Handling a cancer problems with Machine Learning.

Presented by: WPSYG
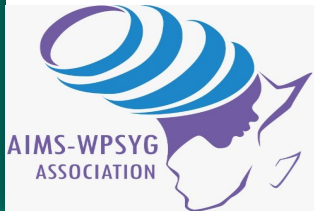
AIMS-WPSYG ASSOCIATION
Women Promoting Science to the Younger Generation

May12
Celebrating Women in Mathematics

Presentation plan:

❖ Context and motivation

❖ Theory

❖ Practical Application

CONTEXT AND MOTIVATION

Context

- Cancer is the second leading cause of death globally
- Breast cancer is the second most common cancer among women in Sub-Saharan Africa,
- Breast cancer could be successfully treated if detected early

Motivation

- ★ Presence of noise in images
- ★ Radiologist's visual perception ability
- ★ inadequate clarity, poor contrast, and the less experienced radiologist
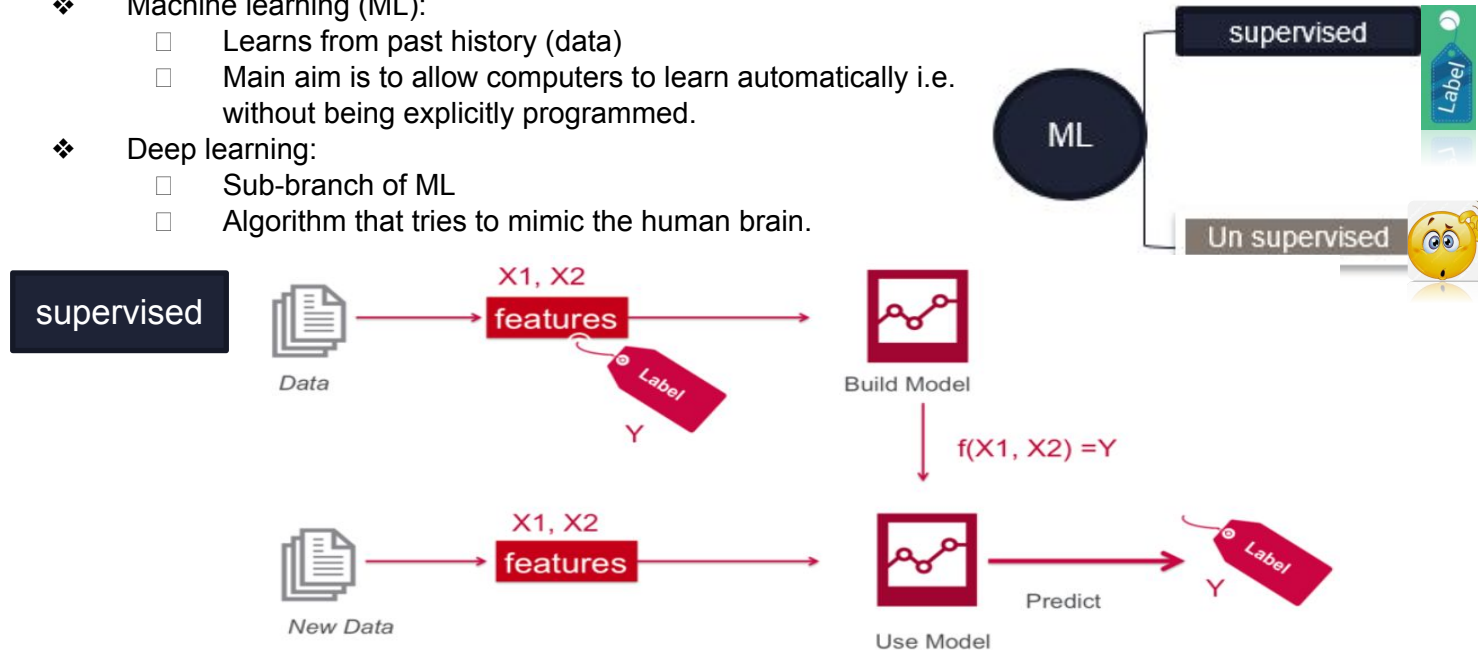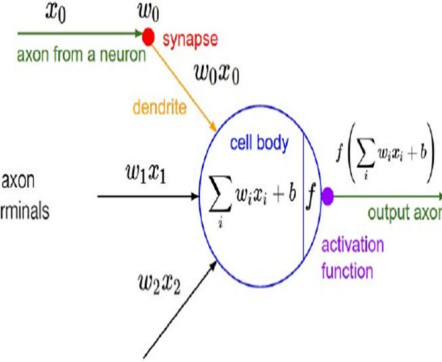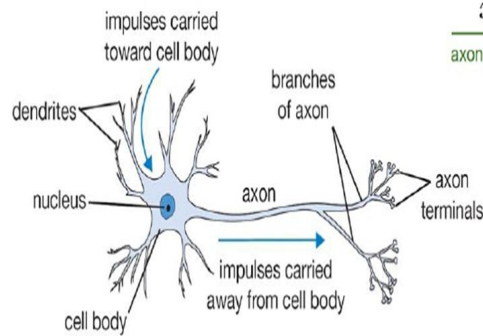
THEORY

