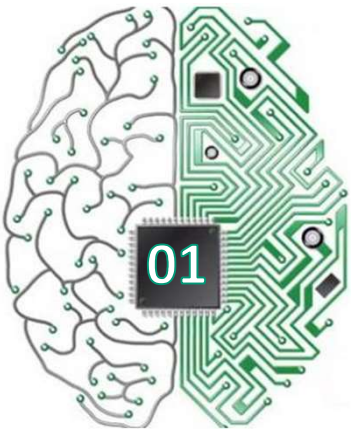
Open Elective Course [OE]

Course Code: CSO507

Winter 2023-24

Lecture#



Deep Learning

Unit-1: Introduction to Deep Learning

Course Details

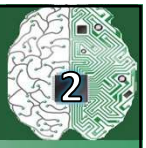
Course Instructor:

Dr. Monidipa Das

Assistant Professor

Department of Computer Science and Engineering

Indian Institute of Technology (Indian School of Mines) Dhanbad, Jharkhand 826004, India



Unit-1: Introduction to Deep Learning (DL)

- What is Learning?
- What is Deep Learning?
- Traditional Machine Learning vs. Deep Learning
- Brief History of Deep Learning
- Why Deep Learning?

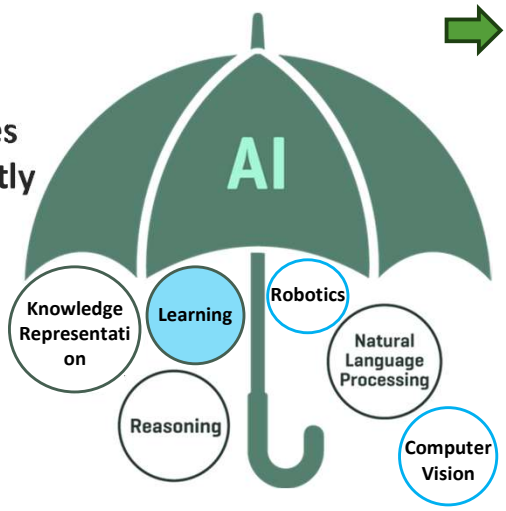
What is Learning

3

- A subfield of artificial intelligence (AI)
- Machine Learning: "the field of study that **gives computers the ability to learn without explicitly being programmed**." [Arthur Samuel, 1950]

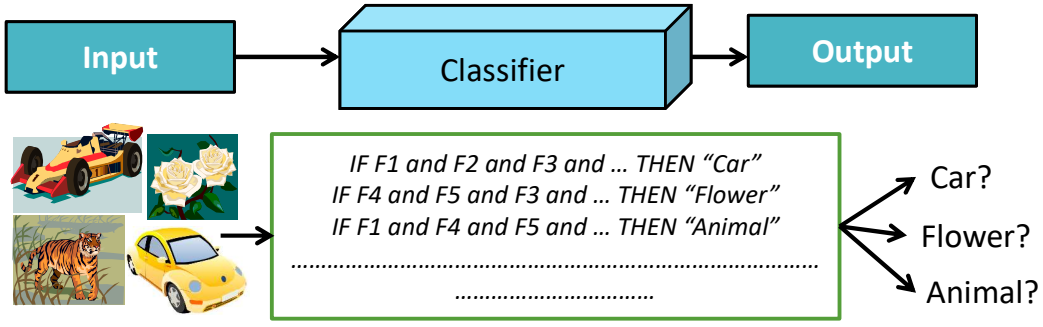
A computer program is said to ***learn*** from ***experience E*** with respect to some class of ***tasks T*** and ***performance measure P***, if its performance at tasks in ***T***, as measured by ***P***, improves with experience ***E***.

---[Tom Mitchell]

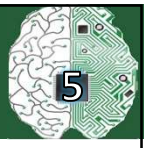


Explicitly Programming

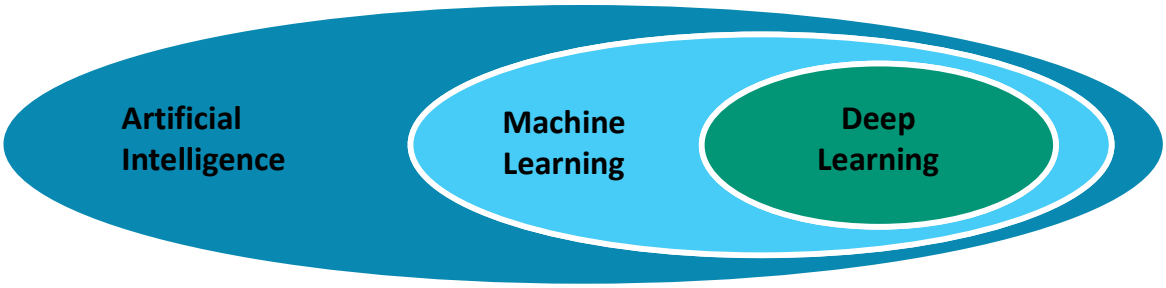
4



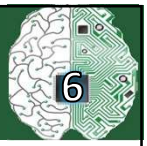
What is Deep Learning (DL)?



