PA 23 Soln

ML

Gr-A

1) What is an artificial nurone?

An artificial nurone is a mathematical func that models biological nurone in an artificial nural n/w. (ANN). It is a tiny part of a computer system that takes i/p, does some calculations, & gives an o/p, similar to how a brain cell works.

(diagram) - $\boxed{\text{Done}}$

2) Describe the multipoint crossover for a nA problem. $\boxed{\text{Done}}$

3) Differentiate the fuzzy sets for the triangular, & trapitoidal membership functions.

Defination $\boxed{\text{Done}}$. Fuzzy sets use mFs to define the degree of belonging for elems within the set.

~~feature~~

Key differences of fuzzy sets on the basis on tri & trapi mF:

Feature	Tri	Trapi
i) No. of Parameters	3 (a, b, c)	4 (a, b, c, d)
ii) Shape	Single peak (triangle)	Flat - top (trapizoid)
iii) Complexity	Simpler, fewer parameters	Slightly more complex
iv) Application	sharp, single-point focus	Broders range of full membership.

4) Compare the fuzzifier & defuzzifier of a fuzzy inference system.

Fuzzifier	Defuzzifier
i) It converts crisp i/p vals. into fuzzy sets (crisp = hard, fuzzy = soft)	i) It converts fuzzy o/p sets into a crisp o/p val.
ii) In FIS, it is input interface for fuzzifying the system's input.	ii) output interface for producing actionable result in FIS.
iii) It's direction is crisp to fuzzy.	iii) It's direction is fuzzy to crisp.
iv) crisp i/p data is given as i/p.	iv) Fuzzy o/p sets from FIS is given as o/p
v) technique used - MFs (tri, trapi, gauss)	v) technique used - Defuzzification methods (centroid WAV, mean of maxima)
vi) It's key func is enabling reasoning with linguistic variables.	vi) It's key func is facilitating practical decision-making / ctrl.

5) Write the basic steps for a GA problem. Done

6) What is reinforcement training?
It refers to a method of learning where desired behaviours are encouraged by providing positive feedback / rewards when a specific action is performed correctly. (or) RL

Gm-B

8) Establish a mamdani fuzzy inference system with proper example. Done

9) ~~Establish a~~ Explain the terms, "chromosome, Gene, Allele, Locus, Genotype, phenotype" with proper examples for a GA problem. Done

b) eg let a string / chromosome is = 101011

i) Chromosome = 101011
ii) Gene = each digit of 101011, eg = 1 or 0
iii) Allele = (val. of gene) for binary chromo. Possible alleles are 0 or 1
iv) Locus = ~~1st~~ the locus of the 3rd gene (val = 1) is position = 3

1 0 1 0 1 1
1 2 3 4 5 6

v) Genotype = the full str 101011
vi) Phenotype = 10 10 11 encoding → A B C

10) Establish a MDC for a 3-class classification problem. Done (from Data sci, both algo & its eg)

11) calculate Specificity & f1-score for the given confusion matrix.

Predicted \ Actual				
	12	1	3	5
	7	45	4	6
	0	2	23	4
	1	5	7	6

↓

	Actual 1	Actual 2	Actual 3	Actual 4
Predicted 1	12	1	3	5
Predicted 2	7	45	4	6
Predicted 3	0	2	23	4
Predicted 4	1	5	7	6

Specificity $= \frac{TN}{TN+FP}$, $= \left(\frac{d}{c+d}\right)$ F1-score $= \frac{2*TP}{2*TP+FN+FP}$ $\left(\frac{2a}{2a+b+c}\right)$

multiclass classification's confusion matrix

↓ Converting it into C.m. for binary classification (binary c.m.)

