# CS-563 Deep Learning

## Introduction to Deep Learning
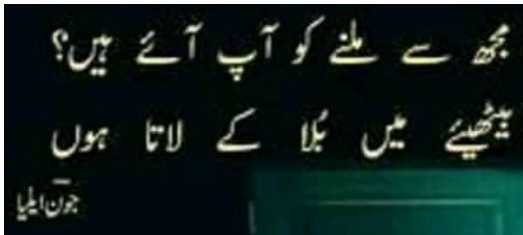
**Nazar Khan**
**Department of Computer Science**
**University of the Punjab**

## Intelligence

Moravec's paradox: The most difficult things to teach a computer are the ones that a two-year old has already learned – talking, listening, seeing, smelling, walking, grasping, memory and recall, thinking.
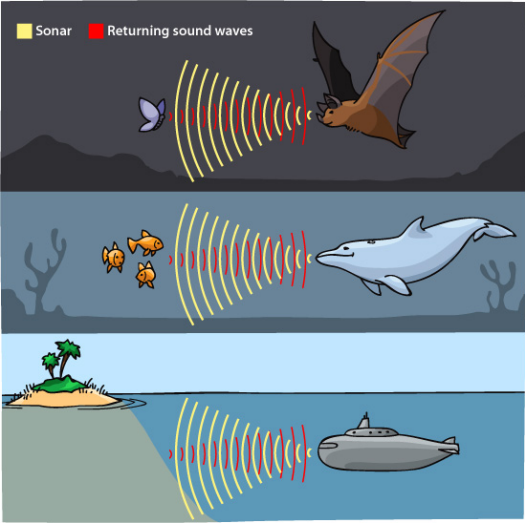
Adult humans have uniquely human attributes — metaphor, poetry, satire, sarcasm.

Birds can fly through very small holes at full-speed *while fighting each other*.

In complete darkness, bats can locate, identify and catch their *flying* prey by sending, receiving and analysing sound waves.

> The ability of biological brains to sense, perceive, analyse and recognise patterns can only be described as stunning.

▶ They also have the ability to learn from new examples with or without being taught.

▶ Mankind's understanding of biological brains and how they operate exactly is embarrassingly limited.

▶ We are clueless regarding the most fundamental questions.

    ▶ What is intelligence?

    ▶ Are you intelligent if you can't make a mistake?

    ▶ Where in our brains does intelligence lie?

    ▶ What is our brain?

    ▶ Are our brains *just* computational devices or do they *do something more*?

    ▶ What is consciousness?

## The Brain

The average human brain has *about* **86 billion neurons** (or nerve cells) and **many more neuroglia** (or glial cells) which serve to support and protect the neurons [and *perhaps* even assist in their functionality]. Each neuron *may* be connected to *up to* **10,000 other neurons**, passing signals to each other via as many as **1,000 trillion synaptic connections**, equivalent *by some estimates* to a **computer with a 1 trillion bit per second processor**. Estimates of the human brain's memory capacity *vary wildly* from **1 to 1,000 terabytes** (for comparison, the 19 million volumes in the US Library of Congress represents about 10 terabytes of data).
Source: https://human-memory.net/brain-neurons-synapses/

- ▶ Claims about the brain are vague.
- ▶ We know *something* about the brain, but we do not know *most* of the crucial functioning.

## So what is this course about?

- Modelling what we do not understand seems foolish.
- However, there do exist numerous *practical* techniques that give machines the *illusion of being intelligent*.
- This is the domain of artificial intelligence, statistical pattern recognition, machine learning and deep learning.
- Instead of attempting to mimic the complex workings of a biological brain, this course
  - aims at explaining mathematically well-founded techniques for analysing patterns and learning from them, and is therefore
  - a *mathematically involved* introduction into the field of pattern recognition and machine learning.
- It will prepare you for further study/research in machine learning, computer vision, natural language processing and others areas attempting to solve AI type problems.

