[Lecture-7]

Q. Unsupervised Learning

→ A type of learning where the model works with the unlabeled data to discover patterns.

Example : Clustering (e.g. k-means)
Dimensionality (e.g. PCA)

Clustering : It is a way of grouping the data points into different cluster, consisting of a similar data points.

Q1.

House size (x_i)	Rent ($1000) ($y_i - \hat{y}_i$)
0.0	1.5
0.4	2.0
0.7	2.5
1.0	3.0

Either error tolerance or no. of steps will be given in the question.

Ans:

we know that,

$$w = w - \alpha \left[\frac{-2}{n} \sum_{i=1}^{n} x_i (y_i - \hat{y}_i) \right]$$

$$b = b - \alpha \left[-\frac{2}{n} \sum_{i=1}^{n} (y_i - \hat{y}_i) \right]$$

cost function $MSE = \left[\frac{1}{n} \sum_{i=1}^{n} (y_i - \hat{y}_i)^2 \right]$

Line for linear regression, $\hat{y} = wx + b$

[Lecture-3]. Supervised Learning

Q. Linear Regression:

→ Linear Regression is a fundamental statistical method for predicting continuous values. It models the relationships between the dependent (labels) and independent variables (features) using a straight line. A model with dependent variable is called multivariate linear regression. A model with independent variable is called multiple linear regression.

Q. Linear Regression Algorithm:

→ i. Initialize weight (w), bias (b) and learning rate (α)

ii. Compute the predicted value, $\hat{y}$.

iii. Compute the loss using MSE, J.

iv. Update the weight and bias using Gradient Descent.

$$w = w - \alpha \frac{\partial J}{\partial w}, \quad b = b - \alpha \frac{\partial J}{\partial b}$$

v. Repeat until convergence.

Q. Polynomial Regression

→ Describes the relationship between the independent and dependent variable (x/y) using an n^{th}- degree polynomial in x.

Types: Linear (degree [1]), Quadratic (degree [2]), Cubic (degree [3+])

[Lecture-4]

Q. Gradient:

→ Gradient of a scalar function is a vector that points in the direction of the greatest rate of increase of the function.

Algorithm:

i. Pick an initial point x_0

ii. Gradient Descent Update Rule

$$x_{t+1} = x_t - \alpha \nabla f(x_t)$$

iii. Iterate until convergence

where,
x_t = current position at iteration t
α_t = step size
$\nabla f(x_t)$ = gradient (first derivative) of the function $f(x)$

iv. $-\nabla f(x_t)$ moves in the negative direction of the gradient to minimize $f(x)$.

1^{st} iteration:

here, $w=0$, $b=0$, $\alpha=0.2$

i	$\hat{y}_i = wx+b$	$y_i-\hat{y}_i$	$(y_i-\hat{y}_i)^2$	$x_i(y_i-\hat{y}_i)$
1	$0\times0.0+0 = 0$	1.5	2.25	$0.00\times1.5 = 0.00$
2	0	2.0	4.00	$0.4\times2.0 = 0.80$
3	0	2.5	6.25	$0.7\times2.5 = 1.75$
4	0	3.0	9.00	$1.0\times3.0 = 3.00$
$n=4$	–	$\sum y_i-\hat{y}_i = 9$	$\sum(y_i-\hat{y}_i)^2 = \cancel{19.25}\ 21.5$	$\sum x_i(y_i-\hat{y}_i) = 5.55$

so, $w = w - \alpha\left[-\frac{2}{n}\sum x_i(y_i-\hat{y}_i)\right]$

$= 0 - 0.2\left[-\frac{2}{4}\times5.55\right]$

$= 0.555$

$b = b - \alpha\left[-\frac{2}{n}\sum(y_i-\hat{y}_i)\right]$

$= 0 - 0.2\left(-\frac{2}{4}\times9\right)$

$= 0.90$

$MSE = \frac{1}{n}\sum(y_i-\hat{y}_i)^2$

$= \frac{1}{4}\times21.5$

$= 5.375$

2nd iteration:

Now $w = 0.555$ and $b = 0.90$ and $\alpha = 0.2$

i	$\hat{y} = wx + b$	$(y_i - \hat{y}_i)$	$(y_i - \hat{y}_i)^2$	$x_i (y_i - \hat{y}_i)^2$
1	$0.555 \times 0.0 + 0.9 = 0.900$	$1.5 - 0.9 = 0.6$	$(0.6)^2 = 0.36$	0
2	$0.555 \times 0.4 + 0.9 = 1.122$	$2.0 - 1.122 = 0.878$	$(0.878)^2 = 0.771$	0.3512
3	$0.555 \times 0.7 + 0.9 = 1.288$	$2.5 - 1.288 = 1.212$	$(1.212)^2 = 1.468$	0.848
4	$0.555 \times 1.0 + 0.9 = 1.455$	$3.0 - 1.455 = 1.545$	2.387	1.545
		4.235	4.986	2.744

$$\therefore w = 0.555 - 0.2 \times \left[-\frac{2}{4} \times 2.7444\right] = 0.8294$$

$$\therefore b = 0.90 - 0.2 \times \left[-\frac{2}{4} \times 4.235\right] = 1.3235$$

$$\therefore MSE = \frac{1}{4} \times 4.986 = 1.2465$$

$$\therefore \hat{y} = wx + b$$

$$= 0.8294x + 1.3235$$

Ans:

Q2.