~~Step 1 TP = 12~~

Step 1

	TP	TN	FP	FN
1	12	45+4+6+2+23+4+		
2	45			
3	23			
4				

Step 1

	TP	TN	FP	FN
1	12	45+4+6+2+23+4+5+7+6 = 102	7+0+1 = 8	1+3+5 = 9
2	45	12+3+5 + 0+2+3+4 + 1+7+6 = 61	1+2+5 = 8	1+4+6 = 11
3	23	12+1+5+ 7+45+6+ 1+5+6 = 88	3+4+7 = 14	0+2+4 = 6
4	6	12+1+3+ 7+45+4 +0+2+23 = 97	5+6+4 = 15	1+5+7 = 13

FP = col. val. add
FN = Row. " "

(no need to draw the binary C.m. (step 2))

∴ Specificity of 1 $= \frac{102}{102+8} = 93\%$

" " 2 $= \frac{61}{61+8} = 88\%$

" " 3 $= \frac{88}{88+14} = 86\%$

" " 4 $= \frac{97}{97+15} = 87\%$

F_1 Score of 1 $= \frac{2*12}{2*12+8+9} = 59\%$

" " 2 $= \frac{2*45}{2*45+8+11} = 83\%$

" " 3 $= \frac{2*23}{2*23+14+6} = 70\%$

" " 4 $= \frac{2*6}{2*6+15+13} = 0.3\%$

7) maximize the function, $f(x) = 4x^2 + 9x + 1$, where $x = 9, 11, 13, 15$ with (chromosome size = 4) such that i) selection operation (Rank selection), ii) uniform crossover, (iii) OP to 2 iterations.

i) selection (Rank selection) → (initial)

String no.	Population	x	$f(x)$ (fitness val)	Prob- count	Expected -count	Actual -c
1	01001 0	9	406 (3rd)	0.21	0.84	1
2	01011 0	11	584 (2nd)	0.3	1.21	1
3	01101 0	13	794 (1st)	0.41	1.64	2
4	00101 0	5	146 (4th)	0.08	0.3	0

$f(x) = 4x^2 + 9x + 1 \Rightarrow 4 * 81 + 9 * 9 + 1 = 406$

$\Rightarrow 4 * 121 + 9 * 11 + 1 = 584$

$\Rightarrow 4 * 169 + 9 * 13 + 1 = 794$

$\Rightarrow 4 * 25 + 9 * 5 + 1 = 146$

$\Sigma f(x) = 1930$ (total fitness)

avg fitness $= \frac{\Sigma f(x)}{n} = 482.5$

$P.C = \frac{f(x)}{\Sigma f(x)}$

$E.C = \frac{f(x)}{\text{avg fitness}}$

$A.C$ = round (E.C)

$S_1 = 10010$
$S_2 = 10110$
$S_3 = 11010$
$S_4 = 11010$

$S_1 = 011010$
$S_2 = 011010$
$S_3 = 010110$
$S_4 = 010010$

ii) crossover (uniform)

$S_1 = 0|1|1|0|1|0$
$S_2 = 0|1|1|0|1|0$

$S_1' = 01\ 1\ 0\ 1\ 0$
$S_2' = 01\ 1\ 0\ 1\ 0$

$S_3 = 0|1|0|1|1|0$
$S_4 = 0|1|0|0|1|0$

$S_3' = 01\ 0\ 0\ 1\ 0$
$S_4' = 01\ 0\ 1\ 1\ 0$

(iii) Calc the table again for best val of x.

it = 1 i) selection

off str.	popu	x	$f(x)$	P.C	E.C	A.C
S_1'	011010	13	794 (2nd)	0.31	1.23	1
S_2'	011010	13	794 (1st)	0.31	1.23	1
S_3'	010010	9	406 (4th)	0.16	0.63	1
S_4'	010110	11	584 (3rd)	0.27	0.91	1

$\Sigma f(x) = 2578$

avg $f(x) = 644.5$

$S_1' = 01101$
$S_2' = 01101$
$S_4' = 01001$
$S_3' = 01011$

ii) cross over

$S_1' = 01101$
$S_2' = 01101$

$S_{11}' = 01101$
$S_{21}' = 01101$

$S_4' = 01001$
$S_3' = 01011$

$S_{31}' = 01011$
$S_{41}' = 01001$

iii) compute table for best val of x

it 2

off str	popu	x	$f(x)$	P.C	E.C	A.C
S_{11}'	01101	13	794	0.31	1.23	1
S_{21}'	01101	13	794	0.31	1.23	1
S_{31}'	01011	11	584	0.27	0.91	1
S_{41}'	01001	9	406	0.16	0.63	1

$\Sigma f(x) = 2578$

avg $f(x) = 644.5$

(same as prev)

∴ the best val. of x after 2 iteration = 13
(chromose = 01101) with fitness val = 7.94

Ans-c

12) consider the fuzzy set Small = {0/0 + 0/2 + 1/3 + 0/4} & negative = {0/1 + 0.7/2 + 1/3 + 0.7/4 + 0/5} & the following fuzzy rule: "Rule 1: If x is small & y is negative Then z is low". Find the firing strength of Rule 1 when $x=3$ & $y=2$ where fuzzy "AND" operation is the minimum operator. What is ELITISM.

