

MANSI DHRUV | Resume

Summary

A meticulous professional adept at crafting, refining, and fortifying highly scalable, secure, and dependable data structures. Collaborates closely with system architects, software architects, data scientists and design analysts to grasp business or industry needs, crafting comprehensive data models in response. Demonstrates proficiency in crafting cloud architectural strategies throughout the entire lifecycle, from design to development and implementation phases, while aspiring to advance into a technical architect role in the near future.

Experience: 4+ Years

Email: mansi.p.dhruv@gmail.com

Github: <https://github.com/msmansidhruv>

LinkedIn: <https://in.linkedin.com/in/mansidhruv>

Core Competencies

Advanced Analytics ELT/ETL	Data Curating Data Engineering	Dev Ops/ML Ops Data Scraping	Cloud Computing	Architecture Design	Infrastructure as Development/Deployment
-------------------------------	-----------------------------------	---------------------------------	-----------------	---------------------	--

Tools and Libraries

AWS Lambda AWS Redshift REST API/FAST API Requests/BS4 PostgreSQL	AWS IAM Role/Policies/User AWS CloudFormation/Terraform GraphQL Scrapy MySQL	AWS EC2 AWS Kinesis Celery Selenium Elasticsearch	AWS ECS Boto3 Django/Jinja/HTML Databricks MongoDB	AWS Glue Docker Javascript Pyspark/Scala	AWS DynamoDb CI/CD Pipelines Python ETL
---	--	---	--	---	--

Soft Skills

Deadline oriented Challenge driven	Leadership Team Player	Decision Making Resilient	Self-Motivated Analytical thinking	Time Management	Strong work ethics
---------------------------------------	---------------------------	------------------------------	---------------------------------------	-----------------	--------------------

Certifications

Aug 22 - [AWS Certified Cloud Practitioner \(CLF\)](#)

Feb 22 - Building your first ETL Pipeline using Azure Databricks

Dec 21 - Big Data Analytics using Spark

Sep 20 - [Machine Learning using Python](#)

Dec 19 - [Docker Essentials](#)

Jun 18 - [Big Data Training](#)

Education

(Integrated) Master of Computer Applications | School of Computer Studies, Ahmedabad University, A'bad - 59, GUJARAT

Score: 3.94 / 4.33 | 2014-2019

H.S.C Science | C. U. Shah Higher Secondary School, A'bad - 59, GUJARAT

Score: 64.46% | March, 2014

S.S.C | Vishwabharati Bhavan, A'bad – 59, GUJARAT

Score: 79.40 % | March, 2012

Achievements

2024, Valuable Award

2023, Gem Award

7th – 8th Oct, 18

October, 18

Sept, 18

Received **Value-able award** for playing a key role thus bringing new business from clients.

Received **GEM Award** for delivering exceptional business value while going extra mile for customer satisfaction.

Awarded Special mention for "**Innovative Project**" in ACM-W Third National Level Hackathon.

Awarded Student Scholarship for attending **Grace Hopper Celebrations India (GHCI – 18)**, India's largest technical conference for women in computing and technology.

Awarded Full Year **Merit Based Scholarship** of INR 73000.

Work Experience

Data Engineer | Oct 21 – Current | Employee

S. G. Analytics | Maharashtra, India

- 2024 • Developed ETL Pipelines in **Glue, Pyspark and Redshift** for a global Fintech company
- Created APIs using FastAPI and deployed the APIs using AWS API Gateway while fetching APIs from third-party sources
 - Performed root cause analysis for SQL insertion durations via fishbone diagram
 - Identified query optimization opportunities to improve time efficiency by 85%
 - Improved development logic by removing duplicates from results using fuzzywuzzy library
- Tech: Pyspark, Python, AWS Glue, FastAPI, AWS API Gateway, AWS Lambda, Fuzzy logic
- 2024 • Mentored juniors in **Python and Scraping**: Trained interns and junior staff in both fundamental and advanced Python programming concepts, as well as web scraping and automation techniques.
- Developed timelines for 4 new initiatives, allotted work to team members, and ensured completion of tasks within stipulated timelines
- 2024 • Created and developed **ETL Pipelines and AWS Infrastructure** for leading Pharma company
- Developed end to end infrastructure for multiple ETL Pipelines connected with AWS Glue, Redshift and Tableau for dashboard purpose. Involved with creating transformations for building multiple machine learning models.
- Automated ETL process by developing scripts to ease data and code integration and reduced data processing time by 45%
- Tech: Python, Selenium, Terraform, YAML, Bitbucket pipelines, Cron Scheduling, AWS
- 2023-2024 • Extracted news articles from IT and Business-related websites by means of automation using Selenium and Scrapy. Automated the process and deployed the end-to-end solution on AWS Cloud.
- Coordinated with clients and reduced the workload by 50%
 - Developed CI/CD pipelines using BitBucket workflows to automate the deployment
 - Created, developed and deployed technical architecture for the developed solution
- Tech: Python, Selenium, Terraform, YAML, Bitbucket pipelines, Cron Scheduling, AWS

