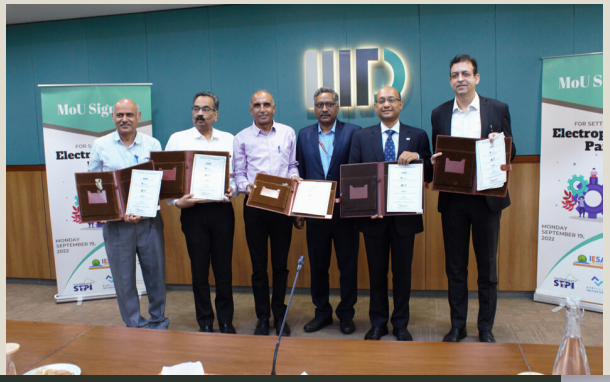
iHub Anubhuti-
IIITD Foundation

Issue No. 08
November 2022

ANUBHAV

A NEWSLETTER BY IHUB ANUBHUTI-IIITD FOUNDATION

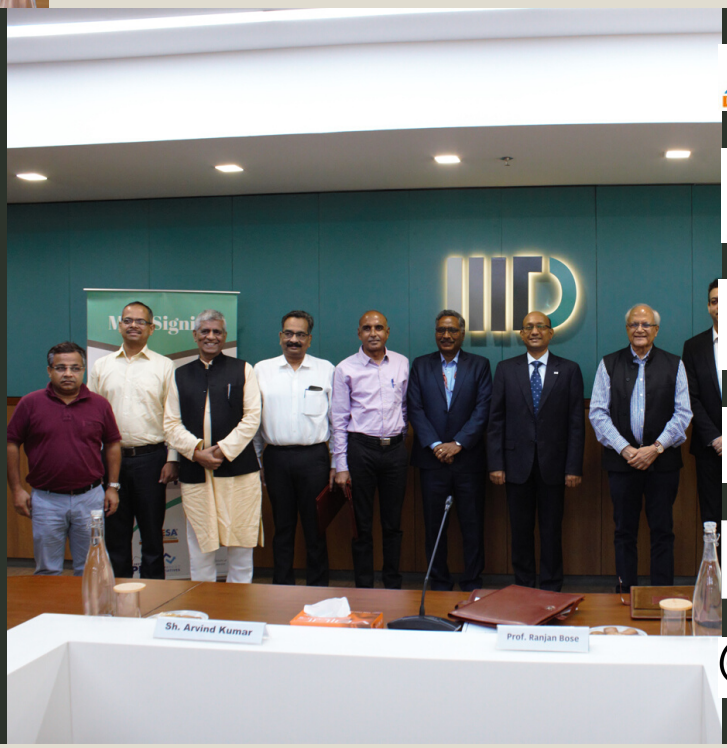


EP 2.0
MoU Signing Ceremony

iHub Anubhuti in collaboration with Software Technology Parks in India (STPI), India Electronics and Semiconductor Association (IESA), Indraprastha Institute of Information Technology (IIITD), and AIC STPINEXT INITIATIVES (STPINEXT) signed an MoU on 19th September 2022 for setting up Electropreneur Park Project 2.0 at IIIT Delhi.

iHUB Anubhuti will jointly establish research labs for prototyping and support incubation & other activities related to Electronics System Design and Manufacturing (ESDM) and Cognitive Computing and Social Sensing (CCSS).

The main objective is to develop an eco-system for research, technology development, generating next generation entrepreneurs and shall have collaborative relationships with other parties in relation to meet Electropreneur Park objectives of creating an ESDM Innovation Hub to promote entrepreneurs, startups and MSMEs.



EP plays its active contribution in the visionary and transformative India Semiconductor Mission (ISM) launched by Honourable Minister of Electronics & IT, Government of India by nurturing Start-Ups, MSMEs and academia through Greater EP programme to contribute to the growth of India ESDM.

MORE ABOUT EP 2.0

EP 2.0, with a vision of creating unicorns in the ESDM sector, to scale up its existing operations to provide a holistic ecosystem – infrastructure, network, mentors and investments, to the StartUps and MSMEs by accelerating the Government of India's endeavour of achieving Atmanirbharatha through robust Make in India, propelled by an equally thriving StartUp India.

Your Models are Vulnerable

The reach of ML in all fields

Machine learning algorithms have advanced to a stage where they can outperform humans when presented with data. Their applications have become increasingly common and widespread today, but their susceptibility to attacks remains a significant concern. When presented with a trivial but adversarial input, these algorithms fail miserably, whereas a human would still be able to perform well.

What is a model stealing/extraction attack?

Typically, to make a "trained" machine learning model accessible to the public, a company hosts it as an inference API. This publicly accessible model will be called the "victim" model. The inference API allows a customer to submit queries to the model and receive the model's prediction/output in return, usually for a minimal monetary cost. E.g. Google's Text-to-Speech API allows you to input sentences and receive the generated sound. This generation is typically performed by a neural network in the background.

The above setup makes trained ML models valuable intellectual property, which serves as motivation for thieves to try and steal these models. A model extraction attack is a way to reverse-engineer the black box victim models and attempt to create a duplicate copy which performs just as well as the victim model.

Akshit Jindal, Ph.D. – IIIT-D under the mentorship of Dr. **Vikram Goyal, Professor, IIIT-D** is researching how an ML model can be efficiently attacked in the least informative setting and still be extracted to a good enough extent.

