MATHS

K-Means:

* Calculate the distance for each data points from existing centroid M_1 and M_2.

$$\text{Euclean distance} = \sqrt{(x_o - x_c)^2 + (y_o - y_c)^2}$$

* Assign the data points to centroids based on the shortest distance to cluster and data points.

* Calculate new centroid M_1 and M_2.

* Again calculate the distance for the new centroid.

* Assign data points to the new Centroid.

4-Means Clustering

$* \; X = \left\{ \begin{pmatrix} 0.5 \\ 1.0 \end{pmatrix}, \begin{pmatrix} 0.0 \\ 4.0 \end{pmatrix}, \begin{pmatrix} 3.5 \\ 2.0 \end{pmatrix}, \begin{pmatrix} 0.5 \\ -4.5 \end{pmatrix}, \begin{pmatrix} -0.5 \\ -3.5 \end{pmatrix} \right\}$

$m_1 = \begin{pmatrix} -2.5 \\ 0.5 \end{pmatrix} \qquad m_2 = \begin{pmatrix} 2.5 \\ 4.5 \end{pmatrix}$

$*$ Euclidean distance $= \sqrt{(x_0 - x_c)^2 + (y_0 - y_c)^2}$

$\Rightarrow$ Iteration 1:

	(0.5, 1.0)	(0.0, 4.0)	(3.5, 2.0)	(0.5, -4.5)	(-.5, -3.5)
distance from $m_1 = (-2.5, 0.5)$	3.04	4.30	6.18	5.83	4.47
distance from $m_2 = (2.5, 4.5)$	4.03	2.54	2.69	9.21	8.54
Assigned group	M_1	M_2	M_2	M_1	M_1

$\sqrt{(0.5 - (-2.5))^2 + (1 - 0.5)^2} \Rightarrow \sqrt{9 + 0.25} = 3.04$

New centroid: M_1, $x = \dfrac{0.5 + -4.0 + 0.5 - .5}{4} = -0.87$

$y = \dfrac{1.0 + 0.5 - 4.5 - 3.5}{4} = -1.62$

For M_2, $x = \frac{0.0+3.5}{2} = 1.75$

$y = \frac{4.0+2.0}{2} = 3$

$\therefore M_2 = (1.75, 3)$

-.5, -3.5
1.91
6.87
M_1

Iteration 2:

	(.5,1)	(0,4)	(-4,.5)	(3.5,2)	(.5,-4.5)
distance from $M_1 = (-0.87, -1.62)$	2.95	5.68	3.78	5.67	3.18
distance from $M_2 = (1.75, 3)$	2.35	2.01	6.26	2.01	7.6
assigned group	M_2	M_2	M_1	M_2	M_1

New, $C_1 = \{x_3, x_5, x_6\}$

$C_2 = \{x_1, x_2, x_4\}$

Gauss elimination

Worksheet 1

Jordan

* Gaussian Elimination method;

Linear equation: $a_1x + b_1y + c_1z = d_1$

$a_2x + b_2y + c_2z = d_2$

$a_3x + b_3y + c_3z = d_3$

$$\begin{bmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} d_1 \\ d_2 \\ d_3 \end{bmatrix}$$

Step 1: find augmented matrix for given matrix

$$C = [A : B]$$

2) Transform C to Echelon form / upper triangler

3) Untill $\left(\begin{array}{ccc|c} 1 & 0 & 0 & \cdots \\ 0 & 1 & 0 & \cdots \\ 0 & 0 & 1 & \cdots \end{array}\right)$

$$\frac{-\frac{1}{2}+8}{\;} \quad \frac{-1+16}{2} = \frac{15}{2}$$

$$3 - \frac{7\times\cancel{7}}{\cancel{6}2} \quad \frac{6-7}{2} =$$

$$-\frac{1}{2} - \frac{7\times\cancel{8}}{\cancel{6}2} = \frac{-1-7}{2} = \frac{-8}{2}$$

