Anugrah Kumar

Sacramento, CA ● (480) 738-0570 ● anugrah18@gmail.com ● LinkedIn: https://www.linkedin.com/in/anugrah-kumar/

Senior Software Engineer

Senior Software Engineer with 9+ years of experience building backend, cloud-native, and distributed systems, with the past several years focused on production-grade AI platforms. Specialized in agentic AI, LLM-powered workflows, offline and online evaluation, and large-scale observability and telemetry-driven systems. Strong background in backend engineering and cloud infrastructure, with a focus on safe, scalable, and data-driven AI deployment.

PROFESSIONAL EXPERIENCE

Senior Software Engineer – AI

LaunchDarkly, Oakland CA (Remote) **Oct-2025 to present**

- Technical lead for integrating offline AI evaluation capabilities into the LaunchDarkly platform, enabling bulk evaluations on golden datasets with support for LLM-as-a-judge, custom judges, and heuristic-based evaluators, allowing scalable, reliable pre-deployment validation of AI-powered features.
- Contributed to shipping an AI experimentation playground, enabling rapid prototyping, prompt iteration, and side-by-side evaluation of model outputs, and serving as a primary interface for developers to experiment with and validate AI behaviors before rollout.
- Integrated DeepEval with G-Eval–based LLM judging into offline evaluation and AI playground systems, enabling standardized evaluation and faster experimentation.
- Worked on Vega, an AI-powered, context-aware observability assistant embedded in the LaunchDarkly platform, enabling developers to understand, debug, and resolve production issues directly from logs, traces, errors, and session replays.

Software Engineer II

Microsoft Corporation, Redmond WA (Remote) **July-2022 to Sep-2025**

- Contributed to the design and development of a RAG-based automated ticket assignment system leveraging Azure OpenAI Service (LLM) for multilevel, multiclass classification and Azure AI Search for intelligent search, retrieval, and team recommendations. This improved MS Teams-related ticket routing efficiency and significantly reduced manual workload and is actively used by support engineers during on-call operations.
- Owned the end-to-end design and implementation of an AI evaluation framework (offline and online) for a RAG-based automated ticket assignment system establishing accuracy, cost, and latency benchmarks and enabling data-driven iteration and safe, scalable production deployment.
- Designed and deployed an autonomous agentic AI system leveraging Semantic Kernel (in production) and LangChain (during prototyping) for orchestration, with execution coordinated via MCP servers and exposed through FastAPI. The agent queries and analyzes 200+ performance metrics per Microsoft Teams tenant, assists in onboarding and validating new performance metrics into monitoring dashboards, uses Azure OpenAI to generate concise summaries and mitigation strategies, applies machine learning to rank high-impact signals, and automates Azure DevOps work item creation — empowering escalation engineers and PM to quickly identify and resolve issues.
- Actively contributed to an agentic AI system for shift-left regression detection in Microsoft Teams, enabling end-to-end automation via multi-agent workflows powered by Semantic Kernel involving automated Azure DevOps PR comments, regression alerts, and post-merge root-cause analysis, improving early detection of performance related issues.
- Led the transition from POC to full production deployment of statistical based anomaly detection algorithms (EMA and MAD) for early performance regression detection in real-time on a pre-merge pull requests (PRs), increasing regression identification accuracy by 27% and improving recall by 18%. Extensively utilized by QA engineers, this algorithm served as the foundation for the pre-merge regression detection heuristics.
- Worked on full stack development of web applications which are used for monitoring and detecting regression of Microsoft Teams performance across multiple rings during canary deployment based on telemetry data.
- Collaborated with UI/UX designers on developing and shipping lite features optimized for faster performance on low-end clients such as VDI (Virtual Desktop Infrastructure), Teams for education etc. for Microsoft Teams.
- Worked on cleaning, processing, transforming and aggregating INP (Interaction to Next Paint) telemetry event data into Parquet format using Apache Spark (PySpark), enabling daily analysis and monitoring of user interaction responsiveness issues through dashboards supporting over 100 million daily active Microsoft Teams users.
- Contributed to reducing MS Teams performance evaluation CI/CD pipeline annual operating costs by approximately $60,000 annually through process optimization, eliminating redundancy and improved resource allocation.

