Adarsh Kuthuru

San Francisco Bay, CA, adarshkuthuru@gmail.com, 857-800-4454, LinkedIn, GitHub, Visa status: H1B, I-140 approved

PROFESSIONAL SUMMARY

Principal Data Scientist and AI Engineer with 9+ years of experience building and scaling production-grade ML, GenAI systems in finance and marketing. Proven record of delivering multi-million-dollar impact through RAG platforms, multi-agentic systems, predictive modeling, and distributed systems. Deep expertise across Python, SQL, PyTorch, Spark, and cloud-native applications on AWS/GCP, spanning research, experimentation, and regulated production. Strong technical leader who translates complex AI into measurable business outcomes while mentoring teams and partnering with senior stakeholders.

TECHNICAL SKILLS

Programming: Python (Scikit-learn, PyTorch), SQL, R, SAS, MATLAB, Shell, Javascript, HTML/CSS, SwiftUI, TeX, VBA
Tools: AWS/GCP suite, Git, Tableau, Snowflake, Databricks, Spark, MLflow, Docker, Streamlit, Flask, REST APIs
AI concepts: Generative AI, Machine Learning (Supervised, Unsupervised, Reinforcement Learning), Deep Learning
Certifications: Deep Learning (Deeplearning.ai), Machine Learning (Stanford University), Agentic Development (Google)

WORK EXPERIENCE

Principal Data Scientist, Capital One | San Francisco Bay, CA **Apr 2025 - present**

- Built and scaled a $2M multi-modal RAG-powered BI assistant using distributed systems for parallel experiments with secure access to LOB-wide big data, achieving 98% model accuracy and reducing model tuning time by 80%.
- Delivered an enterprise-ready multi-agentic system to enhance data quality, feature engineering, and model development; drove a 5% conversion lift and eliminated 70% of ineffective marketing outreach.
- Modernized big data ingestion using IaC-based AWS S3 with enforced IAM and bucket policies, integrating Databricks to enable secure distributed systems and tripling release cadence for ML production workflows.
- Provided technical leadership across cross-functional teams of 5+ members, partnering with stakeholders on scope, roadmap planning, and timelines while accelerating execution and achieving a 95% satisfaction rating.

Senior AI/ML Engineer, Capital One | San Francisco Bay, CA **Sep 2023 - Mar 2025**

- Drove end-to-end machine learning (XGBoost) model development in Python on firmographics and transaction data to predict KPIs, translating scores into go-to-market targeting, increasing annual revenue of a $7B credit card portfolio by 7%.
- Led a $1M cross-functional modernization effort to upgrade PySpark/SQL data pipelines for 80M+ prospects with automated validations, reducing end-to-end processing from 2 weeks to 2 hours, unlocking faster campaign launches.
- Developed a KNN-based predictive model to estimate customer spend potential from transaction data using Python/SQL partnering with stakeholders to activate personalized offers and rewards, driving 15% lift in average monthly spend.
- Redesigned cross-sell targeting using statistical modeling on approval and risk insights; aligned stakeholders on governance and success metrics, increasing monthly applications by 40%.
- Conducted segmentation analysis on 2M+ customers, NLP-based analysis of 100k+ complaints, paired with A/B testing to identify spend attrition drivers, providing actionable insights that reshaped product and marketing strategy.
- Improved predictive modeling performance by driving feature engineering and GPU-based hyperparameter tuning on distributed systems, increasing model accuracy by 50%+ and processing speeds by 10x.

Senior Data Scientist, Discover Financial Services | Boston, MA **Nov 2022 - Aug 2023**

- Drove roadmap planning and execution for complex ML and data science workstreams with clear scope and milestones, clarifying ownership, and removing bottlenecks; shortened delivery timelines by 30%.
- Delivered POCs, detailed plans with effort estimates and timelines, robust technical documentation and tutorials, increasing stakeholder confidence and earned a 90% satisfaction ratings.
- Built a production-ready, regulation-compliant loan payment calculator in Python/SQL modeling 20+ payment scenarios across rate changes, waivers, partial payments, and refunds, eliminating manual handoffs and reducing cycle time by up to 75%.
- Led development and deployment of a RAG-based AI chatbot using embeddings, vector search, and LLM responses to resolve loan application-related questions in real time, reducing application creation-to-submission time by 50%.
- Developed scalable, re-usable Python codebases by enforcing engineering best practices and disciplined code review, reducing bug-related issues by 25%, enabling faster end-to-end project delivery.

