SURYA SHEKHAR CHAKRABORTY

Atlanta, GA · schakraborty76@gatech.edu · 4044090874 · https://github.com/suryashekharc

EDUCATION

Georgia Institute of Technology   Atlanta, GA
MS Computer Science   Aug 2021 - May 2023

Vellore Institute of Technology   Vellore, India
BTech Computer Science and Engineering *GPA: 9.30 / 10*   Aug 2014 - May 2018

EXPERIENCE

Zendrive   Bangalore, India
Software Development Engineer 2   Jul 2019 - Jun 2021

- Geo-data for insurance scoring: Powered large-scale queries with 5x faster speed than Postgres to enrich petabytes of data with geospatial features such as reverse-geocoding and trip-segmentation.
- Stop-sign detection: Scaled and refactored a massive pipeline to detect stop-signs on 17,000,000 US roads with 95+% precision.
- Client dashboard: Owned and designed the Data Flow, Data Model, and APIs to render data of 100,000+ users with sub-second latency.
- Python 3 Migration: Made our massive Data Engineering codebase compatible to enable data scientists to seamlessly integrate modern ML packages.
- Presented data-driven insights directly to the Co-founder's office for pitching to clients.

Zendrive   Bangalore, India
Software Development Engineer 1   Jul 2018 - Jun 2019

- Automated large dataset generation and validation tasks, processing 200M+ rows of driving data.
- Created APIs to query sensor data at scale while abstracting Spark internals and optimizations.
- Technical mentor to 17 new engineers.

Zendrive   Bangalore, India
Intern   Jan 2018 - Jun 2018

- Developed an iOS app "SensorLab" to collect and label empirical data from 9 phone sensors. The app was used to develop 2 new production algorithms and in BMW crash tests in Munich.

Sibia Analytics   Kolkata, India
Intern   Dec 2016 - Dec 2016

- Integrated a client-facing dashboard with NoSQL backend to visualize real-time data for the leading news-provider in West Bengal.

SKILLS

Languages:	Python, R, Objective C, Bash, Java, C++, SQL, Mongo, Scala
Tools:	Git, Github, Gerrit, VS Code, PyCharm, Jupyter, Xcode, RStudio, Airflow, Jenkins
Cloud:	Elastic Beanstalk, EC2, S3, BigQuery, DataStudio
Libraries:	pyspark, pandas, numpy, geospark, flask, sqlalchemy, pytest

PROJECTS

Insta Reddit *Python, pandas, pillow, praw, gspread* https://github.com/suryashekharc/insta_reddit
Created a reusable automated pipeline which fetches top posts from Reddit, creates images from texts, then uploads them on an Instagram page. Wrote a blog post for Better Programming - Medium's best-performing publication of 2019.

Print Schema *Python, json*   https://github.com/suryashekharc/print_schema
Published a package to visualize the data-structure schema of complex Python objects for easy analyses.

AWARDS

Zen of Getting Stuff Done Award   Co-founder award, Jul 2019

Above and Beyond Award   Manager award, Jan 2019