

蒙地卡羅模擬實驗: 學習器的評比

林貫原

January 11, 2024

機器學習方法在前幾章節中被介紹，包括線性判別式分析 (LDA)、二次判別式分析 (QDA)、K-鄰近演算法 (KNN) 以及類神經網路 (ANN) 等四種不同的學習器。每一種學習器都有其適合分析的資料形式。而本章節就是要透過觀察多組資料，比較不同學習器在配適和預測結果上的表現，進而更深入了解哪種學習器適合處理特定類型的資料。

1 資料評比

本節所評比的資料為第五章節中所生成的。因為有些資料是刻意生成不適合做分群的，因此本章節不會將所有資料都做評比，僅挑選合適的幾筆資料，以下會先一一稍微講述使用的資料並放上其散佈圖，之後再放上評比的結果，最後在做總比較。

1.1 兩群資料

1.1.1 資料 1

首先第一筆資料是上課提供的 la-1.txt，圖 1 為其散佈圖，表 1 為透過學習器評比後的訓練集誤判率及測試集誤判率，可以發現到 ANN 中，擁有 20 個神經元的模型在訓練和測試階段都表現出色，誤判率相對較低，顯示其對資料的適應性較強。LDA 和 QDA 在兩個階段均呈現穩定的表現，誤判率相對較低，顯示其對資料的判別能力相當穩健。相對而言，KNN 的表現相對較差，特別是 $K=5$ 時，在測試集上的誤判率相對較高，顯示其對這組資料的適應性可能較弱。因此在這次實驗中，ANN(20) 表現最佳，建議在相似的應用場景中考慮使用。

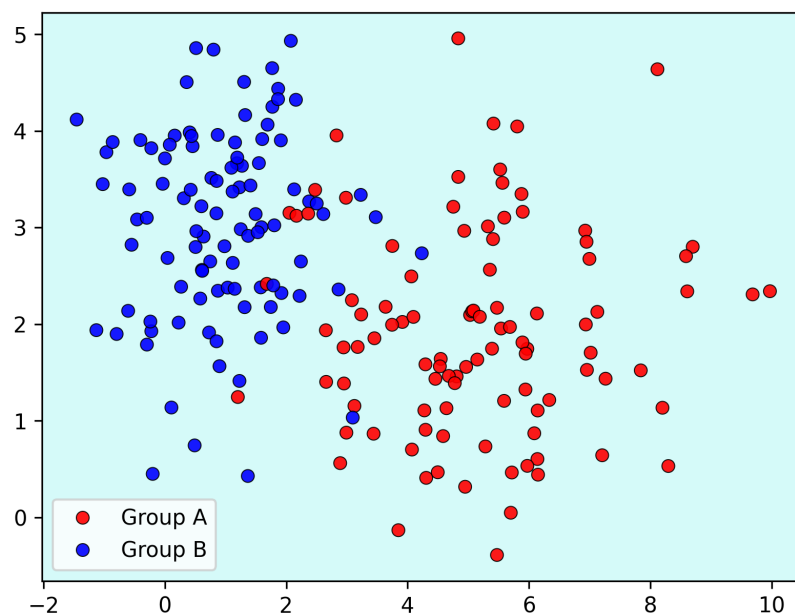


圖 1: 資料 1 散佈圖

表 1: 資料 1 不同學習器評比結果

學習器 誤判率	LDA	QDA	KNN(5)	KNN(15)	ANN(10)	ANN(20)
training error	0.0575	0.0648	0.0542	0.0616	0.0441	0.0395
testing error	0.0632	0.0707	0.0807	0.0680	0.0395	0.0673

1.1.2 資料 4

資料 4 是生成兩組多元常態分佈樣本，第一組的平均值為 $(0, 0)$ ，而第二個群組的平均值為 $(3, 2)$ ，共變異數矩陣皆為 $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ ，且兩組皆包含 200 個樣本，圖 2 為其散佈圖，表 2 為透過學習器評比後的訓練集誤判率及測試集誤判率，根據模擬結果，我們觀察到不同的機器學習模型在訓練和測試階段呈現出相對一致的表現。LDA 和 QDA 在兩個階段都表現優異，誤判率都維持在相對低的水平，顯示其對資料的判別和配適能力較強。KNN 在不同的鄰近數 (K 值) 下呈現出穩定的表現，且相較於其他模型，在測試集上的表現相對較為優越。ANN 的表現也相當穩定，並且在不同的神經元個數下都達到了較低的錯誤率。