MANSI DHRUV | Resume

- **Deployed AI/ML based models** for a leading Fintech firm.
Designed and deployed ETL pipelines as frameworks for fetching streaming and batch model related data from multiple streams i.e Delta Live tables, Redshift and Kafka. Worked with MLFlow to maintain full life cycle for created ML models.
 - Helped client reach 40% reduction in the end-to-end pipeline time
 - Reduced cost by 30% by shifting to AWS infrastructure from legacy Databricks setup.
 - Designed multiple webhooks, workflows in GIT for CI/CD integration with Harness and Shipmunk
 - Developed infrastructure using Cloudformation to leverage Infrastructure as Code (IaC) along with docker containers and port mapping
 - Converted legacy pyspark code into scala and recreated new ETL pipelines using Kafka

Tech: Databricks, PySpark, Scala, Python, Flask, AWS cloud, Docker, JIRA, GIT, Lambda, ECS, SQL, CloudFormation, Kafka, Kinesis, Dynamo, Delta live tables, Scheduling

Backend Engineer | Aug 19 – Oct 21 | Employee

CypherTree Technologies | Maharashtra, India

- 2020
2021
- Developed S.M.A.R.T (Social Media Analytics Reporting Tool) web portal which is an online self-service data science platform that provides insights of various brands and platforms on demand. It is an ecosystem featuring user-friendly apps and advanced analytics
 - Developed REST API for new functionalities. Reduced API Load time by 25 – 30 %
 - Utilized Agile/Scrum work flow and generated effective documentation

Tech: Python, Django, React, Redshift, Celery, JIRA, GIT, Task Scheduling

2019
2020

 - Developed Ecommerce website for fashion brand using Saleor along
 - Developed and designed techstack for an ecommerce website that sells clothes
 - Integrated different payment methods such as PayTM, Debit card, Credit card in existing framework
 - Utilized GraphQL APIs to deliver the required data thereby maintaining data minimization
 - Developed RESTful web service with Json Web Tokens (JWT) Authentication

Tech: Elasticsearch, Python, Django, JavaScript, Docker, JIRA, GIT, GraphQL.

Full Stack Developer | Dec 18 – July 19 | Intern

Security Agency | Gujarat, India

- The system “TALAASH” indicates gathering data from social media and back tracing to find the root cause/source of information using automation and sentiment analysis. Alongside “TALAASH” also matches missing children with the database using facial recognition technique.
- Developed data crawling mechanisms and structured the unstructured social media data
 - Incorporated sentiment analysis to filter the post/message as positive or negative.
 - Collaborated with the team for implementing facial recognition in a client-server architecture.
- Tech: Django, Python, JavaScript, Ionic

Academic Experience

Feature Extraction and Analysis

Developed the system containing two different modules: **Object Detection & Human Detection and Recognition**. It automatically marks student attendance by uploading a group photo of class. It is capable of extracting and recognizing faces from even 15-18 yr old group photos / video.

Tech: OpenCV, Python, Django, Kairos, Haarcascade

Smart Band

Developed a smart Band i.e. a proposed system specifically designed to track people through GPS. Developed the band to provide cashless payment via NFC, check for pulse rate and broadcast special messages. It included a client server architecture with different bands connecting to each other via Bluetooth.

Tech and Hardware: Raspberry Pi 3, Pulse Rate Sensor, GPS, NFC, Python, Django.