Deep Learning Applications for Cyber Security

Rekha R - Researcher SETS

Agenda

01 Why AI to Cyber Security?

02 Introduction to Deep Learning and Overview of algorithms

03 Cyber Security Problems

04 DL Solutions

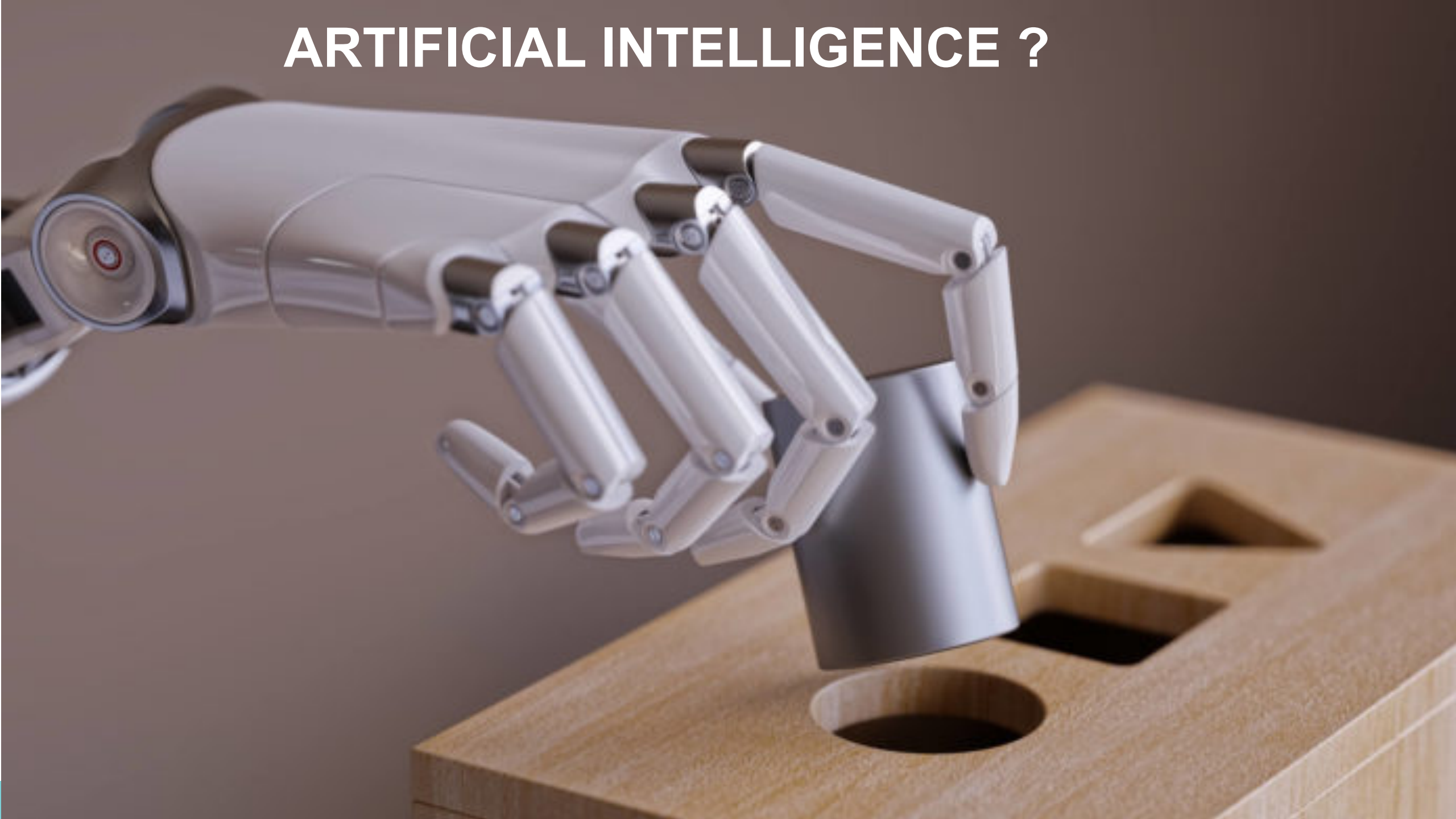
What is Cyber Security?

- *Many Definitions !!!*

- ***Security of** :*
  - *Systems, Devices, Networks, Infrastructure*
  - *Can extend to information, Data*
- ***CyberSecurity***
  - *- The ability to protect or defend the use of cyberspace from cyber attacks*

- ***Cyberspace -** A global domain within the information environment consisting of interdependent network of information systems infrastructures including the internet, telecommunications network, computer systems and embedded processors and controllers*

ARTIFICIAL INTELLIGENCE ?



ARTIFICIAL INTELLIGENCE ?

Any techniques that enable machines to solve task in way like humans do

AI is in Charge

Energy: Nuclear Power Plants

Utilities: Water Plants/ Electrical Grid

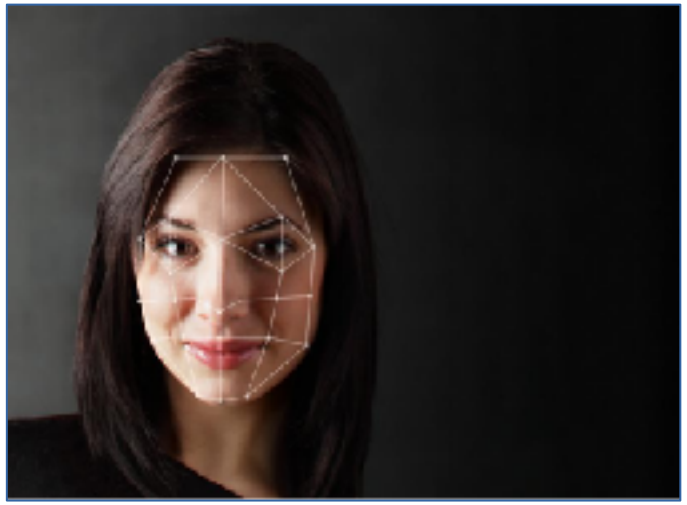
Military: Nuclear Weapons

Communications: Satellites

Stock Market: 75+% of all trade orders generated by Automated Trading Systems

Aviation: Uninterruptible Autopilot System

Few More Domains



Why AI in Cyber Security?

What Challenges we face today ?

INSTRUMENTED & INTERCONNECTED WORLD


COMPLEX ORGANIZATIONS

DEMANDING CITIZENS

DIVERSE, EVOLVING AND SOPHISTICATED THREAT

Sophisticated Malware Spreading

New infection every 3 seconds

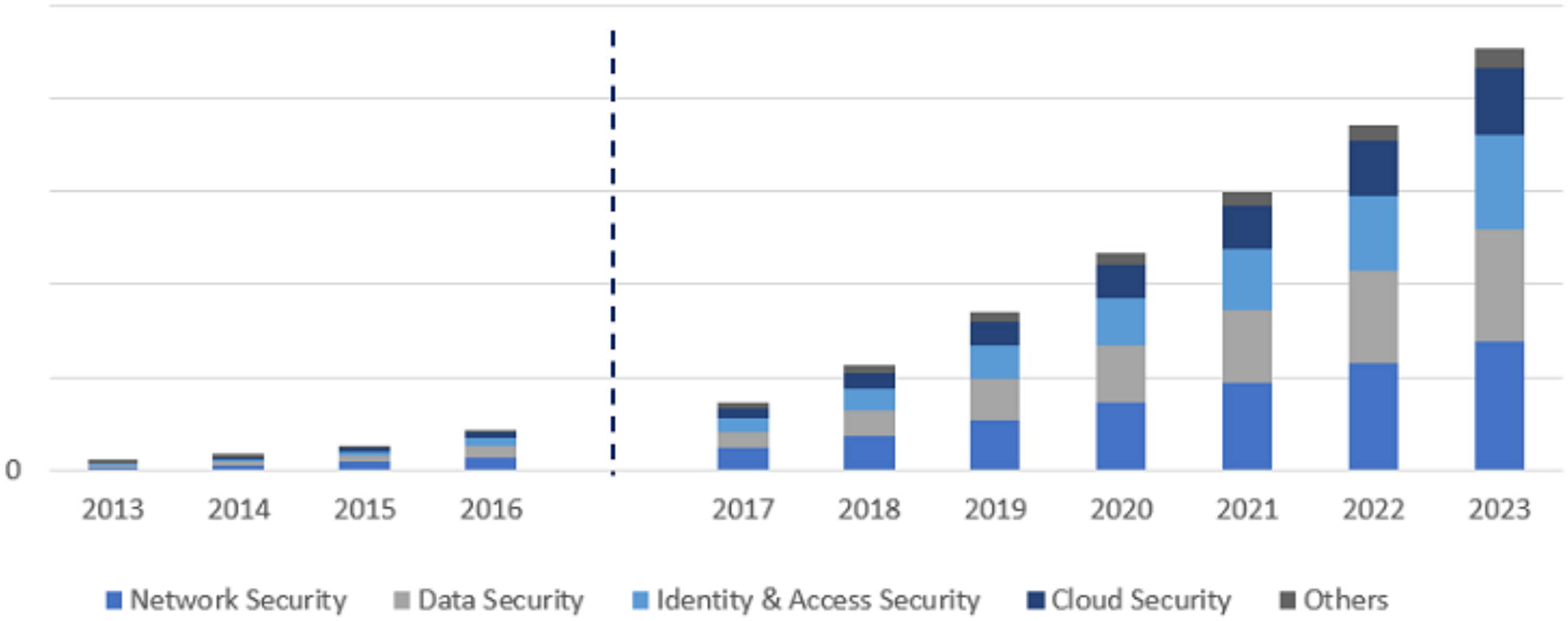
After….

1 minute = 2,021 instances

15 minutes = 9,864 instances

30 minutes = 45,457 instances

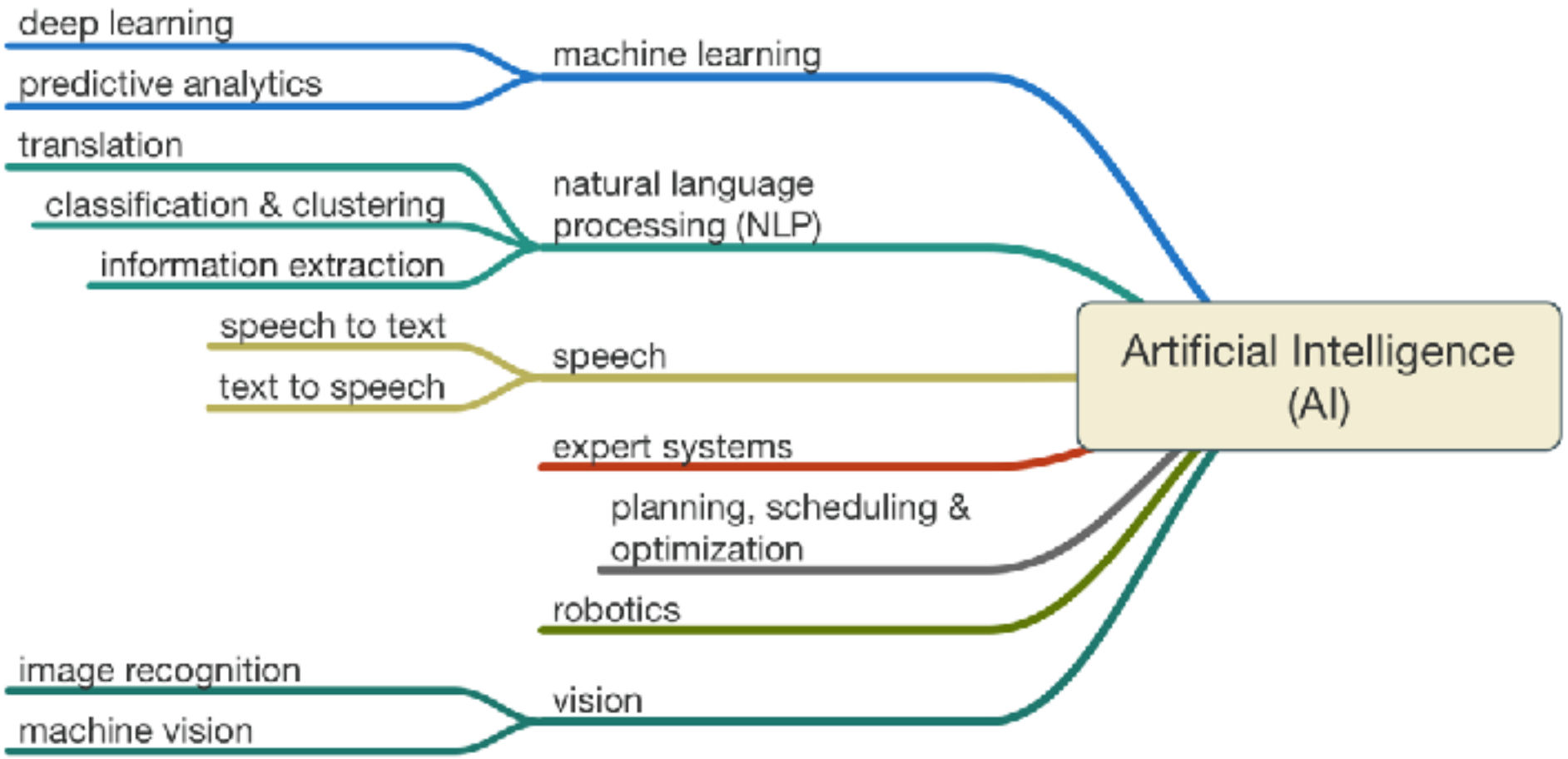
Cyber Security – Future Predictions



HIGHLY AUTOMATED ADVERSARIES

CHANGE CYBER SECURITY

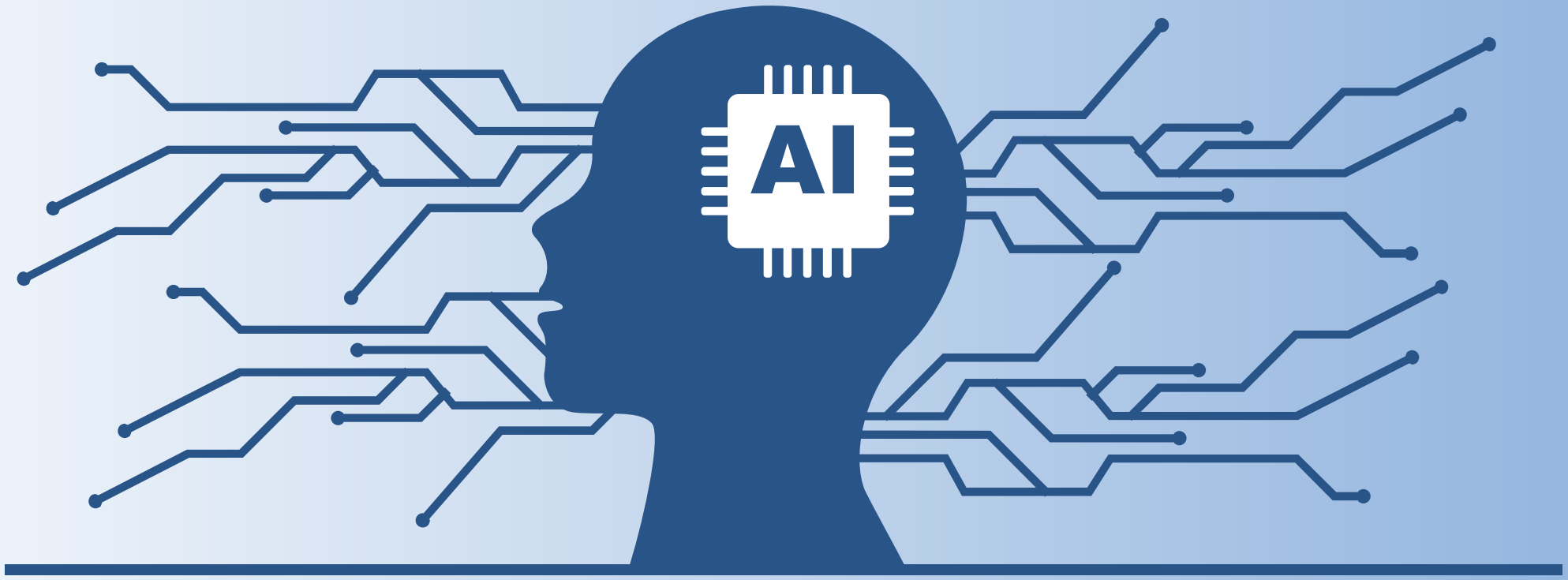
AI – Broad Domain



How AI models learn and understand

What is **Normal** ?

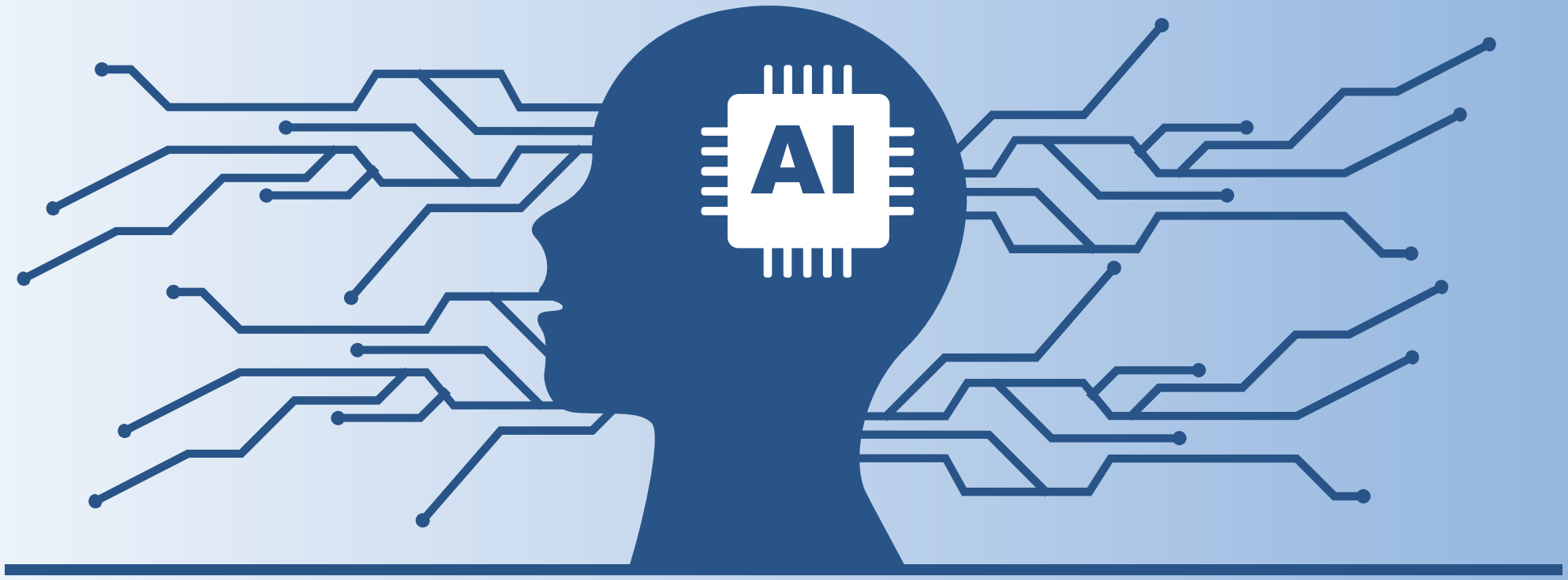
What is **Abnormal** ?



