CSCI 5922: Neural Nets and Deep Learning

Homework 1

Rahul Kumar Mohan Kumar

January 19, 2022

1 AI vs ML vs Deep Learning

Artificial Intelligence:

Artificial Intelligence is a set of rules or algorithms that a machine/computer/robot could use to mimic human intelligence. It is a broad set of study which includes machine learning and deep learning. So whenever a machine completes tasks using a set of rules or an algorithm and solves problems, then such intelligent behavior can be defined as artificial intelligence.

Machine Learning:

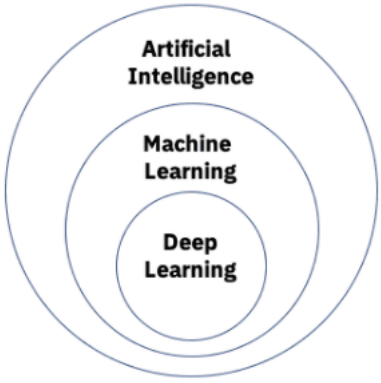
Machine learning can be simply interpreted as enabling machines to learn. The goal of machine learning is to make a machine learn by itself using the provided data and make predictions. The formal definition of Machine Learning is a program that can learn from experience E with respect to some class of task T and a performance measure P if its performance at tasks in T, as measured by P, improves with experience E.

Deep Learning:

Deep Learning is a sub-part of Machine Learning which makes use of Neural Networks which are similar to the neurons in our brain to mimic human brain-like behavior. DL algorithms focus on information processing to possibly identify patterns just like our human brain does and classifies the information accordingly. DL algorithms are more powerful than ML as it can easily work for larger sets of data.

Relation between the three:

As noted earlier, both machine learning and deep learning are subsets of artificial intelligence. Deep learning itself is a subset of machine learning. The relation can be summarised appropriately by the following venn diagram.



2 Supervised Learning Generalization:

The motivation behind splitting data into training and testing data-set is to avoid over-fitting. When we have developed our model using the data we need to make sure that the model is good and can work on unseen data. So we split the data into training and testing data-set so that we can validate our model and improve the accuracy. If we use up all the data to train the model then there won't be any way to decide if the prediction made by the model is accurate or not on new data.

The purpose of **training** data-set is to build up the ml/dl model. The model evaluates the data repeatedly to learn more about the data's behavior and then adjusts itself to serve its intended purpose.

The purpose of **testing** data-set is to validate if the model can make accurate predictions after the model has been built. The test data is unlabeled and provides a simulation of real world data to check if the model has been trained effectively or not. The basic purpose of this data is evaluation and validation.

Question 3

Given Training Data:

Sample	X_1	X_2	X_3	**Y**
1	0	0	0	1
2	1	0	0	1
3	1	1	0	-1

Model:

$$\phi(z) = \begin{cases} 1 & if \;\; z \geq 0 \\ -1 & otherwise \end{cases}$$

$$Where \; z = \sum_{j=0}^{m} x_j \times w_j = w^T x$$

Weight Update:

$$\Delta w_j = \eta(Target^{(i)} - Output^{(i)}) \times x_j^{(i)}$$
$$w_j = w_j + \Delta w_j$$

Part A : Model Training

EPOCH 1 $\implies$ Weights initialised to 0

w_0	w_1	w_2	w_3
0	0	0	0

Output Value for sample 1 = $1(0) + 0(0) + 0(0) + 0(0) => \phi(0) = 1$

Prediction Value:

Sample	X_1	X_2	X_3	**Y**	**Predicted**	w_0	w_1	w_2	w_3
1	0	0	0	1	1	0	0	0	0

Update Weight for Sample 1:

Given Learning Rate = η =0.1
Bias=X_0=1

$\implies \Delta w_0 = 0.1(1-1) \times 1 = 0 \implies w_0 = w_0 + \Delta w_0 = 0 + 0 = 0$
$\implies \Delta w_1 = 0.1(1-1) \times 0 = 0 \implies w_1 = w_1 + \Delta w_1 = 0 + 0 = 0$
$\implies \Delta w_2 = 0.1(1-1) \times 0 = 0 \implies w_2 = w_2 + \Delta w_2 = 0 + 0 = 0$
$\implies \Delta w_3 = 0.1(1-1) \times 0 = 0 \implies w_3 = w_3 + \Delta w_3 = 0 + 0 = 0$

Prediction and weight update for training data:

Sample	X_1	X_2	X_3	**Y**	**Predicted**	w_0	w_1	w_2	w_3
1	0	0	0	1	1	0	0	0	0
2	1	0	0	1	1	0	0	0	0
3	1	1	0	-1	1	0	0	0	0

Update Weights after epoch 1:

$\implies \Delta w_0 = 0.1(-1-1) \times 1 = -0.2 \implies w_0 = w_0 + \Delta w_0 = 0 - 0.2 = -0.2$
$\implies \Delta w_1 = 0.1(-1-1) \times 1 = -0.2 \implies w_1 = w_1 + \Delta w_1 = 0 - 0.2 = -0.2$
$\implies \Delta w_2 = 0.1(-1-1) \times 1 = -0.2 \implies w_2 = w_2 + \Delta w_2 = 0 - 0.2 = -0.2$
$\implies \Delta w_3 = 0.1(-1-1) \times 0 = 0 \implies w_3 = w_3 + \Delta w_3 = 0 + 0 = 0$

