

TANISHK RANE

+91 7030597486 ♦ tanishk.rane@gmail.com ♦ [linkedin.com](https://www.linkedin.com/in/tanishk-rane) ♦ [medium.com](https://medium.com/@tanishk_rane) ♦ [github.com](https://github.com/tanishk-rane)

SUMMARY

Turning raw ideas into smart systems—one model, one prompt, one pipeline at a time. With 3 years of AI/ML experience and a B.E. from PICT, I build AI solutions that do more than just talk, and systems that actually ship. Part engineer, part AI architect and Data strategist—on a mission to build tech that's as useful as it is intelligent.

TECHNICAL SKILLS

- **Languages:** Python, SQL, C++, HTML, CSS
- **Databases:** MySQL, MongoDB, PostgreSQL
- **Frameworks:** PyTorch, TensorFlow, FastAPI, Flask, Gradio, Streamlit, Langchain, FastMCP
- **AI ML Techniques:** Agentic AI, CNNs, LLMs, RAG, NER, Anomaly Detection, Time Series Forecasting, Prompt Engineering
- **Libraries:** OpenCV, Mediapipe, Matplotlib, Seaborn, Pandas, NumPy, PySpark, Scikit-learn
- **Tools:** Git, Docker, Ngrok, HuggingFace Transformers, OpenAI
- **Cloud Platforms:** AWS (EC2, Lambda, S3, Glue, Step Functions, DynamoDB, Athena, RDS, Bedrock), Azure (AI Search, AI Foundry, Functions, CosmosDB)
- **Specialization:** Deep Learning, GenAI, Natural Language Processing(NLP), Computer Vision, Data Engineering, Machine Learning

EXPERIENCE

AIML Engineer II

Ciklum India(Previously Infogen Labs)

July 2022 - Present

Pune, Maharashtra

- Conducted research, developed **Proof of Concepts**, trained Juniors and delivered efficient **AI/ML Solutions**.
- Designed and deployed **end-to-end Data and AI architectures** including **model development**, training, fine-tuning and **cloud deployment** on platforms like **AWS** and **Microsoft Azure**.
- Built **conversational and voice AI applications** using **LLMs (GPT, LLaMA2)** and **Whisper**, boosting efficiency by **30%** through custom audio processing and **prompt engineering** techniques.
- Built **intelligent AI agents** and integrated **RAG pipelines** using **Azure platform** and **OpenAI APIs**.
- Deployed **ML/DL models** using **FastAPI**, **automating workflows** reducing manual effort by **90%**.
- Designed scalable **ETL/ELT pipelines** with **PySpark** and **AWS Services**, cutting data processing time by **80%**.
- Developed and fine-tuned **U²-Net, YOLOv5 models** for **background segmentation**, achieving **90%** accuracy.
- Automated onboarding workflows using **metadata-driven design**, reducing setup time from 6 weeks to 10 days.
- Developed an **Agent to convert natural language into PostgreSQL queries** using **LLMs (LLaMA, DeepSeek)** with **R&D on optimal LLMs and workflows**; added **chart recommendation module** and deployed it on **Azure**.
- Wrote **clean, secure, and maintainable code** following **coding practices** and **documentation standards**.

PROJECTS

Aptara - Storyboard Agent

- Orchestrated the end to end construction of a **RAG based AI Agent** utilizing **Azure AI Search** and **Azure Foundry**, achieving a **90% relevance score** in storyboard content generation and reducing research time.
- Automated agent workflow using Python and **Azure SDKs**, reducing manual configuration efforts by **75%**.
- Performed **prompt tuning**, **prompt engineering**, and parameter tuning to optimize content quality and efficiency.
- Finetuned **GPT-4o model** using **Azure AI Foundry** with JSONL format which helped to increase the output accuracy by **15%**.

AI Based Interior Design Application

- Contributed to the development of an **AI-powered interior design app** for Android and iOS, featuring **style transfer**, **paint simulation**, **object replacement**, and **design rendering** using **diffusion models** and **VLMs**.
- Conducted **RnD** to identify optimal models for each feature , resulting in **25% latency reduction** and smoother pipeline performance using **Stability AI** and **OpenAI models**, deploying inference endpoints on cloud.
- Boosted **visual realism** by **30%** through prompt engineering, **hyperparameter** tuning and workflow optimization.

School Data Solutions

- Led the **customer onboarding and integration** of K–12 school data into a **metadata-driven ETL pipeline**, processing student performance data, reducing customer onboarding time from 6 weeks to 10 days.
- Achieved **82% accuracy** in student risk prediction using **machine learning models**.
- Automated **data validation**, **quality reporting** using **Jira Automations**, and **scalable delta processing** for daily updates, reducing latency by more than **60%**.

FitMatch.ai

- Designed and implemented **deep learning solutions** using **CNNs** and **3D point cloud processing** for **human body part classification**, **segmentation** and **regeneration**, with accuracy of **90%+**.
- Processed and normalized **3D data (PCD, Meshfiles)**; built **model pipelines** and deployed using **AWS Lambda** for **iOS AppClip** integration.
- Optimized **deep learning models** for **serverless deployment** by applying **pruning techniques** and converting formats (**ONNX**, **TFLite**), enabling efficient execution on **AWS Lambda**.
- Developed CV modules including **A-pose detection**, **Rotation detection**, **Human detection** and **Hair Detection**.

Pixlit

- Developed an **AI-powered solution** for processing **images**, **videos**, **audio**, and **PDFs** with **deep learning models** achieving over **90% accuracy**.
- Fine-tuned **NLP models** for **multimedia parsing**, reaching **95%+** accuracy on audio, video, and document data.
- Built and optimized **computer vision algorithms** and **post-processing scripts**, improving output **quality and consistency** by **30%**.
- Deployed **scalable model APIs** on **AWS** and **on-premise servers**, reducing **inference latency** by **35%**.

VoiceCity

- Collaborated in the development of a **state-of-the-art AI-based Voice Cloning software** for applications such as **Text-to-Speech (TTS)** and **Speech-to-Text (STT)**.
- Fine-tuned **deep learning models**, including **GAN** and **transformer-based architectures**, for **high-fidelity voice replication**, achieving **naturalness** and **speaker similarity scores** above **90%**.
- Engineered **model ensemble strategies** to combine multiple outputs, enhancing **audio quality** and reducing **synthesis artifacts** by **25%**.

PUBLICATIONS

-
- | | |
|---|------|
| • An Augmented Reality Based System for Fashion and Beauty Makeover
Proceedings of the IEEE IISec 2022 Conference | 2022 |
|---|------|

EDUCATION

Pune Institute of Computer Technology, affiliated to SPPU, Pune Bachelor of Engineering - Information Technology CGPA : 9.87 (Institute Topper)	July 2018 - June 2022
---	-----------------------