

PROFILE SUMMARY: Computer Science undergraduate with hands-on experience in **GenAI, LLMs, Multimodal AI, Deep Learning, Data Engineering, Full-Stack Web Development, and Cloud Computing**. Proven ability to design scalable ETL/ELT pipelines, Medallion Architecture data platforms, role-based web applications, and production-ready, real-time AI applications on **Azure and AWS**. Skilled in **Python, SQL, Flask, PyTorch, TensorFlow, PySpark, Snowflake, dbt, and Power BI**, with working knowledge of Prompt Engineering, Conversational AI, and Computer Vision. Familiar with **data governance**, dimensional modeling, and cross-functional collaboration. ACM Chapter Chair; spearheaded 20+ technical events and hackathons.

EDUCATION

SRM Institute of Science and Technology	CGPA: 9.85	Tiruchirappalli, India
B.Tech Computer Science and Engineering (Honours in Quantum Computing)		2024–2028
IIT Madras		Online
B.S. in Programming and Data Science (Specialization in AI and Deep Learning)		2024–Present

SKILLS SUMMARY

Programming	Python, SQL, C++, Bash, HTML, CSS, JavaScript
Generative AI & LLMs	Large Language Models (LLMs), Generative AI, Multimodal AI, Prompt Engineering, Conversational AI, Retrieval-Augmented Generation (RAG)
Deep Learning & Computer Vision	PyTorch, TensorFlow, Keras, CNN, Vision Transformer (ViT), Physics-Informed Neural Networks (PINNs)
Data Engineering	ETL/ELT, PySpark, dbt, Snowflake, Azure Databricks, Azure Data Factory, Delta Lake, Airflow, Data Modeling, Dimensional Modeling, Medallion Architecture, Delta Live Tables (DLT), Apache Spark, Data Lake, Unity Catalog
Cloud	Microsoft Azure, AWS, Cloudflare, cPanel
Databases	PostgreSQL, MySQL, MongoDB, SQLite, SQLAlchemy ORM, Relational Schema Design
Machine Learning	Scikit-learn, Pandas, NumPy
Web & APIs	Flask, Jinja2, Bootstrap 5, React, Node.js, REST APIs, Role-Based Access Control (RBAC), Session Management, MVC / Modular Application Design
Analytics & BI	Power BI, DAX, Power Query, Microsoft Excel
Developer Tools	Git, GitHub, Docker, VS Code, Jupyter Notebook

WORK EXPERIENCE

Research Intern – ANRF Funded Project	Feb 2026 – Mar 2026
Department of Mathematics, NIT Tiruchirappalli	Full-time (Paid)
- Developed and benchmarked Physics-Informed Neural Networks (deep learning models) to solve non-linear partial differential equations including Keller–Segel systems, the heat equation, and one- and two-dimensional Fisher–KPP equations. - Diagnosed training bottlenecks affecting PINN convergence under different non-linear initial conditions, evaluated model accuracy using L2 error metrics, solution visualizations, and residual analysis across multiple training configurations, and collaborated with research scholars to validate and reproduce experimental results. - Designed and executed numerical experiments using PyTorch for spatio-temporal PDEs with automatic differentiation, multiple optimizers, and custom physics-informed loss functions.	
DECCAANOW Corp	Dec 2024 – Present
Junior Data Scientist	Hybrid
- Designed relational database schemas and optimized SQL query performance for AI(LLM) applications, ensuring high data integrity, efficient storage, and scalable database performance. - Contributed to the TNHRCE AI Crowd Management project for Samayapuram Mariamman Temple by designing backend data models and enabling a real-time, AI-driven crowd management system. - Built and managed the online registration platform for RASRANG 2025 and RASRANG 2026, successfully handling registrations for more than 15,000+ users. - Built AI-powered Team & Task Management CRM solutions using Python, SQL, REST APIs, Cloudflare,	

and modern database technologies.

- Collaborated with cross-functional engineering teams to deliver scalable solutions for AI-powered, real-time applications.

ACHIEVEMENTS

Runner-up – MedAI-thon 2025(National Level Med-Tech Hackathon) 2025
Secured 20,000 cash prize for developing an AI-based healthcare solution.

IEEE EPICS – Access & Abilities Project 2025
Selected for IEEE EPICS community project; received \$1,200 funding to develop an adaptive assistive solution for children with autism.

PROJECTS

Trekking Management Application – Capstone, IITM Programming & Modern App Development | [LINK](#) | [Video](#) 2026
Tech Stack: Python, Flask, Flask-SQLAlchemy, Jinja2, Bootstrap 5, SQLite

- Developed a full-stack trekking management application with separate workflows for administrators, trek staff, and trekkers, as the capstone project of the IIT Madras Programming and Modern App Development diploma.
- Implemented role-based access control (RBAC) for user and staff approval, blocking, trek management, staff assignment, and booking administration.
- Built trek search and filtering, booking and slot-reservation logic, booking approval, trek history, profile management, and role-specific dashboards.
- Structured the Flask application into modular authentication, administrator, staff, and common/user route modules with SQLAlchemy models and Jinja2 templates; used server-side sessions and flash messages for authentication, validation, and user feedback.

Realtime Multimodal AI Avatar Platform | [LINK](#) 2026
Tech Stack: Python, React, TypeScript, Node.js, Tavus CVI, LLM

- Built a production-ready multimodal conversational AI avatar using Tavus Conversational Video Interface (CVI) and LLM-based persona grounding, enabling real-time text, voice, and live video interactions.
- Developed a React + TypeScript frontend with live avatar streaming and transcript UI.
- Engineered a secure Node.js backend for conversation orchestration and REST API integration.
- Implemented personalized knowledge grounding from structured Markdown for context-aware responses.
- Designed scalable low-latency real-time session management.

