# Sanchit Kabra

sanchit23@vt.edu Sanchit Kabra Sanchit-404

## Education

**Virginia Tech, Blacksburg** **2024 – Present**
*Master of Science in Computer Science by research* *Blacksburg, Virginia*

**Birla Institute of Technology and Science, Pilani** **2020 – 2024**
*Bachelor of Engineering in Computer Science(Honors)* *Pilani, Rajasthan*

## Research Experience

**Virginia Tech – *Research Assistant*** **May 2022 – Present**
*Advisor: Chandan Reddy* *Arlington, VA*

- **Inherent Biases vs. Task-Specific Limitations in Generative Language Models:** Conducted a large-scale evaluation to identify whether stereotypical disparities in model outputs stem from unfair biases against identity groups or models' task-specific limitations
  - Developed a methodology to mitigate stereotypical biases **without using bias-related information**, leading to metric gains from 34% to 70% on average
  - Prompt and instruction-tuned open-source LLMs such as **Llama2 and 3**, **Mistral**, **Phi-2**, and **Gemma** at scale. Leveraged vector quantization and multi-GPU training for Llama 70b and Mixtral models
- **Concept Formation and Alignment in Transformer-Based Language Models:** Developed a methodology using fuzzy **UMAPs** and **Louvain community clustering** algorithm to extract taxonomy of concepts and their hierarchical relationships by leveraging the inherent structure in semantic embeddings of pre-GPT transformer-based models
  - Leveraged the concept taxonomy and token engineering to mitigate stereotypical bias **by 48%** on average
- **Adversarial Learning for Code Tasks:** Implemented **Metropolis-Hastings**, genetic, and greedy attacks to generate adversarial substitutes
  - Evaluated adversarial attacks that leveraged code structure using **ASTs** to generate efficient and imperceptible adversarial code samples on pre-GPT 3.5 models
  - Created a new evaluation metric for summarization tasks that addresses the unreliability of BLEU scores by converting the code to pseudo-code first, and then to target languages

**Virginia Tech – *Graduate Research Assistant*** **Oct 2024 – Present**
*Advisor: TM Murali* *Blacksburg, VA*

- Deployed language models for knowledge extraction to identify patterns in viral behavior and transmissions
- Focused on the classification of rare viral hosts during the jump of viruses from one species to another, aiming to support **pandemic prevention efforts**
- Developed a pipeline using state-of-the-art LLMs for extracting and analyzing biological insights from genomic and epidemiological data, contributing to a better understanding of cross-species virus transmission

**Aalto University – *Research Assistant*** **June 2023 – Jan 2024**
*Advisor: Rohit Babbar* *Finland*

- Reformulated extreme classification, as a multi-label learning task where each item to be recommended is treated as separate label, leading to **12% metric gains** over traditional collaborative filtering and content-based recommendation methods
- Developed a multimodal recommendation model on top of the LightXML pipeline and trained it on a novel dataset of over 5 million data samples curated from AmazonTitles-670K dataset

## Publications

- **Biased or Flawed? Mitigating Stereotypes in Generative Language Models by Addressing Task-Specific Flaws**, Akshita Jha, Sanchit Kabra, Chandan Reddy, 2024.
- **Aligned at the Start: Concept Formation in LLM Embeddings**, Mehrdad Khatir, Sanchit Kabra, Chandan Reddy, 2024.
- **AMaizeD: An End to End Pipeline for Automatic Maize Disease Detection**, ICST 2023 - IEEE Explore, Pawan K Ajmera, Sanchit Kabra, Anish Mall, Ankur Lhila, Aaryan Agarwal.

## Projects

**Road Distress Detection using Aerial Imagery** **April 2022 – May 2023**
*Advisor: Dr. Pratik Narang* *New Delhi, India*

- Developed a novel framework for road damage detection and analysis leveraging drone-based surveillance in low-visibility conditions
- Created a method for object detection on ortho-mosaics spanning tens of kilometers, using **RetinaNet**, with experimentation on **YOLO-v7**, **Vision Transformer**, and **Faster R-CNN**
- Employed an encoder-decoder based **U-Net** architecture for blurred images and a **feature fusion network attention** model for dehazing
- Experimented with hyperspectral imaging, converting 3-channel images to 31-channel images for enhanced feature extraction

**Early Detection of Plant Diseases using Aerial Imagery** **October 2024**
*Advisor: Dr. Pawan Ajmera* *IIT-Indore Drishti Program*

- Developed a pipeline for early detection of plant diseases utilizing aerial imagery and vegetation indices, funded by the IIT-Indore Drishti Program with a grant of $5,000
- Generated various vegetation indices such as NDVI, SAVI, and WVI, enhancing detection accuracy by 12%
- Implemented Inception V3 and VGG for image segmentation, achieving accuracies of 96.4% and 95.1%, respectively
- Utilized SRCNN and SRGAN to address image blurriness caused by shaky aerial imagery

**Compiler Construction for a Custom Language** **Jan 2023 – May 2023**
*BITS Pilani*

- Constructed the compiler from scratch, implementing all five phases of compilation: lexical analysis, syntax analysis, semantic analysis, optimization, and code generation in an end-to-end process
- Implemented features such as static scoping, arithmetic and boolean expressions, and support for multiple return values
- Implemented core functionalities including assignment, I/O, conditionals, loops, and function calls
- Generated intermediate representations and assembly code, ensuring compatibility with NASM and GCC

**Bayesian Network for Graduate School Admissions Prediction** **October 2024**
*Advisor: Prof Jin-Hee Cho*

- Developed a Bayesian Network to predict graduate admissions using features such as GPA, GRE, TOEFL, research experience, and letters of recommendation
- Achieved 97% accuracy and 89% F1 score with an inductive bias model, beating an automatically generated and fully connected model and applied SMOTE to enhance predictions amidst class imbalance
- Reduced the complexity of the models by 80 independent parameters and maximum in-degrees of 2
- Conducted sensitivity analysis to evaluate uncertainty factors affecting model predictions, including noise, bin sizes, and the number of nodes

## Awards

- **Amazon Trusted AI Challenge**, 2024: A global university competition focused on responsible AI and large language model coding security. Part of **HokieTokie**, one of the 10 teams selected globally. Advisor: Rouxi Jia
- **Higher Secondary School Exams State Topper**, 2020: Ranked second in the Maharashtra State Board Examinations, among 1.5 million students. Scored a perfect 100/100 in Math, Chemistry, Physics, and Biology.
- **INSPIRE Scholarship**, 2020: Offered by the Department of Science & Technology (DST), Government of India for finishing in the top 1% in Class XII Board Examinations.

## Skills

- **Programming Languages:** Python, Java, R, Matlab, C/C++, SQL, bash
- **Frameworks and Tools:** PyTorch, TensorFlow, Numpy, Pandas, Snowflake, RAG Huggingface, SciKit, Matplotlib, pyspark
- **Technologies:** AWS, Linux

## Relevant Coursework

- Data Structures and Algorithms
- Design and Analysis of Algorithms
- Operating Systems
- Database Systems
- Compiler Construction
- Computer Architecture
- Deep Learning
- Natural Language Processing
- Multimodal Vision
- Decision Making Under Uncertainity
- Computational Learning Theory