EPOCH 2

w_0	w_1	w_2	w_3
-0.2	-0.2	-0.2	0

Output Value for sample 1 = $-0.2(1) + -0.2(0) + -0.2(0) + 1(0) => \phi(-0.2) = -1$

Sample	X_1	X_2	X_3	**Y**	**Predicted**	w_0	w_1	w_2	w_3
1	0	0	0	1	-1	-0.2	-0.2	-0.2	0

$\implies \Delta w_0 = 0.1(1+1) \times 1 = 0.2 \implies w_0 = w_0 + \Delta w_0 = -0.2 + 0.2 = 0$
$\implies \Delta w_1 = 0.1(1+1) \times 0 = 0 \implies w_1 = w_1 + \Delta w_1 = -0.2 + 0 = -0.2$
$\implies \Delta w_2 = 0.1(1+1) \times 0 = 0 \implies w_2 = w_2 + \Delta w_2 = -0.2 + 0 = -0.2$
$\implies \Delta w_3 = 0.1(1+1) \times 0 = 0 \implies w_3 = w_3 + \Delta w_3 = 0 + 0 = 0$

Output Value for Sample 2 $= 0(1) + -0.2(1) + -0.2(0) + 1(0) => \phi(-0.2) = -1$

Sample	X_1	X_2	X_3	**Y**	**Predicted**	w_0	w_1	w_2	w_3
1	0	0	0	1	-1	-0.2	-0.2	-0.2	0
2	1	0	0	1	-1	0	-0.2	-0.2	0

$\implies \Delta w_0 = 0.1(1+1) \times 1 = 0.2 \implies w_0 = w_0 + \Delta w_0 = 0 + 0.2 = 0.2$
$\implies \Delta w_1 = 0.1(1+1) \times 1 = 0.2 \implies w_1 = w_1 + \Delta w_1 = -0.2 + 0.2 = 0$
$\implies \Delta w_2 = 0.1(1+1) \times 0 = 0 \implies w_2 = w_2 + \Delta w_2 = -0.2 + 0 = -0.2$
$\implies \Delta w_3 = 0.1(1+1) \times 0 = 0 \implies w_3 = w_3 + \Delta w_3 = 0 + 0 = 0$

Output Value for Sample 3 $= 0.2(1) + 0(1) + -0.2(1) + 0(0) => \phi(0) = 1$

Sample	X_1	X_2	X_3	**Y**	**Predicted**	w_0	w_1	w_2	w_3
1	0	0	0	1	-1	-0.2	-0.2	-0.2	0
2	1	0	0	1	-1	0	-0.2	-0.2	0
3	1	1	0	-1	1	0.2	0	-0.2	0

Update weight after epoch 2:

$\implies \Delta w_0 = 0.1(-1-1) \times 1 = -0.2 \implies w_0 = w_0 + \Delta w_0 = 0.2 - 0.2 = 0$
$\implies \Delta w_1 = 0.1(-1-1) \times 1 = -0.2 \implies w_1 = w_1 + \Delta w_1 = 0 - 0.2 = -0.2$
$\implies \Delta w_2 = 0.1(-1-1) \times 1 = -0.2 \implies w_2 = w_2 + \Delta w_2 = -0.2 - 0.2 = -0.4$
$\implies \Delta w_3 = 0.1(-1-1) \times 0 = 0 \implies w_3 = w_3 + \Delta w_3 = 0 + 0 = 0$

w_0	w_1	w_2	w_3
0	-0.2	-0.4	0

Part B : Model Testing

Given test data:

Sample	X_1	X_2	X_3	**Y**
1	1	1	0	-1
2	1	0	1	-1
3	1	1	1	1
4	0	0	0	1

Prediction calculation:

Sample 1: $\implies 0(1) + -0.2(1) + -0.4(1) + 0(0) => \phi(-0.2) = -1$
Sample 2: $\implies 0(1) + -0.2(1) + -0.4(0) + 0(1) => \phi(-0.2) = -1$
Sample 3: $\implies 0(1) + -0.2(1) + -0.4(1) + 0(1) => \phi(-0.6) = -1$
Sample 4: $\implies 0(1) + -0.2(0) + -0.4(0) + 0(0) => \phi(0) = 1$

Sample	X_1	X_2	X_3	**Y**	**Predicted**
1	1	1	0	-1	-1
2	1	0	1	-1	-1
3	1	1	1	1	-1
4	0	0	0	1	1

Part C : Model Evaluation

Confusion Matrix:

	Actual (1)	Actual (-1)
Predicted (1)	1	0
Predicted (-1)	1	2

Part D : Model Evaluation

Precision :

$$\frac{TP}{TP + FP} = \frac{1}{1+0} = \frac{1}{1}$$

Recall :

$$\frac{TP}{TP + FN} = \frac{1}{1+1} = \frac{1}{2}$$