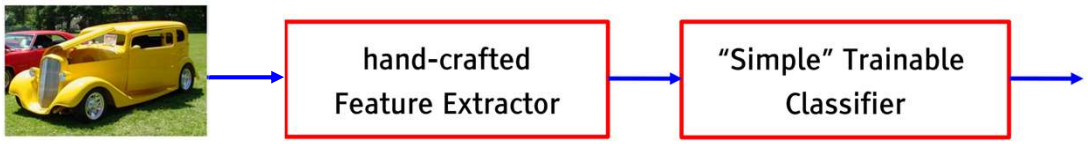
- "….computational models that are composed of **multiple processing layers to learn representations of data** with **multiple levels of abstraction**." ------ *Deep Learning by Y. LeCun et al. Nature 2015*



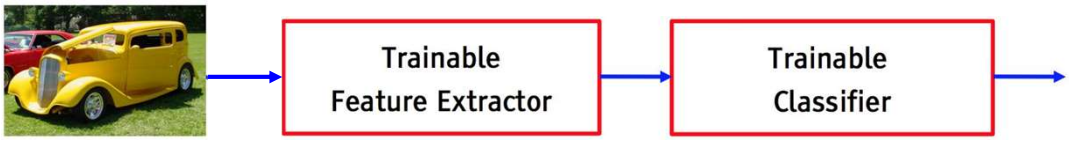
Traditional Machine Learning vs. Deep Learning



- **The traditional model of pattern recognition (since the late 50's)**
  - Fixed/engineered features (or fixed kernel) + trainable classifier

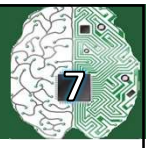


- **End-to-end learning / Feature learning / Deep learning**
  - Trainable features (or kernel) + trainable classifier



Courtesy: "Deep Learning Tutorial" by Yann LeCun [ICML, Atlanta, 2013]

