سلسلة تمارين في التحليل التمييزي

التمرين 1:

لنفترض أن لدينا مجموعة بيانات تحتوي على متغيرين مستقلين X_1 وX_2 ، ونريد تصنيف الحالات إلى فئتين ($Y = 0$ أو $Y = 1$). البيانات الخاصة بالمتوسط لكل متغير في كل فئة هي كالتالي:

- الفئة: $Y = 0$
  - المتوسط لـ X_1: $\bar{X}_{1,0} = 5$
  - المتوسط لـ X_2: $\bar{X}_{2,0} = 3$
- الفئة: $Y = 1$
  - المتوسط لـ X_1: $\bar{X}_{1,1} = 8$
  - المتوسط لـ X_2: $\bar{X}_{2,1} = 7$

باستخدم التحليل التمييزي الخطي (LDA) لبناء دالة تمييزية .ثم قم بتصنيف حالة جديدة لها القيم التالية:

- $X_1 = 6$
- $X_2 = 4$

الحل المفصل:

1.حساب معاملات المتغيرات w_1 وw_2 :

المعادلة العامة لدالة التميز الخطية هي:

$$D(X) = w_1X_1 + w_2X_2 + w_0$$

حيث:

$$w_i = \bar{X}_{i,1} - \bar{X}_{i,0}$$

- بالنسبة لـ: w_1

$$w_1 = \bar{X}_{1,1} - \bar{X}_{1,0} = 8 - 5 = 3$$

- بالنسبة لـ: w_2

$$w_2 = \bar{X}_{2,1} - \bar{X}_{2,0} = 7 - 3 = 4$$

إذن:

$$w_1 = 3 \quad و \quad w_2 = 4$$

2. حساب الثابت: w_0

المعادلة المستخدمة لحساب w_0 هي:

$$w_0 = -\frac{1}{2}\left(\sum_{i=1}^{p} \bar{X}_{i,1}^2 - \sum_{i=1}^{p} \bar{X}_{i,0}^2\right)$$

- حساب: $\sum_{i=1}^{p} \bar{X}_{i,1}^2$

$$\sum_{i=1}^{p} \bar{X}_{i,1}^2 = \bar{X}_{1,1}^2 + \bar{X}_{2,1}^2 = 8^2 + 7^2 = 64 + 49 = 113$$

- حساب: $\sum_{i=1}^{p} \bar{X}_{i,0}^2$

$$\sum_{i=1}^{p} \bar{X}_{i,0}^2 = \bar{X}_{1,0}^2 + \bar{X}_{2,0}^2 = 5^2 + 3^2 = 25 + 9 = 34$$

- حساب الفرق:

$$\sum_{i=1}^{p} \bar{X}_{i,1}^2 - \sum_{i=1}^{p} \bar{X}_{i,0}^2 = 113 - 34 = 79$$

- حساب: w_0

$$w_0 = -\frac{1}{2} \times 79 = -39.5$$

إذن:

$$w_0 = -39.5$$

3. بناء دالة التمييز النهائية:

دالة التمييز تصبح:

$$D(X) = 3X_1 + 4X_2 - 39.5$$

4. تصنيف الحالة الجديدة:

لدينا حالة جديدة مع القيم التالية:

$$X_1 = 6 \quad \text{و} \quad X_2 = 4$$

نحسب قيمة $D(X)$ لهذه الحالة:

$$D(X) = 3(6) + 4(4) - 39.5$$
$$D(X) = 34 - 39.5 = -5.5$$

.5اتخاذ القرار:

- إذا كانت: $D(X) > 0$ الحالة تنتمي إلى الفئة $Y = 1$.
- إذا كانت: $D(X) \leq 0$ الحالة تنتمي إلى الفئة $Y = 0$.

في هذه الحالة:

$$D(X) = -5.5 \leq 0$$

إذن:

الحالة تنتمي إلى الفئة $Y = 0$

تمرين 2:

لنفترض أن شركة تسويقية ترغب في تصنيف العملاء بناءً على سلوكهم الشرائي إلى فئتين:

- $Y = 0$:العملاء غير المستهدفين (غير المهتمين بالشراء).
- $Y = 1$:العملاء المستهدفين (المهتمين بالشراء).

تم جمع البيانات حول عينة من 10 عملاء، وتشمل المتغيرات المستقلة:

- X_1 :متوسط إنفاق العميل الشهري (بالدولار).
- X_2 :عدد زيارات الموقع الإلكتروني للعميل شهريًا.

البيانات:

حالة	X_1 (إنفاق شهري)	X_2(عدد الزيارات)	الفئة(Y)
1	50	3	عميل غير مستهدف
2	60	4	عميل غير مستهدف
3	70	5	عميل غير مستهدف
4	80	6	عميل مستهدف
5	90	7	عميل مستهدف
6	100	8	عميل مستهدف
7	110	9	عميل مستهدف
8	120	10	عميل مستهدف
9	40	2	عميل غير مستهدف
10	130	11	عميل مستهدف

المطلوب:

1. استخدم التحليل التمييزي الخطي (LDA) لبناء دالة تمييزية.
2. قم بتصنيف عميل جديد لديه:
   - $X_1 = 75$ (إنفاق شهري).
   - $X_2 = 5$ (عدد زيارات).

