Curriculum Vitae

iaakashroy.github.io/web

Aakash Roy

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

Email: 21i190007@iitb.ac.in
Gmail: aakashroyiitb@gmail.com
GitHub: https://github.com/iaakashRoy

ABOUT ME

I am a Data Scientist with over a year of professional experience, driven by a deep curiosity to unravel the meanings embedded in mathematical equations and natural patterns. I am fascinated by the mysteries of existence and excited about the future. Regarding AI, I am both optimistic and cautious about its potential.

RESEARCH INTERESTS

I am interested in exploring and comparing the boundaries between existing multimodal models and the human brain. 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
I am currently involved in **Health Insurance Benefit Chatbot Assistant (BeCA) - Orchestration, a multiclass, multilabel in-chat intent detection model development that utilizes Transformers** to classify straightforward and ambiguous follow-up questions according to their order in the chat conversation. This work is informed by the paper **"On the Use of Transformer-Based Models for Intent Detection Using Clustering Algorithms"**. The initial deployment of the model was carried out using **One-vs-Rest Linear SVM, incorporating a probabilistic approach**.
Prior to this, I was part of the development team for the **BeCA - Plan-Level**. Executed various **text chunking, model fine-tuning & prompt engineering techniques** for faster and accurate information retrieval from plan-level user data and concise answer formation. **This solution is already in production.**

Data Science Intern | ***Carelon Global Solutions (Elevance Health Inc.)*** May–Jul 2022
Conducted an 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 insights. Utilized various NLP & ML techniques, including **NER, BoW, TF-IDF, BERT embeddings, PCA, KNN** by implementing **"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**.

Application and Comparison 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.

Industry Project: AI Driven Credit Scoring Model For Farmers (Virenxia, TCAAI) 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 Nonlinear 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

Recipient of the Departmental Teaching Assistantship 2021–2023
Indian Institute of Technology, Bombay

Received Pre-Placement Offer after completion of AI/ML Internship 2022
Carelon Global Solutions(Elevance Health Inc.)

Selected among top 13 candidates in India for MSc Programme in the only Industrial Engineering & Operations Research department in India 2021
Indian Institute of Technology, Bombay

Cleared Chennai Mathematical Institute Master's Entrance Examination 2021
Organized by the Chennai Mathematical Institute, Chennai, India

Selected for the Mathematics Training and Talent Search program 2019
Organized by the Mathematics Training and Talent Search Trust, India

Bagged Certificate in Mathematics Talent and Quiz Competition Analytica 2019
Organized by the Department of Mathematics, St. Xavier's College, Kolkata, India

Recipient of Swami Vivekananda Merit Cum Means Undergraduate Scholarship 2018-2020
Offered by the Government of West Bengal

COURSES

Comprehensive list of pertinent courses completed during my BSc and MSc studies 2017-2023
B.Sc-M.Sc Coursework

Compilation of various other courses studied from different online platforms 2022-2024
Other Relevant Courses

WORKSHOPS AND SEMINARS

Data Science Bootcamp organized by Analytics club Jun 2022
IIT Bombay and DPhi

Python & Introduction to Machine Learning Feb 2019
University of Calcutta

National Seminar and Conference on Contemporary Research in Theoretical and Applicable Mathematics Aug 2018
University of Calcutta

Hands-on workshop on programming languages for Mathematicians Feb 2018
University of Calcutta

CERTIFICATIONS

DS0101EN: Introduction to Data Science, IBM Dec–Jan 2021

Machine Learning A-Z: Hands-On Python & R In Data Science, Udemy Jul–Dec 2021

SKILLS

Machine Learning	TensorFlow, PyTorch
General Software Engineering	Python (Numpy, Pandas, Scikit-learn, Scipy), PySpark
Data Visualization	Matplotlib, Seaborn
Misc	Git, Docker, Visual Studio, Jupyter, PostMan, AWS, LaTeX, Pyomo
Languages	English, Bengali, Hindi