

# Aditeya Baral

+1(551)263-8608 | [aditeyabaral@nyu.edu](mailto:aditeyabaral@nyu.edu) | [linkedin.com/in/aditeyabaral](https://www.linkedin.com/in/aditeyabaral) | [aditeyabaral.com](https://aditeyabaral.com) | [Google Scholar](https://scholar.google.com/citations?user=aditeyabaral)

## EDUCATION

### New York University, Courant Institute of Mathematical Sciences

Masters in Computer Science; GPA - 4.00/4.00

New York, NY

Sep 2024 – May 2026

**Concentration:** Artificial Intelligence

**Courses:** NLP with Representation Learning

### PES University

Bachelor of Technology in Computer Science & Engineering; GPA - 8.71/10.00

Bengaluru, India

Aug 2018 – May 2022

**Specialization:** Machine Intelligence & Data Science

**Courses:** Discrete Math, Linear Algebra, Intro to Data Science, Data Analytics, Big Data, ML, NLP

## RESEARCH EXPERIENCE

### Cisco Systems

Applied Researcher, Webex AI

Bengaluru, India

July 2022 – July 2024

- **Instruction fine-tuned** LLMs like Mistral and Llama-2 on-prem to enable **secure** and **cost-effective** AI solutions such as **translation** and **RAG** for engineers and customers.
- Led the initiative to build a novel **pre-training algorithm** for conversational data using **PyTorch** and **HuggingFace**.
- Developed the **Webex Contextual Search** engine and **improved searching, ranking, recommendations** and **topic modelling** with minimal overhead latency.
- Integrated **OpenAI APIs** and on-prem LLMs with the **Webex AI Assistant** to add **auto-replies, summarisation, querying** and **action-item extraction** to message threads and meeting transcripts.

### Intel Corporation

Applied Research Scientist Intern; Advisors: Anbumani Subramanian, Anay Majee

Bengaluru, India

Aug 2021 – Dec 2021

- Explored **Few-Shot Learning Object Detection (FSOD)** techniques to reduce **catastrophic forgetting** in constrained and heterogenous driving environments.
- Investigated and designed novel **representation learning** and **attention mechanisms** to learn **inter/intra-object relationships** using **PyTorch**.
- Outperformed existing approaches at the time on base and novel classes on the **Few-Shot India Driving Dataset**, a benchmark for FSOD.

### Center for Cloud Computing and Big Data, PES University

Research Assistant; Advisor: Dr. KV Subramaniam

Bengaluru, India

May 2020 – July 2020

- Compiled and used **TailBench** to **simulate and profile** application loads, monitor performance, and analyse results.
- Explored ways to **reduce tail latencies** in latency-critical applications such as translation and image recognition.

## INDUSTRY EXPERIENCE

### Cisco Systems

Big Data Analytics Engineer, Webex Media Quality Analytics

Bengaluru, India

July 2022 – July 2024

- Developed and deployed streaming jobs in **Scala** and **Flink** to compute **real-time metrics** from Calls and Meetings.
- Applied **statistical modelling** techniques to investigate and report **media quality insights** to downstream consumers.
- Led the development of **real-time auditing pipelines** using **Kafka** and **Python** to ensure **data consistency** between streaming jobs and downstream data stores like **Iceberg** and **Pinot**.
- Built graphs and dashboards on the **Webex Media Quality Analytics Dashboard** using **Grafana** and **Kibana** to set up alerts and KPIs for clients and customers.

Big Data Analytics Engineering Intern, Webex VideoMesh Analytics and APIs

Jan 2022 – June 2022

- Migrated the **Meetings Analytics Engine** from Java and Spark to **Scala** and **Flink** to significantly **improve real-time report generation** by over **30%**.
- Built **VideoMesh Developer APIs** using **Java** and globally rolled them out with **customer-facing applications**.