Traditional Machine Learning vs. Deep Learning



Image

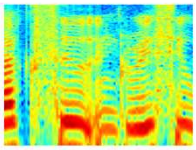
Image

Vision Features

Detection

Audio

Audio

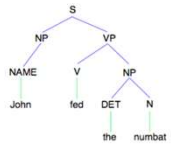


Audio Features

Speaker ID

Text

Text

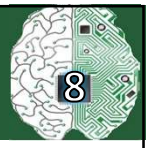


Text Features

Web Search

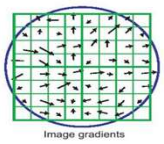
Courtesy: Prof. Andrew Ng

Traditional Machine Learning vs. Deep Learning

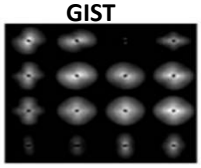


Vision Features

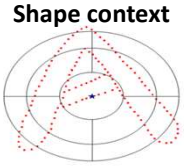
SIFT



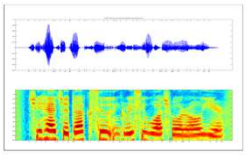
GIST



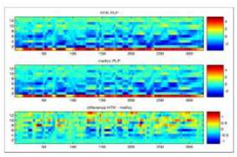
Shape context



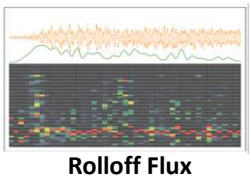
Audio Features



Spectrogram

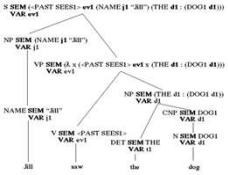


MFCC



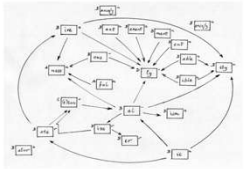
Rolloff Flux

Text Features



Parser

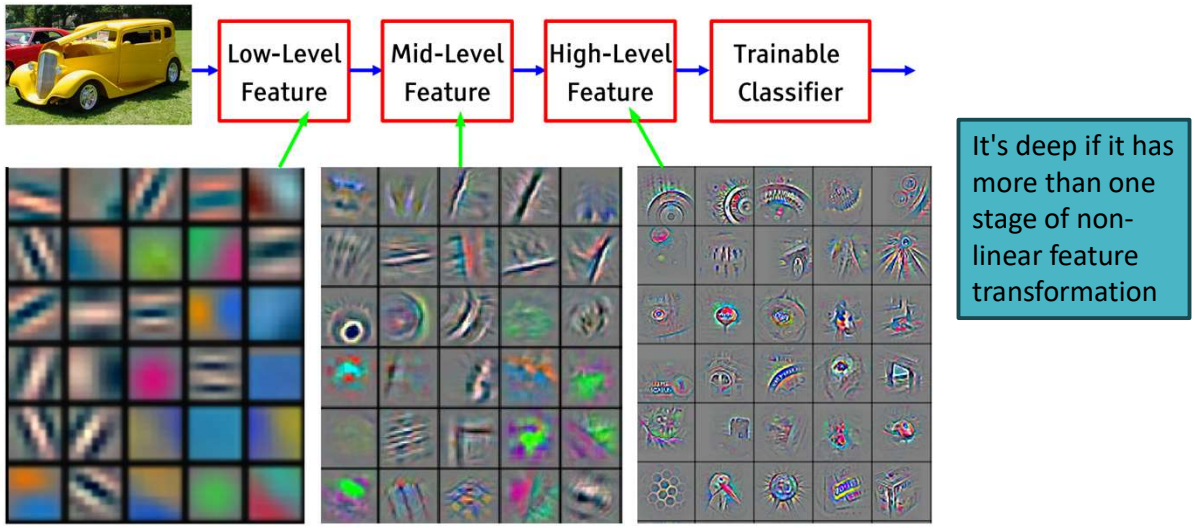
Named entity



Stemming

Deep Learning (DL): Learning Hierarchical Representations

9



It's deep if it has more than one stage of non-linear feature transformation

Courtesy: "Deep Learning Tutorial" by Yann LeCun [ICML, Atlanta, 2013]

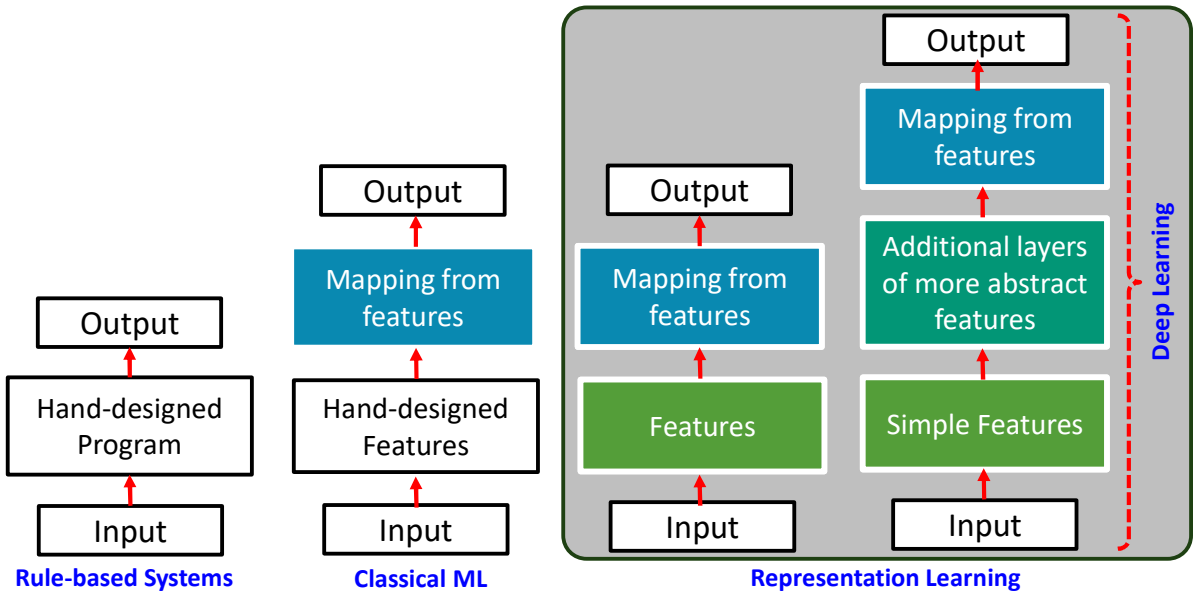
Trainable Feature Hierarchy

10

- Each module transforms its input representation into a higher-level one.
- High-level features are more global and more invariant
- Low-level features are shared among categories

- **Image recognition**
  - Pixel → edge → texton → motif → part → object
- **Text**
  - Character → word → word group → clause → sentence → story
- **Speech**
  - Sample → spectral band → sound → … phone → phoneme → word

