Deep Learning with Perception

Jawwad Shamsi

September 12, 2021

1 Fundamentals of Machine Learning

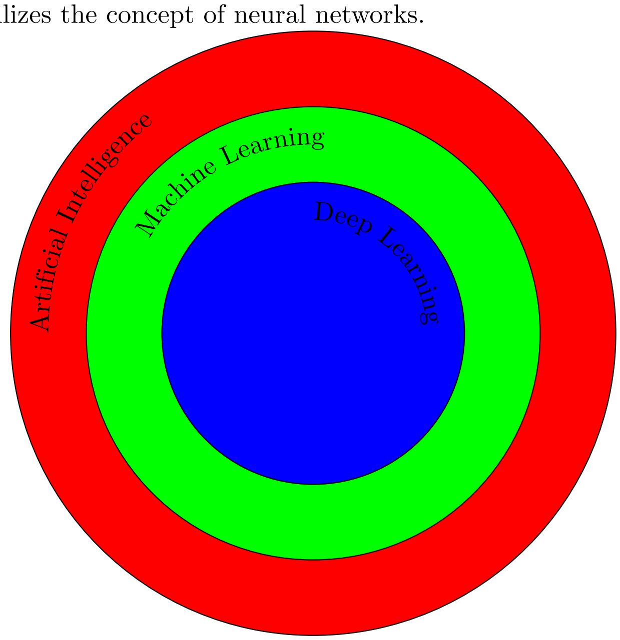
In the first week, we will study fundamentals of Machine Leanring

1.1 Difference between AI vs. ML vs. DL vs. DS

AI: a concept to create intelligent machines that can simulate human thinking capability and behavior.

ML is a subset of AI, which provides us statistical tools to explore the data and gives ability to learn. It is a field of study that gives computers the ability to learn without being explicitly programmed.

DL is a subset of ML, which alleviate the need of feature engineering. It utilizes the concept of neural networks.



1.2 Fundamental Concepts of Machine Learning

We would like to have a program which can learn from experience.

The output of Machine learning would be a program

Machine Learning is a field of study that gives computers the ability to learn without being explicitly programmed (Arthur Samuel, 1959).

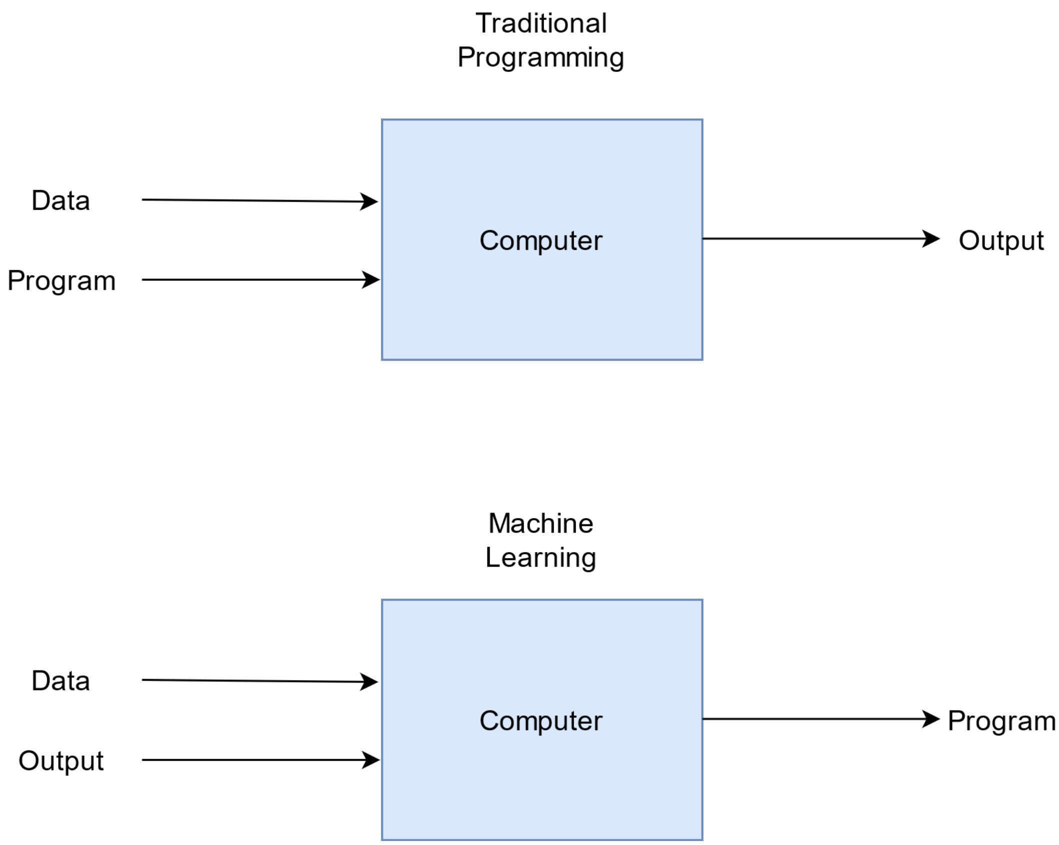


Figure 1: ML vs Prog

1.3 Applications of Machine Learning

Natural Language Processing, Computer Vision, Computational Biology, Computer Vision. Recommender systems by Netlfix, Amazon, Google ads. Alphago by Google Drug Discovery Character recognition by Post Office Assisted Driving Cancer Detection

1.4 Supervised learning

Given a set of features, find the rule which will predicts the label for unseen data Types of Supervised Learning

1. Classification : Predict the class
2. Regression: Predict the value

Example of Classification : Predict grades based on no. of hours studied Example of Regression: Predict price of land according to area.