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 1 & 4 & 9 \\ 1 & 8 & 27 \end{bmatrix} \qquad b = \begin{bmatrix} 3 \\ 2 \\ 7 \end{bmatrix}$$

$$= \left(\begin{array}{ccc|c} 1 & 2 & 3 & 3 \\ 0 & 2 & 6 & -1 \\ 0 & 6 & 24 & +4 \end{array}\right)$$

$$= \left(\begin{array}{ccc|c} 1 & 2 & 3 & 3 \\ 0 & 2 & 6 & -1 \\ 0 & 0 & 6 & 7 \end{array}\right)$$

$$= \left(\begin{array}{ccc|c} 1 & 2 & 3 & 3 \\ 0 & 1 & 3 & -1/2 \\ 0 & 0 & 6 & 7 \end{array}\right)$$

$$= \left(\begin{array}{ccc|c} 1 & 2 & 3 & 3 \\ 0 & 1 & 3 & -0.5 \\ 0 & 0 & 1 & 7/6 \end{array}\right)$$

$$= \left(\begin{array}{ccc|c} 1 & 2 & 0 & -1/2 \\ 0 & \cancel{0} & 0 & -4 \\ 0 & 0 & 1 & 7/6 \end{array}\right)$$

$$= \left(\begin{array}{ccc|c} 1 & 0 & 0 & 7.5 \\ 0 & 1 & 0 & -4 \\ 0 & 0 & 1 & 7/6 \end{array}\right)$$

Eigenvalue and vector:

* $A \cdot X = \lambda \cdot X$ → Eigenvector
  (λ) → Eigenvalue

$\Rightarrow Ax - \lambda x = 0$

$\Rightarrow \boxed{(A - \lambda I)X = 0}$

* Write characteristic equation for Matrix A

$$|A - \lambda I| = 0$$

$$\Rightarrow \quad A = \begin{bmatrix} 8 & -8 & -2 \\ 4 & -3 & -2 \\ 3 & -4 & 1 \end{bmatrix}$$

$$\Rightarrow \begin{vmatrix} 8-\lambda & -8 & -2 \\ 4 & -3-\lambda & -2 \\ 3 & -4 & 1-\lambda \end{vmatrix} = 0$$

$$\Rightarrow (8-\lambda)\{(-3-\lambda)(1-\lambda)-(8)\}+8\{4(1-\lambda)$$
$$-(-2\times 3)\}+(-2)\{-16-3(-3-\lambda)\}$$

$$\Rightarrow (8-\lambda)(-3+3\lambda-\lambda+\lambda^2-8)+8(4-4\lambda+6)$$
$$-2(-16+9+3\lambda)$$

$$\Rightarrow -24+24\lambda-8\lambda+8\lambda^2-64+3\lambda-3\lambda^2$$
$$+\lambda^2-\lambda^3+8\lambda+32-32\lambda+48+32-18-6\lambda$$

$$\Rightarrow -\lambda^3+6\lambda^2-11\lambda+6=0$$

$$\Rightarrow -\lambda^2(\lambda-1)+5\lambda(\lambda-1)-6(\lambda-1)=0$$

$$\Rightarrow (\lambda-1)(-\lambda^2+5\lambda-6)=0$$

$$\Rightarrow (\lambda-1)(\lambda^2-5\lambda+6)=0$$

$$\Rightarrow (\lambda-1)(\lambda^2-2\lambda-3\lambda+6)=0$$

$$\Rightarrow (\lambda-1)\{(\lambda(\lambda-2)-3(\lambda-2)\}=0$$

$$\Rightarrow (\lambda-1)(\lambda-2)(\lambda-3)=0$$

$$\boxed{\lambda = 1, 2, 3}$$

Shortcut: $\lambda^3 - \begin{bmatrix}\text{Sum of diagonal} \\ \text{elements}\end{bmatrix}\lambda^2 + \begin{bmatrix}\text{Sum of} \\ \text{diagonal minor}\end{bmatrix}\lambda - |A| = 0$

Now for $\lambda = 1$

$$\begin{bmatrix} 7 & -8 & -2 \\ 4 & -4 & -2 \\ 3 & -4 & 0 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

* Now apply cramer's rule

$$7x_1 - 8x_2 - 2x_3 = 0 \quad — \quad (1)$$
$$4x_1 - 4x_2 - 2x_3 = 0 \quad — \quad (2)$$
$$3x_1 - 4x_2 - 0 = 0 \quad — \quad (3)$$

$$\frac{x_1}{\begin{vmatrix} -8 & -2 \\ -4 & -2 \end{vmatrix}} = \frac{-x_2}{\begin{vmatrix} 7 & -2 \\ 4 & -2 \end{vmatrix}} = \frac{x_3}{\begin{vmatrix} 7 & -8 \\ 4 & -4 \end{vmatrix}}$$

$$\Rightarrow \frac{x_1}{8} = \frac{+x_2}{+6} = \frac{x_3}{4}$$

$$X = \begin{bmatrix} 8 \\ 6 \\ 4 \end{bmatrix} = \begin{bmatrix} 4 \\ 3 \\ 2 \end{bmatrix}$$ [Eigenvector for $\lambda_1 = 1$]

* Do the same for λ_2 and λ_3.

