Anunay Kumar

anunay.kumar2002@gmail.com +1 236 865 2403 anunaykumar03 anunaykumar03

Education

University of Waterloo, *BASC in Computer Engineering, Honours*

- **GPA: 4.0/4.0 - Dean's Honors List**
- Algorithms and Data Structures, Real-Time Operating Systems, Embedded Microprocessor Systems, Systems Programming and Concurrency
- Advanced Discrete Mathematics, Probability and Statistics, Digital Hardware Systems

Skills

LANGUAGES: Python, C++, C, Javascript, SQL, Java, Flutter/Dart, Verilog
FRAMEWORKS: Tensorflow, PyTorch, Sci-kit Learn, Matplotlib, NodeJS, Express, ReactJS, Hugging Face, Pandas
TOOLS: Git, Linux, Bash Scripting, Firebase, Docker, AWS, Jenkins, JIRA, MongoDB, MySQL, Jupyter

Experience

Incoming RTOS Software Developer - Wind River VXWorks

Data Engineering - Purolator (Python, SQL, AWS, Spark) Sept. 2022 - Dec. 2022

- In charge of the Analytics team migration to **AWS and Power BI.**
- Built the data pipeline to track discrepancies between generated orders and picked up orders.
- Designed a cost-effective and fast event-based architecture using **S3, AWS Lambda, Event-Bridge, Glue and Redshift**.
- Achieved 80% data processing speed increase by writing efficient SQL and using **Spark+AWS's MPP** architectures.

Project X - Waterloo Data Science Team Researcher Sept. 2022 - Current

- Created a Pytorch ensemble model to perform language detection using video and masked audio files.
- Experimented with multiple model architectures that included Spacio-Temporal CNN, 3D CNN and 2D CNN + LSTM.
- Designed the preprocessing pipeline that included normalizing video frames and lip-extraction.
- Currently writing a research paper that documents literature review and project's advancements.

Undergraduate Research - Social and Intelligent Robots Laboratory (C++, Python) May 2022 - Sept. 2022

- Implemented **C++ Point Cloud Library** in conjunction with **OpenCV** for semantic segmentation of 3D Point Cloud Data to generate dimensions for the Fetch Robot.
- Created ROS nodes for an initial prototype that will input point cloud data, process it and perform segmentation in real-time.

Machine Learning Engineer - Blackberry (Python, Groovy, Docker, SQL) Jan. 2022 - Apr. 2022

- Used Docker and Bach Scripts to deploy a VM resource monitoring system in Linux with Grafana and Prometheus.
- Designed and optimized a static analysis tool checker to improve false-positive code classifications to improve developer productivity by **85%** using an efficient Random Forest and Category Boosting blended classifier through Tensorflow.
- Used **AutoML** infrastructure and **complex SQL joins** and **dynamic pivots** for data engineering.
- Identified critical data logic flaws that were mislabeling the training data and **designed a safe and fault-free data processing pipeline** using **SQL and Pandas.**
- Optimised Jenkins Pipeline running an NLP model for open-source license classification.

AI/ML/NLP Data Scientist - Genellipse Inc. (TensorFlow, ReactJS, SQL) May 2021 - Aug. 2021

- Designed and optimized a **neural network using TensorFlow and AutoML** that predicts potential client customers, increasing sales efficiency and resource allocation.
- Improved prediction using **hyperparameter tuning**, custom metrics, **data stratification** and early stopping, resulting in a **94.8%** accurate model.
- Created complex **DAX and SQL Queries and generated PowerBI reports** highlighting important data analytics for the client board.

Projects

Real Time OS (C, ARM) Jan. 2023 - Apr. 2023

- Learnt RTOS by implementing a lightweight RTX on Intel DE01-SOC Board.
- Designed a free list first fit memory allocation scheme.
- Implemented a preemptive, multi-tasked kernel for task management with memory ownership functionality.
- Developed Inter-Task Communication and UART I/O through mailboxes and a logarithmic time scheduler with a Red-Black Tree.

Systolic Array Implementation of a Matrix Multiplier (Verilog) Jan. 2023 - Apr. 2023

- Designed a Systolic Matrix Multiplier using Verilog, System Verilog and debugged using Vivado.
- Generalized it to multiply higher dimension matrices than the array size by slicing and cascaded counters.
- Used packed multiplication and DSP Space Analysis to optimize and increase operating frequency on the Xilinx FPGA board.

Multi-Threaded Web Scraper (C) Aug. 2022 - Dec. 2022

- Built a multi-threaded web scraper to learn multi-threading in C.
- Used concurrency paradigms like Readers-Writers Problem and Bounded Buffers.

FPGA Audio Player (C) - Embedded Microprocessors Aug. 2022 - Sept. 2022

- Used Altera NIOS tools to debug and build a media player to play *.wav* files from onboard memory.
- Implemented button debouncing using state machines for push button functionality.
- Implemented Stereo and Mono 16 Bit audio playback capabilities at normal, half and double speed.