Dhruv Belai

Boston | MA | (857) 763-9077
belai.d@northeastern.edu | linkedin.com/in/dhruv-belai | github.com/DeeyBeey | deeybeey.vercel.app

EDUCATION

Northeastern University, *Boston, MA* *September 2023 – April 2025*
Master of Science in Computer Science
Relevant Coursework: Scalable Distributed Systems, Algorithms, Database Management Systems, Web Development

University of Mumbai, *Mumbai, IN* *January 2020 – June 2023*
Bachelor of Engineering in Computer Engineering
Relevant Coursework: Data Structures, Big Data Analytics, Software Engineering, Cloud Computing, Operating Systems

SKILLS

Languages: Python, R, Java, JavaScript, Bash, SQL, HTML, CSS
Data &Analytics: MySQL, PostgreSQL, SQLite, MongoDB, DynamoDB, Tableau, PowerBI (DAX), Qlik, Excel
Cloud & Tools: AWS, EC2, Lambda, S3, RDS, SageMaker, Glue, Git, GitHub, Docker, VMWare, Jupyter, VS Code, IntelliJ
Frameworks: Pandas, NumPy, Matplotlib, Seaborn, Keras, TensorFlow, PySpark, Scikit-learn, Transformers, OpenCV, PyTorch, googletrans, tidyverse, ggplot2, dplyr, tidyr, shiny, Express, React, Node.js
Certifications: AWS Academy Cloud Architecting, Cloud Foundations, Data Engineering, IBM – ML with Python

WORK EXPERIENCE

Data Scientist Teaching Assistant, *Northeastern University* *September 2024 – April 2025*
- Provided constructive feedback while grading 15 assignments and 6 practicums for a class of 300+ students.
- Guided 100+ students in hypothesis testing, correlations, predictive modeling, and assessing statistical significance.
- Led a team of 15 TAs and assisted students for 10+ hours weekly in office hours, maintaining 99% grading accuracy.

Data Scientist, *Nibodh Educare* *July 2022 – August 2023*
- Preprocessed and wrangled student engagement and performance data using PySpark to identify dropout risks, driving a 10% improvement in student retention – a key KPI – through targeted resource allocation and interventions.
- Developed K-Means clustering for student segmentation and Random Forest models for dropout prediction using SageMaker, leading migration to AWS, which cut infrastructure costs by 30% and improved scalability.
- Built Tableau dashboards to visualize key trends, driving data-backed decisions on student retention strategies.
- Delivered insights to stakeholders, enhancing cross-functional collaboration and data-driven decision-making.

Software Engineer Intern, *Nibodh Educare* *January 2022 – June 2022*
- Built a full-stack web portal using HTML, CSS, JavaScript, and the MERN stack, integrating front-end and back-end with object-oriented principles and reusable components – reducing redundant code and saving development effort.
- Boosted system responsiveness by 20% via AWS, microservices, and scalable RESTful APIs, ensuring high availability.
- Drove agile development, using Git for version control and automating testing with Cypress and Jest.
- Liaised between educators and developers, turning teacher feedback into product specs that led to a user-friendly grade tracking module and strong pilot feedback.

PROJECTS

Paymate, *Automate Bill Splitting* *January 2025 – April 2025*
- Built a MERN stack web app to automate bill splitting, replacing manual Excel workflows, and saving users significant time, integrating Tesseract OCR for receipt scanning, and OpenAI LLM API to extract and summarize expense details.

NBA Game Outcome Prediction Using Neural Networks *October 2024 – December 2024*
- Built and optimized PyTorch models to predict NBA outcomes from 25 years of SQL-extracted data, achieving 65% accuracy – a 15% gain over baseline via feature engineering, hyperparameter tuning, and early stopping.

Distributed Key-Value Store *June 2024 – August 2024*
- Built a distributed key-value store in Java using TCP, UDP, and RMI, with custom architectures supporting get/put/delete operations, 2PC for fault-tolerant transactions, and PAXOS for high availability and consistency.