Linear PCA

* Mean of the variables (features)

$$\bar{X} = 1 \sum (i)$$

$$\frac{1}{N}\sum_{i=1} x$$

* Centered data $Y = x^{(ij)} - \bar{x}$

* Compute covariance matrix

$$C = \frac{1}{N}\sum_{i=1}^{N} Y \cdot Y^T$$

* Calculate eigenvalue and eigenvector of covariance matrix.

[for PCA we will always take the highest eigenvalue, λ]

[Qus এ বলা থাকবে λ কয়টা নিতে হবে। বলা না থাকলে সবচেয়ে বড় λ (Eigen value) নিতে হবে eigen vector এর জন্য]

* Transformation matrix, $T = (v_1, v_2, v_3)$

$$= \begin{pmatrix} \lambda_1 & \lambda_2 & \lambda_3 \\ \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots \end{pmatrix}$$

* Transformed data point $Z_{(ij)} = T^T \cdot y^{(ij)}$

$Z_1 = T^T \cdot y^1$; $Z_2 = T^T \cdot y^2$ ....

Example:

Linear PCA calc (10 Punkte)

Sie haben die folgende Antwort gegeben:

You are provided with the following data, regarding nutrition scores of some vegetables, containing the amount of magnesium, calcium and potassium per $100g$:

	Mg	Ca	Ka
Carrots	49	42	57
Cabbage	45	40	52
Beetroot	45	37	55
Tomatoes	45	34	49
Onions	49	36	54
Leek	49	39	51

Apply the PCA algorithm to transform the vegetable's features into a one-dimensional data space. Give all necessary calculation steps as result.

Let, Carrots, cabbage, beetroot, tomatoes, onions, leek as $x_1, x_2, x_3, x_4, x_5, x_6$

$\therefore X = \{x_1, x_2, x_3, x_4, x_5\}$

$$= \left\{ \begin{matrix} 49 & 45 & 45 & 45 & 49 & 49 \\ 42 & 40 & 37 & 34 & 36 & 39 \\ 57 & 52 & 55 & 49 & 54 & 51 \end{matrix} \right\}$$

mean value, $\bar{X} = \left\{ \begin{matrix} 47 \\ 38 \\ 53 \end{matrix} \right\}$

Centered data, $Y = \begin{Bmatrix} 2 & -2 & -2 & -2 & 2 & 2 \\ 4 & 2 & -1 & -4 & -2 & 1 \\ 4 & -1 & 2 & -4 & 1 & -2 \end{Bmatrix}$

Covariance matrix $C = \frac{1}{N}\left(Y \cdot Y^T\right)$

$$\Rightarrow \frac{1}{N}\begin{pmatrix} 2 & -2 & -2 & -2 & 2 & 2 \\ 4 & 2 & -1 & -4 & -2 & 1 \\ 4 & -1 & 2 & -4 & 1 & -2 \end{pmatrix}\begin{pmatrix} 2 & 4 & 4 \\ -2 & 2 & -1 \\ -2 & -1 & 2 \\ -2 & -4 & -4 \\ 2 & -2 & 1 \\ 2 & 1 & -2 \end{pmatrix}$$

$$\Rightarrow \frac{1}{N}\begin{pmatrix} 4+4+4+4+4+4 & 8-4+2+8-4+2 & 8+2-4+8+2-4 \\ 8-4+2+8-4+2 & 16+4+1+16+4+1 & 16-2-2+16-2-2 \\ 8+2-4+8+2-4 & 16-2-2+16-2-2 & 16+1+4+16+1+4 \end{pmatrix}$$

$$\Rightarrow 1/6\begin{pmatrix} 24 & 12 & 12 \\ 12 & 42 & 24 \\ 12 & 24 & 42 \end{pmatrix}$$

$$\Rightarrow \begin{pmatrix} 4 & 2 & 2 \\ 2 & 7 & 4 \\ 2 & 4 & 7 \end{pmatrix}$$

$$\Rightarrow \begin{pmatrix} 4-\lambda & 2 & 2 \\ 2 & 7-\lambda & 4 \\ 2 & 4 & 7-\lambda \end{pmatrix} = 0$$

