HARNESSING ARTIFICAL INTELLIGENCE

Harnessing Artificial Intelligence (AI) is a top DOD priority. With this course, you will understand the principles and limitations of AI systems, and the benefits and risks of using them in military operations.

LEARNING OBJECTIVES

- To understand the nature and history of AI systems.
- To understand the kinds of AI systems that exist, their military applications, their strengths, their vulnerabilities, their strategic value, and their risks.
- To understand the four bedrock principles of AI in the military and apply them to AI projects.
- To think critically about AI, distinguishing hype from reality.

LECTURES

Each session will be a 5-minute introduction, 30-minute lecture, and 15-minute Q&A covering varied aspects of Artifical Intelliigence, incluiding:

- What is AI?
- Classifying AI by Type of Learning
- Critical Domains
- Moving Forward/AI Futures

Course: CS4000 / Harnessing Artificial Intelligence
Quarter: Fall AY2020 (Sept. 2019)
Credits: 0-2
Scheduling: M+W 1200-1300, IN-122
Grading: Pass-Fail

For more information, contact Christine Beck, christine.l.beck.ctr@nps.edu

NAVAL POSTGRADUATE SCHOOL

Supervised Learning

CS4000: Harnessing AI

Fall 2019

Marko Orescanin
marko.orescanin@nps.edu

What is Supervised Learning?

- Visual Recognition

Cat or Dog

https://towardsdatascience.com/image-classifier-cats-vs-dogs-with-convolutional-neural-networks-cnns-and-google-colabs-4e9af21ae7a8

What is Supervised Learning?

