

# Clay Ndugga

SOFTWARE ENGINEER

☎ (+1) 403-542-4358 | ✉ clayndugga@gmail.com | 🏠 clayndugga.github.io | 📱 clayndugga | 🌐 clay-ndugga

## Education

### Applied Mathematics and Engineering

Kingston, ON

QUEEN'S UNIVERSITY

2024

- Similar to **Software Engineering** with a focus on advanced mathematics
- **Relevant Coursework:** Statistical Modelling, Data structure's and Algorithms, Stochastic Processes, Operating Systems, Embedded Systems, Control Theory, Information Theory, Optimization and Control of Stochastic Systems, Data Compression, Computer Networks

### AWS Certified Solutions Architect (SAA-CO3)

2024 Certification

## Work Experience

### Software Engineer Intern

Calgary, AB

LONGVIEW SYSTEMS

May 2023–Sep 2023, May 2021–Sept 2021

- Built an **end-to-end IoT Azure cloud solution** encompassing data ingestion, storage, and real-time decision-making through an **ML model**
- Used a Large-Language-Model (**LLM**) and DevOps API to automate **Azure** wiki documentation saving 10+ weekly hours of manual writing
- Designed and implemented a robust data ingestion system using **REST APIs**, achieving a 70% improvement in ingestion times.
- Modernized an existing codebase with best-practice **system design** patterns and migrated to **Azure/Databricks cloud architecture**
- **Technologies:** Azure, Databricks, Python, PySpark, MLOps, MLFlow, SQL

### Machine Learning Engineer Intern

Calgary, AB

BOARDWALK REAL ESTATE

May 2022 – Sept 2022

- Built an **ML model** using **Tensorflow** to predict tenant lease renewals with 85% accuracy, reducing profit loss to empty properties
- Collected and processed over 1,000,000 data points from various sources including **databases**, surveys, and property management reports.
- Identified significant factors that effect the probability of a tenant renewing their lease
- **Technologies:** Python, SQL, Jupyter, Pandas, NumPy, Scikit-Learn, Tensorflow

## Thesis

### Deep Learning for Point Cloud Compression

[🔗 Link](#)

3D COMPUTER VISION MACHINE LEARNING

- Integrated cutting-edge machine learning developments to establish a Non-linear Transform Coder for pointcloud data compression.
- Optimized TensorFlow code to efficiently process large-scale point cloud datasets, achieving significant improvements in resource utilization.
- **Technologies:** Python, Tensorflow, LiDAR

## Projects

### Chat with PDF

[🔗 Link](#)

WEB APPLICATION

- Deployed a web app for interactive PDF chat with clickable references.
- Developed a scalable REST API with Node.js/Express, containerized with Docker, and deployed on Google Cloud Run with autoscaling.
- Implemented embedding generation with OpenAI and boosted query accuracy through dynamic re-ranking with Pinecone.
- **Technologies:** Typescript, **Docker**, **NodeJS**, express, React, NextJS, Google Cloud, Firebase, Pinecone Vector DB, OpenAI

### Youtube Comment Finder

[🔗 Link](#)

SERVERLESS WEB APPLICATION

- Used a CDN to deploy a web application that lets users search for youtube comments and insights
- Used **DynamoDB** to managed the execution of jobs across multiple **lambda functions**
- Built a scalable, serverless backend architecture using AWS to handle 1,000+ concurrent requests
- **Technologies:** AWS, AWS Lambda, DynamoDB, Next.js, **React**, **TypeScript**, Python

## Skills

**ML Topics:** Tensorflow | Pytorch | Scikit-learn | Pandas | XGBoost | Computer Vision | NLP | RAG

**Dev-ops & Infrastructure:** Azure | Databricks | AWS | Docker | Git | Linux

**Other:** Python | SQL | MATLAB | Javascript | C | HTML/CSS | React | NodsJS | ExpressJS | MongoDB