

INSAF ISMATH

Abu Dhabi (UAE Golden Visa) | i.m.insaf@gmail.com | +971 566071157 | [in linkedin.com/in/iminsaf](https://www.linkedin.com/in/iminsaf) | insafismath.com

PROFILE

AI/ML Engineer and Researcher with 5+ years of experience in AI, translating business problems into end-to-end production-grade AI systems that deliver measurable impact, and publishing research on large multimodal models at top-tier AI conferences including EMNLP and CVPR. Skilled across the full generative AI stack alongside classical ML and scalable data-led solutions.

PROFESSIONAL EXPERIENCE

2PointZero – Energy and Consumer investment platform under IHC

AI Engineer

Abu Dhabi, UAE
Oct 2025 – Present

- **Geospatial intelligence platform (beauty & wellness portco):** Drives consolidation and expansion decisions across a multi-brand salon network (130+ branches, 8 brands) via a branch-level Risk Index and hex-level Opportunity Index over an H3 grid. Shipped as an interactive dashboard, 15-endpoint REST API, and 13-tool MCP server for natural-language executive queries. *React, MapLibre, Cesium; FastAPI + MCP server*
- **Live Avatar AI Agent (Group Growth Day):** Delivered the opening and closing remarks and engaged C-suite of 20+ portfolio companies live. Grounded on per-company briefs and the Group's portfolio context, it posed tailored follow-up questions after each presentation and fielded audience Q&A within scoped guardrails. *STT-LLM-TTS pipeline, LiveKit, Anam 3D avatar, ElevenLabs*
- **FTE (Full-Time Equivalent) benchmarking agent:** Compresses weeks of peer benchmarking into a single conversational query. Custom ReAct loop with multi-provider LLM fallback returning citation-anchored answers, exposed as 7 FastMCP tools over a hybrid retriever on PostgreSQL. Upstream pipeline scrapes peer disclosures and runs per-segment iterative LightGBM imputation across 12 segments, validated by multi-LLM consensus. *FastAPI, PostgreSQL (pgvector + tsvector + pg_trgm), FastMCP, LiteLLM, Qwen3, LightGBM, Redis*
- **Group-wide custom GPT rollout (8,000+ GPTs):** Led portfolio rollout on the Group's enterprise agentic AI platform, compressing months of manual config into an automated pipeline that LLM-maps employees to GPTs by role, function, and hierarchy and provisions at scale via Playwright. Ran technical rollout sessions with portfolio-company teams. *Prompt Engineering, Playwright*
- **Strategic research partnership (2PointZero × MBZUAI):** Represented Innovation Labs in scoping a multi-year partnership, pitched business use cases to faculty across multilingual financial NLP, multimodal document intelligence, causal decision intelligence, and AI evaluation & governance, paving the path to joint publications, sponsored PhDs, and co-developed IP.
- **Agent evaluation pipeline:** LangGraph harness that auto-generates ground-truth Q&A datasets, runs multi-provider parallel answer generation, and scores via LLM-as-Judge with Langfuse traces, turning ad-hoc agent testing into reproducible benchmarking. *LangGraph, Langfuse*
- **AI Recruiter App (university career fairs):** Mobile-first candidate-research app scans a candidate's QR, resolves it to their CV, runs a source-aware agent that ingests the CV and enriches it with live Google Search grounding, and outputs a structured candidate summary with real-time recruiter notes. *React + TypeScript, Tailwind, Supabase, Gemini/Groq/OpenAI, Vercel*

Intelligent Visual Analytics Lab (IVAL) – MBZUAI

AI Research Engineer

Abu Dhabi, UAE
May 2025 – Oct 2025

- **Calibration of prompt-tuned CLIP (CVPR 2026):** Contributed to a dual-regularization framework addressing underconfidence on base classes and overconfidence on novel classes; ran experiments under the standard base-to-novel protocol across 11 datasets and assisted with manuscript writing and revision. *PyTorch, CLIP, prompt-tuning frameworks (CoOp, MaPLe, KgCoOp)*

VisionLabs – Human and Object Recognition Technologies

AI Engineer

Dubai, UAE
May 2024 – Aug 2024

- **Synthetic-data face augmentation pipeline:** Trained IR-50 + ArcFace baseline on MS1M-RetinaFace-T1 (5M faces, 90K identities) and extracted IR-101 CurricularFace descriptors to compute per-identity class centers; generated per-identity synthetic faces with Arc2Face conditioned on those centers, and retraining on mixed real+synthetic data improved face-recognition accuracy by 30% across 7 benchmarks. *PyTorch; IR-50/IR-101, ArcFace/CurricularFace; Arc2Face (diffusion)*

Intelligent Visual Analytics Lab (IVAL) – MBZUAI

Graduate Research Assistant

Abu Dhabi, UAE
Dec 2023 – May 2025

- **Prompt-sensitivity benchmark for LMMs:** Built an evaluation harness on 20× NVIDIA A100s, orchestrating 1,500+

inference runs across 8 open-source and 2 proprietary LMMs over 3 MCQA benchmarks in under 2 months. Published at EMNLP 2025 Findings. *PyTorch, Hugging Face*