- Boston Housing Prices

average number of rooms

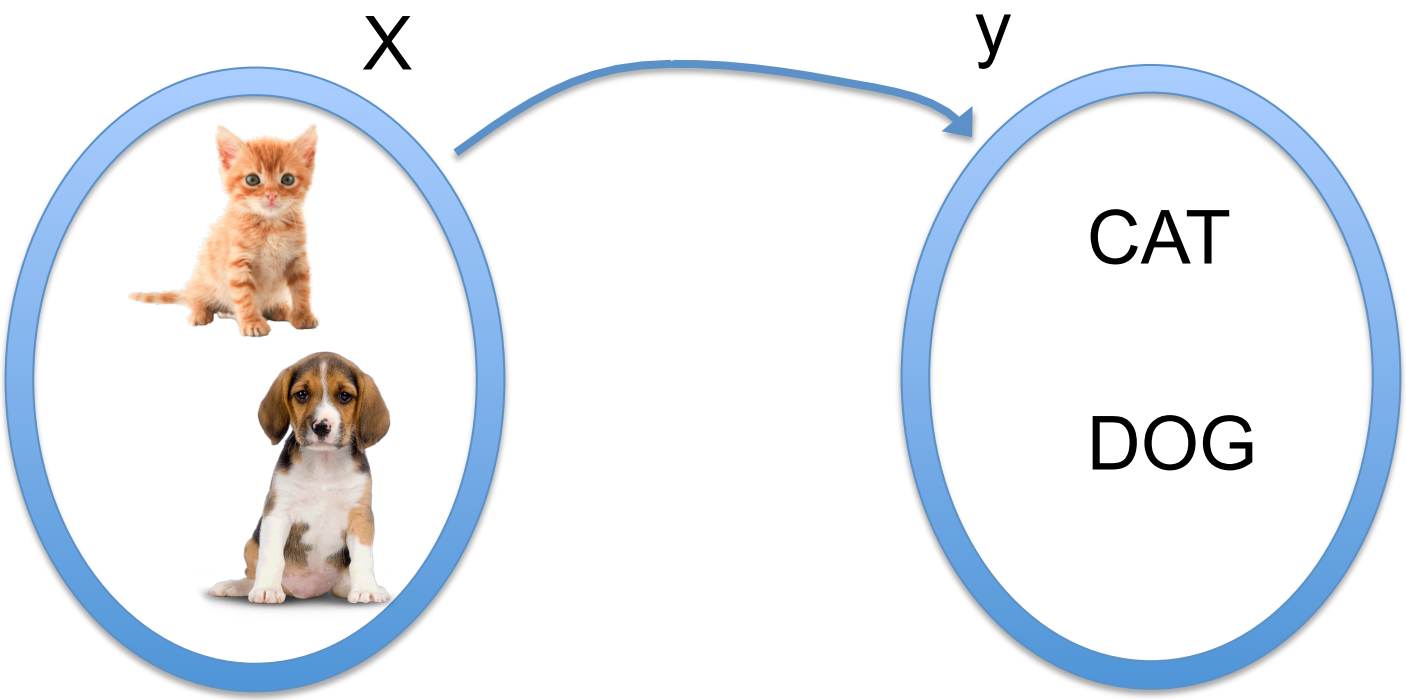
price in $1000's

	CRIM	ZN	INDUS	CHAS	NOX	RM	AGE	DIS	RAD	TAX	PTRATIO	B	LSTAT	PRICE
0	0.00632	18.0	2.31	0.0	0.538	6.575	65.2	4.0900	1.0	296.0	15.3	396.90	4.98	24.0
1	0.02731	0.0	7.07	0.0	0.469	6.421	78.9	4.9671	2.0	242.0	17.8	396.90	9.14	21.6
2	0.02729	0.0	7.07	0.0	0.469	7.185	61.1	4.9671	2.0	242.0	17.8	392.83	4.03	34.7
3	0.03237	0.0	2.18	0.0	0.458	6.998	45.8	6.0622	3.0	222.0	18.7	394.63	2.94	33.4
4	0.06905	0.0	2.18	0.0	0.458	7.147	54.2	6.0622	3.0	222.0	18.7	396.90	5.33	36.2

Can we predict value of the house based on the average number of rooms per dwelling ?

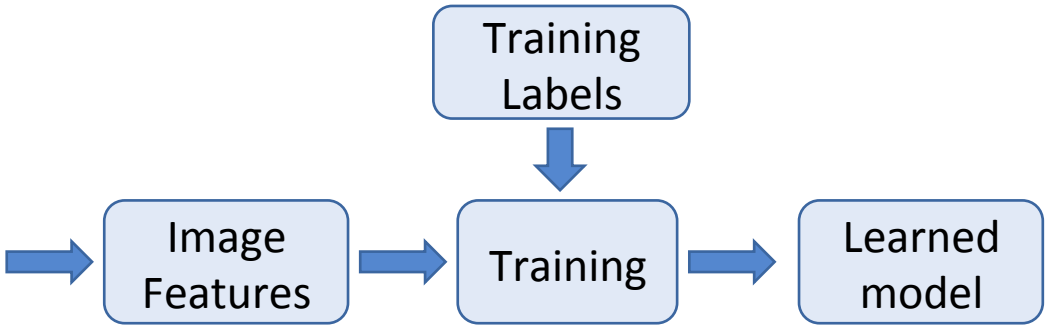
More Formal Definition

- Learn mapping between the pairs *(**x**, y)*, where **x** is input and *y* is label

