# Ankit Patel

765-775-8913 • ankit.patel03@gmail.com • linkedin.com/in/ankitpt/ • https://github.com/ankitpt

Oct'23 - Present

Mar'22 – Aug'23

Jul'20 – Mar'22

#### WORK EXPERIENCE

Founding Machine Learning Engineer, SigIQ.ai (EdTech startup with \$10M+ raised)

1. PadhAI - Android App for Indian Civil Services exam preparation

- Architected daily news feature with LLM-based relevance filtering and summarization, paired with semantic and keyword-based retrieval of relevant previous year questions, driving **2.5M+ reads and 43% of app usage**
- Trained article relevance scoring model with GPT-labeling of handcrafted features; 3.5x increase in user engagement
- Developed mock test analytics feature identifying subject-wise strengths/weaknesses, driving 10k+ test attempts
- 2. Evertutor.ai Webapp for GRE exam preparation
  - Designed probabilistic user knowledge model for GRE webapp using MLE estimation for personalized learning paths
  - Built 82% accurate decision-tree **college recommender system** using user profiles and QS/USnews university ranking features, driving **50% conversion of paying users**
- Led cross-functional collaboration on feature prioritization and tech specs resulting in 160k+ users base across two products

#### Machine Learning Engineer, Matterport

- Trained CNN room classifier with ResNet, boosting F-score by 15% across 20+ classes and reducing inference time by 30%
- Engineered data pipeline (Deeplake+Pydantic) ingesting 500k+ images 60% faster, enabling rapid model development cycles
- Trained semantic segmentation model (ViT) to detect floor material, achieving 0.8 mIoU across 8 classes on indoor scenes
- Engineered data annotation pipeline leading to 2x rare object labels improving performance by 25% on object detection task

#### AI Data Scientist, Blue Wave AI Labs

• Trained CNN models to predict neutron multiplication in nuclear reactors, enabling **cost savings of 320k USD** in cycle load design per nuclear reactor cycle

- Utilized SHAP and GradCAM for model explainability, highlighting key parameters in nuclear reactor operation
- Quantified prediction uncertainty via MC-dropout, enabling data-driven optimization of reactor operation
- Led presentations to the nuclear station engineers, informing feature design that improved model performance by 18%

# TECHNICAL SKILLS

Programming and Utilities	Python, Django, SQL, Bash, Git, Docker, AWS, GCP, Postman, Supabase
Libraries	Pytorch, Tensorflow, scikit-learn, pandas, llama-index, streamlit

# EDUCATION

Purdue University, West Lafayette, IN, USA   MS in Geomatics   GPA: 3.90/4.00	Aug'18 - May'20
Indian Institute of Technology (IIT), Kanpur, India   B.Tech. in Civil Engineering   GPA: 8.3/10.0	Aug'14 - May'18

# **PROJECTS & PUBLICATIONS**

# 1. Deep Learning based Road Marking Extraction from LiDAR Intensity Images

- Achieved state of art F-score of 85% by training U-net (Keras) with soft dice loss function to segment lane markings
- Fine-tuned the U-net encoder to achieve F-score of 86% on a different LiDAR sensor dataset with 5x less data

- **Patel, A.**, et al. (2020). Intensity Thresholding and Deep Learning Based Lane Marking Extraction and Lane Width Estimation from Mobile Light Detection and Ranging (LiDAR) Point Clouds. Remote Sensing, 12(9), 1379.

- **Patel, A.**, et al. Transfer Learning for LiDAR-Based Lane Marking Detection and Intensity Profile Generation. Geomatics 2021, *1*, 287-309. <u>https://doi.org/10.3390/geomatics1020016</u>

#### 2. Discovery of Novel Links in COVID-19 Knowledge Graph using Graph Embedding Techniques

- Utilized HOPE and SDNE embeddings to predict new links in COVID-19 knowledge graph with F-score of 88%
- Won best solution award in advanced category at Oak Ridge National Lab's 2021 SMC Data Science Challenge

- Patel, A., et al. Finding Novel Links in COVID-19 Knowledge Graph Using Graph Embedding Techniques. In Smoky Mountains Computational Sciences and Engineering Conference (pp. 430-441). Cham: Springer International Publishing