TR

Tanya G. Roosta

Saratoga, CA 95070 1-510-220-0047 Tanya.roosta@gmail.com
Personal Website , LinkedIn

PROFESSIONAL SUMMARY

Highly motivated, self-starter individual with more than 15 years of work experience in data science, analysis and machine learning in high tech and quantitative finance industries. Detail-oriented team player with strong organizational skills. Ability to handle multiple projects simultaneously with a high degree of accuracy.

SKILLS

Computer Skills

- Proficient in: Pytorch, Python, Java, C, C++, R, Matlab
- Familiar with: SAS
- SQL, Tableu
- PATENTS
- Written Communication

Interpersonal Skills:

- Conflict Resolution
- Team Building
- Problem-Solving
- Project Planning
- Organization
- Self-Directed and innovative

WORK HISTORY

SENIOR APPLIED SCIENCE MANAGER, AMAZON *06/2019 to Current*

Amazon Search: 01/2024 to *CURRENT*

Working on providing improved search experiences to the customers, through improving the existing retrieval methods, as well as using personalization with generative AI techniques. In addition, my team works on optimizing the retrieval and ranking models to improve latency and save on costs in production.

Amazon AGI: 06/2019-12/2023

Leading a team of scientists, machine learning, and software engineers working on scalable generative AI techniques to increase customer interaction with their Alexa devices through improving their conversational experiences
In this role, I devise strategic roadmap for product release, work with various teams to scope requirements, do research and help build the prototype technical solution, and oversee the end-to-end implementation in production.

In both roles, I have used my technical and interpersonal skills to accomplish the following:

AI Foundation Model Development:
Successfully built and deployed AI foundation models from design through training to implementation.

High-Impact Applied Research:
Conducted cutting-edge applied research, incorporating the latest AI developments into customer experiences.

Technology Stack Utilization:
Leveraged diverse technology stack including Pytorch, AWS, HuggingFace, Lightning, and VectorDBs.

Insight Extraction from Data:
Utilized advanced technologies to extract meaningful insights from vast volumes of numeric and textual data.

Interpersonal Communication:
Effectively translated complex technical concepts into tangible business goals. Enhanced team collaboration by bridging the gap between technical and business perspectives.

LECTURER *07/2022 to CURRENT*
UNIVERSITY OF CALIFORNIA, BERKELEY
Teaching fundamentals of machine learning, GenAI, and Statistics courses.

LEAD MACHINE LEARNING SCIENTIST *03/2018 to 05/2019*
SUMUP ANALYTICS
As a key contributor at this early-stage dynamic Fintech startup, I led impactful research initiatives while simultaneously showcasing my leadership skills. My role encompassed not only the development of robust code and REST APIs, hosted on AWS API Gateway, but also doing cutting-edge research in topic modeling, text summarization, document recommendation, sentiment analysis, and streaming text analytics.

In addition to hands-on technical contributions, I took charge of managing a team of skilled scientists and engineers across the entire technical stack. My leadership ensured seamless collaboration and synchronization, resulting in scalable and reliable feature development that aligned with our organizational objectives.

Furthermore, I actively engaged with prospective clients, demonstrating my ability to understand their unique needs. I played a pivotal role in developing proof-of-concept solutions tailored to address specific client requirements, showcasing my ability to transform research outcomes into practical solutions that meet client expectations.

DIRECTOR OF DATA SCIENCE AND MACHINE LEARNING

05/2014 to 03/2018

MOODY'S ANALYTICS

I spearheaded diverse engagements within the banking and financial sector, assuming dual roles as the quantitative model lead and team manager. I successfully navigated a spectrum of projects, overseeing the development and implementation of quantitative risk measurement products and services.

In my capacity as a team manager, I fostered a collaborative environment, steering a team of professionals toward shared objectives. My leadership extended to projects encompassing domains, such as credit risk modeling for commercial wholesale portfolios, commercial real estate, residential mortgages, and counterparty credit risk.

In the execution of these projects, I demonstrated the ability to communicate complex quantitative concepts to stakeholders, ensuring a comprehensive understanding across all levels. My interpersonal communication skills played a pivotal role in contributing to the success of engagements and strengthening relationships with clients and stakeholders.

SENIOR QUANTITATIVE ANALYST

06/2012 to 04/2014

FEDERAL RESERVE BANK OF SAN FRANCISCO

In my capacity as the lead quantitative analyst overseeing supervisory exams, I played a central role in the examination processes of major bank holding companies regulated by the Federal Reserve Bank of San Francisco, including industry leaders, such as Wells Fargo, Union Bank of California, and Zions Bank. Additionally, I actively contributed to critical system-wide initiatives at the Federal Reserve, including the Comprehensive Capital Analysis and Review (CCAR) and BASEL II quantification exams, working with investment institutions such as Morgan Stanley and Goldman Sachs.

