Shrey Bhadiyadara

602-301-9242 | shreybhadiyadara33@gmail.com | linkedin.com/in/shrey-bhadiyadara | github.com/shrey333

Education

Arizona State University — Tempe, AZ
Master of Science in Computer Science — *August 2023 – May 2025*

Dharmsinh Desai University — Nadiad, India
Bachelor of Technology in Computer Engineering — *August 2018 – May 2022*

Technical Skills

Languages: Python, JavaScript, TypeScript, HTML, CSS, SQL, NoSQL
Frameworks: Django, DRF, React, NodeJS, Redux, Express.js, Celery, jQuery, Socket.io
DB/Storage: MongoDB, PostgreSQL, DynamoDB, Pinecone, ElasticSearch, Memcached, Redis
Tools: Git, Docker, AWS CDK, AWS SDK, Kubernetes, GitHub Actions, AWS (EC2, S3, Lambda, API Gateway, EKS, Fargate, CloudFront, Cognito, CloudWatch)
AI/ML: PyTorch, AutoGen, LangChain, OpenAI API, Google Gemini, Scikit-learn, Hugging Face

Experience

Software Engineer — August 2022 – Present
Neliti Pte. Ltd. — *Remote*

- As the founding engineer, led the development of Neliti Dashboard, building a responsive and scalable frontend using React, Redux, TailwindCSS, and TypeScript, coupled with a robust Django Rest Framework (DRF) backend
- Architected end-to-end CI/CD pipeline using GitHub Actions, Docker, and AWS Fargate/EKS, reducing deployment time from 1 hour to 14 minutes and achieving 99.9% application availability
- Improved search efficiency for 12 million users by transitioning data processing from Solr to Elasticsearch, enhancing search performance by reducing query time
- Led the integration of Xendit and Square payment gateways, collaborating closely with the CTO to define technical requirements, align with business objectives, and guide the implementation process
- Automated journal crawling with Python-based scripts, integrating Crossref API for DOI registration and configuring domain setups in AWS Route 53, reducing average registration time from 2 hours to under 20 minutes
- Designed and implemented a token-based pricing model and a role-based access control (RBAC) system, collaborating with the team on pricing strategy, user needs, and security requirements
- Built an internal page builder using a Python-based data model to streamline content creation and launched Neliti's ad platform, integrating with frontend technologies like AlpineJS for a seamless user experience

Full Stack Developer — January 2022 – June 2022
Suvit Fintech Pvt. Ltd. — *Surat, India*

- Collaborated with cross-functional team to automate accounting data entry using Node.js, React, DynamoDB, and GraphQL, achieving high-throughput processing (10,000+ rows in less than 2 minutes) and reducing manual effort
- Improved React Table rendering speed by 27% by leveraging React DevTools Profiler to identify bottlenecks, applying useMemo and useCallback for memoization, and optimizing cell rendering logic

Projects

NotebookLM

- Developed a multi-capability Q&A document intelligence agent with Python/FastAPI, React, LangChain, Pinecone, Redis, and Google Gemini 2.0 Flash, enabling advanced cross-document queries, automated source citations, and a scalable agentic RAG design

Multi-Agent Review Analysis for Restaurant Scoring

- Built a multi-agent workflow using AutoGen and GPT-4o-mini to extract restaurant ratings from unstructured text reviews, processing over 10,000 reviews with a 96.67% accuracy rate in identifying food and service quality scores

Tuning Large Language Model for GSM8K(Grade School Math 8K) Dataset

- Achieved 94.24% test accuracy fine-tuning Llama-2-7b-hf on GSM8K utilizing QLoRA and PEFT with A100 GPUs, cutting training time from 10+ hours to under 30 minutes; used WandB for monitoring

Elastic Face Recognition Application

- Developed a scalable face recognition application using Python/FastAPI on AWS, leveraging EC2, S3, CloudWatch, and Elastic IP to optimize resources and handle 1000 concurrent requests with deep learning models