Gautam Chauhan

Bloomington,Indiana|812-606-3239|gchauha@iu.edu

EDUCATION

Indiana University Bloomington **Aug 2021-May 2023**
Master of Science in Data Science GPA: 3.43
Primary Specialization: Computational and Analytical Data Science
Coursework: Applied Machine Learning, Advanced Database Systems, Data Mining, Network Science, Applied Algorithms, Statistics

CMR Institute of Technology **Aug 2015-June 2019**
Bachelor of Technology in Information Science & Engineering GPA:7.22 First Class
Coursework: Data Warehousing and Mining, Database Design and Development, Software Engineering, Computer Networks, Object Oriented programming in C, C++ and Java, Data Structures, Engineering Mathematics, Entrepreneurship

University of Copenhagen – Copenhagen Centre for Social Data Science **Aug 2018-Sept 2018**
Social Data Science Summer Training
Coursework: Textual Data Mining, Data Warehousing, Ethics & Methodologies of Data Acquisition, Techniques of Data Wrangling, Machine Learning.

TECHNICAL SKILLS

Languages/ Developer Tools: Python, Shell, C++, Java, R, Git, Docker
Databases: MySQL, PostgreSQL
Software: PowerBI, Tableau, Excel
Libraries: Pandas, Numpy, Scikit-Learn, OpenCV, Keras, TensorFlow, Seaborn, SciPy, Statsmodels
Machine Learning: Classification, Regression, Clustering, Neural Networks, Self-Supervised Learning,Unsupervised Learning
Applications: Image Processing, Computer Vision, NLP

PROFESSIONAL EXPERIENCE

Optimal Strategix Group (Bangalore, India): Data Scientist / Python Developer / Python Developer Intern **Jan 2019 – Aug 2021**

- Analyzed data to build predictive machine learning models as client deliverables for various Fortune 500 companies.
- Built end to end models using AWS S3 storage, sagemaker and lambda services which in turn was capable of feeding insights directly to the client dashboard.
- Designed and developed a smart nudging algorithm for Patient Engagement Platform capable of segmenting patients and delivering nudges based on Segment of Patient *x* Behavior of Patient.
- Programmed and trained a Convolutional Recurrent Neural Network for text detection and OCR on videos having high accuracy.
- Formulated and developed a lighter and efficient version of the statistical technique 'ASEMAP' for conjoint analysis.
- Built and Deployed 20+ analytical models API's using Python, Flask, Django in collaboration with Product team.
- Delivered 10+ analytical real time dashboards using Dash, Flask and Streamlit.
- Increased performance and accuracy of existing segmentation and sentiment analysis by 20%.
- Developed and maintained python scripts for server maintenance.

Veda Labs (Delhi, India): **Computer Vision – ML Intern** **Sept 2018 – Dec 2018**

- Developed a CNN model for detection of abnormal behaviors using cctv footage.
- Helped raise 1 million INR in a seed round by going on Live TV.

PROJECTS

Multi-Modal Brain Tumor Segmentation using 3-D Convolution Neural Network (Python, Keras)

- Trained 100GB+ 3-D Brain images on a U-Net Architecture and a novel U-net made by replacing skip connections with Gating Signal.
- Novel Architecture gave similar accuracy to U-net by only using 20% of the parameters which was further increased by 3% using transfer learning from VGG-19.

Customer Aware Recommender System (Python, Laravel, Angular, Tensorflow):

- Designed a nudging algorithm which won first prize in National Level Hackathon.
- Algorithm reduced redundant nudges and had increased 30% Click-Through-Rate and is currently used by Indian e-commerce giant 'Paytm'.

Severity of Accidents in Great Britain (Python, Sci-Kit-Learn, Statsmodels)

- Performed Statistical Analysis and identified causes of severe accidents under guidance from postdoc researchers at Copenhagen Centre for Social Data Science.
- Final deliverable included importance of risk factors by crash injury severity of over 2 decades of data.

Aspect Based Sentiment Analysis for CX (Python, Tensorflow, AWS S3):

- Developed an extension of BERT Model for detecting sentiments towards key aspect terms for a customer experience setting capable of co-extracting sentiments for each keyword having an increased accuracy.