

# Curriculum Vitae

## Aakash Roy

Associate Data Scientist III, Elevance Health Inc.  
Indian Institute of Technology Bombay, India  
LinkedIn: [linkedin.com/in/aakash-roy-18723610b](https://www.linkedin.com/in/aakash-roy-18723610b)

Gmail: [aakashroyiitb@gmail.com](mailto:aakashroyiitb@gmail.com)  
Web: <https://iaakashroy.github.io/web/>  
GitHub: <https://github.com/iaakashRoy>

## ABOUT ME

I am a Data Scientist with over a year of experience, passionate about uncovering insights from mathematical patterns. With a strong foundation in mathematics, I have explored diverse AI applications, including NLP, recommendation systems, and computer vision.

## RESEARCH INTERESTS

Inspired by challenges in AI, my research interest focuses brain-inspired architectures. I am particularly interested to develop learning algorithms that not only mimic brain-like processing, reasoning, and memory but also serve as tools to uncover brain functions and cognitive mechanisms by combining neuroscience, deep learning, and reinforcement learning for efficient, safe, and aligned models.

## EDUCATION

**Master of Science (M.Sc.) in Industrial Engineering and Operations Research** Jul 2021–Jul 2023  
Advisor: Prof. Jayendran Venkateswaran | *Indian Institute of Technology, Bombay*  
CGPA: **8.29/10**  
Major Courses: Deep Learning for NLP, Distributed Machine Learning, Online Machine Learning, Introduction to Machine Learning, Optimization Models, Engineering Statistics, Stochastic Processes, Linear Systems

**Bachelor of Science (B.Sc.) in Mathematics** Jul 2017–Oct 2020  
The Bhawanipur Education Society College | *University of Calcutta, India*  
Percentage: **75.5%, 1st Class (Honors)**  
Major Courses: Abstract Algebra, Linear Algebra, Real Analysis, Differential Equations, Probability, Statistics

**West Bengal Council of Higher Secondary Education** Apr 2016–May 2017  
Barrackpore Government High School | *West Bengal, India*  
Grade: **90%** in Mathematics

**West Bengal Board of Secondary Education** Apr 2014–May 2015  
Nona Chandanpukur Manmatha Nath Higher Secondary School | *West Bengal, India*  
Grade: **93%** in Mathematics

## PROFESSIONAL EXPERIENCES

**Associate Data Scientist III | Carelon Global Solutions (Elevance Health Inc.)** Jul 2023–Present

- **Current Project: Health Insurance Benefit Chatbot Assistant (BeCA) - Orchestration**
  - Developed a **multiclass, multilabel in-chat intent detection model, utilizing Transformers to classify sequence of follow-up questions** within chat conversations, distinguishing between straightforward and ambiguous ones, based on, “*On the Use of Transformer-Based Models for Intent Detection Using Clustering Algorithms.*”
  - **Deployed** the initial model using **One-vs-Rest Linear SVM with a probabilistic approach.**
- **Previous Project: BeCA - Plan-Level**
  - Developed advanced **text chunking, model and prompt tuning techniques**, optimized for **efficient and accurate retrieval of information from user plan-level data, enabling concise answer formation (Deployed).**

**Data Science Intern | Carelon Global Solutions (Elevance Health Inc.)** May–Jul 2022

- Conducted analysis on the **Adjustment Inventory of health insurance claims**, focusing on classifying which claims require adjustment by performing **Exploratory Data Analysis (EDA) on large-scale data** to identify trends and patterns for actionable insights.
- Utilized various NLP & ML techniques, like **NER, BoW, TF-IDF, BERT embeddings, PCA, and KNN**, implementing concepts from the paper “*A Complete Process of Text Classification System Using State-of-the-Art NLP Models.*”

## ACADEMIC PROJECTS & READINGS

---

### Image Caption Generator Using Deep Learning

Jan–May 2023

*Instructor: Prof. Pushpak Bhattacharyya, IIT Bombay*

Utilized **Convolutional neural network (CNN)** [VGG19] to process image attributes of Microsoft COCO Dataset. Employed the “**Encoder-Decoder**” **Attention Mechanism from the “Show, attend and Tell: For Image Captioning” paper, combined with Word2Vec embeddings**, to establish relationship between text tokens and image pixels by dynamically focusing on relevant parts of an image. Constructed model using **Long short-term memory (LSTM)**, **Beam Search** for generating the image caption tokens. **Achieved BLEU-1 score 72.49(70.7 in paper)**.

### Class-Prior Estimation From Positive Labeled Data

Jan–May 2023

*Supervisor: Prof. P Balamurugan, IIT Bombay*

Executed the paper “**Class-prior estimation for learning from positive and unlabeled data**”, considering the problem of **learning a classifier using only positive and unlabeled samples**. Estimated the class-prior on an unlabeled dataset with additional samples coming only from positive class. Utilized **penalized L1 divergence for model fitting** to cancel the error caused by the absence of negative samples. Experimentally demonstrate the usefulness of **penalized f-divergences method on MNIST hand-written digit dataset**.

### Optimal Algorithms for Non- Smooth Distributed Optimization in Networks

Jan–May 2023

*Instructor: Prof. Mayank Baranwal, IIT Bombay*

Reviewed and presented the “**Distributed optimization of non-smooth convex functions**” paper, focusing on **Lipschitz continuity** and providing optimal convergence rates and algorithms.