الحل المفصل:

1. حساب المتوسط لكل متغير داخل كل فئة:

الفئة: $Y = 0$

- المتوسط لـ: X_1

$$\bar{X}_{1,0} = \frac{\sum_{i=1}^{n_0} X_{1,i}}{n_0} = \frac{50 + 60 + 70 + 40}{4} = \frac{220}{4} = 55$$

- المتوسط لـ: X_2

$$\bar{X}_{2,0} = \frac{\sum_{i=1}^{n_0} X_{2,i}}{n_0} = \frac{3+4+5+2}{4} = \frac{14}{4} = 3.5$$

الفئة: $Y = 1$

- المتوسط لـ: X_1

$$\bar{X}_{1,1} = \frac{\sum_{i=1}^{n_1} X_{1,i}}{n_1} = \frac{80+90+100+110+120+130}{6} = \frac{630}{6} = 105$$

- المتوسط لـ: X_2

$$\bar{X}_{2,1} = \frac{\sum_{i=1}^{n_1} X_{2,i}}{n_1} = \frac{6+7+8+9+10+11}{6} = \frac{51}{6} = 8.5$$

إذن:

$$\bar{X}_{1,0} = 55, \quad \bar{X}_{2,0} = 3.5, \quad \bar{X}_{1,1} = 105, \quad \bar{X}_{2,1} = 8.5$$

2. حساب معاملات المتغيرات w_1 و w_2

المعادلة العامة لمعاملات المتغيرات هي:

$$w_i = \bar{X}_{i,1} - \bar{X}_{i,0}$$

- بالنسبة لـ: w_1

$$w_1 = \bar{X}_{1,1} - \bar{X}_{1,0} = 105 - 55 = 50$$

- بالنسبة لـ: w_2

$$w_2 = \bar{X}_{2,1} - \bar{X}_{2,0} = 8.5 - 3.5 = 5$$

إذن:

$$w_1 = 50 \quad \text{و} \quad w_2 = 5$$

3. حساب الثابت: w_0

المعادلة المستخدمة لحساب w_0 هي:

$$w_0 = -\frac{1}{2}\left(\sum_{i=1}^{p} \bar{X}_{i,1}^2 - \sum_{i=1}^{p} \bar{X}_{i,0}^2\right)$$

- حساب: $\sum_{i=1}^{p} \bar{X}_{i,1}^2$

$$\sum_{i=1}^{p} \bar{X}_{i,1}^2 = \bar{X}_{1,1}^2 + \bar{X}_{2,1}^2 = 105^2 + 8.5^2 = 11025 + 72.25 = 11097.25$$

- حساب: $\sum_{i=1}^{p} \bar{X}_{i,0}^2$

$$\sum_{i=1}^{p} \bar{X}_{i,0}^2 = \bar{X}_{1,0}^2 + \bar{X}_{2,0}^2 = 55^2 + 3.5^2 = 3025 + 12.25 = 3037.25$$

- حساب الفرق:

$$\sum_{i=1}^{p} \bar{X}_{i,1}^2 - \sum_{i=1}^{p} \bar{X}_{i,0}^2 = 11097.25 - 3037.25 = 8060$$

- حساب w_0:

$$w_0 = -\frac{1}{2} \times 8060 = -4030$$

إذن:

$$w_0 = -4030$$

4. بناء دالة التمييز النهائية:

دالة التمييز تصبح:

$$D(X) = 50X_1 + 5X_2 - 4030$$

5. تصنيف الحالة الجديدة:

لدينا حالة جديدة مع القيم التالية:

$$X_1 = 75 \quad \text{و} \quad X_2 = 5$$

نحسب قيمة $D(X)$ لهذه الحالة:

$$D(X) = 50(75) + 5(5) - 4030$$
$$D(X) = 3750 + 25 - 4030 = -255$$

6.اتخاذ القرار:

- إذا كانت: $D(X) > 0$ الحالة تنتمي إلى الفئة $Y = 1$
- إذا كانت: $D(X) \leq 0$ الحالة تنتمي إلى الفئة $Y = 0$

في هذه الحالة:

$$D(X) = -255 \leq 0$$

إذن:

الحالة تنتمي إلى الفئة $Y = 0$

تمرين 3:

شركة تسويقية ترغب في تصنيف العملاء بناءً على سلوكهم الشرائي إلى فئتين:

- $Y = 0$:العملاء غير المهتمين بالعرض.
- $Y = 1$:العملاء المهتمين بالعرض.

تم جمع البيانات حول عينة من 10 عملاء، وتشمل المتغيرات المستقلة:

- X_1 :عدد الإعلانات التي شاهدها العميل خلال الشهر الماضي.
- X_2 :متوسط الوقت الذي يقضيه العميل على الموقع الإلكتروني (بالدقائق).

البيانات:

حالة	X_1 (عدد الإعلانات)	X_2 (متوسط الوقت)	الفئة(Y)
1	3	10	0
2	4	12	0
3	5	15	0
4	6	20	1
5	7	25	1
6	8	30	1
7	9	35	1
8	10	40	1
9	2	8	0
10	11	45	1

المطلوب:

1. استخدم التحليل التمييزي الخطي (LDA) لبناء دالة تمييزية.
2. قم بتصنيف عميل جديد لديه:
   - $X_1 = 7$ (عدد الإعلانات).
   - $X_2 = 22$ (متوسط الوقت).