How AI models learn and understand

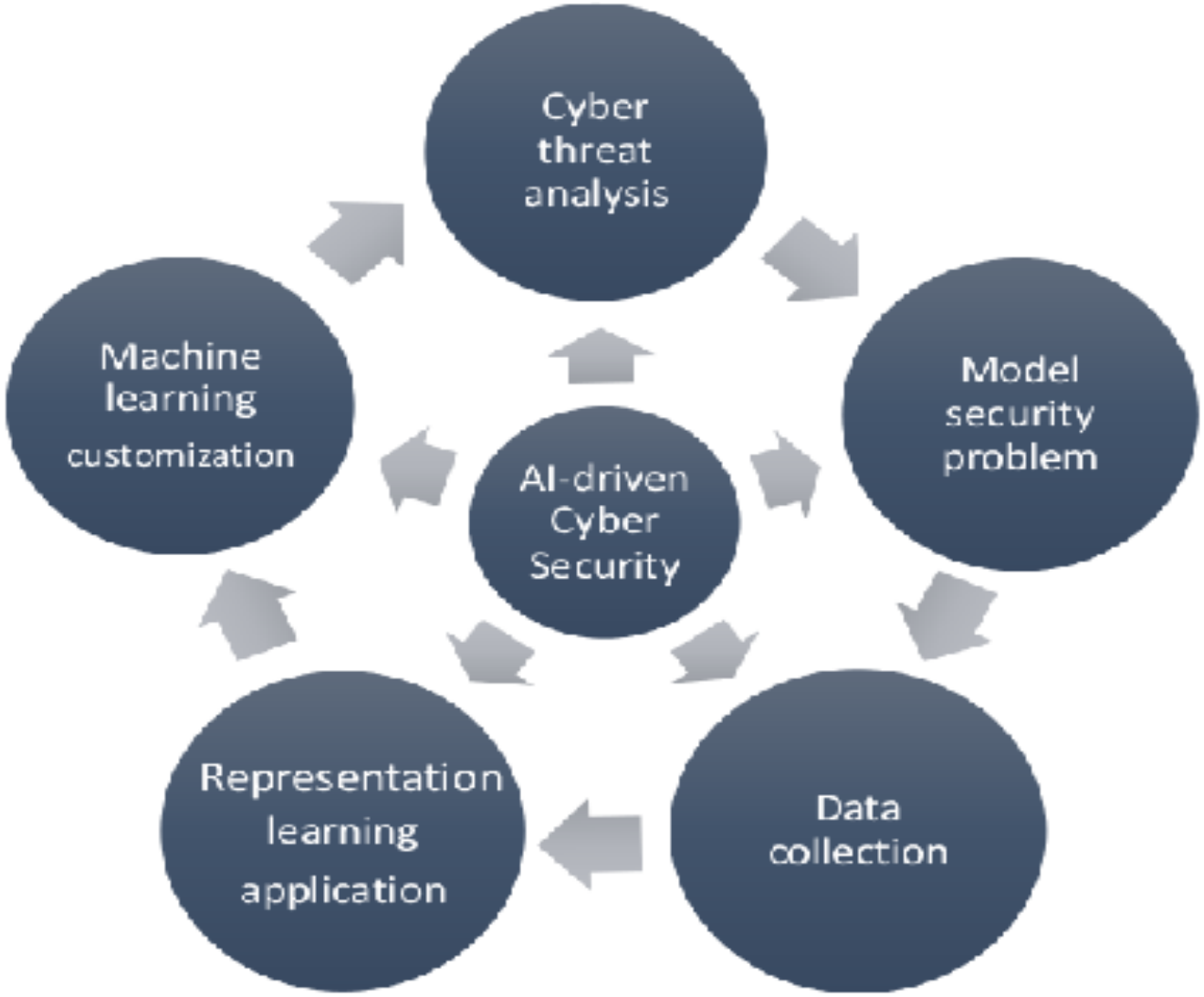
What is **Normal** ?

What is **Abnormal** ?



Secured Data Modelling & Reasoning

Research Methodology



What is AI, ML & DL ?

Artificial Intelligence

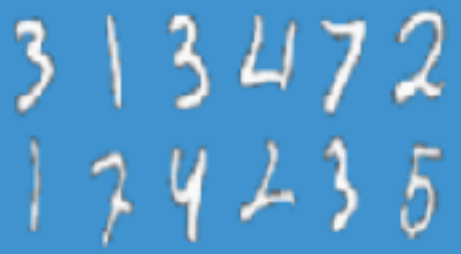
Any technique that enables computers to mimic human behavior

Machine Learning

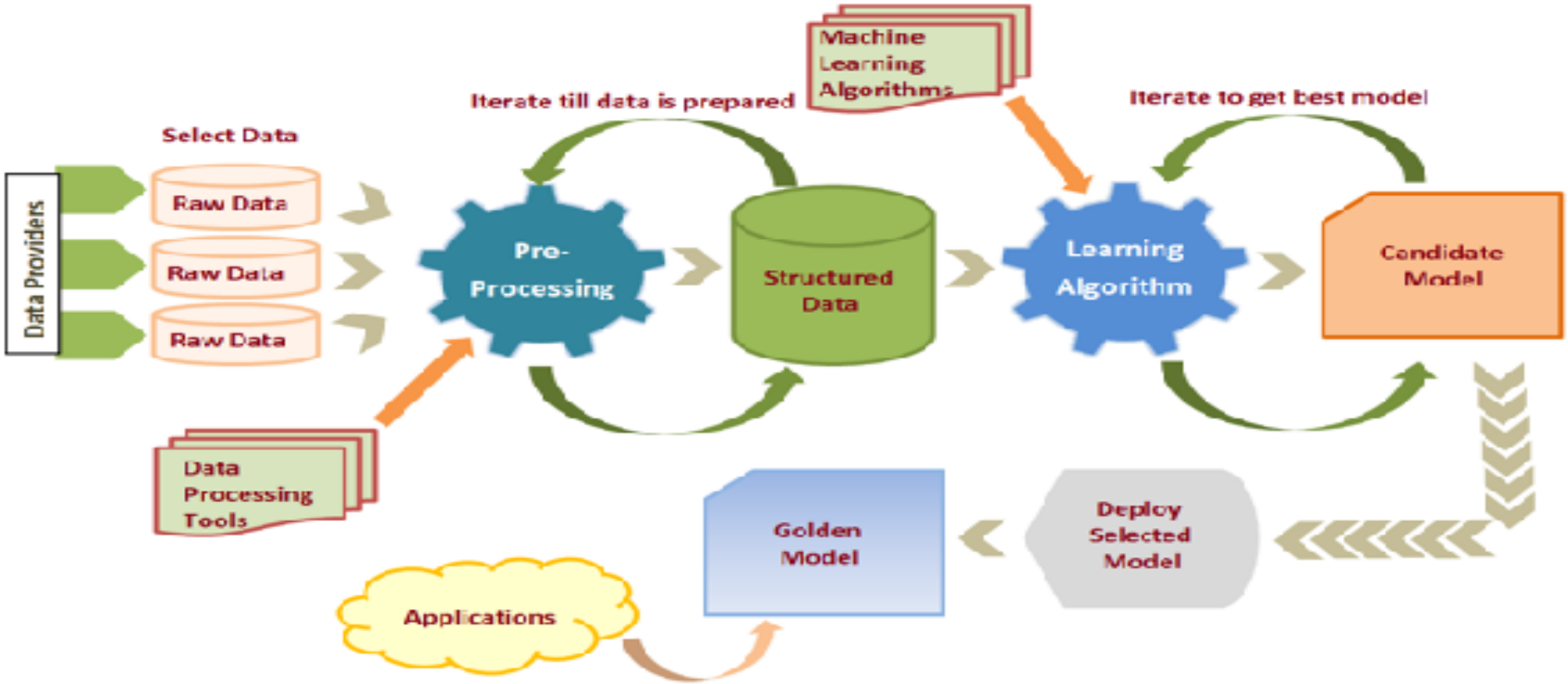
Ability to learn without explicitly being programmed

Deep Learning

Extract patterns from data using neural networks



ML Model Creation



Types of Machine Learning

Supervised

Learning with a labelled training set

Example:
Email classification with already labeled emails

Unsupervised

Discover patterns in unlabelled data

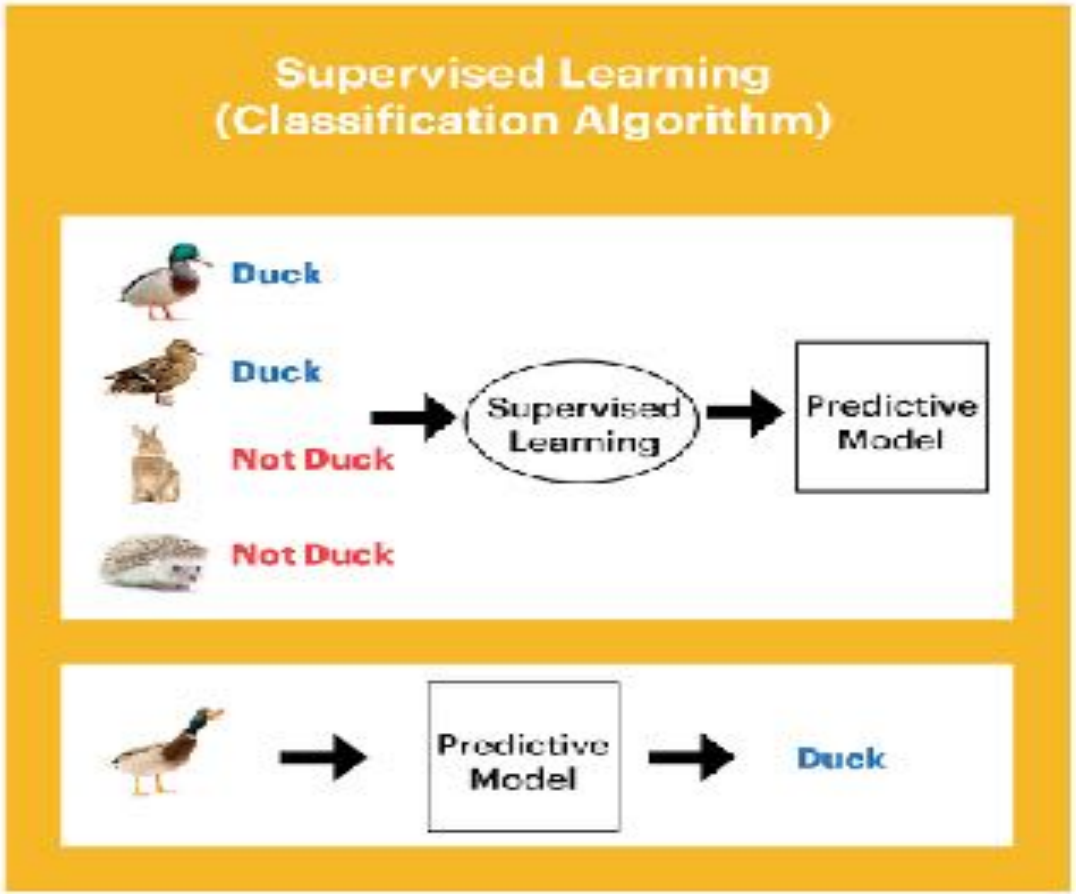
Example:
Cluster similar documents based on text

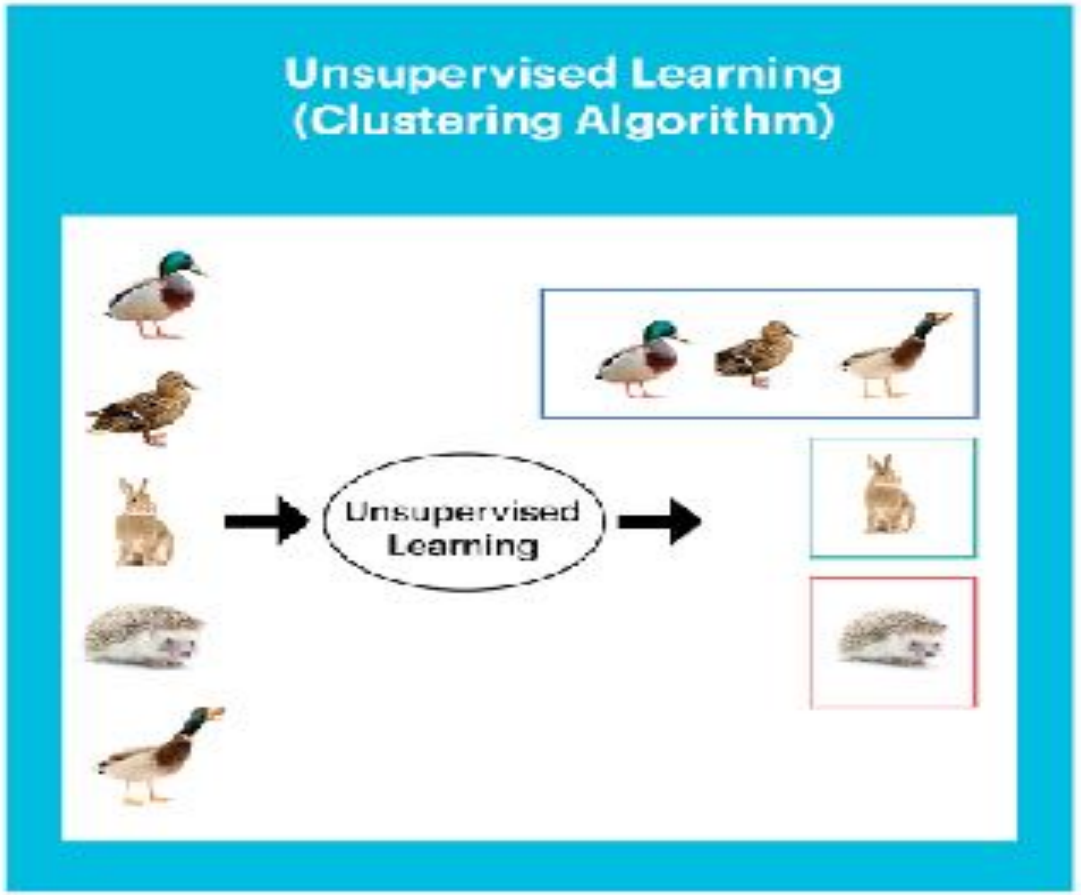
Reinforcement

Learn to act based on feedback/reward

Example:
Learn to play Go, reward:
win or lose

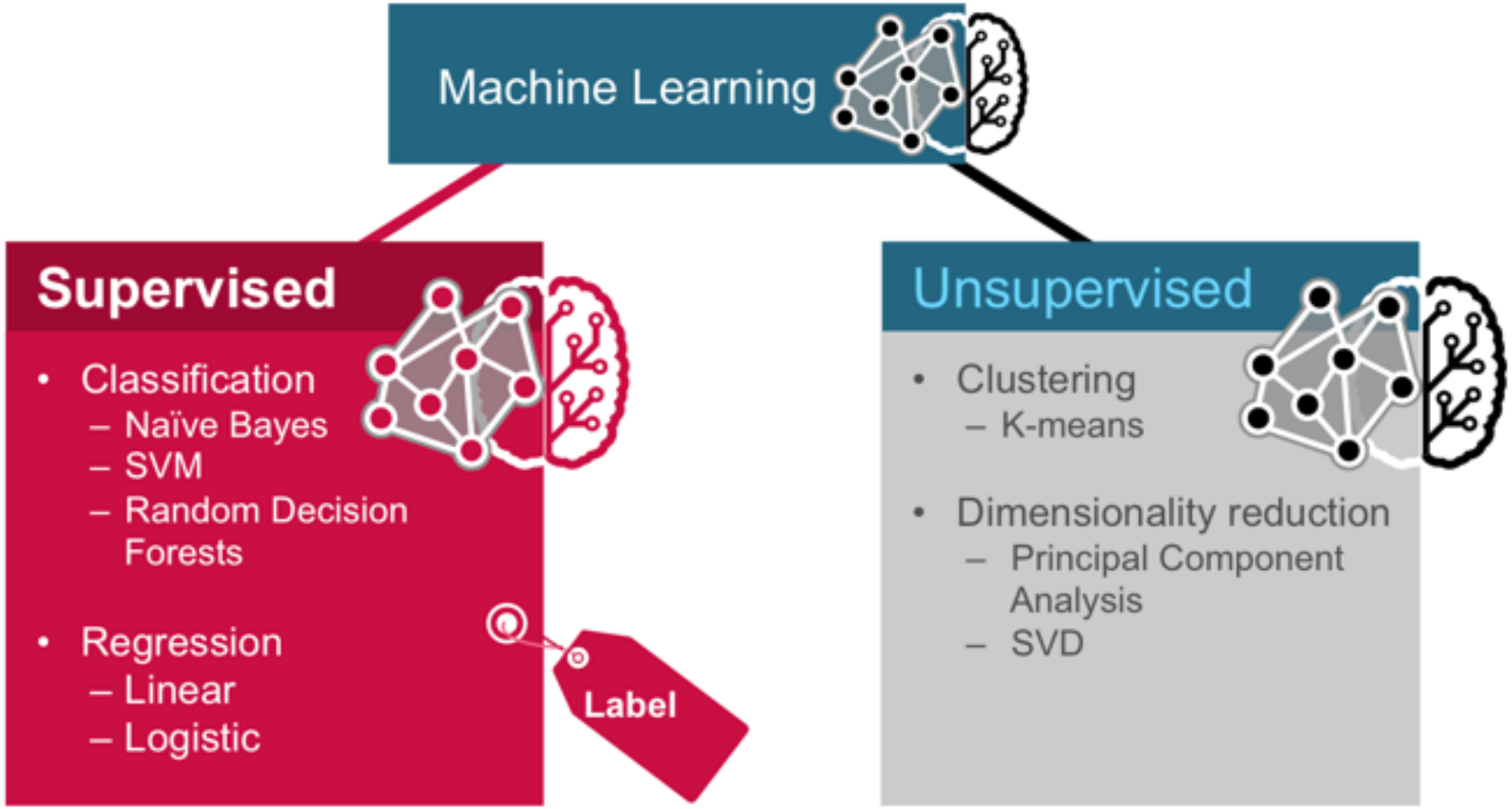
A Simple Example



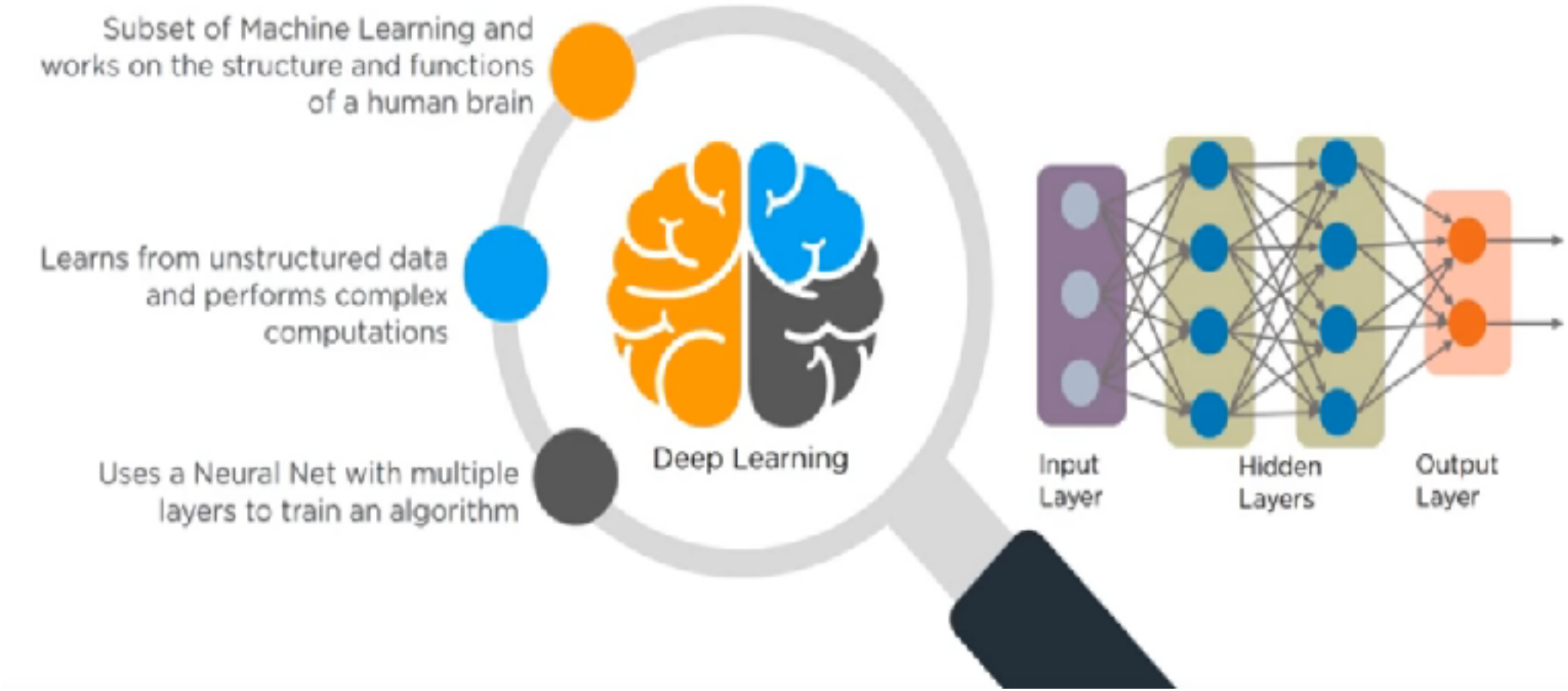


Western Digital.

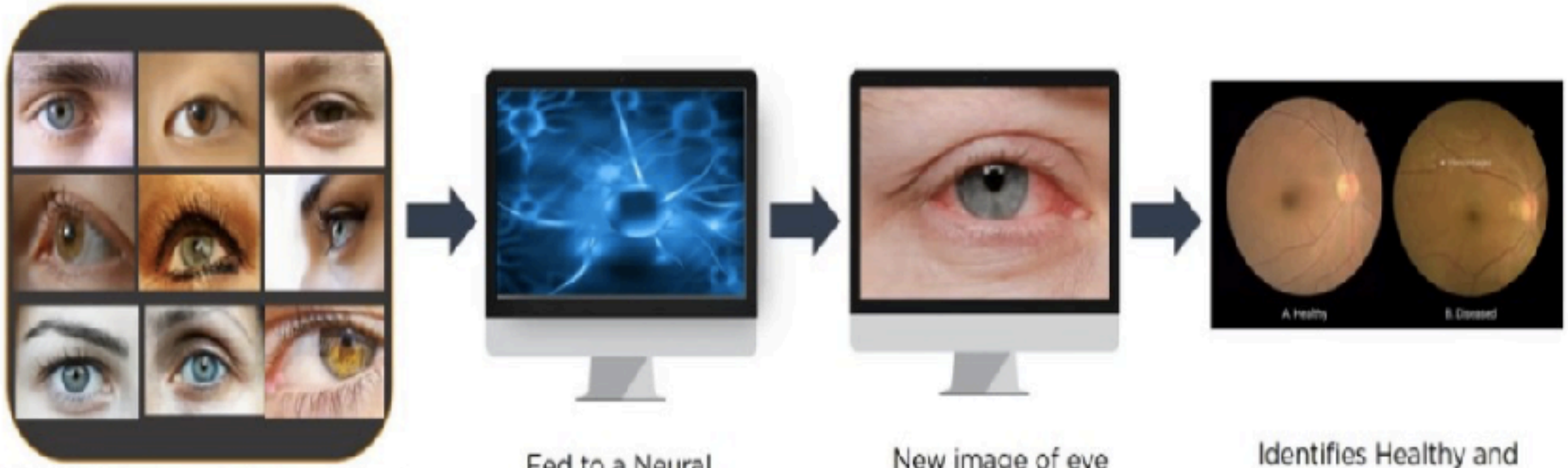
Learning Methods – Few Algorithms



What is Deep Learning ?

