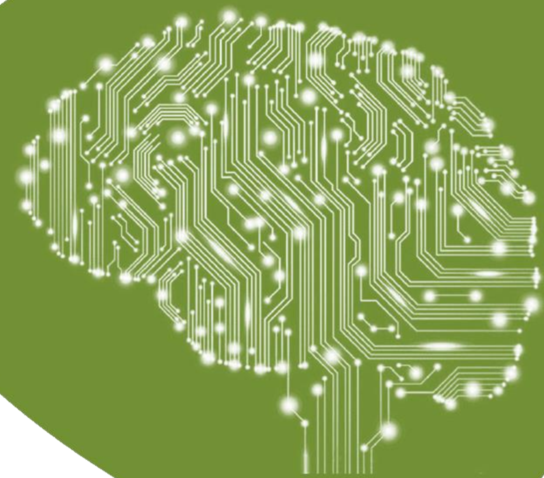
Deep Learning – white paper

August 2024

saiwa

simple artificial intelligence web application



Deep learning refers to a sub-category of machine learning methods based on artificial neural networks that mimic the way human beings acquire certain types of knowledge. In deep learning high-level features are extracted *directly* from the raw input data without AI expert involvement. Figure 1 shows artificial intelligence, machine learning and deep learning relations.

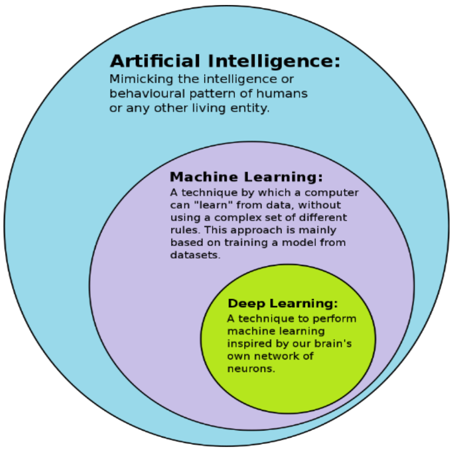


Figure 1. Scheme of the relationship between artificial intelligence, machine learning and deep learning (printed from [1])

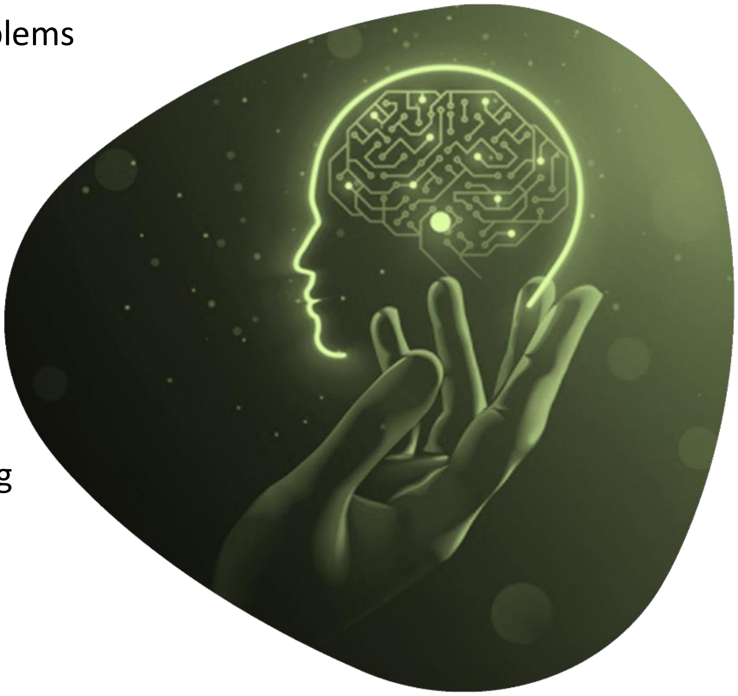
DEEP LEARNING

Compared to classic machine learning methods, deep learning has both advantages and disadvantages and so both methods are widely used in different applications. The primary advantages of deep learning include:

- ✓ Automatic extraction and optimization of features
- ✓ Applicable to several applications and data types
- ✓ Flexible and easily adaptable to new problems
- ✓ High validation speed

and the main disadvantages are:

- ✗ Very large dataset requirement
- ✗ High computational complexity of training



Different types of deep learning models include: convolutional neural networks (CNNs), autoencoders, deep belief networks, recurrent neural networks and reinforcement learning [2]. In saiwa deep learning service three CNNs are provided for training on user specific dataset (i.e. **Detectron2** and **Yolov5** and **Yolov7**). Therefore, we will explain this type of deep learning models in more details in the next section. For information on Detectron2 and Yolov5 and yolov7 please consult the saiwa object detection white paper.

