XGBoost for multiclass classification, improving precision by **12%** on imbalanced dataset
- **Conducted comparative analysis** of 6+ state-of-the-art LLM architectures (GPT variants, BERT, RoBERTa) documenting trade-offs between performance, latency, and resource consumption

Data Science Specialist | Treeleaf Technologies | Kathmandu, Nepal **May 2023 – Sep 2023**

- **Developed comprehensive EDA pipeline** using Python (Pandas, NumPy, Matplotlib, Seaborn) on 500K+ records, uncovering **key feature correlations** that improved downstream ML model accuracy by **18%**
- **Designed and deployed real-time monitoring dashboards** using Grafana and Prometheus, tracking system performance KPIs for 5+ production services with alert thresholds reducing incident response time by **35%**
- **Created interactive Power BI visualizations** for C-level stakeholders, enabling data-driven decision-making across 3 business units with 95%+ dashboard adoption rate
- **Optimized ML model pipeline** through rigorous hyperparameter tuning and cross-validation, benchmarking 8+ algorithms (Scikit-learn, XGBoost, LightGBM) and documenting performance metrics for production deployment
- **Implemented proactive alerting system** using Grafana for log data analysis, reducing mean-time-to-detection (MTTD) of critical issues by **40%**

Python Developer | Omistics Technologies | Kathmandu, Nepal **Jan 2023 – Jul 2024**

- **Built scalable Django REST API backend** serving 10K+ daily requests, implementing authentication, caching strategies, and database optimization (PostgreSQL) reducing response latency by **22%**
- **Spearheaded R&D initiatives** evaluating cutting-edge AI/ML frameworks and architectures for internal product integration; documented technical recommendations across **12+ research areas**
- **Conducted data analytics projects** using Tableau, Weka, KNIME, and RapidMiner on diverse datasets (e-commerce, healthcare, IoT); generated **15+ actionable insights** presented to non-technical stakeholders
- **Created technical documentation bridge** between engineering and product teams, authoring **25+ technical guides** that improved cross-functional communication and reduced developer onboarding time by **30%**
- **Collaborated with ML researchers** to prototype AI-powered features, translating complex technical concepts into production-ready code with measurable business impact

Technical Writing Specialist | Omistics Technologies | Kathmandu, Nepal **Apr 2022 – Dec 2022**

- Authored comprehensive technical documentation suite including API specs, user manuals, and troubleshooting guides following industry best practices
- Partnered with engineering teams to translate complex technical concepts into accessible documentation for diverse audience levels
- Implemented feedback loop from users, iteratively refining documentation reducing support tickets by **20%**

MACHINE LEARNING PROJECTS

Automated Essay Grading System **Jan 2024 – Apr 2024**

Technologies: PyTorch, HuggingFace Transformers, Python, scikit-learn, XGBoost - Developed end-to-end automated grading model comparing **traditional ML algorithms** vs. **deep learning approaches** (BiLSTM with Attention, Fine-tuned BERT, RoBERTa) - Engineered custom attention mechanism to highlight essay features impacting grade prediction, improving model interpretability by **35%** - Achieved **91% accuracy** on test set with ensemble approach (Transformer + XGBoost), validated on 2K+ manually graded essays - GitHub: Automated-Grading-Models

Microplastic Particle Classification **Oct 2023 – Dec 2023**

Technologies: PyTorch, TensorFlow, OpenCV, Python - Built deep learning pipeline using **VGG16, ResNet-50, and Vision Transformers (ViTs)** for automated particle identification on microscopy images - Preprocessed dataset of **1.2K+ annotated images**, implemented data augmentation achieving **88% accuracy** with ResNet-50 (baseline: 72%) - Optimized model inference: quantized ResNet-50 reducing model size by **75%** with <2% accuracy drop, enabling edge deployment - GitHub: Microplastic-Classification

