

EDUCATION

Birla Institute of Technology and Science, Pilani - Hyderabad Campus

Work Experience

Founding Developer – Zenalyst.ai, Bangalore

May 2025 – Present

- Led a **20-member tech team**; delivered a **POC in 7 days** and **MVP in 15 days**, accelerating **GTM by ~3x**.
- Architected and deployed a scalable LLM analytics platform** on **AWS (EC2/ECS/EKS)** serving financial insights across **microservices** with **>99% operational uptime**.
- Built and integrated an **LLM caching layer**, **reducing monthly inference cost by 10%** across high-frequency workloads.
- Engineered the External KnowledgeBase system that scraped, cleaned, and structured unbounded external sources into **RAG-ready datasets**, **increasing retrieval accuracy by ~25%** (internal benchmark).
- Fine-tuned multiple **Financial LLMs** (LoRA/QLoRA + SFT) for **reasoning**, **SQL execution**, and report interpretation, achieving measurable gains in accuracy on internal QA + NL \leftrightarrow SQL tests.
- Defined and implemented the MVP & Pilot architecture**, enforcing **secure** data flows, service boundaries, and reproducible ML pipelines.
- Conducted a **company-wide hackathon**, designing problem statements & evaluation criteria; resulted in 6+ strong candidate leads and improved team cohesion.

LLM Systems, Retrieval & Research Engineering

- Integrated **GraphRAG**, **knowledge-graph** relationships, and **structured retrieval pipelines** enabling multi-hop reasoning over financial documents.
- Optimised context-window utilisation**, **chunking strategies**, and embedding **reranking**, reducing retrieval latency and **improving LLM response relevance** in long-form financial analysis.
- Developed FE prototypes (Streamlit) and BE (FastAPI microservices) for pilot clients; implemented secure sessioning, audit logging, and horizontal scalability.
- Automated ML workflows using Docker + GitHub Actions, enabling reproducible training runs, CI for data pipelines, and frictionless deployment.
- Reviewed 20+ **research papers** across LLM reasoning, tabular QA, financial modelling, and other in-code optimization techniques.

Summer Research Intern – IEEE SPS, Gujarat Section

May 2025 - Present

- Conducted literature review of 8-10 **Research papers** on **Self-Supervised Learning (SSL)**, identifying **Barlow Twins**' 18-25% superior performance in **low-data regimes** compared to **SimCLR** and **BYOL** in **Medical imaging applications**.
- Established baseline **Faster R-CNN** model on **VinDr-CXR** dataset (18,000 annotated images) achieving **0.32 mAP**, with **SSL-pretrained** backbone showing **15-25% improvement** over random initialization in preliminary tests.
- Pretrained a **Barlow Twins** encoder on **112,120 unlabeled CheXpert** chest X-rays using domain-relevant **augmentations** (crop, blur, solarize), achieving **>90%** similarity between augmented view embeddings.

Data Analyst Intern – Schiffer and Menezes, Goa

May 2024 - July 2024

- Increased machine uptime by **8%** by developing a **predictive maintenance** model using **XGBoost**, enabling early detection of potential breakdowns based on historical sensor data.
- Processed and **analyzed over 30,000 equipment logs** using **Python** (Pandas, NumPy, Matplotlib) to identify patterns and correlations in failure events.
- Conducted **data analysis** to identify and correlate patterns, enabling predictive maintenance strategies.

Projects

Developing Banking Score from Imbalanced Dataset

Feb 2025 - Mar 2025

- Performed **data cleaning** and **imputation** using **KNN** with GPU acceleration, handling missing values and irrelevant columns.
- Applied **Random Forest Feature Selection** and **SMOTE** to address feature relevance and class imbalance.
- Trained an **XGBoost** classifier on a large dataset with nearly **1,200 columns and 97,000 rows**, achieving an **accuracy of 99%** in predicting the binary target variable **bad_flag**.

Momentum vs Value: Factor Investing Analysis for Indian Equity Markets

- Analyzed **459,582 stock price** observations and **5,910 fundamental metrics** across NSE stocks to compare momentum vs value strategies.
- Demonstrated that **momentum outperforms value by 26.46%** with **p = 0.0001**, providing statistically defensible investment insights for markets representing **₹4+ trillion in capitalization**.
- Engineered complete **data pipelines** for cleaning, feature generation, and portfolio construction; **validated** findings using **two independent methodologies** (122-stock and 2,217-stock datasets).
- Delivered reproducible code + documentation via **Python** (pandas, numpy, scipy), **Jupyter**, **Git**, enabling end-to-end factor research replicable by analysts and quant teams.

Football Analysis (Personal Project)

- Developed a robust **Computer Vision** pipeline to accurately **detect** and continuously **track players**, referees, and football in **real-time** football **video** using **YOLO** for dynamic scene understanding.
- Quantified overall team performance by **calculating ball possession percentage** and **identifying key playmakers** through in-depth **data analysis** using **OpenCV** and custom tracking algorithms.
- Employed **K-means clustering** for precise pixel-wise segmentation and utilized advanced optical flow techniques for **detailed motion analysis** of the ball and players throughout the game.

Skills

- **Programming:** Python, C++, JavaScript (Node.js), PHP, OOP, Computer Vision, Object Detection, Predictive Maintenance
- **ML & AI:** Deep Learning, NLP, LLMs, RAG/GraphRAG, Financial QA, Self-Supervised Learning (SSL), CNNs, XGBoost, Random Forest, K-Means, KNN Imputation, SMOTE, Feature Selection
- **Libraries & Frameworks:** PyTorch, TensorFlow, Scikit-Learn, Hugging Face Transformers, spaCy, OpenCV, YOLO, Pandas, NumPy
- **Backend Development:** FastAPI, Flask, Django, Node.js, REST APIs, JWT Auth
- **Web Development:** HTML5, CSS3, JavaScript, React.js
- **Data & Databases:** SQL, MySQL, MongoDB, Microsoft Excel, Data Preprocessing, Data Cleaning, EDA, Feature Engineering
- **Cloud & DevOps:** Docker, GitHub Actions (CI/CD), AWS (EC2/ECS/EKS), Azure, Linux, Ubuntu

Positions of Responsibility

- **Events Head - Student Alumni Relation Cell (SARC):** June 2024 - July 2025
Managed over 20+ events including Alumni talks like Gaurav Chaudhary (**Technical Guruji**).
- **Sponsorship Head - Sanskrit and Foreign Language Club (SaFL)** Aug 2024 - June 2025
Secured and overlooked many sponsorship deals worth **Rs. 2.5 Lakhs** for college events.
- **Hostel Representative, Shankar Bhavan:** Sep 2024 - March 2025
Was elected as the hostel representative of Shankar Bhawan, BITS Pilani, Hyderabad Campus