

Praneet Bomma

+447833882533 | praneetbomma@gmail.com | London, UK

LinkedIn: [linkedin.com/in/praneet9](https://www.linkedin.com/in/praneet9) | GitHub: github.com/Praneet9 | Portfolio: praneet.me

WORK EXPERIENCE

Research Engineer

April 2024 – Present

Hadean Supercomputing Ltd

London, UK

- Developed military training simulation scenario creation and orchestration with LLMs. Added support for function calling, platoon formations, directional cues and lat-long based movement.
- Enhanced existing doctrine violation checking by implementing a Retrieval-Augmented Generation (RAG) solution, transitioning from an instruction-based LLM with a 100k context window to a more efficient Gemma 2 model.
- Deployed the RAG-based solution in a local air-gapped environment, achieving a 80% reduction in operational costs by eliminating reliance on Gemini 1.5 Pro calls, while maintaining the accuracy of violation flagging.
- Streamlined CI/CD processes by dockerizing 11 systems and developing robust Git workflows for Docker builds, deployed via Docker Compose with built-in health checks and failure handling.
- Designed and implemented the UI and backend using ReactJS and Flask, which was successfully demoed to clients

Machine Learning Engineer

August 2021 – September 2022

Deep Learning Analytics

Canada, Remote

- Developed data ingestion pipeline and a deep branched architecture from scratch, integrating 3D Convolutions and LSTMs for Rain Attenuation Prediction on spatial and sensor data.
- Trained AutoEncoder and Conditional GAN for floor pattern transfer. Implemented illumination-transfer research paper and experimented with Pyvista and Blender for adding shadow effects on floor patterns.
- Enhanced object detection, segmentation, jersey classification, and pose-estimation models on 4K stream data. Developed 3D pose transformation using multiple-view geometry. Initiated development of data ingestion pipelines and implemented classification of football events like passing, goals, blocks, freekicks, etc on live video using XGBoost. Engineered additional features and achieved 90%+ AUC score and automated hours of manual work.
- Orchestrated an end-to-end Realtime Football Video Analytics pipeline, efficiently processing 12 4K streams with 5 models using Deepstream, Triton and TensorRT. Leveraged Libtorch and CUDA kernels for instance segmentation post-processing. Helped build end-to-end streaming pipeline with C++ and CUDA streams on A100 servers.
- Assisted in developing visualization tools in C++ to render real-time 3D graphics (OpenGL) overlaid on UHD video with Cinder and Unreal Engine. Visualization included rendering 3D pitch with player tracking graphics, performance analytics, and the results of the deep learning models.

Data Scientist / ML Engineer

September 2019 – August 2021

Blackstraw.ai

Mumbai, India

- Refined YOLOv3 and DeepSORT tracking for a Real-time Risk Monitoring Video Analytics system. Improved facemask detection with colour mode changes to training data. Optimized models with TensorRT, achieving a 2x throughput. Deployed the system with Deepstream 5 for 40 cameras to process streams in real-time and reduced the server cost by 80%. Streamlined CI/CD with Docker and Jenkins on AWS.
- Built object detection and tracking for UAVs. Quantized models to half-precision and optimized the threading pipeline to run real-time on edge devices like Nvidia Jetson Nano and Xavier.
- Designed real-time detection and instance segmentation using Yolact++, experimented with SLAM for path planning, and crafted a novel Neural Path Planner with branched architecture incorporating human priors for achieving intermediate waypoints independent of GPS.
- Developed Autonomous Navigation System for mini-trucks. Achieved 5x optimization in Occupancy Grid generated using LiDAR, increasing the speed from 4FPS to 20FPS. Integrated all components and established seamless inter-process communication using Robot Operating System.
- Automated deployment of ventilator software on Raspberry Pi, in partnership with PhDs from Stanford University and the University of Utah. Designed custom system services for automating software components and implemented robust security measures for system integrity and failure handling.

EDUCATION

Master's in Autonomous and Intelligent Systems

September 2022 – September 2023

University of Sheffield, Distinction

Sheffield, UK

B.E. in Computer Science

August 2015 – May 2019

Mumbai University, 7.23 CGPA

Mumbai, India

TECHNICAL SKILLS

Development: Python, C, C++, CUDA, SQL, Flask, Django, Redis, Kafka, ROS, Bash Scripting, FastAPI
Machine Learning and Deep Learning: PyTorch, Tensorflow, Keras, Scikit-Learn, PyTorch Distributed
Deployment: Git, Docker, Deepstream, GStreamer, TensorRT, Triton, ONNX, OpenVINO, Docker, AWS
Libraries: NLTK, Libtorch, HuggingFace, OpenCV, cuDF, Pandas, NumPy, Matplotlib

PROJECTS

MAFAT

Data Science competition by The Israeli Ministry of Defense and Directorate of Defense Research & Development (DDR&D) to classify whether a radar signal segment represents a human or an animal

- Attained 23rd rank globally on the competition public leaderboard, showcasing a 0.9028 ROC AUC score.
- Developed a CRNN architecture from the ground up for classification tasks, effectively addressing class-imbalanced data challenges through the use of Focal Loss and Hyperparameter tuning.
- Employed an ensemble technique to optimize performance by selecting the superior score between two well-performing trained architectures. Additionally, explored the potential of transformers to enhance the AUC score.

ReLIE - Research Paper Implementation

Implemented paper by Google Research - "Representation Learning for Information Extraction from Form-like Documents" - [Link to paper](#)

- Constructed the neural network outlined in the paper from scratch, addressing gaps and unknown elements that were omitted in the paper's implementation
- Incorporated a self-annotated dataset alongside a publicly available dataset to enrich the training process and enhance model performance
- Successfully enhanced results by effectively applying Focal Loss to handle imbalanced data challenges

Travel Mode Classifier - Research Paper Implementation

Implemented paper "Trajectory-as-a-Sequence: A novel travel mode identification framework" - [Link to paper](#)

- Engineered a 5-step data aggregation and filtering process, leveraging Pandas vectorization and multiprocessing, to handle a sizable 5GB dataset efficiently.
- Developed a feature engineering pipeline extracting relevant features from GPS data, optimizing model performance for sequence classification tasks focused on travel modes (walk, bike, car, taxi) using a publicly available dataset.
- Constructed and trained a sequence-to-sequence model, comprising a convolutional encoder and bidirectional LSTM decoder from scratch, achieving accurate classification of travel modes.

OPEN-SOURCE CONTRIBUTIONS

- Core contributor in the deployment team for [pez-globo/pufferfish-software](#), the software repository for the cost-effective and portable pufferfish ventilator.
- Contributed a feature for [kornia](#), a differentiable computer vision library for PyTorch.
- Contributed two features for [unifyai/ivy](#), an open-source machine learning framework.
- Fixed a bug in the TensorRT conversion code for YOLOv3 in the [jkjung-avt/tensorrt_demos](#) repository.

CERTIFICATIONS

- Deep Learning Specialisation on Coursera

PUBLICATIONS

- "Disorder Detection of Tomato Plant using IoT & Ensemble Techniques" AFITA conference at IIT Bombay, Maharashtra, India. [Link to paper](#)

ACHIEVEMENTS

- **Ranked 23rd Globally** in the prestigious MAFAT Challenge, hosted by The Israeli Ministry of Defense. Distinguished as one of the top 25 teams out of 300 with an exceptional AUC score surpassing 0.90.
- Recognized as an **Outstanding Performer** for Q2, Q3 - 2020 & Q1 - 2021 at Blackstraw.ai