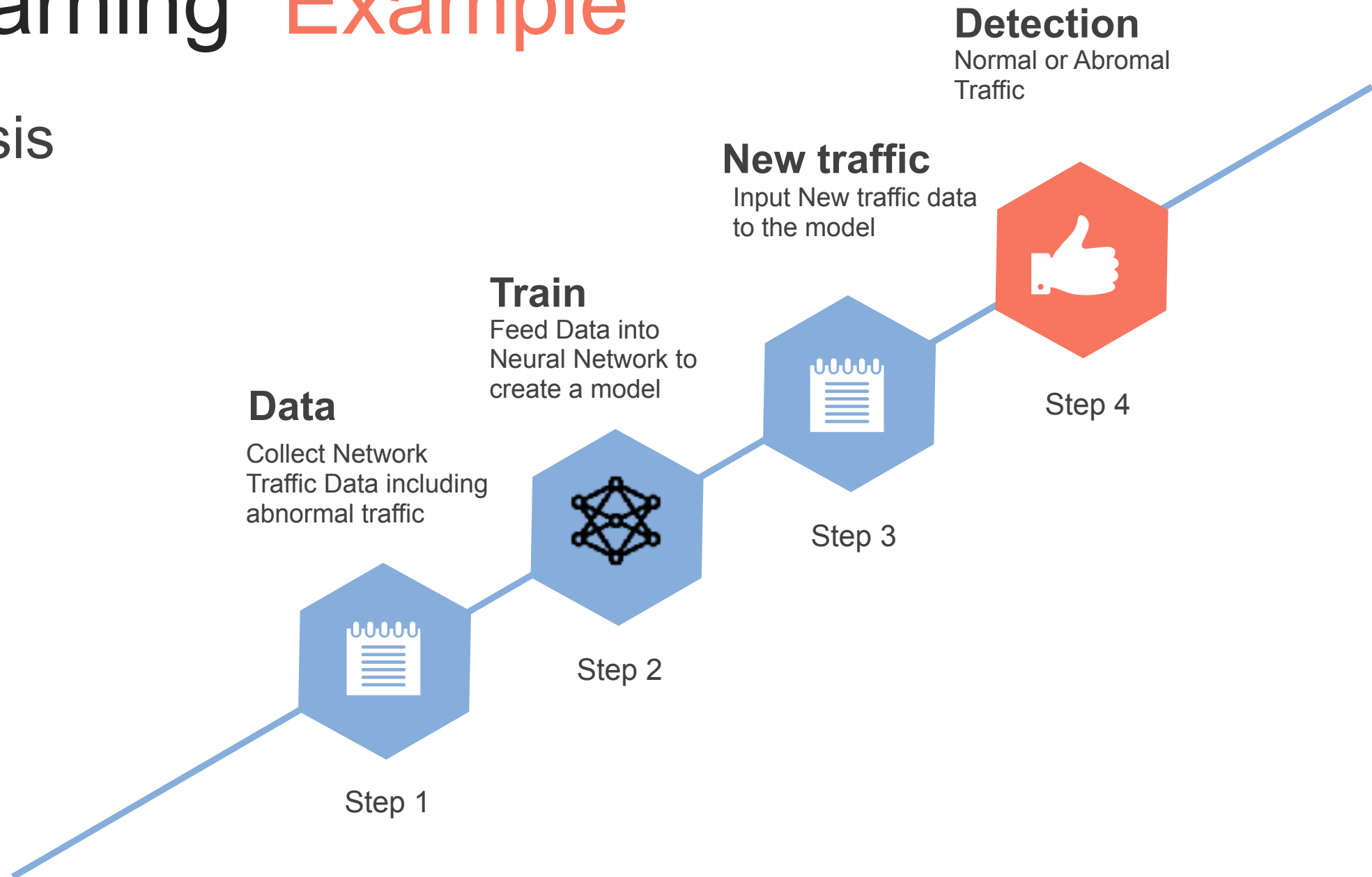


Deep Learning - Real time Example

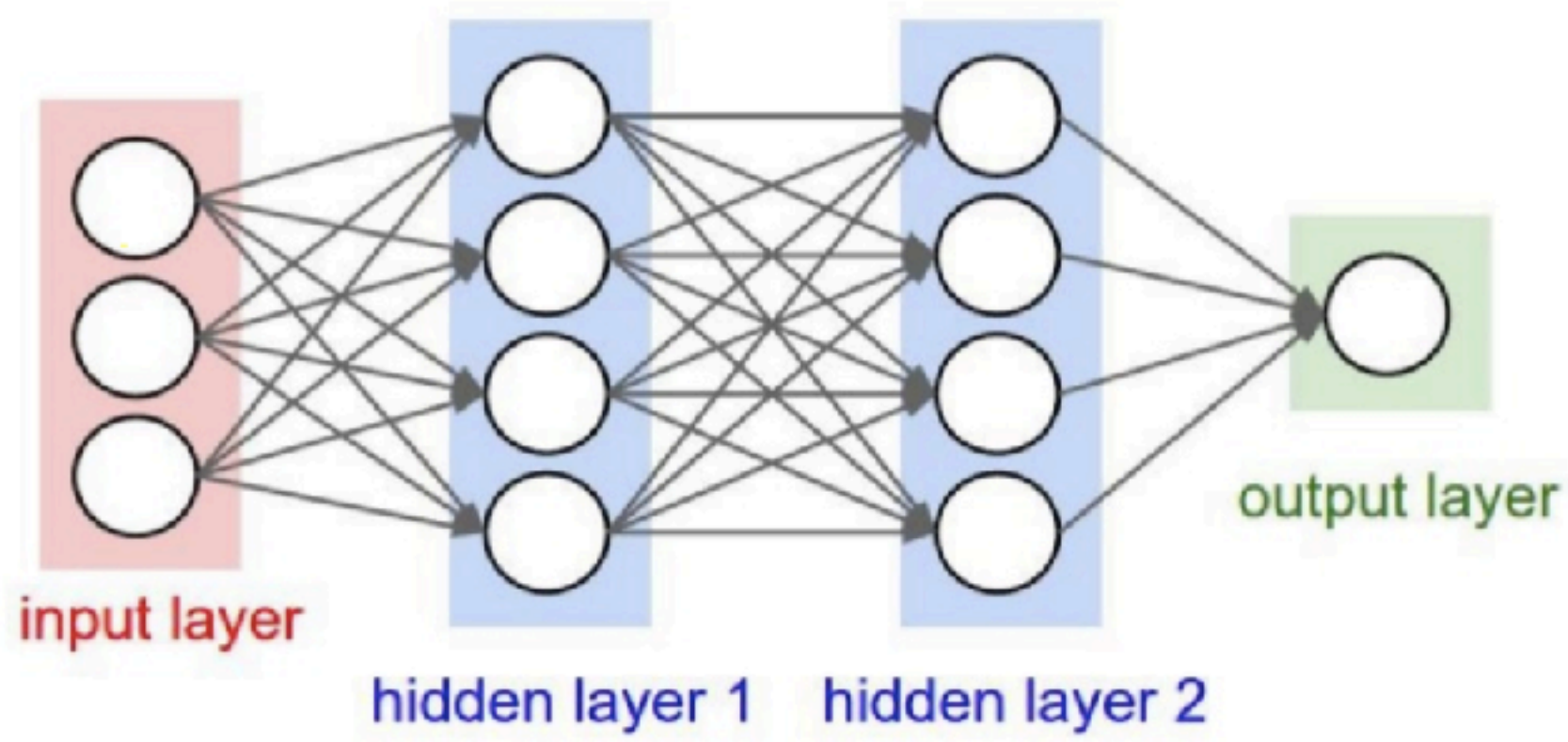


Deep Learning Example

Traffic Analysis

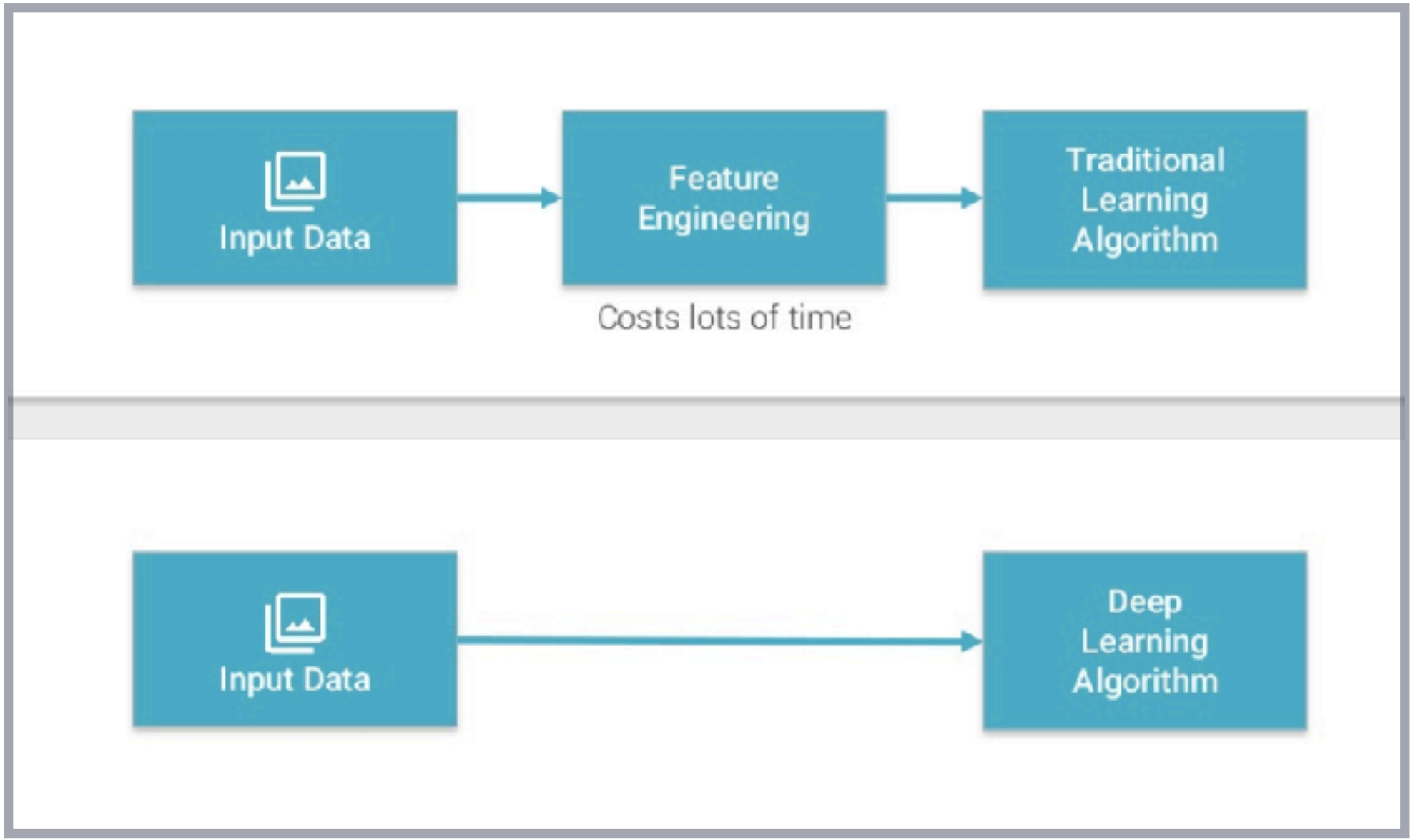


Deep Learning - Artificial Neural Network

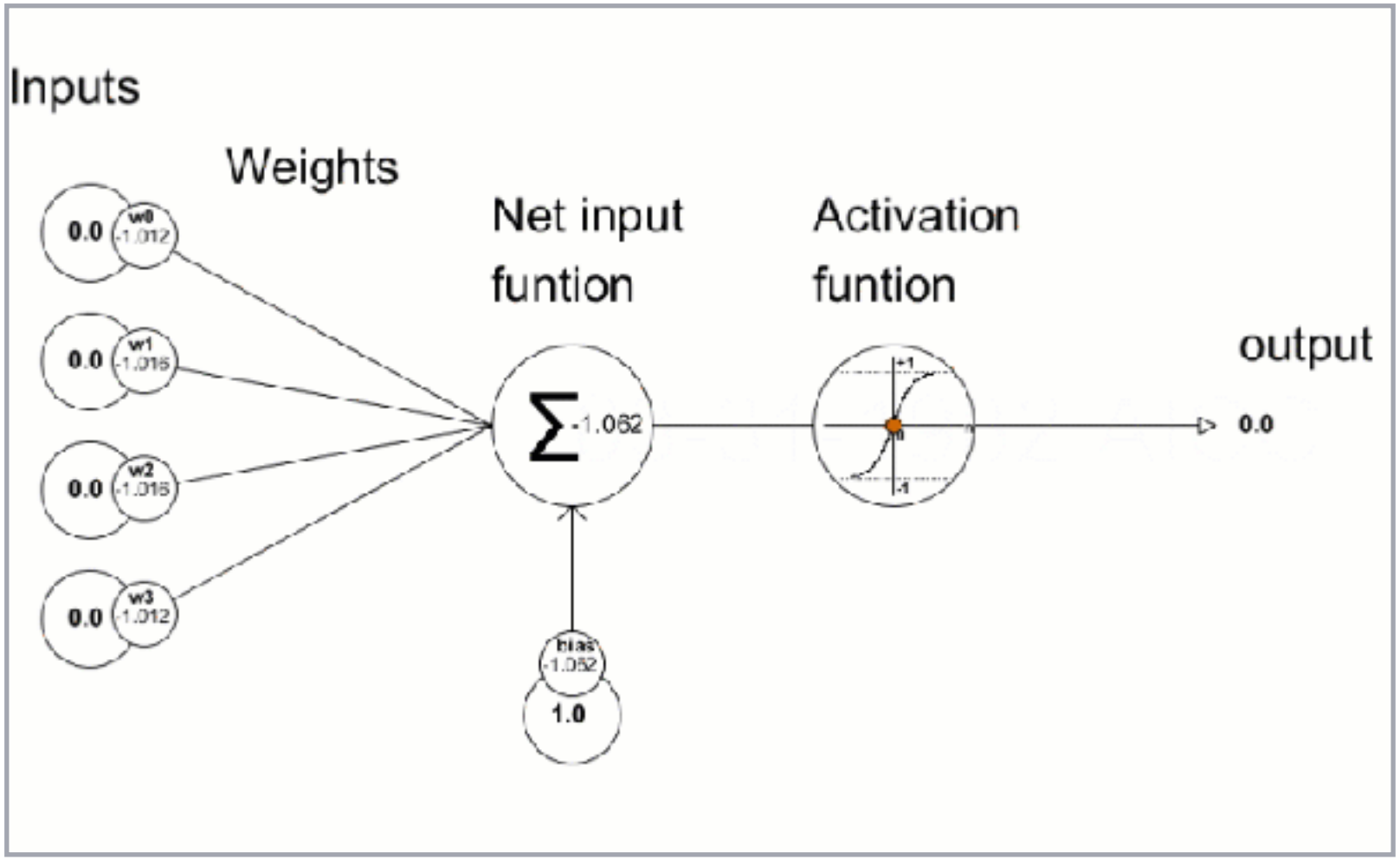


Consists of one input, one output and multiple fully-connected hidden layers in-between. Each layer is represented as a series of neurons and progressively extracts higher and higher-level features of the input until the final layer essentially makes a decision about what the input shows. The more layers the network has, the higher-level features it will learn.

No more Feature Engineering



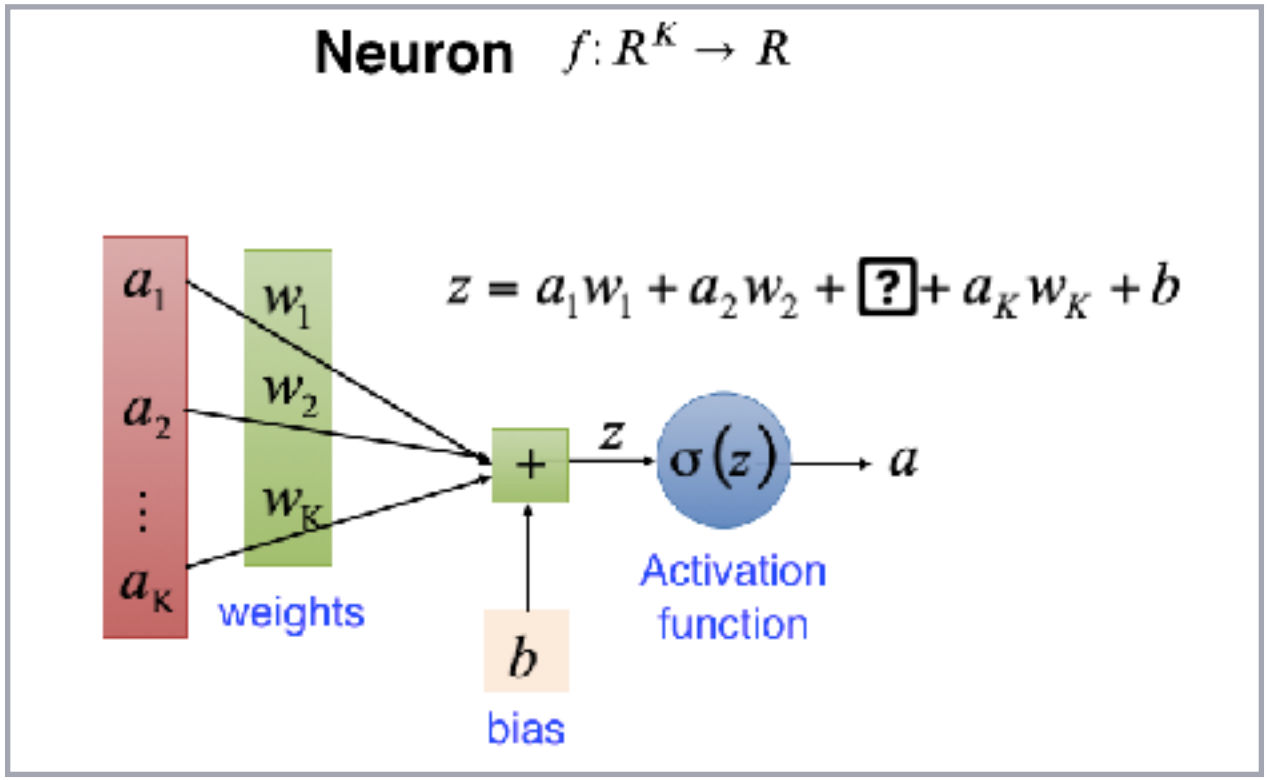
What happens in a Neuron ?



