APURVE YAJNIK

yapurve@gmail.com • Gurugram, India • +91 8437826504

https://github.com/apurveyajnik - https://linkedin.com/in/apurveyajnik

For more information : yapurve.com/about

Skills

Programming

- Python
  - pandas, numpy, scipy, matplotlib
  - scikit-learn, spacy, lightgpm, h2o, xgboost
  - django, flask
  - pytorch
  - huggingface transformers, openai, claude, google-genai, llamaindex, langchain

Other tools/architectures

- Git, copilot, cursor
- SQL DBs - MySQL, SQL Server
- NoSQL DBs - Mongodb, Redis, Neo4j
- R - dyplyr, ggplot2
- Docker, Kubernetes
- Kafka
- Spark
- AWS (EC2, S3, Sagemaker etc.)
- Llama, YOLO
- GCP Vertex AI

Work Experience

Principal Data Scientist

Drivebuddyai - A division of Roadzen

Dec 2022 - Present

- Developed AI modules detecting driver distraction, drowsiness, collisions warning, and hard braking events in logistics vehicles and cabs with 85-95% accuracy, processing 10k+ events daily.
- Event context creation and video proof generation using sensor fusion.
- Built scalable data analytics pipelines handling 50TB+ of telematics data monthly.
- Implemented geospatial data integration for real-time tracking and road environment profiling.
- Working as an individual contributor and data science team manager, while taking the role of scrum master for the whole engineering team.
- Showcased our various features of our product Driver Monitoring and ADAS systems at multiple trade shows.

Data Scientist

Self-Employed Freelancer

Sep 2020 - Nov 2022

- Built Transformer-based LLM model for extracting Climate Change disclosures from 10-K reports with 85% accuracy.
- Created feature stores on GCP for various aggregate features for Machine Learning tasks.
- Built a Discord bot to provide customer support information and manage toxic comments.
- Created the backend APIs for hotel group booking system.

Data Scientist

Housing.com

Nov 2019 - June 2020

- Created a house rent price suggestion model for brokers.
- Built and deployed an ensemble containing object detection models and Image classifiers for image filtering for India's biggest real estate marketplace Housing.com.
- Created a module to generate rental listing descriptions using GPT-2.

Applied Research Engineer
Netomi (formerly msg.ai)

July 2017 - Oct 2019

- Implemented and trained Deep Learning based models from scratch for NLP tasks with a main focus on Intent Classification, Text Similarity, and Question Answering.
- Built and shipped the end-to-end Natural Language Understanding pipeline for customer support chat and email bots from research to production for clients like Westjet, Tmobile, Singtel, etc.
- Created an automated multi-tenant fine-tuning pipeline for intent classification models for 10+ enterprise clients.

Education

Bachelor's Degree - Computer Engineering
Thapar Institute of Engineering and Technology, India

2012 - 2016