MACHINE LEARNING

❖ Machine learning (ML):
  - Learns from past history (data)
  - Main aim is to allow computers to learn automatically i.e. without being explicitly programmed.
❖ Deep learning:
  - Sub-branch of ML
  - Algorithm that tries to mimic the human brain.

"Machine Learning is the science of getting computers to learn and act like humans do, and improve their learning over time in autonomous fashion, by feeding them data and information in the form of observations and real-world interactions."
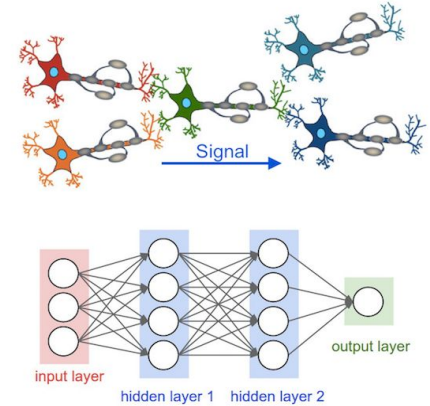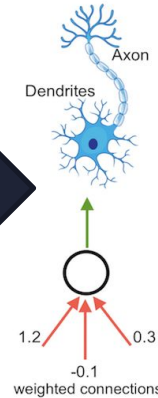*Ref:*https://emerj.com/ai-glossary-terms/what-is-machine-learning/

DEEP LEARNING DEMYSTIFIED

Ref1: https://greydanus.github.io/2016/08/05/what-is/

Ref2: https://opensourceforu.com/2017/11/a-quick-look-at-image-processing-with-deep-learning/
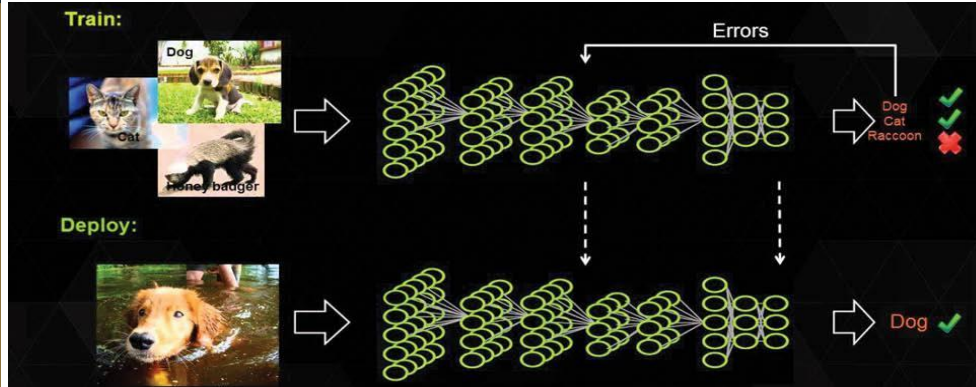
Advantages:
- ✔ Best at image processing than humans.
- ✔ Best at Alpha Go.

Challenges:
- ☐ Need huge amount of data.
- ☐ Computationally expensive.

Literature review

- **Classification of images using 2D CNN.**

  - To make good use of the advances in computer vision and image processing.

  - Challenges: Computationally expensive.

- **Using image classification for cancer detection:**

  - Medical diagnosis is one of the most important area in which image processing procedures are usefully applied due to its increase in diagnosis accuracy.

  - The work of Nan wu et al propose a DNN to improve radiologist performance in breast cancer screening(1).