Convolutional neural networks (CNNs) are among the most successful and commonly used architectures in deep learning applications. They have been extensively applied in different applications such as NLP, speech processing, and computer vision e.g. medical image segmentation, anomaly detection, and robot vision [3]. CNNs mainly consist of four type of layers:

1. Convolutional layers
2. Nonlinear (ReLU) layers
3. Pooling layers
4. Fully connected layers

Convolutional layers are fundamental for feature extraction. ReLU (rectified linear units) layers apply an activation function on feature maps (results of convolutional layers); and then in the pooling (subsampling) layers, feature maps are downsampled to decrease the network parameters, speed-up the training process, and also control overfitting. Finally, for classification problems the CNNs employ fully connected layers. Figure 2 shows the overall architecture of a typical CNN.

Some of the most well-known CNN architectures include: AlexNet, VGGNet, ResNet, GoogLeNet, MobileNet, DenseNet, and YOLO. For detailed information on these CNNs, please refer to [3].

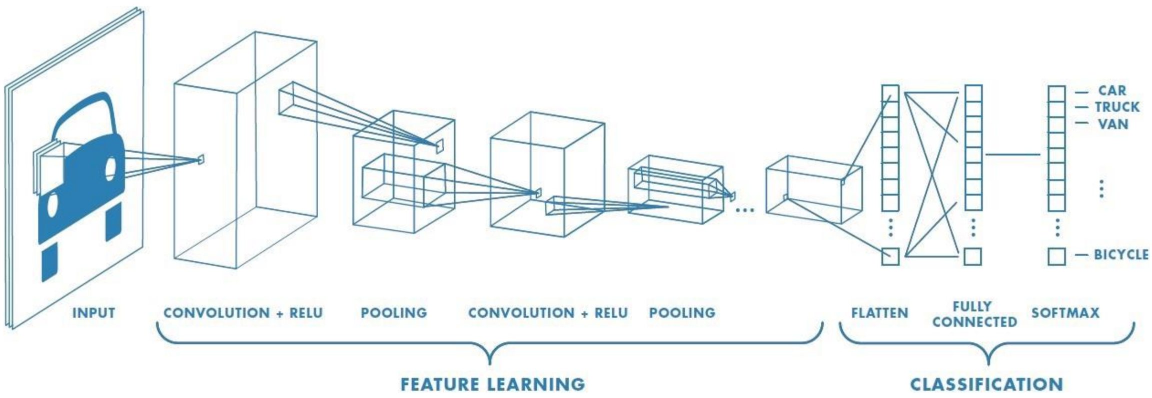


Figure 2. Overall architecture of a typical CNN with four types of layers (printed from [5])

References:

[1] https://en.wikipedia.org/wiki/Deeplearning

[2] Pouyanfar, Samira, et al. "A survey on deep learning: Algorithms, techniques, and applications." ACM Computing Surveys (CSUR) 51.5 (2018): 1-36.

[3] Li, Zewen, et al. "A survey of convolutional neural networks: analysis, applications, and prospects." IEEE Transactions on Neural Networks and Learning Systems (2021).

[4] Minaee, Shervin, et al. "Image segmentation using deep learning: A survey." IEEE transactions on pattern analysis and machine intelligence (2021).

[5] https://towardsdatascience.com/a-comprehensive-guide-to-convolutional-neural-networks-the-eli5-way-3bd2b1164a53

info@ saiwa.ai

720 Guelph Line Burlington, ON L7R 4E2 Building
+15148131809
www.linkedin.com/company/saiwa
www.instagram.com/saiwa.ai