

EDUCATION

- **Master's in data science**, Rutgers University New Brunswick May 2024
- **Bachelor's in computer science**, Indian Institute of Technology Indore May 2019

TECHNICAL SKILLS

- **Programming Languages:** Python, R
- **Libraries and Frameworks:** Git, Spark, PyTorch, TensorFlow, MongoDB, Postgres SQL
- **Cloud:** lambda, S3, CloudWatch, Sage maker, IAM, EC2
- AWS certified Cloud practitioner
- AWS certified Machine Learning specialist

EXPERIENCE

Unit Computing Specialist (student worker), Rutgers UCM

Feb 2023 - Present, New Brunswick

- Wrote Python programs to clean Excel data, eliminating over 50 hours of manual labor every quarter.
- Created Tableau dashboards for budget analysis and IT ticket trends, helped identify and decrease recurrent issues by 70%.

Data Analyst, Deloitte

June 2019 - Jan 2022, Bangalore

- Reduced banking customer churn by 33% during lockdown by identifying churn customers and key factors using logistic regression, while helping the client target strategies toward online features.
- Cut down on over 35 hours of manual labor per week to identify boats on a dock. used AWS Lambda and CloudWatch to automate the development of a change detection system with OpenCV and Mask R-CNN.
- Boosted a restaurant chain's take-away orders by over 80% during COVID-19 through customer segmentation using DBSCAN clustering that drove a tailored marketing approach. Also developed a Tableau dashboard to present findings.
- Enhanced medical record query speed by over 110% for 800k+ records by setting up an ETL pipeline with AWS Glue and using Amazon Comprehend for Named Entity Recognition. Utilized identified entities as tags for improved indexing.
- Mentored interns, reviewed code of peers and followed Agile methodology for project management.
- Was recognized with applause award twice for my client centric work approach and timely deliverables.

Database Management Intern, MAQ Software

May 2018 - July 2018, Hyderabad

- Implemented an ETL pipeline using SQL Server Management Studio and SSIS, consolidating three large data marts with over 2 million records into one structured format. Developed triggers and stored procedures to identify and rectify table errors and inconsistencies during transfer.

PROJECTS

Chatbot model with a personality: [\[Github \]](#)

- Built a generative AI (Gen AI) model chatbot to replicate Chandler Bing's dialogue style from "Friends," utilizing an extensive dataset of 8,700 dialogues. The model, featuring a seq2seq with 2-layer LSTM network with a dropout layer, achieved a BLEU score of 0.63.
- Parallely fine-tuned Google's T5 small model with the same dataset for benchmarking purposes.

2024 Travelers Insurance Analytics University Contest:

- Performed Tweedie regression on a zero-inflated dataset with over 29,000 records, fine-tuning parameters via grid search and assessing model efficacy with the Gini index. This methodical approach garnered us third place among 200+ teams.

Twitter Search: [\[Github \]](#)

- Designed a web application with a local cache of 200 trending tweets, leveraging a combination of Postgres (relational) and MongoDB (non-relational) to query a dataset of about 120,000 tweets from 13,000 users.
- Applied NLP techniques for efficient search, including synonym search and Levenshtein distance, and managed API requests and the web app using Flask.

Ear image recognition:

- Utilized a CNN model with SIFT features to implement image recognition on ear pictures for biometric verification; 91% accuracy and 73% precision were attained. high-resolution photos of the pupils that were taken and prepared for the dataset.

VOLUNTEERING

- **Club Head, AVANA:** Led AVANA, the campus social welfare club at IIT Indore, spearheading initiatives such as cleanliness drives, weekend teaching workshops for underprivileged children, and blood donation drives.
- **Volunteer Team Lead, Breathe India:** Facilitated access to medical resources during the COVID-19 second wave by updating our website daily with information on available beds, nebulizers, and medicines, and promptly connecting individuals to necessary healthcare services.