

Sanket Patole

LinkedIn - Sanket Patole | GitHub - SanketPatole
Gmail - snkt.pat2@gmail.com | +919561333389

SKILLS

PROGRAMMING

- Python • Unix Shell • Excel-VBA
- Gitlab-CICD • AWS-Cloudformation
- Rest APIs

DATA SCIENCE

- Supervised Learning
- Unsupervised Learning
- Computer Vision • Scikit-Learn
- Keras • Statistics • EDA
- Feature Engineering

DATA ENGINEERING

- SQL • Snowflake
- Pyspark • AWS-S3 • AWS-Lambda
- AWS-State Machine • Docker
- AWS-ECR/AWS-EKS
- SnowPipe • AWS-Fargate
- Abinitio ETL Tool • Abinitio Metadata Hub

OTHERS

- Jira • Scrum • Client Communication
- Technical Interviewing

CERTIFICATIONS

- AWS Certified Cloud Practitioner

EDUCATION

MIT PUNE

B.E. (COMPUTER SCIENCE)
May 2016 | Pune, India

NOWROSJEE WADIA COLLEGE

HSC (10+2)
May 2011 | Pune, India

COURSEWORK

- Ubiquitous Computing
- Operating Systems
- Artificial Intelligence
- Theory of Computation
- Data Structures and Algorithms
- Database Management Systems
- Compiler Designing

PROFESSIONAL EXPERIENCE

PERSISTENT SYSTEMS

Engineering Lead

PUNE, INDIA

Apr 2022 - Present

Python | PostgreSQL | Snowflake | AWS | Gitlab CICD | DBT | Pytest

- Built marketing data pipeline for over 100+ schools in USA using SQL, DBT and the cloud based in-house ETL application.
- Worked on the enhancements and automated deployment of the cloud based in-house ETL application suite (using Python, AWS and Gitlab CICD).

BITWISE INDIA

Programmer Analyst

PUNE, INDIA

Oct 2018 - Apr 2022

Pyspark | Abinitio ETL tool | Abinitio Metadata Hub | Shell Scripting | Python | Excel-VBA | Teradata | Snowflake | Rest APIs | Pyspark

- Built and deployed the entire data governance pipeline from scratch.
- Built and deployed the entire metadata quality-check pipeline from scratch.
- Migrated 2500+ Teradata tables along with ETL operations from on-prem to AWS cloud platform with Pyspark.

FIGMD INDIA

Associate Software Engineer

PUNE, INDIA

Nov 2016 - Oct 2018

MSSQL | Python | Advanced Excel | SSIS ETL tool

- MSSQL | Python | Advanced Excel | SSIS ETL tool.
- Worked on extracting and cleaning raw clinical data for 200+ hospitals in USA.
- Wrote sql mapping scripts for 1000+ complex clinical data elements and for 200+ hospitals.

DATA SCIENCE PROJECTS

VEGETABLE IMAGE CLASSIFIER | CNN & TRANSFER LEARNING

- Github Link: [Ninjacart - CNN & Transfer Learning](#)

DRIVER ATTRITION PREDICTOR | ENSEMBLE LEARNING

- Github Link: [OLA - Ensemble Learning](#)

WIKIPEDIA VIEWS PREDICTOR | TIME SERIES FORECASTING

- Github Link: [Adease - Time Series Forecasting](#)

MOVIE RECOMMENDER | RECOMMENDER SYSTEMS

- Github Link: [Zee - Recommender Systems](#)

PUBLICATIONS

- [1] S. Z. T. S. S. Sanket Patole, Mangesh Bedekar. Analyzing bookmarks in web-browsers: Inferring user interests from bookmarks using tag clouds. ICCUBE, 2017.