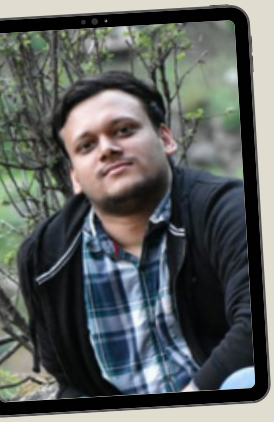
How is an extraction attack carried out?

The process of model extraction is quite similar to knowledge distillation. Attackers collect a large number of unlabelled data samples, which are then sent to the victim model as queries. The victim model outputs a prediction for each query, which the thief treats as the ground truth label for the query. The prediction might range from the confidence scores (softmax probabilities) to just the hard label (class name). These query-output pairs serve as training data for the thief's own model, allowing them to create a copy of the victim model, which we'll call the thief model. The attacks are relatively cheaper than training a model from scratch and do not require extensive training and parameter tuning

Who can try to steal a model and why?

- A malicious competitor or adversary might steal a model to craft adversarial examples. These examples can be used to "break" the victim model and showcase its shortcomings.
- The victim themselves can try to gauge the security level of their model by performing such attacks.
- A thief can try to steal the model for monetary gain, ideally by rebranding it as their own model and exposing it via a much cheaper API

DR. VIKRAM Goel



AKSHIT jindal

Project's scope and future goals

- Provide a tool/framework for performing such attacks on trained models
- Coming up with new methods for performing attacks in more and more information-restricted settings
- Building defences to prevent such attacks in the future, showcasing the efficiency of such defences via our tool
- As these methods require a deep dive into the inner workings of ML/DL models, our work can also contribute to the explainability aspect.

Blog is written by Akshit Jindal - Ph.D. , IIIT Delhi.

WEBINARS

Dr. Suman Jana
Columbia University

Title: Scalable, Accurate, Robust Binary Analysis with Transfer Learning

Suman Jana is an associate professor in the department of computer science and the data science institute at Columbia University. His primary research interest is at the intersection of computer security and machine learning. His research has received six best paper awards, a CACM research highlight, a Google faculty fellowship, a JPMorgan Chase Faculty Research Award, an NSF CAREER award, and an ARO young investigator award.

Title: Towards Autonomous Driving in Dense, Heterogeneous, and Unstructures Environments

Rohan Chandra is currently a postdoctoral researcher at the University of Texas, Austin, hosted by Dr. Joydeep Biswas. Rohan obtained his B.Tech from the Delhi Technological University, New Delhi in 2016 and completed his MS and PhD in 2018 and 2022 from the University of Maryland advised by Dr. Dinesh Manocha. His doctoral thesis focused on autonomous driving in dense, heterogeneous, and unstructured traffic environments. He is a UMD'20 Future Faculty Fellow, RSS'22 Pioneer, and a recipient of a UMD'20 summer research fellowship. He has published his work in top computer vision and robotics conferences (CVPR, ICRA, IROS) and has interned at NVIDIA in the autonomous driving team.

Dr. Rohan Chandra
University of Texas, Austin

Call for Start-Up

Anubhav Seed Funding

Anubhav Seed Funding Program was rolled out by iHub Anubhuti - IIITD Foundation to leverage entrepreneurship in the field of Cognitive Computing and Social Sensing whose primary focus areas are Healthcare, Legal Information Systems, Education, and Sustainability.

- **17 applications** were received in the first batch of the Anubhav Seed Funding Program till 29th October 2022.

- A Screening Committee meeting was held on 01/11/2022 (Tuesday) to review the applications for the Initial screening of the applications.

- Out of the 17 applications, 6 are found appropriate and recommended for the Investment Committee to present the idea in the next round of Evaluation as per the eligibility, theme and terms of the program. 11 applications will represent their pitch in front of the Screening Committee to get clarity about their idea.

To Apply: Scan the code
or
click the link to the website

https://ihub-anubhuti-iiitd.org/SeedFund/

The program offerings can be leveraged by entrepreneurs to advance and scale up their venture to the next stage.

Anubhav seed funding is for start-ups working in vertical COGNITIVE COMPUTING & SOCIAL SENSING with focus areas **HEALTHCARE**, **EDUCATION**, **LEGAL INFORMATION SYSTEM** & **ENVIRONMENT SUSTAINABILITY.**

Anubhav Seed Funding is an initiative of iHub Anubhuti to support & nurture tech start-ups with access to seed funding.

17

Applications till now

6

selected for next round

Mr. Bhat Dittakavi, CEO and Mr. **Debanjan Chatterjee**, Secretary of **AI4ICPS**, Technology Innovations Hub of IIT, Kharagpur had visited iHub Anubhuti-IIITD Foundation and had a small interaction with the iHub Anubhuti team and IIITD Faculty members.

Newsletter Editorial Team

SWATI MANGLA

Junior Manager
iHub Anubhuti-IIITD Foundation

Contact Us

info@ihub-anubhuti-iiitd.org

@Anubhutilhub

@ihubanubhuti.iiitd

@iHubanubhuti

@ihub_anubhuti_iiitd

@iHubanubhuti

5th floor, Old Academic Building, IIITD, GB Pant Polytechnic Extension, Okhla Phase-III, Delhi 110020, Ph: 011-26907335

https://ihub-anubhuti-iiitd.org