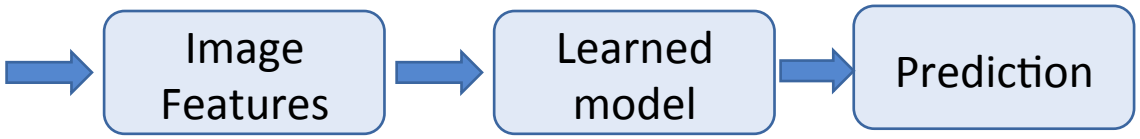


Supervised Learning Framework

Training



Testing

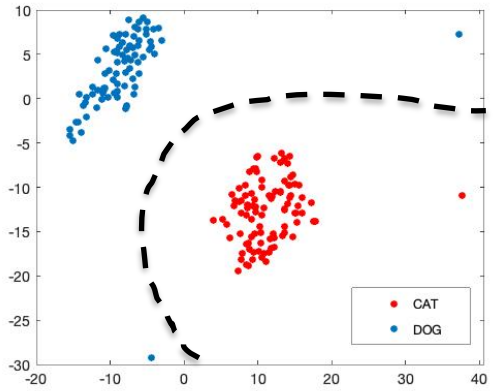


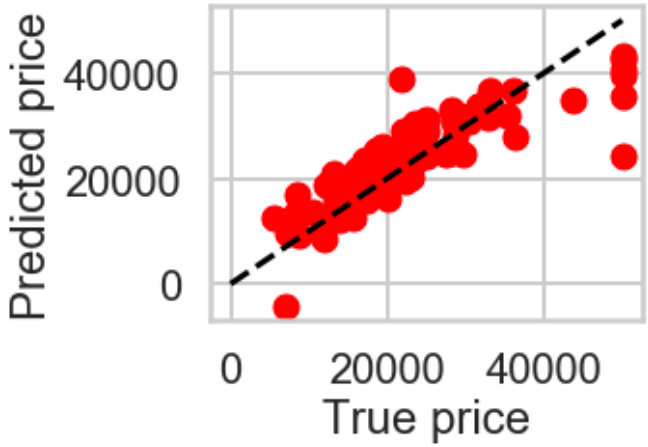
Test Image

Supervised Learning

CLASSIFICATION
DISCRETE

REGRESSION
CONTINUOUS





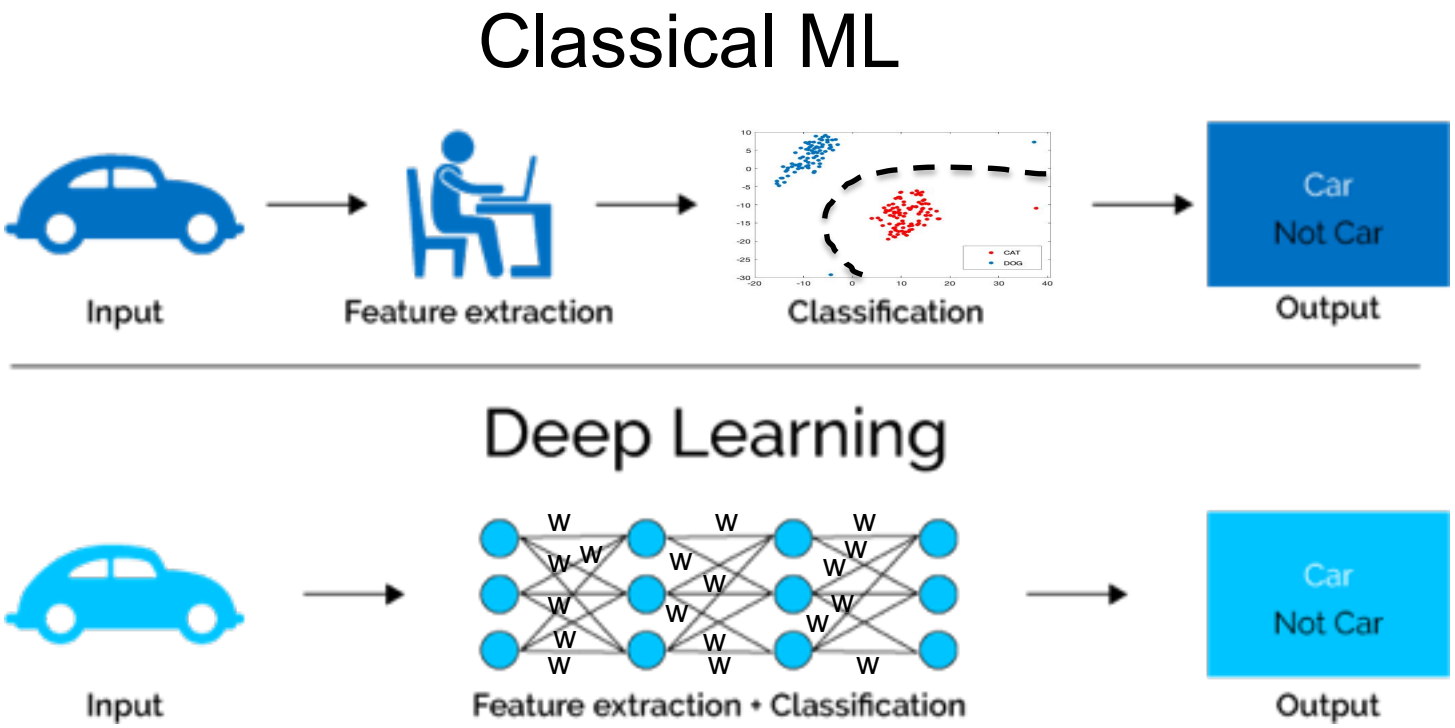
ML Classical Algorithms

- Support Vector Machines
- Logistic Regression
- Linear Regression
- Decision Trees
- Random Forests
- k Nearest Neighbor

Require experts to engineer features !

Some can be effective on smaller datasets!

