Himanshu Singh

https://www.himanshu.phd/

Email : himanshus@iiitd.ac.in
Mobile : +91-9716812719

Education

- **Indraprastha Institute of Information Technology** Delhi, India
  Doctor of Philosophy (Ph.D.), Computer Science, CGPA: 8.24 *Aug. 2020 – Present*
  Advisor: Dr. A. V. Subramanyam
  Coursework: Advanced Machine Learning, Deep Learning, Machine Learning, Probability and Random Processes, Convex Optimization, Reproducing Kernel Hilbert Spaces and Applications, Meta Learning, Digital Image Processing

- **Guru Gobind Singh Indraprastha University** Delhi, India
  Bachelor of Technology in Computer Science and Engineering *Aug. 2013 – Jul. 2017*

Experience

- **National University of Singapore** Singapore
  Visiting Scholar *May 2025 – Oct. 2025*
  - Worked on robustness and alignment of foundation models with focus on jailbreak simulation, and adversarial prompt defense mechanisms.

- **Indraprastha Institute of Information Technology** Delhi, India
  Research Scholar *Aug. 2020 – Present*
  - Defending deep learning models against adversarial attacks on image classification using purification techniques with diffusion models.
  - Developed Language Guided Adversarial Purification method for adversarial robustness.(Patent filed)
  - Investigating adversarial robustness by combining foundation models from computer vision and natural language for image classification.

  Teaching Assistant *Aug. 2020 – May 2024*
  - Advances in Deep Learning [M'23]
  - Computer Vision [W'21]
  - Digital Image Processing [M'21, M'22]
  - Natural Language Processing [M'20]
  - Statistical Machine Learning [W'22, W'23, W'24, W'25]

- **Medway Technologies Pvt. Ltd.** Remote
  Consultant *Dec. 2021 – Dec. 2022*
  - Developed an OCR pipeline to extract text from images of printed and handwritten medical prescriptions.
  - Compare extracted text with database to identify prescribed medicines and patient information.

- **Animaker Inc.** Bangalore, KA, India
  Research Scientist (Machine Learning and Computer Vision) *Jul. 2017 – Aug. 2020*
  - Contributed in building patented text to video platform Steve AI
  - Developed and deployed a machine learning pipeline for background subtraction in images and videos.
  - Designed and implemented algorithms to identify and rank relevant sections of long videos based on activity detection, facial recognition, attention, and saliency analysis.
  - Built a recommendation system for layout optimization, enabling intelligent placement of characters and text in visual content.
  - Developed color recommendation models for scene components, utilizing user-preferred hue templates from large datasets and creating quantitative scoring models to assess the quality of color combinations.
  - Provided dynamic color palette suggestions based on existing scene components and background images or videos, enhancing visual coherence and aesthetic appeal.

Publications

- **H. Singh**, Z. Xu, A. V. Subramanyam, M. Kankanhalli, Do Prompts Guarantee Safety? Mitigating Toxicity from LLM Generations through Subspace Intervention, *Preprint.*
- **H. Singh**, A. V. Subramanyam, S. Rajput, M. Kankanhalli, Nearest Neighbor Projection Removal Adversarial Training, *Preprint.*
- **H. Singh** and A. V. Subramanyam, Language Guided Adversarial Purification, *ICASSP 2024 - 2024 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Seoul, Korea, Republic of, 2024, pp. 7685-7689.

Patents

- V.C. Bhagath, **H. Singh**, System and Method for Language Learning, *Indian Patent No. 567815. Granted, 24 Jun. 2025.*

Projects

- **Image Segmentation** Aug. 2016 – May 2017
  Amity School of Engineering and Technology
  - Performed image segmentation on Brain MRI scans to detect tumors by the use of machine learning methods. I started with implementing K-means and FCM, improved their accuracy using kernel methods. In the later part, I implemented Firefly Algorithm for optimization and getting consistent results.
- **Migration of Tinxsys from Pramati to JBoss** Jun. 2016 – Jul. 2016
  Goods and Services Tax Network (Intern)
  - Currently Tinxsys handles the whole indirect tax infrastructure. It is deployed on Pramati Server with an Oracle database and is managed by Wipro. The task was to identify necessary changes to the Tinxsys system so that it can be migrated to WildFly server and give a proof of concept.
- **Online Attendance Management System** May 2015 – Jul. 2015
  Amity School of Engineering and Technology (In-house Summer Training)
  - System developed for maintaining the attendance of the student on a daily basis in the college. The system also provides functionality for uploading of notes and sending email/notification to individual students or to a whole branch. This system implements accountability and ensures transparency.

Technical Skills

- **Languages**: Python, C++, SQL
- **Databases**: MongoDB, MySQL
- **Libraries and Software Packages**: PyTorch, Hugging Face, OpenCV, Numpy, Pandas, Scikit-learn, Matlab

Awards and Achievements

- Best TA Award - Advances in Deep Learning
- Best Performance - Research award for outstanding performance at Animaker Inc
- Bagged 1st Position at ”Syntaxomania” a coding event organized by the IEEE group at Amitech 2015
- First position in street play at Conversance’14 (Cultural fest AIACTR, Delhi)
- Recipient of scholarship for the year 2011-2012 from Shanti Gyan Niketan Sr. Sec. School
- Recipient of Certificate of Merit from CBSE for securing 10 CGPA
- School Level Topper, 65 State Rank, 556 Olympiad rank in International Olympiad of Science 2010
- Gold medalist in inter-district Taekwondo Championship, Mathura

Community Service

- Reviewer, IEEE MLSP 2025
- Reviewer, IEEE ICASSP 2025
- Volunteer, IEEE ICASSP 2024

Certifications and Trainings

- Deep Learning Specialization by Deeplearning.ai on Coursera
- Machine Learning by Stanford University on Coursera
- Machine Learning Foundations: A Case Study Approach by University of Washington on Coursera
- Machine Learning: Regression by University of Washington on Coursera
- The Data Scientist's Toolbox by Johns Hopkins University on Coursera