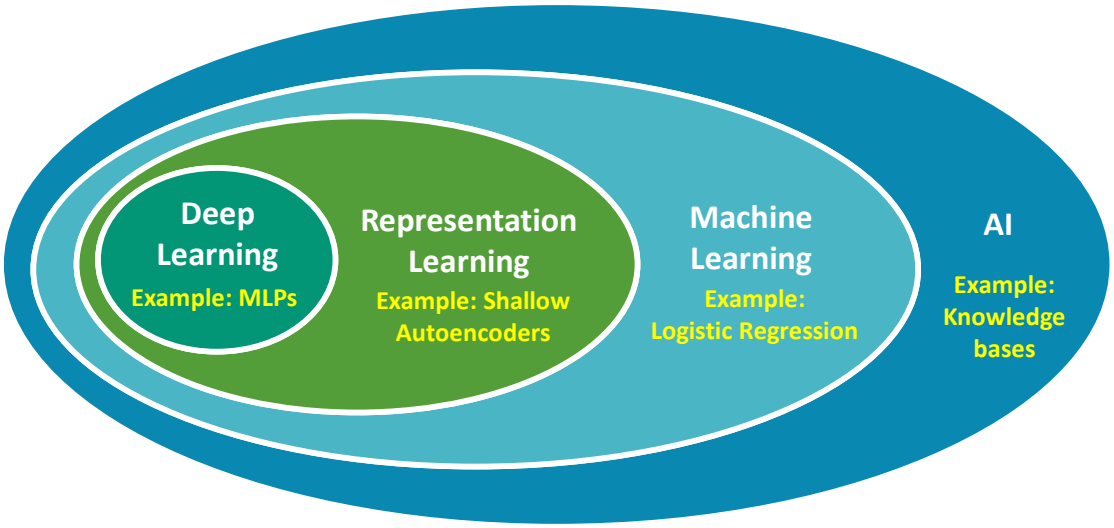
Three Paradigms of AI



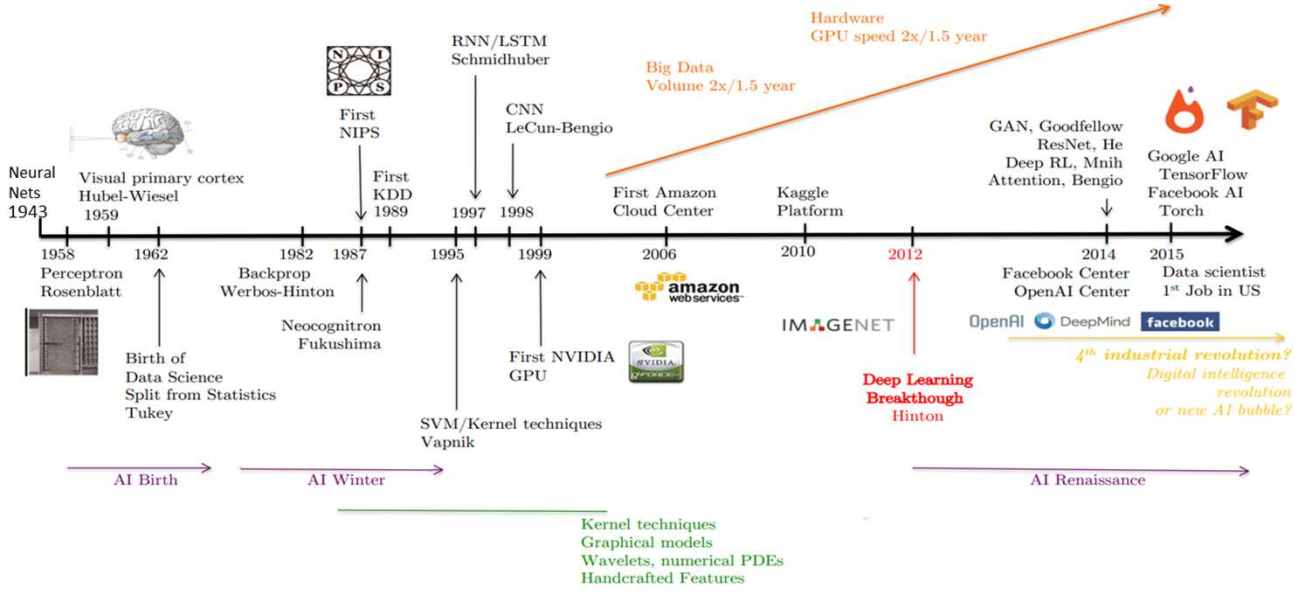
AI Paradigm Shifts

- **Knowledge-based systems**
  - Cannot perform simple recognition tasks
- **Simple machine learning methods**
  - Cannot perform complex recognition tasks
- **Deep Learning methods**

