YOGHES WARAN

+91 7010370291 ◇ Tamil Nadu, India

ywa0291@gmail.com ◇ [LinkedIn](#)

EDUCATION

B.Tech in Control Engineering, National Institute of Technology Tiruchirappalli   Jul 2016 - Jun 2020
Minor in ECE
CGPA: 8.22 / 10

EXPERIENCE

Data Scientist   Sep 2022 - Nov 2024
SprigHub INC   *Remote, India*

- Experienced Data Scientist and Product Strategist with a proven track record of driving innovation, defining product road maps, and delivering impactful features and services. Played a key role in client acquisition and helped secure significant funding by aligning data-driven solutions with business growth objectives.

- **Data Analyst Agent**

  - Developed an intelligent agent using open-source LLMs and the vLLM inference engine to accelerate insight generation and dashboard creation from marketing data, featuring core modules for planning, code execution, error correction, and solution ranking based on user queries.

  - Reduced server costs while maintaining performance by optimizing a custom LLM with Q8 and NF4 quantization on AWS infrastructure.

  - Enhanced domain-specific performance in marketing analytics by fine-tuning various LLMs using LoRA for text-to-SQL, and leveraging tools like Unsloth AI and DeepSpeed for efficient training and inference.

- **Marketing Budget Optimization**

  - Developed a Bayesian causal inference-based Marketing Mix Model (MMM) to measure campaign impact and inform budget allocation, including a causal discovery pipeline and large-scale feature engineering to improve media effectiveness. Conducted what-if simulations to forecast outcomes.

  - Developed Conditional Average Treatment Effect models to identify high-ROI customer segments for coupon targeting. Applied advanced causal inference techniques including Double Machine Learning (DML), meta-learners, GRF, and IPW. Built a scalable end-to-end pipeline covering data preprocessing, model training, hyperparameter optimization, and model evaluation using metrics like PEHE and policy risk.

  - Developed and assessed multi-channel marketing attribution models including Markov Chain, LSTM, tree-based, and first/last-touch approaches to analyze and optimize channel performance. Engineered a modular, end-to-end pipeline encompassing data preprocessing, feature selection, class imbalance handling, and model evaluation using metrics like log loss, AUC, and ROI.

  - Collaborated with clients to optimize marketing budgets, resulting in an ROI increase of at least 18%.

- **Experimentation Platform**

  - Designed and co-developed an A/B testing engine to identify optimal campaign groupings across funnel stages. Integrated fixed-sample frequentist, sequential testing, and Bayesian approaches.

  - Applied a range of statistical techniques,including t-tests, Mann-Whitney U, Wilcoxon tests, CUPED, SRM detection, and the delta method to analyze metrics such as conversion lift, CTR lift, and week-over-week sales. These methods were widely adopted by clients to optimize marketing performance, contributing to a 34% quarter-over-quarter increase in marketing ROI for a key client.

- **Day to Day work**

- Created ETL processes using Python, PySpark, and SQL to transfer data from Google Cloud Platform to AWS, ensuring seamless data integration across platforms.

- Provided automated actionable insights, forecasted KPIs, and performed anomaly detection and root cause analysis of marketing KPIs using causal inference techniques with DoWhy and Scikit-Learn.

Engineer   Feb 2021 - Sep 2022

Technip Energies   *Mumbai, India*

- **Control Engineer**

    - Coordinated with Datasheer Digital compmay (sponsored by Technip Energies) to develop a computer vision model to identify and categorize symbols in PIDs and associate them to lines and text labels and to automate the whole process, which led to savings of 100 man hours per week.

    - Developed a preliminary design of the control system architecture and control logic for a polybutadiene rubber plant,IOCL Panipat.

    - Designed process control loops like cascade, spilt range, ratio control, etc, to regulate the process parameters.

SKILLS

Languages	Python, C++, Cuda C++, SQL (Postgres), R.
Frameworks	Spark, PyTorch.
Developer Tools	Git, Docker, AWS Services, Google Cloud Platform.
Libraries	Deepspeed, Unsloth AI, Transformers, Scipy, DoWhy, Causal-learn, Scikit-Learn.

BLOGS

Cuda C++ Blog series, a comprehensive set of posts covering CUDA programming, from basic to advanced topics like Flash-Attention.