الحل المفصل:

1. حساب المتوسط لكل متغير داخل كل فئة:

الفئة $Y = 0$

- المتوسط لـ:X_1

$$\bar{X}_{1,0} = \frac{\sum_{i=1}^{n_0} X_{1,i}}{n_0} = \frac{3+4+5+2}{4} = \frac{14}{4} = 3.5$$

- المتوسط لـ:X_2

$$\bar{X}_{2,0} = \frac{\sum_{i=1}^{n_0} X_{2,i}}{n_0} = \frac{10+12+15+8}{4} = \frac{45}{4} = 11.25$$

الفئة $Y = 1$

- المتوسط لـ:X_1

$$\bar{X}_{1,1} = \frac{\sum_{i=1}^{n_1} X_{1,i}}{n_1} = \frac{6+7+8+9+10+11}{6} = \frac{51}{6} = 8.5$$

- المتوسط لـ: X_2

$$\bar{X}_{2,1} = \frac{\sum_{i=1}^{n_1} X_{2,i}}{n_1} = \frac{20 + 25 + 30 + 35 + 40 + 45}{6} = \frac{195}{6} = 32.5$$

إذن:

$$\bar{X}_{1,0} = 3.5, \quad \bar{X}_{2,0} = 11.25, \quad \bar{X}_{1,1} = 8.5, \quad \bar{X}_{2,1} = 32.5$$

2. حساب معاملات المتغيرات w_1 وw_2:

المعادلة العامة لمعاملات المتغيرات هي:

$$w_i = \bar{X}_{i,1} - \bar{X}_{i,0}$$

- بالنسبة لـ: w_1

$$w_1 = \bar{X}_{1,1} - \bar{X}_{1,0} = 8.5 - 3.5 = 5$$

- بالنسبة لـ: w_2

$$w_2 = \bar{X}_{2,1} - \bar{X}_{2,0} = 32.5 - 11.25 = 21.25$$

إذن:

$$w_1 = 5 \quad \text{و} \quad w_2 = 21.25$$

3. حساب الثابت: w_0

المعادلة المستخدمة لحساب w_0 هي:

$$w_0 = -\frac{1}{2}\left(\sum_{i=1}^{p} \bar{X}_{i,1}^2 - \sum_{i=1}^{p} \bar{X}_{i,0}^2\right)$$

- حساب: $\sum_{i=1}^{p} \bar{X}_{i,1}^2$

$$\sum_{i=1}^{p} \bar{X}_{i,1}^2 = \bar{X}_{1,1}^2 + \bar{X}_{2,1}^2 = 8.5^2 + 32.5^2 = 72.25 + 1056.25 = 1128.5$$

- حساب: $\sum_{i=1}^{p} \bar{X}_{i,0}^2$

$$\sum_{i=1}^{p} \bar{X}_{i,0}^2 = \bar{X}_{1,0}^2 + \bar{X}_{2,0}^2 = 3.5^2 + 11.25^2 = 12.25 + 126.5625 = 138.8125$$

- حساب الفرق:

$$\sum_{i=1}^{p} \bar{X}_{i,1}^2 - \sum_{i=1}^{p} \bar{X}_{i,0}^2 = 1128.5 - 138.8125 = 989.6875$$

- حساب: w_0

$$w_0 = -\frac{1}{2} \times 989.6875 = -494.84375$$

إذن:

$$w_0 = -494.84375$$

4. بناء دالة التمييز النهائية:

دالة التمييز تصبح:

$$D(X) = 5X_1 + 21.25X_2 - 494.84375$$

5. تصنيف الحالة الجديدة:

لدينا حالة جديدة مع القيم التالية:

$$X_1 = 7 \quad \text{و} \quad X_2 = 22$$

نحسب قيمة $D(X)$ لهذه الحالة:

$$D(X) = 5(7) + 21.25(22) - 494.84375$$
$$D(X) = 35 + 467.5 - 494.84375 = 7.65625$$

اتخاذ القرار:

- إذا كانت: $D(X) > 0$ الحالة تنتمي إلى الفئة $Y = 1$.
- إذا كانت: $D(X) \leq 0$ الحالة تنتمي إلى الفئة $Y = 0$.

في هذه الحالة:

$$D(X) = 7.65625 > 0$$

إذن:

الحالة تنتمي إلى الفئة $Y = 1$

تمرين 4:

شركة تسويقية ترغب في تصنيف العملاء بناءً على سلوكهم الشرائي إلى فئتين:

- $Y = 0$: العملاء غير المهتمين بالعرض.
- $Y = 1$: العملاء المهتمين بالعرض.

تم جمع البيانات حول عينة من 12 عميلًا، وتشمل المتغيرات المستقلة:

- X_1 :متوسط إنفاق العميل الشهري (بالدولار).
- X_2 :عدد المنتجات التي قام بشرائها خلال الشهر الماضي.

البيانات:

حالة	X_1 (انفاق شهري)	X_2 (عدد المنتجات)	الفئة (Y)
1	50	2	0
2	60	3	0
3	70	4	0

حالة	X_1 (انفاق شهري)	X_2 (عدد المنتجات)	(Y) الفئة
4	80	5	1
5	90	6	1
6	100	7	1
7	110	8	1
8	120	9	1
9	40	1	0
10	130	10	1
11	140	11	1
12	30	2	0

المطلوب:

1. استخدم التحليل التمييزي الخطي (LDA) لبناء دالة تمييزية.
2. قم بتصنيف عميل جديد لديه:
   - $X_1 = 95$ (إنفاق شهري).
   - $X_2 = 6$ (عدد المنتجات).

