Anish Dhane

Sheffield, United Kingdom | anishdhane@gmail.com | +447747144103 | linkedin.com/in/anish-dhane | github.com/AnuDi3745

Experience

Data Science Intern, School of Inspirational Leadership – Pune, India Sep 2022 - Mar 2023

- Designed and deployed Python scripts integrating multiple data sources, improving data processing efficiency and project reliability by **25–40%**.
- Analyzed real-world company datasets to optimize sales forecasting models, achieving forecast accuracy of **85–90%**.
- Implemented a hands-on approach to developing a speech-to-text feature for a virtual assistant, boosting overall project delivery speed by **20–30%** during the internship.
- Developed and maintained clean, structured datasets for the question-answering model, improving data usability and assisting model training.

Education

University of Surrey, MSc Artificial Intelligence Sept 2024 – Sept 2025

- **Aggregated Mark Achieved**: 67.33 (Merit)
- **Coursework:** Computer Vision and Pattern Recognition, Fundamentals of Machine Learning, AR, VR, and Metaverse, AI and Sustainability, NLP, and Applied ML.
- **Dissertation:** *Enhancing ARC-Solver with a Modular Reasoning Core* – built a **Reasoning Core** that solved **32 ARC tasks** while providing **step-by-step reasoning traces** to explain model decisions. Contributed insights for **AGI research** by making model approaches transparent and interpretable.

Projects

Enhancing ARC-Solver with Modular Reasoning (MSc Dissertation)

- Built a **Reasoning Core** that solved 32 ARC benchmark tasks with **step-by-step reasoning traces**.
- Enabled **interpretability and transparency** in decision making, moving towards **explainable AGI systems**.
- Provided research insights to develop future reasoning-capable AI systems.

Prompt Tuning for Vision-Language Models (CLIP) github.com/oz-e/applied-ml

- Benchmarked tuning techniques (CuPL, CoOp, CoCoOp, and Tip-Adapter) on CLIP for image classification.
- Improved **few-shot and zero-shot performance** across datasets (Caltech101, Flowers102, UCF101).
- Developed Tip-Adapter for training-free adaptation using cache-based prediction; **contributed to CuPL improvement** via filtering strategies.

Abbreviation and Long-Form Detection using Token Classification comm061nlp.streamlit.app.

- Built a token classification system to identify abbreviations and their long forms using BERT and RoBERTa models on domain-specific datasets.
- Achieved **F1 = 0.90** using RoBERTa with augmented data.
- Built and deployed a **Streamlit web app** with automatic logging through Google Sheets API.

Skills

Programming: Python (pandas, NumPy, scikit-learn, PyTorch, TensorFlow)

Machine Learning: Feature engineering, data preprocessing, model training, evaluation

Frameworks and Tools : PyTorch, TensorFlow, Scikit-learn, Streamlit, Git, Jupyter, Google Colab

Deep Learning: Vision-Language models, Transformers (BERT, RoBERTa), LSTMs