## PAPERS AND PUBLICATIONS

[1] **ChatBERT - Multi-task approach to Pre-Training for Structured Conversations**

Webex AI 2023

Authors: Aditeya Baral, Work done as part of Cisco Webex AI Research

[2] **CalBERT - Code-mixed Adaptive Language Representations using BERT**

AAAI-MAKE 2022

Authors: Aditeya Baral, Aronya Baksy, Ansh Sarkar, Deeksha D, Ashwini M Joshi

### [3] Information Maximization to Overcome Catastrophic Forgetting in Few-Shot Object Detection

Authors: Aditeya Baral, Anay Majee, Anbumani Subramanian: Work done as part of Intel Research

### [4] MAPLE - MAsking words to generate blackout Poetry using Seq2Seq LEarning ACL-ICNLSP 2021

Authors: Aditeya Baral, Himanshu Jain, Deeksha D, Mamatha H R

### [5] Analysis of Kepler Objects of Interest using ML for Exoplanet Identification IEEE CONIT 2021

Authors: Ameya Rajendra Bhamare, Aditeya Baral, Saarthak Agarwal

## RESEARCH PROJECTS

---

### CMLFormer - Pre-training for Code-Mixed Languages

*In progress*

A novel mechanism for pre-training code-mixed language models by attending to linguistic features.

### Can LLMs understand Math? - Exploring the Pitfalls in Mathematical Reasoning

*In progress*

An evaluation framework to assess mathematical reasoning abilities of LLMs through step-by-step quantitative grounding.

### MAPLEv2 - Multi-task Approach for generating blackout Poetry with Linguistic Evaluation

An approach to more coherent and creative blackout poetry generation by leveraging multiple linguistic evaluation tasks.

### Automated Essay Grading using Knowledge Graph Decomposition

An approach to rapid automated grading of essays using self-supervised knowledge-graph decomposition.

### Object-Aware Video Indexing

A self-supervised mechanism of tagging and querying video databases using object detection and semantic querying.

## TEACHING EXPERIENCE

---

### Teaching Assistant, CS322: Big Data

Bengaluru, India

PES University; Professors: Dr. KV Subramaniam, Dr. Prafullata K Auradkar, Animesh Giri

June 2021 – Dec 2021

Designed and graded coursework, assignments and projects, and delivered hands-on sessions on Hadoop and Spark for a class of 600+ enrolled students for the undergraduate Big Data course.

## AWARDS AND SCHOLARSHIPS

---

**Webex Analytics Datathon**, Cisco Systems

2<sup>nd</sup> / 20+ teams

2024

**Webex IDEA Hackathon**, Cisco Systems

1<sup>st</sup> / 300+ teams

2023

**Webex Playtime Hackathon**, Cisco Systems

Top 20 Globally & Regional Winner / 300+ teams

2023

**Undergraduate Researcher Award**, PES University

One among 900+ students

2022

**Prof. CNR Rao, MRD & DAC Scholarship Awards**

Top 20% among 900+ students

2022

**Finalist**, Intel Technovation, Flipkart, IBM, IISc Hackathons

One among 200+ teams

2022

**National newspaper coverage**, ToI

Remodelled garbage collection tracking in Bengaluru

2017

## SKILLS

---

**Languages:** Python, Scala, Java, C, R, Groovy, Octave, SQL, L<sup>A</sup>T<sub>E</sub>X

**ML/Stats:** PyTorch, Tensorflow, HuggingFace, NLTK, pandas, NumPy, scikit-learn, seaborn, matplotlib

**Artificial Intelligence Techniques:** Representation Learning, Transfer Learning, Few-Shot Learning, Language Models, Natural Language Understanding

**Big Data/Cloud:** Hadoop, Kafka, Zookeeper, Spark, Flink, Iceberg, Pinot, ELK

**Frameworks/Tools:** Git, Docker, Flask, Grafana, PSQL, MongoDB, AWS, Linux

## SERVICES AND VOLUNTEERING EXPERIENCE

---

### Speaker, Guest Lecture on - Building Foundation Models using Transformers

Bengaluru, India

PES University

Sep 2023

Delivered a guest lecture to undergraduate students on the advancements in representation learning techniques for language and highlighted the importance of interdisciplinary research.