My analytical skills were prominently featured in my responsibilities, which extended to conducting empirical research in areas, such as credit risk modeling, systematic risk indicators utilizing Credit Default Swap (CDS) spreads, and Commercial Real Estate (CRE) modeling and analysis. Employing a comprehensive research methodology, I gathered data from diverse sources, including Bloomberg, RCA, FRED, FR Y-9C, and 14A bank schedules. My proficiency in data cleaning processes and the application of sophisticated time series models allowed me to extract actionable insights from complex datasets.

SENIOR QUANTITATIVE RESEARCHER
10/2010 to 05/2012

ALLIANZ GLOBAL INVESTOR CAPITALS

I was responsible for the development of a statistical risk model for emerging markets, which allowed me to showcase my strong quantitative and analytical skills by employing techniques, such as principal component analysis (PCA). This work involved a meticulous approach to data analysis, enabling the creation of a robust model to assess and manage risks associated with emerging market dynamics.

Furthermore, I crafted a strategy centered around the minimum variance portfolio. Rigorous testing of this strategy culminated in the establishment of the AllianzGI U.S Managed Volatility Fund (NGWAX). My quantitative skills played a key role in designing and implementing this strategy, contributing to the fund's inception and success in navigating market volatility effectively.

SOFTWARE ENGINEER
08/2008 to 02/2010

CISCO SYSTEMS

I was responsible for the development of software in C language, specifically for the Group Encrypted Transport Virtual Private Network (GETVPN) solution. This project showcased my proficiency in programming and my ability to navigate complex coding tasks within a robust technology stack.

In addition, I actively collaborated on crafting the functional specification document for the SmartGrid project. This endeavor required not only a deep understanding of project requirements but also the application of my technical skills to articulate complex functionalities. These experiences underscore my capability to leverage diverse technology stacks and contribute effectively to multifaceted software development initiatives.

EDUCATION

Ph.D. | Electrical Engineering and Computer
2002-2008

University of California, Berkeley, CA

Dissertation: Using Statistical Methods for Anomaly and Attack Detection in Wireless Sensor Networks. Advisor: Prof. Shankar Sastry

Masters | Statistics

University of California, Berkeley, CA

Thesis: Convergence Analysis of Reweighted Sum-Product Algorithm. Advisor: Martin Weignwright

Masters | Electrical Engineering and Computer Science

University of California, Berkeley, CA

Thesis: Power aware Routing in Wireless Ad-hoc Networks

Masters | Financial Engineering
Haas School of Business

Bachelors | Electrical Engineering and Computer Science
University of California, Berkeley, CA

PUBLICATIONS

Federated Multimodal Learning with Dual Adapters and Selective Pruning for Communication and Computational Efficiency. CCGrid 2025.
How Well Do LLMs Represent Values Across Cultures? Empirical Analysis of LLM Responses Based on Hofstede Cultural Dimensions. Agentic & GenAI Evaluation KDD2025.
The Order Effect: Investigating Prompt Sensitivity in Closed-Source LLMs. Feb 2025, Arxiv.
AuditLLM: A Tool for Auditing Large Language Models Using Multiprobe Approach. The 33rd ACM International Conference on Information and Knowledge Management (CIKM), 2024.
ClaimVer: Explainable Claim-Level Verification and Evidence Attribution of Text Through Knowledge Graphs. EMNLP 2024.
What is Lost in Knowledge Distillation. NeurIPS 2023 (ENSLP)
Once-for-All Federated Learning: Learning from and Deploying to Heterogeneous Clients, FL4Data-Mining, 2023
Quantifying Catastrophic Forgetting in Continual Federated Learning ICASSP, 2023
Learning from Federated Learning in Real World. ICASSP, 2022
Training Mixed-Domain Translation Models via Federated Learning. NAACL, 2022
PerFedSI: A Framework for Personalized Federated Learning with Side Information. FL NeurIPS, 2022.
Communication-Efficient Federated Learning for Neural Machine Translation. NeurIPS 2021 (ENSLP)
Convergence Analysis of Reweighted Sum-Product Algorithms International Conference on Acoustic, Speech, and Signal Processing 2007 (**best student award**)
For the full list of publications, please see Personal Website.

PATENTS

Dynamic Group Creation for Managed Key Servers. Granted: 6/2014
Sender-specific Counter-Based Anti-Replay for Multicast Traffic. Granted: 6/2011
Protection of Control Plane Traffic against Replayed and Delayed Packet Attack. Granted: 2/2014
Updating Machine Learning Models Across Devices. Filed, Number P75341-US01
Updating Machine Learning Models. Filed, Number P75359-US01

AWARDS

Best student paper award, the International Conference on Acoustic, Speech and Signal Processing, 2007

National Science Foundation Graduate Student Research Fellowship