

Sai Akhil Kogilathota

(934)799-0867 | saiakhilkogilathota@gmail.com | Stony Brook, NY | [LinkedIn](#)

SUMMARY

Passionate Data Scientist with 5+ YOE, driving impactful outcomes through advanced analytics, machine learning, and automation. Proven track record in enhancing decision-making, optimizing operations, and delivering actionable insights across Medical and Fintech, Telecom industries.

WORK EXPERIENCE

Data Scientist, MARC

May 2023 — Present

- Clinical RAG Assistant:** Built and deployed a clinical **multimodal RAG** system using **LangChain, Pinecone, and AWS (EC2, S3)** to extract patient data and augment diagnostic processes, enabling faster and more accurate clinical decision support.
- Multi-Agent Healthcare System:** Developing a **Multi-Agent System** for Patient Condition and Disease Diagnosis with Treatment Planning Using **OpenAI Agents SDK**.
- EHR Data Pipeline Automation:** Built scalable ETL pipelines for electronic health records (EHR) using Python, Spark, and PostgreSQL, ensuring HIPAA-compliant handling of **5M+ patient records with <2s query latency**.
- Patient Monitoring Portal:** Built **Tableau dashboards** integrating EHR and wearable data streams (ECG, glucose, oxygen levels), providing hospital staff with a single pane of glass for patient monitoring.
- Conversational AI Chatbot:** Built a **GPT-4 powered chatbot** with transformer-based topic modeling, automating 92% of query routing and saving **6+ hours/week in manual triage**.
- Trial Recruitment Platform:** Built a swipe-based matching interface and integrated recommendation system, improving match accuracy for rare disease trials, **reducing participant search time by ~40%**, and **accelerating trial enrollment by up to 30%**.

Research Assistant, Stony Brook University, New York, United States

Aug 2022 — May 2023

- Multimodal Research:** Introducing a lightweight probing framework that predicts **hallucinations in VLM** image captioning by analyzing hidden-states pre-generation; achieved strong results on 4 VLMs, culminating in a **first-authored EMNLP 2025 submission**.
- Self-Finetuning LLMs:** Developing a novel framework for **self-finetuning LLMs using Agent-to-Agent (A2A)** protocols orchestrated with **LangGraph** in controlled environments, enabling adaptive model improvement without manual intervention.

Data Scientist, Lendingkart Finance Limited

Jun 2019 — Aug 2021

- LLM Fine-tuning:** Fine-tuned a BERT-Base model on internal search data using **Hugging Face & PyTorch**, improving document retrieval accuracy by **70%**. Deployed the model on AWS SageMaker for scalable, real-time search capabilities.
- NLP Analytics:** Performed sentiment analysis on sales conversations by converting audio to text with **Google STT API**. **Applied NLP techniques** to extract key concerns, reasons for customer drop-off, and actionable insights, enabling targeted interventions for sales teams.
- KYC Automation:** Developed and deployed an OCR-based and face-matching system to accelerate onboarding, reducing onboarding time by **80%**; **leveraged PyTorch, YOLOv8** for real-time face detection, with end-to-end hosting and scaling on AWS.
- Borrower Segmentation:** **Increased campaign conversions by 25%** by uncovering high-value borrower segments using **K-Means clustering**, after identifying and resolving two critical issues through in-depth data analysis.
- Engagement Optimization:** Deployed a personalized discount engine with A/B testing on **10K+ users**, **raising loan completions by 35%** and reducing onboarding drop-offs.

Data Scientist, Comviva Technologies

Aug 2018 — Jun 2019

- Revenue & Retention Optimization:** Built Recharge Likelihood and Churn models with XGBoost on AWS SageMaker, driving over **\$160K USD in revenue uplift** and **retaining 12K+ users** per quarter through targeted engagement.
- Workflow Automation & Dashboarding:** Automated KPI tracking and model scheduling via **Python, SQL, and Power BI**, **saving 30+ analyst hours/month** and streamlining insights delivery across 3 teams.
- Data Infrastructure Optimization:** Redesigned telecom data schema in **PostgreSQL**, **reducing storage by 20%** and **cutting dashboard load time by 60%**, boosting data access for analytics teams.
- Usage Segmentation:** Used LightGBM and Ensemble models to model international and multi-SIM user behavior with **75% accuracy**, optimizing network provisioning and **reducing infra costs by \$50k/quarter**.

PROJECTS

Revize | React.js, Django, SQL, AWS, [Link](#)

- Developed a smart web application for spaced repetition learning, enabling users to add topics, auto-schedule revision dates, track daily tasks, and review progress with flashcard-based content.
- Integrated **OpenAI SDK** for automated summarization into flashcards, and deployed the application on **AWS EC2 with a Postgres database**.

Stock Recommendation System | NLP, LSTM, Pytorch, Galformer, [Link](#)

- Developed a novel **PyTorch**-based architecture leveraging **Galformer**, a specialized transformer for stock systems and **NLP**, to analyze stock sentiment and generate buy/sell recommendations based on purchase price and quantity.
- Deployed the model with a **Streamlit** front end, achieving **15% better** prediction accuracy compared to a baseline **LSTM model**.

EDUCATION

Stony Brook University, Stony Brook, United States - Master's, Data Science

Aug 2021 — May 2023

Ramaiah Institute of Technology, Bengaluru, India - Bachelor's, Electronics and Communication

Aug 2016 — Jul 2020

SKILLS

Programming Languages	: Python, Java, R, MySQL
Software & Technology	: AWS, Sagemaker, MongoDB, Elasticsearch, Docker, Kubernetes, CI/CD
Data Processing & Visualization	: SQL, Spark, MS SQL Server, T-SQL, NoSQL, Tableau, Power BI
Machine Learning Techniques	: Regression, Classification, Decision Trees, Dimension Reduction (SVD, PCA), Deep Learning, CNN
Natural Language Processing	: BERT, Sentiment Analysis, Topic Modeling, Semantic Search, Large Language Models, Hugging Face, NLTK
Statistical Analysis	: Causal Inference, Hypothesis Testing, Chi-Square, T-test, A/B Testing, Experiment Design