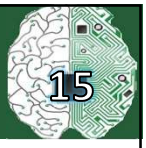
DL in AI System



Brief History of Deep Learning

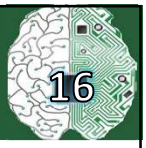


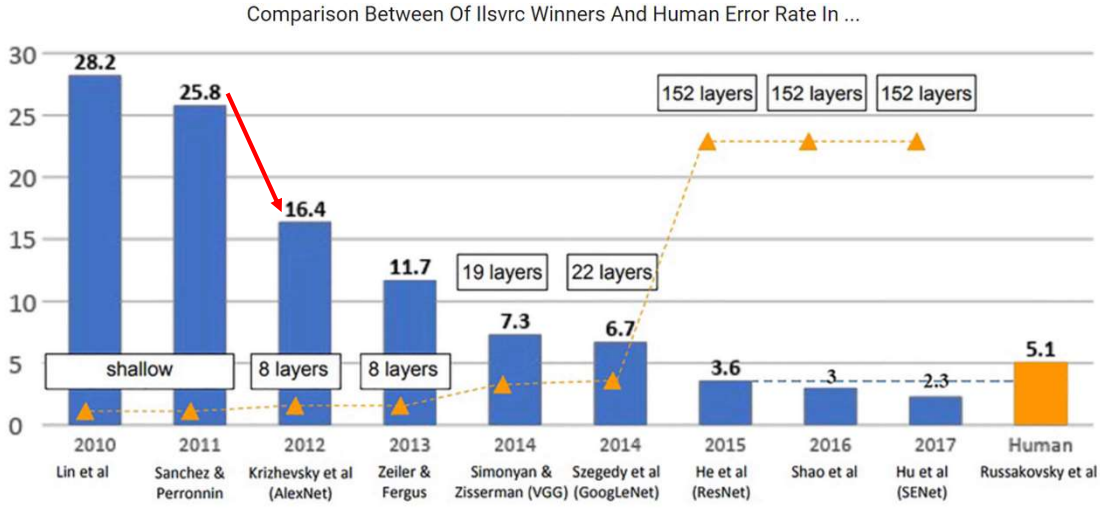
2012: The year of DL breakthrough



- **ImageNet Large Scale Visual Recognition Challenge (ILSVRC):**
  - International Image Classification Challenge
  - 1.4 million images Collected from Flickr and other search engines
  - 1000 object categories
  - Each image was manually labelled by a human being
  - They use a crowdsourcing approach to label the pictures with Amazon Mechanical Turk.
  - Training set: 1.2 million labelled images
  - Test set: 0.2 million labelled images
  - Each team use the training set to "train" their algorithm, then they submit it to the organizers.
  - The organizers evaluate each algorithm on the test set.

2012: The year of DL breakthrough

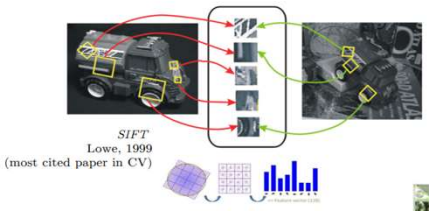




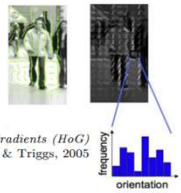
Before 2012

17

Old Paradigm: Input data ➔ Engineer Features ➔ Build Model

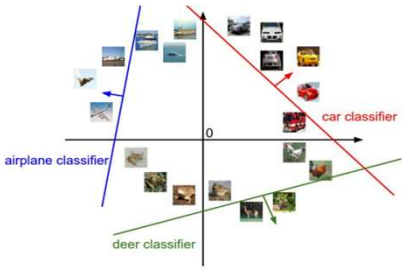


SIFT features



HOG features

⇒



SVM classifier

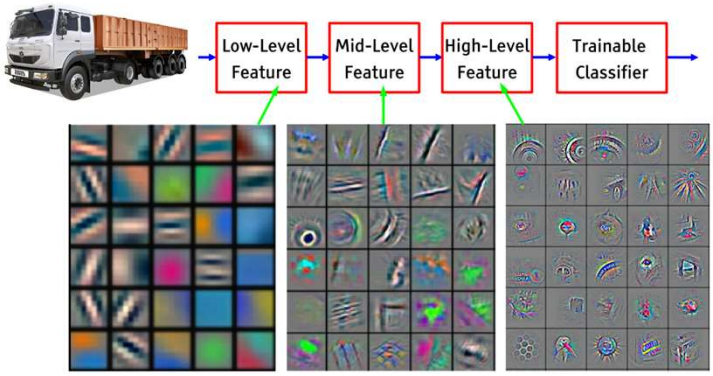
Good Features make good models, but what are good features?
What if one could automatically learn features from the input data?

- Needs domain knowledge
- Time consuming
- Adaptability to changing environment

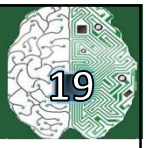
After 2012

18

- **New paradigm: Input data ➔ Automated feature learning ➔ Build Model**
  - Learn data features and classifier together, a.k.a. end-to-end systems.
  - Neural networks are the first class of models that can train end-to-end systems with large learning capacity by using multiple layers.



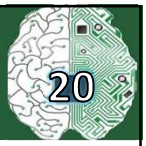
Key successful ingredients



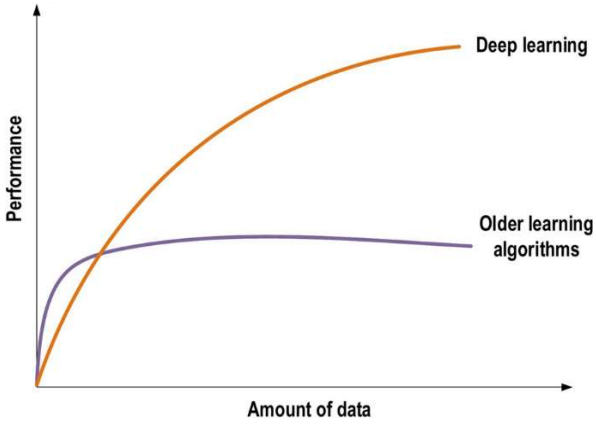
- **Deep Learning success =**

Big data	+	Hardware	+	Software
• Dataset with billion points • Better storage and access		• GPUs, Parallellism • TPUs		• Improved techniques/ algorithms • New tools and models

Why Deep Learning?