الحل المفصل:

1. حساب المتوسط لكل متغير داخل كل فئة:

الفئة $Y = 0$

- المتوسط لـ X_1:

$$\bar{X}_{1,0} = \frac{\sum_{i=1}^{n_0} X_{1,i}}{n_0} = \frac{50 + 60 + 70 + 40 + 30}{5} = \frac{250}{5} = 50$$

- المتوسط لـ X_2:

$$\bar{X}_{2,0} = \frac{\sum_{i=1}^{n_0} X_{2,i}}{n_0} = \frac{2 + 3 + 4 + 1 + 2}{5} = \frac{12}{5} = 2.4$$

الفئة $Y = 1$

- المتوسط لـ X_1:

$$\bar{X}_{1,1} = \frac{\sum_{i=1}^{n_1} X_{1,i}}{n_1} = \frac{80 + 90 + 100 + 110 + 120 + 130 + 140}{7} = \frac{770}{7} = 110$$

- المتوسط لـ X_2:

$$\bar{X}_{2,1} = \frac{\sum_{i=1}^{n_1} X_{2,i}}{n_1} = \frac{5 + 6 + 7 + 8 + 9 + 10 + 11}{7} = \frac{56}{7} = 8$$

إذن:

$$\bar{X}_{1,0} = 50, \quad \bar{X}_{2,0} = 2.4, \quad \bar{X}_{1,1} = 110, \quad \bar{X}_{2,1} = 8$$

2. حساب معاملات المتغيرات w_1 وw_2

المعادلة العامة لمعاملات المتغيرات هي:

$$w_i = \bar{X}_{i,1} - \bar{X}_{i,0}$$

- بالنسبة لـw_1:

$$w_1 = \bar{X}_{1,1} - \bar{X}_{1,0} = 110 - 50 = 60$$

- بالنسبة لـw_2:

$$w_2 = \bar{X}_{2,1} - \bar{X}_{2,0} = 8 - 2.4 = 5.6$$

إذن:

$$w_1 = 60 \quad \text{و} \quad w_2 = 5.6$$

3. حساب الثابت w_0

المعادلة المستخدمة لحساب w_0 هي:

$$w_0 = -\frac{1}{2}\left(\sum_{i=1}^{p} \bar{X}_{i,1}^2 - \sum_{i=1}^{p} \bar{X}_{i,0}^2\right)$$

- حساب: $\sum_{i=1}^{p} \bar{X}_{i,1}^2$

$$\sum_{i=1}^{p} \bar{X}_{i,1}^2 = \bar{X}_{1,1}^2 + \bar{X}_{2,1}^2 = 110^2 + 8^2 = 12100 + 64 = 12164$$

- حساب: $\sum_{i=1}^{p} \bar{X}_{i,0}^2$

$$\sum_{i=1}^{p} \bar{X}_{i,0}^2 = \bar{X}_{1,0}^2 + \bar{X}_{2,0}^2 = 50^2 + 2.4^2 = 2500 + 5.76 = 2505.76$$

- حساب الفرق:

$$\sum_{i=1}^{p} \bar{X}_{i,1}^2 - \sum_{i=1}^{p} \bar{X}_{i,0}^2 = 12164 - 2505.76 = 9658.24$$

- حساب: w_0

$$w_0 = -\frac{1}{2} \times 9658.24 = -4829.12$$

إذن:

$$w_0 = -4829.12$$

4. بناء دالة التمييز النهائية:

دالة التمييز تصبح:

$$D(X) = 60X_1 + 5.6X_2 - 4829.12$$

5. تصنيف الحالة الجديدة:

لدينا حالة جديدة مع القيم التالية:

$$X_1 = 95 \quad \text{و} \quad X_2 = 6$$

نحسب قيمة $D(X)$ لهذه الحالة:

$$D(X) = 60(95) + 5.6(6) - 4829.12$$
$$D(X) = 5700 + 33.6 - 4829.12 = 904.48$$

6. اتخاذ القرار:

- إذا كانت: $D(X) > 0$ الحالة تنتمي إلى الفئة $Y = 1$.
- إذا كانت: $D(X) \leq 0$ الحالة تنتمي إلى الفئة $Y = 0$.

في هذه الحالة:

$$D(X) = 904.48 > 0$$

إذن:

الحالة تنتمي إلى الفئة $\boldsymbol{Y = 1}$

تمرين 5:

شركة تسويقية ترغب في تصنيف العملاء بناءً على سلوكهم الشرائي إلى فئتين:

- $Y = 0$: العملاء غير المهتمين بالعرض.
- $Y = 1$: العملاء المهتمين بالعرض.

تم جمع البيانات حول عينة من 10 عملاء، وتشمل المتغيرات المستقلة:

- X_1: متوسط عدد الرسائل النصية التي يرسلها العميل شهريًا.
- X_2: متوسط الوقت الذي يقضيه العميل على التطبيق (بالدقائق).

البيانات:

حالة	X_1 (عدد الرسائل)	X_2 (متوسط الوقت)	الفئة(Y)
1	5	10	0
2	6	12	0
3	7	15	0
4	8	20	1
5	9	25	1
6	10	30	1

حالة	X_1 (عدد الرسائل)	X_2 (متوسط الوقت)	الفئة(Y)
7	11	35	1
8	12	40	1
9	4	8	0
10	13	45	1

المطلوب:

1. استخدم التحليل التمييزي الخطي (LDA) لبناء دالة تمييزية.
2. قم بتصنيف عميل جديد لديه:
   - $X_1 = 9$ (عدد الرسائل).
   - $X_2 = 22$ (متوسط الوقت).