- **Long-form video understanding benchmark:** Built a human-in-the-loop QA generation pipeline sourcing egocentric (Ego4D) and non-egocentric videos (3–30 min) where existing benchmarks fell short; implemented inference and evaluation code across 7 open-source and 2 proprietary Video-LMMs. *PyTorch, Hugging Face*
- **Selective Channel Recalibration Attention (SCRA):** Added to ConvNeXt-Large-V2; improved fine-grained image classification accuracy on CUB Birds, FGVC Aircraft, and FoodX while staying within 5% of baseline FLOPs. *PyTorch, ConvNeXt V2*

Faculty of Engineering – University of Peradeniya
Research Assistant

Sri Lanka
Jan 2021 – Aug 2023

- **Real-time Multiple Dyadic Interaction Detection:** Built a real-time pipeline for detecting co-occurring two-person interactions in crowded surveillance footage, combining YOLOv7 + SORT for spatiotemporal localization, YOLO-Pose for skeleton extraction, and X3D-M + attention for action classification. Published at IEEE ICIIIS 2023. *PyTorch, YOLOv7, SORT, X3D-M*
- **Solar GHI Nowcasting from Sky Images:** Built a physics-informed two-stage system where a classical CV sky-condition classifier routes inputs to regime-specialized ResNet sub-models, achieving 10.80% nRMSE and 98.88% Pearson correlation. Published at IEEE ICIIIS 2021. *PyTorch, OpenCV, ResNet*

EDUCATION

Mohamed Bin Zayed University of Artificial Intelligence (MBZUAI)
Master of Science in Computer Vision

Abu Dhabi, UAE
Aug 2023 – May 2025

- GPA: 3.95/4.0 (Rank 2/36); MBZUAI Full Scholarship (4.6% acceptance rate).
- Relevant Coursework: Mathematical Foundations of AI, Deep Learning, Probabilistic & Statistical Inference, Visual Object Recognition and Detection, Advanced Vision and Language, Human and Computer Vision

University of Peradeniya
Bachelor of Science in Electrical & Electronic Engineering

Sri Lanka
Nov 2017 – Mar 2023

- GPA: 3.95/4.0, First Class Honours (Rank 3/100)

Chartered Institute of Management Accountants (CIMA)
CIMA Diploma in Management Accounting (CIMA Dip MA)

UK
Nov 2016 – Feb 2018

PUBLICATIONS

- **Promptception: How Sensitive Are Large Multimodal Models to Prompts?** *EMNLP 2025 (Findings)*. [\[Paper\]](#)
A 61-prompt evaluation framework across 15 categories measuring prompt sensitivity in 10 LMMs on 3 MCQA benchmarks, revealing up to 15% accuracy deviation from phrasing alone.
- **Towards Calibrating Prompt Tuning of Vision-Language Models.** *CVPR 2026*. [\[Paper\]](#)
Dual-regularization calibration framework (mean-variance margin loss + text moment-matching) for prompt-tuned CLIP, reducing Expected Calibration Error across 11 datasets and 7 prompt-tuning methods without sacrificing accuracy.
- **Real-time Multiple Dyadic Interaction Detection in Surveillance Videos in the Wild.** *IEEE ICIIIS 2023*. [\[Paper\]](#)
Real-time multi-dyadic interaction detection in crowded surveillance video via a skeleton-based X3D-M model with attention.
- **Global Horizontal Irradiance Modeling from Sky Images Using ResNet Architectures.** *IEEE ICIIIS 2021*. [\[Paper\]](#)
Hybrid classification-regression framework combining physics-based sky-condition labeling with condition-specific ResNet nowcasters for solar irradiance prediction.

SKILLS & LEADERSHIP

GenAI: LangGraph, LangChain, LlamaIndex, CrewAI, Model Context Protocol (MCP), Langfuse, LLM APIs, Retrieval-Augmented Generation (RAG), structured prompting, LLM-as-Judge evaluation, AI agents

Deep Learning & ML: PyTorch, TensorFlow, OpenCV, Hugging Face, CUDA, distributed multi-GPU inference, PEFT, vision-language models, CNNs, AutoML (PyCaret), XGBoost, scikit-learn, K-means clustering, pandas, NumPy

Full-Stack: Python, FastAPI, Pydantic, Celery; PostgreSQL, pgvector, Qdrant, Redis, SQL; React, TypeScript, MapLibre, CesiumJS, H3, Streamlit; Docker, AWS (SageMaker, S3), Git, GitHub Actions, pytest, Playwright, $\LaTeX$

AI Coding Tools: Claude Code, Codex, GitHub Copilot, Cursor

Competitive Programming: IEEEExtreme 15.0 (2021) - Global Rank 161/2,403, Country Rank 9/300

Founding President - MBZUAI Consulting Club : Established the UAE's first Inter-university case competition.