# Prerequisites

- The course is designed to be self-contained.
- *Required mathematical details will be covered in the lectures*.
- However, this is a *math-heavy course*. Students are encouraged[1] to brush up on their knowledge of
  - Probability (Bernoulli, Binomial, Gaussian, Discrete, Continuous)
  - Calculus (Differentiation, Partial derivatives, Chain rule)
  - Linear Algebra (Vectors, Matrices, Dot-product, Orthogonality, Eigenvectors)
- This is also a *code-heavy* course. Be ready to become good at coding.

---

[1]ordered

## Prerequisites

The only way to benefit from this course is to be prepared to *spend lots of hours reading the slides, textbooks, tutorials, and attempting exercises* preferably alone or with a class fellow.

Learn to be mostly alone for the next five months.
Learn to be (reasonably) selfish.
Your social life will be adversely affected. It should be! Time to grow up.
You will need to work harder than ever before.
And even that might not be enough!

## Administrivia

Passing this course with at least a B+ grade is necessary for students planning to undertake researchin the CVML group.

**Course web-page**:
http://faculty.pucit.edu.pk/nazarkhan/teaching/CS563/CS563.html
**Texts and resources**:

- *Deep Learning: Foundations and Concepts* by Chris Bishop and Hugh Bishop https://www.bishopbook.com/
- *Deep Learning* by Ian Goodfellow and Yoshua Bengio and Aaron Courville http://www.deeplearningbook.org/
- *Pattern Recognition and Machine Learning* by Christopher M. Bishop (2006)
- https://pytorch.org/tutorials/

## Administrivia

**Lectures**:
Tuesday and Thursday, 8:15 am – 9:45 pm, AlKhwarizmi Lecture Theater
**Google Classroom**:
https://classroom.google.com/c/ODA1MjAOMTA2NjI4?cjc=v5vss5ir

## Administrivia

**Grading scheme**:

| | |
|---|---|
| Assignments | 35% |
| Project | 15% |
| Quizzes | 5% |
| Programming Tests | 5% |
| Mid | 20% |
| Final | 20% |

# Online Quizzes

- ▶ Online quiz after every 2 lectures.
- ▶ Quiz will *automatically* start at 10 am every Friday.
- ▶ 5 minute duration.
- ▶ Student's responsibility to take the quiz on time. No retakes.
- ▶ No quiz will be dropped. Everything counts.
- ▶ Quizzes will be easy.

## Recitations

- *Extremely important* part of the course.
- Will be conducted every Friday at
  https://meet.google.com/kqu-rbny-acz.
- You will need a reasonably good computer. Contact the instructor or TA
  *as soon as possible* if you don't have access to a good computer.
- Each recitation will be accompanied by a small programming test.
- No test will be dropped. Everything counts.
- Tests will be easy.

## Assignments

- ▶ There will be 5 programming assignments.
- ▶ You are encouraged to help each other but not to cheat.
- ▶ *Do not cheat.*
- ▶ No assignment will be dropped. Everything counts.

# Project

- Solve an interesting problem via Deep Learning or implement a research paper.
- Recorded video presentation.
- Place code on Github.
- Create web-page containing everything about the project.
- Report prepared in LaTeX.
- Templates will be provided to help you out.

## Programming Environment

We will be using

- ▶ Python as our programming environment.
- ▶ PyTorch as our deep learning framework.
- ▶ Google Colaboratory as our GPU enabled machine on the cloud.
- ▶ Jupyter notebooks as our interactive tutorials.

# Deep Learning vs. Machine Learning vs. AI
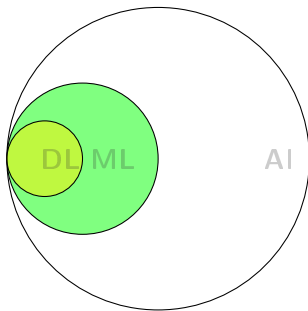


ML and AI problems are increasingly being solved using DL.
Laymen have started considering DL to be the same as AI.

