

# Shrenik Jain

+1-(858)-241-1904 | [shrenik-jain.github.io](https://shrenik-jain.github.io) | [shrenikkjain81@gmail.com](mailto:shrenikkjain81@gmail.com) | [linkedin/shrenik-jain9](https://linkedin/shrenik-jain9) | [github/shrenik-jain](https://github/shrenik-jain)

## SUMMARY

Experienced Machine Learning Engineer with 3 years of success in building AI-driven systems and leading technical initiatives. Expertise in Python, PyTorch, and MongoDB, specializing in the implementation and performance enhancement of Computer Vision and Natural Language Processing applications.

## EDUCATION

**University of California San Diego** May 2026

Master of Science, Electrical and Computer Engineering (Machine Learning and Data Science) - GPA: 3.7/4.0

*Coursework:* Recommender Systems & Web Mining, Statistical Learning, Linear Algebra & Application

**University of Pune** May 2022

Bachelor of Technology, Electrical Engineering - GPA: 4.0/4.0

*Coursework:* Data Structures & Advanced Algorithms, Machine Learning, Deep Learning, Image & Video Processing

## EXPERIENCE

**Machine Learning Researcher, Spatiotemporal ML Lab** Oct 2024 - Present

- Conducted research for the advancement of Diffusion models by optimizing temporal dynamics and spatial dimensions to expedite computation and enhance algorithmic efficiency.

**Machine Learning Engineer, Pivotchain Solutions** Jul 2022 - Aug 2024

- Led a cross-functional team for the agile development of an AI System, designing Computer Vision algorithms to provide a scalable and real-time interface for autonomous event monitoring achieving 70% faster threat containment (MTTC).
- Implemented a Spatiotemporal Autoencoder model to validate AI-generated video clips, precisely classifying 15,000+ daily security events, reducing false positives by 30%, and saving \$250K annually in operational costs.
- Optimized MongoDB queries to efficiently handle over 10 Million records of video surveillance JSON data, reducing data retrieval times by 55% and improving overall system performance for end-users across multiple regions.

**Software Development Intern, Qualys Inc.** Jan 2022 - July 2022

- Streamlined the orchestration of high-availability microservices deployment using Kubernetes, achieving 99% uptime and handling 1 Million+ daily requests across 3 production environments.
- Architected end-to-end CI/CD pipelines with Groovy and Docker, reducing deployment time from 30 to 10 minutes and enabling 200+ successful deployments monthly across multiple environments.

**Machine Learning Researcher, University of Pune** Jul 2021 - Dec 2021

- Spearheaded the development of a research paper summarization system using a BERT-based encoder, improving ROUGE-1 scores from 0.35 to 0.46 compared to baseline extractive summarization methods.
- Conducted research on transformers and multi-head self-attention mechanisms, for enhanced language modeling.

**Machine Learning Intern, Validus Analytics LLP** Feb 2021 - Dec 2021

- Implemented ConvVAE-based generative modeling for dataset enhancement, expanding a critical training dataset from 50,000 to 150,000 samples while maintaining high perceptual integrity (SSIM > 0.85).
- Analyzed Vector-Quantized VAEs (VQ-VAEs) and Convolutional VAEs (Conv-VAEs) for unsupervised learning of complex data via latent representations and generative modeling.

## CONSULTING EXPERIENCE

**Machine Learning Consultant, Pixstory** Aug 2023 - Mar 2024

- Contributed to building a RAG-based Conversational Search System using Large Language Models (LLMs), improving search relevance and increasing average user session duration by 2 minutes.
- Optimized system throughput and hardware efficiency 3x by implementing asynchronous requests and parallel execution strategies, which cut the average query response time from 2 seconds to 600 milliseconds.

**Software Development Consultant, AI For Rural** Sep 2021 - Nov 2021

- Collaborated with organization owners to engineer the end-to-end backend architecture for a website, including designing APIs, database schemas, and core services, resulting in a 45% improvement in user engagement.
- Developed REST APIs and integrated them with various data sources, enabling real-time data updates and reducing data retrieval time by 40% for critical information.

## TECHNICAL SKILLS

**Languages:** Python, C++, Java, JAX, SQL, JavaScript, Bash, HTML, CSS

**Machine Learning:** TensorFlow, PyTorch, MLFlow, Hugging Face Transformers, LangChain, Scikit-learn, OpenCV, CUDA, spaCy, NLTK, ONNX, TFLite, TorchServe, TritonServer, TF-Serving

**Frameworks & Technologies:** PySpark, Apache Spark, Docker, Kubernetes, Flask, Django, SpringBoot, Git, Jenkins, Linux/Unix, REST, Apache Kafka, MQTT, RabbitMQ, FFmpeg, ONVIF

**Cloud & Databases:** AWS, GCP, Azure, MongoDB, Elasticsearch, SQL, Cassandra, Milvus, Vector Stores