

Job Details

Old Resume

New Resume

Guardrails AI

→

...

Senior Machine Learning Engineer

San Francisco Bay Area · Reposted 3 months ago · Over 100 applicants

Starting at \$180K/yr + Stock options · Hybrid · Full-time

1-10 employees

4 of 4 skills match your profile - you may be a good fit

See how you compare to over 100 other applicants. [Retry](#) [Premium for \\$0](#)

Apply

Save

About the job

About the role

As a senior **Machine Learning** Engineer at Guardrails AI, you will be contributing to a critical component of the company's roadmap.

- Your primary responsibilities will be to own end-to-end ML projects. A key component of your work will be to scope out and build whatever is the best-possible ML solution (high performance, low cost, **low latency**) to green-field problems. This includes scoping out the ML use case, designing and **curating** appropriate datasets, **shortlisting algorithms**, training models (including finetuning), experimentation, evaluation and deployment.
- You'll build general solutions to improve the reliability and robustness of **LLMs**, as well as industry specific solutions focusing on finance, healthcare, etc. You'll work closely with customers on understanding their problems.
- You'll think about the best way to generalize your work and build abstractions to improve ML and experimentation infrastructure.

About you

- You have minimum 3 years of experience building **Machine Learning** models in production settings. Your role is not just the research and experimentation in order to build a SoTA models, but also to ship the product to a production environment.
- You're excited about being cross-functional and owning the lifecycle of building ML solutions – from **model development** to deployment.
- You're extremely comfortable with **Python**, sklearn, **numpy** and a **deep learning** framework such as **Pytorch / Keras / TensorFlow**.
- You stay up-to-date with the latest research in your area of interest in ML. You're comfortable reading research papers, figuring out what the key ideas from the paper are and reproducing the results in a production environment.
- You've previously worked with **deep learning** systems on **unstructured data** such as text, images, **LIDAR**, etc.
- You're really excited about Large Language Models, and excited about building applications that leverage **LLMs** for solving problems.
- You enjoy fast-paced environments, and exhibit a high degree of ownership and self-sufficiency. Unstructured environments are exciting to you because they represent opportunities for growth and leadership.

ML Stack

- Python
- Pytorch
- Huggingface transformers
- Numpy
- Sklearn
- Pandas
- Ray

Saurabh Bhausahab Zinjad

saurabhjinjad@asu.edu

+14809135544

linkedin.com/in/saurabhjinjad

github.com/Ztrimus

WORK EXPERIENCE

Senior Machine Learning Engineer
Tiger Analytics

06/2022 - 07/2023Bangalore, India

Achievements/Tasks

- Led a team of 8 analysts to develop Data & CI/CD pipelines and Constraint-based ML Models, enhancing MSP Value Opti- mization in the Petcare sector, resulting in a 15% increase in predictive accuracy.
- Pioneered the MLCORE product, an end-to-end MLOps platform, by translating research into practical applications, which secured four new significant clients and improved deployment efficiency by 20%.
- Crafted Interactive Dashboards and a Web App with comprehensive documentation, streamlining data visualization and user interaction, leading to a 25% improvement in client engagement.

Software Engineer
Winjit Technologies

01/2020 - 06/2022Pune, India

Achievements/Tasks

- Built 10+ RESTful APIs with well-managed, optimized databases.
- Constructed ML POCs, implemented research ideas, and Systematized through Prototyping.
- Created low-latency responsive UI/UX user workflows (30-40) with high-quality web architecture and integrated ML solutions.
- Handle a technical team of 6-7 developers.
- Analyzed business needs & delivered an apt solution.
- Devise a generic solution for dynamic forms module with CSS customizing ability that used in-house, which reduces development time by 1x.

Machine Learning Engineer
Automation Teknix

08/2019 - 01/2020Pune, India

Achievements/Tasks

- Build a cars inspection system on an assembly line for mobiles using computer vision, Thus reducing survey error by 22%.
- My abilities involved doing Initial research, prototyping deep learning flow, Conceptualization POC, training, monitoring models in terms of accuracy and time.
- 7% increase in accuracy and reduce inference time by 2x.
- Used SSD CNN, Tensorflow, MXNet, TFLite, Image Processing.