- Actively engaged in rotational on-call duties, addressing customer escalations and bugs related to front-end performance on Microsoft Teams.

Software Engineer
Polysign, San Francisco CA (Remote) **July-2021 to May-2022**

- Developed various microservices using Flask (Python) for a trustless peer-to-peer network secured by a private blockchain, used for various financial services such as trading digital assets, KYC etc.
- Designed and developed an asynchronous observer service to monitor and track the status of records submitted to a private blockchain for memorialization. The service ensured real-time updates and reliable processing of record statuses
- Mentored junior developers for developing various blockchain based client utilities used for signing and authentication of orders that are broadcasted among various nodes.
- Designed and developed microservice using Flask (Python) for management and transaction of digital assets (bitcoin) held by custodian.
- Developed a block explorer UI using React to display detailed information about transactions recorded in a private blockchain. The interface included features for viewing transaction metadata, block details, and real-time updates, providing users with an intuitive and seamless way to explore and verify blockchain data

Software Engineer
Intel Corporation, Folsom CA **June-2016 to June-2021**

- Spearheaded the development of a globally utilized automation tool employing React and Django (Python) for packaging, testing, and deploying firmware components to external customers, resulting in an estimated monthly time savings of 600 hours.
- Responsible for development of backend API for Chatbot application used by customers for inquiring firmware information using Django (Python) and Redis for caching.
- Designed and implemented a React and Node.js web application for debugging purposes, facilitating the parsing, decoding, and visualization of tracing logs associated with various firmware components.
- Built a PyTorch based computer vision microservice for detecting hexadecimal status codes on hardware boards to enable remote firmware debugging.
- Developed and maintained tools for IT service workers to monitor, report and communicate incidents in real-time, leveraging employee telemetry data to efficiently identify and resolve issues using React, ASP.NET MVC 5, and Node.js.
- Developed a prototype to automate the data collection process for various computer peripherals using pattern recognition in machine learning.
- Developed web-based application using jQuery and D3.js for data visualization of the GPU test content data which improved the overall analysis and debugging process of the team.

TECHNICAL SKILLS

Front End: React, Apollo Client, Redux, Tkinter, WinForms, jQuery, HTML/CSS
Back End: FastAPI, Flask, Django, Node.js, Express, .NET, SocketIO
API Protocols: RESTful APIs, GraphQL, WebSocket, MCP (Model Context Protocol)
Database: PostgreSQL, Microsoft SQL server, MongoDB, Redis, KQL, Clickhouse
Programming Languages: Python, JavaScript, TypeScript, C#, C++
Cloud Technologies: Azure, Amazon Web Services (AWS)
AI Tools & Frameworks: Azure Open AI, Azure AI Search, LangChain, Semantic Kernel, DeepEval, Amazon Bedrock
Data Science/Machine Learning: NumPy, Pandas, Scikit-learn, PyTorch
Distributed System: Apache Kafka, Apache Spark (PySpark), Azure Databricks
DevOps: Azure DevOps Services, Docker**,** Kubernetes, GitHub Actions, CI/CD
Tools: Git, JIRA, Chrome Developer tools, Vite, Webpack, Postman, D3.js, ECharts, Power BI, Cursor, GitHub Copilot
Achievements: Microsoft Global Hackathon 2022 Award Winner
Certifications: Azure AI Fundamentals, Azure Fundamentals, AWS Certified Cloud Practitioner

EDUCATION

Master of Science in Computer Science 2016
Arizona State University, Tempe, AZ

Bachelor of Technology in Computer Science and Engineering 2014
Vellore Institute of Technology, India