

KAMLESH CHHIPA

Bangalore, India • +91-8432386605 • kamleshchhipa480@gmail.com • [linkedin.com/in/kamlesh-chhipa/](https://www.linkedin.com/in/kamlesh-chhipa/)

PROFESSIONAL EXPERIENCE

BrightEdge, Bangalore

Jun 2025 – Present

Software Development Engineer 2 (SDE-II)

- Architected and scaled the core LLM processing infrastructure for AI Hyper Cube using Python, ClickHouse and MySQL, processing 5B+ tokens per month to run scoped LLM prompts that extracted intent, brand entities, sentiment, and citations from AI-generated prompts at production scale.
- Engineered a GPU-based LLM inference platform using vLLM, Vast.ai, and open-source models for large-scale entity extraction, reducing monthly AI processing costs by 62% (\$40K → \$15K).
- Built pre- and post-processing pipelines for AI Hyper Cube, processing 100M+ keyword SERPs using Python and BigQuery for prompt generation and Spark for post-processing with brand-level hash partitioning (MD5 modulo), enabling reliable analytics and UI consumption.
- Built a scalable Trino-Iceberg-ClickHouse data pipeline leveraging Python, GCS, Hive metadata, JavaScript UDFs, and Kubernetes. Created a custom memory-aware job scheduler and replaced the legacy BigQuery UDF process, achieving \$18K-\$20K in monthly savings while processing 1B+ events per month.
- Led the development of DCX Collector V2 leveraging Python, FastAPI, RabbitMQ, Redis, Docker, Argo Workflows, and BigQuery. Achieved a 4x faster processing speed, enabled seamless processing of 100M+ keywords per month across multiple vendor integrations.
- Implemented an automated AIO stitching system using Python, JavaScript, BigQuery UDFs, and Argo, mitigating Google num=100 deprecation and reducing DCX keyword collection costs ~60% across 25+ locales.

Hevo Data, Bangalore

Feb 2022 – May 2025

Full Stack Developer

- Engineered the Snowflake Partner Connect feature by exposing a secure POST API with a JWT fingerprint for automated destination creation, driving an additional 10-15 MQLs per month.
- Optimized the Signup API, reducing response time from 9 seconds to 2-2.5 seconds (~75% decrease) and increasing signup conversion rates by 4-5%, enhancing the overall user experience.
- Led end-to-end development of a secure referral portal using JavaScript Promises to sync job submissions across the database, webhooks, and Lever API, increasing referral engagement by 40%.
- Migrated the Django website frontend to a Next.js app, incorporating TypeScript and Tailwind CSS, and Dockerized the app, achieving a 15% improvement in application load times.
- Implemented A/B testing on high-conversion pages (signup, demo) for a Django website using Jinja templates, HTML, SCSS, and JavaScript, resulting in a 20% increase in demo requests.
- Enhanced Core Web Vitals by deploying Varnish in-memory caching (improving good LCP URLs from 55.73% to 65.23%) and migrating to the CloudFront CDN, further raising good LCP URLs to 83.85%.

DUIT Technologies, Delhi

Sep 2020 – Jan 2022

Software Engineer Intern

- Created an ETL pipeline using Google Apps Script that fetched real-time data from a stock API and loaded it into Firebase and Google BigQuery, resulting in a 20% reduction in project costs.
- Evaluated 10+ Machine Learning APIs, including Google Video Intelligence API, and Improved backend architecture with Python and MySQL, achieving ~90% accuracy in video analysis.

EDUCATION

Jaypee Institute of Information Technology, Noida

Jul 2018 – May 2022

Bachelor of Technology [B.Tech] (Computer Science and Engineering)

SKILLS & ACHIEVEMENTS

Backend: Data Structures and Algorithms, Python, FastAPI, Django, REST APIs, Django REST Framework, ORM, MySQL, PostgreSQL, Redis, Authentication & Authorization, OOPs, JSON, RabbitMQ, Trino, Iceberg, ClickHouse, Hive metadata, Argo Workflows, Apache Spark, Bigquery

Frontend: HTML, CSS, JavaScript, TypeScript, React.js, Next.js, Vue.js, Redux, Tailwind CSS, Ajax

DevOps: Git, Docker, CI/CD, Jenkins, Ansible, CircleCI, AWS (EC2, S3, CloudFront, ECS), Google Cloud, Kubernetes

Achievements: Recognized as Star Player (Oct 2025) at BrightEdge; 1st place, 2023 Hackathon; 5-star rating in SQL on HackerRank; Intern of the Month, DUIT Technologies.