Small = {0/0 + 0/2 + 1/3 + 0/4}

Ans-c, negative = {0/1 + 0.7/2 + 1/3} + 0.7/4 + 0/5}

Rule 1: " If x is small & y is negative then z is low.

given, $x=3$, $y=2$ & fuzzy "AND" operation is the minimum operator.

$x=3$, comes.

i) Finding the membership value of x & y in the respective fuzzy sets–

Small(3) = 1, negative(2) = 0.7

ii) Applying AND operator –

small (3) AND negative (2)

= minimum {small(3), negative(2)}

= minimum {1, 0.7}

= 0.7

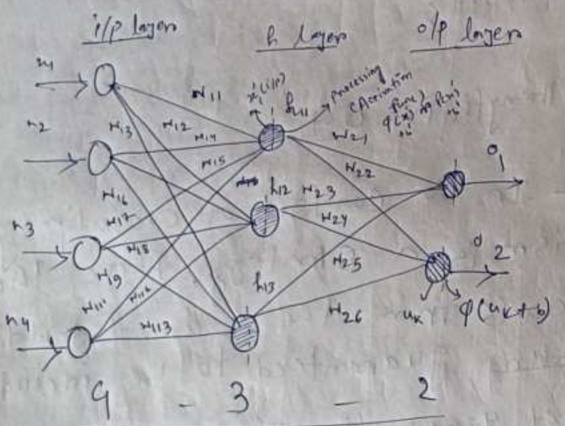
∴ Firing strength is = 0.7

ELITISM – In the context of evolutionary algos, elitism is a strategy, where the best individuals from the current population are ~~generated~~ guaranteed to be included in the next generation. This ensures that the best solutions found so far, are not lost & can contribute to the search for even better sol^n in ~~furth~~ future generations.

13) What is Gradient-Descent? Draw a very clear 4-3-2 ANN architecture with explaining all its components. What is a self-organizing feature map?

a) Gradient-Descent – It is an iterative optimization algo used to find the minimum of a function. It works by repeatedly adjusting the parameters of a model in the direction of the negative gradient, that indicates the direction of the steepest descent.

b) 4-3-2 ANN



x_1, x_2, x_3, x_4 = i/p feature vectors

o_1, o_2 = output class

h_{11}, h_{12}, h_{13} = hidden nodes

W = weights

= u_k = Processing node

b = bias

$\phi(\cdot)$ = Activation func $[\phi(x)]$

$[x = \text{input}]$

$x = \cancel{\text{input}}\ u_k + b$

$= \sum_{j=1}^{n} w_{kj} * x_j + b$ or $\sum_{i=1, j=1}^{h, n} w_{kj} * h_{ki} + b$

[eg, $\underline{h_{11}}$, $x_1' = u_k + b$

$= \cancel{w_{11} + w_{14} + w_1}$ (let $b = 0$)

$= w_{11} \cdot x_1 + w_{14} * x_2 + w_{15} * x_3 + w_{111} * x_4 + 0$

$h_{11} = \phi(x_1') = \tanh(x_1')$ [Say A.F = $\tanh(x)$]

$= \tan(w_{11} \cdot x_1 + w_{14} * x_2 + \cdots w_{111} * x_4)$

= outcome]

[eg, $\underline{O_1}$, $x_1'' = u_k + b$ (let $b = 0$)

$= w_{21} * h_{11} + w_{23} * h_{12} + w_{25} * h_{13} + 0$

$o_1 = \phi(x_1'') = \tanh(x_1'')$

$= \tanh(w_{21} * h_{11} + \cdots + h_{13})$

= outcome]

I/p layer - The 1st layers of ANN, consisting 4 neurons / nodes. Each nodes are feature. I/p layers receives i/p data.