الحل المفصل:

1. حساب المتوسط لكل متغير داخل كل فئة:

الفئة $Y = 0$

- المتوسط لـ: X_1

$$\bar{X}_{1,0} = \frac{\sum_{i=1}^{n_0} X_{1,i}}{n_0} = \frac{5 + 6 + 7 + 4}{4} = \frac{22}{4} = 5.5$$

- المتوسط لـ: X_2

$$\bar{X}_{2,0} = \frac{\sum_{i=1}^{n_0} X_{2,i}}{n_0} = \frac{10 + 12 + 15 + 8}{4} = \frac{45}{4} = 11.25$$

الفئة $Y = 1$

- المتوسط لـ: X_1

$$\bar{X}_{1,1} = \frac{\sum_{i=1}^{n_1} X_{1,i}}{n_1} = \frac{8 + 9 + 10 + 11 + 12 + 13}{6} = \frac{63}{6} = 10.5$$

- المتوسط لـ: X_2

$$\bar{X}_{2,1} = \frac{\sum_{i=1}^{n_1} X_{2,i}}{n_1} = \frac{20 + 25 + 30 + 35 + 40 + 45}{6} = \frac{195}{6} = 32.5$$

إذن:

$$\bar{X}_{1,0} = 5.5, \quad \bar{X}_{2,0} = 11.25, \quad \bar{X}_{1,1} = 10.5, \quad \bar{X}_{2,1} = 32.5$$

2. حساب معاملات المتغيرات w_1 ***و*** w_2

المعادلة العامة لمعاملات المتغيرات هي:

$$w_i = \bar{X}_{i,1} - \bar{X}_{i,0}$$

- بالنسبة لـ: w_1

$$w_1 = \bar{X}_{1,1} - \bar{X}_{1,0} = 10.5 - 5.5 = 5$$

- بالنسبة لـ: w_2

$$w_2 = \bar{X}_{2,1} - \bar{X}_{2,0} = 32.5 - 11.25 = 21.25$$

إذن:

$$w_1 = 5 \quad \text{و} \quad w_2 = 21.25$$

3. حساب الثابت: w_0

المعادلة المستخدمة لحساب w_0 هي:

$$w_0 = -\frac{1}{2}\left(\sum_{i=1}^{p} \bar{X}_{i,1}^2 - \sum_{i=1}^{p} \bar{X}_{i,0}^2\right)$$

- حساب: $\sum_{i=1}^{p} \bar{X}_{i,1}^2$

$$\sum_{i=1}^{p} \bar{X}_{i,1}^2 = \bar{X}_{1,1}^2 + \bar{X}_{2,1}^2 = 10.5^2 + 32.5^2 = 110.25 + 1056.25 = 1166.5$$

- حساب: $\sum_{i=1}^{p} \bar{X}_{i,0}^2$

$$\sum_{i=1}^{p} \bar{X}_{i,0}^2 = \bar{X}_{1,0}^2 + \bar{X}_{2,0}^2 = 5.5^2 + 11.25^2 = 30.25 + 126.5625 = 156.8125$$

- حساب الفرق:

$$\sum_{i=1}^{p} \bar{X}_{i,1}^2 - \sum_{i=1}^{p} \bar{X}_{i,0}^2 = 1166.5 - 156.8125 = 1009.6875$$

- حساب: w_0

$$w_0 = -\frac{1}{2} \times 1009.6875 = -504.84375$$

إذن:

$$w_0 = -504.84375$$

4. بناء دالة التمييز النهائية:

دالة التمييز تصبح:

$$D(X) = 5X_1 + 21.25X_2 - 504.84375$$

5. تصنيف الحالة الجديدة:

لدينا حالة جديدة مع القيم التالية:

$$X_1 = 9 \quad \text{و} \quad X_2 = 22$$

نحسب قيمة $D(X)$ لهذه الحالة:

$$D(X) = 5(9) + 21.25(22) - 504.84375$$
$$D(X) = 45 + 467.5 - 504.84375 = 7.65625$$

6.اتخاذ القرار:

- إذا كانت: $D(X) > 0$ الحالة تنتمي إلى الفئة $Y = 1$.
- إذا كانت: $D(X) \leq 0$ الحالة تنتمي إلى الفئة $Y = 0$.

في هذه الحالة:

$$D(X) = 7.65625 > 0$$

إذن:

الحالة تنتمي إلى الفئة $Y = 1$

تمرين 6:

شركة تسويقية ترغب في تصنيف العملاء بناءً على سلوكهم الشرائي إلى فئتين:

- $Y = 0$: العملاء غير المهتمين بالعرض.
- $Y = 1$: العملاء المهتمين بالعرض.

تم جمع البيانات حول عينة من 12 عميلًا، وتشمل المتغيرات المستقلة:

- X_1: متوسط إنفاق العميل الشهري (بالدولار).
- X_2: عدد المنتجات التي قام بشرائها خلال الشهر الماضي.

البيانات:

حالة	X_1 (إنفاق شهري)	X_2 (عدد المنتجات)	الفئة(Y)
1	50	2	0
2	60	3	0
3	70	4	0
4	80	5	1
5	90	6	1
6	100	7	1
7	110	8	1
8	120	9	1
9	40	1	0
10	130	10	1
11	140	11	1
12	30	2	0

المطلوب:

1. استخدم التحليل التمييزي الخطي (LDA) لبناء دالة تمييزية.
2. قم بتصنيف عميل جديد لديه:
   - $X_1 = 95$ (إنفاق شهري).
   - $X_2 = 6$ (عدد المنتجات).