AI: software that solves problems by itself.
ML: algorithms and models that *learn* from processed data.
DL: *neural networks* that *learn better* from *less processed* data.

## Introduction

▶ Machine Learning is concerned with automatic discovery of regularities in data.

▶ Regularity implies order.

▶ Learning implies exploiting order to make predictions.

| Input | Output |
|-------|--------|
| 13 | 26 |
| -4 | -8 |
| 3 | 6 |
| 6 | 12 |
| 10 | 20 |
| -23 | -46 |
| 47.6813 | ? |
| -1.3 million | ? |

## Machine Learning
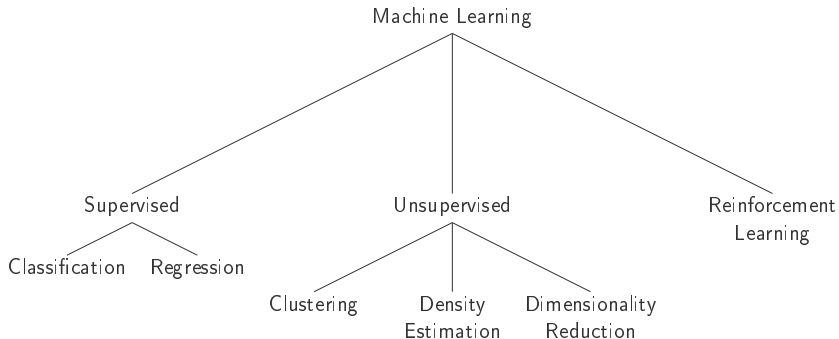
# Supervised Learning

- **Classification**: Assign **x** to *discrete* categories.
  - Examples: Digit recognition, face recognition, *etc.*.
- **Regression**: Find *continuous* values for **x**.
  - Examples: Price prediction, profit prediction.
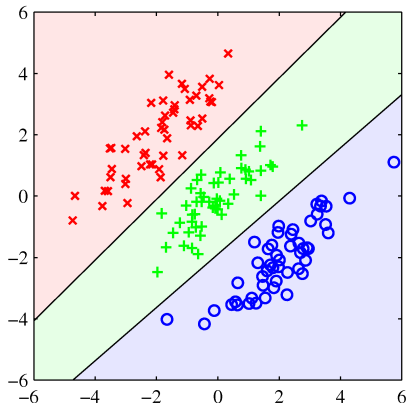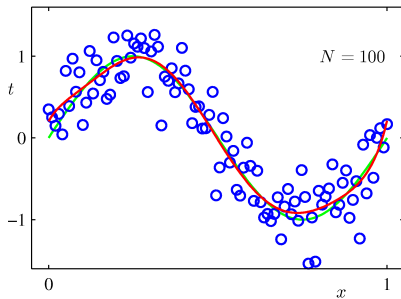


Classification                                                    Regression

# Unsupervised Learning

- **Clustering**: Discover groups of similar examples.
- **Density Estimation**: Determine probability distribution of data.
- **Dimensionality Reduction**: Map data to a lower dimensional space.
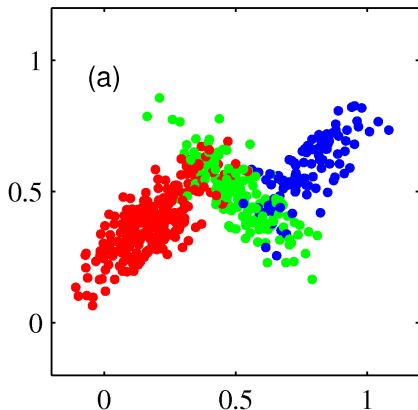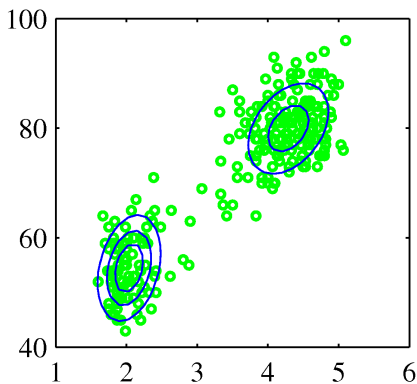


