

SATYAM GOVILA

govila.satyam@gmail.com ♦ (+91) 844-584-1444 ♦ New Delhi, India
satyamgovila.github.io/portfolio/ ♦ GitHub ♦ LinkedIn

WORK EXPERIENCE

- **Concirus Pvt Ltd** Full-time
Data Engineer(SDE-2) July 2019 - Present
 - **Role:** Worked as a Data Engineer in the Data Science team at Concirus.
 - **APIs for ETL pipeline:** Built APIs required for ETL process of data pipelines from different sources of data for both real-time and batch streaming processes, using Python, Spark and AWS technologies.
 - **Design Architecture:** Designed and developed LLDs and HLDs for developing optimal data pipeline architecture.
 - **Reduce Costs:** Performed auditing of AWS S3 buckets, EMR clusters and EC2 instances and assisted in implementing different methods using Lambda, removing Spark clusters and migration of scripts to EMR in order to save cost ; saved AWS cost in a quarter by 8%
 - **Discussion and Planning:** Worked with stakeholders and product team, to assist them in data related technical issues and regarding ad-hoc data requests.
 - **Real Time Stream Processing:** Refactored existing data pipelines to handle large amount of streaming JSON data , integrate error handling and automate it using Apache Airflow.
 - **SQL and Athena for Data query & analytics:** Worked on creating and developing a complete end to end process to extract and transform data and make SQL tables for analytics ,using AWS Athena.
 - **Feature Development:** Developed an automated end-to-end data pipeline to build a feature that gives insight into a vessel's history, showing specific events over the course of the lifetime of the vessel from scratch to meet the product requirements.
 - **Integrating Alerts:** Integrated automated Alerting System for Job Monitoring.
- **AAIC Technologies Private Limited , www.appliedroots.com** Remote
Mentor, AI and ML (Part-time, Contractual) July 2019 - Present
 - **Role:** Responsible for mentoring cohort by assisting them in their ML/DL projects , holding project sessions, monitoring discussion groups and grading submissions.
- **Remedo Clinitech Pvt. Limited, New Delhi** Intern
Front End Developer June 2017 - July 2017
 - **Role:** On-site internship under a leading health-tech startup, worked on web automation and data analytics using Javascript and Python

EDUCATION

- **University of Hyderabad**
PGD - Machine Learning and AI February 2021 - April 2022
Courses: Data Structures, Analysis Of Algorithms, Deep Learning, Machine Learning, Networking, Databases
- **ZHCET , Aligarh Muslim University**
Bachelor of Technology - Electronics Engineering July 2015 - July 2019

PROJECTS

- **Prediction of Defaulters in Banking org using different ML techniques:** Built a ML model using attributes such as general profile & historical records of a borrower to predict whether one is likely to have serious delinquency in the next 2 years. Tech: Python, ML, Model Monitoring
- **Developing Website Monitoring Sytem Architecture using AWS:** Built an architecture to analyse the log data in real time by using multiple data streaming and processing technologies. Also, integrated web app to visualise the complete process and analytics.Tech:Python ,Docker , Apache Spark, Apache NiFi, Apache Kafka, Plotly
- **Real Time Log Data Analytics using Lambda Architecture , Spark and Kafka:** Built an AWS architecture for real time website monitoring system that provides functionality to test availability and performance of website by streaming and analysing logs, and trigger Lambda in case of any event.Tech: Python, AWS services- Kinesis , Lambda , Aurora , Dynamo DB , SNS
- **Designing and building Pipeline for Auto Reply Twitter Application using LSTM :** Designed and developed an end-to-end pipeline for deep learning application that automatically replies to query-related tweets with a trackable ticket ID generated based on the query category predicted using DL models. Tech: Python, Kafka , Spark Streaming , Flask ,Deep Learning, AWS EC2 , Tweepy API
- **MLOPs Project : Building Auto-Regressor Model for Time-Series Analysis and Deploying on GCP :** Developed an end-to-end ML process which helps to analyse the time series data by using an auto-regressor model for forecasting, deployed on Google Cloud Platform(GCP). Tech: Python, Docker , Kubernetes , Flask , Google Cloud Platform
- **Thesis Project: ML Based Smartphone based Virtual Gym Assistant :** Built a Virtual Gym Assistant application as a part of thesis project , that provides a smart and yet affordable AI solution which not only helps people work out effectively at the comfort of their homes but also provides them with real-time feedback on their posture, and act as a personal virtual trainer that will help them to do their exercises in efficient manner. Tech: Python, AWS EC2 , Pose Estimation Techniques , Streamlit
- **Blogs (<https://satyamgovila.medium.com/>) :** Published various technical blogs on Medium regarding system design and deep learning.

SKILLS SUMMARY

- **Languages & Frameworks:** Python/SQL , Scikit, NLTK, SpaCy, Keras, Flask
- **Tools:** Kubernetes, Docker, GIT, PostgreSQL, MySQL
- **Big Data Frameworks:** Hadoop , Spark , Kafka , Linux, AWS Big Data technologies, GCP , Airflow