

Education

Master of Science in Data Science

University of Colorado Boulder, Colorado

GPA: 3.9/4.0

Aug 2024 – May 2026

Bachelor of Technology in Information Technology / Minor in Data Science

Dwarkadas J. Sanghvi College of Engineering, University of Mumbai, India

GPA: 8.93/10.0

Aug 2020 – May 2024

Work Experience

Projxon

Aug 2025 – May 2026

AI Specialist Intern

Remote, USA

- Built automated ETL pipeline from LinkedIn and Windsor.ai APIs into BigQuery across 5 workstreams; designed structured SQL schemas, KPI transformations, and data quality checks powering Looker Studio dashboards.
- Applied Association Rule Mining to 6+ months of behavioral data to surface content engagement patterns, delivering recommendations that reduced manual reporting cycles by 40%.
- Engineered end-to-end HR automation pipeline using Google Apps Script: dynamic document generation, role-based templating, batch PDF processing, and automated email workflows for 3 document types.
- Co-designed OrcaAI, an AWS-powered AI optimization platform for SMBs using predictive analytics, classification, and recommendation systems; secured 2nd place and \$25K in AWS credits at pitch competition.
- Designed Gemini API-powered HR chatbot with structured knowledge integration and prompt engineering, improving response accuracy across 30+ organizational policy categories.

Exaplus Technologies

Dec 2023 – Jul 2024

AI/ML Engineer Intern

- Built LLM-powered document intelligence pipeline using OpenAI API to extract structured JSON from unstructured contracts and financial reports; engineered format-agnostic prompts exposed via FastAPI endpoint, cutting analyst turnaround by 70%.
- Developed automated financial reconciliation system ingesting invoices and bank statements from 3+ sources; applied NLP extraction and anomaly detection to flag mismatches, delivering structured JSON reports to client dashboards.
- Engineered K-Means clustering pipeline segmenting clients by transaction behavior patterns, auto-routing records to report templates across 3+ formats and removing human categorization from the full reporting cycle.

E4 Software Services

Aug 2023 – Nov 2023

Machine Learning Engineer Intern

- Built end-to-end resume parsing pipeline using SpaCy NER and custom statistical NLP logic to extract structured entities across 5 fields, returning normalized JSON outputs validated with precision/recall benchmarking.
- Developed enterprise document parsing pipeline using Tesseract OCR and Canny edge detection to parse structured and semi-structured PDFs into normalized JSON key-value outputs for downstream business reporting.
- Designed preprocessing pipelines for text cleaning, tokenization, and entity validation; built reproducible evaluation scripts and documented notebooks for team handoff.

Technical Skills

Languages: Python, R, SQL, JavaScript, Dart.

ML / AI: Machine Learning, Deep Learning (CNN, RNN, MLP, 3D-CNN), NLP, Computer Vision, LLM Integration, Prompt Engineering, RAG, OpenAI API, Gemini API, Gradient Boosting, K-Means, Naive Bayes, SHAP, LIME, ALE, Scikit-learn.

Cloud & Backend: AWS (Lambda, API Gateway, SQS, RDS, S3), GCP, BigQuery, FastAPI, Docker, REST APIs, CI/CD, ETL Pipelines.

Databases & Visualization: PostgreSQL, MySQL, SQLite, Pandas, Power BI, Tableau, Looker Studio, DAX, GGplot (R).

Tools: Git, Google Apps Script, Liquibase, Agile, System Design.

Projects

Meridian: Marketing Decision Intelligence Platform | Python, Scikit-learn, SciPy, FastAPI, OpenAI API, React, Docker

- Built end-to-end Marketing Mix Modeling platform implementing geometric adstock decay, Hill saturation, and OLS/Ridge regression to decompose revenue into channel contributions with R^2 and MAPE diagnostics.
- Developed SciPy SLSQP constrained budget optimizer under diminishing returns; built LLM analyst agent using OpenAI function calling grounded on verified model outputs with zero hallucination, deployed across 7 REST endpoints.

AgentSquared | FastAPI, Gemini API, RAG, SQLAlchemy, SQLite, Next.js, Vultr

- Built no-code AI agent platform at HackCU 2026 in under 48 hours enabling SMBs to deploy production-ready support agents from website URLs in under 60 seconds, with RAG-based knowledge ingestion and async web crawling.
- Engineered FastAPI backend with JWT auth, Gemini-powered persona generation, and 15+ REST endpoints; deployed multi-agent workflows for customer support, sentiment monitoring, and brand posting on Vultr with Nginx and PM2.

CloudSense | AWS Lambda, API Gateway, SQS, RDS, S3, FastAPI, GPT-4o-mini, PostgreSQL

- Built serverless LLM-powered code review platform with webhook-triggered Lambda pipelines processing 1,000+ commit events asynchronously at sub-2 second latency; integrated GPT-4o-mini for structured code-quality findings.
- Deployed full-stack on AWS with JWT-secured APIs, VPC-connected RDS, S3-hosted frontend, and CloudWatch monitoring with CI/CD throughout the deployment lifecycle.

Explainable AI in Diabetes Prediction | Python, CNN, RNN, MLP, SHAP, LIME, ALE

- Benchmarked CNN, RNN, and MLP architectures on clinical data; applied SHAP, LIME, and ALE to audit model behavior and detect demographic bias; published peer-reviewed research on responsible AI in healthcare.