Hidden layer - The 2nd layer, consists of 3 nodes / neurons. It processes i/p data from i/p layer.

o/p layer - the final layers, consists of 2 nodes. It produces the final output of the ANN. (Note - the o/p is actual o/p, not predicted o/p)

Weights - It is the connections b/w nodes in diff. layers. the values of w always lie b/w [0-1] $(0 \le w \le 1)$. the w vals. are adjusted during training.

Activation func - they introduce non-linearity into the n-n, allowing it to learn complex patterns.

Bias - A.F. contains some errors, to stop it a small val. is added to the processed i/p (u_k), that is called bias.

Sofm → self-organizing feature map

It is also known as Kohonen map. It is a type of neural n/w that learns to represent the i/p data in a lower-dimensional space, while preserving the topological relationships b/w the data points.

14) what is clustering? what are the main parameters for a good clustering technique? what are conventional & fuzzy set theories? Define uniform crossover & single point crossover in GA. [Done](a), [done](c), [done](d)

(b) main parameters for good clustering

i) No. of clusters (too few clusters = overgeneralization, too many = fragmentation)
ii) Distance metric (euclidean, manhattan etc)
iii) clustering algo (choosing the ~~no~~ right clustering algo. is very imp) (Depends on data & desired outcome)

(c) Conventional set theory - Here, an element either belongs to a set or it doesn't. This is binary classification.

Fuzzy set theory - It allows for partial membership, where an element can belong to a set to a certain degree. This is more flexible & can handle uncertainty & ambiguity better.

15) what is PCA & why is it important? Describe each step of PCA by considering a proper example. Give some real time application of neural networks.
[Done](a) [Done](b)

(c) Real time applications of NN

i) NLP → natural language processing
ii) Health care
iii) Speech recognition
iv) Facial recognition
v) computer vision & DIP (Digital Image Processing)
vi) Finance
vii) ~~Seo~~ Social media
viii) Chatbots
ix) Deep learning
x) Fraud detection

17 a) why is naive Bayes so 'naive'? (b) you came to know that your model is suffering from low bias & high variance. which algo should you use to tackle it? why? (c) what do you understand about Type I & Type II errors? (d) what is non linear classification of supervised learning? Explain with an eg.

c) done

a) Naive Bayes is called 'naive' as, it makes a strong assumption that all features are conditionally independent given the class label. This means it assumes that the presence or absence of a particular feature doesn't influence the presence or absence of another feature, that is often not true in real-world scenarios.

b) To tackle low bias & high variance ~~one~~ ~~s~~ we should use (a regularization method or ensemble method like) RF or Random forest or Gradient Boosting.

As these methods help to reduce overfitting, that is the cause of high variance, while still maintaining low bias.

d) non-linear classification is a type of supervised learning where the decision boundary b/w classes is not a ~~straig~~ straight line. For eg, a SVM with a non-linear kernel such as RBF (Radial basis func), that can create complex decision boundaries to seperate data points belonging to diff. classes.

16) Describe ~~some~~ Generative & discriminative m.l. techniques. Suppose a genetic algorithm uses chromosomes of the from $x = abcdefgh$ with a fixed length of eight genes. Each gene can be any digit b/w 0 & 9. Let the fitness of individual x be calculated as:

$f(x) = (a+b)*(c+d)+(e+f)-(g+h)$. Let the initial population consist of 4 individuals with the following chromosomes: $x_1 = 72413532$; $x_2 = 97121601$; $x_3 = 53221285$; $x_4 = 71852494$. Use the following (i) Evaluate the fitness of each individual, (ii) Cross the ~~fitnes~~ fitest 2 individuals using one-point crossover at the middle point, (iii) Evaluate the fitness of the new population with the best 4 chromosomes (2 old & 2 new) (iv) Perform (ii) to (iii) upto 3 iterations.

Generative ml techniques - It focus on learning the underlying patterns of data to create new data instances that resemble the original data set. It often uses the unsupervised learning for training.

Discriminative ml. techniques - It aims to distinguish b/w diff. categories of data, Primarily used for classification tasks ~~like~~. It often uses supervised learning for the training.

