

# Privacy breach in Deep Learning through Computer Architecture

Presented by  
Kandi Vishnu Vardhan Reddy  
18CS30022

# Motivation:

- The rapidly emerging Deep Learning technology has recently triggered a substantial amount of interests in the computer security community.
- Since the Success of such applications depends on the amount of Data used to efficient DL models, Data breaches are one of the top cybersecurity problems affecting the digital economy.
- Sensitive data such as healthcare diagnostics, Face recognition, Browser history,.. is fed to the DL models during training, Which brings additional problems to protect the user identity.




## Problem definition:

- There are lots of possible threats to deep learning for intentional or unintentional exposure of sensitive information. This information can be the training data, inference queries or model parameters or hyperparameters.
- In this project we investigate the attacks on DL applications which use computer architectural explorations.
- Here our adversary does not need the ability to query the victim model; instead, he monitors the host machine where the victim's deep learning (DL) system is running and tries to retrieve hyper-parameter information by using side-channel information leakage through architectural footprints.



# Reading List (papers/books) :

- [PRIVACY IN DEEP LEARNING: A SURVEY](#)
  - [SECURITY ANALYSIS OF DEEP NEURAL NETWORKS OPERATING IN THE PRESENCE OF CACHE SIDE CHANNEL ATTACKS](#)
  - [STEALING NEURAL NETWORKS VIA TIMING SIDE CHANNELS](#)
  - [MASTIK : A Micro-Architectural Side-Channel Toolkit \(Yuval Yarom: The University of Adelaide and Data, CSIRO Adelaide, Australia\)](#)
- 

## Work Done:

- Learnt different attacks in cryptography.
- Read the required topics in neural networks to understand different research papers.
- Went through different research papers to get to know about the project.
- Tried to implement some attacks and tools to understand and gain further experience.