الحل المفصل:

1. حساب المتوسط لكل متغير داخل كل فئة:

الفئة $Y = 0$

- المتوسط لـ: X_1

$$\bar{X}_{1,0} = \frac{\sum_{i=1}^{n_0} X_{1,i}}{n_0} = \frac{50 + 60 + 70 + 40 + 30}{5} = \frac{250}{5} = 50$$

- المتوسط لـ: X_2

$$\bar{X}_{2,0} = \frac{\sum_{i=1}^{n_0} X_{2,i}}{n_0} = \frac{2 + 3 + 4 + 1 + 2}{5} = \frac{12}{5} = 2.4$$

الفئة $Y = 1$

- المتوسط لـ: X_1

$$\bar{X}_{1,1} = \frac{\sum_{i=1}^{n_1} X_{1,i}}{n_1} = \frac{80 + 90 + 100 + 110 + 120 + 130 + 140}{7} = \frac{770}{7} = 110$$

- المتوسط لـ: X_2

$$\bar{X}_{2,1} = \frac{\sum_{i=1}^{n_1} X_{2,i}}{n_1} = \frac{5 + 6 + 7 + 8 + 9 + 10 + 11}{7} = \frac{56}{7} = 8$$

إذن:

$$\bar{X}_{1,0} = 50, \quad \bar{X}_{2,0} = 2.4, \quad \bar{X}_{1,1} = 110, \quad \bar{X}_{2,1} = 8$$

2. حساب معاملات المتغيرات w_1 ***و*** w_2***:***

المعادلة العامة لمعاملات المتغيرات هي:

$$w_i = \bar{X}_{i,1} - \bar{X}_{i,0}$$

- بالنسبة لـ: w_1

$$w_1 = \bar{X}_{1,1} - \bar{X}_{1,0} = 110 - 50 = 60$$

- بالنسبة لـ: w_2

$$w_2 = \bar{X}_{2,1} - \bar{X}_{2,0} = 8 - 2.4 = 5.6$$

إذن:

$$w_1 = 60 \quad و \quad w_2 = 5.6$$

3. حساب الثابت: w_0

المعادلة المستخدمة لحساب w_0 هي:

$$w_0 = -\frac{1}{2}\left(\sum_{i=1}^{p} \bar{X}_{i,1}^2 - \sum_{i=1}^{p} \bar{X}_{i,0}^2\right)$$

- حساب: $\sum_{i=1}^{p} \bar{X}_{i,1}^2$

$$\sum_{i=1}^{p} \bar{X}_{i,1}^2 = \bar{X}_{1,1}^2 + \bar{X}_{2,1}^2 = 110^2 + 8^2 = 12100 + 64 = 12164$$

- حساب: $\sum_{i=1}^{p} \bar{X}_{i,0}^2$

$$\sum_{i=1}^{p} \bar{X}_{i,0}^2 = \bar{X}_{1,0}^2 + \bar{X}_{2,0}^2 = 50^2 + 2.4^2 = 2500 + 5.76 = 2505.76$$

- حساب الفرق:

$$\sum_{i=1}^{p} \bar{X}_{i,1}^2 - \sum_{i=1}^{p} \bar{X}_{i,0}^2 = 12164 - 2505.76 = 9658.24$$

- حساب: w_0

$$w_0 = -\frac{1}{2} \times 9658.24 = -4829.12$$

إذن:

$$w_0 = -4829.12$$

4. بناء دالة التمييز النهائية:

دالة التمييز تصبح:

$$D(X) = 60X_1 + 5.6X_2 - 4829.12$$

5. تصنيف الحالة الجديدة:

لدينا حالة جديدة مع القيم التالية:

$$X_1 = 95 \quad و \quad X_2 = 6$$

نحسب قيمة $D(X)$ لهذه الحالة:

$$D(X) = 60(95) + 5.6(6) - 4829.12$$
$$D(X) = 5700 + 33.6 - 4829.12 = 904.48$$

6. اتخاذ القرار:

- إذا كانت: $D(X) > 0$ الحالة تنتمي إلى الفئة $Y = 1$.

- إذا كانت: $D(X) \leq 0$ الحالة تنتمي إلى الفئة $Y = 0$.

في هذه الحالة:

$$D(X) = 904.48 > 0$$

إذن:

الحالة تنتمي إلى الفئة $Y = 1$

تمرين 7:

شركة تسويق إلكتروني ترغب في تصنيف العملاء بناءً على سلوكهم على الإنترنت إلى فئتين:

- $Y = 0$: العملاء غير المهتمين بالعروض.
- $Y = 1$: العملاء المهتمين بالعروض.

تم جمع البيانات حول عينة من 8 عملاء، وتشمل المتغيرات المستقلة:

- X_1: عدد النقرات على الإعلانات خلال الشهر الماضي.
- X_2: متوسط الوقت الذي يقضيه العميل على صفحة المنتج (بالدقائق).