1.5 clustering or unsupervised learning

Data is unlabelled. Lets group them based on features and see the pattern.

K-Means = A popular clustering algorithm Given a set of unlabelled feature vectors group them into clusters.

begin

1. Suppose there are K clusters
2. pick k items
3. cluster remaining samples by minimizing distance between the data points in the same cluster
4. Find median point

Let's suppose we have a few data points and we arranged them in to two clusters. Later, we have two new data points, one of them fits well into one of the cluster, whereas the other one is at the edge of the circle. In such a case, we have two options. First, either to re-cluster all the data points or to have more than two clusters.

1.6 Data

Some questions:

1. What is the training data and how are we going to evaluate it
2. what is the training data and how are we going to evaluate the success of the system
3. How do we figure what are the right features and how do we measure distances between them. How to find similarities?

Data can be split into training and testing

1.6.1 training data

Observe set of examples. Infer something about the data and generate the model

1.6.2 validation data

: Used inference model to make predictions about the data. Inference on labelling things

1.7 Features

Features are attributes about data point. Selection of right features are very important in Machine learning.

"All models are wrong but some are useful". George Box

Suppose I would like to predict grades in this course. Features such as prior programming experience, existing CGPA, no. of lectures attended are important. some other features such as color of eyes or residential address may seem irrelevant and may lead to overfitting.

We can input all the features and sorts them to find out which are useful. However, it may lead to overfitting. So goal is to maximize the features which are likely to be useful and minimize those which are not useful.

Let's take an example:

Table 1: Feature Selection Example

Name	Egg-Laying	Scales	Poisonous	Cold-blooded	no. of legs	reptile
Cobra	True	True	True	True	0	Yes
Rattlesnake	True	True	True	True	0	Yes
Boa Constrictor	False	True	False	True	0	Yes
Chicken	True	True	False	False	2	No
Alligator	True	True	False	True	4	Yes
Dart Frog	True	False	True	False	4	No
Salmon	True	True	False	True	0	No
Python	True	True	False	True	0	Yes

Consider the above table. For the first row, we start by selecting all the five features. These are highlighted in red. This continued for row 2. In row 3, we have a different case. Bona Constrictor does not lay eggs, it is not poisonous and it is a reptile. So we refined the selected features to three. In row 5, we refined the selected features by changing the requirement of no. of legs to either 0 or 4. The problem arises in row 7 and row 8. They both have similar features but the response variable is different. We consider this as an anomaly or outlier. In the final selection, we have reduced the no. of features to two, i.e. scales and Cold-blooded. Design Choice: NO False negatives. That is, no instance, which is a reptile and we call it not a reptile. We may have false positives, i.e., an instance that is not a reptile and we call it as a reptile (example:salmon).

1) Features 2)Distance will be used to how to group items or how to separate them a part. 3) Relative weights of different dimensions.

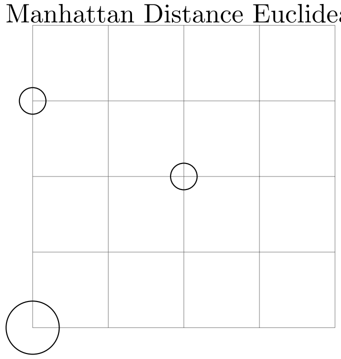
Another way to separate reptiles and non-reptiles is to measure the distance between different data points.

A simple way is to convert them into feature vectors. Goal is to either classify them through a divider or to cluster them accordingly.

Cobra [1,1,1,1,0] rattlesnake[1,1,1,1,0] boa constrictor[0,1,0,1,0] dart frog[1,0,1,0,4]

How to measure distance between these feature vectors?

Manhattan Distance Euclidean Distance



Feature Engineering: choice of features matters too many features may lead to overfitting weights may have an impact

comparing distances using the following feature vector Alligator [1,1,0,1,4] rattlesnake [1,1,1,1,0] boa Constrictor [0,1,0,1,0] dart frog [1,0,1,0,4]

Table 2: Distance Computation

	rattlesnake	boa constrictor	dart frog	alligator
rattlesnake	-	1.414	4.243	4.123
boa Constrictor	1.414	-	4.472	4.123
dart frog	4.243	4.472	-	1.732
alligator	4.123	4.123	1.732	-

In table 2 distance between Alligator and rattlesnake and alligator and boa constrictor is quite high. Although all the three animals are reptiles.

Using binary features in table 3, i.e. an animal has either got legs or no legs, we may get a more realistic value

Table 3: Distance Calculation using binary

	rattlesnake	boa constrictor	dart frog	alligator
rattlesnake	-	1.414	1.732	1.414
boa Constrictor	1.414	-	2.236	1.414
dart frog	1.732	2.236	-	1.732
alligator	1.414	1.414	1.732	-

1.8 Confusion Matrix

A Confusion Matrix provides an overview of the performance of a classification model. It has NXN entries, where N denotes the no. of classes.

1. **True Positive**: Both the actual and the predicted values are positive.
2. **True Negative**: Both the actual and the predicted values are negative
3. **False Positive**: The actual value is negative and the predicted value is positive.
4. **False Negative**: The actual value is positive and the predicted value is negative.

Prediction outcome

		p	**n**	**total**
actual value	$\mathbf{p'}$	True Positive	False Negative	P'
	$\mathbf{n'}$	False Positive	True Negative	N'
	total	P	N	

amsmath

Problem 1. Computing Confusion Matrix
Suppose you would like to develop a classifier which can distinguish if a student will proceed to MS, immediately after completing his/her studies. Feature variables include CGPA and monthly income of parents.