$f(x) = x^2 \Rightarrow 2x \qquad \therefore \nabla f(x_t) = 2x$

initial, $x_0 = -4$

step size, $\alpha = 0.8$

Ans:

we Know that,

$x_{t+1} = x_t - a \nabla f(x_t)$

$\therefore x_1 = -4 - 0.8 \times (2 \times -4) = 2.4$

$y = x^2 = 2.4 \times 2.4 = 5.76 \qquad \therefore (x_1, y_1) = (2.4, 5.76)$

$\therefore x_2 = 2.4 - 0.8 \times (2 \times 2.4) = -1.44$

$y = 2.07 \qquad \therefore (x_2, y_2) = (-1.44, 2.07)$

$\therefore x_3 = -1.44 - \{0.8 \times 2 \times (-1.44)\} = 0.864$

$y = (0.864)^2 = 0.746. \qquad \therefore (x_3, y_3) = (0.864, 0.746)$

Q.

Linear	Polynomial	Logistic
1. Predicts continuous outcomes, assumes a straight relationship between the independent and dependent variables	1. Extends linear regression to model non-linear relationships, using polynomial terms of the independent variables	1. Predicts probabilities for binary outcomes, using the logistic function to squeeze the output between 0 and 1. It is used for classification rather than regression

Q. K means Clustering

→ It is a method for grouping n observations into K clusters.

Algorithm:

i. Choose the number of cluster K.

ii. Select K random points from the data as centroids.

iii. Assign all the points to the closest cluster centroids.

iv. Recompute the centroids of newly formed clusters.

v. Repeat steps 3 and 4.

Q. point (x, y) A(1,2) B(2,3) C(3,3) D(6,7) E(7,8) F(8,8)

Ans:

Let, K = 2 [means divide the dataset into two clusters]

Centroid 1 (C1) = A(1,2)

Centroid 2 (C2) = E(7,8)

By using Manhatan distance:

	C1	C2	Decision
A	$\|1-1\|+\|2-2\| = 0$	$\|7-1\|+\|8-2\| = 12$	C_1
B	$\|1-2\|+\|2-3\| = 2$	$\|7-2\|+\|8-3\| = 10$	C_1
C	$\|1-3\|+\|3-3\| = 3$	$\|7-3\|+\|8-3\| = 9$	C_1
D	$\|1-6\|+\|2-7\| = 9$	$\|7-6\|+\|8-7\| = 2$	C_2
E	$\|1-7\|+\|2-8\| = 12$	$\|7-7\|+\|8-8\| = 0$	C_2
F	$\|1-8\|+\|2-8\| = 13$	$\|7-8\|+\|8-8\| = 1$	C_2

Recompute Centroids:

$$C_1 = \left(\frac{1+2+3}{2}, \frac{2+3+3}{3}\right) = (2, 2.67)$$

$$C_2 = \left(\frac{6+7+8}{3}, \frac{7+8+8}{3}\right) = (7, 7.67)$$

Again by using Manhattan distance

	C1	C2	Decision
A	$\lvert 2-1 \rvert + \lvert 2.67-2 \rvert = 1.666$	$\lvert 7-1 \rvert + \lvert 7.67-2 \rvert = 10.666$	C_1
B	$\lvert 2-2 \rvert + \lvert 2.67-3 \rvert = 0.334$	$\lvert 7-2 \rvert + \lvert 7.67-3 \rvert = 9.666$	C_1
C	$\lvert 2-3 \rvert + \lvert 2.67-3 \rvert = 1.334$	$\lvert 7-3 \rvert + \lvert 7.67-3 \rvert = 8.666$	C_1
D	$\lvert 2-6 \rvert + \lvert 2.67-7 \rvert = 5.334$	$\lvert 7-6 \rvert + \lvert 7.67-7 \rvert = 1.666$	C_2
E	$\lvert 2-7 \rvert + \lvert 2.67-8 \rvert = 5.334$	$\lvert 7-7 \rvert + \lvert 7.67-8 \rvert = 0.334$	C_2
F	$\lvert 2-8 \rvert + \lvert 2.67-8 \rvert = 6.334$	$\lvert 7-8 \rvert + \lvert 7.67-8 \rvert = 1.334$	C_2

Recompute centroids:

$$C_1 = \left(\frac{1+2+3}{3}, \frac{2+3+3}{3}\right) = (2, 2.666)$$

$$C_2 = \left(\frac{6+7+8}{3}, \frac{7+8+8}{3}\right) = (7, 7.67)$$

The C_1 and C_2 remain unchanged; so the solution is,

$C1 =$ A, B, C

$C2 =$ D, E, F