البيانات:

حالة	X_1 (عدد النقرات)	X_2 (متوسط الوقت)	الفئة(Y)
1	3	5	0
2	4	6	0
3	5	7	0
4	6	10	1
5	7	12	1
6	8	15	1
7	9	18	1
8	10	20	1

المطلوب:

1. استخدم التحليل التمييزي الخطي (LDA) لبناء دالة تمييزية.
2. قم بتصنيف عميل جديد لديه:
   - $X_1 = 7$ (عدد النقرات).
   - $X_2 = 11$ (متوسط الوقت).

الحل المفصل:

1. حساب المتوسط لكل متغير داخل كل فئة:

الفئة $Y = 0$

- المتوسط لـ X_1:

$$\bar{X}_{1,0} = \frac{\sum_{i=1}^{n_0} X_{1,i}}{n_0} = \frac{3+4+5}{3} = \frac{12}{3} = 4$$

- المتوسط لـ: X_2

$$\bar{X}_{2,0} = \frac{\sum_{i=1}^{n_0} X_{2,i}}{n_0} = \frac{5+6+7}{3} = \frac{18}{3} = 6$$

الفئة $Y = 1$

- المتوسط لـ: X_1

$$\bar{X}_{1,1} = \frac{\sum_{i=1}^{n_1} X_{1,i}}{n_1} = \frac{6+7+8+9+10}{5} = \frac{40}{5} = 8$$

- المتوسط لـ: X_2

$$\bar{X}_{2,1} = \frac{\sum_{i=1}^{n_1} X_{2,i}}{n_1} = \frac{10+12+15+18+20}{5} = \frac{75}{5} = 15$$

إذن:

$$\bar{X}_{1,0} = 4, \quad \bar{X}_{2,0} = 6, \quad \bar{X}_{1,1} = 8, \quad \bar{X}_{2,1} = 15$$

2. حساب معاملات المتغيرات w_1 **و** w_2

المعادلة العامة لمعاملات المتغيرات هي:

$$w_i = \bar{X}_{i,1} - \bar{X}_{i,0}$$

- بالنسبة لـ: w_1

$$w_1 = \bar{X}_{1,1} - \bar{X}_{1,0} = 8 - 4 = 4$$

- بالنسبة لـ: w_2

$$w_2 = \bar{X}_{2,1} - \bar{X}_{2,0} = 15 - 6 = 9$$

إذن:

$$w_1 = 4 \quad \text{و} \quad w_2 = 9$$

3. حساب الثابت: w_0

المعادلة المستخدمة لحساب w_0 هي:

$$w_0 = -\frac{1}{2}\left(\sum_{i=1}^{p} \bar{X}_{i,1}^2 - \sum_{i=1}^{p} \bar{X}_{i,0}^2\right)$$

- حساب: $\sum_{i=1}^{p} \bar{X}_{i,1}^2$

$$\sum_{i=1}^{p} \bar{X}_{i,1}^2 = \bar{X}_{1,1}^2 + \bar{X}_{2,1}^2 = 8^2 + 15^2 = 64 + 225 = 289$$

- حساب: $\sum_{i=1}^{p} \bar{X}_{i,0}^2$

$$\sum_{i=1}^{p} \bar{X}_{i,0}^2 = \bar{X}_{1,0}^2 + \bar{X}_{2,0}^2 = 4^2 + 6^2 = 16 + 36 = 52$$

- حساب الفرق:

$$\sum_{i=1}^{p} \bar{X}_{i,1}^2 - \sum_{i=1}^{p} \bar{X}_{i,0}^2 = 289 - 52 = 237$$

- حساب: w_0

$$w_0 = -\frac{1}{2} \times 237 = -118.5$$

إذن:

$$w_0 = -118.5$$

4. بناء دالة التمييز النهائية:

دالة التمييز تصبح:

$$D(X) = 4X_1 + 9X_2 - 118.5$$

5. تصنيف الحالة الجديدة:

لدينا حالة جديدة مع القيم التالية:

$$X_1 = 7 \quad \text{و} \quad X_2 = 11$$

نحسب قيمة $D(X)$ لهذه الحالة:

$$D(X) = 4(7) + 9(11) - 118.5$$
$$D(X) = 28 + 99 - 118.5 = 8.5$$

6. اتخاذ القرار:

- إذا كانت: $D(X) > 0$ الحالة تنتمي إلى الفئة $Y = 1$.
- إذا كانت: $D(X) \leq 0$ الحالة تنتمي إلى الفئة $Y = 0$.

في هذه الحالة:

$$D(X) = 8.5 > 0$$

إذن:

الحالة تنتمي إلى الفئة $Y = 1$

تمرين 8:

شركة تسويق ترغب في تصنيف العملاء بناءً على سلوكهم الشرائي إلى فئتين:

- $Y = 0$: العملاء غير المهتمين بالعرض.
- $Y = 1$: العملاء المهتمين بالعرض.

تم جمع البيانات حول عينة من 8 عملاء، وتشمل المتغيرات المستقلة:

- X_1 :متوسط إنفاق العميل الشهري (بالدولار).
- X_2 :عدد المنتجات التي قام بشرائها خلال الشهر الماضي.
- X_3 :عدد الرسائل التسويقية التي استقبلها العميل.

البيانات:

حالة	X_1 (إنفاق شهري)	X_2 (عدد المنتجات)	X_3 (عدد الرسائل)	الفئة(Y)
1	50	2	3	0
2	60	3	4	0
3	70	4	5	0
4	80	5	6	1
5	90	6	7	1
6	100	7	8	1
7	110	8	9	1
8	120	9	10	1

المطلوب:

1. استخدم التحليل التمييزي الخطي (LDA) لبناء دالة تمييزية.
2. قم بتصنيف عميل جديد لديه:
   - $X_1 = 95$ (إنفاق شهري).
   - $X_2 = 6$ (عدد المنتجات).
   - $X_3 = 7$ (عدد الرسائل).