Clustering                                         Density Estimation

# Reinforcement Learning

▶ Find actions that maximise a reward within an environment.
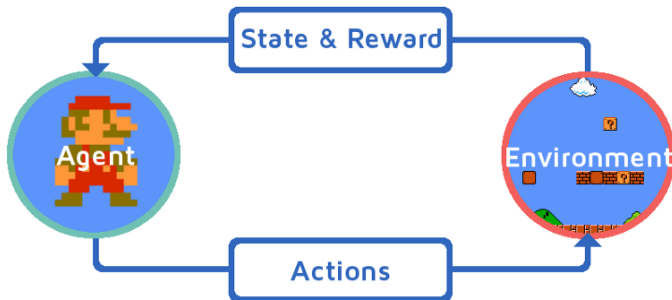


**Figure:** Based on the current state of the game (environment), each action of the player changes the state and yields a reward – points or death. The player learns to reinforce taking actions that lead to positive reward and not taking actions that lead to negative reward. Source: https://www.freecodecamp.org/news/a-brief-introduction-to-reinforcement-learning-7799af5840db/

# Applications of Deep Learning



https://doi.org/10.1145/3234150

# DL Applications in CVML Group
*Recognition of handwritten Urdu text*[2][3]



CRR: 100.00, WRR: 100.00

---

[2] Tayaba Anjum and Nazar Khan. 'An attention based method for offline handwritten Urdu text recognition'. In: *International Conference on Frontiers in Handwriting Recognition (ICFHR)*. 2020.

[3] Tayaba Anjum and Nazar Khan. 'CALText: Contextual Attention Localization for Offline Handwritten Text'. In: *Neural Processing Letters* (2023). URL: https://doi.org/10.1007/s11063-023-11258-5.

# DL Applications in CVML Group
*Handwritten Text Line Extraction*[4]
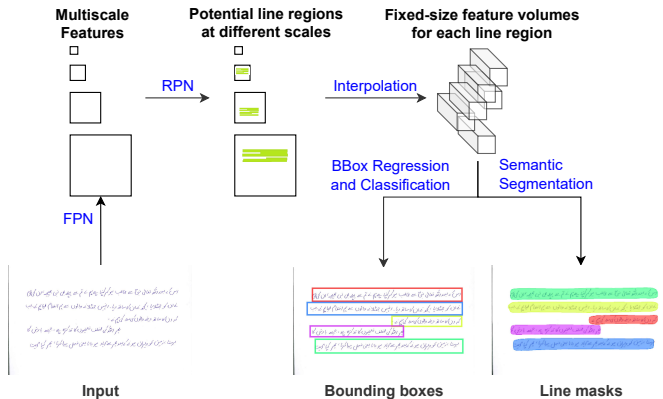


[4] Adeela Islam, Tayaba Anjum, and Nazar Khan. 'Line Extraction in Handwritten Documents via Instance Segmentation'. In: *International Journal on Document Analysis and Recognition* (2023).
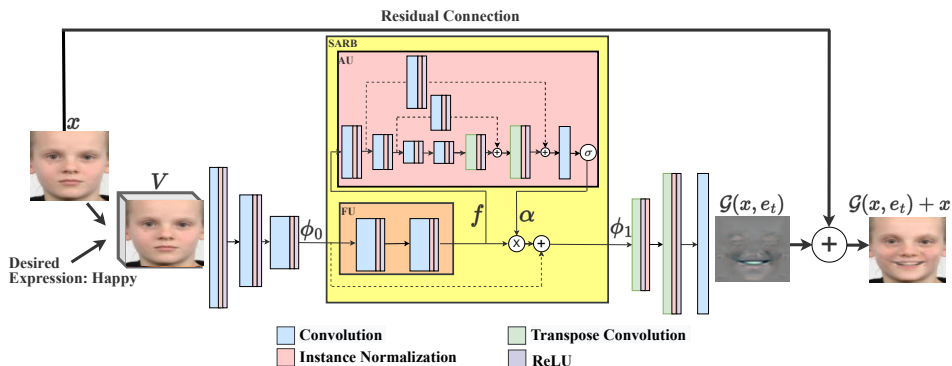
# DL Applications in CVML Group
*Handwritten Text Line Extraction*[5]