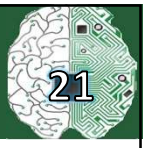


- **Limitations of Conventional ML**
  - Only deep learning systems are able to absorb (exploding) big data
  - Limited in ability to process natural data in raw form
  - Require careful engineering and domain expertise to transform raw data
  - Simple ML algorithms depend heavily on representation of given data



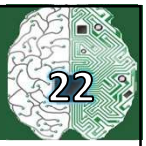
Deep Learning allow a machine to be fed with raw data to automatically discover representations needed for detection or classification

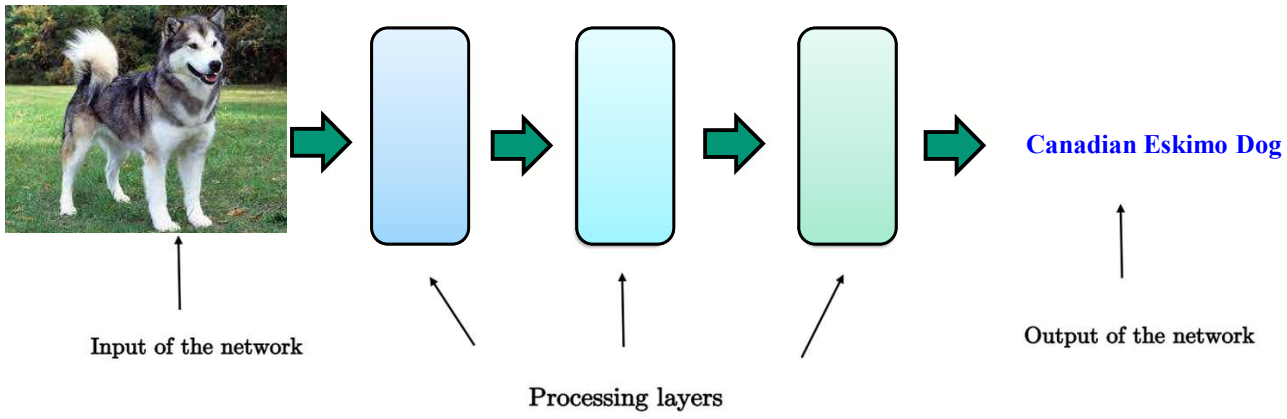
Why Deep Learning?



- DLs are universal learning techniques that can be applied to very distinct fields like computer vision, natural language processing, physics, chemistry, healthcare, biology, etc.
- Tasks where deep network models provide the best performance (SOTA):
  - Image recognition and captioning
  - Object detection
  - Voice recognition
  - Speech synthesis
  - Translation
  - Recommender systems

Image recognition





Courtesy: Prof. Xavier Bresson

Image captioning

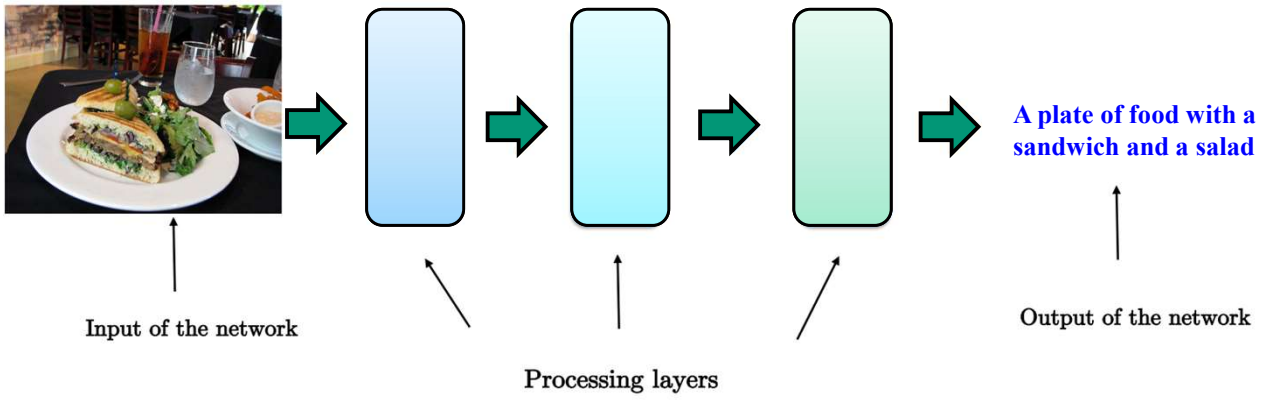


Image captioning (more examples…)

