

# Abhimanyu Hans

College Park, MD · [ahans1@umiacs.umd.edu](mailto:ahans1@umiacs.umd.edu) · (929) 832-1886 · <https://ahans30.github.io/>

## RESEARCH INTERESTS

My research interests revolve around ensuring the security, efficiency and robustness of generative models. I particularly care about memorization and privacy related problems around language models.

## EDUCATION

### University of Maryland – College Park

MS Computer Science (Advised by **Prof. Tom Goldstein**)

Fall 2022 – Dec 2024 (Expected)

### International Institute of Information Technology Bangalore

Post Graduation Diploma in Software Development

Bangalore, India

2020 – 2021

### University of Delhi

BS Mathematics (Honours)

2014 – 2017

## PUBLICATION

- Spotting LLMs With Binoculars: Zero-Shot Detection of Machine-Generated Text [\[ICML\]](#) [\[code\]](#) [\[demo\]](#)  
**A. Hans**, A. Schwarzschild, V. Cherepanova, H. Kazemi, A. Saha, M. Goldblum, J. Geiping, T. Goldstein
- Be like a Goldfish, Don't Memorize! Mitigating Memorization in Generative LLMs [\[NeurIPS\]](#) [\[code\]](#)  
**A. Hans**, Y. Wen, J. Kirchenbauer, H. Kazemi, P. Singhanian, S. Singh, G. Somepalli, J. Geiping, A. Bhatele, T. Goldstein

## AWARDS

- Invention of the Year 2024 [\[Maryland Innovate, 2024\]](#)
- Top Performer award for highest academic achievements [\[IIIT Bangalore, 2021\]](#)

## RELEVANT WORK EXPERIENCE

### University of Maryland Institute for Advanced Computer Studies

Graduate Research Assistant

College Park, MD

May 2023 - Present

- Working on language models from security and alignment perspective.

### PayPal

Machine Learning Engineer

Remote Work (India)

Oct 2021 - Aug 2022

- Automated model monitoring framework, saving 70+ FTE weeks, and deploying interactive dashboards for multiple model-specific teams (spanning 100+ model use cases)
- Employed quantitative metrics (AUC, Accuracy, KS-statistic, population stability index) and qualitative analysis (LIME) to enhance model performance and feature interpretability.

### HDFC Bank

Machine Learning Engineer

Mumbai, India

Jan 2020 - Oct 2021

- Created an anomalous transaction detection system using AutoEncoder and Isolation Forest algorithms, enhancing fraud detection capabilities to achieve 25% recall at 0.1% FPR
- Developed a user-friendly model repository CRUD application using Flask, Jinja Templates, and PostgreSQL, streamlining internal model management processes.

### Think360.ai

Associate Data Scientist

Mumbai, India

Oct 2018 - Jan 2020

- Developed a deep tree models for predicting repossession values of construction equipment.
- Developed an automated business intelligence pipeline using Tableau, R, and Bash, serving over 5000 FTE with dashboards and email reports.

## SKILLS

- Languages & Tools:** Python, Java, JavaScript, SQL, Git, Docker, Linux, AWS, Shell
- Frameworks:** PyTorch, HuggingFace suite, Flask, React, Tableau, Pandas, NumPy, Sklearn, SpringMVC

## PATENTS

---

Method for Detecting Text Generated by Large Language Models.

## PROJECTS (MACHINE LEARNING & SOFTWARE ENGINEERING)

---

### **Multi-Armed Bandit Algorithms**

(*code*)

A *NumPy*-based implementation of Greedy, Epsilon-Greedy, and Upper Confidence Bound algorithms along with their empirical performance analysis.

### **Active Learning**

(*code*) (*slides*)

An approach to ML wherein the learner queries the user with best instance for a label.

### **Model Repository**

(*code*)

Full-stack Flask app serving as a repository for models and their performance metadata.