| English | Arabic | German | Bengali | Chinese | French | Hindi | Japanese |
|---------|--------|--------|---------|---------|--------|-------|----------|



---

[5] Adeela Islam, Tayaba Anjum, and Nazar Khan. 'Line Extraction in Handwritten Documents via Instance Segmentation'. In: *International Journal on Document Analysis and Recognition* (2023).

# DL Applications in CVML Group
*Handwritten Text Line Extraction[6]*



| Input image | Extracted text lines | Recognized text | CER | WER |
|---|---|---|---|---|

---

[6] Adeela Islam, Tayaba Anjum, and Nazar Khan. 'Line Extraction in Handwritten Documents via Instance Segmentation'. In: *International Journal on Document Analysis and Recognition* (2023).

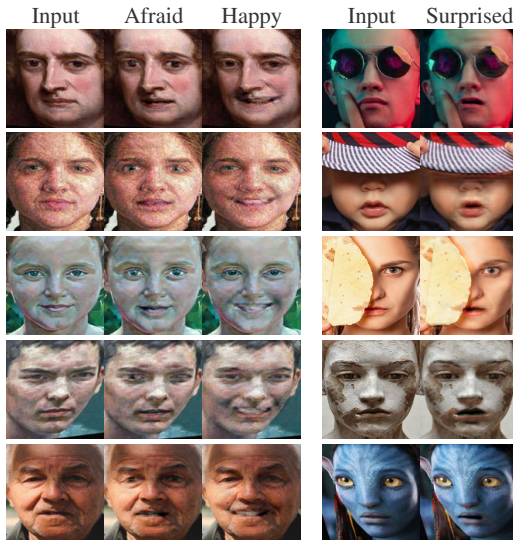# DL Applications in CVML Group
*Facial Expression Synthesis*[78]



[7] Arbish Akram and Nazar Khan. 'US-GAN: On the importance of Ultimate Skip Connection for Facial Expression Synthesis'. In: *Multimedia Tools and Applications* (2023).