Neuron Representation

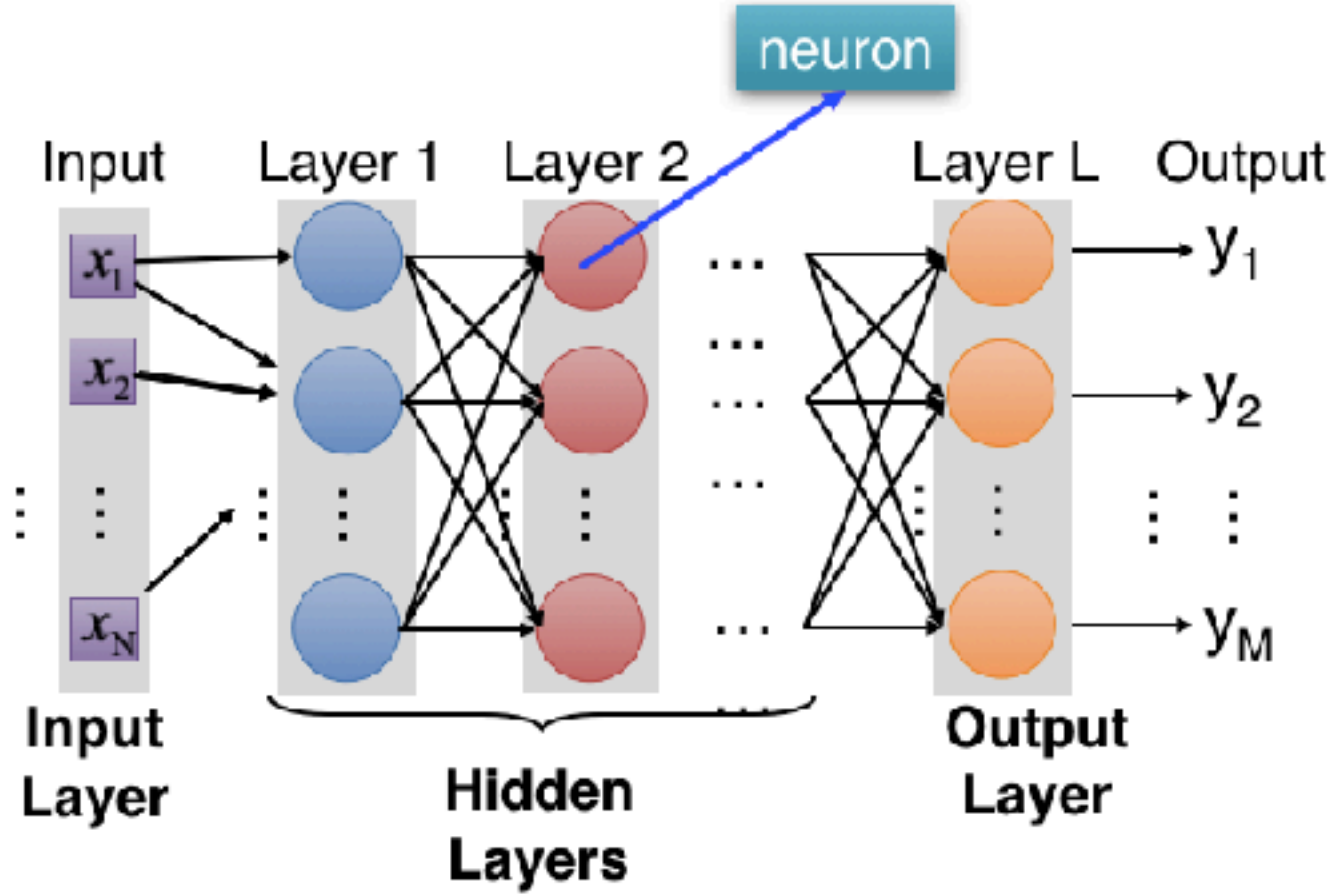
Weights *show the strength of the particular node/input. Initialise the weights randomly and update weights with model training.*

Bias *allows you to shift the activation function by adding a constant (i.e. the given bias) to the input.*



Activation function *convert a input signal of a node in A-NN to output signal. In short, used to map the input to (0, 1) or (-1, 1)*
Ex: ReLu, Leaky/Randomized ReLu, Softmax (Output probabilities) , Sigmoid, tanh

Deep Neural Network



Deep means many hidden layers

Training the Network

1. *Randomly **initialise** the network weights and biases*
2. ***Training**:*
   - *Get a ton of training data (e.g. Network Traffic)*
   - *For every piece of training data(bytes/pixel/words/wave forms- which has little meaning), feed it into the network*
3. ***Testing**:*
   - *Check whether the network gets it right*
   - *If not, how wrong was it? Or, how right was it? (What probability did it assign to guess?)*
4. ***Tuning /Improving**: Nudge the weights a little to increase the probability of the network more probability getting the answer right.*

Machine Learning Vs Deep Learning

Factors	Deep Learning	Machine Learning
Data Requirement	Requires large data	Can train on lesser data
Accuracy	Provides high accuracy	Gives lesser accuracy
Training Time	Takes longer to train	Takes less time to train
Hardware Dependency	Requires GPU to train properly	Trains on CPU
Hyperparameter Tuning	Can be tuned in various different ways.	Limited tuning capabilities

Percentage of Content Analysis and Correlation differs

Deep Learning Approaches

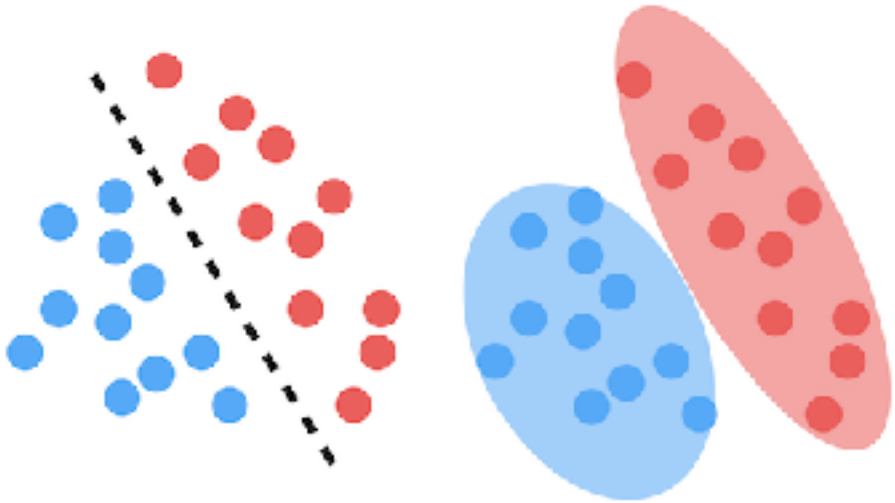
Discriminative Model

- Model P(y|x)
- Learn the **boundary** between classes
- Usually better performance
- E.g. logistic regression, SVM

Generative Model

- Model P(x,y)
- Model the **distribution** of individual classes
- Can "generate" synthetic data points
- E.g. Hidden Markov Model

Discriminative **Generative**



Deep Learning Approaches

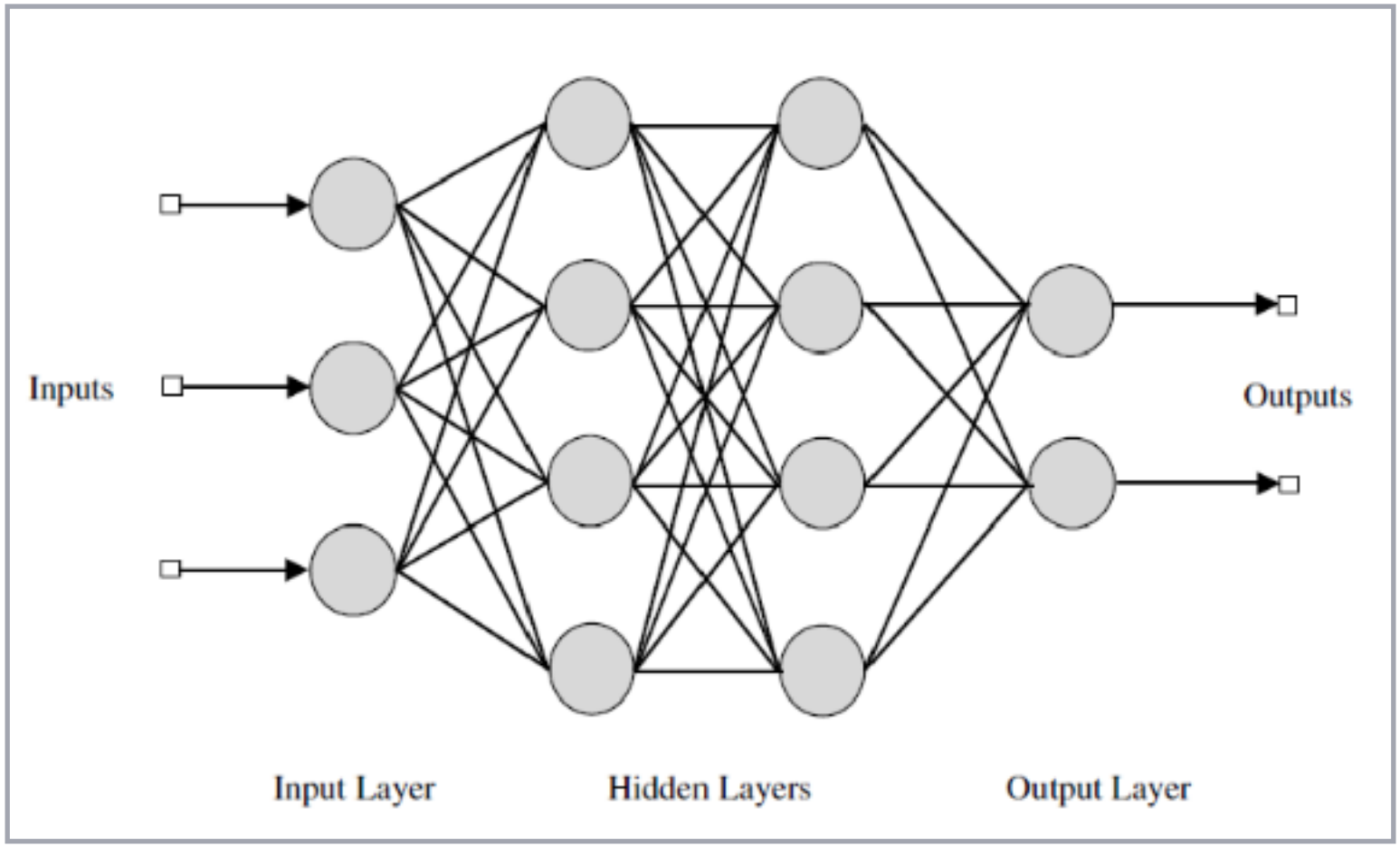
Deep Discriminative Model - Supervised

- *Convolutional Neural Network - Classification*
- *Recurrent Neural Network - Regression (LSTM)*
- *Deep Neural Network - Regression/Classification*

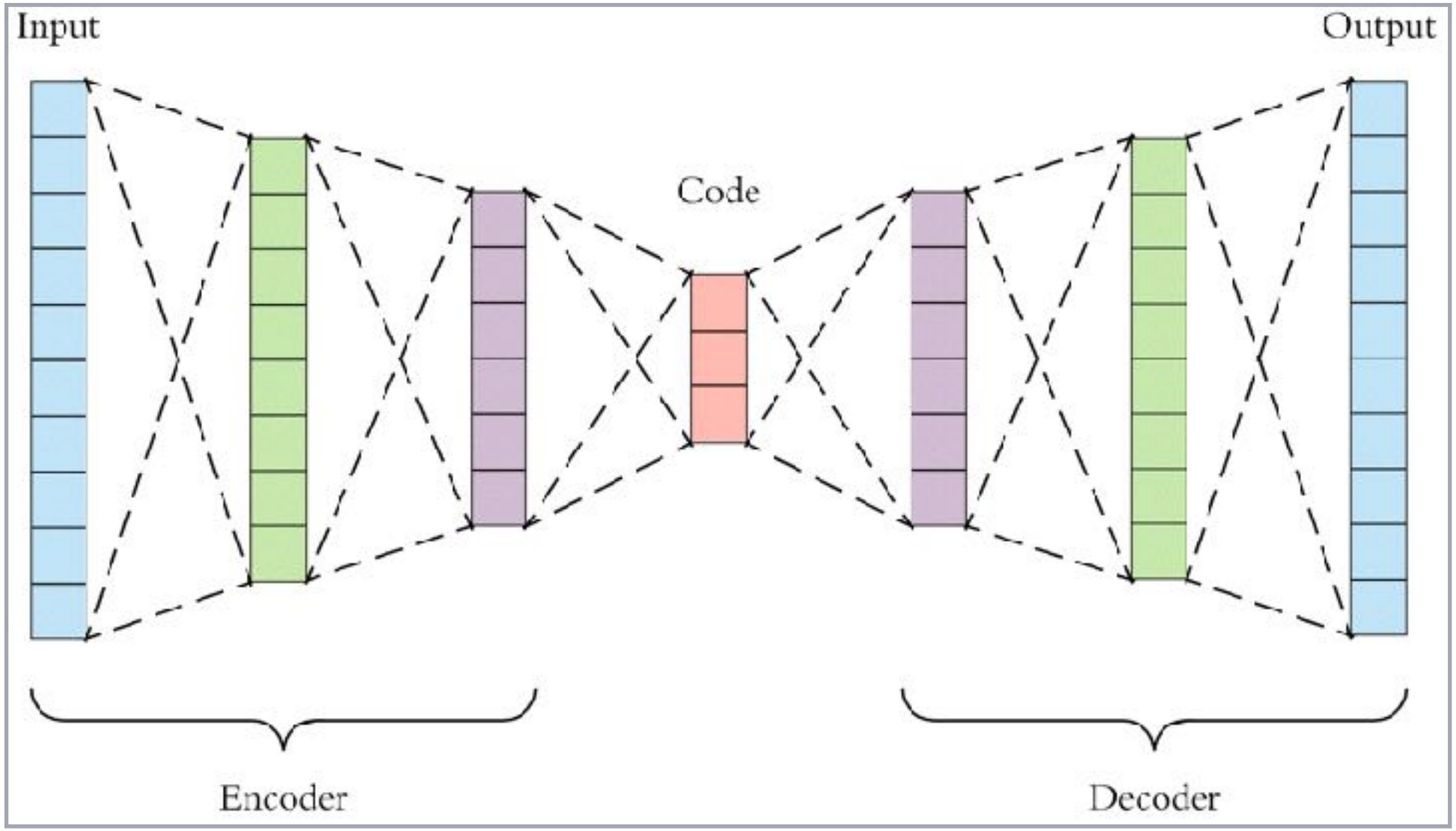
Generative Model - Unsupervised

- *Auto Encoders - Association*
- *Restricted Boltzmann Machines - Association*
- *Deep Belief Networks - Association*
- *Self Organised Map / Kohenen - Clustering*
- *Generative Adversarial Networks (GANS)*