$=$

$$(4-\lambda)\left(49 - 14\lambda + \lambda^2 - 1\right) - 2\left(14 - 2\lambda - 8\right) + 2\left(8 - 14 + 2\lambda\right)$$

$$= (4-\lambda)\left(\lambda^2 - 14\lambda + 33\right) - 12 + 4\lambda + 4\lambda - 12$$

$$= 4\lambda^2 - 56\lambda + 132 - \lambda^3 + 14\lambda^2 - 33\lambda - 24 + 8\lambda$$

$$\Rightarrow -\lambda^3 + 18\lambda^2 - 81\lambda + 108$$

$$\Rightarrow \lambda^3 - 18\lambda^2 + 81\lambda - 108$$

$$\Rightarrow \lambda^2(\lambda - 3) - 15\lambda(\lambda - 3) + 36(\lambda - 3)$$

$$\Rightarrow (\lambda - 3)(\lambda^2 - 15\lambda + 36)$$

$$\Rightarrow (\lambda - 3)(\lambda^2 - 12\lambda - 3\lambda + 36)$$

$$\Rightarrow (\lambda - 3)(\lambda - 3)(\lambda - 12) = 0$$

$\lambda = 12 \quad \lambda = 3 \quad \lambda = 3$

$\lambda_1 = 12 \qquad \lambda_2 = 3 \qquad \lambda_3 = 3$

for $\lambda_1 = 12$

$$\begin{pmatrix} 4-12 & 2 & 2 \\ 2 & 7-12 & 4 \\ 2 & 4 & 7-12 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = 0$$

$$= \begin{pmatrix} -8 & 2 & 2 \\ 2 & -5 & 4 \\ 2 & 4 & -5 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = 0$$

$$\Rightarrow \; -8x_1 + 2x_2 + 2x_3 = 0 \quad — \; 1$$

$$2x_1 - 5x_2 + 4x_3 = 0 \quad — \; 2$$

$$2x_1 + 4x_2 - 5x_3 = 0 \quad — \; 3$$

$$\Rightarrow \frac{x_1}{\begin{vmatrix} 2 & 2 \\ -5 & 4 \end{vmatrix}} = \frac{-x_2}{\begin{vmatrix} -8 & 2 \\ 2 & 4 \end{vmatrix}} = \frac{x_3}{\begin{vmatrix} -8 & 2 \\ 2 & -5 \end{vmatrix}}$$

$$\Rightarrow \frac{x_1}{18} = \frac{\cancel{-}x_2}{\cancel{-}36} = \frac{x_3}{36}$$

$$X = \begin{pmatrix} 18 \\ 36 \\ 36 \end{pmatrix} = \begin{pmatrix} 2 \\ 4 \\ 4 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \\ 2 \end{pmatrix}$$

$\therefore$ Transformation matrix $T = \begin{pmatrix} 1 \\ 2 \\ 2 \end{pmatrix}$

$\therefore$ first PCA : $z = T^T . y^{(i)}$

$$= \begin{pmatrix} 1 & 2 & 2 \end{pmatrix} \begin{Bmatrix} 2 & -2 & -2 & -2 & 2 & 2 \\ 4 & 2 & -1 & -4 & -2 & 1 \\ 4 & -1 & 2 & -4 & 1 & -2 \end{Bmatrix}$$

$$= \begin{pmatrix} 18 & 0 & 0 & -18 & 0 & 0 \end{pmatrix}$$

Reduced to one dimension.

Kernel PCA

* Polynomial kernel, kp = $\left((x^i)^T . x^j + a \right)^d$
matrix

$\Rightarrow$ Given, $a = 1$, $d = 2$. Calculate Polynomial kernel, kp.

$$kp_{11} = \left\{ \begin{pmatrix} 3 & 2 & 1 & 2 \end{pmatrix} \begin{pmatrix} 3 \\ 2 \\ 1 \\ 2 \end{pmatrix} + 1 \right\}^2$$

$$= (9+4+1+4+1)^2$$

$$= 361$$

$$kp_{22} = \left\{ \begin{pmatrix} 2 & -2 & -1 & 0 \end{pmatrix} \begin{pmatrix} 2 \\ -2 \\ -1 \\ 0 \end{pmatrix} + 1 \right\}^2$$

$$= (4+4+1+1)^2$$

$$= 100$$

$$kp_{33} = \left\{ \begin{pmatrix} 3 & 4 & 2 & 1 \end{pmatrix} \begin{pmatrix} 3 \\ 4 \\ 2 \\ 1 \end{pmatrix} + 1 \right\}^2$$

$$= (9+16+4+1+1)^2$$

$$= 961$$

$$kp_{12} = \left\{ \begin{pmatrix} 3 & 2 & 1 & 2 \end{pmatrix} \begin{pmatrix} 2 \\ -2 \\ -1 \\ 0 \end{pmatrix} + 1 \right\}^2$$