SKILLS

Python, Java/Script, C#, C++, SQL, R, Java

Pythons, tensorflow, PySpark, PyTorch, MXNet

OpenCV, Scikit Learn, Pandas, Matplotlib, Keras

Azure, AWS, Docker, Kubernetes, MLFlow, Jupyter Notebook, GIC

Django, Flask, FastAPI, MongoDB

Angular, React, .Net Core, NodeJs, SQL Server

MySQL, Postman

ACHIEVEMENTS

1st Runner Up Prize in Prompt Engineering For Humanities

Big Data Awards - Sunhack Hackathon

PROJECTS

GenAI's Capabilities and Boundaries Exploration - Prompt Engineering For Humanities (10/2023 - 10/2023)

- 1st Runner Up Prize in crafted AI persona, to explore LLM's subtle contextual understanding and create innovative colla- orations between humans and machines.
- Addressed limitations in narrative flow, simplicity, emotional depth, and hallucinations through innovative approaches.
- Demonstrated creative mindset and ability to navigate complex tasks and adapt to evolving requirements during hackathon.

LLM-facilitated Pipeline for Personalized Resume Generation (12/2023 - 02/2024)

- Automated resume and cover letter generation by developing a Python library, directly addressing the need for efficiency in job application processes, resulting in a 50% reduction in application time for users.
- Enhanced job application workflow using Prompt Engineering and Web Scraping, integrating Large Language Models (LLMs) to tailor documents to specific rules, aligning with the job's requirement for natural language processing expertise.
- Implemented the application with a focus on user experience, leading to a streamlined job hunting process that mirrors the duties of deploying and maintaining machine learning models in a production environment.

Search Engine for All File Types - Sunhack Hackathon - Meta & Amazon Sponsored (11/2023 - 11/2023)

- Awarded Big Data Awards by developing a Python FAST API with Angular, enabling efficient data access and retrieval, which is critical for handling large datasets in machine learning.
- Implemented a search engine using BERT, OCR, and ReAct50 for vector embeddings of diverse file types, ensuring low-latency search capabilities, pertinent to the job's focus on data science and machine learning techniques.
- Contributed to an Elasticsearch implementation, achieving millisecond response times, demonstrating the ability to deploy ML models into production environments as required by the job description.

Page 1 of 2

EDUCATION

Arizona State University(ASU), Tempe

Masters of Science in Computer Science - Thesis Advised by Dr. Huan Liu, Amrita Bhattacharjee

06/2021 - 01/2023Tempe AZ

Courses

- Operationalizing Deep Learning, AI Safety, Social Media Mining
- Statistical Machine Learning

Pune Institute of Computer Technology(PICT), Pune University

Bachelor of Engineering

07/2015 - 06/2019Pune, India

Courses

- DSA, OOP, OS, System Design, AI, ML
- Digital Video, Image Processing, Computer Networks

CERTIFICATIONS

Deep Learning Specialization - DeepLearning.AI

Server-side Backend Development

Microsoft Azure Fundamentals

MLOps for AI Engineers and Data Scientists

Front-End Web UI Frameworks and Tools

SAURABH BHAUSAHEB ZINJAD

+1480-913-5544

saurabhjinjad@asu.edu

github.com/Ztrimus

linkedin.com/in/saurabhjinjad

EDUCATION

Arizona State University(ASU), Tempe

Masters of Science in Computer Science - Thesis Advised by Dr. Huan Liu, Amrita Bhattacharjee

06/2021 - 01/2023Tempe AZ

Courses

- Operationalizing Deep Learning, AI Safety, Social Media Mining, Statistical Machine Learning
- Statistical Machine Learning

Pune Institute of Computer Technology(PICT), Pune University, India

Bachelor of Engineering

07/2015 - Jun 2019Pune, India

Courses

- DSA, OOP, OS, System Design, AI, ML, Digital Video and Image Processing, Computer Networks
- Digital Video, Image Processing, Computer Networks

