

Chinmay Sultanpuri

Bachelor of Engineering, KLE Technological University

[LinkedIn](#) • [GitHub](#) • csultanpuri@gmail.com • + 91 7676953569

ABOUT ME

Final-year CS(AI) undergraduate at KLE Tech with a focus on AI systems, evaluation, and infrastructure.
Led 3 published research projects with 5 more in progress, and mentored 100+ juniors on research workflows.

EDUCATION

Degree	Institute/Percentage	CGPA	Year
B.E.	KLE Technological University, Hubballi	9.57/10	2022-2025
Class 12	Prism PU Independent College, Dharwad	92%	2022

RESEARCH AND INDUSTRY EXPERIENCE

ELECTRIC VEHICLE INNOVATION CENTER, KLE TECH

Vidhyanagar, Hubballi

Undergraduate Research Student

July 2025 - Present

- Selected for REU; **only CSAI student** from department chosen. Working under Dr.Kiran Patil.
- Investigating optimal EV charging station placement in power distribution networks to minimize grid congestion and cost using genetic algorithms.
- Exploring load forecasting, heuristic placement, and power flow analysis under real-world constraints

KNIT SPACE: ACTION-BASED LLM TESTING HARNESS ([CERTIFICATE](#), [GITHUB](#))

Sai Nagar, Hubballi

AI Engineering Intern

February - June 2025

- Developed a novel evaluation framework for LLMs inspired by **Lagrangian mechanics**, modeling language generation as motion through latent space.
- Defined **semantic velocity**, **kinetic energy**, **token surprisal**, and a unified **Action (S)** metric to assess fluency, compute efficiency, and solution elegance across 239+ models.
- Introduced **Efficiency of Language Output (ELO = 1/S)** as a model-agnostic diagnostic, penalizing verbosity and reward multi-path solutions in tasks like math, reasoning, and hallucination detection.
- Deployed as a modular benchmarking tool supporting **major model APIs** (OpenAI, Claude, Gemini, DeepSeek, Mistral) and inference platforms (HF, OpenRouter).

PROJECTS

YET ANOTHER GAME-N-GEN: ACCEPTED IN 9TH INT'L JOINT CONF. ON ADVANCES IN COMPUTATIONAL INTELLIGENCE

[Github](#)

- Designed the **first unified simulator** for multi-game generation using a modified **Stable Diffusion v1.5**, learning frame transitions for Chess, Game of Life, Snake, and Car Avoidance without game engines or symbolic rules.
- Reformulated next-frame prediction as a **self-supervised latent prediction task**, conditioning on prior frame + action token, avoiding expensive diffusion denoising while preserving 29.8 dB PSNR.
- Built and open-sourced a **670GB+ dataset**, using **RL agents (DQN)** for environment interaction, custom tokenization, and domain-adapted VAE decoder finetuning.
- Led a 4-person team, implemented modeling pipeline, managed training on DGX H100s, and published pretrained models on GitHub + Hugging Face.

MOBILE-FRIENDLY MAXViT :ACCEPTED IN 9TH INT'L JOINT CONF. ON ADVANCES IN COMPUTATIONAL INTELLIGENCE [Github](#)

- Proposed the first MaxViT-based supernet compatible with the Once-for-All (OFA) NAS framework for efficient mobile medical imaging.
- Trained a MaxViT-Small teacher on the HAM10000 dataset using imbalance-aware loss; performed knowledge distillation to student models with 66.7% compute reduction, while maintaining 82.74% macro F1.
- Deployed distilled models to iOS via CoreML; open-sourced full pipeline (PyTorch). Led a 4-member team across data prep, NAS search, distillation, and deployment on DGX H100.

GAN-BASED RIB FRACTURE SEGMENTATION: PUBLISHED AT CRM 2025, WARAGNAL, INDIA

- Led team to train a SwinNet3D-based GAN on RibFrac; achieved 71.8% Dice, surpassing State-of-the-Art by 4%.
- Led a 4-member team, deployed training on **DGX Tesla V100** infrastructure.

KLEIN: GLOBALLY DISTRIBUTED AI INFERENCE : ([PORTFOLIO](#)) [Github](#)

- Designed a self-hostable AI inference platform architecture with support for versioning, multi-modal inputs, and scalable deployment (load balancing, Merkle trees, LSM caches, etc.).

PUBLICATION & PREPRINTS

- *A Diffusion-Based Interactive Simulator for Multi-Game Generation*, **IJCACI 2025** – First author
- *Mobile-Friendly MaxViT: Efficient Skin Lesion Classification via NAS and Distillation*, **IJCACI 2025** – Second author
- *Transformer-GAN Enhanced Rib Fracture Segmentation: Integrating SwinUNET3D with Adversarial Learning*, **CRM 2025** – Second author
- 5 additional manuscripts in progress (all first-author), including 2 to A* venues.
- [View all Papers & Preprints](#)

MENTORSHIP & TEACHING EXPERIENCE

Mentor – EDA Course & WiDS Datathon

July 2025 - Present

- Developed instructional materials for the Exploratory Data Analysis (EDA) course; authored Python notebooks and lab templates for undergraduate instruction.
- Mentored **100+** juniors during WiDS 2025 via model evaluation walkthroughs, code reviews, and data prep sessions — mentee teams placed **4th, 8th, and 15th globally** in the final leaderboard.

SKILLS

Languages: Python, C, C++, SQL (MySQL)

Frameworks & Libraries: PyTorch, TensorFlow, Hugging Face, NumPy, scikit-learn

Concepts: OOP, Data Structures & Algorithms, Web Development, REST APIs

Tools & Deployment: Docker, Git, CoreML (iOS), Linux, DGX H100/V100 servers

Concepts: LLM Evaluation, Diffusion Models, Semantic Search, Distributed Systems, Model Compression

ACHIEVEMENTS

- Selected as 1 of 8 interns (and only CSAI branch student) at KNITSPACE, a highly selective research company (150+ applicants).
- Maintained highest CGPA in CSAI for 6 consecutive semesters; only student permitted CSAI branch transfer
- WiDS Datathons: Global Ranks – 2024 (**#14**), 2025 (**#38**); All India Ranks – 2024 (**#2**), 2025 (**#1**)