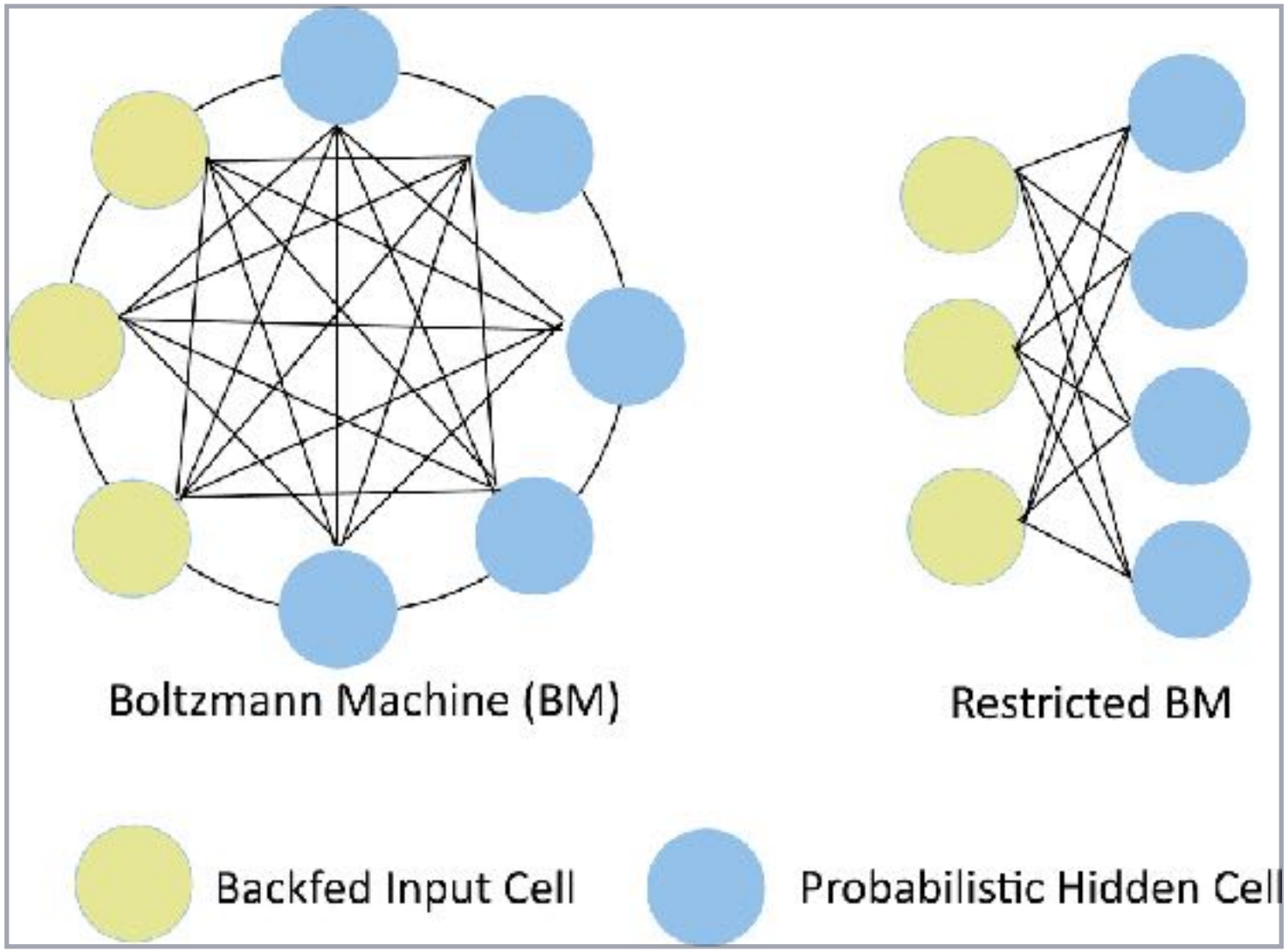
DNN - Feed Forward Neural Network



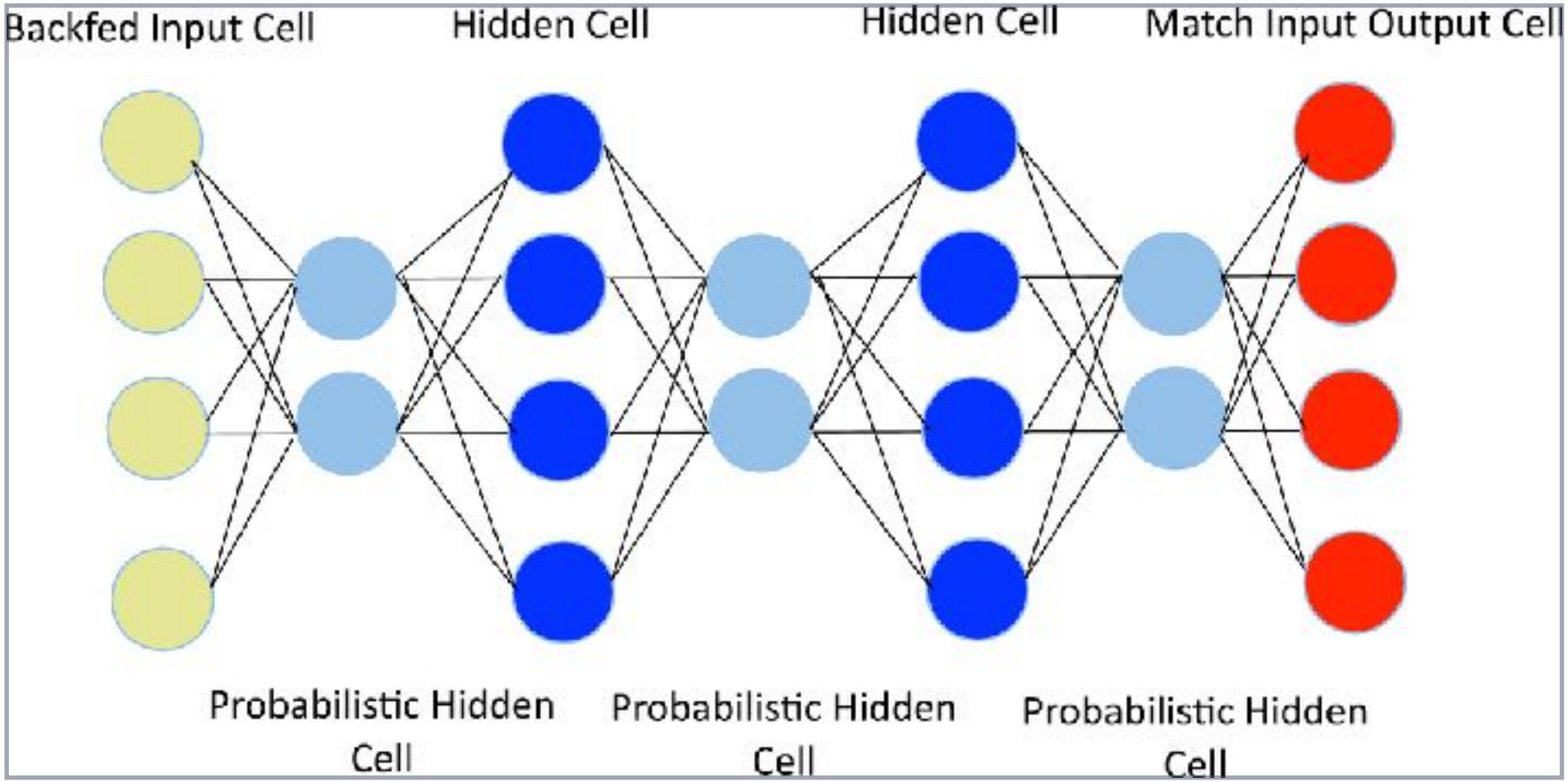
AutoEncoders



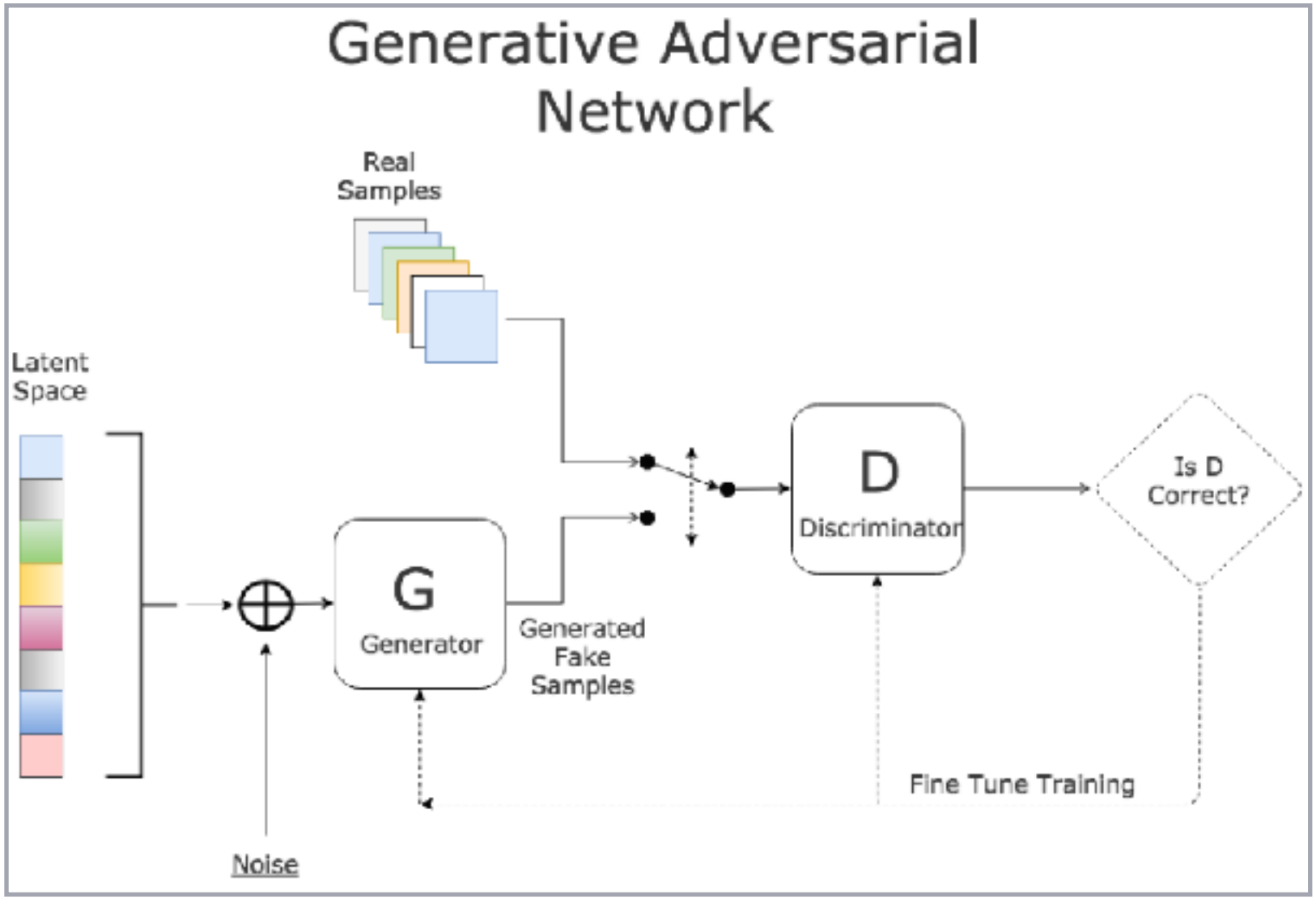
Boltzmann Machines



Deep Belief Network



Generative Adversarial Network



Deep Learning - Pros & Cons

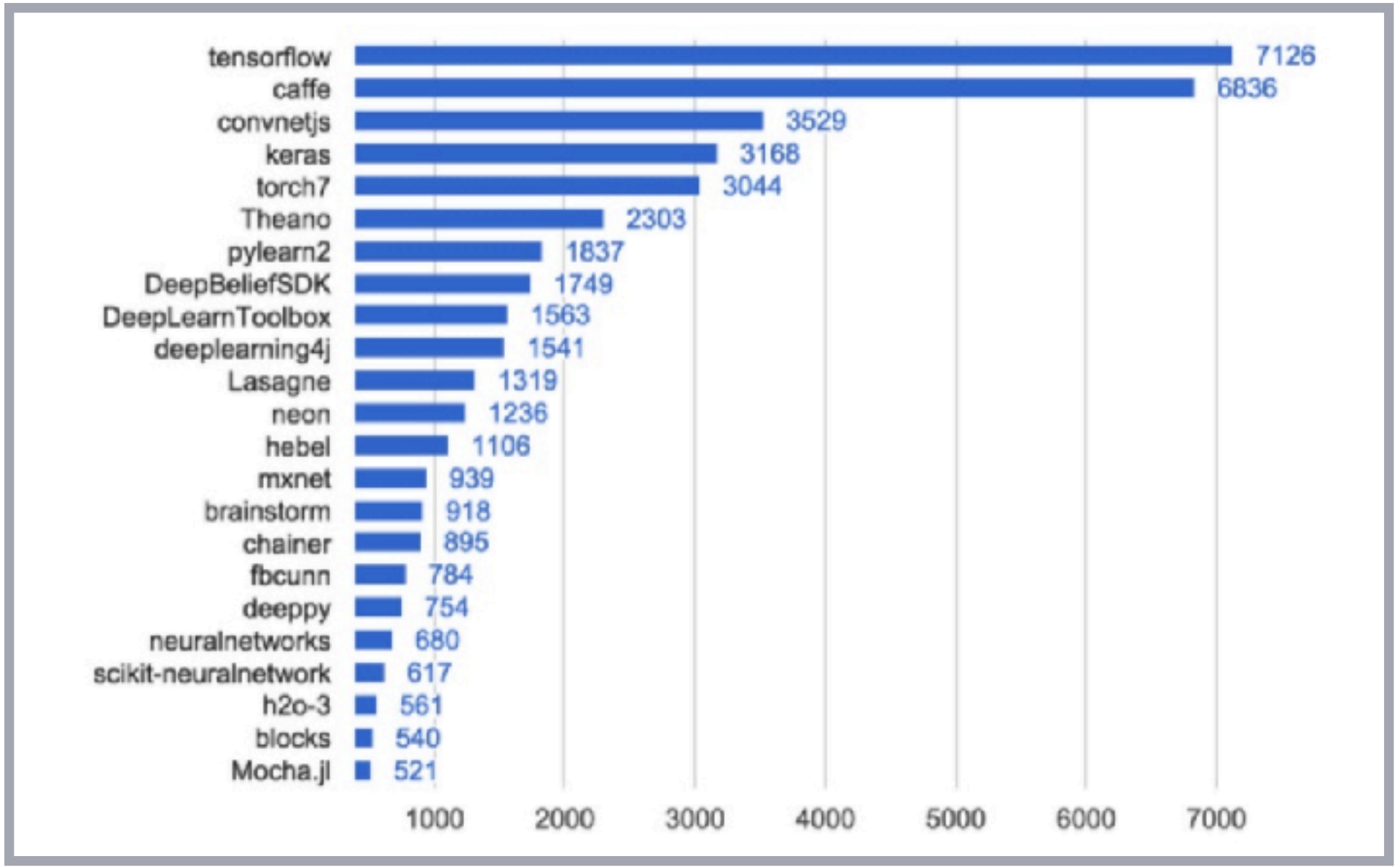
Pros

- *Best in class performance*
- *Reduce the need for feature engineering*
- *Is an architecture that can be adapted to new problems relatively easily*

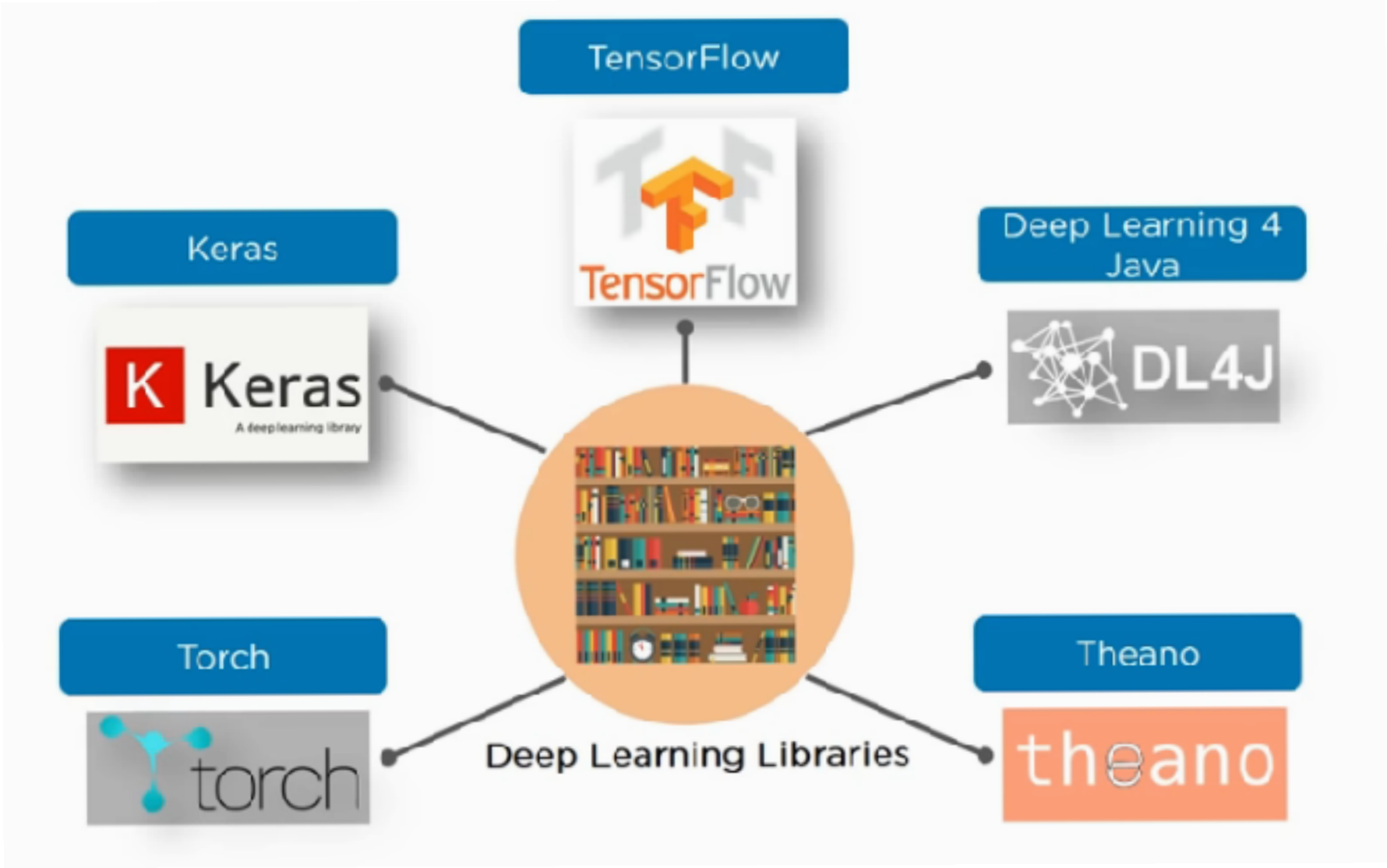
Cons

- *Requires a large amount of data*
- *Is extremely computationally expensive to train [GPUs, TPUs]*
- *What is learned is not easy to comprehend*

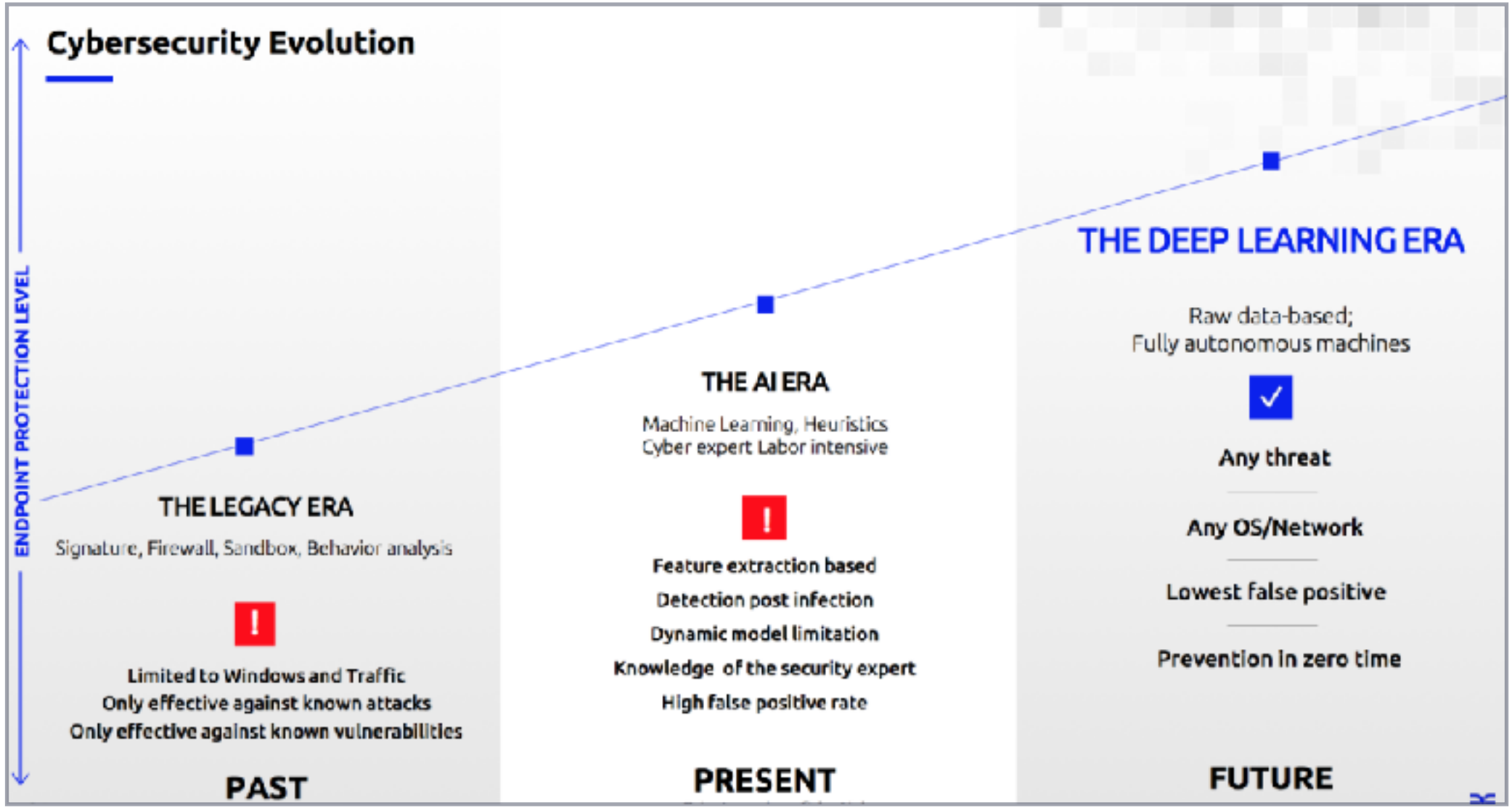
Deep Learning Tools - Its all open source



Popular libraries for DL



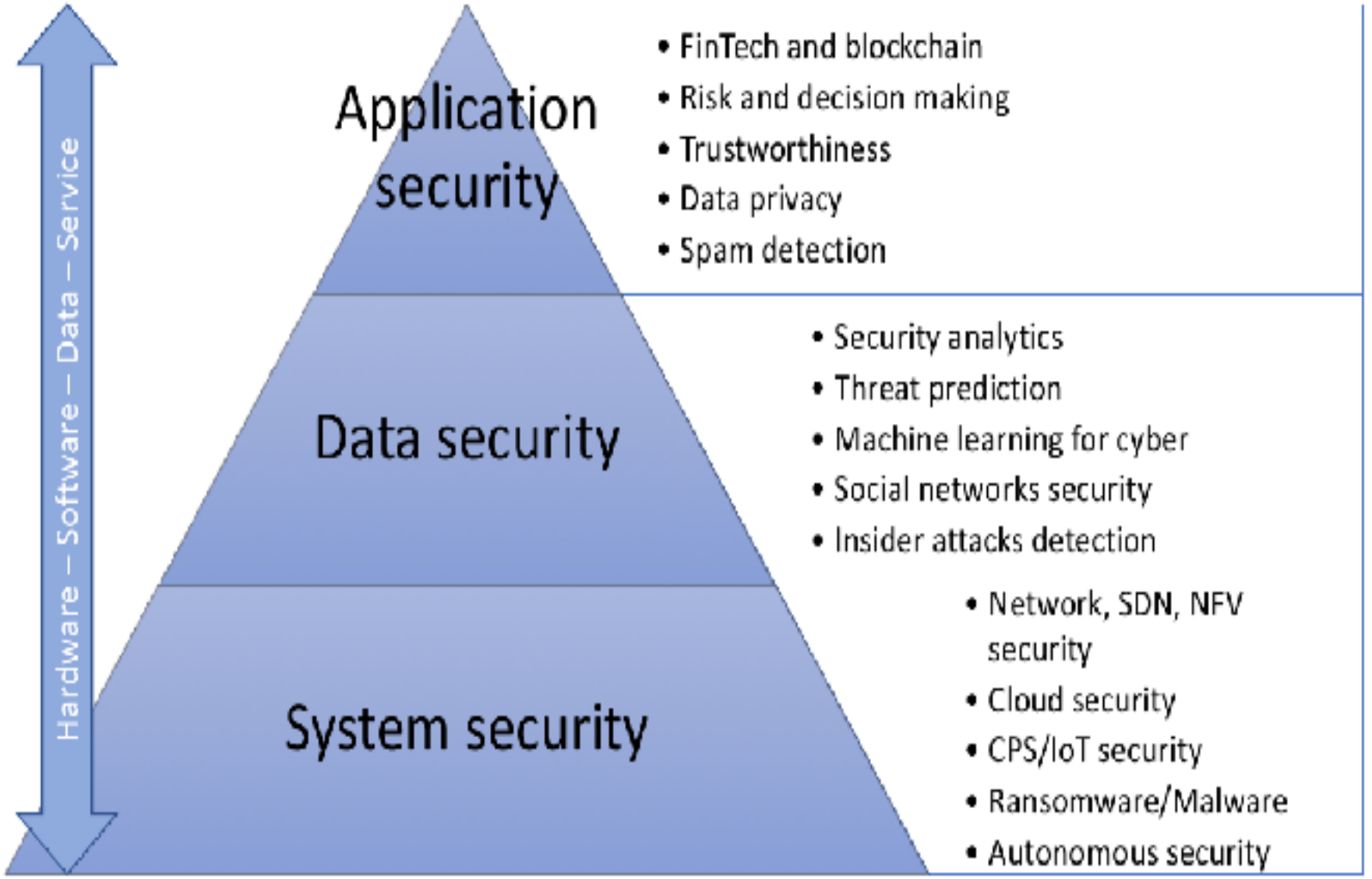
Cybersecurity Evolution



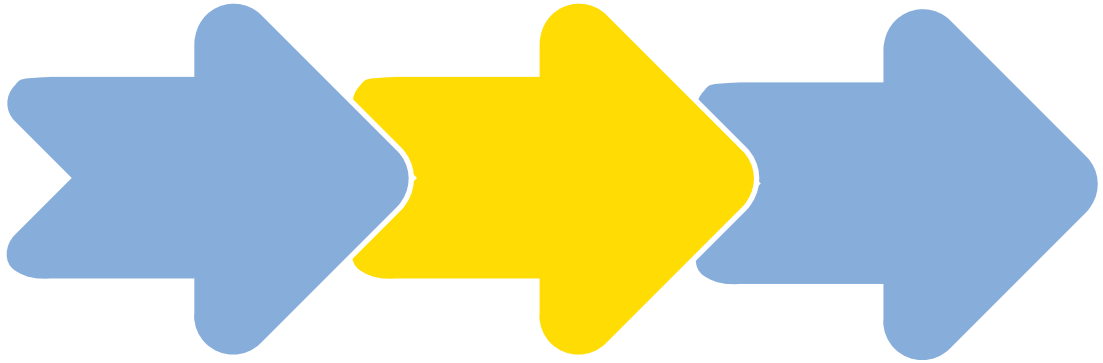
Deep Learning in Real world situation

	Deep Learning	Traditional Machine Learning
Computer Vision	98%	2%
Speech Recognition	80%	20%
Text Understand	65%	35%
Cyber Security	2%	98%

