Anand Raj

Email: anand.raj@gwu.edu | +1 (703) 457-0873 | LinkedIn | GitHub
Home address: 850 N, Randolph Street, Arlington, VA – 22203, USA
Domain Skills: Software Engineering, Machine Learning, Computer Vision, Natural Language Processing,

EDUCATION

Master of Science, **Data Science -** The George Washington University ***May 2025 (Expected)***
Relevant Courses: Data Mining, Machine Learning, Natural Language Processing, Cloud Computing **GPA: 3.95**

Bachelor of Engineering, **Electrical and Electronics -** RNS Institute of Technology ***Sep 2021***
Relevant Coursework: Data Structures, Mathematics, Object Oriented Programming Using C++, Python Programming

TECHNICAL SKILLS

Programming languages and Databases: R, Python, C++, SQL, MongoDB, Neo4j

Libraries and Tools: NumPy, Pandas, Matplotlib, Sklearn, Folium, Plotly, PyTorch, Keras, TensorFlow, NLTK, spaCy, genism, Hugging Face, LangChain, Databricks, Tableau, Flask, AWS, GCP

Product Development: Agile Methodology, Product Life Cycle, Jira, Confluence, Git, GitHub, DevOps, Docker

WORK EXPERIENCE

Data Science Intern | AARP | Washington DC, United States ***June – Dec 2024***
Skills Used: *SQL / Python | PySpark | Databricks | AWS S3 | Hugging Face |Llama 70B | Fine-Tuning | Prompt Engineering*

- Designed and developed scalable AI/ML solutions for dynamic article labeling, automating customer query resolution and generating personalized marketing messages in **Databricks** using **PySpark** by leveraging **AWS S3** for data storage.
- Developed a dynamic labelling system using **Llama 70B**, iteratively generating new labels for articles based on previous batches, resulting in a **90% reduction** in manual labelling efforts.
- Automated customer query resolution by **fine-tuning a pre-trained LLM** (Flan-T5) on 34,000 customer queries and resolutions, applying Parameter Efficient Fine Tuning (**LoRA**) to optimize model performance and deployed the model into **production** on databricks leading to a **60% reduction** in the customer care team's workload by improving response accuracy and efficiency.

Software Engineer | Continental Automotive Group | Bangalore, India ***Sep 2021 - Aug 2023***
Skills Used: *C++ | Python | Product Development | Computer Vision | DevOps |ADAS | Agile | Tensorflow |Git*

- Worked on Advanced Driving Assistance Systems (ADAS) and developed autonomous vehicle systems like Emergency Brake Assist, and Rear Pre-Crash Predict in Agile Methodology.
- Designed and tested algorithms in **C/C++** at L3 Level and adhered to **CI/CD pipelines** for continuous integration and testing. Automated simulation scenario generation in Carmaker IPG by scripting in **Python**, significantly reducing manual efforts and streamlining testing processes by **60%**.
- Enhanced **pedestrian detection** performance to **87%** for Mercedes Benz VS30 platform vans by fine-tuning Region of Interest (ROI) parameters, significantly improving feature representation.
- Analyzed recorded vehicle data, extracting and preprocessing frames for model training. Developed machine learning model using **HOG** (Histogram of Gradients) and **SVC** in Python using **Tensorflow**, transitioning successful models into real-time inference through implementation in C++ on vehicle ECUs.

Data Science Intern | Innodatatics | Bangalore, India ***Jun – Aug 2020***
Skills Used: *Python | SQL | Tableau | Data Analysis | Machine Learning | Customer Segmentation*

- Collaborated with a dynamic team to conduct in-depth data analysis and utilized **data mining techniques** in **SQL, Python** and **Tableau**, providing valuable insights into client's sales data.
- Conducted behavioral segmentation of 500,000+ users, identifying key patterns in user engagement, temporal trends, and conversion rates between free and paid users, leading to an **8% increase in customer retention**.
- Formulated data-driven recommendations and compelling narratives and communicated to our client, resulting in a **10%** uplift in paid user conversions**.**

Research Intern | Defense Research Development Organization | Bangalore, India ***Jan – April 2020***
Skills Used: *Python | Tensorflow | Deep Learning | BERT | AWS | Flask | Git | GitHub | Docker*

- Developed and deployed a high-performing multi-label classification model using **BERT**, **Flask**, and **AWS EC2** to automatically categorize NLP research papers, improving categorization performance by **20%**, and streamlined the research process and deployed the system on **AWS** for inference.

PROJECTS

Medical Report Generation using Vision Language Models ***December 2024***
Skills Used: *Python | PyTorch | VLM | Vision Transformers | LLM | CUDA | GPU Optimization | ML Pipeline | Fine-Tuning*

- Designed and implemented a scalable solution for generating medical reports from chest X-ray images, using **PyTorch**.
- Employed BioViLT (**Vision Transformer**) for image features extraction and developed an alignment model to align image features with text data. **Fine-tuned** a Medical **Large Language Model** (BioGPT) to produce accurate, context-aware medical reports, optimizing workflows for diagnostic efficiency and reliability. **Optimized** the model and leveraged maximum **GPU utilization** to ensure peak performance and efficiency in processing.