

Arujit Pradhan

+91-9938207651(M) | arujit17@gmail.com | arujit.github.io

EDUCATION

NATIONAL INSTITUTE OF TECHNOLOGY, ASSAM
B.TECH - COMPUTER SCIENCE
2014 - 2018 | CGPA: 7.94

COURSEWORK

Operating Systems
Object Oriented Programming
Artificial Intelligence + ML
Data Structures and Algorithms
TOC and Compiler Design
Distributed Computing

SKILLS

PROGRAMMING

Strong:

- Java • Scala • Go
- Python • SQL

Familiar:

- C/C++ • CSS/HTML • NodeJS
- Matlab • Shell • LaTeX

FRAMEWORKS

- Flink • Spark • Kafka • Airflow
- Terraform • Presto • Scikit

PLATFORMS

- Unix/Linux • AWS • GCP

TOOLS AND DATABASE

- Docker • Kubernetes • git
- Redshift/Bigquery • Bigtable
- Influx • Prometheus • Chef
- Postgres • Redis • DynamoDB

LINKS

Github:// [arujit](#)
StackOverflow:// [arujit](#)
LinkedIn:// [arujit-pradhan](#)
Leetcode:// [arujit](#)
Twitter:// [@arujit](#)

INTERESTS

- Big-Data • Devops • Data-Science
- Data structures • Algorithms
- Board Games • Music/movies
- Trivia/Quizzing

VOLUNTARY

- Teaching Assistant, NITS
- Open Source Contributions
- Coding Club, NITS

SUMMARY

Junior Software Engineer building platforms for Data since day one. Passionate about Distributed Systems, Software Architecture, Open Source and Clean code.

EXPERIENCE

GOJEK | PRODUCT ENGINEER

Aug 2019 - Present | Bangalore, India

- **Dagger** Developed framework for end to end run-time custom java code injection in in-house Flink based Data Streaming Platform. Built Smoke and Integration testing pipelines and added multiple other features.
- **Longbow** Implemented Key-value based state backend on top of Bigtable for complex Flink Streaming queries requiring up to 3 months of production Data.
- **Thor** Built a CLI tool for all in-house Data products. Using thor exposed git-based Data platform to customers.
- **Firehose** Incorporated multiple features to in-house Kafka-sink platform like Redis and Influx sinks, Templatised JSON body support to HTTP sinks and sidecar based monitoring using influx reporter.

MAKEMYTRIP | SOFTWARE ENGINEER

July 2018 – Aug 2019 | Bangalore, India

- **GoMemory** Worked on a centralized real-time user profiling Database on DynamoDB and Kafka-streams to provide user details < 200ms. Built a continuous integration testing framework to ensure correctness and a lightweight web API serving user personalization data on top of it.
- **KShift** Added Avro and Protobuf data support in Kafka to Redshift/S3 ETL tool.
- **Data Infrastructure** Worked on S3 + Delta-based Data warehouse. Maintained and continuously enhanced all Data infrastructures like Kafka, Yarn, Airflow etc. Established a fault-tolerant CDC pipeline using Debezium.

MEDCORDS | SOFTWARE ENGINEERING INTERN

April 2017 – July 2017 | Pune, India

- Worked on OCR module for medical reports. Experimented with basic ML models on patient health data.
- Created a basic JWT based auth framework for user login flow.

COGNITIVESCALE | SOFTWARE ENGINEERING INTERN

April 2016 – June 2016 | Hyderabad, India

- Worked on a Geospatial Data visualization tool built on top of ELK stack.

PROJECTS

Open Source Contributions

- Minor feature contributions and bug fixes in various Open Source projects like *Apache Flink*, *Beast*, *Airflow*, *kaf* and *Azkarra Streams*.

Freelances

- **Kepi.ai | Vehicle Detector** Worked on end-to-end vehicle detection and tracking module for a Traffic Assistance IOS app.
- **IIT BBSR | Plant Disease Detection** Worked on Data analysis, forecasting of diseases in crops with ARIMA and ANN.
- **The Social Street | Retail Analyser** Worked on image classifications and heat map generation modules from CCTV feed for user behavior extraction.