Chetana Sharma

sharma.c@northeastern.edu | +1(857)250-1186 | Boston, MA, 02215
Github: schetana12 | LinkedIn: schetana12 | Blog
Available: Jan 2021 - Aug 2021 for Co-op

EDUCATION

Northeastern University, Boston, MA *Jan 2020 - Dec 2021 (Expected Graduation)*
MS in Data Science—3.9/4 (CGPA)
Related Courses—Data Management and Processing, Algorithms, Deep Learning, Reinforcement Learning, Supervised ML

Indira Gandhi Institute of Technology, Delhi, India *Aug 2012 - June 2016*
B.Tech in Computer Science—82/100
Related Courses—Object Oriented Programming, Data Structures and Algorithms, Database System, Data Mining

TECHNICAL SKILLS

- **Programming Languages** - R, Python, Golang, SQL
- **Libraries and Frameworks** - Keras, Streamlit, Matplotlib, Numpy, Pandas, Scikit-learn
- **Technologies** - AWS, Airflow, Docker, MongoDb, Terraform, Elasticsearch

EXPERIENCE

Developer Consultant, ThoughtWorks (Gurugram) *July 2016 - Dec 2019*
Client: ThoughtWorks' Recruitment Team
- Led a team of 3 people to build a semi-supervised solution to score Github profiles from GHTorrent
- Performed statistical tests of independence to construct a reliably negative dataset to amplify the training data. Used SVM to model profiles and prioritize potential hires
- Developed and hosted an Elasticsearch powered backend on an AWS EC2 instance for analyzing results of model

Client: Largest OOH advertising business in Europe
- Collaborated in a team of 30 people to construct a modified Genetic Algorithm and Linear Programming as part of a digital trading platform. Reduced asset availability time from 3 hours to seconds and increased asset occupancy by 10% by optimizing sale of inventory based on audience impressions
- Integrated Airflow to create workflows for data ingestion and orchestration of other batch jobs. Implemented AWS Batch for parallel execution of large workflows
- Developed a suite of core business capabilities exposed via RESTful APIs based on Microservices architecture using Golang. Worked on terraform to spin up various services of AWS

Graduate Teaching Assistant, Northeastern University (Boston) *Sep 2020 - Dec 2020*
Course: DS5110 Introduction to Data Management and Processing
- Assisting professor in grading assignments and helping students in topics like data analysis, statistical modeling, and machine learning

PROJECTS

Speech emotion based face generation using Conditional GAN [link] *July 2020 - Aug 2020*
- Identified emotion from audio with a test accuracy of 60.8% by employing a 1-D 6 layer CNN, followed by
- Generated human faces conditioned on emotion identified from audio using conditional GANs
- Built a Streamlit based web app connecting pipelines for emotion detection and conditioned face generation

Image Captioning [link] *Sep 2020 - Dec 2020*
- Explored inject and merge architectures for captioning images which vary with respect to the stage at which image information is fed into the model. Implemented merge architecture utilising Keras on Flickr8k dataset.

Instacart Market Basket Analysis *March 2020 - April 2020*
- Conducted data preprocessing, and exploratory data analysis by utilizing R packages
- Generated frequent item sets with 10% confidence to provide insights on customer purchase patterns
- Analyzed and compared various recommendation algorithms such as association rules, popular items and collaborative filtering techniques where user-based CF recommended 10 relevant products with ROC score of 0.15

Text editor for visually challenged *July 2014 - Aug 2014*
- Created an application using C# to convert the audio message of a user into text and enable the user to listen, save and email the text using speech commands.

RESEARCH WORK AND TALKS

- **Privacy Preserving Data Mining -** Published a research paper in Journal of King Saud University - Computer and Information Sciences, that proposed geometric data perturbation using 3-D rotation. It improved privacy preservation of data over 2-D rotation technique and yielded better mining results
- **ThoughtWorks' XConf -** Presented in 2018 edition of ThoughtWorks' conference named XConf on Machine Learning: Discovery to Delivery, focused on how to take a data science project from its inception phase to delivery