[8] Arbish Akram and Nazar Khan. 'SARGAN: Spatial Attention-based Residuals for Facial Expression Manipulation'. In: *IEEE Transactions on Circuits and Systems for Video Technology* (2023), pp. 1–1. DOI: 10.1109/TCSVT.2023.3255243.
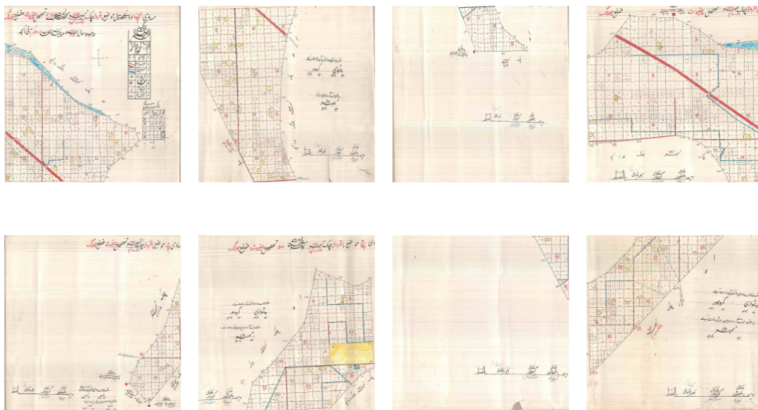
# DL Applications in CVML Group
*Facial Expression Synthesis*
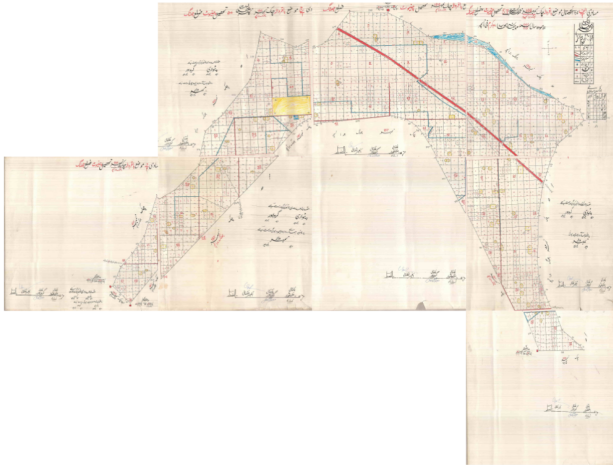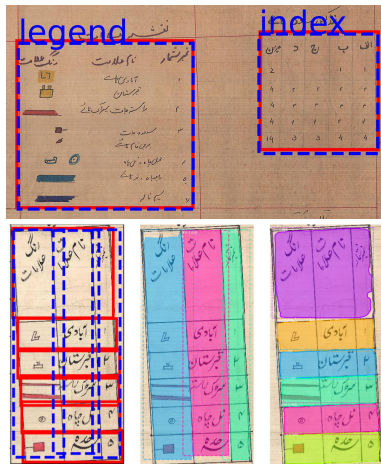
# DL Applications in CVML Group
*Learning to solve jigsaw puzzles*[9]



[9]Adeela Islam. 'Learning to Solve Jigsaw Puzzles'. MPhil Thesis. University of the Punjab, 2020.

# DL Applications in CVML Group
*Learning to solve jigsaw puzzles[10]*



---

[10] Adeela Islam. 'Learning to Solve Jigsaw Puzzles'. MPhil Thesis. University of the Punjab, 2020.

# DL Applications in CVML Group
*Table detection and understanding*[11]



---

[11] Asmat Batool. 'Detection and Recognition of Tabular Structures in Historical Documents'. MPhil Thesis. University of the Punjab, 2020.
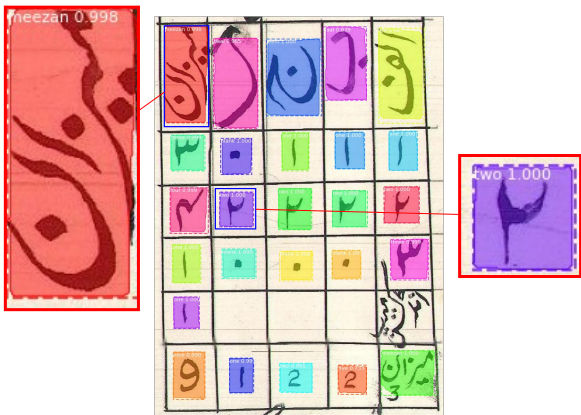
# DL Applications in CVML Group
*Table detection and understanding*[12]



---
[12] Abubakar Siddique. 'Detection, Recognition, and Spotting of Hand-drawn Metadata in Historical Maps'. MPhil Thesis. University of the Punjab, 2022.

# DL Applications in CVML Group
*Table detection and understanding*[13]

[13] Abubakar Siddique. 'Detection, Recognition, and Spotting of Hand-drawn Metadata in Historical Maps'. MPhil Thesis. University of the Punjab, 2022.

## Summary

- ▶ DL represents the future of AI.
- ▶ The hype is real.
- ▶ Learning from data is now considered an automatic choice for solving problems.
- ▶ Ignoring DL is not an option.
- ▶ Next lecture: a gentle introduction to neurons and neural computation.