

NAVEEN PRAKASH R

+353 0892119224 ✉ naveenprakash25@gmail.com

[linkedin.com/in/np25](https://www.linkedin.com/in/np25)

github.com/lightshinichi

Visa Status: Stamp 1G (Eligible to work full-time in Ireland)

Education

Masters in Computer Science (Intelligent Systems) - Trinity College Dublin (Sept 2023-Sept 2024)

Bachelors in Electronics and Communication Engineering - SRM University (Jun 2017-June 2021)

Experience

Research Assistant - School of Computer Science ,University of Galway (June 2026 -Present)

Skills: Computer Vision, ROS2, Vision Language Models, Jetson Orin NX, Optical Flow, Image Processing

- Designed Built TerraBase, a real-time autonomous navigation platform (ROS 2 Humble, Linux) for a Unitree Go2 quadruped, integrating a self-supervised foundation model (DINOv2-based visual traversability) to flag unfamiliar terrain in unseen environments relevant to Search-and-Rescue;
- Deployed and optimized a DINOv2 Vision Transformer for real-time on-device inference by converting it to ONNX for an NVIDIA Jetson Orin NX (JetPack 6.2, CUDA 12.6, PyTorch), reducing per-frame latency to fit the embedded edge compute budget
- Integrated the Jetson compute unit with the physical Go2 robot and ZED 2i stereo camera. Improved real-time performance of a foundation model on edge hardware

Machine Learning Engineer (Remote) - SiPra

(Oct 2024 – Nov 2025)

Skills: Computer Vision, Python, Optical Flow, Image Processing

- Designed a **deep learning pipeline** to predict early Alzheimer's risk from **gait patterns** using 2D and 3D skeletal key points across 100+ subjects and multiple walking variations.
- Built a multi-stage architecture combining **YOLO** and **openpose** based pose extraction, **Vision Transformer feature encoding, relative spatial encoding, and LSTM-based temporal modeling** for 3D gait trajectory analysis.
- Modeled spatiotemporal gait dynamics by lifting 2D keypoints to 3D representations and learning sequence-level biomarkers correlated with cognitive decline.

Machine Learning Engineer - Multicoreware

(Dec 2021 – Apr 2023)

Skills: Python, C++, Jenkins, Linux, CMake, Pytorch, LSTM, Paraview, Density Based Clustering, CV

- Designed modular ML inference services with **REST APIs** enabling integration of **radar-based activity recognition models** into downstream applications.
- Implemented **CI/CD** workflows using **Jenkins** to automate compilation, testing, and deployment of **PyTorch** operator optimizations, enabling continuous integration of performance improvements for accelerator-based inference.
- Built and orchestrated a radar data pipeline using **Airflow** for ingestion, cleaning, and preprocessing of radar point cloud signals. Implemented **FFT** based feature extraction and clustering with DBSCAN to prepare structured inputs for downstream **BlazePose** and **LSTM** based activity and identity models achieving up to 95% accuracy.
- Built a production ML deployment pipeline on **AWS**. Trained and deployed models using **AWS SageMaker** and packaged inference services with **Docker** containers. Stored and versioned container images in **Amazon ECR** and deployed scalable **SageMaker inference endpoints**. Managed model versioning and experiment tracking with **MLflow**. Provisioned infrastructure using **Terraform** to support repeatable deployments across environments.

AI Developer (Freelance) Kyrios Ark World Logistics

(Dec 2020 – Dec 2021)

Skills: Computer Vision, TensorFlow, Python, Tesseract OCR, Spacy, Flask, Twilio API, Google Maps API, Pandas, scikit-learn, OpenCV, NLP, Google Sheets API

- Developed an end-to-end **OCR + NLP** automation pipeline using **Tesseract** and **Spacy** to extract invoice metadata (e.g., HS codes, quantities, CIF values) from scanned documents, streamlining customs documentation and reducing human error by 70%.
- Deployed a WhatsApp-based shipment tracking bot using **Twilio** and **Flask**, enabling clients to query real-time container status through **natural language**, backed by automated updates from Google Sheets.
- Designed a truck **ETA prediction system** leveraging Google Maps API and traffic-aware logic to estimate delivery times from manufacturing hubs to port terminals, significantly improving dispatch accuracy and reducing missed port cut-off incidents.

