

DHANYA SRI KOTHAPALLY

857-492-2032 · dhanyasrikothapally@gmail.com · linkedin.com/in/dhanyasrikothapally ·
dhanyasrikothapally.github.io · medium.com/@dhanyasrikothapally

SUMMARY

Product builder and AI engineer who ships AI products for enterprise clients and owns a product from concept through development and deployment. Brings genuine fluency in how AI products work at the technical level, together with command of the product fundamentals: requirements, roadmap, prioritization, and competitive evaluation. Grounds every decision in customer problems and market signals, takes end-to-end ownership, and communicates with equal clarity to engineers, clients, and leadership. Actively targeting both product management (APM/PM) and AI engineering roles.

EDUCATION

Master of Professional Studies in Data Analytics (Applied Machine Intelligence)

Northeastern University, Boston

Aug 2026

GPA: 3.8/4.0

Courses: Data Science Methods & Tools, Data Warehouse & Mining, Project Management, ML in Finance

Bachelor of Technology in Artificial Intelligence

Anurag University, India

Jun 2024

GPA: 9.2/10.0

Courses: Database Design & Management Systems, Deep Learning, Machine Learning, Data Visualization, Big Data

WORK EXPERIENCE

Data Engineer & MLOps Co-op (Product-focused)

June 2026

Crewasis AI, New York

- Drove the product direction for three client-facing AI products (StatAgent, Winston Pro, Model 3), interviewing enterprise clients to identify their business problems, translating competing needs into prioritized product requirements, and partnering with engineering to deliver features that removed 4–6 hours/week of manual work per client.
- Launched all three products into production using **AWS Lambda, ECS Fargate, Step Functions, S3, and Bedrock**, automating client data workflows end to end and establishing a reusable serverless pattern across the engineering team.
- Owned Winston Pro, a conversational AI product that let clients self-serve answers from their own data; built its **RAG pipeline** by ingesting YouTube transcripts, chunking content, generating embeddings via AWS Bedrock Titan, and storing vectors in PostgreSQL pgvector for semantic search, then ran feedback sessions to refine accuracy and the user experience, driving product adoption and reducing clients' dependence on manual analyst support.
- Built StatAgent to validate whether client-reported pain points were real market opportunities: engineered a **statistical engine using VADER sentiment analysis and TF-IDF keyword extraction** on private TikTok/social datasets to quantify how many users actually faced each problem, producing auditable deliverables where every statistic traced directly to source data, then translated the findings into statistics-backed insights that told clients where demand was strong enough to justify building, tuning, or repositioning a product for market fit.
- Architected Persona Builder for consumer brands needing to know who actually adopts their products: built a **social media segmentation pipeline using LLM-based feature extraction, embeddings, and clustering** on TikTok/Douyin data to generate client-facing persona profiles, analyzed audience demographics and age cohorts to identify the segments driving adoption, and enabled per-client reuse through prompt customization alone.
- Collaborated across data science, engineering, and client stakeholders through each build cycle, documenting product specifications, demoing iterations, and adapting scope as client requirements evolved to keep delivery aligned with business needs.
- Led intern recruitment for the next engineering cohort by screening candidates, conducting technical interviews, and contributing to onboarding curriculum design to scale the data engineering team.

Data Analytics Apprentice

2025

Intelmatix

- Delivered the market analysis behind a consumer brand's store-expansion decisions, analyzing **3M+ SafeGraph mobility records and U.S. Census data** to identify which markets represented the strongest growth opportunity.
- Built a **composite market-scoring model in Python (pandas, NumPy)** using weighted metrics for engagement and visits, ranking candidate locations and uncovering 4 high-opportunity Census Block Groups (CBGs) with strong customer demand, turning raw geographic data into a prioritized, defensible recommendation for where to expand.
- Presented executive-ready findings and **interactive Tableau dashboards** visualizing location footprints, scoring dis-

tributions, and opportunity maps, framing the data as a clear go/no-go decision for non-technical stakeholders, and delivered a business intelligence report with model validation and strategic recommendations for retail site planning.

