Arpit Solanki

Delhi, India
8238059340
solankiarpit1997@gmail.com
https://github.com/arpit1997

I have experience with data, infrastructure, backends, and managing projects like <u>DISHA</u>. *I have worked with data cataloging, versioning, workflow orchestration, data ingestion, efficient data storage, monitoring infrastructure, on-premise deployment, writing backends, one to one mentoring/teaching.*

Technical Skills

Likes: python airflow prestodb hive dask apache-spark parquet workflow docker data-ingestion rabbitmq linux prometheus git

---

Experience

Data Engineer – <u>SocialCops</u>   *Jun 2018 → Current*
airflow, prestodb, parquet, apache-spark, dask, hive, monitoring, prometheus, rabbitmq, docker

- Worked on deployments and development for the project <u>DISHA</u>
- Worked on data cataloging using Presto as a query engine and Hive as metastore.
- Worked on automated data reports to help data scientists and business users see how is data.
- Worked on a simple versioning system for data to keep track of changes in data.

Mentor – <u>Codementor</u>   *Jan 2019 → Mar 2019*
python, workflow, github, apache-spark, api

- Provide one to one live help on given problems/bugs to mentees
- Teach and guide on new technologies, fields or approaches on problem-solving.
- Have done over <u>100 sessions</u> with people with 5 star rating.

Data Engineering Intern – <u>SocialCops</u>   *Jan 2018 → May 2018*
python, celery, deployment, web-scraping, data-ingestion, workflow, airflow

- Worked on deployments and development for the project <u>DISHA</u>.
- Worked on data ingestion through APIs, databases, Object stores etc.
- Worked on ETL pipelines and workflow management with Airflow

Backend Developer Intern – <u>Vstv.io</u>   *Sep 2017 → Nov 2017*
node.js, express, mongodb, jwt, bcrypt, video-streaming, video-capture

I am working as a Backend developer intern with MEAN stack. My responsibilities includes creating web services, handling data streams to Backend and collaboration with frontend team.

Backend Developer – <u>MateLabs</u>   *May 2017 → Jul 2017*
python-2.7, django, django-rest-framework, orm, ifttt, postgresql, amazon-ec2, amazon-s3, boto3, docker, rabbitmq, celery, django-celery, luigi

Creating web services in Django, Integrating machine learning and deep learning pipelines, Improve Mateverse back end.

Full Stack Developer and Code Reviewer (Remote) – <u>BrainPlay Learning Solutions</u>   *May 2017 → Jul 2017*
node.js, express, mongodb, rest, pm2, mocha, chai

Database designs and optimisation, Testing functionality of rest-apis, finding possible bugs.

Education

B.Tech – Indian Institute of Information Technology Vadodara   *Aug 2014 → Apr 2018*

Human Detection and Tracking

In this Project we have worked on the problem of detecting and recognizing a human in a video which can be useful in surveillance systems and security applications.

MobileCloudIR

MobileCloudIR is a distributed system which stores the whole Wikipedia dump in different smartphones. Users can query and get related articles. This system includes an Android application and a coordinator web server running in the local network

---

Projects & Interests

Stack Overflow – https://stackoverflow.com/users/5250746/arpit-solanki   *Aug 2015 → Current*

Written 238 answers. Active in django, python, django-templates, node.js, python-3.x and 3 other tags.

Public Artifacts

Testing Celery Tasks – Arpit Solanki – Medium – https://medium.com/@solankiarpit1997/testing-celery-tasks-a8a3568d0213   *Feb 2018*

python, celery, task, scheduled-tasks

Recently I came across the problem of writing unit tests for Celery tasks. Writing unit tests for celery tasks can be painful since they…

Apps & Software

PQusic – http://github.com/arpit1997/pqusic   *Dec 2016*

python-3.4, django, beautifulsoup, html5, css3, plyr, html5-audio, javascript, jquery, materializecss

A music web application

Others

Won 2nd prize at Hackathon-2016 IIIT Vadodara – Acomplishment   *Mar 2016*

python-3.x, google-drive-sdk, command-line-interface, request, oauth-2.0

SyncIt is a cross-platform client to synchronize work with your Google Drive folder specifically targeting public PC's