AI in Cyber Security



Cyber Security Tasks - 3 Dimensions



01 Why ? - Goal

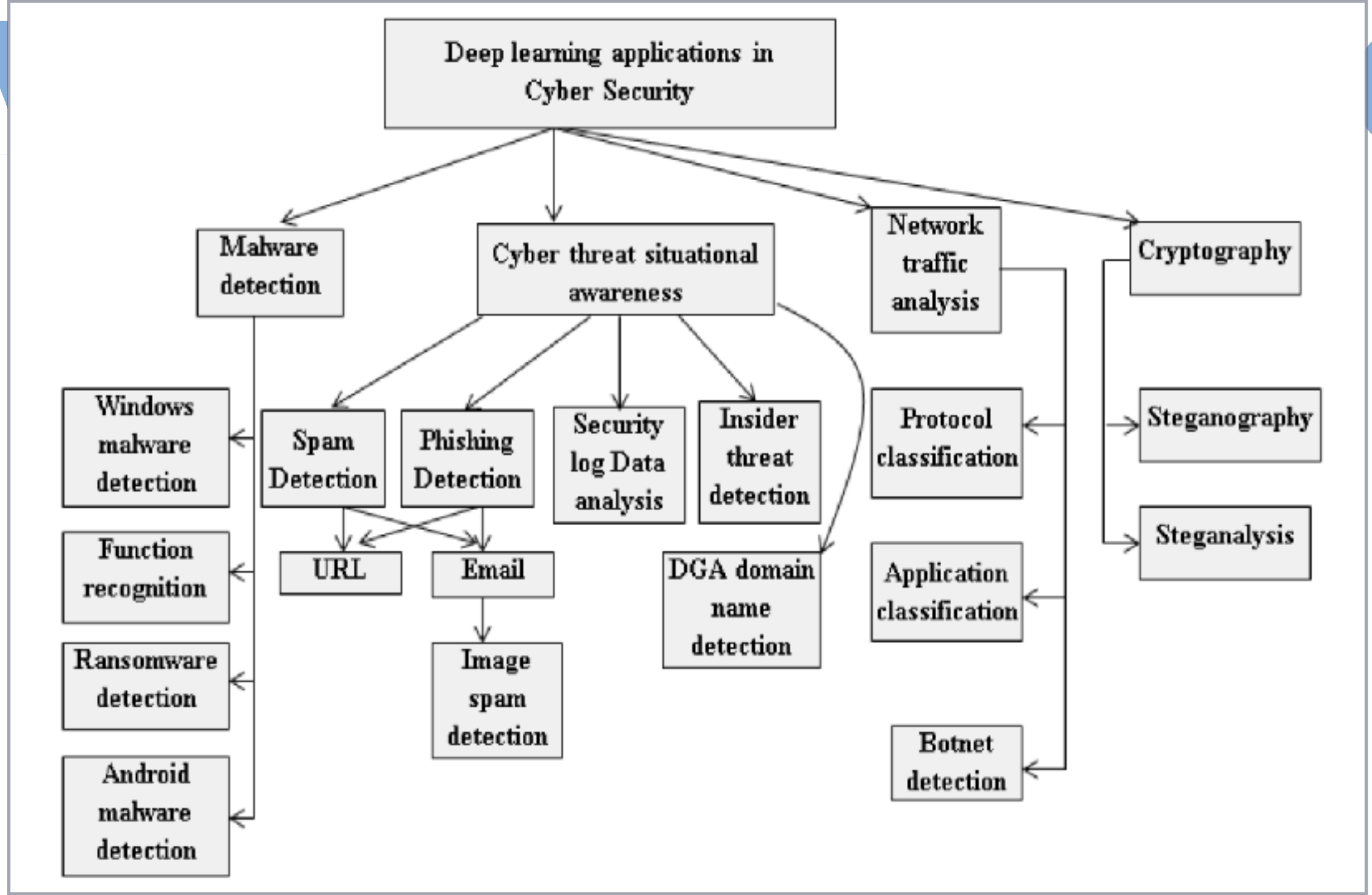
- Prediction
- Prevention
- Detection
- Response
- Monitoring

02 What ?- Technical Layer

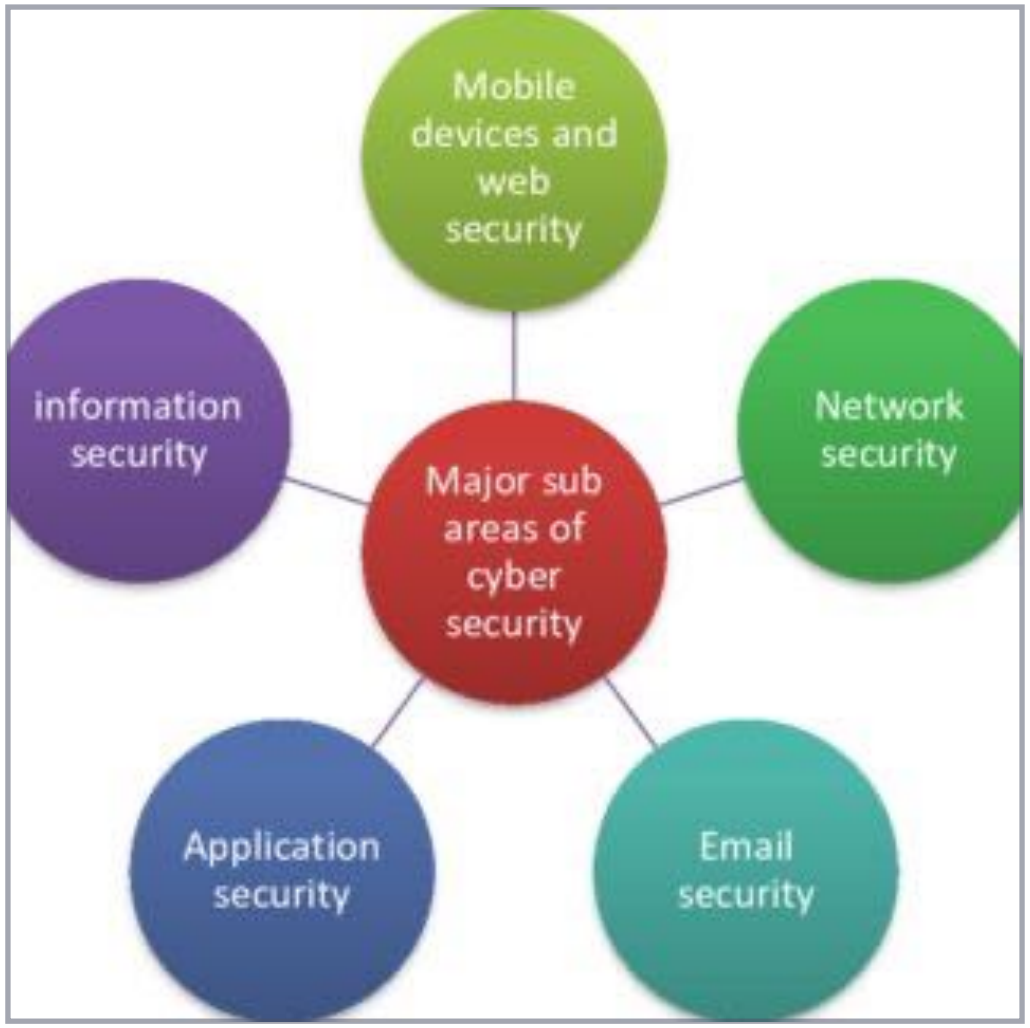
- Network - (Traffic Analysis / Intrusion detection)
- Endpoint (Anti-malware)
- Application (Web Application Firewalls / Database Firewalls)
- User (User Behaviour Analytics)
- Process (Anti fraud)

03 How ? - Area

- in transit in real time
- at rest
- historically


Deep learning applications in Cyber Security
Malware detection
Cyber threat situational awareness
Network traffic analysis
Cryptography
Windows malware detection
Function recognition
Ransomware detection
Android malware detection
Spam Detection
Phishing Detection
Security log Data analysis
Insider threat detection
URL
Email
Image spam detection
DGA domain name detection
Protocol classification
Application classification
Botnet detection
Steganography
Steganalysis

Cyber Security - Major Areas



Cyber Security with Deep Learning

Complex Decision Boundaries: *Protocols and payloads are complex in their variety and structure. Deep learning can make sense of the complexities of threats and identify all types of threats if trained correctly.*

Large training sets: *Enormous threat data sets with hundreds of millions of samples are already available*

GPUs *: Recent technology advancements have made it possible for deep learning model training and validation to be performed in hours, or even minutes when it used to take weeks.*

Deep Learning Applications

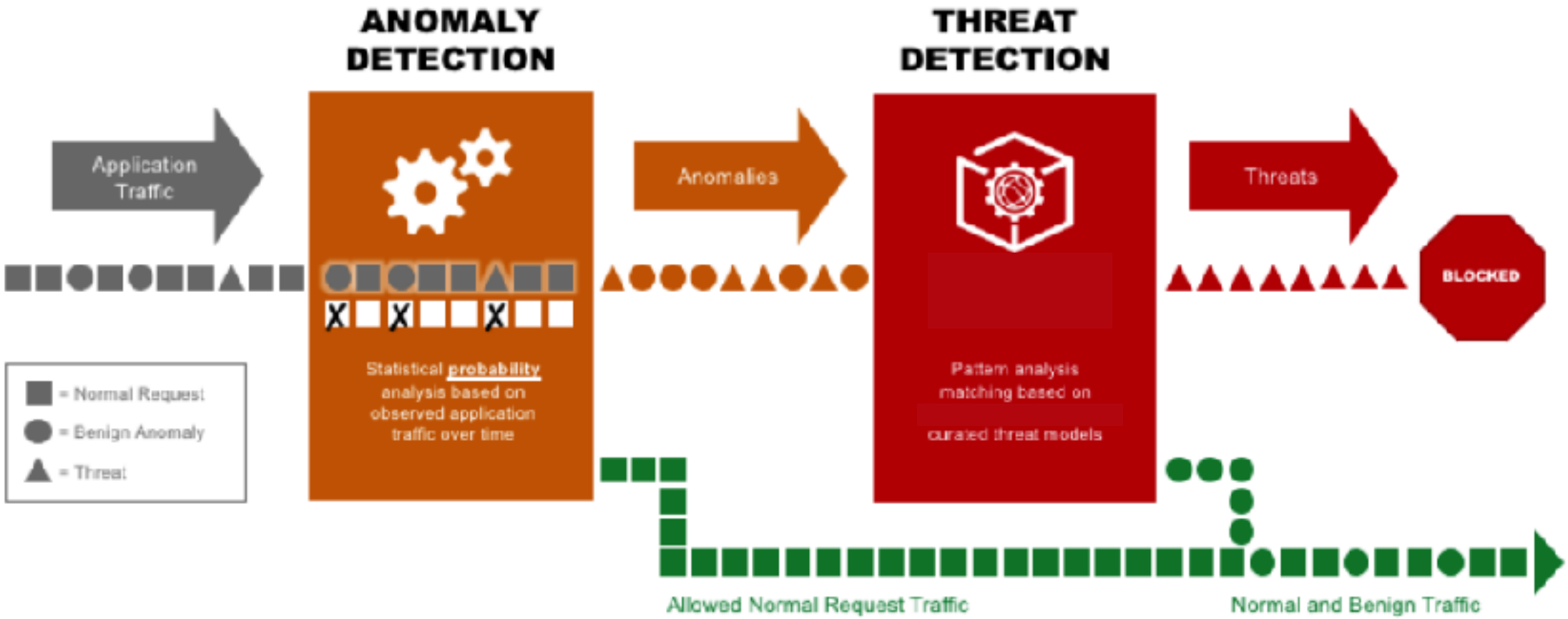
- ***On Endpoint traffic:***
  - *Has access not only to Payload, but also to the runtime behaviour of the specific malware.*
  - *But limited processing and memory resources to meet the computational requirements of a deep learning system.*
  - *Also visibility of threats is limited to one specific end point.*

- ***On Security Incident and Event Management Systems:***
  - *collates logs and alerts from a variety of network and security devices in the network. Applying DL to SIEM traffic has the benefit of lots of interesting data across the enterprise, but the detection is significantly delayed by the time it gets to the SIEM.*
  - *Additionally, due to the myriad of data available, scattered data points may lead to unclear threat verdicts.*

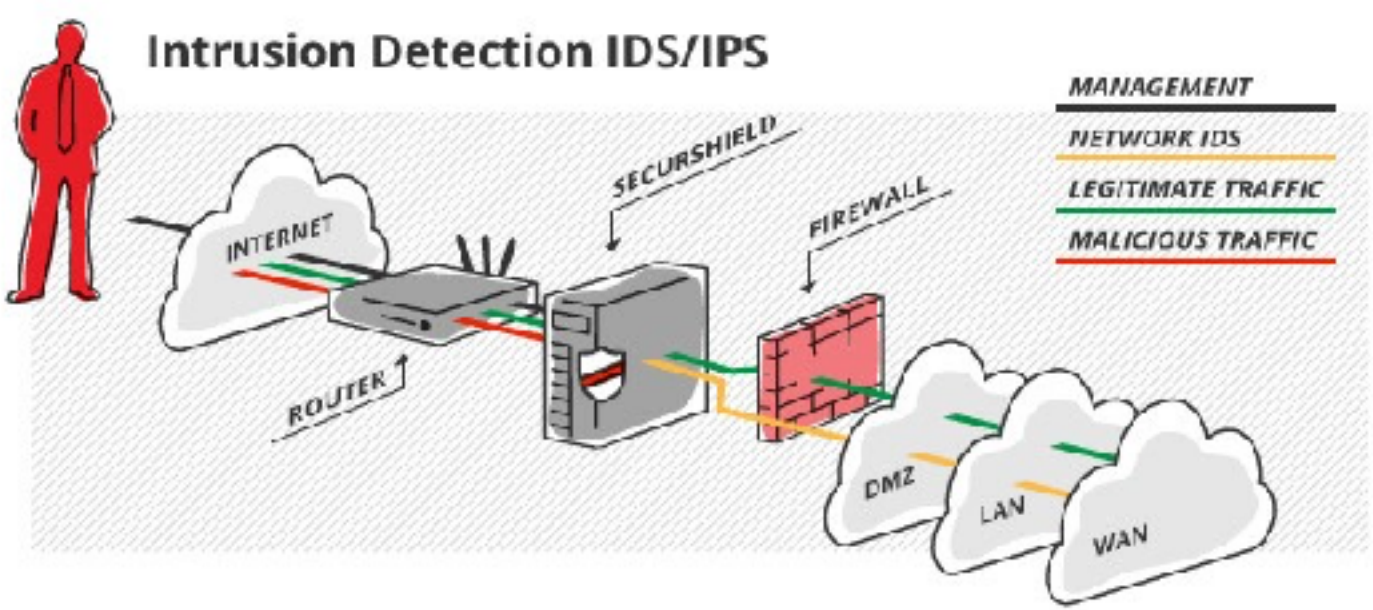
Deep Learning Applications

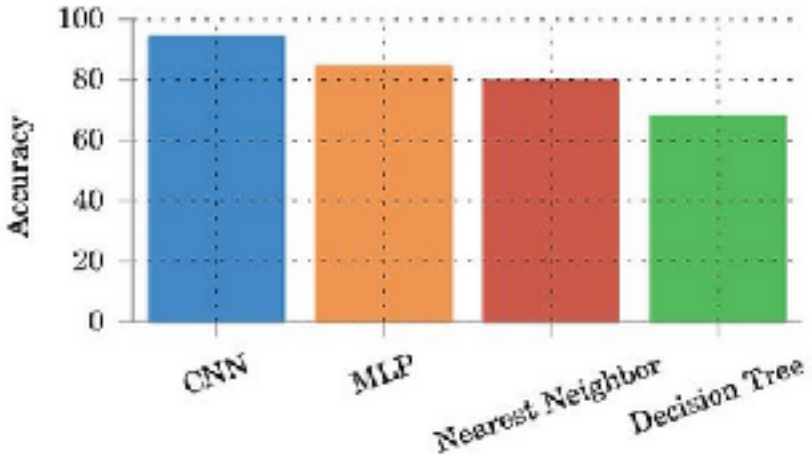
- ***On Network traffic:***
  - *Network payloads and headers brings complexity because of the variety in the structure.*
  - *There are multiple ways to identify malicious intent. - Multiple AI models*
  - *When applying DL at the perimeter of the enterprise brings the benefit of stopping the threat closest to the source of entry before it has the opportunity to move laterally through the enterprise.*
- ***Key metrics for the DL Cyber security solutions:***
  - *Detection speed*
  - *Accuracy and Reliability*
  - *Known and Unknown threats*
  - *Orchestrate Prevention*
  - *Performance*

Threat Detection Example



Intrusion Detection System - DL

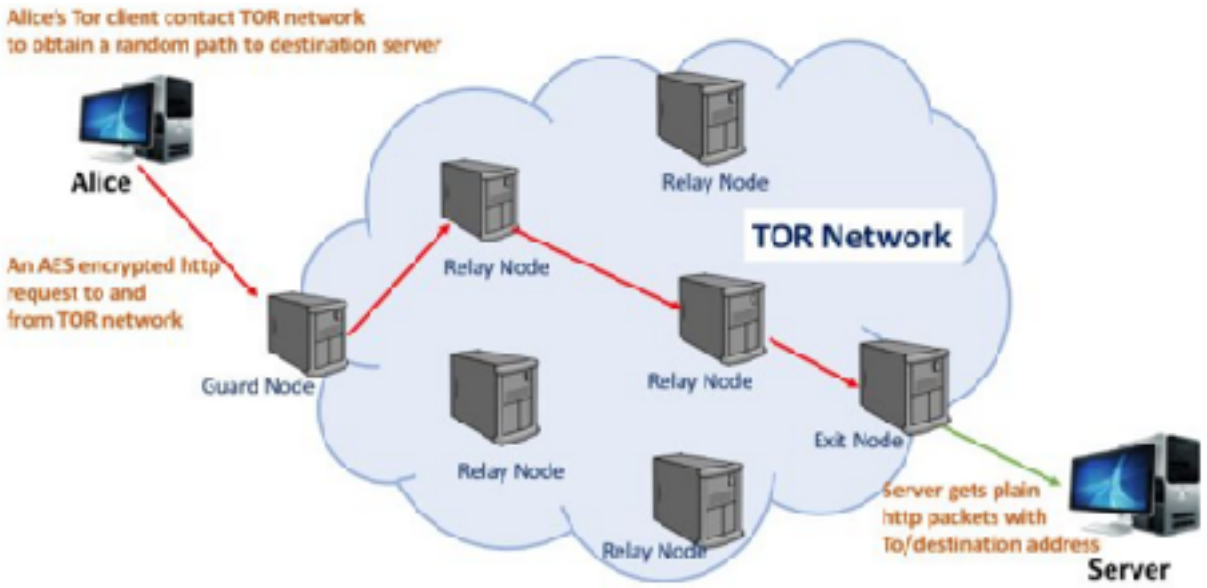


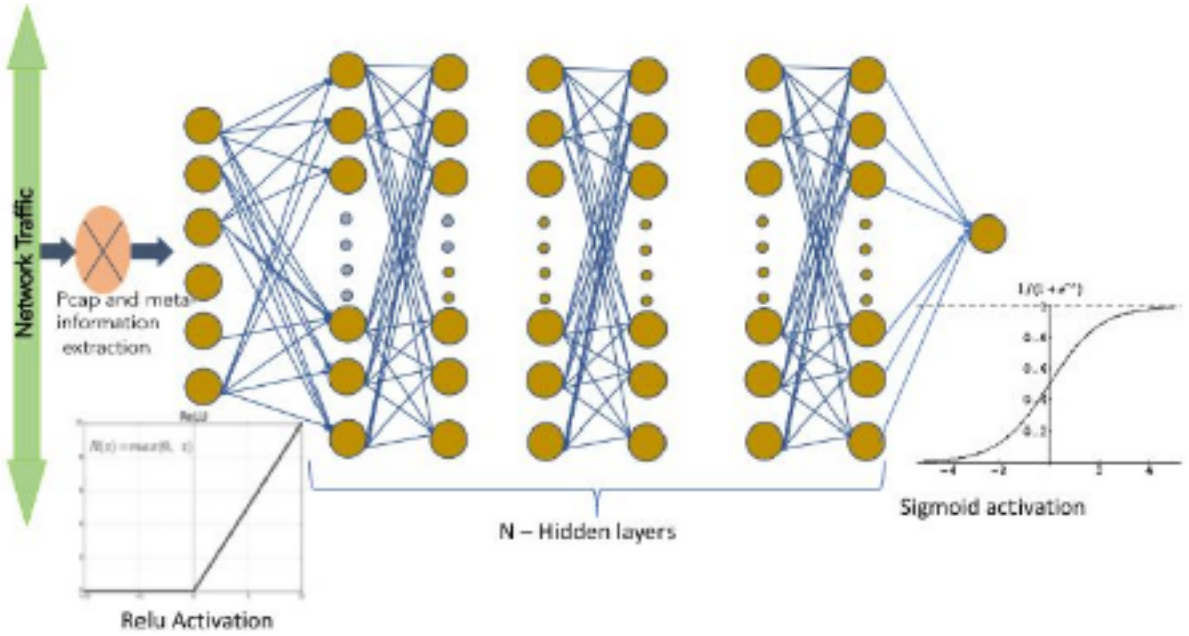


"Advanced AI learns to understand cyber security, recognise patterns and connect dots between threats"