$f(x) = (a+b) * (c+d) + (e+f) - (g+h)$

$x_1 = 72413532$

$x_2 = 97121601$

$x_3 = 53221285$

$x_4 = 71852494$.

i) $f(x_1) = (7+2) * (4+1) + (3+5) - (3+2) = 48$

$f(x_2) = (9+7) * (1+2) + (1+6) - (0+1) = 54$

$f(x_3) = (5+3) * (2+2) + (1+2) - (8+5) = 22$

$f(x_4) = (7+1) * (8+5) + (2+4) - (9+4) = 97$

ii) | | Rank |
|---|---|
| $x_4 \longrightarrow$ | 1 |
| $x_2 \longrightarrow$ | 2 |
| $x_1 \longrightarrow$ | 3 |
| $x_3 \longrightarrow$ | 4 |

iii) $x_4 : 7185 \mid 2494$

$x_2 : 9712 \mid 1601$

$\downarrow$

$x_1' : 71851601$

$x_2' : 97122494$

iiii) $x_1' : 71851601$ (new)

$x_2' : 97122494$

$x_3' :$ ~~53221285~~ 72413532

$x_4' : 53221285$ (old)

$f(x_1') = (7+1) * (8+5) + (1+6) - (0+1) = 110$

$f(x_2') = (9+7) * (1+2) + (2+4) - (9+4) = 41$

$f(x_3') = (7+2) * (4+1) + (3+5) - (3+2) = 48$

$f(x_4') = (5+3) * (2+2) + (1+2) - (8+5) = 22$

v) $it = 1$

	Rank
$x_1' \longrightarrow$	1
$x_3' \longrightarrow$	2
$x_2' \longrightarrow$	3
$x_4' \longrightarrow$	4

$x_1' : 7185 \mid 1601$

$x_3' : 7241 \mid 3532$

$\downarrow$

$x_1'' : 71853532$ (new)

$x_2'' : 72411601$

$x_3'' : 97122494$ (old)

$x_4'' : 53221285$

$f(x_1'') = (7+1) * (8+5) + (3+5) - (3+2) = 107$

$f(x_2'') = (7+2) * (4+1) + (1+6) - (0+1) = 51$

$f(x_3'') = (9+7) * (1+2) + (2+4) - (9+4) = 41$

$f(x_4'') = (5+3) * (2+2) + (1+2) - (8+5) = 22$

$it = 2$

	R
$x_1'' \longrightarrow$	1
$x_2'' \longrightarrow$	2
$x_3'' \longrightarrow$	3
$x_4'' \longrightarrow$	4

$x_1'' : 7185 \mid 3532$

$x_2'' : 7241 \mid 1601$

$\downarrow$

$x_1''' : 71851601$ (new)

$x_2''' : 72413532$

$x_3''' : 97122494$

$x_4''' : 53221285$ (old)

$f(x_1''') = (7+1)*(8+5) + (1+6) - (0+1) = 110$
$f(x_2''') = (7+2)*(4+1) + (3+5) - (3+2) = 48$
$f(x_3''') = (9+7)*(1+2) + (2+4) - (9+4) = 41$
$f(x_4''') = (5+3)*(2+2) + (1+2) - (8+5) = 22$

it = 3

	R
$x_1''' \longrightarrow$	1
$x_2''' \longrightarrow$	2
$x_3''' \longrightarrow$	3
$x_4''' \longrightarrow$	4

x_1''' : 7 1 8 5 | 1 6 0 1
x_2''' : 7 2 4 1 | 3 5 3 2

x_1^{iv} : 7 1 8 5 3 5 3 2
x_2^{iv} : 7 2 4 1 1 6 0 1
(new)

x_3^{iv} : 9 7 1 2 2 4 9 4
x_4^{iv} : 5 3 2 2 1 2 8 5
(old)

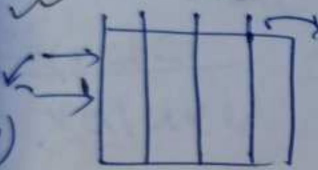
$f(x_1^{iv}) = (7+1)*(8+5) + (3+5) - (3+2)$
$f(x_2^{iv}) = (7+2)*(4+1) + (1+6) - (0+1) = 107$
$f(x_3^{iv}) = (9+7)*(1+2) + (2+4) - (9+4) = 51$
$f(x_4^{iv}) = (5+3)*(2+2) + (1+2) - (8+5) = 41$
$= 22$

...) & when there will be a i/p variables
new house it will do the same. i,e it.
continiously ~~tto~~ try to find out the numeric values.

