Dinesh Arasavalli

Detroit, Michigan, United States | dinesharasavalli19@gmail.com | +1 313 761 9952 | LinkedIn | GitHub

EXPERIENCE

Data Science Intern — *Pacific Northwest National Laboratory (PNNL)* 05/2024 – 07/2024

- Engineered a time-lagged feedforward neural network (TLFN) using Python and TensorFlow to predict Locational Marginal Prices (LMP), achieving 15% higher accuracy than benchmarks and enhancing grid stability.
- Formulated and solved a linear programming model in PyTorch and NumPy to optimize energy storage dispatch, reducing operational costs by 10%.
- Analyzed 500GB of energy market data (PJM, ERCOT) using Pandas, SQL, and Matplotlib to identify key patterns and inform model development.
- Built algorithms for real-time market participation, improving response times by 20% through efficient data processing and optimization techniques.
- Collaborated with cross-functional energy analysts to deliver 6 data visualization reports and presentations, reducing stakeholder review cycles from 5 days to 2 days and accelerating project decision-making.

Data Analyst — *TherDose Pharma Private Limited, Hyderabad, India (Hybrid)* 03/2021 – 03/2023

- Designed interactive Power BI and Tableau dashboards, reducing reporting time by 40% and enhancing data accessibility for R&D, operations, and marketing teams.
- Conducted data cleaning, preprocessing, and statistical analysis in Python and SQL, improving data reliability by 25%.
- Applied predictive modeling techniques to oncology datasets, influencing a $900k investment decision through trend analysis and regression insights.
- Collaborated with cross-functional teams to translate complex analytical findings into actionable insights, improving alignment and decision-making efficiency by 30%.
- Enhanced data management and governance processes, reducing compliance-related errors by 20% and ensuring regulatory adherence across data systems.
- Delivered 3 automated KPI dashboards and delivered 9 executive presentations, minimising reporting by 50% and driving data adoption across 3 business units; promoted to lead analytics initiatives within 6 months.

Projects

Trained a EfficientNet-B0 on ImageNet-LT with LDAM Loss and Quantization

- Trained EfficientNet-B0 on long-tailed ImageNet-LT (115K images, 1000 classes) using LDAM loss and AdamW optimizer, achieving 68% top-1 accuracy on minority classes with single GPU.
- Applied Langevin dynamics and class-balanced weighting to address 100:1 class imbalance, improving rare class recall by 15% over baseline through hyperparameter tuning.
- Deployed post-training quantisation (FP32→INT8) and HSwish activation, reducing model size from 20MB to 5MB and improving inference speed by 35% while maintaining 67% accuracy.

Cardiovascular Disease Trends and Prediction Dashboard using Plotly Dash.

- Conducted EDA on 10 years of cardiovascular data from US Chronic Disease Dataset, identifying critical correlations; developed 96% accurate predictive model to forecast disease cases and optimize healthcare resource planning.
- Built a web application featuring 3 dashboards, 4 interactive tabs, and 12 dynamic Plotly graphs using Plotly Dash, providing real-time data exploration and visualization for cardiovascular health insights.

Scalable Chest X-Ray Analysis with Pretrained CNNs.

- Trained and Customised ResNet-50 and VGG-16 on 20K+ chest X-ray images using multi-GPU data pipelines and distributed dataloaders.
- Applied CLAHE preprocessing and batch-augmented training to detect COVID anomalies with 95% accuracy, reducing false positives by 35%.

Technical Skills

Programming Languages: Python, R, SQL, YAML, CUDA
Frameworks & Libraries: PyTorch, TensorFlow, Keras, Hugging Face, Transformers, OpenCV, scikit-learn, NLTK, spaCy
Generative & Foundation Models: GPT, BERT, GANs, VAEs, RLHF, LoRA, PEFT
Computer Vision & Robotics: Image Classification, Object Detection (YOLO, Swin Transformer), Semantic Segmentation.
Model Optimisation : Quantization, Knowledge Distillation, Neural Architecture Search (NAS), Pruning, Mixed Precision.
Distributed Systems & HPC: CUDA, PySpark, Apache Spark, Hadoop
Data Analysis & Visualisation: NumPy, Pandas, Matplotlib, Seaborn, Plotly, Power BI
Math & Optimisation: Linear Algebra, Multivariate Calculus, Optimisation Algorithms, PCA

EDUCATION

Wayne State University Detroit, MI, USA
Master of Science in Data Science and Business Analytics *10/2023 – 12/2024*

Lovely Professional University Punjab, India
Bachelor of Technology in Civil Engineering and Computer Science(hybrid program) *08/2018 – 07/2022*

Certifications

Python for Data Science and AI (Coursera) — **Data Analysis in Python** (IBM)