Informationssysteme 5. Jahrgang Skriptum

Machine Learning Basics v1.0

Nikolaus Hofer

Contents

Introduction

What is Machine Learning?

Machine learning (ML), artificial intelligence (AI), deep learning (DL) and Neural Networks (NN) all those terms are commonly used interchangeably but how are those terms defined and how are they differentiated?

> Field of study that gives computers the ability to learn without being explicitly programmed.
> - *Arthur Samuel, 1959*

> 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 M. Mitchell, 1997*

Those definitions while already quite old still apply today and show that ML learns from data. This leads to the conclusion that without any data there is no machine learning.

To really understand the difference between the terms we first need to understand the individual terms.

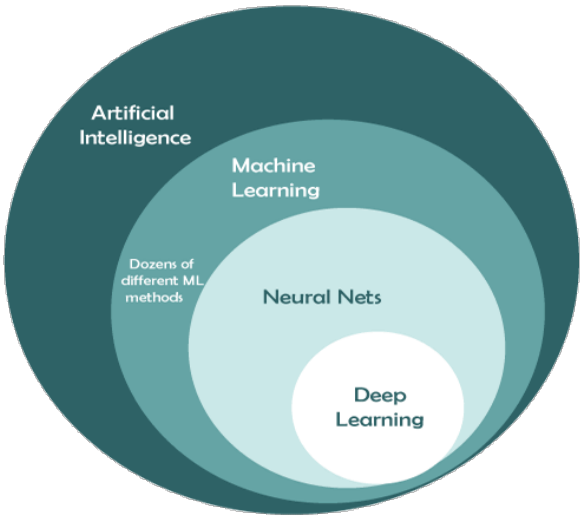

Figure 1: Difference between ML, AI, DL, NN (javatpoint.com)

For all those terms shown there are different definitions available. Here are some of them to make the differentiation easier:

Artificial Intelligence

> It is the science and engineering of making intelligent machines, especially intelligent computer programs. - *John McCarthy, 2004*

AI tries to make computers/machines imitate human intelligent behavior. One example for AI is an E-Mail Spam-Filter which tries to filter incoming E-Mails using different characteristics. (e.g. sender address, combination of words in the subject) This artificial intelligence can be achieved with traditional algorithmic programming.

Machine Learning

> Field of study that gives computers the ability to learn without explicitly being programmed. - *Arthur Samuel, 1959*

Machine Learning is a subfield of Artificial Intelligence still trying to imitate human behavior.

Data

ML is not programmed in the traditional sense but learns based on provided data. The ML algorithms used try to search for patterns and correlations in the data provided. This step is called **training** and results in a so-called **model**.

The data is split into samples and features, each sample has a number of defining features. Using the spam filter example: each E-Mail is a sample and the sender, number of words, size, attachments etc. are features defining each sample.

The features have to be selected by hand and they are very domain specific.

Goal

The resulting model should be able to handle previously unknown data and create as generic correlations as possible. So when a spam filter receives a similar E-Mail with one or two words exchanged (e.g. the Name of the recipient) it should still detect a spam E-Mail.

Main groups of ML-algorithms

There are three main groups of ML-algorithms:

1. **Supervised:** uses prerecorded datasets which are separated into input and an expected output. This expected output (also called labels) has to be prerecorded or labeled by a human.
    1. **Classification:** A common example of supervised ML is classification which tries to assign a class to the provided input data. Classification models usually have integer outputs representing the different classes. The E-Mail spam filter from the previous example could be one application for a supervised ML Model that is trained on E-Mails a user has already marked as Spam. The difference to the AI approach mentioned before is the lack of figuring out the criteria yourself to define the filtering rules. Other common applications are: image segmentation, medical diagnosis, error prediction
    2. **Regression:** By contrast to the classification, regression has a floating point value as an output. This can be used to predict numerical values e.g. the price of a car, or the failure rate of a specific component.
2. **Unsupervised:** has no specific label as a goal but tries to find correlations in the data. The big advantage over supervised learning is, that no labels are needed to train the model. Finding those patterns in the data can help with the clustering/association of different data points. Examples: anomaly detection, recommendation systems, network analysis *There are also mixed approaches*

like semi-supervised machine learning which uses mostly unlabelled data but needs some labeled data for training

