

Vineet Kukreti

+91-7983207219 | vineetkukreti34@gmail.com | [linkedin.com/in/vineetkukreti](https://www.linkedin.com/in/vineetkukreti) | github.com/vineetkukreti34

PROFESSIONAL SUMMARY

AI Engineer with 1.5+ years of experience building production-grade healthcare AI systems across dermatopathology and oncology. Specialized in knowledge graphs, LLM-driven pipelines, and scalable medical data platforms enabling faster clinical and research insights.

EXPERIENCE

AI Engineer | *Southguild Technologies, Coimbatore*

May 2024 – Present

- Built production-grade healthcare AI platforms across dermatopathology and oncology, supporting real-world clinical and biomedical research workflows
- Developed an LLM-driven dermatopathology retrieval system using Neo4j and NLP-to-Cypher pipelines, achieving **98.8% accuracy** and enabling **5× faster** clinical insight retrieval
- Designed unified ETL pipelines integrating multi-source biomedical datasets to model gene–protein–pathway–mutation interactions for cancer research
- Implemented LLM-assisted oncology data pipelines, reducing manual data processing effort by over **80%**
- Deployed scalable microservices using FastAPI and Docker for real-time medical AI applications

PROJECTS

PathAssist: Dermatopathology Retrieval System

- Designed and built an end-to-end, clinician-validated dermatopathology retrieval system using Neo4j knowledge graphs for diagnoses and differential diagnoses
- Engineered an NLP-to-Cypher pipeline leveraging Cohere, Claude, and OpenAI, achieving **98.8% accuracy (88/89)** and outperforming baseline LLM-only approaches
- Enabled **5× faster** clinical insight retrieval, reducing clinician research time by over **60%**
- Deployed a production-ready microservice architecture using FastAPI and Docker; system validated by clinicians and presented at a medical conference
- **Tech Stack:** Python, Neo4j, FastAPI, Docker, Cohere API, Claude API, OpenAI

Target Discovery ETL Pipeline

- Built a unified ETL pipeline integrating heterogeneous biomedical databases to support gastric cancer drug-target discovery
- Modeled end-to-end gene interaction journeys across proteins, pathways, DNA elements, mutations, and regulatory mechanisms
- Enabled graph-based analytics over **100k+ records**, improving biomedical data accessibility by **65%** and accelerating biomarker discovery
- **Tech Stack:** Python, Neo4j, PostgreSQL, MongoDB

Navya Medical AI Pipeline

- Developed an LLM-assisted oncology data pipeline to automate patient data extraction, structuring, and classification
- Reduced manual processing effort by over **80%** while improving data consistency across clinical datasets
- Delivered a scalable pipeline enabling faster and more reliable clinical research workflows
- **Tech Stack:** Python, Neo4j, MongoDB, Pandas, AsyncIO, Docker, Claude API

SKILLS

- **Programming:** Python
- **AI / ML:** Large Language Models (LLMs), NLP, Knowledge Graphs, Prompt Engineering
- **Databases:** Neo4j, PostgreSQL, MongoDB
- **Backend & Infrastructure:** FastAPI, Docker, AsyncIO
- **Domains:** Healthcare AI, Oncology, Dermatopathology

EDUCATION

B. Tech (Hons.) in Computer Science Engineering | *Graphic Era University, Dehradun* | **8.5 CGPA**
Aug 2020 – Present
Specialization in Machine Learning & Artificial Intelligence
Intermediate/+2 | *Shri Guru Ram Rai Public School, Kotdwar* May 2018 – June 2020

ACHIEVEMENTS

- Hold **three patents** in AI/ML applications developed during undergraduate research
- Published **two research papers** in IEEE International Conferences on Artificial Intelligence and Machine Learning
- Awarded **PMSS Scholarship** (Prime Minister's Scholarship Scheme) for academic excellence