$$= 4$$

$$kp_{13} = \left\{ \begin{pmatrix} 3 & 2 & 1 & 2 \end{pmatrix} \begin{pmatrix} 3 \\ 4 \\ 2 \\ 1 \end{pmatrix} + 1 \right\}^2$$

$$= 484$$

$$kp_{21} = kp_{12}$$

$$kp_{23} = \left((2\ -2\ -1\ 0)\begin{pmatrix}3\\4\\2\\1\end{pmatrix}+1\right)^2$$

$$= (6-8-2+1)^2$$

$$= 9$$

$$kp_{32} = kp_{23}$$

$$kp_{31} = kp_{13}$$

$\therefore$ Polynomial kernel $Kp = \begin{pmatrix}361 & 4 & 484\\ 4 & 100 & 9\\ 484 & 9 & 961\end{pmatrix}$

* Gaussian kernel matrix, $kg = \exp\left(-\frac{\|x^i - x^j\|^2}{2\sigma^2}\right)$

$\Rightarrow$ Given, $\sigma = 0.5$. Calculate gaussian kernel kg.

$$\Rightarrow kg_{11} = \exp\left(-\frac{\left\|\begin{pmatrix}3\\2\\1\\2\end{pmatrix}-\begin{pmatrix}3\\2\\1\\2\end{pmatrix}\right\|^2}{2(.5)^2}\right)$$

$$= \exp\left(-\frac{0}{0.5}\right) = \exp(0) = 1$$

$$\Rightarrow kg_{12} = \exp\left(- \frac{\left\| \begin{pmatrix} 3 \\ 2 \\ 1 \\ 2 \end{pmatrix} - \begin{pmatrix} 2 \\ -2 \\ -1 \\ 0 \end{pmatrix} \right\|^2}{2(.5)^2} \right)$$

$$\exp\left(- \frac{\left\| \begin{pmatrix} 1 \\ 4 \\ 2 \\ 2 \end{pmatrix} \right\|^2}{0.5} \right)$$

$$= \exp\left(- \frac{1^2 + 4^2 + 2^2 + 2^2}{0.5} \right) = \exp\left(\frac{-25}{0.5} \right) = \exp(-50)$$

$$\Rightarrow kg_{13} = \exp\left(- \frac{\left\| \begin{pmatrix} 3 \\ 2 \\ 1 \\ 2 \end{pmatrix} - \begin{pmatrix} 3 \\ 4 \\ 2 \\ 1 \end{pmatrix} \right\|^2}{2(.5)^2} \right)$$

$$= \exp\left(- \frac{(0+4+1+1}{0.5} \right) = \exp(-12)$$

$\Rightarrow kg_{21} = kg_{12}$ $\Rightarrow kg_{22} = 1$ $\Rightarrow kg_{33} = 1$

$kg_{31} = kg_{13}$ $\Rightarrow kg_{32} = kg_{23}$

$$\therefore kg = \begin{pmatrix} 1 & \exp(-50) & \exp(-12) \\ \exp(-50) & 1 & \exp(-94) \\ \exp(-12) & \exp(-94) & 1 \end{pmatrix}$$

* Kernel PCA:

* Find polynomial kernel matrix, K_P.
* kernel PCA $\tilde{K} = (K - 1_N K - K 1_N + 1_N K 1_N)$

  where, $1_N = \frac{1}{3}\begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}$

* Find eigenvalues and eigenvector.

  [if asked one-dimensional PCA take higest λ,
  if asked two-dimensional PCA take two λ]

* Transformed data point / two dimensional

  Kernel PCA $z = \lambda^T \cdot K$

  kernel matrix.

  transposed eigenvector

$$\Rightarrow \begin{pmatrix} 361 & 4 & 484 \\ 4 & 100 & 9 \\ 484 & 9 & 961 \end{pmatrix} - \begin{pmatrix} 283 & 37.6 & 484.6 \\ 283 & 37.6 & 484.6 \\ 283 & 37.6 & 484.6 \end{pmatrix}$$

$$\boxed{K - 1_N K}$$

$$\Rightarrow \begin{pmatrix} 78 & -33.6 & .6 \\ -279 & 62.4 & -475.6 \\ 201 & -28.6 & 476.4 \end{pmatrix} - K1_N + 1_N K 1_N$$

$$K1_N = \begin{pmatrix} 283 & 283 & 283 \\ 37.6 & 37.6 & 37.6 \\ 484.6 & 484.6 & 484.6 \end{pmatrix}$$

$$1_N K 1_N = \frac{1}{3} \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} 283 & 283 & 283 \\ 37.6 & 37.6 & 37.6 \\ 484.6 & 484.6 & 484.6 \end{pmatrix}$$

$$= \begin{pmatrix} 268.4 & 268.4 & 268.4 \\ 268.4 & 268.4 & 268.4 \\ 268.4 & 268.4 & 268.4 \end{pmatrix}$$

$$\widetilde{K} = \begin{pmatrix} 63.4 & -48.2 & -14 \\ -48.2 & 293.2 & -244.8 \\ -15.2 & -244.8 & 260.2 \end{pmatrix}$$