AI/ML Engineer

Sep 2023

Technologics Global Pvt. Ltd., India

- Developed a high-accuracy multi-class classification model (95%) using **Support Vector Machine (SVM)** on the Iris dataset to categorize flower species based on petal and sepal dimensions.
- Implemented a structured ML pipeline involving **data wrangling, feature selection, and data normalization** on a dataset of 150 samples across 3 classes and 4 features, followed by exploratory data analysis (EDA), model training, and performance evaluation using Python, pandas, scikit-learn, and NumPy.
- Compared model performance across linear, polynomial, and RBF kernels, visualizing decision boundaries using matplotlib to understand class separation, and validated model quality through classification reports and visual insights, reinforcing understanding of supervised learning and non-linear classification techniques.

PROJECTS / PRODUCTS

Resume Intelligence System (Self-Initiated AI Product)

Ongoing

- Owning the full product lifecycle of an AI resume-tailoring product from ideation onward: defined the problem, identified the target customer segment, wrote product requirements after validating the pain point against real market demand, and conducted **market and competitive analysis** (Jobscan, Teal, Resume Worded) with a pricing strategy and ROI rationale, then built a phased product roadmap that prioritizes features by customer value and delivery effort.
- Advancing the product through the development phase of the SDLC, managing the backlog and tracking each stage (requirements, development, QA, and testing) in **JIRA** to keep delivery structured and on schedule.
- Translating product requirements into a working **MVP (Python, FastAPI, RAG with ChromaDB, Claude API)** and defining a validation plan that measures real outcomes (user callbacks) over vanity metrics.

Document Intelligence Automation

Outamiation Externship

- Built a document-intelligence product that automated extraction and validation of key information from 200+ page mortgage documents, targeting the slow, error-prone manual review that bottlenecked lenders' processing, by integrating **OCR, PDF parsing (PyMuPDF), and RAG workflows** that convert large documents into searchable, structured information.
- Designed the product's retrieval pipeline using **LlamaIndex**, applying chunk tuning, metadata filtering, and prompt optimization so answers were accurate and source-traceable, improving the precision and reliability of RAG-based search across large mortgage datasets.
- Benchmarked multiple language models (**Mistral, Phi-2**) on accuracy and cost to recommend the best configuration, then packaged the result as a working demo interface and a stakeholder-facing technical report.

EV Adoption & Charging Infrastructure Optimization

- Analyzed **179K+ EV registrations across 700+ Washington ZIP codes** and Alternative Fuel Station data to find where charging demand outstripped supply, identifying the highest-opportunity markets for new infrastructure.
- Mapped adoption trends from 2015–2024 (including Tesla's 42.94% market share) and built **geographic pressure maps in Tableau**, identifying high-demand but underserved areas (e.g., Mill Creek with 4,209 EVs per station) to visualize where demand was concentrated and underserved.
- Translated the analysis into specific location recommendations for new charging stations, turning raw registration data into a prioritized, decision-ready investment plan.

Mental-Health Emotion & Risk Classification Using NLP and Deep Learning

- Trained a full **end-to-end NLP pipeline** on 40K mental-health Reddit posts, performing advanced data preprocessing, class balancing, and linguistic normalization to enable high-quality cognitive-state inference aligned with emotional-AI and clinical-risk modeling use cases.
- Tested both traditional (**TF-IDF + Logistic Regression**) and deep learning architectures (ULMFiT AWD-LSTM), improving prediction accuracy from 61.75% to 65.31% and generating actionable insights through confusion-matrix-based error analysis and model diagnostics.
- Adopted **transfer-learning workflows**, fine-tuned pretrained language models on domain-specific data, and conducted metric-driven model validation, mirroring real-world development cycles for sentiment analysis, behavior insight, and anomaly detection pipelines.

Conversational AI Project on Financial Data Analysis

BCG GenAI Job Simulation

- Designed a **conversational AI chatbot** that analyzed multi-year SEC 10-K data from Microsoft, Apple, and Tesla to

provide financial insights for strategic decision-making.

- Applied data extraction and analysis techniques using **pandas and Excel** to compute key financial KPIs, including multi-year net income trends, revenue growth, and total profitability rankings.

Malicious URL Detection

- Directed a team project developing an advanced malicious URL detection system using **Convolutional Neural Networks (CNN)** for feature extraction and Random Forest for classification, applying machine learning techniques to address cybersecurity challenges.