Source: https://github.com/danieljl/keras-image-captioning

Object detection & recognition

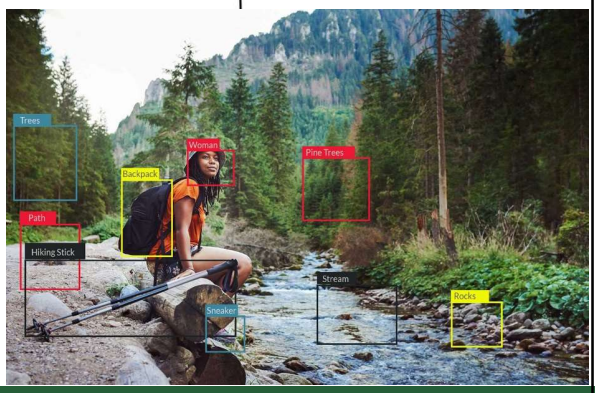
25

Input of the network

Processing layers

Multiple bounding boxes
And their labels

Output of the network



Courtesy: Prof. Xavier Bresson

Voice recognition (speech-to-text)

26

Input of the network

Processing layers

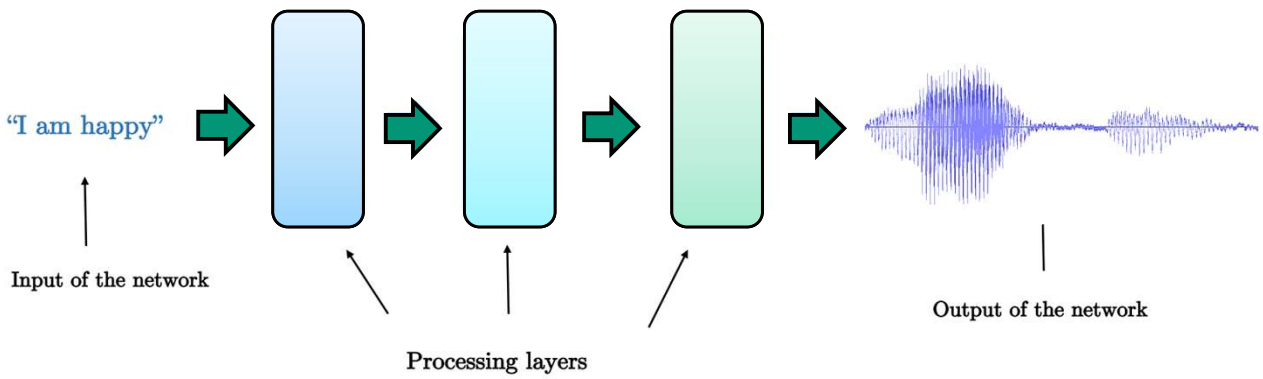
"What is your name?"

Output of the network

Courtesy: Prof. Xavier Bresson

Speech synthesis (text-to-speech)

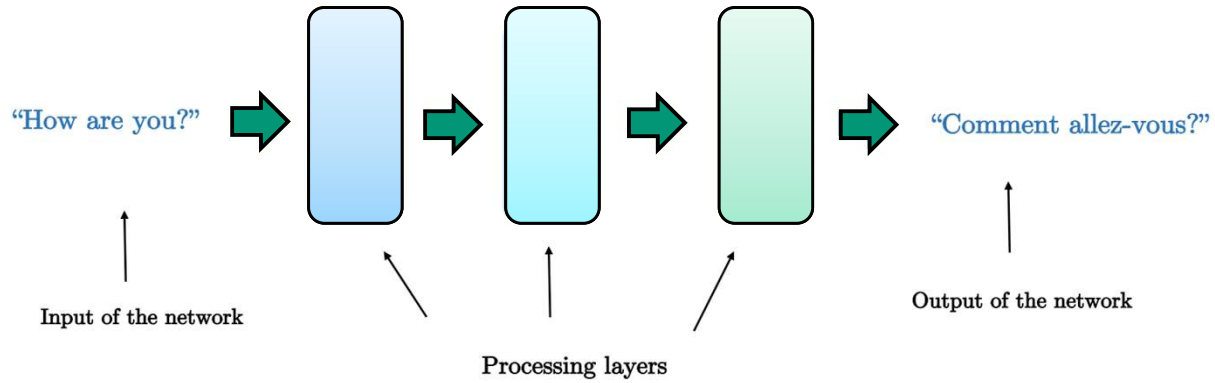
27



Courtesy: Prof. Xavier Bresson

Translation (sequence-to-sequence)