Data PreProcessing: i) Remove noisy data (outliers)
(Process of transforming raw data into a clean data set)
[eg. height 6.11, forget to give '.', so suppose then the height - 611 ft this is outlier]

ii) Data imputation (missing value imputation)
[we try to estimate the values of the missing tuples, we can apply in various statistical computation]

iii) Dimentionality reduction (Feature extraction / Feature selection) (attributes)


(p)
(s)

pg 23 (line)

When P>>>S the m.l model fails to give us the desired result. This problem is called curse of dimentionality. This can be solved by dimentionality reduction or F.E or F.S. [we will remove the very less imp attribute/feature] that will not give any impact to the target attribute]

we will try to find out the less imp feature & will remove it, this is called feature selection (F.S). There are many algos for this.

we'll try to approximate/map n no. of attributes (the original features) into k no. of attributes (another no. of features) to get desired ~~attrib~~ result [e.g. 10,000 ~ 100], this called feature execution (F.E). There is an algo for it that is called principal component analysis (PCA).

iv) Data transformation (Data normalization)

In ds, diff. attributes have diff. ranges. If on this we apply the process it will not give us the desired result. So we have to take them all in same range.

It has 2 types - i) min-max
ii) Standard scaler (Z-score)

i) try to transform the data in b/w (0,1)
ii) try to ... in b/w (-1,1)

i) min-max

x_i = original valu
x_i' = updated value

$$\boxed{x_i' = \frac{x_i - X_{min}}{X_{max} - X_{min}}}$$

ii) Standard scal

$$\boxed{x_i' = \frac{x_i - \mu}{s}}$$

formula for z-score normalizati

we have to pen to the other se

2nd TYPES of (D

numeric (either floatin

Continuous — Disc

Interval scale — Ratio scale

Continuous - range the values

Discrete - values in

PCA

Steps

1. Subtract the mean from each data dim. & find the mean adjusted data.

Suppose 2d → 1d

	x	y
1	x_1	y_1
2	x_2	y_2
n	x_n	y_n

original data —transform→

$x-\mu_x$	$y-\mu_y$
$x_1-\mu_x$	$y_1-\mu_y$
$x_2-\mu_x$	$y_2-\mu_y$
$x_n-\mu_x$	$y_n-\mu_y$

mean(μ) substraction

mean adjusted data

2. Compute the covariance matrix 'A'.
(this is a type of mat that is used to represent the cov values b/w pairs of elems given in a random vector)

Covariance mat p-dim (no of features)

2d data

$$A = \begin{pmatrix} var(x) & cov(x,y) \\ cov(y,x) & var(y) \end{pmatrix}$$

$cov(x,y) = cov(y,x)$ always

$var(x) = \frac{\sum(x_i-\mu_x)^2}{n-1}$ (for sample data)

$cov(x,y) = \frac{\sum(x_i-\mu_x)(y_i-\mu_y)}{n-1}$

3.

i) Calculate the eigen values & eigen vectors of the covariance matrix 'A'.

It is applicable for square mat.
If n no of eigen values & vectors are given then n = dim of data set

ii) $A\,X = \lambda X$

$(A-\lambda)X = 0$

if consider,

$A-\lambda = 0$ we will get diff λ values

$\lambda_1, \lambda_2, \dots, \lambda_p$

As the data is 2d we will get 2 eigen values — λ_1, λ_2

4. Arrange the eigen values in descending order from largest to smallest.

5. From eigen values we will find out the eigen vectors.

$Ax = \lambda x$

$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} u_1 \\ u_2 \end{pmatrix}$ (λ_1) eigen vectors

$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}$ (λ_2) eigen vector

As we have 2 eigen vals λ_1, λ_2 we have to solve this exp for 2 times.

$Ax = \lambda_1 x$ — (i)

$Ax = \lambda_2 x$ — (ii)

we have to find out the value of x & y by using (i) (ii)

$A = \begin{pmatrix} x \\ y \end{pmatrix}$ → cov mat.