1)   Ref: https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=8861376

Invasive ductal carcinoma (**IDC) detection**
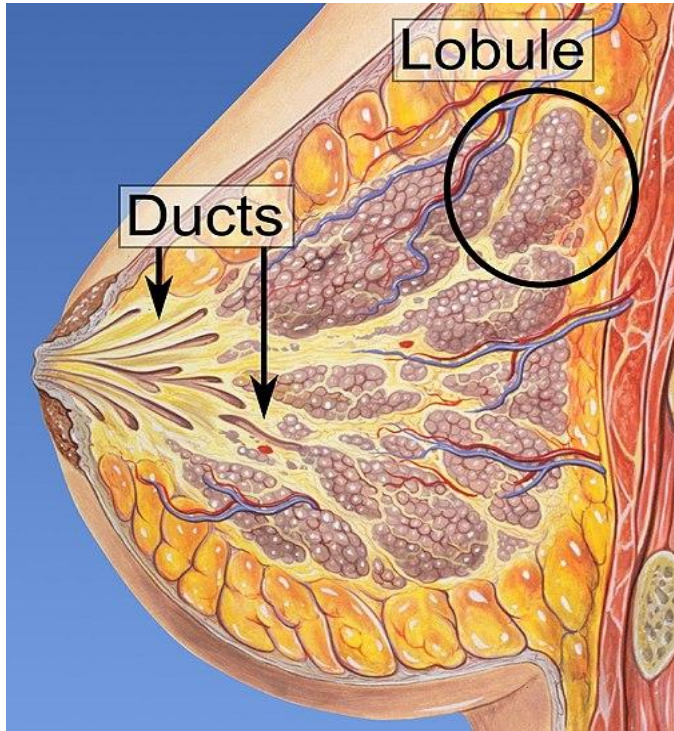
- IDC is - with ~ 80 % of cases - one of the most common types of breast cancer

- Often a biopsy is done to remove small tissue samples

- In addition sick cells need to be located to find out how advanced the disease is and which grade should be assigned.

- This has to be done manually and is a time consuming process.

- Furthermore the decision depends on the expertise of the pathologist and his or her equipment

Ref: https://commons.wikimedia.org/wiki/File:Lobules_and_ducts_of_the_breast.jpg

Concept of transfer learning

- **Pre-training:** The network is generally trained on a large-scale benchmark dataset representing a wide range of categories.
- **Fine-tuning:** Pre-trained network is further trained on the specific target task of interest, which may have fewer labelled examples than the pre-training dataset.

Ref: https://www.sciencedirect.com/science/article/abs/pii/S0167865520300891
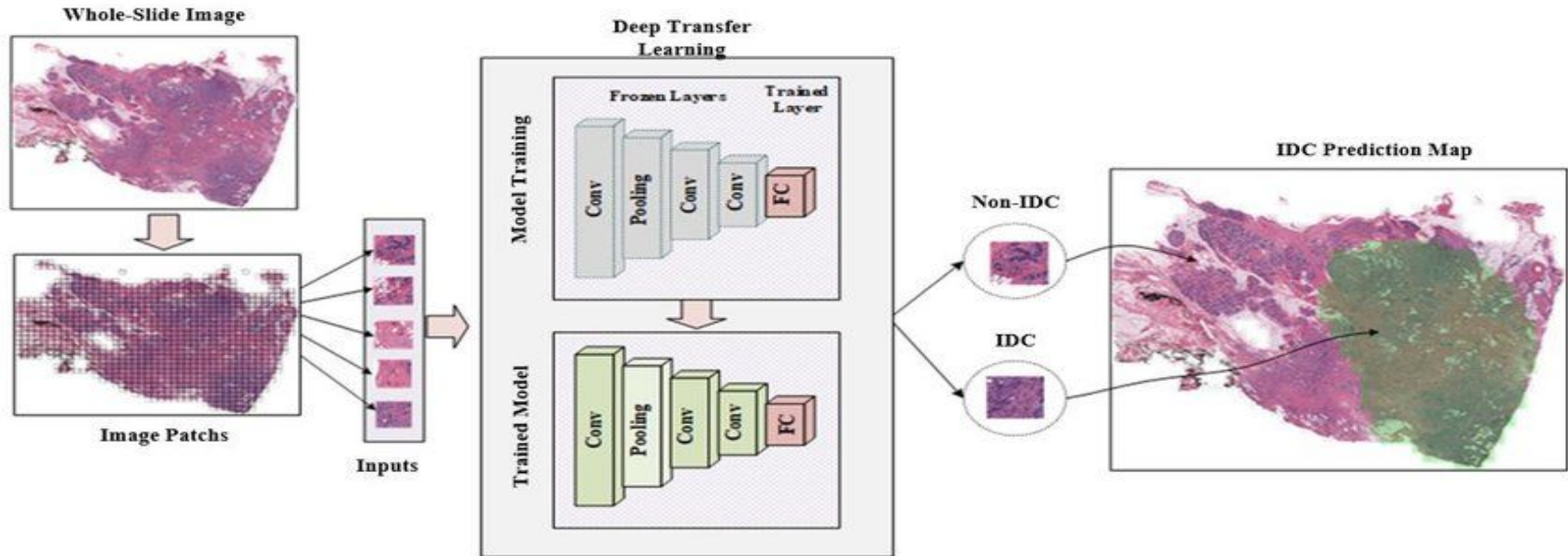
PRACTICAL APPLICATION

Enough of theory, let's move to a concrete example in the notebook.