

# Aayush Bajaj

+1-(438)-778-6965 | [aayushbajaj58@gmail.com](mailto:aayushbajaj58@gmail.com) | [LinkedIn://Aayush](https://www.linkedin.com/in/Aayush) | [Github://Aayush](https://github.com/Aayush)

## EDUCATION

### University of Montreal

Montreal, Canada

*Master of Science - Computer Science (MILA Lab)*

*Sept 2023 - April 2025 (Expected)*

**Coursework:** Foundations of Machine Learning, Data Science, Deep Learning(CNN, RNN, Transformers), Mathematics for Computer Science, Probabilistic Graphical Models

**Grade:** 3.8/4

### Maharaja Agrasen Institute of Technology

New Delhi, India

*Bachelor of Technology - Information Technology*

*Aug 2016 – Sept 2020*

**Coursework:** Operating Systems, Data Structures, Analysis Of Algorithms, Artificial Intelligence, Computer Networks, Database Management System, Theory of Computation, Computer Graphics

**Grade:** 7.9/10

## EXPERIENCE

### Momentum Health

May 2024 – Present

*Machine Learning Intern*

*Montreal, Canada*

- Created a training dataset for predicting scoliosis progression from 3D meshes using ray tracing to calculate change. Applied a modified version of PointNet++ with SFT and FFT. Achieved 83% accuracy on OOD test with FFT, surpassing SOTA.
- Enhanced the accuracy of predictions with an average error of less than 10 by 4% using a cross-attention mechanism that integrates both image and tabular data for predicting Cobb angle.
- Built and maintained crucial ML infrastructure involving data, training, and evaluation pipelines.

### NatWest Group

Nov. 2021 – Aug 2023

*Data Scientist*

*Bangalore, India*

- Developed Unsupervised Anomaly Detection Models for millions of cash deposits using Spark, SQL and MLlib leading to mitigation of millions of pound worth of potential money laundering risk.
- Worked on Graph Algorithms in Quantexa to run on Ownership structure of the companies. Managed Sagemaker pipelines for experimentation and model training.
- Conducted an analysis to unearth customers who deviate from their declared risk profile using statistics and knowledge graphs. 32% of the customers were found suspicious leading to millions of pounds worth of risk mitigation.

### Avantari Technologies

Sept. 2020 – Oct. 2021

*Machine Learning Engineer*

*Hyderabad, India*

- Built and trained Tacotron2+WaveRNN Autoregressive/Non-Autoregressive TTS models using PyTorch for custom voice synthesis with high fidelity.
- Deployed and maintained MLOps using Docker, Terraform (IaC) and Serverless Framework. Reduced the TTS inference time by 88%.

## SKILLS

**Languages:** Python, Scala, C/C++, SQL, R

**Frameworks:** PyTorch, TensorFlow, Transformers, Flask, Spark

**Developer Tools:** Git, Docker, Linux, Snowflake, Amazon Web Services, Google Cloud Platform, Hue, Teradata

**Libraries:** Pandas, NumPy, Matplotlib, Statsmodels, PySpark, NLTK, Spacy

## PUBLICATIONS

- [1] Ahmed Masry, Megh Thakkar, **Aayush Bajaj**, Aaryaman Kartha, Enamul Hoque, and Shafiq Joty. *ChartGemma: Visual Instruction-tuning for Chart Reasoning in the Wild*. 2024. arXiv: 2407.04172 [cs.AI]. URL: <https://arxiv.org/abs/2407.04172>.
- [2] **Aayush Bajaj**, Abhishek Jha, Lakshay Vashisth, and K.C. Tripathi. "Comparative Wavelet and MFCC Speech Emotion Recognition Experiments on the RAVDESS Dataset". In: *Mathematical Statistician and Engineering Applications* 71.3 (2023), pp. 1288–1293. ISSN: 2094-0343. URL: <https://www.philstat.org.ph/index.php/MSEA/article/view/468>.