Debangshu Bhattacharjee

Computer Science Student

+91 9748625028

debangshu133@gmail.com

Debangshu Bhattacharjee

github.com/DebangshuB

Kolkata, India

Technical Skills

Languages

C/C++ | SQL | Python | R | Markdown | LaTeX

Developer Tools

VS Code | Docker | Git | Jupyter Notebook | GitHub | PyCharm | Kubernetes (Learning)

Libraries

Matplotlib | NumPy | Pandas | Heroku | SciKit Learn | Tensorflow | Flask | Tweepy | BeautifulSoup | Plotly | Dask & Dask-ML | PySpark

Tools / Frameworks

Apache Kafka | Apache Airflow | AWS EC2 | Dagster | Apache Spark | MongoDB | MySQL | Data Build Tool

Coursework

Academic

AI | Data Structures | C/C++ | Cloud Computing | DMDW | Mathematics | DBMS | NLP | Machine Learning

Other

Deep learning | Machine learning | NLP | Data Engineering | Tableu | Amazon Web Services | Big Data

Objective

A motivated student pursuing Computer Science at KIIT University, 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.

Education

Kalinga Institute of Industrial Technology — Bhubaneswar, India
B.Tech. in CSE, CGPA : 9.6 — *June 2019 - Present*

Delhi Public School Ruby Park — Kolkata, India
Senior Secondary Education — *April 2017 - April 2019*

Experience

Home.LLC | *Data Science Intern* — *March 2022 - Present*

- Analysed large datasets using *SparkSQL* and *Pandas* with *Python*.
- Created *data pipelines* for monthly data collection from various *APIs*.
- Created *visualizations* to gain insights on the housing market using *Tableau & Python (Seaborn and Matplotlib)*.

Projects

HashTag Live Analysis | *Apache Kafka, Apache Spark, MongoDB, Tweepy*

- Computes sentiment of tweets over a time period using *Apache Spark*.
- Could be used to monitor popularity of marketing campaigns live.

MLOps with Dagster | *Dagster, SciKit Learn, Pandas, BeautifulSoup*

- End-to-End MLOps pipeline with *Dagster* (data orchestration platform).
- Includes *downloading data* to *preprocessing* to *hyperparameter tuning*.

A Simple DBT Project | *Data Build Tool, MySQL, Python*

- A straightforward project demonstrating the use of *data build tool*.
- Transforms tables in *MySQL* for downstream analysis.

ETL with Airflow | *Apache Airflow, MySQL, Python*

- Automates *ETL* task on log data.
- Uses Python to create *DAGs* which run on *Apache Airflow*.

HashTag Sentiment Analysis | *Python, Flask, Tensorflow, Tweepy*

- Full-stack machine learning web application using *Flask* and *Tensorflow*.
- Utilized *RNNs* to perform *Sentiment Analysis*.

Cardio UI | *Python, Flask, SciKit Learn, Heroku*

- Full-stack machine learning web application using *Flask* and *Scikit Learn*.
- Utilizes user health data and habits to predict chances of heart disease.

Extracurricular Activities

Flipkart Grid 3.0 Hackathon

Trust and Affluence Score Prediction

Developed an application that could predict a person's trust and affluence score from their Twitter data. Uses *Python* and *SciKit Learn* along with *Numpy* and *Pandas* for processing. Plans to expand to Instagram and Facebook data.

Assistant Coordinator

KIIT Film Society

Have to regularly manage people and coordinate various activities. Organized multiple events and workshops and have to work with people on a regular basis as the Sound Team head.