

HIMANSHU SURANA

+91 9328412428 | himanshusurana1998@gmail.com

India

EDUCATION

Dharmsinh Desai University, Nadiad
Bachelor of Technology, Information Technology

July 2015 - May 2019
Cumulative GPA: 8.2/10

TECHNICAL SKILLS

Programming Languages	Python, Scala, C, C++, Java, GoLang
Databases	MySQL, MongoDB, SQLite, DynamoDB, neo4j, Graph Database, RDS, Redshift, S3, Glue, Athena, Hive
Framework	Airflow, Kafka, Hudi, Sqoop, Hadoop, Spark, Flask, FastAPI, Azure Databricks, Redis, Snowflake
Miscellaneous	AWS Services, GCP, Azure, Git, BitBucket, Docker, Kubernetes

WORK EXPERIENCE

Data Engineer | KnowledgeFoundry, India

June 2019 - Present

- Automated process for writing hive queries for ETL of multiple tables (both one time and incremental) by generating this automated scripts.
- The process reads csv files from folder location and creates tables, then does the incremental load sequentially.
- In our project, we set up Snowflake as the go-to spot for storing our structured data, while we put DBT to work for our ETL needs. With DBT, we crafted SQL-based models that laid out the transformation game plan and business logic. We made sure to keep things nimble with incremental loading and version control using DBT, making our data processing slick and teamwork seamless.
- Now, our transformed data is ready for analysis through cool business intelligence tools, letting us dig into insights effortlessly. We keep a close eye on things, running checks in both Snowflake and DBT to make sure our data is top-notch and our pipeline stays in good shape.
- Working closely with product team to improve Dunning model for increase ship rate and higher recover, resulting increased revenue.
- Designed and developed data pipelines to extract, transform, and load data from various sources into a centralized data warehouse.

Data Engineer Consultant (Contract) | InfoGain, India

June 2021 - Apr 2022

- Created Data Warehouse solution using AWS Redshift and AWS glue. This was OLAP database migrated from MSSQL server.
- Created DBT pipeline for ETL to move data between MySQL warehouse and activity database to neo4j graph database. This was done in native python programming. Setup was done on AWS linux box with neo4j running as a docker container.
- Built ETL pipeline to perform market basket analysis and other marketing statistics on millions of rows of transactional data. Redshift was used as a transactional database and was populated in serverless fashion using Amazon lambda function in real time.

Data Engineer (Contract) | ZS Associates, Pune, India

Oct 2020 - Apr 2021

- Pipeline to convert the data to structured form and be servable to prodigy for doing ML related tagging. Entire pipeline was built in modular fashion using pure python and shell scripting. Data transformations were done in Python and data was dumped in s3.

KEY PROJECTS

Realtime Employee Management | [Link](#)

- Runner up of EY GDS hackathon. Used AI for better skill matching from job request to employee pool. This algorithm can potentially save hours of Manual work for enterprise hiring needs.
- Data was gathered from linkedin, scraping pipeline built on python to gather relevant information from profiles. This was later used to serve as data for the Artificial Intelligence algorithm.

Realtime Football Stats (AWS Lambda, API Gateway) | [Link](#)

- Hosted a cloud serverless API endpoint on AWS to serve football stats generated from Machine Learning algorithm.
- Implemented using SSD and yolo. System capable of running on CPU at 35fps. System can detect passes, kick. This system can be used to generate metrics (potentially used by fantasy sport sites, betting agencies, clubs).

Video Streaming Platform (NoSQL, cron & Flask)

- Python scraper using BeautifulSoup and urllib to fetch the streamable video links from the internet and insert it into NoSQL database.
- Used Redis caching to reduce database hits while serving the videos. Entire system was encapsulated in Flask website running as microservice on a Docker container and nginx container used for proper routing of these requests.

Find Meritorious Students (PhpMyAdmin, PHP, AWS, MySQL) | [Github](#)

- Designed a web-app that handles data of a million high school students across state.
- Defined a point system to find meritorious students based on achievements, age, etc.
- Made a dashboard to search using specific criteria for specific results

AWARDS AND HONORS

- [Winner](#) - EY Smart AI Hackathon- 2020, judged by entrepreneurs from E&Y.
- [Winner](#) - Regional round and **Finalists** - National round of the SSIP SMART INDIA HACKATHON 2017. [Link](#)
- [All India Rank 24](#) in National Creativity Aptitude Test, INDIA. (NCAT). [NCAT](#)