$\lambda_1 = 523.41 \quad \lambda_2 = 93.259 \quad \lambda_3 = 0$

eigen vector $\bar{\lambda} = \begin{pmatrix} 0.001372 & -0.048706 \\ -0.018491 & 0.021538 \\ 0.017119 & 0.027168 \end{pmatrix}$

two KPCA $= \begin{pmatrix} 0.001372 & -0.018491 & 0.017119 \\ -0.048706 & 0.021538 & 0.027168 \end{pmatrix}$

$\cdot \begin{pmatrix} 361 & 4 & 484 \\ 4 & 100 & 9 \\ 484 & 9 & 961 \end{pmatrix}$

$= \begin{pmatrix} 8.706924 & -1.689541 & 16.948988 \\ -4.347402 & 2.203488 & 2.728586 \end{pmatrix}$

Polynomial Regression

You get the following data regarding the taste of wine, depending on acid (pH), sugar (mg/100ml), and alcohol (vol%), measured using a star rating from 0 to 5.

	Sugar	Acid	Alcohol	Stars
Bordeaux	4000	5.0	6.0	4
Messanges	8000	5.0	8.0	5
Cabernet	6000	2.5	8.0	3
Veron	2000	7.5	10.0	2

Use linear regression to predict the taste labels of the following unlabeled wines:

	Sugar	Acid	Alcohol	Stars
Veronais	3000	6.25	9.0	?
Vidure	4000	5	12.0	?
Trouchet Noir	4000	2.5	15.0	?

Give the prediction of the taste labels as well as all necessary calculation steps as a result.

Let, $x_1 =$ Sugar

$x_2 =$ Acid

$x_3 =$ Alcohol

and $y =$ Stars

$$Y = B_0 + B_1 x_1 + B_2 x_2 + B_3 x_3$$

$$\begin{array}{ccccc} b_0 & b_1 & b_2 & b_3 & y \\ 1 & 4000 & 5 & 6 & 4 \\ 1 & 8000 & 5 & 8 & 5 \\ 1 & 6000 & 2.5 & 8 & 3 \\ 1 & 2000 & 7.5 & 10 & 2 \end{array}$$

$$\Rightarrow \left(\begin{array}{cccc|c} 1 & 4000 & 5 & 6 & 4 \\ 0 & 4000 & 0 & 2 & 1 \\ 0 & 2000 & -2.5 & 2 & -1 \\ 0 & 2000 & -2.5 & -4 & 2 \end{array}\right)$$

$$\Rightarrow \left(\begin{array}{cccc|c} 1 & 4000 & 5 & 6 & 4 \\ 0 & 4000 & 0 & 2 & 1 \\ 0 & 0 & +5 & -2 & 3 \\ 0 & 0 & 5 & 10 & -3 \end{array}\right)$$

$$\Rightarrow \left(\begin{array}{cccc|c} 1 & 4000 & 5 & 6 & 4 \\ 0 & 4000 & 0 & 2 & 1 \\ 0 & 0 & 5 & -2 & 3 \\ 0 & 0 & 0 & -12 & 6 \end{array}\right)$$

$\therefore -12 b_3 = 6 \Rightarrow \boxed{b_3 = -0.5}$

$\therefore 5b_2 - 2(-0.5) = 3 \Rightarrow \boxed{b_2 = 0.4}$

$\therefore 4000 b_1 + 2 b_3 = 1 \Rightarrow b_1 = \frac{1 - 2.(0.5)}{}$

$$\overline{}$$
$$4000$$

$$\boxed{b_1 = 0}$$

$\therefore b_0 + 4000 \cdot 0 + 5(0.4) + 6(-0.5) = 4$

$\Rightarrow \boxed{b_0 = 5}$

$\therefore$ For sugar $x_1 = 3000$, Acid $x_2 = 6.25$
Alcohol $x_3 = 9$,

Stars $y = B_0 + B_1 x_1 + B_2 X_2 + B_3 X_3$
$= 5 + 0 + 2.5 - 4.5$
$\boxed{= 3}$