What is Deep Learning

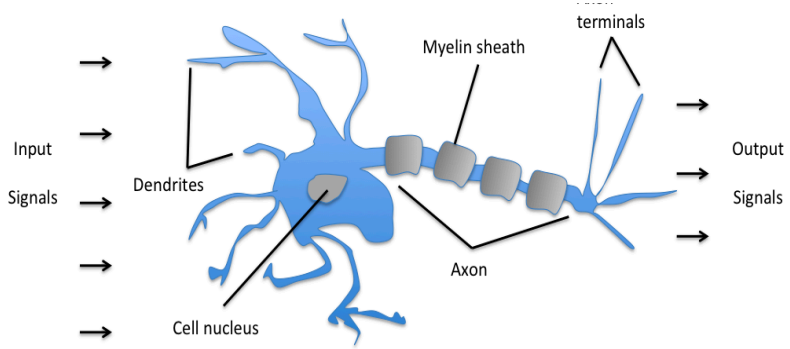


https://www.xenonstack.com/blog/static/public/uploads/media/machine-learning-vs-deep-learning.png

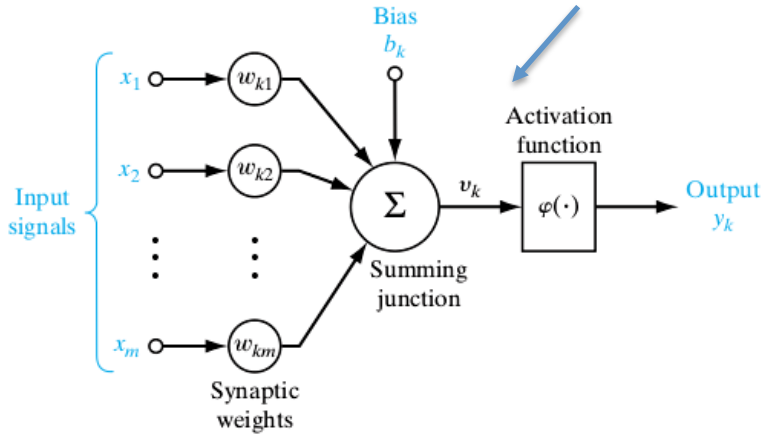
Neural Networks

- First emerged in 60's
- In 2010's started achieving human like performance on many tasks

