

Pranesh Mukhopadhyay

praneshmukherjee7@gmail.com | 96747-70912

 github.com/Mukhopadhyay  linkedin.com/in/praneshmukhopadhyay  kaggle.com/praneshmukhopadhyay

EXPERIENCE

BOXWISH MEDIA LLP. / SPIRITZONE | CONSULTANT - DATA ENGINEERING

Dec 2021 – Present | Navi Mumbai

ORDER ASSIGNMENT SYSTEM

- Real time order assignment system made using **Python**, **FastAPI** and hosted and container orchestration handled using **Docker compose**, supporting Webhook listener & dynamic trip creation.
- Intelligent algorithm used to dynamically perform order batching based on delivery partner availability and order volumes. Maintaining a fair and unbiased delivery agent ranking system.
- Added support for centralized logging using MongoDB and end to end unit testing using Pytest & Flake8

PRODUCT SEARCH API

- An incrementally trainable search API, which can index products even with their attributes developed using **Python** & **Scikit-Learn** using Bag of Characters model
- Hosted using **gunicorn** for maximizing concurrent requests
- 80%+ Unit test coverage using & Passed swarm test conducted using Locust

ORDER ETA CALCULATION API

- 52 out of 2500 applicants chosen to be a KPCB Fellow 2014.
- 52 out of 2500 applicants chosen to be a KPCB Fellow 2014.
- 52 out of 2500 applicants chosen to be a KPCB Fellow 2014.

QUOSPHERE INFOSOLUTIONS PVT. LTD. | JUNIOR DATA SCIENTIST

Jul 2020 – Dec 2021 | Navi Mumbai

Spiritzone Product Recommender System

• Python • Flask • Collaborative Filtering • Content based recommender system

- A hybrid recommender system trained using user's product browsing patterns and feedback path drawn from customer engagement platforms where the models are incrementally trained using the customer interaction data.
- Highly scalable and easily maintainable stateless microservices developed using Python and Flask. Capable of handling 30+ requests per second asynchronously.
- End to end unit & integration tested using python's built-in unit test library

Market Basket Analysis

Python • FastAPI • Apriori Algorithm • Docker compose • FP-Growth Algorithm

- Developed using the **Apriori** and **FP-Growth** algorithm for efficient, scalable and maintainable stateless microservice.
- Used FastAPI and Docker to make the deployment process easier alongside a complete build pipeline, capable of training the the models on a daily basis.
- Unit tests with 90%+ code coverage & successfully handled load testing with 30+ concurrent requests per second (conducted using Locust).

Clickstream Data Analytics Pipeline

Azure Data Explorer • Python • Grafana

- 52 out of 2500 applicants chosen to be a KPCB Fellow 2014.
- 52 out of 2500 applicants chosen to be a KPCB Fellow 2014.
- 52 out of 2500 applicants chosen to be a KPCB Fellow 2014.

PROJECTS

MONOLG | AN EASY TO USE CENTRALIZED MONGO DB LOGGING LIBRARY

A Simple easy to use centralized logging library for Python using MongoDB. Deployed at Python Packaging Index aka. PyPI at pypi.org/project/monolg. The library was written so that any Python developer who uses the default logging library can get started with Monolg as easily. With complete documentation and user guide hosted in Read the docs at monolg.readthedocs.io.

 github.com/Mukhopadhyay/monolg

YOUTUBE SUBTITLE DATASET | THE MOST VERSATILE YOUTUBE TRANSCRIPTION DATASET

Mar 2012 – May 2013 | Ithaca, NY

Dataset containing 2500+ unique popular YouTube video subtitle, with manually annotated video category, that can be used as the dependent variable. Alongside the video categories we have various different video related attributes such as duration, channel name, subscriber count etc.

[kaggle.com/datasets/praneshmukhopadhyay/youtubers-saying-things](https://www.kaggle.com/praneshmukhopadhyay/youtubers-saying-things)

EDUCATION

ST. XAVIER'S COLLEGE | M.Sc IN COMPUTER SCIENCE

Aug 2020 | Kolkata, WB • Cum. GPA: N/A

Curriculum

Artificial Intelligence • Data Structures • Algorithms • Data Warehousing • Compiler Designing • Database Management System

ASUTOSH COLLEGE | B.Sc IN COMPUTER SCIENCE

June 2018 | Kolkata, WB

Dean's List (All Semesters) • Cum. GPA: 3.92 / 4.0 • Major GPA: 3.94 / 4.0

Curriculum

Data Structures • Algorithms • Database Management System • Automata

SKILLS

PROGRAMMING

Python • C++ • JavaScript

LIBRARIES & FRAMEWORKS

TensorFlow • PyTorch • Pandas • Numpy • Scipy • FastAPI

DATABASES

MSSQL • PostgreSQL • MongoDB • MySQL

DEVOPS AND OTHER TOOLS

Docker • Kubernetes • Github Actions • Azure Pipelines • Grafana

CERTIFICATES

Machine Learning by Stanford University

🌟 Coursera • Certificate ID: 636WHFRBMH5V

Data Science Math Skills by Duke University

🌟 Coursera • Certificate ID: AFQSU5CAT4HD

Introduction to TensorFlow for Artificial Intelligence by DeepLearning.AI

🌟 Coursera • Certificate ID: E3XDATA99F3Y