

# Anand Raj

Email: [anand.raj@gwu.edu](mailto: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, 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

---

## 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** and frameworks like **Hugging Face** for efficient text processing.
- 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 T5 model on 34,000 customer queries and resolutions, applying Parameter Efficient Fine Tuning (PEFT) to optimize model performance.
- Optimized a query resolution model, leading to a **60% reduction** in the customer care team's workload by improving response accuracy and efficiency.

**Software Engineer | Continental AG | Bangalore, India**

**Sep 2021 - Aug 2023**

**Skills Used:** C++ / Python / Product Development / Testing / Code Quality Test / ADAS / Requirements Engineering / Agile / 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.
- Engineered and validated **safety-critical functions** for multiple OEMs, driving robust **requirements engineering** and compliance with industry standards (Euro NCAP), integrating elements of planning, and controls.
- Designed and tested algorithms in **C/C++** at L3 Level using **GTest** for Advanced Range Sensor-5th Gen and Short Range Radar, ensuring reliability with **QAC** compliance and version control via **Git/GitHub**.
- Automated simulation scenario generation in **Carmaker IPG** by scripting in **Python**, significantly reducing manual efforts and streamlining testing processes by **85%**.
- Provided problem-solving solutions to customer-reported bugs in the simulation environment.

**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

- Developed and deployed a high-performing multi-label classification model using **BERT**, **Flask**, and **AWS EC2** to automatically categorize NLP research papers, improving categorization accuracy (micro F1 score) by **20%**, and streamlined the research process, enabling senior scientists to identify key research papers **30% faster**.

---

## PROJECTS

**Medical Report Generation**

**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.