6. We will transform the eigen vectors into unit length Eigen vectors.

$$\begin{pmatrix} u_1' \\ u_2' \end{pmatrix} = \begin{pmatrix} \frac{u_1}{\sqrt{u_1^2+u_2^2}} \\ \frac{u_2}{\sqrt{u_1^2+u_2^2}} \end{pmatrix}$$

7. derive the new dataset.

$$\boxed{\text{new dataset} = \text{Row Feature vector} * \text{Row Data Adjacent}}$$
(dim - 2 x n)

Feature vector (p x p - dim) [∴ P eigen values → P eigen vectors] $= \begin{pmatrix} u_1' & v_1' \\ u_2' & v_2' \\ \vdots & \vdots \\ u_p' & v_p' \end{pmatrix}$

For 2d, $\boxed{\begin{pmatrix} u_1' & v_1' \\ u_2' & v_2' \end{pmatrix}}$

Row Feature vector (p x p) = (Feature vector) T → transpose (→ p x p)

Row Data Adj (2 x n) = (mean Adj Data) T → transpose (n x 2)

For 2d, 2 x 2
R.F.V. = 2 x 2 (Feature vector)
R.D.A = 2 x 2

Eigen value of a matrix - It is a special set of scalar values that is associated with the set of linear eqns most probably in the matrix eqns.

Eigen vectors of a matrix - It is a vector that is associated with a set of linear eqns. These are also known as latent vectors, proper vectors or characteristic vectors.

①

CW

PCA (Dimen...

let we have 2D data

x	y
2.5	2.4
0.5	0.7
2.2	2.9
1.9	2.2
3.1	3
2.3	2.7
2	1.6
1	1.1
1.5	1.6
1.1	0.9

$\bar{x} = \frac{1.81}{10}$,

∴ $M_x = \frac{18.1}{10} = 1.8$

∴ A = covariance matrix

var(x) =

var(y) =

, weight, gender
data But dis clas
n algo (ML algo).
of these are given, but not y)
rical levels/O/PS.

(prov)

alues / continuous values

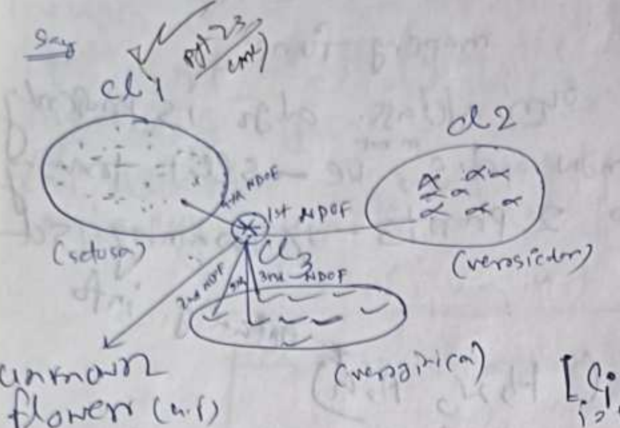
pervised learning

[known, GA → Actual y info]

Compare
to see the accuracy
of our mapping func / class. algo

KNN (K-nearest neighbour classification algo)

K= the no. of neighbours info that we use to build our classification model.
→[A supervised m.l. algo that uses a distance-based approach to classify/predict the grouping of a data point]||



[cl = class]

let $K = 5$

[N.D.O.F = nearest dist of flower]

$[C_i = C_1, C_2, C_3]_{i=1-3}$ [cl ≈ C]

As the, NDOF towards cl_3 is the most, so we'll considers that unknown obj falls into cl_3.

(demonstration)

- Cal. the dis b/w the u.f & all other flowers
- see, the ~~shortest dists b/w u.f &~~ K no. of n.n (nearest neighbours), here 5 n.n.
- Among the N DOF (n.n), we can see, {3, 3, 3, 3, 1} i.e the majority of the nearest flower is belonging to cl_3, so that u.f blongs to 3rd catagory | cl_3.

for this eg
$$\begin{cases} C_1 = \frac{1}{5} = 0.2 \\ C_2 = \frac{1}{5} = 0.2, \; C_3 = \frac{3}{5} = 0.6 \end{cases}$$

- If we have a ~~tg~~ tie among the no. of n.n b/w the clases we can increase the value of K & ~~see the~~ check again same (way 1)

eg – {3, 2, 3, 2, 1}
d_1 d_2 d_3 d_4 d_5

[d = distance]