AI Capstone Project - DeepLearning | [LINK](#) 2026
Tech Stack: Python, PyTorch, TensorFlow/Keras, CNN, ViT, Data Augmentation

- Engineered an end-to-end deep learning pipeline that integrated CNN feature extraction with Vision Transformers (ViT) to maximize spatial and contextual representation for land-use classification.
- Designed and trained custom CNN and ViT architectures from scratch, implementing data augmentation, regularization, and dynamic early stopping to reduce overfitting by 15%.
- Optimized hyperparameter tuning sweeps (learning rates, batch sizes, and optimizer schedules), accelerating model training convergence times by 35%.
- Achieved 97.2% classification accuracy on the EuroSAT dataset, benchmarking performance trade-offs between traditional convolutional backbones and transformer-based latencies via detailed confusion matrices.

Scalable Azure Data Engineering Pipeline with Medallion Architecture | [LINK](#) 2026
Tech Stack: Azure, Azure Databricks, PySpark, Python, Delta Lake, DLT, ADLS Gen2

- Built an end-to-end Azure data engineering pipeline using Azure Databricks and PySpark to automate data ingestion, cleansing, and transformation through Bronze, Silver, and Gold Medallion layers, producing analytics-ready datasets.
- Implemented Delta Live Tables (DLT) for pipeline orchestration, incremental data processing, Star Schema models with SCD Type 2 for historical tracking, enforced data governance using Unity Catalog, and optimized the Gold layer for Power BI reporting.

UPI Transaction Analytics Dashboard | [LINK](#) 2026
Tech Stack: Power BI, DAX, Power Query, SQL, Microsoft Excel

- Designed an Power BI analytics dashboard using official NPCI UPI transaction statistics to analyze and monitor India's digital payments ecosystem at scale. Built a dimensional data model, created reusable DAX measures, and designed interactive dashboards featuring KPIs, trend analysis, MoM and YoY growth, rolling averages, cumulative metrics, and transaction distribution insights. Applied Power Query for data prepa-

ration and delivered executive-ready self-service BI reports that enable data-driven analysis of transaction performance across multiple time periods.

CERTIFICATIONS

Advanced Certificate in Programming and Modern App Development — LINK	2025
<i>IIT Madras (Diploma Program) — 1.5 Years</i>	
Foundation in Data Science — LINK	2024
<i>IIT Madras (Foundations Program) — 1 Year</i>	
SAP - Strategic Analysis and Solution Design — LINK	2026
<i>SAP Career Certificates</i>	
Google Advanced Data Analytics Professional Certificate — LINK	2025
<i>Google Career Certificates</i>	
Google People Management Essentials Specialization — LINK	2025
<i>Google Career Certificates</i>	
Tools for Data Science – IBM — LINK	2025
Machine Learning Master & Data Engineering Professional — LINK — LINK	2025
<i>RapidMiner</i>	
Oracle Cloud Infrastructure Certified Foundations Associate — LINK	2026
<i>Oracle</i>	

LEADERSHIP & VOLUNTEERING

Chair – Association for Computing Machinery (ACM) Student Chapter	Mar 2026 – Present
<i>Promoted from Vice Chair (Mar 2025 – Mar 2026) to Chair; currently leading the chapter. Organized 20+ technical events, including workshops, coding competitions, and technical initiatives.</i>	
Mentoring Coordinator – Institution’s Innovation Council (IIC)	Sep 2024 – Sep 2025
<i>Coordinated student mentoring initiatives, innovation programs, startup awareness activities, and technical workshops while supporting entrepreneurship and interdisciplinary collaboration across the institution.</i>	
Joint Secretary – Technical Council, SRM IST - Trichy	Sep 2024 – Sep 2025
<i>Coordinated university-level technical events, student activities, and inter-department collaborations while supporting the planning and execution of academic and technical initiatives.</i>	
Secretary(Hostels) – Student Council(SRMIST)	July 2026 – Present
<i>Represent student residents and coordinate hostel administration, student welfare initiatives, grievance resolution, cultural and technical activities, and inter-hostel engagement while collaborating with wardens and university management to enhance campus residential life.</i>	

CURRENT RESEARCH

KV-Cache Optimization for Efficient LLM Inference	2026
<i>Focus: Transformer Attention, Memory-Bound Inference, PyTorch</i>	
• Studying Key-Value (KV) cache mechanisms in autoregressive transformer inference, analyzing how caching prior attention states avoids redundant recomputation and reduces inference latency in LLMs. • Exploring memory-reduction techniques including Multi-Query Attention (MQA), Grouped-Query Attention (GQA), and quantized KV-cache representations to lower GPU memory footprint during long-context generation. • Analyzing trade-offs between cache eviction strategies (sliding window, attention-sink retention) and generation quality for extended context-length inference.	
Principal Component Analysis for Dimensionality Reduction	2026
<i>Focus: PCA, Feature Engineering, High-Dimensional Data</i>	
• Applying Principal Component Analysis (PCA) to reduce dimensionality of high-dimensional feature sets while preserving maximum variance for downstream model training. • Evaluating explained-variance thresholds and eigenvalue decomposition to balance information retention against model complexity and overfitting risk.	