JAY SHAH

Surat, GJ | **Contact:** +91 9638581669 |**Mail**: J.59shah@icloud.com |
LinkedIn: https://www.linkedin.com/in/jay-shah-742751161 **| GitHub**: https://github.com/JayyShah

SUMMARY

Experienced Software Developer with a strong background in Cloud Computing, Data Engineering, and Computer Vision. Passionate about emerging technologies and their real-world applications, particularly interested in AI-driven innovation. Seeking a challenging role as Computer Vision Engineer/Developer, Data Engineer, and Data Scientist/Machine Learning to contribute to innovative product development projects.

EDUCATION

UKA TARSADIA UNIVERSITY Bardoli, Gujarat
B. Tech Mechanical Engineering 2016-2020

SKILLS

Programming: Python
Data Analysis: Numpy, Pandas, Matplotlib, Scikit-Learn
Cloud Services: AWS (Compute, Storage, Network, Databases), Terraform
Cloud Analytics: Kinesis Stream, Analytics, EMR, Glue, Athena, Quick Sight, IOT analytics, RedShift, Hadoop Fundamentals, Serverless Analytics
Containers: Docker, Kubernetes
Machine Learning: Algorithms, Linear Algebra, Hugging Face, Transformers (BERT, SAM, Fast-Sam, Yolo Series, DETR, VIT, Grounding DinoV2, CLIP)
Certifications & Training: Machine Learning Operation (MLOPS) Challenge – Microsoft,
Machine Learning Challenge – Microsoft,
ChatGPT Prompt Engineering – DeepLearning.Ai,
Training – AWS Cloud Practitioner

WORK EXPERIENCE

Raj Vijtech Pvt Ltd Surat, GJ
Data Analytics Project Intern Jan 2022 – March 2022

- Collaborated with the business development team to support Marketing analysis through data analysis.
- Conducted Exploratory Data analysis and employed visualization techniques to uncover correlations.
- Developed a Recency, Frequency, and Monetary (RFM) analysis model in Python to analyse churn rate, customer lifetime value, and Customer Segmentation.

Triveni Global Software Services LLP Surat, GJ
Associate Software Developer March 2022 – July 2023

- **Factory Big Data:**
  - Created AWS/Azure Infrastructure using Terraform.
  - Actively contributed to MLOps to align expertise with business KPIs/goals.
  - Designed Data Streaming Pipelines on Databricks using PySpark.
  - Ensured data quality and delivery.
  - Developed and deployed a layer-4 Load Balancer (TCP) RKE2 Cluster in Kubernetes using Terraform.
  - Utilized MLFlow for Experiment Tracking and Model Registry, along with BentoML.
- **HRIS Bot:**
  - Hands-on improving NLP skills and learning new NLP concepts con-currently.
  - Demonstrated Fine-tuning Albert QA on custom dataset.
  - Implemented Rasa-Framework for handling pre-defined actions and unmatched inputs.
  - Inference the Chatbot by creating a Web Application in Flask.

PROJECTS

CosmoVision

- CosmoVision is a One-Stop Solution for all the Computer Vision related Problems, which includes a Copilot which uses Blog posts, document pages to answer questions about Computer Vision. A Semi-automated Data Annotation tool which is emphasized on bridging the gap between Human-Machine Intelligence (Currently under Development).

Automated Data Pipelines Using Terraform

- The process involved three steps, controlled by a local Airflow instance. These steps included downloading and uploading the PDF file to S3 storage, extracting texts from the PDFs using a Lambda function, and segmenting the extracted text into questions using a Glue Job.

CI/CD Pipeline In AWS:

- This Project Aims to create a CI/CD Pipeline in AWS For ECS cluster using the following services: - AWS CodeCommit - AWS CodeBuilt - AWS CodeDeploy - AWS CodePipeline