SUJAL T S

+91 9353750004 sujal2004ts@gmail.com

LinkedIn GitHub LeetCode GeekForGeeks

Education

- **KLE TECHNOLOGICAL UNIVERSITY** *Hubli, India*
  B.E in CSE-AI | Current CGPA: 8.68/10.00 (2022 - 2026)
- **PACE UM PU COLLEGE** *Shivamogga, India*
  Pre-University Education | Percentage: 96.00% (2020-2022)

Experience

- **Amazon Machine Learning Summer School 2025** *(Aug 2025 – Present)*
  Competitive Machine Learning and coding program organized by Amazon scientists

- Selected among the **top 3,000 candidates nationwide** through a rigorous ML & coding assessment conducted by Amazon.
- Currently undergoing training in **Machine Learning, Deep Learning, Generative AI, and Probabilistic ML**, with hands-on coding challenges and real-world problem-solving sessions.
- Collaborating with peers and mentors to implement **industry-grade AI solutions** and participate in live coding assessments.

Technical Skills

- **Languages:** C++, Python, C, SQL
- **Frameworks & Libraries:** PyTorch, OpenCV, NLTK, Pymoo, ExpressJs, NodeJs
- **Tools & Platforms:** Docker, Git, GitHub, Linux, Cisco Packet Tracer
- **Databases:** MySQL
- **Soft Skills:** Problem Solving, Leadership, Team work, Self-learning, Risk Management, Adaptability

Publications

Dravidian T5 - Machine Translation on Dravidian Languages
Accepted for Springer presentation at the 5th International Conference on Emergent Converging Technologies and Biomedical Systems(ETBS 2025)

KnParaphraser - Kannada Paraphrase Generation
Accepted for Springer publication at the 12th International Conference on Signal Processing and Integrated Networks (SPIN 2025)

XplainOA - Knee Arthritis Detection with XAI
Accepted for Springer presentation at the 5th International Conference on Emergent Converging Technologies and Biomedical Systems(ETBS 2025)

Projects

- **Neural Machine Translation for Dravidian Languages**
  Developed a robust Machine Translation system for low-resource Dravidian languages

- Developed a Dravidian-specific **T5 Transformer** model using a pivot-based data augmentation framework combined with diverse decoding strategies to enhance multilingual translation quality.
- Created a high-quality parallel corpus **DraPara** of **100k sentence pairs** across Kannada, Tamil, Telugu, and Malayalam by leveraging model-based filtering techniques using BLEU and Sementic similarity thresholds.
- Trained and optimized the Dravidian-T5 model, reducing the original mT5 size by **over 20%**, resulting in **23% fewer FLOPs per inference**, enabling efficient deployment on resource-constrained hardware.
- Achieved an average **BLEU** score of **0.69** and **similarity of 0.7764** on unseen datasets; released both the Dravidian-T5 model and the DraPara corpus publicly on Hugging Face.

• **Neural Architecture Search for Image Deblurring**

Multi-objective NAS framework for optimizing PSNR vs. FLOPs

- Developed a **genetic encoding scheme** to represent architectures and implemented a decoding mechanism to convert them into trainable deep learning models.
- Designed novel **crossover (Uniform + SBX)** and **mutation (Swap + Increment-Decrement)** operators to enhance search efficiency.
- Optimized architectures using **NSGA-II**, balancing performance (**PSNR**) and computational cost (**FLOPs**) for efficiency.
- Currently evaluating on the **GoPro** image deblurring dataset, aiming for competitive results against **state-of-the-art models**.

• **Automated Knee Arthritis Detection Using X-Ray Imaging and Explainable AI**

Deep learning project for multi-stage OA grading with mobile deployment and XAI

- Developed a pipeline to classify knee osteoarthritis severity into **KL grades 0–4** using **9,786 X-ray images** from the OAI dataset.
- Fine-tuned an **Xception CNN** with augmentation and class weighting, achieving **71% accuracy** and balanced accuracy of **0.74**.
- Integrated **Grad-CAM** to visualize key joint regions, improving model interpretability for clinical use.
- Optimized the model with **TensorFlow Lite quantization**, reducing size by **75%** and accelerating inference by **30%** for mobile deployment.

• **Women Employment Web Portal**

Developed an AI-powered job portal to assist women in finding employment opportunities

- Developed a full-stack job portal using Nodejs, Expressjs, and integrated AI-based filtering with the **LinkedIn Jobs API** to automate retrieval and recommend opportunities.
- Implemented AI-driven job matching using **OpenAI** to filter jobs by salary, experience, and age range, designing an NLP-based system.
- Conducted **Selenium-based testing**, verifying search functionality, API responses, and AI filtering, achieving a **90% pass rate** post-optimization.
- Enhanced response handling for a more efficient user experience, ensuring seamless integration between **AI recommendations** and real-time job listings.

• **Kannada Paraphrase Generation (KnParaphraser)**

Developed a novel framework for generating diverse paraphrases in low-resource Kannada

- Developed a Kannada paraphrasing model **KnParaphraser** using a novel data augmentation framework that combined round-trip translation with advanced sampling methods.
- Created and publicly released **KanPara**, a high-quality corpus of **100k paraphrase pairs** for low-resource NLP, generated via NLLB-200 backtranslation and rigorous alignment.
- Fine-tuned and pruned the T5 model by **50%**, optimizing it for Kannada to enable efficient training while preserving paraphrase fidelity.
- Employed multiple decoding strategies (Greedy, Beam, Top-K, Top-P, Combined Sampling) with semantic filtering to achieve paraphrases that balanced lexical diversity and meaning preservation.

Achievements

- Shortlisted for the technical interview round in **HackWithInfy 2025**, among **2,00,000+ participants** in Infosys's national-level coding competition.
- Released open-source AI models on Hugging Face, including **Dravidian-T5**, **DraPara**, and **XplainOA**, contributing to multilingual NLP and medical XAI research.
- Solved **500+ coding problems** across LeetCode and GeeksforGeeks, strengthening skills in data structures and algorithms.