Yogheswaran A

(+91) 7010370291 — ywa0291@gmail.com — linkedin.com/in/yogheswaran — Personal Website

ABOUT

Data scientist with 3^+ years of experience designing and deploying advanced AI solutions, including AI agents, Generative AI, and causal inference models. Proven ability to leverage machine learning, statistical analysis, A/B testing, and full stack big data pipelines to drive product innovation. A strategic and collaborative problem solver who transforms complex challenges into scalable solutions that boost business performance and profit.

SKILLS, CERTIFICATES AND BLOGS

Technologies: Python, C++, Cuda C, SQL, R, DoWhy, Causal-learn, SK-learn, Pandas, SciPy, NumPy, PyTorch, Transformers, DeepSeed, Airflow, Hive, PySpark, AWS Services, GCP, Docker, Git, Asana, Jira.
Certificates: Hugging Face Agents, ML, NN & Deep Learning, Convolutional NN, Improving Deep NN.
Blogs: Cuda C++: Basics to flash attention, Math behind diffusion models, Stable diffusion - 3 theory.

WORK EXPERIENCE

SPRIGHUB INC Gen AI, Agents, Machine Learning, Causal Inference, Experimentation, POC Sep 2022 - Nov 2024
DATA SCIENTIST Statistcal Methods, Data-driven recommendations, Analytics and Insights, ETL Remote, India

- Collaborated with the CEO, CTO, and global teams to define roadmaps and implement solutions for marketing and customer analytics.
- Designed and deployed advanced statistical and machine learning methods that provided actionable insights for 5 clients with over $50M^+$ users, generating more than $\$5M^+$ in revenue.
- Led the complete development of an AI agent to automate marketing analytics and dashboard generation. Integrated core modules, optimized models with Q8/NF4 quantization, fine-tuned with LoRA and DeepSpeed, and deployed using AWS Lambda and AWS Glue, saving 100^+ man-hours per week and enabling real-time insights.
- Developed a bayesian MMM model to optimize digital marketing budgets. Automated weekly model tuning and execution of data pipelines using Airflow, AWS Glue, AWS Lambda, PySpark, and SQL, leading to a $14\%^+$ ROI per client.
- Researched and deployed scalable causal inference models for user segmentation, churn, and email targeting, helping a pilot customer save \$10K.
- Used DML, Meta-Learners, GRF, and IPW for promotions targeting and evaluated models with PEHE and policy risk. Deployed on AWS SageMaker, potentially saving $\$50k^+$ for a client.
- Co-developed an experimentation platform to optimize campaign groupings by integrating frequentist and bayesian statistical methods.
- Applied techniques such as t-tests, Mann-Whitney U, CUPED, and the delta method to measure lift, driving a 34% QoQ increase in marketing revenue for a key client.
- Led a team of five to create a standardized template for migrating $300M^+$ records from GCP to AWS using SQL & PySpark. Streamlined a 10-step manual process into a single automated flow, enabling faster insights.
- Built anomaly detection, root cause analysis, and KPI forecasting models, extensively used by a client with 15,000 locations and $\$69M^+$ in assets, resulting in a \$10K per week increase in revenue.

TECHNIP ENERGIES Feb 2021 - Sep 2022
ENGINEER Machine Learning, Deep Learning, Computer Vision, Advanced Control Theory Mumbai, India

- Coordinated with Datasheer Digital to develop a computer vision model for identifying, categorizing, and linking symbols in P&IDs, automating the process and saving over 500 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.

EDUCATION

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