



Anisha Bhatnagar

✉ ab10945@nyu.edu ☎ +1 5513447023  <https://www.linkedin.com/in/anishabhatnagar/>  <https://github.com/anishabhatnagar>
🏠 <https://scholar.google.com/citations?user=MoTdTKsAAAAJ&hl=en>

Education

09/2022 – 05/2024 New York, USA	Masters in Computer Science , <i>Courant Institute of Mathematical Sciences, New York University</i> GPA : 3.806/4
07/2017 – 05/2021 Noida, India	Bachelors in Technology(Computer Science and Engineering) , <i>Amity University Uttar Pradesh, India</i> CGPA : 8.79/10




Professional Experience

08/2024 – Present New York, USA	Researcher, NYU Courant <ul style="list-style-type: none">Research and engineering on pathology extraction from MRI reports by leveraging open-source Large Language Models (LLMs) to improve clinical data analysis.Achieved retrieval recall of 86% for pathology extraction by using a state-of-the-art biomedical retriever for Retrieval-Augmented Generation (RAG).
09/2023 – 05/2024 New York, USA	Teaching Assistant, New York University <ul style="list-style-type: none">Design & Innovation Graduate Course (Fall '23), Department of Computer ScienceLean Launchpad Graduate Course (Spring '24), Department of Computer ScienceNatural Language Understanding (Spring '24), Center of Data Science.
05/2023 – 08/2023 New York, USA	Student Researcher (NYU ITP Capstone Project), New York Public Library <ul style="list-style-type: none">Designed and built a web application for managing 890,000+ digital items at NYPL, enhancing organization and accessibility.Implemented automated asset tagging and metadata generation for 3 asset types with a custom Vision Transformer and Flan-T5 LLM, improving search retrieval.Enabled large-scale data processing and reduced processing time for complex queries by 45% through HPC tunneling.
08/2021 – 06/2022 Pune, India	Application Development Associate, Accenture <ul style="list-style-type: none">Developed web applications in C# and .NET for clients in the Insurance industry, balancing legacy code enhancements with innovative features.Collaborated cross-functionally with architects and QA experts for seamless solution integration.Supervised weekly production deployments, utilizing automated testing with 100+ test cases to ensure successful builds.
04/2020 – 05/2020 Noida, India	Student Intern, AT&T <ul style="list-style-type: none">Established a comprehensive deep-learning LSTM pipeline for sentiment analysis on over 10,000 mobile phone reviews, identifying key themes in user experience feedback.Enhanced customer segmentation accuracy to 97.3% by incorporating sentiment features as a key factor for segmentation and applying Spectral Clustering and Random Forests.



Skills

Data Structures & Algorithms • Deep Learning/ Machine Learning • Artificial Intelligence (AI) • Natural Language Processing • Computer Vision • Python • PyTorch • HuggingFace • Sci-kit learn • Large Language and Vision Models (LLVMs) • Large Multi-modal Models (LMMs) • Numpy & Pandas • Git • C/ C++ and Java • SQL / RDBMS • Web Development • REST APIs • Django • High Performance Computing (HPC) • Kubernetes • Docker • AWS EC2 • Google Cloud Platform (GCP) • DevOps

Projects

Hate Content Detection in Videos <ul style="list-style-type: none">Introduced Hate-LLama, a multimodal audio-visual language model, based on LLaMA-7B, finetuned for hate speech detection in online videos, utilizing techniques such as Data and Model Parallel Training.Hate-LLama analyzes both visual frames and audio to classify hate speech, achieving an accuracy of 71%.Proposed a benchmark dataset of 300 videos with 33% hate and 67% non-hate content to address the scarcity of labels.Github Link: https://github.com/anishabhatnagar/Hate-LLaMA 	
Analysis of transformer models on Hindi-English Code-Switched text <ul style="list-style-type: none">Analyzed performance changes of BERT-style models in sentiment analysis for Romanized code-switched inputs.Generated Hindi and English translations, and Hindi transliterations using GPT-3.5 and IndicXLIT models to support evaluation.Evaluated TwHIN-Bert, mBERT, and XLM-T in a zero-shot setting, noting a consistent 4-9% performance degradation.Github link: https://github.com/anishabhatnagar/hi-en-senti 	
Autonomous Racing with Reinforcement Learning <ul style="list-style-type: none">Devised AI agents to play the Trackmania F-1 racing game using the Soft Actor-Critic (SAC) algorithm, LIDAR inputs, and RNNs.Experimented with pure LIDAR, LIDAR with track progress, and hybrid environments to optimize configurations.Increased training efficiency by experimenting with sensory data integration and scored the best lap time of 35 seconds, approaching the 30-second human best.Github Link: https://github.com/anishabhatnagar/RL-Racing 	

Publications

03/2021	A Sentiment Analysis Based Approach for Customer Segmentation , <i>Recent Patents on Engineering</i> 
07/2019	Machine Learning Techniques to Reduce Error in the Internet of Things , <i>IEEE</i>  Presented at the 9th International Conference on Cloud Computing, Data Science & Engineering (Confluence), Noida, India, 2019, indexed in IEEE