NanoGPT: Language Model Hyperparameter Study **Aug 2023 – Oct 2023**

Technologies: PyTorch, Python, Transformer Architecture - Implemented character-level GPT model from scratch to study training dynamics across varying hyperparameters (learning rate: 0.001–0.1, batch size: 8–128) - Quantified convergence behavior and text quality metrics across **50+ experimental configurations**, documenting trade-offs between training speed and model performance - Generated visualizations showing optimal hyperparameter ranges, reducing experimental time for future model iterations by **40%** - GitHub: NanoGPT-Study

Music Streaming Application with Genre Classification **Jun 2023 – Aug 2023**

Technologies: Django, PyTorch, CNN, Python, SQLite - Developed full-stack music streaming web app with **integrated CNN-based music genre classifier** (Accuracy: 87% on 10-genre dataset) - Engineered audio preprocessing pipeline using librosa, extracting mel-spectrogram features from **5K+ tracks** for model training - Deployed application handling **500+ concurrent users**, integrated Celery for asynchronous classification tasks reducing UI latency - GitHub: Music-Streaming-App

Unsupervised Data Mining in Heart Disease Dataset **May 2023 – Jul 2023**

Technologies: Python, scikit-learn, Pandas - Applied **association rule mining (Apriori algorithm)** and **clustering (K-Means)** on UCI Heart Disease dataset (303 samples, 13 features) - Discovered **15 high-confidence association rules**

linking patient attributes to disease risk, generating actionable medical insights - Optimized clustering with elbow method and silhouette analysis; visualized patterns improving data interpretability for healthcare analysts

Netflix Content Analysis & Stock Correlation Study **Apr 2023 – Jun 2023**

Technologies: Tableau, Python, Pandas, Matplotlib - Analyzed **8K+ Netflix titles** across genres, release dates, ratings; created interactive Tableau dashboards exploring content trends - Engineered hypothesis: Netflix content diversity impact on stock performance; conducted statistical correlation analysis (Pearson, Spearman) - Generated **5 data-driven insights** presented to stakeholders, identifying content categories with highest viewer engagement and ROI potential

---

TECHNICAL SKILLS

- **Core ML Frameworks:** PyTorch, TensorFlow, Keras | **Deep Learning:** Convolutional Neural Networks (CNN), Recurrent Neural Networks (RNN), Transformers, Attention Mechanisms, Generative Adversarial Networks (GAN), Vision Transformers (ViT)
- **NLP & LLMs:** Hugging Face Transformers, BERT, RoBERTa, GPT models, Text Classification, Fine-tuning, Embedding Strategies, Attention Mechanisms
- **Computer Vision:** OpenCV, Image Classification, Object Detection, Semantic Segmentation, Data Augmentation, Transfer Learning
- **Machine Learning:** Supervised Learning, Unsupervised Learning, Reinforcement Learning, Feature Engineering, Hyperparameter Optimization, Model Evaluation Metrics, A/B Testing
- **Data Tools:** Pandas, NumPy, Scikit-learn, Matplotlib, Seaborn, XGBoost, LightGBM
- **Data Visualization & Analytics:** Tableau, Power BI, Grafana, Prometheus, KNIME, Weka, RapidMiner
- **Backend & Web:** Django, Django REST Framework, Python, SQL, PostgreSQL, SQLite
- **Programming Languages:** Python (Expert), R (Intermediate), Bash/Shell Scripting
- **Tools & Platforms:** Git, GitHub, VS Code, Anaconda, Google Colab, PyCharm, Jupyter Notebook
- **Additional:** Model Deployment Concepts, API Development, Data Pipeline Design, Documentation

---

CERTIFICATIONS & ACHIEVEMENTS

- **Machine Learning using Python Programming** 2022
- **Data Science Professional Certificate** 2022
- **Introduction to Packet Tracer Networking** 2021
- **Hult Prize Innovation Challenge Finalist** 2021
  Selected among top 10 teams for sustainability-focused tech solution
- **Python Skill Development Workshop** 2021

ADDITIONAL INFORMATION

- **Languages:** English (Fluent), Nepali (Native)
- **Open to:** Relocation, Remote opportunities in AI/ML Engineering
- **Portfolio:** GitHub profile showcasing all projects with detailed documentation and reproducible experiments