Skills

Programming Languages: Python, C++, C, ReactJS, SQL, Pyspark

Deep Learning Libraries: Pytorch, TensorFlow, OpenCV

Machine Learning Libraries: Scikit-Learn, NumPy, Pandas, Matplotlib, Seaborn

Tools and Tech: AWS, ROS 2, Jetson Orin, Langchain, LoRA, LLM Streamlit, Django, Git, Docker, Kubernetes, AWS Sagemaker, Terraform, Flask, Jira, MongoDB, Paraview, Jetson Nano, Arduino, ESP32, Raspberry Pi, Image Processing, Vision Algorithms, Quantization, OCR, Object Recognition

Certification

AWS Certified Data Engineering Associate (DEA C01)

- Designed secure, scalable batch and streaming data pipelines using S3, Glue, EMR, Kinesis, and Redshift
- Applied data modelling, partitioning, performance optimization, and monitoring best practices for analytics workloads

Projects

Ireland Manufacturing Site Selection Agent (ITag)

Tools & Algorithms: Python, Anthropic Claude API, Agentic Orchestration, Tool-Use, Parallel Execution, Constraint Parsing, Weighted Scoring, IDA/CSO/EPA/SEAI APIs, argparse, python-dotenv

- Architected a **multi-tool agentic LLM system** using the **Anthropic Claude API** with a custom **agentic orchestration loop** that parses natural-language manufacturing requirements into structured **SiteRequirements** objects via **LLM-driven constraint extraction** (NLP intent parsing, structured output generation).
- Engineered **six concurrent data-fetching tools** integrating real-world datasets from **IDA Ireland, CSO, EPA, and SEAI** to evaluate Irish counties across industrial land availability, workforce depth, logistics infrastructure, EPA permit complexity, and grid energy costs.
- Designed a **weighted composite scoring engine** (land 25%, IDA incentives 20%, labour 20%, logistics 20%, EPA 10%, energy 5%) to rank counties and synthesize **top-5 site recommendations** with per-factor breakdowns, risk annotations, and CLI-exportable text reports.

LLaMA 3.2 Fine-tuning for Text-to-SQL Generation

Tools & Algorithms: PyTorch, HuggingFace Transformers, PEFT, LoRA, LLaMA 3.2, torchao, Kaggle, MLflow

- Fine-tuned **LLaMA 3.2 (3B Instruct)** using **LoRA/PEFT** for Text-to-SQL generation, enabling accurate conversion of natural language queries into structured SQL statements.
- Applied **parameter-efficient fine-tuning (PEFT)** with LoRA adapters to minimize trainable parameters while preserving model performance, significantly reducing GPU memory requirements during training.
- Built a GPU-accelerated inference pipeline on **Kaggle** integrating **HuggingFace Transformers** and **torchao** for optimized model loading, quantization, and real-time SQL query inference.
- Managed model versioning and secure credential handling via **HuggingFace Hub** and **Kaggle Secrets**, packaging fine-tuned LoRA weights as distributable datasets for reproducible deployment.

Accomplishments

- Finalist on Smart India Hackathon 2020 (**Hardware Edition**)
- Designed, Developed and patented:
 - Designed and patented a smart accident prevention helmet integrating bone conduction audio and a heads-up display, with an embedded NVIDIA Jetson Nano running a real time traffic sign and traffic signal recognition model for on device inference and rider alert generation (**Patent Application Number: 202141060755**)
 - A fully automated smart kitchen robot for making stuffed parathas (An Indian Bread variety) (**Patent Application Number: 202141060759**)
- Developed and published a paper on Automated inventory management system <https://ieeexplore.ieee.org/document/9752546>