

Nakul Chamariya

+91 7021590832 — nakulchamariya74@gmail.com — LinkedIn — GitHub — Upwork

TECHNICAL SKILLS

Languages: Python, Tensorflow, **Data Science:** Preprocessing, Feature engineering, Visualization, Statistical analysis, **Machine Learning:** Supervised, Unsupervised, Semi-Supervised, **Deep Learning:** ANN, CNN, RNN, **Gen-ai:** LLAMA 2, GPT-3, GPT-4, Stable diffusion, **NLP:** Transformers, Chatbot, **Computer vision:** Object detection, Object segmentation, OCR, **Speech:** Speech-to-text, Text-to-speech, **Recommendation Systems:** Content based, Collaborative filtering, Matrix factorization, **MLops:** Github Actions, CircleCI, Docker, DVC, **Web App:** Plotly Dash, Streamlit, PyWebIO, Gradio, **Web scrapping:** Selenium, BeautifulSoup, Scrapy, **API:** Flask, FastAPI, **Database:** MySQL, MongoDB, SQLite, **AWS:** EC2, ECR, S3, Lambda, ECS, App runner, API Gateway, **Other:** Linux, Git

LIBRARIES USED

Machine Learning: Numpy, Pandas, Scikit-learn, XGBoost, LightGBM, **Deep Learning:** TensorFlow, Keras, PyTorch, OpenCV, **NLP:** NLTK, SpaCy, Gensim, Hugging Face Transformers, **Computer Vision:** OpenCV, Pillow, Matplotlib, **Speech-to-Text:** OpenAI Whisper, fairseq, **Gen AI:** OpenAI, Langchain, Llamaindex

WORK EXPERIENCE

MySellerCentral

Feb'24 – Till date

Senior Data Scientist

Mumbai, India

Skills Used: AWS, Deep Learning, Computer Vision, NLP, Gen-AI, Web Scraping, MLOps, API

- Led the setup and management of **production** and **development environments**, including **GitHub repositories** with stage, dev, and prod branches. Developed **CI/CD pipelines** using **Jenkins** and **GitHub Actions** for seamless deployment across environments.
- Implemented **containerization** for all modules with **Docker**, optimizing with **multi-stage builds** and **Alpine images**, achieving up to **5x cost savings** and **10x latency improvements**.
- Configured **monitoring** with **Grafana** and **Prometheus** at both instance and module levels, setting up automated **alerts** for issue detection.
- Developed **AWS SNS** and **Lambda functions** to automate notifications for **API key exhaustion** (OpenAI, Keepa, Zyte) and improved security using **AWS Secrets Manager**, eliminating exposure of sensitive access credentials.
- Optimized **resource management** with **Celery** and **SQS**, implementing **queue-based request handling** to **prevent server-side bottlenecks**, and integrated **Ray** for enhanced resource utilization in both production and development environments.
- Designed advanced **scraping solutions** to gather product reviews from Amazon, utilizing **proxy rotation** for efficiency.
- Built and deployed **AI-based content generation modules** on **ECS Fargate**, enabling sellers to upload **audio** or **text** in any local language to create optimized content (titles, descriptions, keywords) for different marketplaces such as **Amazon**, **Flipkart**, and **Shopify**. Set up **autoscaling**, **load balancing**, and **CI/CD workflows** for these deployments.
- Developed **content grading modules** to assess and provide improvement suggestions for marketplace listings, along with **AI-powered scraping tools**.
- Engineered **infographic**, **lifestyle image**, and **video generation modules** for sellers, deploying on **RunPod serverless GPU instances**, saving over **\$1,000 monthly**.
- Created an **SQL-based sales analytics module** that responds to **audio queries** in various local languages, providing **data-driven insights** with **visualizations** and **detailed reports**.
- Developed a **sentiment analysis module** to evaluate product attributes from customer reviews, enhancing prediction speed from **20 minutes to 2 minutes** through **multithreading** and **multiprocessing** with LLM API calls.

- Fine-tuned **LLaMA 3 8B** to build a custom **chatbot** trained on company-specific data, deployed on **AWS SageMaker** to handle seller queries efficiently.
- Extensively leveraged **Bedrock** and **OpenAI GPT** for module development, implementing end-to-end pipelines with comprehensive **monitoring** and **alert systems** to ensure high availability and performance.

goGLOCAL

Data Scientist

Feb'23 – Feb'24

Mumbai, India

Skills Used: AWS, Deep Learning, Computer Vision, NLP, Gen-AI, Web Scraping, MLOps, API

- Extracted **real-time insights** from various websites using **advanced web scraping** techniques.
- Designed a **listing tool** integrating **object detection** and **OCR** for evaluating brand listing strength.
- Optimized listings with **Gen-AI** for better **title**, **description**, and **bullet points** scores.
- Improved image-based listings with **object segmentation** and **background generation**, enhancing listing scores.
- Conducted **bidding optimization** for improved product rank and visibility.
- Developed an **attribute-based sentiment analysis** model for actionable product feedback.
- Created a **multilingual translation tool** for brand SKUs across different countries.
- Implemented a **question-answering bot** for compliance and guidelines using **Llama2** and **Langchain**.
- Enhanced an **e-commerce chatbot** with **conversation insights** and **history modules**.
- Engineered a **real-time price tracker** with **automated email notifications**.
- Developed a **sales forecasting model** for predicting future sales and inventory.
- Designed a **strategic insights dashboard** for category-level marketplace insights.
- Implemented **image translation** to automatically translate text on images using **object detection** and **OCR**.
- Engineered **APIs** for models with **Celery** and **Redis** for concurrent request handling and set up a **CI/CD pipeline**.

Edvak Technologies INC

Data Scientist

Dec'21 – Jan'23

Hyderabad, India

Skills Used: AWS, Machine Learning, Deep Learning, NLP, Speech-to-Text, MLOps, API

- Revolutionized **lab test and medication recommendations** with a **multi-label classification** system.
- Enhanced text accuracy with **Diagnosis Prediction** using **Clinical-Bert** and **Clinical-NER** for medical coding.
- Improved **Speech-to-Text** precision with a custom **spell checker**.
- Developed a **patient attendance prediction** model for clinic show-up insights.
- Streamlined information with a **coreference resolution** model.
- Predicted outpatient visit volumes using a **time-series forecasting** model with **Facebook Prophet**.
- Anticipated consultation durations with a **service time prediction** model.
- Orchestrated **deployment** to **AWS ECR** through Dockerization for scalable access.

ACHIEVEMENTS

- **100 percent success rate** on **upwork** with **5 star ratings** from all clients

EDUCATION

- **2018 - 2022:** B.Tech. in **Electronics and Telecommunication Engineering**, K.J. Somaiya College Of Engineering
- **2016 - 2018:** HSC in Science, Swami Vivekanand Junior College Of Science & Commerce