Ramy Jaber

www.ramyjaber.com | [in] linkedin.com/in/ramyj | ramyij@pm.me

PROFESSIONAL EXPERIENCE

Intel, Remote, NJ   *Oct 2022 - Present*
Cloud Solution Architect - LLMs
○ Led technical GTM integration for $70M/3-year partnership with SeekrFlow Enterprise LLM (seekr.com); developed demo Python notebooks, authored technical sales materials, driving 7 enterprise customer engagements and two committed POVs within 3 months
○ Lead CUDA-to-Gaudi migration strategy for InflectionAI inference workload using Intel's Optimum-Habana PyTorch extension. Identified CUDA-optimized functions and substituted with HPU equivalent, ensuring functional parity while achieving hardware-specific optimizations, resulting in $8M ARR contract
○ Deployed Open Sourced LLMs on vLLM and optimized inference pipeline performance through profiling, identifying bottlenecks in resource utilization and collaborating with engineering teams to optimize low-level operators such as GEMM, GELU, Softmax
○ Implemented distributed training proof of concept using HCCL DDP across multi-node Gaudi clusters, achieving near-linear TFLOP scaling (1-8 nodes) on GPT-2 training workloads and demonstrating 5% performance advantage over Nvidia H100 baseline resulting in technical win within 6 weeks of first engagement

Cloud Solution Architect - Performance Engineering (Granulate.io)
○ Drove $3M ARR as technical account owner, advising customers on optimal infrastructure decisions across data solutions (Databricks, EMR, Spark on K8s, etc) for Digital Native and Enterprise customers, leading to average of 20% cost reduction
○ Automated onboarding, improving velocity from from 20 to 300 workloads per day

DataRobot, Remote, NJ   *Jul 2021 - Oct 2022*
Pre-sales Data Scientist
○ Led technical presales across accounts in Financial Services, Retail, and Telecom sectors in NYC region, delivering tailored demos, including a novel credit rating prediction solution for private business loans resulting in three additional POCs
○ Prevented $2M account churn by identifying and incubating new use cases across two departments, developing a novel model compliance documentation automation

Udacity, Remote, NY   *Feb 2019 - Apr 2021*
Director, Solution Architects - Global Enterprise
○ Scaled Solutions Architecture team from 1 to 10 SA, supporting a global sales force of 40 Account executives while exceeding revenue targets: Y1: $13M (1.7x quota), Y2: $33M (1.8x quota), Y3: $40M projected (1.1x quota)
○ Led development of enterprise skills transformation program for Big 4 consultancy, designing curriculum and custom projects to upskill 400+ employees in data analytics
○ Established comprehensive "GTM Nanodegree" learning program for all Sales, Customer Success, and Solution Engineers, incorporating Value Messaging, success stories, and foundational data science knowledge to enable consistent technical storytelling

Senior Solution Architect
○ First pre-sales technical resource on hyper growth Enterprise Sales team. Defined the Solution Architect role and directly supported $13M in sales in 2019
○ Expanded key account revenue by $2.1M through executive relationship building and technical partnership with Shell Oil & Gas (🔥)
○ Evangelized content knowledge of Data Science, AI/ML, and Cloud by presenting content "deep dives" and other enablement on a monthly basis

Appian Corporation, Reston, VA   *Feb 2015 - Jul 2017*
Lead Solution Engineer - Senior Solution Engineer - Solution Engineer
○ Promoted to team lead in one year, facilitating scaling of team from 8 to 29 engineers
○ Led project team to analyze performance of 600 sites and developed Python scripts to ingest monitoring alerts reducing noise of hundreds of hourly emails by 85%
○ Partnered with Product Development team to resolve high impact software bugs by providing analysis of performance log data with recommended actions

EDUCATION

Columbia University in the City of New York
December 2018 - New York, NY
Masters of Science in Data Science

Stevens Institute of Technology
May 2012 - Hoboken, NJ
Bachelors of Engineering in Engineering Management
Minor in Economics
Minor in Pure and Applied Mathematics

SKILLS

LEADERSHIP & STRATEGY
Executive Relationship Management
Technical Team Management
Sales Enablement & Training
Enterprise Account Strategy
Cross-functional Collaboration

DATA SCIENCE / ML / GEN AI
AI Accelerators (GPU, Intel Gaudi)
Distributed training/inference (HCCL)
Transformers
HuggingFace
Model performance analysis
Inference Optimization
Infrastructure Evaluation
Model deployment and monitoring
Model Serving (vLLM)

CLOUD ADMINISTRATION
Spark - Databricks, EMR, Dataproc
Infrastructure - VMs, storage, serverless
Kubernetes Orchestration

SOLUTION ARCHITECTURE
Product demos
Use Case identification
POV Execution
Product feedback

SALES
MEDDPICC
Value based selling
Force Management