Data Scientist, Discover Financial Services | Boston, MA **Dec 2021 - Oct 2022**

- Built and scaled a regulation-compliant GenAI mortgage underwriting software in Python that automates document analysis and triggers rule-based workflow, accelerating application processing by 60%.
- Developed an NLP-based recommendation system using SBERT transformer and density-based clustering to identify similar underwriting cases, improving consistency and reducing onboarding time for new underwriters by 30%.
- Optimized end-to-end ticket workflows in GitHub/Jira with clear descriptions, ownership, resources, test evidence, and acceptance criteria, enhancing release management and reducing delays by 25%.

- Strengthened model and data reliability through rigorous validation and concise stakeholder reporting, supporting 24/7 production stabilization and reducing critical incidents by 25%.

Graduate PhD Researcher, 74 Capital Management (Boston College) | Boston, MA **Aug 2019 - Jun 2021**

- Built an end-to-end NLP and causal modeling pipeline using sentiment extraction and instrumental variable regression in Python to isolate true news shocks and convert signals into a systematic trading strategy, delivering 50% annualized returns.
- Developed a fund manager skill score using unsupervised machine learning with PCA and DBSCAN clustering on financial risk exposures, translating clusters into an investable strategy that delivered 30% annual ROI.
- Led end-to-end statistical modeling with Difference-in-Differences and Propensity Score Matching to isolate true treatment effects from selection bias, implementing portfolio controls that mitigated 20% annual downside.
- Transformed Econometrics and Data Analytics curriculum for 500+ MBA/PhD students by integrating Python-based ML workflows (scikit-learn, statsmodels) and real-world case studies in predictive modeling, achieving 5/5 ratings.

Senior Research Associate, Indian School of Business | India **Jul 2017 - Jul 2019**

- Owned the end-to-end model development pipeline including data preparation, identification strategy, estimation, robustness checks to estimate the impact of earnings forecast misses on share buybacks, delivering 15% annual returns.
- Engineered a PySpark big data pipeline to process 7B+ USPTO patent/citation records and build network-based innovation metrics, translating signals into trading strategies delivering 30% annualized returns.
- Executed large-scale multi-variate regressions with fixed effects and clustered standard errors on 25B+ records of BAB risk exposure across 25 countries, improving timing and allocation for a 10% annual return uplift.
- Designed statistical indicators for political-regime shifts and converted them into allocation signals for trading strategies across 25 countries, improving annual returns by 10%.
- Led 5+ technical workshops on AI/ML-based signal generation, backtesting, and portfolio optimization, mentoring 10+ junior researchers and maintaining a 100% performance rating.

Research Associate, Indian School of Business | India **Jun 2016 - Jun 2017**

- Built a non-linear portfolio optimization engine using ARIMA time-series forecasting for risk/return and genetic algorithms for dynamic rebalancing, delivering 25% average annual risk-adjusted returns.
- Owned an end-to-end time-series risk pipeline using GARCH to forecast volatility and co-movement across 25 countries, improving hedging effectiveness and reducing annual average downside by 30%.
- Developed logistic regression models to predict tail-risk event probabilities across credit, market, and operational risks, integrating outputs into position-sizing and reducing expected annual losses by 20%.

EDUCATION

PhD in Quantitative Finance (Paused), **Boston College***, Boston, USA*

Master of Technology in Engineering, **Indian Institute of Technology (IIT)***, Kharagpur, India*

Bachelor of Technology in Engineering, **Indian Institute of Technology (IIT)***, Kharagpur, India*