Commute Time Optimization Analysis

Hypothesis Testing

- Captured 64 real-time commute samples across two transit modes over two weekdays and performed time-series comparison using **z-tests in R**, confirming that subway travel was 22.7% faster than bus commutes (11.36 min vs. 14.7 min), executing the full data analysis lifecycle including data wrangling, outlier filtering, statistical testing, and visualization to enhance precision and ensure reproducibility.

LEADERSHIP, ACTIVITIES & VOLUNTEER EXPERIENCE

AI Product Reviewer & Writer

Medium · Ongoing

- Write product reviews and head-to-head comparisons of AI tools, evaluating them with a structured method: running the same task across multiple LLMs, stress-testing edge cases, and surfacing non-obvious features.
- Document real user friction points and translate hands-on evaluations into clear, honest write-ups for a general audience, building an evaluator's eye for what makes a product genuinely useful versus well-marketed.

McKinsey Forward Program

McKinsey & Company · June 2026

- Selected for McKinsey.org's competitive 10-week program; completed modules in structured problem-solving, communicating for impact, adaptability & resilience, and digital & AI essentials across self-paced courses and mandatory live webinars, earning a McKinsey.org Forward digital badge upon completion.
- Applied **hypothesis-led problem solving** using the SMART framework, issue decomposition, and evidence-based recommendation development to a real-world case assignment, strengthening the ability to break down ambiguous problems and synthesize findings into actionable decisions.
- Practiced audience-centered messaging, structured storytelling, and responsible GenAI adoption strategies, structuring communication that moves an audience to action and reframing challenges into opportunities.

Student Facilitator, Graduate Leadership Institute

Northeastern University · May 2025

- Competitively selected as a facilitator for a 6-week leadership program; led sessions and workshops for a cohort of 30+ graduate students across 12 leadership competencies.
- Facilitated group exercises and reflective discussions, strengthening communication, active listening, and consensus-building across diverse perspectives.

Proctor, Residential Security Office

Northeastern University · Dec 2024 – Present

- Acted as the first point of contact for 400+ residents and visitors across 428+ hours, fielding requests and resolving inquiries with empathy and clear communication while enforcing access protocols.
- Took initiative to onboard new proctors and surface recurring operational issues to supervisors, escalating incidents through the right channels to drive timely resolution.

Blind Coding Competition Coordinator

Anurag University, India · Mar 2022

- Took end-to-end ownership of a campus-wide coding competition, coordinating venue, registration, and technical setup across multiple stakeholders, and resolved live logistical issues under time pressure, delivering an event that earned a 95% participant satisfaction rating.

SKILLS

Product & Strategy: Product Requirements, User Stories, Product Roadmap, Product Lifecycle / SDLC, Backlog Management, Prioritization, Voice of Customer, Product Adoption, Market & Competitive Analysis, Product Evaluation, Pricing / ROI Analysis, Experimentation & Hypothesis Testing, KPI Dashboards, Customer Insight, Go-to-Market Strategy

AI / ML & Data Engineering: Generative AI, Large Language Models (LLMs), Retrieval-Augmented Generation (RAG), Natural Language Processing (NLP), Deep Learning (PyTorch, TensorFlow), Predictive Modeling, Supervised & Unsupervised Learning, Feature Engineering, Statistical Analysis, Hypothesis Testing, Data Storytelling, Neural Networks, Transfer Learning, Embeddings, Vector Databases (pgvector, ChromaDB)

Cloud & MLOps: AWS (Lambda, ECS Fargate, S3, Bedrock, Step Functions, ECR, RDS, DynamoDB, CloudWatch, Secrets Manager), Docker, boto3, FastAPI, LlamaIndex

Programming & Tools: Python (pandas, NumPy, Matplotlib, seaborn, scikit-learn), R (ggplot2, dplyr), SQL, MongoDB, C, MS Excel, Tableau, Power BI, JIRA, Git/GitHub, MS Office (Word, PowerPoint)

Collaboration & Leadership: Agile, Workflow & Process Management, Cross-functional Collaboration, Requirements Gathering, Stakeholder Communication, Training & Facilitation, Client Service, Leadership, Mentorship

Certifications: Claude 101 (Anthropic), AI Fluency: Framework & Foundations (Anthropic), McKinsey Forward (McKinsey & Company), AI (Cognizant), Cybersecurity (Cisco), MongoDB, Python (Infosys), AI Foundations (IBM)