الحل المفصل:

1. حساب المتوسط لكل متغير داخل كل فئة:

الفئة $Y = 0$

- المتوسط لـ: X_1

$$\bar{X}_{1,0} = \frac{\sum_{i=1}^{n_0} X_{1,i}}{n_0} = \frac{50 + 60 + 70}{3} = \frac{180}{3} = 60$$

- المتوسط لـ: X_2

$$\bar{X}_{2,0} = \frac{\sum_{i=1}^{n_0} X_{2,i}}{n_0} = \frac{2 + 3 + 4}{3} = \frac{9}{3} = 3$$

- المتوسط لـ: X_3

$$\bar{X}_{3,0} = \frac{\sum_{i=1}^{n_0} X_{3,i}}{n_0} = \frac{3 + 4 + 5}{3} = \frac{12}{3} = 4$$

الفئة $Y = 1$

- المتوسط لـ: X_1

$$\bar{X}_{1,1} = \frac{\sum_{i=1}^{n_1} X_{1,i}}{n_1} = \frac{80 + 90 + 100 + 110 + 120}{5} = \frac{500}{5} = 100$$

- المتوسط لـ: X_2

$$\bar{X}_{2,1} = \frac{\sum_{i=1}^{n_1} X_{2,i}}{n_1} = \frac{5 + 6 + 7 + 8 + 9}{5} = \frac{35}{5} = 7$$

- المتوسط لـ: X_3

$$\bar{X}_{3,1} = \frac{\sum_{i=1}^{n_1} X_{3,i}}{n_1} = \frac{6 + 7 + 8 + 9 + 10}{5} = \frac{40}{5} = 8$$

إذن:

$$\bar{X}_{1,0} = 60, \quad \bar{X}_{2,0} = 3, \quad \bar{X}_{3,0} = 4, \quad \bar{X}_{1,1} = 100, \quad \bar{X}_{2,1} = 7, \quad \bar{X}_{3,1} = 8$$

2. حساب معاملات المتغيرات w_1, w_2, w_3 **و:**

المعادلة العامة لمعاملات المتغيرات هي:

$$w_i = \bar{X}_{i,1} - \bar{X}_{i,0}$$

- بالنسبة لـ: w_1

$$w_1 = \bar{X}_{1,1} - \bar{X}_{1,0} = 100 - 60 = 40$$

- بالنسبة لـ: w_2

$$w_2 = \bar{X}_{2,1} - \bar{X}_{2,0} = 7 - 3 = 4$$

- بالنسبة لـ: w_3

$$w_3 = \bar{X}_{3,1} - \bar{X}_{3,0} = 8 - 4 = 4$$

إذن:

$$w_1 = 40, \quad w_2 = 4, \quad w_3 = 4$$

3. حساب الثابت: w_0

المعادلة المستخدمة لحساب w_0 هي:

$$w_0 = -\frac{1}{2}\left(\sum_{i=1}^{p} \bar{X}_{i,1}^2 - \sum_{i=1}^{p} \bar{X}_{i,0}^2\right)$$

- حساب: $\sum_{i=1}^{p} \bar{X}_{i,1}^2$

$$\sum_{i=1}^{p} \bar{X}_{i,1}^2 = \bar{X}_{1,1}^2 + \bar{X}_{2,1}^2 + \bar{X}_{3,1}^2 = 100^2 + 7^2 + 8^2 = 10000 + 49 + 64 = 10113$$

- حساب: $\sum_{i=1}^{p} \bar{X}_{i,0}^2$

$$\sum_{i=1}^{p} \bar{X}_{i,0}^2 = \bar{X}_{1,0}^2 + \bar{X}_{2,0}^2 + \bar{X}_{3,0}^2 = 60^2 + 3^2 + 4^2 = 3600 + 9 + 16 = 3625$$

- حساب الفرق:

$$\sum_{i=1}^{p} \bar{X}_{i,1}^2 - \sum_{i=1}^{p} \bar{X}_{i,0}^2 = 10113 - 3625 = 6488$$

- حساب: w_0

$$w_0 = -\frac{1}{2} \times 6488 = -3244$$

إذن:

$$w_0 = -3244$$

4. بناء دالة التمييز النهائية:

دالة التمييز تصبح:

$$D(X) = 40X_1 + 4X_2 + 4X_3 - 3244$$

5. تصنيف الحالة الجديدة:

لدينا حالة جديدة مع القيم التالية:

$$X_1 = 95, \quad X_2 = 6, \quad X_3 = 7$$

نحسب قيمة $D(X)$ لهذه الحالة:

$$D(X) = 40(95) + 4(6) + 4(7) - 3244$$
$$D(X) = 3800 + 24 + 28 - 3244 = 608$$

6. اتخاذ القرار:

- إذا كانت: $D(X) > 0$ الحالة تنتمي إلى الفئة $Y = 1$.
- إذا كانت: $D(X) \leq 0$ الحالة تنتمي إلى الفئة $Y = 0$.

في هذه الحالة:

$$D(X) = 608 > 0$$

إذن:

الحالة تنتمي إلى الفئة $\boldsymbol{Y = 1}$