PROFESSIONAL EXPERIENCE

Senior Machine Learning Engineer

Tiger Analytics

Jun 2022 - Jul 2023Bangalore, India

- Spearheaded development of Data & CI/CD pipelines, Interactive Dashboards, Constraint-based ML Models, Web App, and Comprehensive Documentation for MSP Value Optimization in Petcare sector with a team of 8 analysts. (Team Leadership)
- Developed MLCORE product (end-to-end MLOps platform) by implementing research ideas, organizing through prototyping, and Integrating it with numerous cloud services, attracting an additional four significant clients. (Interpersonal Skills)

Software Engineer

Winjit Technologies

Jan 2020 - Jan 2022Pune, India

- Engineered 10+ RESTful APIs Architecture and Distributed services; Designed 30+ low-latency responsive UI/UX application features with high-quality web architecture; Managed and optimized large-scale Databases. (Systems Design)
- Initiated and Designed a standardized solution for dynamic forms generation, with customizable CSS capabilities feature, which reduces development time by 8x; Led and collaborated with a 12 member cross-functional team. (Idea Generation)

Deep Learning Engineer

Automation Teknix

Sep 2019 - Jan 2020Pune, India

- Devised a Lightweight Object Recognition Engine with a low computational cost by leveraging an SSD algorithm with MobileNetV2 architecture, which decreased survey error by 22%. (Problem Identification)
- Conducted thorough Initial research; prototyped neural network flow; conceptualized POC, training, and monitoring of models. This resulted in a 7% increase in accuracy and reduced inference time by 2x. (Experiment Design)

PUBLICATIONS AND RESEARCH

LLM-facilitated Pipeline for Personalized Resume Generation - Website - GitHub

Dec 2023 - Feb 2024

- As the First Author, Led the proposal and development of 'ResumeFlow,' a pioneering tool powered by Large Language Models (LLMs), enabling users to generate personalized resumes tailored to specific job postings in seconds.
- Designed an innovative pipeline utilizing the advanced capabilities of LLMs (e.g., GPT-4, Google's Gemini) for information extraction and language understanding, applying prompt engineering and web scraping to extract job and resume details, and generate tailored resumes.
- Developed Python library to enable seamless integration and optimized with any job application workflow and system.

PROJECTS

Search Engine for All file types - Sunhack Hackathon - Meta & Amazon Sponsored

Nov 2023 - Nov 2023

- Big Data Awards in providing efficient data access and retrieval using FAST APIs and angular development.
- Converted and stored every file type data as vector embeddings, ensuring low-latency search capabilities.
- Used Machine Learning techniques such as BERT, OCR, ResNet50, and Image Captioning to parse Image features.

GenAI's Capabilities and Boundaries Exploration - Prompt Engineering For Humanities

Oct 2023 - Oct 2023

- 1st Runner Up Prize in crafted AI persona, to explore LLM's subtle contextual understanding and create innovative collaborations between humans and machines.
- Addressed limitations in narrative flow, simplicity, emotional depth, and hallucinations through innovative approaches.
- Demonstrated creative mindset and ability to navigate complex tasks and adapt to evolving requirements during hackathon.

TECHNICAL SKILLS

Programming Languages: Python, JavaScript, C#, C++, SQL, R, Java, Shell Scripting

Data Science: Databricks, PySpark, TensorFlow, PyTorch, MXNet, OpenCV, Scikit Learn, Pandas, Matplotlib, Keras

Cloud and DevOps: Azure, AWS, Docker, Kubernetes, MLFlow, Jupyter Notebook, Git

Full-Stack Tech: Angular, React, .Net Core, NodeJs, Django, Flask, FastAPI, MongoDB, SQL Server, MySQL, Postman

Certifications: Deep Learning Specialization, Server-side Backend Development, Microsoft Azure Fundamentals, MLOps for AI Engineers and Data Scientists, Front-End Web UI Frameworks and Tools

Cosine Similarity

User Personalization Score

0.880

↑ [resume_master_data]

Job Alignment Score

0.456

↑ [resume_JD]

Job Match Score

0.493

↑ [master_data_JD]

Resume Metrics

Overlap Coefficient

User Personalization Score

0.979

↑ [resume_master_data]

Job Alignment Score

0.384

↑ [resume_JD]

Job Match Score

0.493

↑ [master_data_JD]