~~way 2~~ $d_1 + d_3$ Compare $d_2 + d_3$

if $(d_1 + d_3) < (d_2 + d_4)$, then $cl_3 \leftarrow$ u.k

if $(d_1 + d_3) > (d_2 + d_4)$, then $cl_2 \leftarrow$ u.k

pg 23
KNN Algo

1) Store a
2) For co
3) search
input obj
(euclidean
4) For (
for each
of objs
-ging to
5) The
class

Eg-①
Dataset

S_1	
1	
2	
3	
4	
5	
6	

pyq 23
KNN Algo sugg

pyq 23 (imp)

1) Store all the input data in the training set
2) For each object in the test set
3) search for K nearest neighbours to the input obj's / pattern using any distance matrix (obj's) (euclidean etc)
4) For classification we'll compute the confidence for each classes as c_i/K where c_i is the no. of obj's among the K-nearest neighbours belonging to class label 'i'
5) The classification for the i/p obj is the class with the heighest confidence.

Eg-① Imp sugg

Dataset

(eg of binary classification) class label attribute

Sl	CGPA	Assi m	Proj m	Result
1	9.2	8.5	8	Pass
2	8	8.0	7	Pass
3	8.5	8.1	8	pass
4	6	4.5	5	Fail
5	6.5	5.0	4	Fail
6	8.9	7.2	7	Pass
7	5.8	3.8	5	Fail
8	8.9	9.1	9	pass

Training d.s (available d.s)

vi) model designing

<u>catagory of machine learning model</u>

i) unsupervised learning [~~18~~ clusters] → clustering

ii) supervised learning

iii) Rein~~forming~~cement lea~~r~~ning

as similar as possible

this are different

and as similar as possible

• <u>iris data</u> — setosa, versicolor, virginica,

— 50 rows — 50 rows — 50 rows (total 150 rows)

Petal length & width, sepal length & width

<u>Unsupervised Learning</u> – we don't know which obj falls into which catagory. eg—
In above eg data there will be clusters of 3 types of flowers but we can't distinguish that which cluster is of which flower.

<u>Supervised learning</u> – Here we basically train our model, & if there is a unknown obj in the model it ~~th~~ can correctly predict ~~that~~ in what catagory it falls into.

<u>Reinforcement learning</u> – Semisupervised learning. Here step by step we try to improve our ml model to get more accurate result. If the model isn't giving more accurate result then

Penalize it otherwise give it reward (Save it)
Eg- self driving

Iris data - The iris data are a data frame of 150 measurements of iris petal & sepal length & width, with 50 measurements for each species of 'setosa', 'versicolor' & 'virginica'.

Unsupervised learning - It is a type of m.l that learns from data without human supervision. Unlike supervised learning, unsupervised m. l. models are given unlabeled data & allowed to discover patterns & insights without any explicit guidance / instruction.

Supervised learning - It is a category of m.l that uses labeled datasets to train algorithms to predict outcomes & recognize patterns. Unlike unsupervised l., s. l. algos are given labeled training to learn the relationship b/w the i/p & o/ps.

Reinforcement learning - It is a m.l. training method based on rewarding desired behaviours & punishing undesired ones. In general, a reinforcement learning agent - the entity being trained - is able to perceive & interpret its environment, take actions & learn through trial & error.

12:57 0.53 KB/S 98

solution

2. **Draw a very clear 4-3-2 ANN architecture with explaining all its components.**

Component	Explanation
Input Layer	The first layer of the ANN, consisting of 4 neurons. It receives the input data.
Hidden Layer 1	The second layer, consisting of 3 neurons. It processes the input data from the input layer.
Hidden Layer 2	The third layer, consisting of 2 neurons. It further processes the data from the previous layer.
Output Layer	The final layer, consisting of 2 neurons. It produces the final output of the ANN.
Weights	The connections between neurons in different layers. They represent the strength

Are these results useful? Yes No ×

(c) What do you understand about Type I & Type II errors?

Type I error (False Positive):

This occurs when you reject a true null hypothesis. In other words, you predict something to be true when it's actually false.

Type II error (False Negative):

This occurs when you fail to reject a false null hypothesis. In other words, you predict something to be false when it's actually true.