### Masters Thesis Project: Application of Augmented Reality in Industry

Aug–Dec 2022

*Advisor: Prof. Jayendran Venkateswaran, IIT Bombay*

Developed an AR application to enhance visualization and problem-solving in industrial processes. **Integrated real-time data and 3D model rendering using WebXR, Three.js, and ARCore**, compatible with web browsers and mobile devices. **Reviewed and demonstrated the potential use cases of AR for improving layout planning, bottleneck identification, and machine maintenance in industrial settings**.

### Analysis of LinUCB and DivLinUCB in News Recommendation

Jan–May 2022

*Supervisor: Prof. Manjesh Kumar Hanawal, IIT Bombay*

Implemented both the LinUCB, DivLinUCB algorithms with **exploration-exploitation trade-off** and the paper “**Diversity in News recommendation using contextual bandits**” to find out whether it is effective in building recommendation engines. Instead focusing solely on **maximizing click through rate**, which may overexpose some articles whereas the remaining ones may be recommended very rarely, we considered the “**historical frequency**”, **or the number of times it has been recommended as the “cost”** of recommending it. This project is a comparison of performance between the above mentioned algorithms. **Achieved CTR value of 0.9508 with alpha value 0.001**. Validated the accuracy and correctness by observing the UCB values of predicted arms.

### TCAAI Project: AI Driven Credit Scoring Model For Farmers (Virenxia)

Jan–May 2022

*Supervisor: Prof. Usha Anantakumar, IIT Bombay*

Developed a credit scoring model for farmers, addressing challenges such as low financial inclusion, limited access to credit, and unstructured data. Machine learning techniques, including **K-Means and Hierarchical Clustering**, were used to **group farmers based on behavior, with key factors being crop production, land area, and selling price**.

### Optimization Modeling and Data Visualization

Jul 2021–Apr 2022

*Instructor: Prof. P Balamurugan, IIT Bombay*

Explored various Optimization models using **pyomo**, solved them using **cbc, glpk**. Implemented **Gradient descent, Line search, Newton’s algorithm and BFGS algorithm** for solving Unconstrained Non-linear optimization problems & performed loss minimization. Visualized data using **seaborn, matplotlib** and summarized using **descriptive statistics**. Performed **Column normalization, Clustering by k-means, Dimensionality reduction by PCA**.

## ACHIEVEMENTS

<b>Recipient of the Departmental Teaching Assistantship</b> <i>Indian Institute of Technology, Bombay</i>	2021–2023
<b>Received Pre-Placement Offer after completion of AI/ML Internship</b> <i>Carelon Global Solutions(Elevance Health Inc.)</i>	2022
<b>Selected among top 13 candidates in India for MSc Programme in the only Industrial Engineering &amp; Operations Research department in India</b> <i>Indian Institute of Technology, Bombay</i>	2021
<b>Cleared Chennai Mathematical Institute Master's Entrance Examination</b> <i>Organized by the Chennai Mathematical Institute, Chennai, India</i>	2021
<b>Selected for the Mathematics Training and Talent Search program</b> <i>Organized by the Mathematics Training and Talent Search Trust, India</i>	2019
<b>Bagged Certificate in Mathematics Talent and Quiz Competition Analytica</b> <i>Organized by the Department of Mathematics, St. Xavier's College, Kolkata, India</i>	2019
<b>Recipient of Swami Vivekananda Merit Cum Means Undergraduate Scholarship</b> <i>Offered by the Government of West Bengal</i>	2018-2020

## COURSES

<b>Comprehensive list of pertinent courses completed during my BSc and MSc studies</b> <i>B.Sc-M.Sc Coursework</i>	2017-2023
<b>Compilation of various other courses studied from different online platforms</b> <i>Other Relevant Courses</i>	2022-2024

## WORKSHOPS AND SEMINARS

<b>Data Science Bootcamp organized by Analytics club</b> <i>IIT Bombay and DPhi</i>	Jun 2022
<b>Python &amp; Introduction to Machine Learning</b> <i>University of Calcutta</i>	Feb 2019
<b>National Seminar and Conference on Contemporary Research in Theoretical and Applicable Mathematics</b> <i>University of Calcutta</i>	Aug 2018
<b>Hands-on workshop on programming languages for Mathematicians</b> <i>University of Calcutta</i>	Feb 2018

## CERTIFICATIONS

<b>DS0101EN: Introduction to Data Science, IBM</b>	Dec–Jan 2021
<b>Machine Learning A-Z: Hands-On Python &amp; R In Data Science, Udemy</b>	Jul–Dec 2021

## SKILLS

<b>Machine Learning</b>	Scikit-learn, Scipy, TensorFlow, PyTorch, DeepEval
<b>General Software Engineering</b>	Python (Numpy, Pandas, Langchain), PySpark
<b>Data Visualization</b>	Matplotlib, Seaborn
<b>Misc</b>	Git, Docker, Visual Studio, Jupyter, PostMan, AWS, L <sup>A</sup> T <sub>E</sub> X, Pyomo
<b>Languages</b>	English, Bengali, Hindi