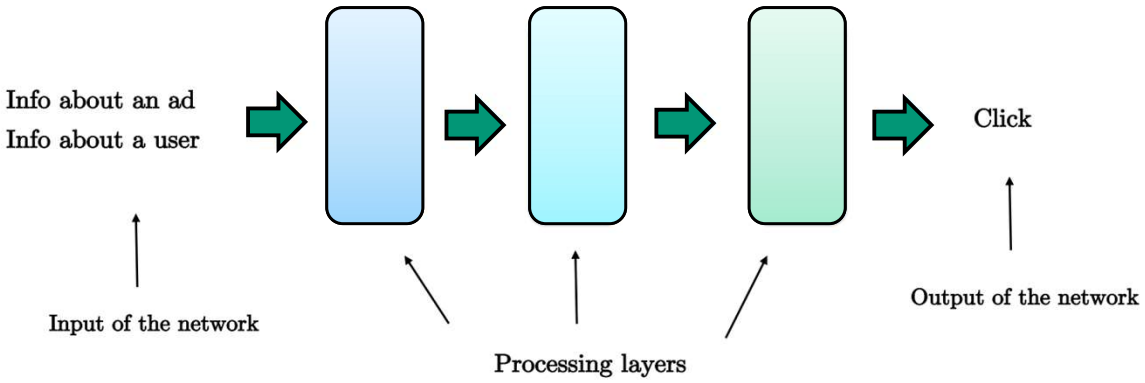
28



Courtesy: Prof. Xavier Bresson

Online advertisement

- Will this user click on this ad?

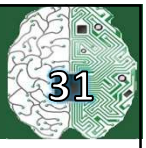


Courtesy: Prof. Xavier Bresson

Course Details

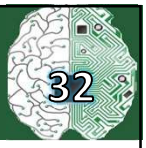
- Course Structure/Syllabus
- Course Materials
- Course Evaluation/Tests

Course Structure/Syllabus



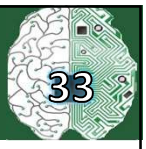
- **Deep Learning**
  - **UNIT-1: Introduction**
    - Course overview
    - Basic concepts in linear algebra, probability and vector calculus
    - Optimization: Types of errors, bias-variance trade-off, overfitting-underfitting
    - Brief review of concepts from optimization, variants of gradient descent, momentum based methods.
  - **UNIT-2: Linear and Logistic Regression**
    - MSE in linear regression, squared loss error from white noise
    - Logistic regression, properties of sigmoid function, cross-entropy error
  - **UNIT-3: Artificial neural network**
    - Basic concepts of artificial neurons, single and multi layer perceptrons, perceptron learning algorithm, its convergence proof, different activation functions, softmax cross entropy loss function

Course Structure/Syllabus



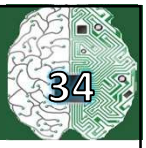
- **Deep Learning**
  - **UNIT-4: Convolutional neural network**
    - Basic concepts of CNN, Famous CNN architectures in computer vision, Regularization, Dropout, Batch normalization, CNN in language processing
    - Applications of CNN: Detection, segmentation problem definition, challenges, Evaluation, Datasets and Localization by regression.
    - Discussion on detection as classification, region proposals, RCNN and YOLO architectures, fully convolutional segmentations, Mask-RCNNs
  - **UNIT-5: NLP basics, Recurrent Architectures, Transformers, NLP Applications, Vision Transformers:**
    - Word and sentence embeddings, Language Models, Discussion on Recurrent Neural Networks (RNNs), Long-Short Term Memory (LSTM) architectures, Attention, Transformers, NLP applications and vision transformers.

Course Structure/Syllabus



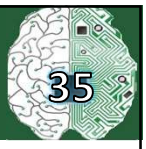
- **Deep Learning**
  - **UNIT-6: Representation Learning**
    - Unsupervised representation learning, Transfer learning, Domain adaptation.
  - **UNIT-7: Structured Probabilistic Models for Deep Learning**
    - Using graphs to describe models, directed and undirected models, energy-based models, restricted Botlzman machine.
  - **UNIT-8: Generative Adversarial models**
    - Stochastic AutoEncoders, Variational Autoencoders, Generative Adversarial Networks.
  - **UNIT-9: Graph Neural Networks**
    - Graph Embeddings, Shallow and deep graph embeddings, Graph neural networks (GNNs), Applications of GNN

Course Materials



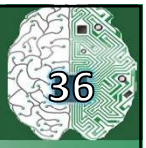
- **Text Book**
  - **Deep learning -** Bengio, Yoshua, Ian J. Goodfellow and Aaron Courville - MIT Press, 2016.
  - **Probabilistic Machine Learning: An Introduction** - Kevin P. Murphy - 1st Edition, MIT Press, 2022
- **Reference Books**
  - **Neural Networks for Pattern Recognition** - Christopher Bishop - Clarendon Press 1st Edition 1996
  - **Neural Networks and Learning Machines** - Simon Haykin, Pearson Education; 3rd Edition 2016
  - **Pattern Recognition and Machine Learning** - Christopher Bishop - Springer, 1st Edition, 2006

Evaluation/Tests



- **Distribution of Marks/Weightage:**

  - ✓ **Quiz-1 Weightage:** 10%
    (Before Mid-Sem; Tentatively on 14-FEB-2024; 07:30AM-8:00AM)
  - ✓ **Quiz-2 Weightage:** 10%
    (After Mid-Sem; Tentatively on 16-APR-2024; 07:00PM-7:30PM)
  - ✓ **Mid-Sem Exam Weightage:** 32%
  - ✓ **End-Sem Exam Weightage:** 48%



Questions?

Acknowledgement:
- Prof. S. N. Srihari
- CSO507 Winter 2022-23 Course Material @iitism