3. **Reinforcement Learning:** uses a feedback-method to learn from past experience. The ML-model tries out different approaches and depending on the outcome of those approaches the model will get a reward. Gaming, navigation and robotics are common tasks for reinforcement learning algorithms. (A.I. Learns to Drive From Scratch in Trackmania - YouTube)

This course focuses on machine learning.

Neural Networks

Artificial neural networks try to imitate the function of the human brain using mathematical functions. The human brain consists of so-called neurons which are interconnected.

The artificial neuron combines multiple numerical weighted inputs to a single output value. When combining multiple artificial neurons an artificial neural network is formed.

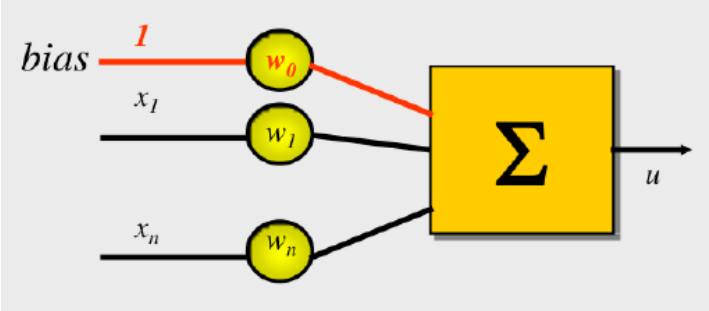

Figure 2: Single artificial neuron (Witold Jacak, 2012)

Neural networks have a big advantage over traditional machine learning algorithms: There is no need to pre-select and pre-process the features used as input. The structure and learning approach used for NN can extract complex correlations in the data.

Tryout Neural Networks for yourself: Tensorflow Playground

More information on Neural Networks will be discussed in future chapters.

Deep Learning

Neural networks themselves don't have to be deep learning algorithms if they only combine a small number of neurons. The deep in deep learning stands for the number of layers in a neural network. So the more layers a NN has the deeper it is.

With the increased depth and complexity of a neural network the more complex, the problem can be which has to be solved. A really complex problem which is a common task for deep neural networks is image classification. To be able to distinguish a tiger and a striped cat the neural network has to learn the facial features or overall appearance of a tiger from multiple different angles.

Common problems solved with deep learning: image recognition, natural language processing (NLP), handwriting recognition, translation, large language models (LLM, e.g. ChatGPT)

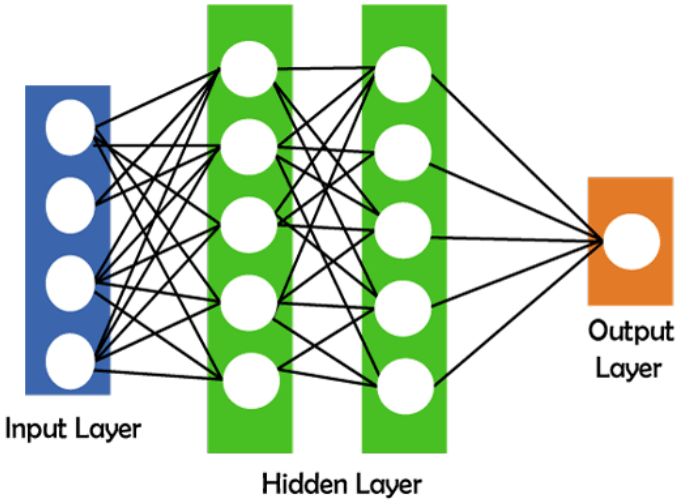

Figure 3: Layered artificial neural networks (javatpoint.com)