Raunit Kohli

Boston, MA | 774-262-9233 | raunitkohli@gmail.com | raunit-kohli | kohlir2020

WORK EXPERIENCE

Rocket Software **Boston, MA**

Machine Learning Software Engineer 09/2024 – Present

- Design and implement novel Large Language Model Citation Generator using multi-Gaussian clustering algorithms and dynamic content similarity techniques, yielding 80% more relevant document references for production-grade RAG model.
- Deploy RESTful API endpoint for user query feedback in parallel collaboration with cross-functional teams against time-critical deadline, allowing for immediate customer feedback necessary for creating gold-standard datasets.
- Perform experimental process of fine-tuning, testing, and evaluating multiple transformer-based sentence embedding models with custom tokenizers to balance keyword relevance and semantic similarity for improved document reranking.
- Develop proof-of-concept few-shot LLM prompting framework to translate natural language queries into executable SQL, integrating regex-based syntax validation, enhancing table-based document retrieval and question answering.
- Augment state-of-the-art AI Code Assistant tools to create a novel LLM-powered SAS to python IT chain management code translation tool. YADA YADA YADA
- Utilize AWS CloudWatch and develop multi-purpose scripts to parse and analyze on-going trial logs, enabling proactive troubleshooting and creating consistently responsive data to customer requests, leading to higher user satisfaction.

Robotics Using Bayesian Inference Lab - UC San Diego **San Diego, CA**

Machine Learning Research Consultant 09/2024 – Present

- Spearheaded implementation of Reinforcement Learning algorithms in PyTorch for robotic PiRat with ROS utilizing extensive neural and positional data from 2500+ rat experiments to mimic and predict multi-agent interactions.
- Led technical engineering of network synchronized multi-modal Embedded-System aimed to host AI-powered classroom.

Staff Machine Learning Research Engineer (Published) 01/2021 – 09/2024

- Spearheaded implementation of Reinforcement Learning algorithms in PyTorch for robotic PiRat with ROS utilizing extensive neural and positional data from 2500+ rat experiments to mimic and predict multi-agent interactions.
- Led technical engineering of network synchronized multi-modal Embedded-System aimed to host AI-powered classroom.
- Developed signal processing applications for live and post analysis, leveraging advanced filtering techniques to enhance identification of stress-inducing behaviors from neural and physiological signals (HRV, PPG) of project subjects.
- Built real-time Homebase Detection Algorithm with convolution feature extractor for novel experiment apparatus.
- Enhanced and scaled production-deployed Computer Vision-based live Pose-Estimation Software in C++ to validate statistical accuracy of Kinect Tracker, introducing new features and parameter logging, leading to 300% processing speed.

IBM (International Business Machines) **San Jose, CA**

Automation and Testing Developer Intern 06/2023 – 09/2023

- Scripted in Ruby to automate API endpoint testing for entire IBM Aspera on Cloud Workflows App designed to automate workflows of global content transfer and exchange across on-premises and multi-cloud environments.
- Developed and documented comprehensive automated regression framework with 780+ API tests covering all endpoints.
- Identified 15+ critical issues and bugs on live production server, communicated with developers to determine ideal expected behavior, and raised issues in project management tools to expertly report relevant failure details.
- Executed full development cycle to outline and implement API Endpoint Test Controller for collaborative QA repository.

UC San Diego Cognitive Science **San Diego, CA**

Data Science Teaching Assistant 09/2022 – 12/2023

- Created weekly section presentations, podcasted to 800+ students, for content review and interactively guiding students at all coding-experience levels through programming assignments by demonstrating live debugging and problem-solving.
- Formulated weekly exams, held tutoring sessions, scripted automated grading pipelines to optimize multi-LMS system.

TECHNICAL SKILLS

Languages: Python, Ruby, Java, C++, SQL, Unix; **Libraries:** Pandas, PyTorch, NumPy, Scikit, HuggingFace, Dspy
Agile/Tools: Atlassian (Jira, Confluence, Bitbucket), Version-Control (Git, Github, Gitlab), Excel, Outlook, Webex, Teams
Skills: Policy-Optimization Algorithms, NLPs, Data Visualization, API Development, Neural Networks, Data Structures, ML Classification/Regression, Markov Decision Processes, Kubernetes Deployment, AWS Suite Supervision, Signal Processing

EDUCATION

University of California – San Diego **09/2020 – 12/2023**

Bachelor of Science, Machine Learning and Neural Computation Major GPA: 3.91/4.0

Minor: Computer Science; **Program of Concentration:** Philosophy

Awards: President & Recruitment Advisor – Phi Sigma Pi National Honor Fraternity; Provost Honors (Warren College)