SUDIP DAS

· sudipdsf@gmail.com · (+91) 7980633486 · https://www.linkedin.com/in/sudip-das-0

EDUCATION

M.Tech in Computer Science   Gandhinagar, Gujarat
IIT Gandhinagar, *GPA: 8.44*   August 2020 - August 2022
B.Tech in Computer Science   Salt Lake, West Bengal
Techno Main Salt Lake, *GPA: 8.35*   August 2016 - August 2020

PUBLICATIONS

- Athulkumar R, Aarchi Agrawal, K S Ashin Shanly, **Sudip Das** and Nidhin Harilal
  Image Caption Generator using Siamese Graph Convolutional Networks and LSTM
  ACM IKDD CODS 2022

- Vridhi Rohira, Shivam Chaudhary, **Sudip Das**, Krishna Prasad Miyapuram
  Automatic Epilepsy Detection from EEG signals
  ACM IKDD CODS 2023

- **Sudip Das**, Pankaj Pandey, Krishna Prasad Miyapuram
  Improving self-supervised pretraining models for epileptic seizure detection from EEG data
  Master Thesis project (Pre-print)

SKILLS

Programming & Scripting:	Python (Advanced), Bash, Jupyter, Git (BitBucket, GitLab), RESTful APIs
AI/ML Frameworks:	PyTorch, TensorFlow, scikit-learn, vLLM, TGI, Ollama, Hugging Face Transformers
Generative AI:	LLM Pretraining & Finetuning, Prompt Engineering, RAG, Langchain, LangGraph, LangSmith, Agentic Systems, Model Evaluation, SOTA Open-source LLMs, Model Context protocol
NLP:	Transformers, Tokenization, Named Entity Recognition, Text Summarization, Guardrails, Evaluation Pipelines
Computer Vision:	Layout Parser, Object Segmentation, Vision Transformers, OpenCV, OCR
Data Engineering:	Data Pipeline Design, Dataset Management, ETL, Feature Engineering, Data Validation
Database & Messaging Queues:	PostgreSQL, MySQL, MongoDB, ElasticSearch, Qdrant, RabbitMQ
Cloud & MLOps:	Docker, Kubernetes, Azure AI Foundry, AWS, S3, SQS, EC2, Bedrock, Sagemaker, CI/CD, MLflow, Model Deployment, Monitoring
SDLC & Leadership:	Agile (Scrum, Kanban), Technical Mentorship, Project Management, Cross-functional Collaboration

INDUSTRY EXPERIENCE

smartSense consulting solutions Pvt. Ltd.   Gandhinagar, India
Associate AI/ML Engineer → AI/ML Engineer I → II   June 2022 – Present

AI/ML Engineer II   April 2025 – Present

- **AI-Powered Legal Document Automation Platform (Outsourced.co, NZ LegalTech SaaS)**
  Tech Stack: Python, FastAPI, MongoDB, Milvus, Azure AI Foundry

  - Developed and deployed an **LLM-driven platform** automating the generation of affidavits, statements of claim, and legal notices, streamlining compliance for New Zealand legal workflows.

  - Designed and implemented a robust **preprocessing pipeline** to extract data from PDFs, HTML, and other formats, feeding structured information into a reasoning model for advanced context understanding.

  - Engineered and productionized a **modular prompt templating system** leveraging structured assets (guides, examples, templates, guardrails) to dynamically assemble context-rich prompts for the LLM, ensuring high-quality and legally compliant document generation.

  - Conducted extensive R&D and benchmarking to select the optimal reasoning model, ultimately choosing **O4 Mini** for its superior performance in legal document automation.

  - Implemented a **Milvus-powered RAG workflow** for semantic invoice search, enabling rapid retrieval and contextualization of legal and financial documents.

AI/ML Engineer I   July 2023 – April 2025

- **Graph-Based AI Agents (Second Brain, Internal B2B Product)**
  Tech Stack: LangGraph, Langchain, FastAPI, Celery, Azure ML Studio, ElasticSearch, MongoDB, Qdrant, RAGAS

  - Architected and implemented Leave and Onboarding Agents using LangGraph, orchestrating multi-agent workflows based on user queries.
  - Designed an AI-powered knowledge management system that organizes departmental documents into secure vectorized knowledge bases, also known as brains.

- **Generative AI-Powered PDF2HTML for Legal Documents (New Zealand Legal Firm)**
  Tech Stack: EC2, Detectron2, Anthropic Sonnet 3.5, Gemini 1.5 Flash

  - Developed a GPU-accelerated, on-premise OCR pipeline using Layout Parser to convert **100K+ legal documents** from PDF to structured HTML, enabling downstream automation and analytics.
  - Leveraged **Gemini 1.5 Flash** to further process extracted tables of contents into precise HTML code, and analyzed images to generate rich metadata for legal workflows.
  - Achieved significant operational efficiency and delivered **$400K+ in cost savings** by replacing proprietary cloud solutions with an optimized on-premise pipeline.

- **Patent and Image Drafting Automation (Client: One of the World's Largest Chemical Companies, Germany)**
  Tech Stack: vLLM, QLoRA, Finetuning, Pretraining, Prompt Engineering, Model Benchmarking

  - Led a research initiative to enhance patent drafting efficiency by finetuning open-source LLMs (Falcon, LLaMA 2, Mistral, Mixtral) for legal text generation.
  - Increased context window of LLMs with SOTA methods (YARN, RoPE), enabling concise and effective patent drafts and boosting attorney productivity by **30%**.
  - Designed an LLM-powered system to generate accurate SVG diagrams from text, achieving **94% accuracy** and automating image drafting.

Associate AI/ML Engineer   June 2022 – July 2023

- **Data Science Platform for Banks, Professionals & Remote Nodes (Bangalore Data Science Startup)**
  Tech Stack: FastAPI, PostgreSQL, MongoDB, TensorFlow, PyTorch

  - Led a team of 5 to develop a distributed ML training platform, connecting banks, data scientists, and remote nodes.
  - Conducted hackathons with non-sensitive datasets, selecting top talents for banks' AI initiatives.
  - Implemented secure dataset access using a private blockchain-based validation.

- **IntelliBot - Scalable Conversational AI Bot (Client: One of the World's Largest Chemical Companies, Germany)**
  Tech Stack: RASA, Hugging Face, MongoDB

  - Built a transformer-based chatbot handling **50+ intents** with secure departmental access controls.
  - Developed an ML-driven spell checker integrated with NER models, efficiently managing **2000+ entities**.

ACHIEVEMENTS

All India Rank 8 JEST 2020	IMSc
Also cleared Round 2	March 2020
All India Rank 616 GATE 2020	IIT Delhi
With a percentile of 99.4%	April 2020
NPTEL Course DBMS	IIT Kharagpur
Gold Medalist	April 2019