- Jeb Linton, Chief Security Architect, IBM Watson

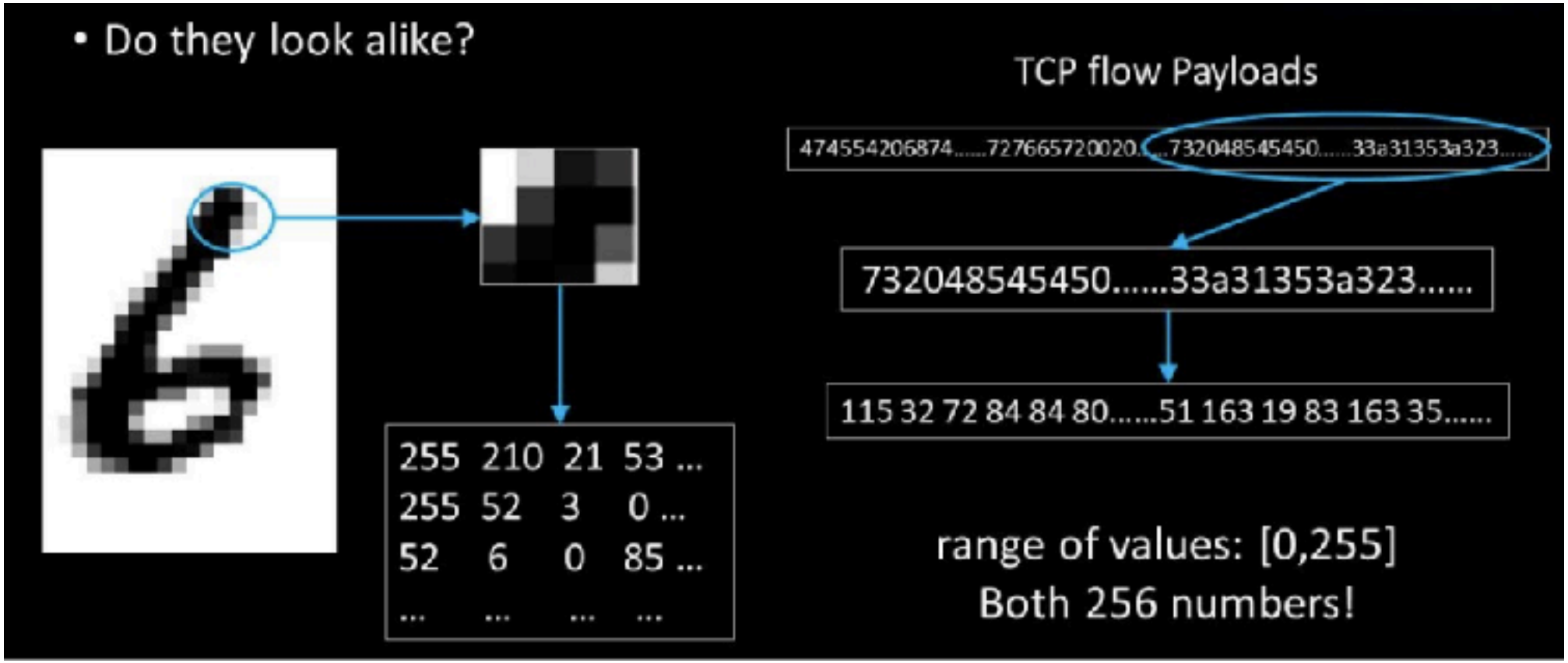
TOR Network Detection - DL



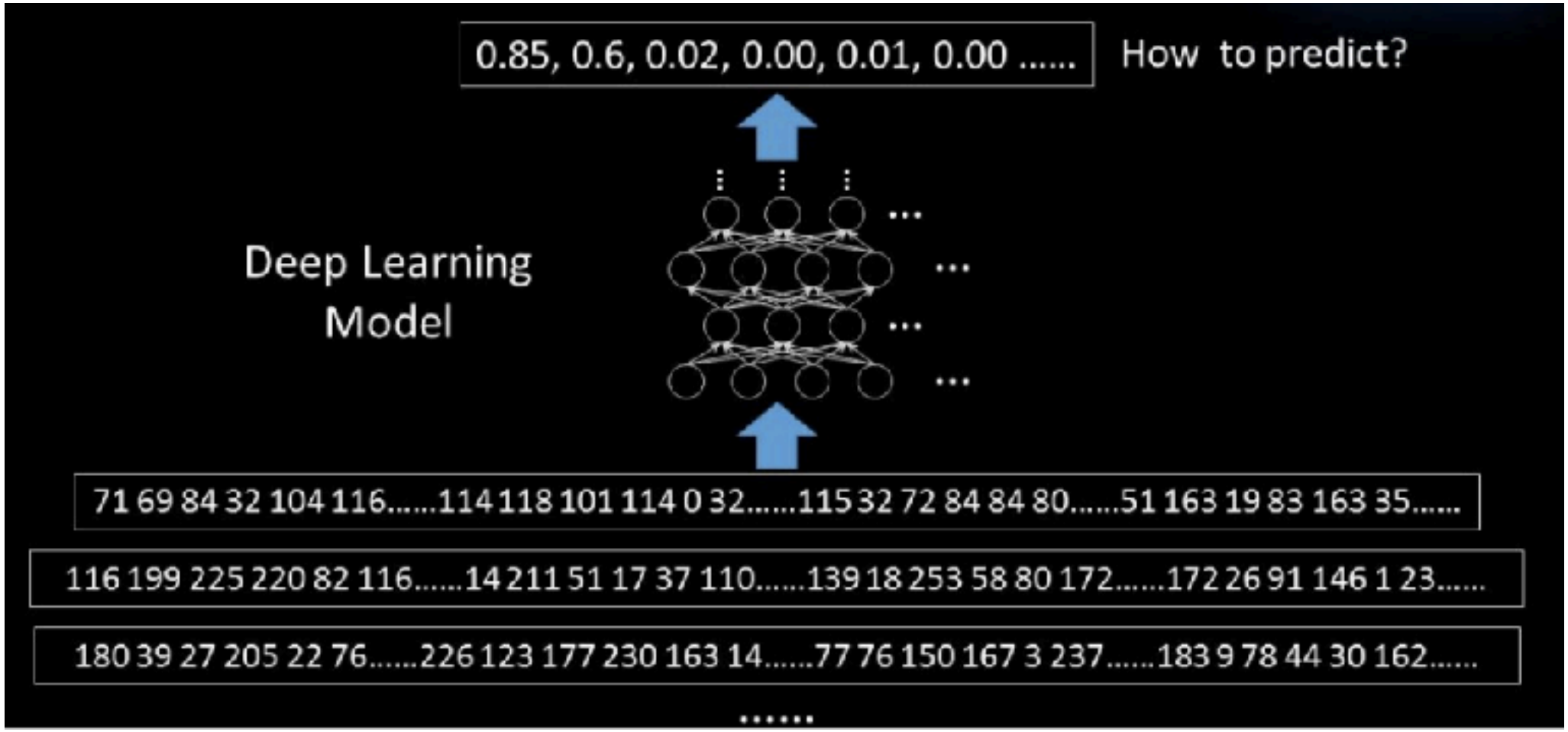


Deep Learning method has higher accuracy in detecting TOR network than Machine Learning

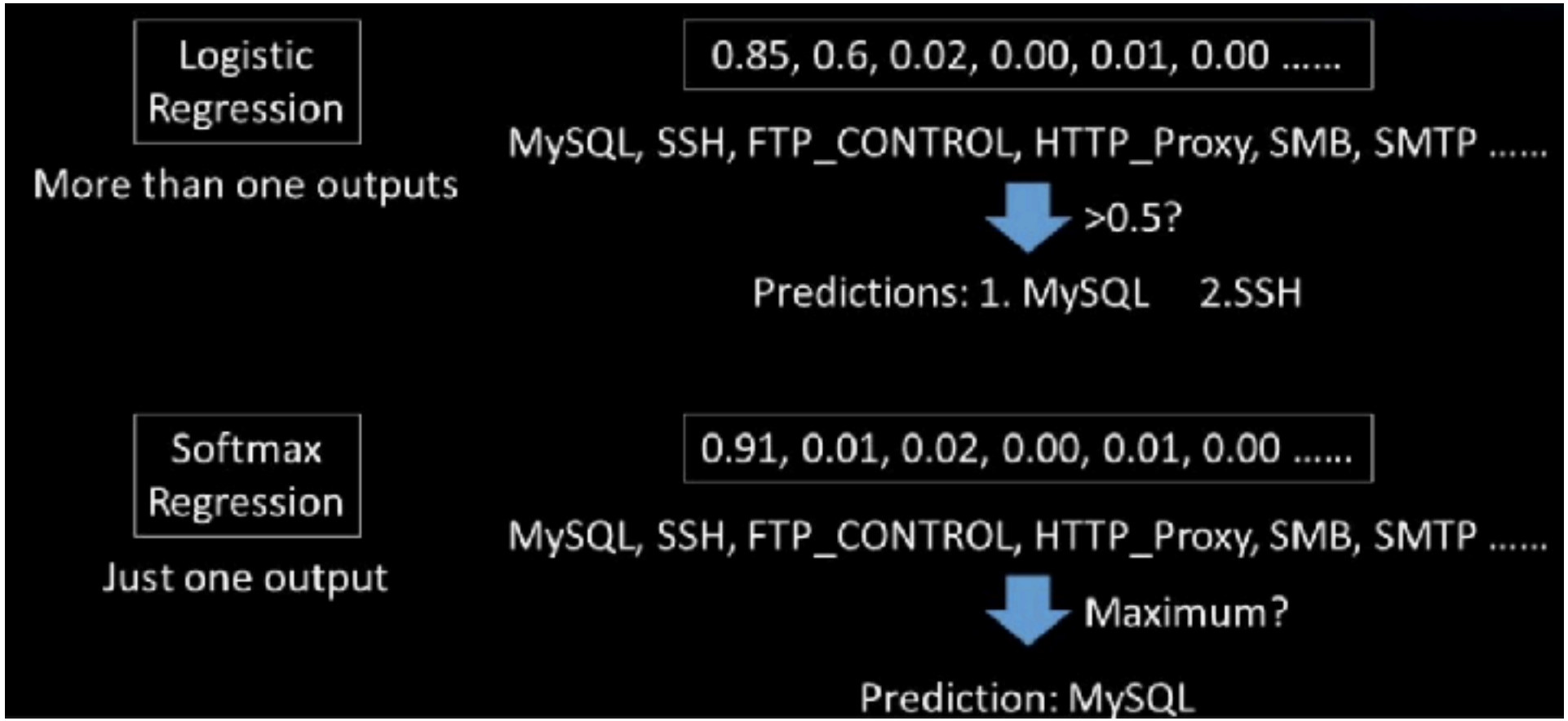
Image Vs Payload



Process of Traffic Classification



Traffic Classification at output layer



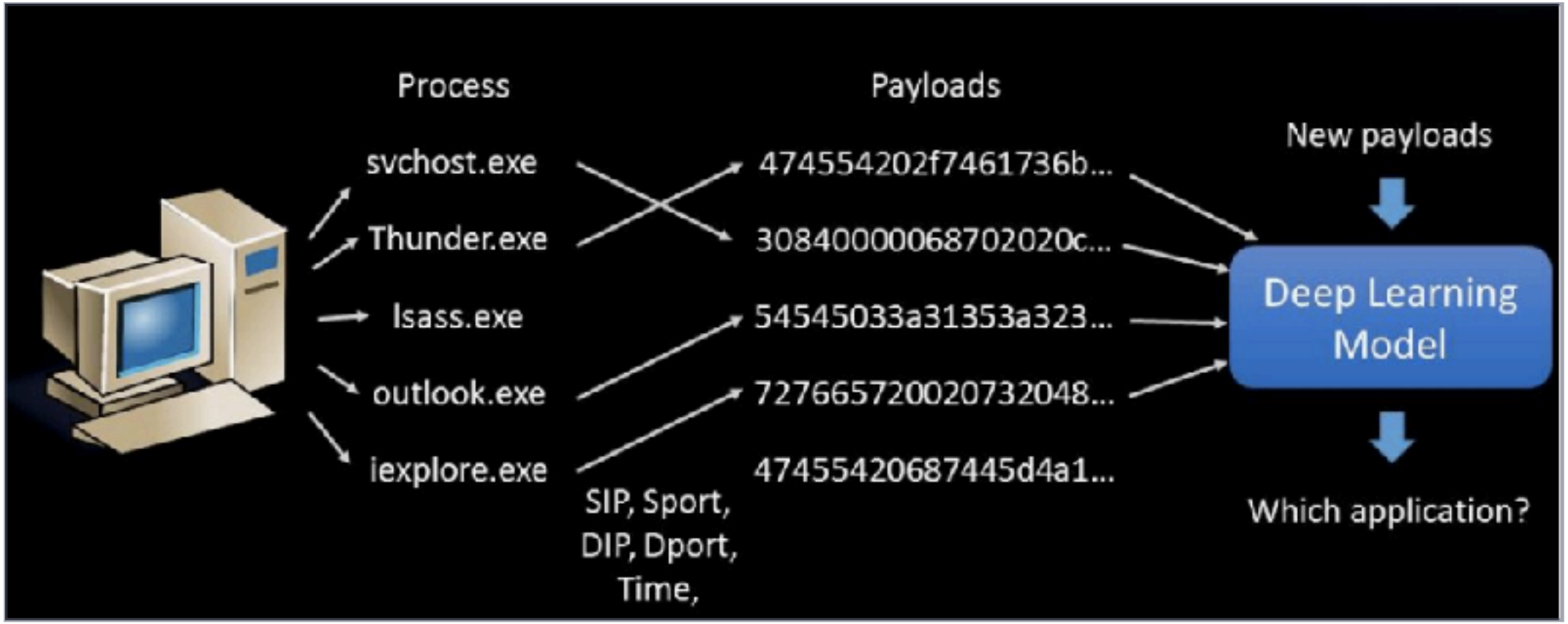
Discover vulnerabilities using DL

Deep Learning techniques to discover vulnerabilities in source and/or binary code bases.

- ***Deep Source Analysis**: New approaches, classification with representation learning and deep learning with multiple sources for code analysis.*

- ***Deep Binary Analysis**: Innovatively convert binary code to different data representation such as image, and employ deep neural networks to assist binary analysis.*

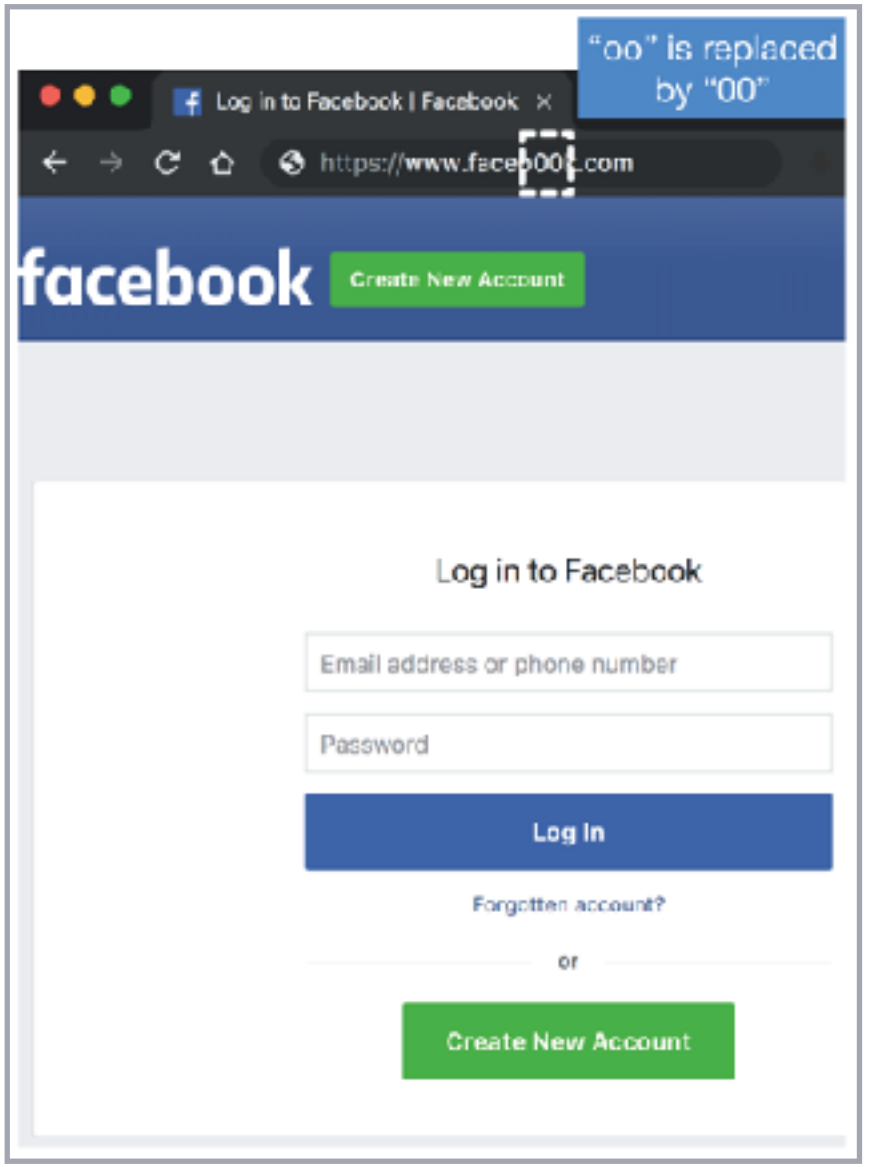
Application Identification



5 Cyber Security Threats – DL can protect against

Threats	Drawbacks	Deep Learning	DL Method
Spear Phishing	Lack of speed, accuracy to reliably catch all the malicious links	URL prediction	models are trained to identify micro behaviors (email headers, subsamples of body-data, punctuation patterns, etc.
Watering hole	Track the sites that users visit often from external network	Path detection	Detect malicious domains, Monitor redirect patterns to and from host
Webshell	Online payments, Redirect to hacker servers	Identification and isolation of webshell	Statistics of a normal shopping cart behavior can be detected and DL models can be trained to identify normal behavior from malicious behavior.
Ransomware	File, Computer	Detect unknown ransomware	A large set of ransom files & an even larger set of clean files are to be trained to identify micro behaviors
Remote exploitation	DDOS, DNS poisoning, Port scanning	Detect exploitation payload	Analyze system behavior and identify abnormal instances

