3213619962
achauhan@umass.edu
linkedin.com/in/anshumaan-chauhan/
github.com/Anshumaan-Chauhan02

ANSHUMAAN CHAUHAN

EDUCATION

Master's in science (MS)	**University of Massachusetts Amherst**	**3.9/4.0**	Exp. Graduation: May 2024
Relevant Coursework: Algorithms for Data Science, Systems for Data Science, Machine Learning, Artificial Intelligence, Natural Language Processing, Neural Networks: Neuroscience and Engineering			
Bachelor's in engineering (B.E.)	**BITS Pilani Dubai Campus**	**9.83/10 (CGPA)**	2018-2022

- Awarded merit scholarship of 64,640 AED on total fees based on my GPA.
- Awarded Bronze Medal for an outstanding academic performance and standing third amongst the batch of 2018.

Relevant Coursework: Artificial Intelligence, Neural Networks & Fuzzy Logics, Probability and Statistics, Object Oriented Programming, Compiler Construction, Database Systems, Data Structures and Algorithms, Operating Systems, Computer Architecture, Computer Networks, Design and Analysis of Algorithms

PROFESSIONAL EXPERIENCE

Machine Learning Researcher
Florida Institute of Technology, United States

Jan 2022 – Jun 2022

1. Natural Language Processing for Generating System Architecture
   - **Analyzed and extracted** the representation of the specifications in a subset of English language using Natural Language Processing (NLTK library) and **designed** a compiler for translating it to AADL.
   - **Accepted** in International Design Engineering Technical Conferences & Computers and Information in Engineering Conference 2022.
2. Neural Architecture Search using Reinforcement Learning
   - **Proposed** an approach using Double Deep Q Networks for the automated generation of Convolutional Neural Network architectures. **Minimized** the scalability and time complexity problems without having effect on Search Space by implementing One Shot Training and Prioritized Experience Replay.

Software Engineer Intern
Tata Communications, India

Jun 2020 – Aug 2020

- **Developed automated system** in Python using Flask, Urllib and requests libraries/frameworks, to improve the customer targeting and user experience based **on clicks per second and user heatmap on the website.**
- Performed **cross functional evaluation and strategy testing** along with a team of 5 developers and marketing analysts, to **increase the SEO rankings of the websites** while **reducing hosting costs and marketing spend by >14%**.

PROJECTS/PUBLICATIONS

Visual Story Telling (2023)

- **Developed Visual Story Telling framework** to address issues in story generation with large language models like GPT2, GPT3, PaLM, and Llama, focusing on coherence and consistency.
- Fine-tuned text generation models (DistilGPT and T5) on a **custom dataset called Plot Summary Dataset**, leveraging content conditioning and hierarchical story generation methods.
- Utilized Stable Diffusion models for **sentence-by-sentence visual conversion** based on the generated stories, **highlighting limitations** such as T5 model repetition, generation of new characters not in input, impact of PEFT methods on downstream task performance, and the inability of Stable Diffusion models to perform scene transitions.

Recipe Infusion (2023)

- **Developed Recipe Infusion** framework with Recipe Generation and Style Transfer components.
- **Fine-tuned DistilGPT** model on GPU for 15 epochs after preprocessing and concatenating the RecipeNLG and RecipeBox datasets, resulting in **improved BLEU and Perplexity** scores compared to the non-finetuned model.
- **Implemented Style Transfer** for celebrities, including Donald Trump, Taylor Swift, William Shakespeare, and Michael Scott, training T5-small models on synthetic datasets and Shakespeare's parallel corpora, showcasing the **effectiveness of rephrasing recipes in a specific style**.

Scalability Check for Machine Learning System Predicting Flight Delays (2022)

- **Developed an end-to-end pipeline** for flight delay prediction, utilizing industry-standard systems such as MySQL for data storage and Spark for data querying.
- **Evaluated the scalability** of the pipeline by measuring response time of MySQL and SparkSQL on different dataset sizes, ensuring non-exponential latency increase.
- **Applied feature extraction, engineering, and normalization techniques,** and employed various ML algorithms to accurately predict flight delays, achieving promising results with linear response time increase under increased load.

Constraint-Based Multi-Organ Identification in CT Images using Unsupervised Learning (2022)

- **Proposed** an unsupervised learning approach using Density-Based Spatial Clustering of Applications with Noise (**DBSCAN**) for avoiding the large, labeled training dataset, expense of acquisition and data anonymization.
- **Implemented knowledge-based framework** to rule out infeasible segmentations. Proposed approach showcased Dice Coefficient values of 0.784 and 0.88 for kidneys and lungs respectively.
- **Accepted** in IEEE Nuclear Science Symposium, Medical Imaging Conference and RTSD Conference 2022.

SKILLS

Programming Languages	Python, Java, SQL, HMTL, CSS, Matlab and C.
Frameworks/Technologies	Jupyter, PyCharm, VS Code, Wireshark, MySQL, Airflow, AWS, Tableau
ML Libraries	NumPy, Pandas, Scikit-learn, Keras, PyTorch, Transformers, Tensorflow, Matplotlib, Seaborn