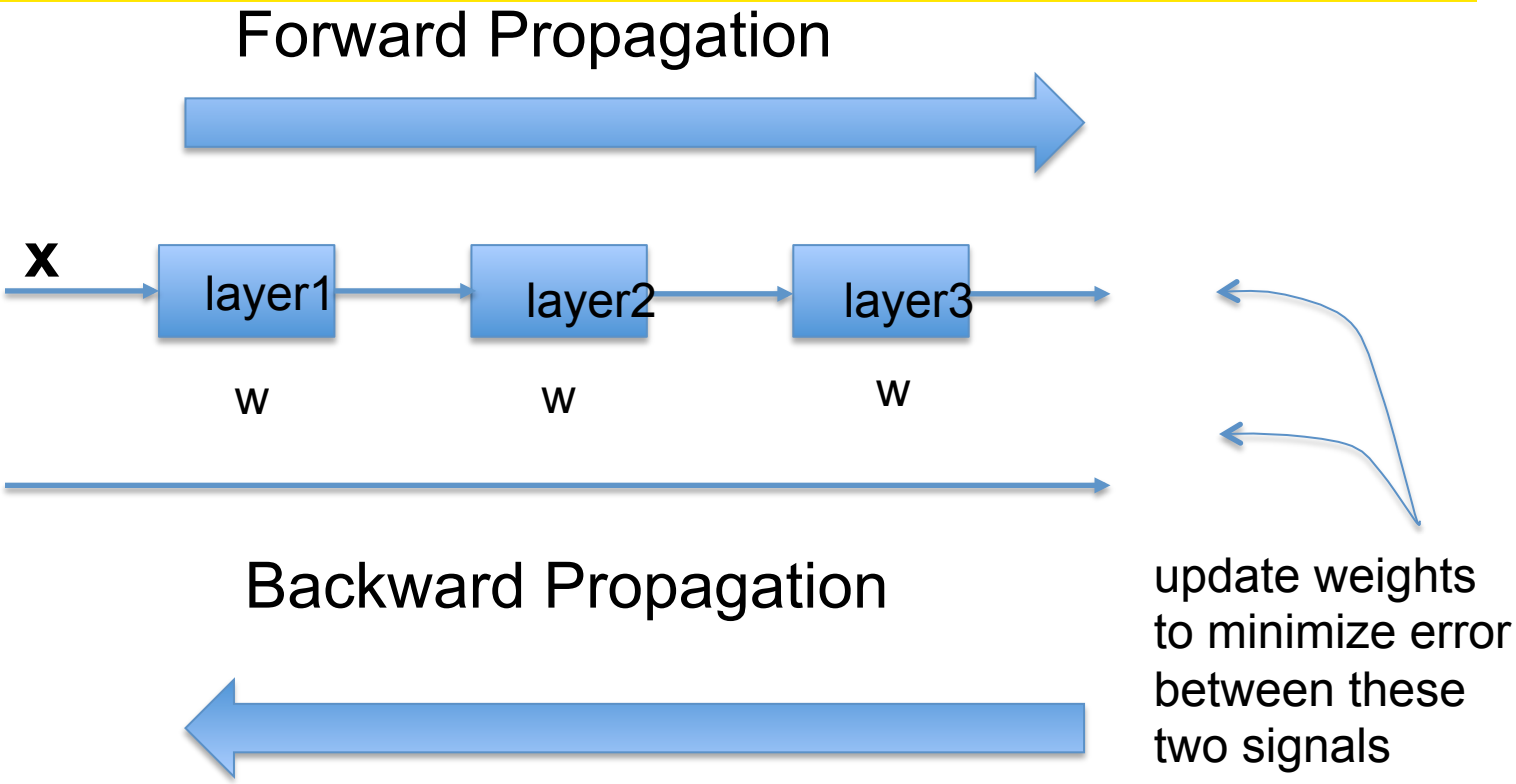
Biological Neuron



Mathematical Neuron



How do we train them?



Deep Neural Networks - CNN

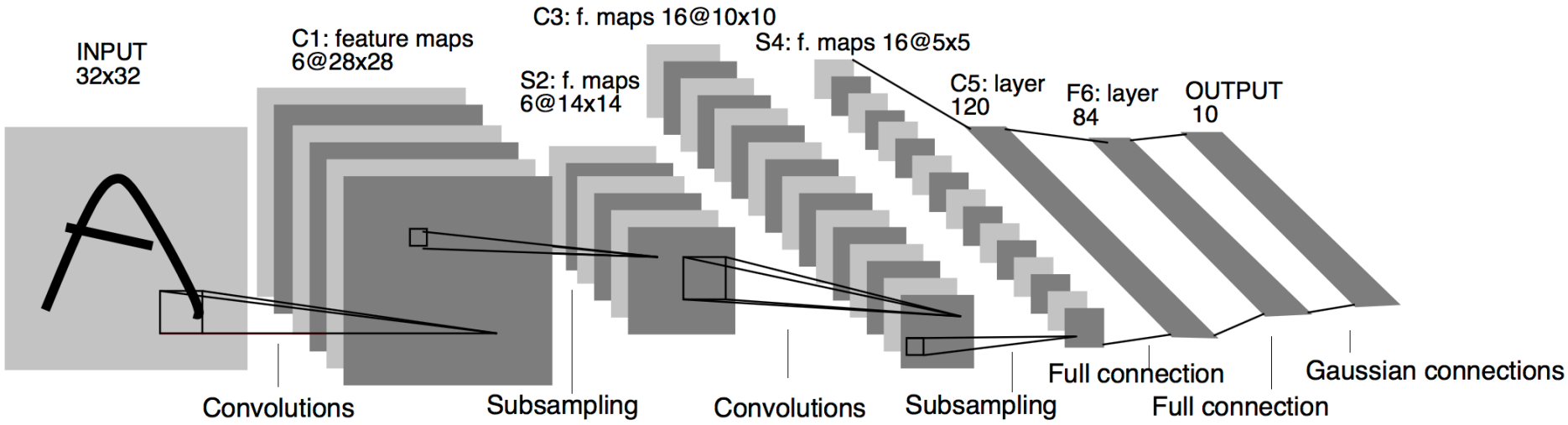
Input

CNN processing

Face Detectio

image from : https://facedetection.com/datasets/

CNN Architecture Example



~60k parameters, 1998
Today architectures use from 100k to >100M parameters

LeNet-5, LeCun 1998

So what is an output from NN?

