



Debangshu Bhattacharjee

DATA ENGINEER

Objective

Passionate about learning technologies that run the data-driven world. Looking forward to being part of a team where I can learn and grow in my technical capabilities.

Experience

Data Engineering Intern, Vahan

JULY 2022 – DECEMBER 2022

- Added functionalities aiding the visibility of running queries and other features that enhance user experience to Vahan's fork of *Redash*.
- Migrated Vahan's on-premise *PostgreSQL* database to *AWS RDS* using *AirByte*.
- Added *OAuth* to Vahan's internal data processing tool.
- Wrote multiple scripts ranging from utility scripts, and automating downloads using *Selenium* to calculating critical data quality metrics.

Data Science Intern, Home.LLC

MARCH 2022 – JUNE 2022

- Analyzed large datasets using *Spark-SQL* and *Pandas* with *Python* to gain state-level insights about housing market factors and trends.
- Created data pipelines for monthly data collection from various APIs reducing the manual intervention needed to none.
- Created visualization to gain insights on the housing market using *Tableau* & *Python* (*Seaborn* and *Matplotlib*).
- Contributed to our Social Media Team (500+ Tweets) and Hiring Team (30+ Interviews) over a period of 3 months.

Projects

HashTag Live Analysis

- Uses *Apache Kafka*, *Apache Spark*, *MongoDB* & *Tweepy*.
- Computes sentiment of tweets over a window using *Apache Spark*.

MLOps with Dagster

- Used *Dagster*, *Sci-Kit Learn*, *Pandas*, and *BeautifulSoup*.
- End-to-End MLOps pipeline including downloading data to preprocessing to hyperparameter tuning.

Education

B.Tech. in CSE, Kalinga Institute of Industrial Technology

JUNE 2019 – APRIL 2023

Extra-curricular activities

Assistant Coordinator, KIIT Film Society

AUGUST 2021 – AUGUST 2022

Regularly managed a team of **100+** people, organized workshops with **40+** attendees on various parts of Film making, and a fest with **200+** attendees.

Details

Kolkata

India

+91 9748625028

debangshu133@gmail.com

Links

[LinkedIn](#)

[GitHub](#)

Skills

Python

SQL

Research / EDA

Docker

Machine Learning / Data Mining

Javascript

React

NodeJS

Tableau

Excel / Google Sheets

Git / GitHub

Apache Airflow

Apache Kafka

Apache Spark

Data Build Tool

Dagster

Amazon Web Services

Hobbies

Listening to Music, Guitar, Photography, Music Production, TV Series & Movies, Computer Games, Football, Cricket