

Prakrit Timilsina

Curriculum Vitae

Personal Information

Full Name:	Prakrit Timilsina
Date of Birth:	8 November 2001
Nationality:	Nepalese
Address:	Kathmandu, Nepal
Email:	prakrittimilsina23@gmail.com
LinkedIn:	linkedin.com/in/prakrit338
GitHub:	github.com/prakrit338

Education

2020–2024	Bachelor of Technology in Computer Science & Engineering Vellore Institute of Technology (VIT), Vellore, India
-----------	--

Professional Experience

Dec 2023–Apr 2025	AI Engineer , Tulips Technologies, Kathmandu, Nepal <ul style="list-style-type: none">Led AI accessibility platform development (Google-funded), achieving 87% improvement in accessibility metrics through automated DOM classification and semantic navigationBuilt scraping and ETL pipelines for Anslagstavlan.se, normalizing postings from 150+ public sources with language detection, metadata mapping and deduplication; indexed content for combined keyword and semantic retrieval. System searches index (keyword + vector similarity), ranks relevant passages, and uses lightweight NLP (rule-based entity extraction and extractive answer selection) to assemble concise, multilingual responsesDesigned production RAG pipeline with Gemini fine-tuning, reducing hallucination rates by 34% with <200ms latency
Apr 2023–Sep 2023	AI Intern , THINK 4 TECH (Remote) <ul style="list-style-type: none">Built domain-specific QA systems and fine-tuned open-source LLMsDeveloped chat interfaces and performed benchmark evaluations

Research & Publications

2025	Novel Hierarchical Integration Method for Efficient Model Merging in Medical LLMs <i>First Author</i> , Submitted to IEEE 9th Conference on SmartIoT <i>Collaboration with Deakin University, Australia</i> <ul style="list-style-type: none">Developed hierarchical merging strategies for medical LLMs derived from Mistral-7BBenchmarked six merging approaches across five medical benchmarksDemonstrated competitive performance of simple averaging for resource-constrained settings
2024	Collaborative Large Language Models: Methodological Framework <i>Bachelor Thesis, First Author</i> <ul style="list-style-type: none">Proposed weighted linear merging of attention and feedforward layers for LLM combinationEvaluated merged models across 14 NLP benchmarks with improved performance on LAMBADA and OpenBookQA

Technical Skills

AI & ML:	LLM fine-tuning, RAG, NER, Transformer architectures, NLP, Computer Vision, Deep Learning
Programming:	Python (NumPy, Pandas, Scikit-learn), PyTorch, TensorFlow, Docker, Git, Linux, Bash
Cloud & Systems:	AWS (S3, EC2), Azure ML, ONNX Runtime, distributed inference architectures
Frameworks:	Hugging Face, LangChain, LlamaIndex, Weights & Biases, OpenCV, spaCy

Certifications

Jul 2025:	AWS Educate – Introduction to Generative AI
Aug 2023:	Google – Introduction to Large Language Models
Jul 2023:	Cognizant – AI Virtual Experience Program
Jun 2023:	IBM – Artificial Intelligence Analyst
Aug 2020:	Coursera – App Development
Aug 2020:	Great Learning – Java Programming

Languages

English:	C1 (Fluent)
Nepali:	Native
Hindi:	Conversational
German:	A1 (Learning)