- Depends on your task and model!
- For Image Classification it is a probability of a class label.
- ResNet50, ImageNet.
- 1000 classes
- Image resized to 224x224x3

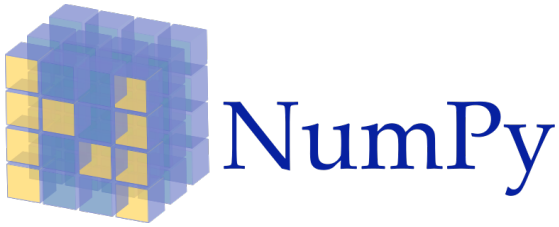
Why is Deep Learning on the rise

- A few years ago someone noticed that a single layer of the mathematical neuron could be described by a matrix multiplication applied to its input vector. Graphics chips (GPU) are really good at such matrix calculations.
- GPU supercomputers can run millions of photos through CNN and train the network (100M parameters!) in few hours.
- Once trained, the GPU-powered network can label an image in milliseconds.

Challenges of Deep Learning

- Size of training dataset enormous (10M photos)
  - How to label?
    - Manually?
    - Another machine?
  - Untrustworthy training datasets
- Bias in training data
- Interpretability --> Explainable AI

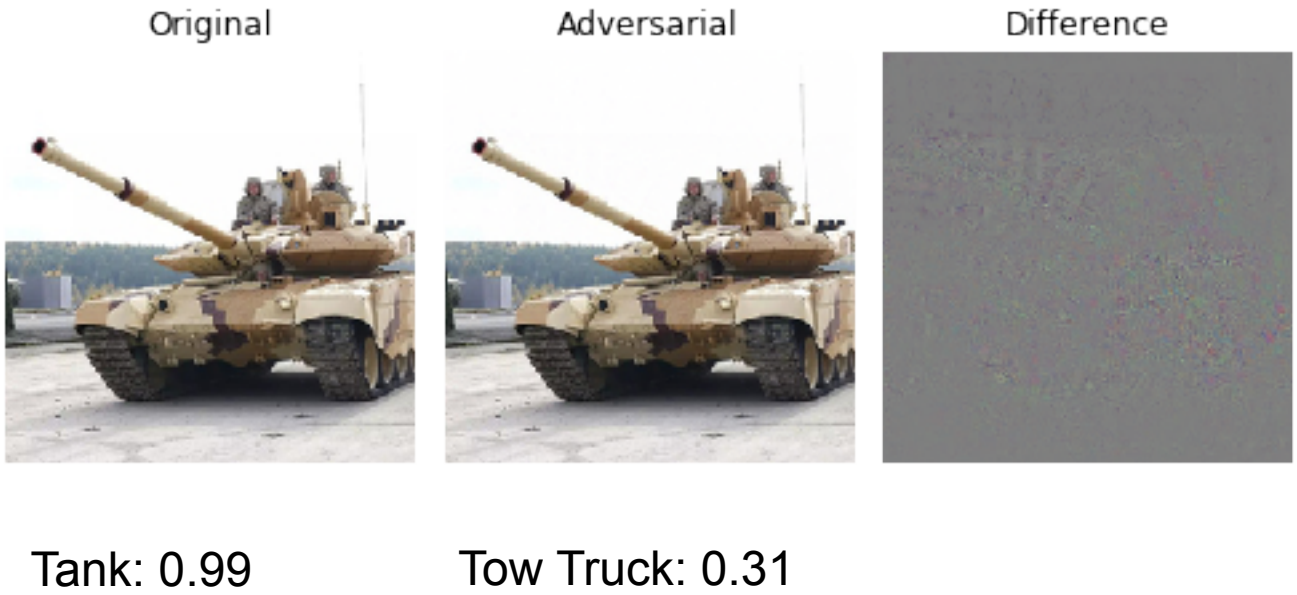
Deep Learning Technologies



AND MANY OTHERS!

Adversarial Examples

- Adversary created small perturbations on inputs can degrade performance of Neural Networks
- In inference
- In training



Physical Adversarial Attacks

Evtimov et al., 2018

Sharif et al. 2019

AI Everywhere

- Speech Recognition
- Image recognition
- Object localization (where things are in the image, tracking)
- Gesture Recognition
- Optical character recognition
- Translation
- Text Classification
- Key Word Spotting (OK Google)
- Activity recognition from sensors (IMU units on cellphones)