Yogendra Yatnalkar

Mobile: **9167483290/7400110228** | Email: **yogenyat@gmail.com** | Medium: **@yogenyat**

Education:

BE in Information Technology
Mumbai University (Fr. CRCE) – **(2016 - 2020)**
CGPA: 8.75

Experience:

Quantiphi Analytics:
- Worked on **7 POC's, 1 Production deployment, 1 Product & 4 RnD.**
- Other responsibilities include: **Sales Engineer (8+ sales deal)**, mentor and interviewer.
  - **Sr. Machine Learning Engineer** (March '22 - Present)
  - **Machine Learning Engineer** (August '20 - Feb '22)

Budsta Analytics:
Anomaly and breakdown detection on industrial machine video sequences
- **Computer Vision Intern** (January '19 - March '19)

Certification:

AWS:
- **Machine Learning Speciality** (Dec '21)
- Solution Architect Associate (Nov '20)

GCP:
- Associate Cloud Engineer (Dec '20)

Nvidia:
- **2 certificates** –→Numba, Triton

Awards and Recognition:

Quantiphi:
- **Trained 16+ interns** in 3 years
- **Top 5** campus **ML Interviewe**r
- **Award:** Q's Think-Tank (3 times)

BE College (Fr.CRCE):
- Excellence in **Applied Mathematics**
- **Team-Leader** in **SIH** 2019 (**Worlds Largest Hackathon**)

Extra-Curricular:
- **AWS Community Builder** (2022 - Present)

Skills:
Core Skills: **Computer Vision**, ML on **AWS**, **MLOps**, **SageMaker,** Drift & Explainability, SQL, **Churn** Prediction, Docker
Experience on: Python, **Pytorch**, TF2, **OpenCV**, **XGBoost**, Shapley, AWS, GCP

Industry Projects: (POC - Proof of Concept, CV - Computer Vision)

Car Image Processing - Saliency segmentation, GAN, Matting & more:
- **Segment** the salient car from the parking lot –→ **add shadow** below it –→ **register** it on any background template.
- **Tagging:** Led a team of **4 taggers** for **5 weeks**. Tagged **7000+ images**.
- **Trained** segmentation model **(U2Net)** for **7 days** on a **single GPU.**
- Researched and trained a custom **GAN model for shadow generation.**
- **Multi-GPU training** using Horovod and Pytorch DP/DDP
- **Turned 80 thousand dollars POC to 1 Million+ dollars project**
- **Keywords**: CV-POC, Segmentation, GAN, **OpenCV**, **Pytorch**, **AWS**, Tagging

NeuralOps - Quantiphi's MLOps Product on AWS cloud:
- **Core Member** during product development. Developed **end-to-end MLOps components** and **pipelines** for image classification tasks using **AWS** (mainly **SageMaker**) and **Airflow**.
- **MLOps Components**: Processing, Training, Monitoring & Explainability
- **MLOps Pipelines**: Training, Batch & Real-time inference, Drift Detection
- Used **KS Statistic test** and **Entropy** for Computer Vision **drift detection**
- **Guiding** the **Q's GCP team** for migrating this product to GCP cloud
- Technology Stack: AWS (**SageMaker**), **Airflow**, SageMaker Pipelines, TF2

Churn Prediction for a global Ed-Tech firm:
- **Problem:** Predict if student will fail the exam in next 2 weeks and churn
- Processed more than **1.5 TB of CSV data** using **SQL on AWS Athena**
- Developed multiple **Gradient Boosting models** like XGBoost, LightGBM and achieved **82% accuracy, 81% recall and 90% precision**
- Generated **explainability** using **Shapley values**

Document Translation Pipeline on GCP Cloud: **(Production Deployed)**
- Input Document –→ XML Parsing to extract text –→ Translate using **GCP translate API** –→ **document processing** for translated document to have the exact same UI
- Developed a **unique** solution for **server crashing** using **Python Multiprocessing.**
- Developed a solution to **process "docx" files** which client preferred over **GCP's advanced Translation API**
- **Keywords**: LXML parsing and processing, **GCP** Vision and Translate API

Media Industry: Large Scale Audio-Video Deduplication
- **Deduplication of** audio and video assets across **90 TB of data**
- Extensively used **FFMPEG on GPU** for separating data-streams, video-transcoding and converting audio to stereo-audio to mono-audio.
- Used **FAISS library** for efficient storing and searching of frame embeddings
- Performed clustering on **similarity-matrix of 2 videos**

Few Other POC's:
- Web-Page **web-elements identification** and its **contrast validation**
- **Brand identification** on wide ad boards in sports game **video streams**
- **AWS Immersion Day:** Training client-partners on AWS SageMake

Industry Research:
- **Multi-Task Learning:** Tensorflow 2 vs Pytorch
- **Backtracking AWS** Lookout For Vision service: (AWS recognized me for this)
- Explored **Tensorflow 3D**, **Entropy** for Drift-Detection
- **NLP:** Hosting of LMs on **AWS Inferentia Instance vs Nvidia GPUs**
- Researching capabilities for **DinoV2, SAM and ImageBind ML models**