For $x_1 = 4000$, $x_2 = 5$, $x_3 = 12$

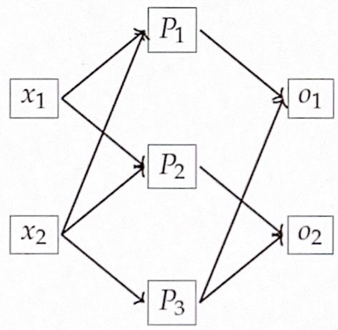
$y = 5 + 0 + 2 - 6$
$\boxed{= 1}$

For $x_1 = 4000$, $x_2 = 2.5$, $x_3 = 15$

$y = 5 + 0 + 1 - 7.5$
$\boxed{= -1.5}$

Neural Network

Exercise 1 [Neural networks - Calculations].
Given the following network:



The initial weights are all 1, the biases of the output layers are 0, the biases of P_i are 0. Activation functions for all layers are $\psi(t) := \frac{1}{1+e^{-t}}$. The input points are

$$X = \left\{ \begin{pmatrix} 2 \\ 5 \end{pmatrix}, \begin{pmatrix} 3 \\ 2 \end{pmatrix}, \begin{pmatrix} 7 \\ -2 \end{pmatrix}, \begin{pmatrix} 8 \\ 4 \end{pmatrix}, \begin{pmatrix} -4 \\ -2 \end{pmatrix}, \begin{pmatrix} 4 \\ 3 \end{pmatrix} \right\}, Y = \left\{ \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \end{pmatrix} \right\}.$$

For,

$x_1 = 2, \quad x_2 = 5,$

$P_1 = w_1 x_1 + w_2 x_2 + b = 7 \;\Big/\; Z_{P_1} = \dfrac{1}{1+e^{-7}} = 0.999$

$P_2 = w_1 x_1 + w_2 x_2 + b = 7 \;\Big/\; Z_{P_2} = \dfrac{1}{1+e^{-7}} = 0.999$

$P_3 = w_2 x_2 + b = 5 \;\Big/\; Z_{P_3} = \dfrac{1}{1+e^{-5}} = 0.9931$

Now computing output layer,

$$O_1 = w_1 P_1 + w_3 P_3 + b = 0.999 + 0.9931 \Rightarrow 1.9924$$

$$ZO_1 = \frac{1}{1+e^{-1.9924}} = 0.88$$

$$O_2 = w_2 P_2 + w_3 P_3 + b = 0.999 + 0.9931 \Rightarrow 1.9924$$

$$ZO_2 = \frac{1}{1+e^{-1.9924}} = 0.88$$

$$\left[\text{Loss function, } C(\theta) = \frac{1}{2}\left\|f_\theta(x^1) - y^1\right\|^2\right]$$

$$\frac{dc}{db_{o_1}} = \text{Error} \times \text{derivative of activation func}$$

$$= (ZO_1 - y^1) \cdot \frac{1}{1+e^{-t}} \cdot \left(1 - \frac{1}{1+e^{-t}}\right)$$

$$= (0.88 - 0) \cdot 0.88 \cdot (1 - 0.88)$$

$$= 0.0929$$

$$\frac{dc}{db_{o_2}} = (ZO_2 - y^2) \cdot \frac{1}{1+e^{-t}} \cdot \left(1 - \frac{1}{1+e^{-t}}\right)$$

$$= (0.88 - 1) \cdot 0.88 \cdot (1 - 0.88)$$

$$= -0.0127$$

Now calculating the weight for output layer.

$$\frac{dc}{dw_{o_1}p_1} = \frac{dc}{db_{o_1}} \cdot Z_{p_1} = 0.0929 \times 0.999 \Rightarrow .0928$$

$$\frac{dc}{dw_{o_2}p_2} = \frac{dc}{db_{o_2}} \cdot Z_{p_2} = -0.0127 \times 0.999 \Rightarrow -0.0127$$

$$\frac{dc}{dw_{o_1}p_3} = \frac{dc}{db_{o_1}} \cdot Z_{p_3} = 0.0929 \times 0.9933 \Rightarrow 0.0923$$

$$\frac{dc}{dw_{o_2}p_3} = \frac{dc}{db_{o_2}} \cdot Z_{p_3} = -0.0127 \times 0.9933 \Rightarrow -0.0126$$