Phishing with Deep Learning



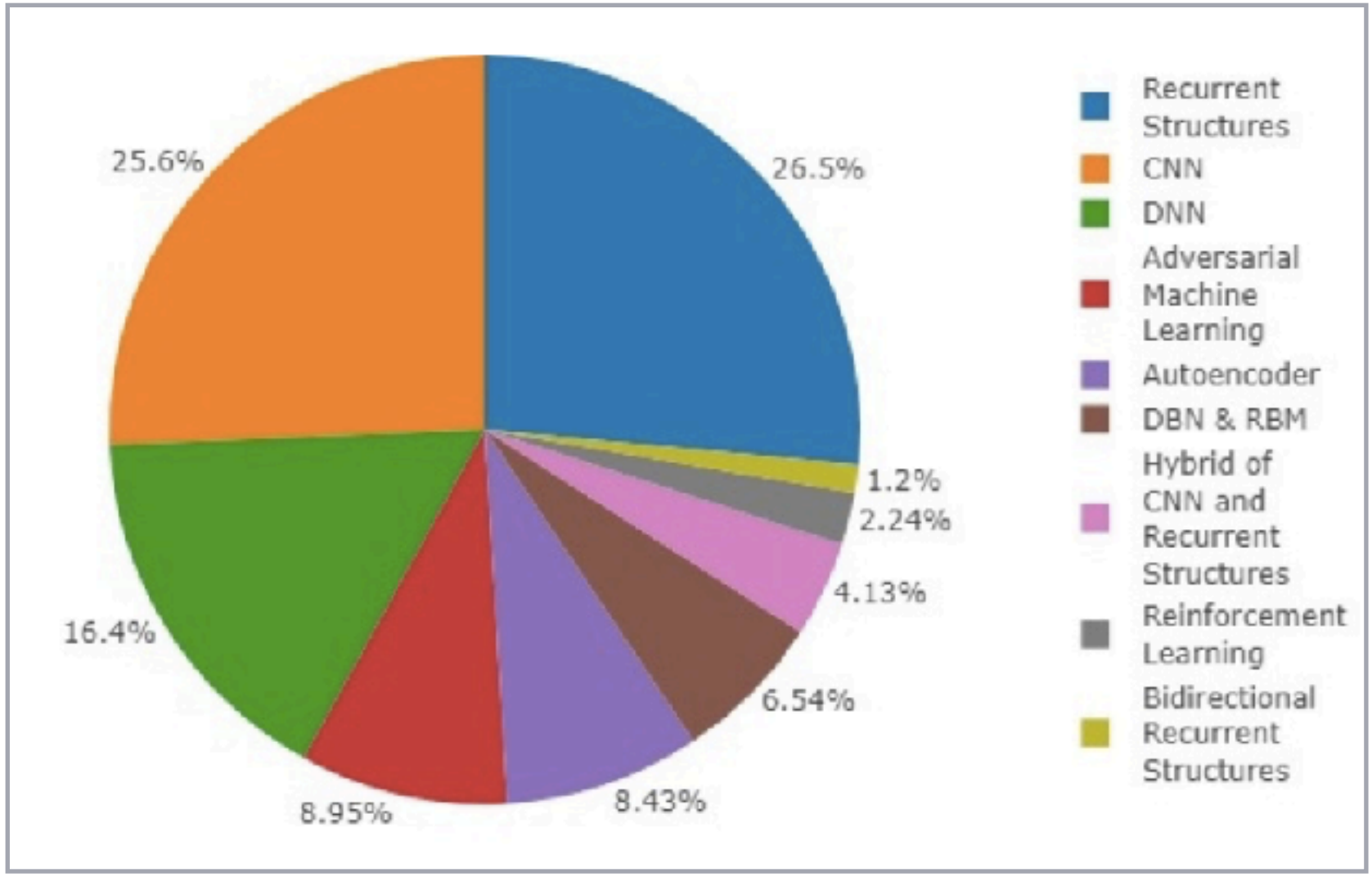
Deep Neural Network

DL Algorithms for Cyber security areas

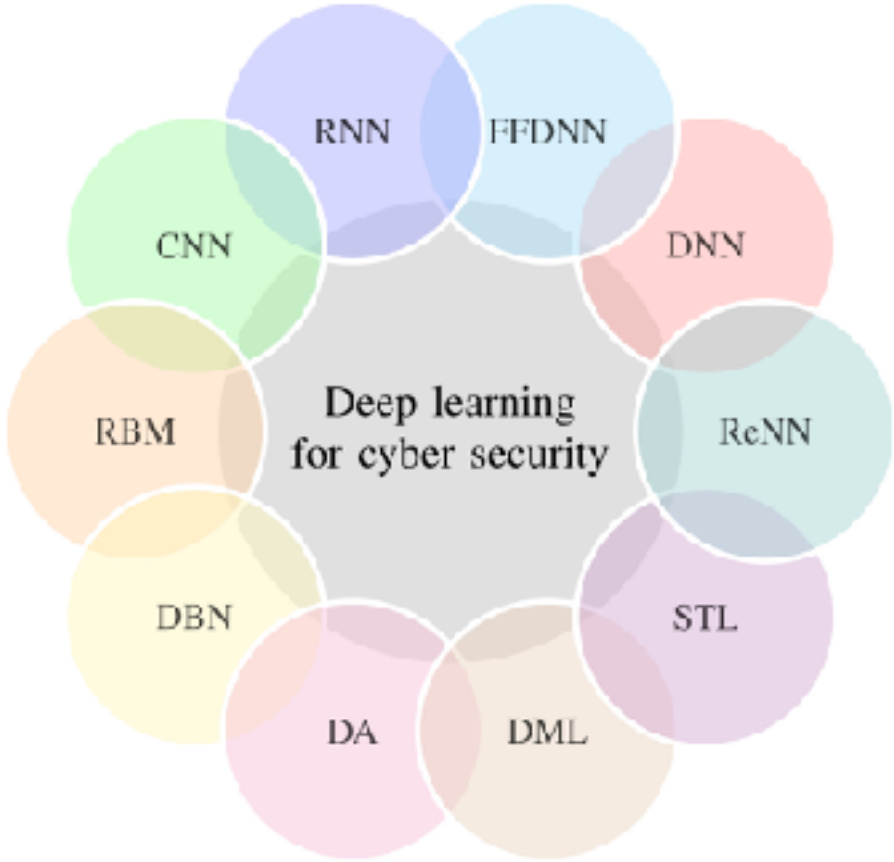
	Network IDS	Botnet	Malware Analysis	Spam Detection
Supervised	RNN CNN	RNN	FNN CNN RNN	
Unsupervised	DBN SAE		DBN SAE	DBN SAE

RNN – Recurrent Neural Network, FNN – Feed Forward, CNN – Convolutional, SAE - Stack Auto Encoder , DBN – Deep Belief Network

Statistics of Deep Learning Architectures in Cyber Security

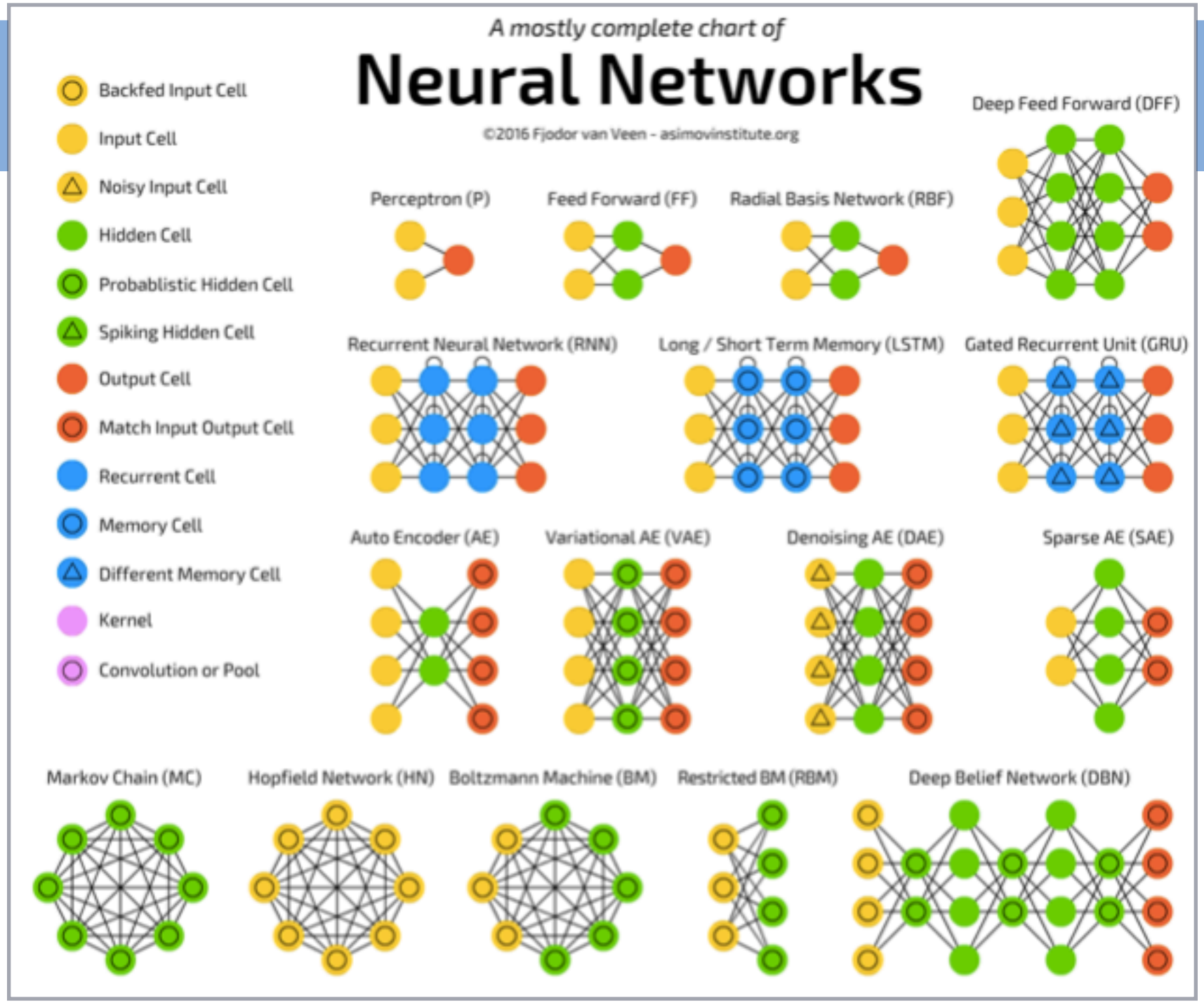


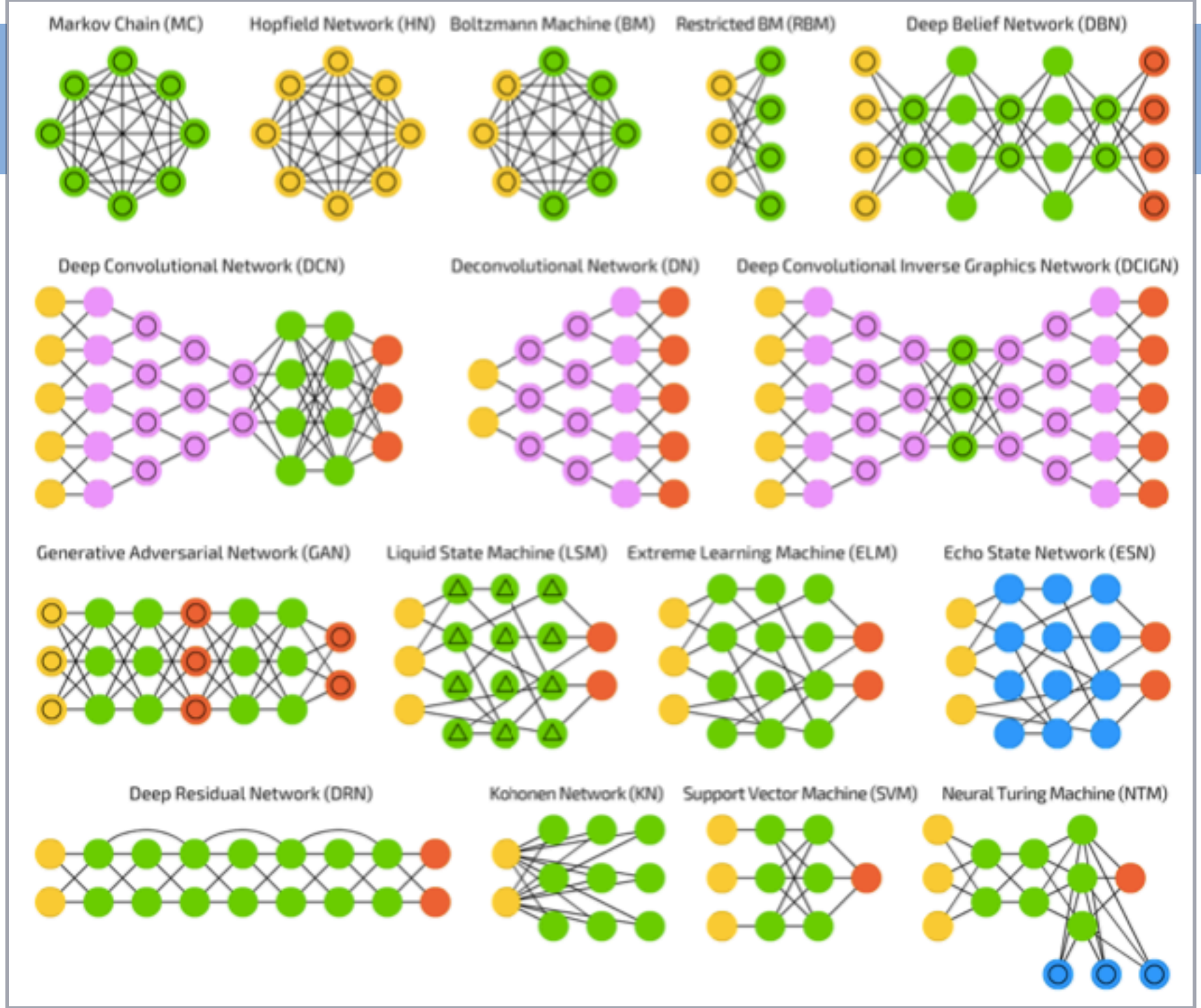
Intrusion Detection System - DL



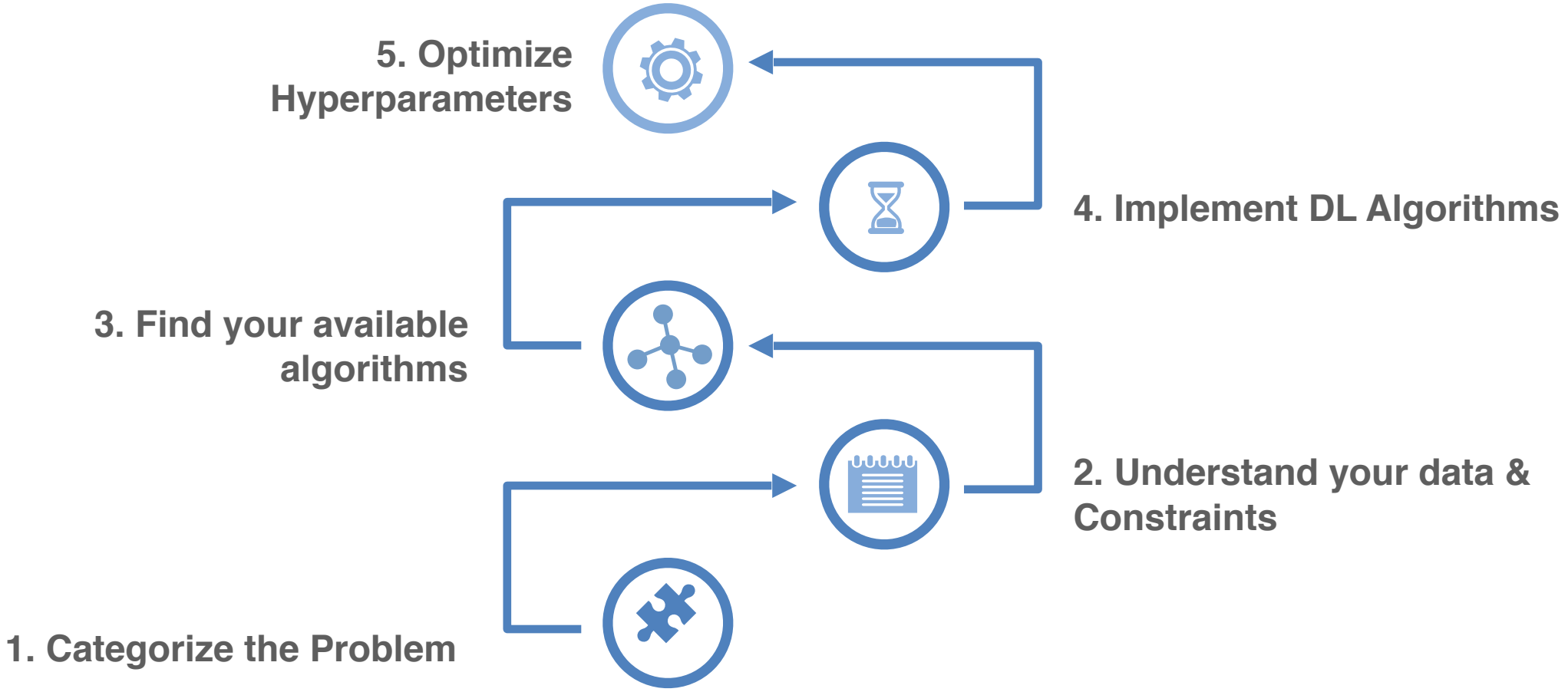
"Advanced AI learns to understand cyber security, recognise patterns and connect dots between threats"

- Jeb Linton, Chief Security Architect, IBM Watson


A mostly complete chart of
Neural Networks
©2016 Fjodor van Veen - asimovinstitute.org
Backfed Input Cell
Input Cell
Noisy Input Cell
Hidden Cell
Probablistic Hidden Cell
Spiking Hidden Cell
Output Cell
Match Input Output Cell
Recurrent Cell
Memory Cell
Different Memory Cell
Kernel
Convolution or Pool
Perceptron (P)
Feed Forward (FF)
Radial Basis Network (RBF)
Deep Feed Forward (DFF)
Recurrent Neural Network (RNN)
Long / Short Term Memory (LSTM)
Gated Recurrent Unit (GRU)
Auto Encoder (AE)
Variational AE (VAE)
Denoising AE (DAE)
Sparse AE (SAE)
Markov Chain (MC)
Hopfield Network (HN)
Boltzmann Machine (BM)
Restricted BM (RBM)
Deep Belief Network (DBN)


Markov Chain (MC)
Hopfield Network (HN)
Boltzmann Machine (BM)
Restricted BM (RBM)
Deep Belief Network (DBN)
Deep Convolutional Network (DCN)
Deconvolutional Network (DN)
Deep Convolutional Inverse Graphics Network (DCIGN)
Generative Adversarial Network (GAN)
Liquid State Machine (LSM)
Extreme Learning Machine (ELM)
Echo State Network (ESN)
Deep Residual Network (DRN)
Kohonen Network (KN)
Support Vector Machine (SVM)
Neural Turing Machine (NTM)

Steps to choose the right DL Algorithm



Categorize your Problem - Input

Labeled data → Supervised learning problem

Unlabeled data → Unsupervised learning problem

If the solution implies to optimize an objective function by interacting with an environment → Reinforcement Learning Problem

Categorize your Problem - Output

Output	Problem	Example
Number	Regression	Rate of attacks, Error predictions
Class	Classification	Yes/No, HTTP/FTP
Set of Input Groups	Clustering	Cyber-profiling

Understand your Constraints

- ***What is your data storage capacity?***
- ***Does the prediction have to be fast?***
- ***Does the learning have to be fast?***

Analyse the available Algorithms

- *Identify the algorithms that are applicable and practical to implement in a reasonable time.*

- *Some of the elements affecting the choice of a model are.*
  - ✓ *The accuracy of the model.*
  - ✓ *The interpretability of the model.*
  - ✓ *The complexity of the model.*
  - ✓ *The scalability of the model.*
  - ✓ *How long does it take to build, train, and test the model?*
  - ✓ *How long does it take to make predictions using the model?*
  - ✓ *Does the model meet the business goal?*

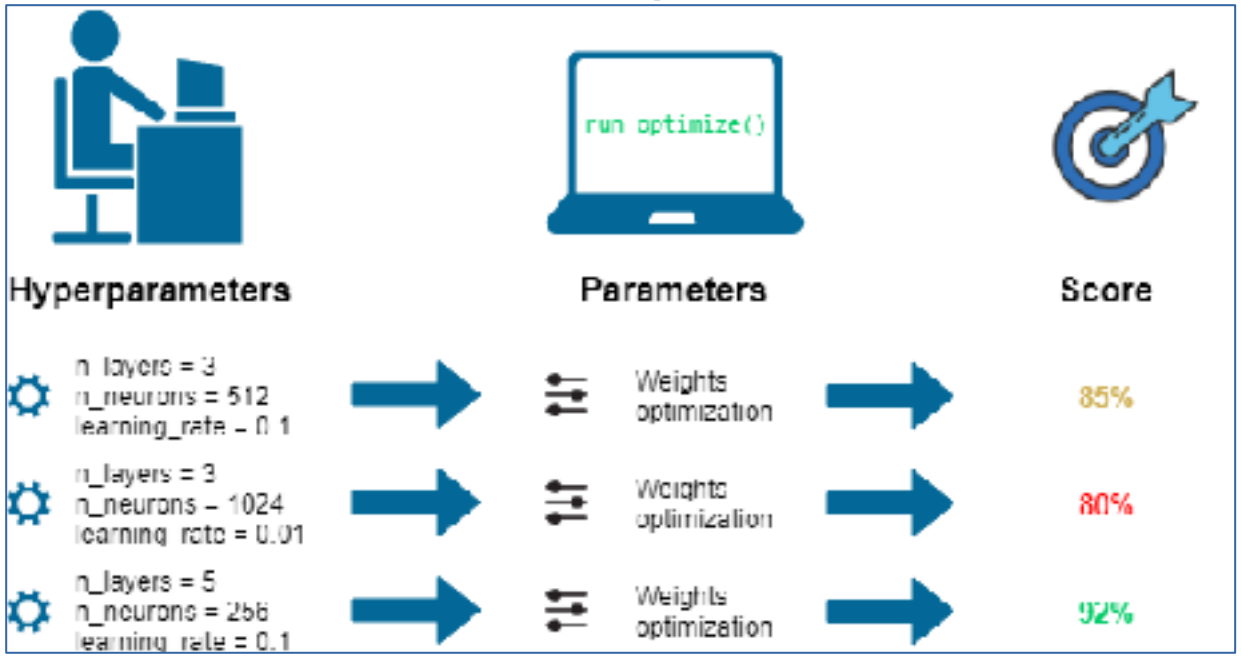
Implement Deep Learning Approach

- ***Approach 1:***
  - ✓ *Use multiple approaches on a single dataset.*
  - ✓ *Compare the performances of each approach on a dataset using a set of carefully selected evaluation criteria.*
- ***Approach 2:***
  - ✓ *Use single approach on different subgroups of datasets.*
  - ✓ *The best solution for this is to do it once or have a service running that does this in intervals when new data is added.*

Optimization of Hyperparameters

- *Hyperparameter is a parameter whose value is set before the learning process begins. Whereas, the values of other parameters are derived via training.*

- *Optimization – Finding the best values for the hyperparameters of your model. Example: K-means clustering, Neural network hidden layers*



Future of AI in Cyber Security

Predictive Protection

AI that **anticipates attacks** and **automatically reconfigures** for protection

AI take away jobs, but Security Domain create jobs for people with AI expertise

Prediction – global cybercrime would cost $6 trillion annually by 2021

Global shortage of two million cyber security professionals in 2019 - **ISACA**

India alone will need 1 million cybersecurity professionals by 2020 – **NASSCOM**

Cybersecurity Ventures predicts there will be 3.5 million cybersecurity job openings by 2021.

There is a **zero-percent unemployment rate in cybersecurity** and the opportunities in this field are endless – **Herjavec Group**

✓ **CAREERS**

Penetration Tester
Network Engineer
Cyber Security Engineer
Information Security Analyst
Cyber Security Architect
Forensic Analyst
Network Analyst
Cyber Security Specialist/ Technician

Thank you for your attention