Now calculating the bias of the hidden layer

$$\frac{dc}{db_{p_1}} = \frac{dc}{db_{o_1}} \times w_{o_1}p_1 \times Z_{p_1}(1 - Z_{p_1})$$

$$= 0.0929 \times 1 \times 0.999(1 - 0.999)$$

$$= 8.446 \times 10^{-5}$$

$$\frac{dc}{db_{p_2}} = \frac{dc}{db_{o_2}} \times w_{o_2}p_2 \times Z_{p_2}(1 - Z_{p_2})$$

$$dwp_2 \quad dbo_2$$

$$= -0.0127 \times 1 \times 0.999(1-0.999)$$

$$= -1.155 \times 10^{-5}$$

$$\frac{dc}{db_{p_3}} = \left(\frac{dc}{db_{o_1}} \cdot w_{o_1 p_3} + \frac{dc}{db_{o_2}} \cdot w_{o_2 p_3}\right) \times Z_{p_3}(1-Z_{p_3})$$

$$= (0.0929 \cdot 1 - 0.0127 \cdot 1) \times 0.9931(1-0.9931)$$

$$= 5.329 \times 10^{-4}$$

Now calculating the remaining weight from hidden layer to input layer.

$$\frac{dc}{dw_{p_1 x_1}} = \frac{dc}{db_{p_1}} \cdot x_1 = 8.446 \times 10^{-5} \times 2 \Rightarrow 1.6892 \times 10^{-4}$$

$$\frac{dc}{dw_{p_1 x_2}} = \frac{dc}{db_{p_1}} \cdot x_2 = 8.446 \times 10^{-5} \times 5 \Rightarrow 4.223 \times 10^{-4}$$

$$\frac{dc}{dw_{p_2 x_1}} = \frac{dc}{db_{p_2}} \cdot x_1 = -1.155 \times 10^{-5} \times 2 \Rightarrow -2.31 \times 10^{-5}$$

$$\frac{dc}{dw_{p_2 x_2}} = \frac{dc}{db_{p_2}} \cdot x_2 = -1.155 \times 10^{-5} \times 5 \Rightarrow -5.775 \times 10^{-5}$$

$$\frac{dc}{dw_{p_3 x_2}} = \frac{dc}{db_{p_3}} \cdot x_2 = 5.329 \times 10^{-4} \times 5 \Rightarrow 2.6645 \times 10^{-3}$$

Support Vector Machine

Support Vector Machine - Linear Example Solved

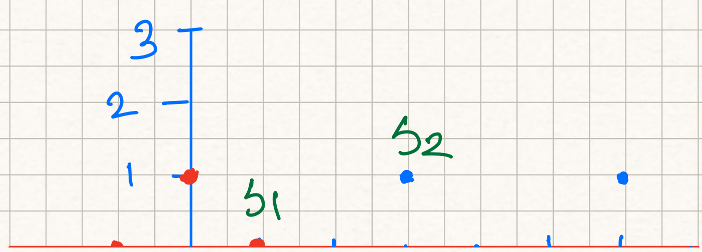
Suppose we are given the following positively labeled data points,

$$\left\{ \begin{pmatrix} 3 \\ 1 \end{pmatrix}, \begin{pmatrix} 3 \\ -1 \end{pmatrix}, \begin{pmatrix} 6 \\ 1 \end{pmatrix}, \begin{pmatrix} 6 \\ -1 \end{pmatrix} \right\}$$

and the following negatively labeled data points,

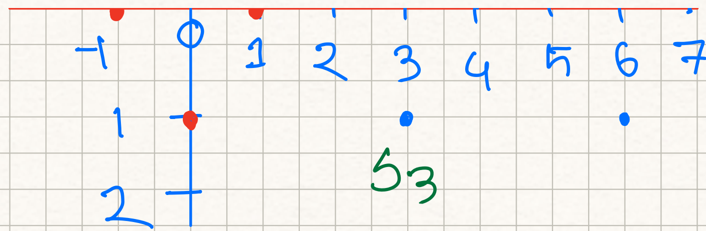
$$\left\{ \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ -1 \end{pmatrix}, \begin{pmatrix} -1 \\ 0 \end{pmatrix} \right\}$$

(*) Plot the data



● → negatively level

● → positively level



$\therefore\ S_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad S_2 = \begin{pmatrix} 3 \\ 1 \end{pmatrix} \quad S_3 = \begin{pmatrix} 3 \\ -1 \end{pmatrix}$

Support vector

* Now augment bias input 1 to each Support Vector.

$$\tilde{S}_1 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} \quad \tilde{S}_2 = \begin{pmatrix} 3 \\ 1 \\ 1 \end{pmatrix} \quad \hat{S}_3 = \begin{pmatrix} 3 \\ -1 \\ 1 \end{pmatrix}$$