總體而言，這些模擬結果表明所選用的機器學習模型在這組資料上均取得了令人滿意的效果。

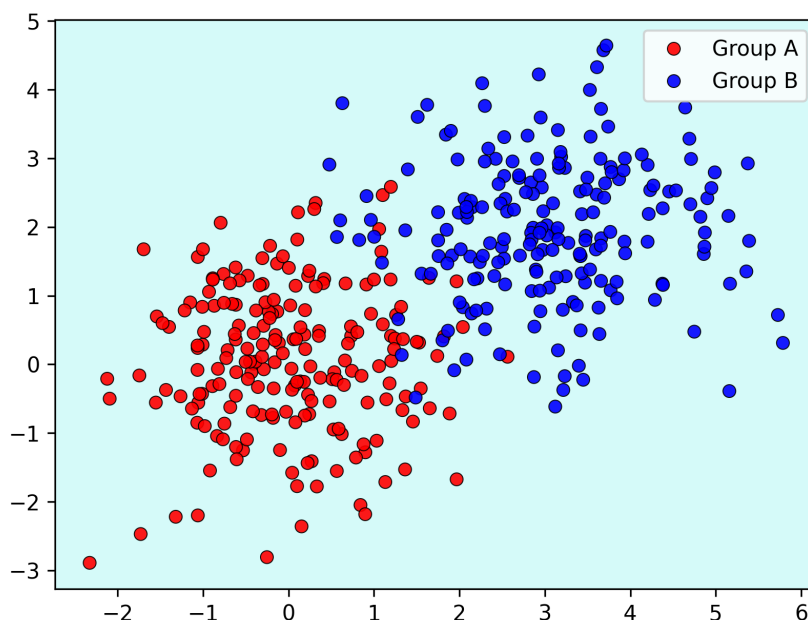


圖 2: 資料 4 散佈圖

表 2: 資料 4 不同學習器評比結果

學習器 誤判率	LDA	QDA	KNN(5)	KNN(15)	ANN(10)	ANN(20)
training error	0.0402	0.0405	0.0386	0.0454	0.0437	0.0447
testing error	0.0455	0.0461	0.0593	0.0523	0.0446	0.0453

1.1.3 資料 5

資料 5 是生成兩組多元常態分佈樣本，第一組的平均值為 $(0, 0)$ ，包含 300 個樣本，共變異數矩陣為 $\begin{bmatrix} 3 & 2 \\ 2 & 3 \end{bmatrix}$ ，而第二個群組的平均值為 $(3, 2)$ ，包含 100 個樣本，共變異數矩陣為 $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ ，圖 3 為其散佈圖，表 2 為透過學習器評比後的訓練集誤判率及測試集誤判率，觀察模擬結果，我們發現不同的機器學習模型在

訓練和測試階段呈現出一致的趨勢。LDA 和 QDA 在兩個階段的誤判率都相對較高，顯示對於這組資料，這兩個模型的配適能力可能受到一定的挑戰。KNN 在訓練集上呈現相對低的誤判率，但在測試集上的誤判率相對提升，顯示可能存在過度配適的現象。ANN 的表現也呈現出較高的誤判率，這可能是因為模型過於複雜或者需要進一步的參數調整。

總體而言，這些結果提示我們在使用這些模型時需要謹慎選擇，並且可能需要進一步的調優和優化。

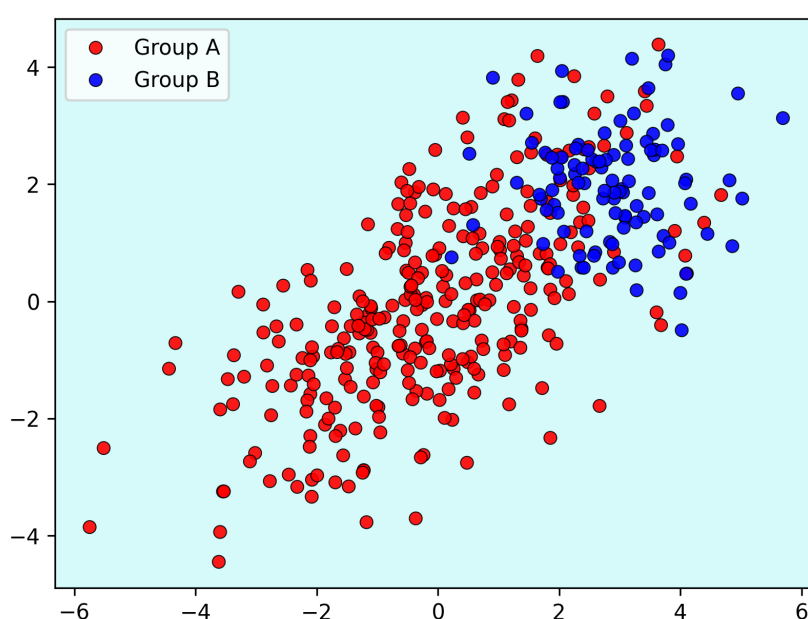


圖 3: 資料 5 散佈圖

表 3: 資料 5 不同學習器評比結果

學習器 誤判率	LDA	QDA	KNN(5)	KNN(15)	ANN(10)	ANN(20)
training error	0.1051	0.1068	0.0888	0.1062	0.1097	0.1091
testing error	0.1119	0.1146	0.1166	0.1163	0.1181	0.1153

1.1.4 資料 6

資料 6 是生成兩組多元常態分佈樣本，第一組的平均值為 $(-1, 0)$ ，共變異數矩陣為 $\begin{bmatrix} 4.5 & 4 \\ 4 & 4.5 \end{bmatrix}$ ，而第二個群組的平均值為 $(4, 2)$ ，共變異數矩陣為 $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ ，且兩組皆包含 200 個樣本，圖 4 為其散佈圖，表 4 為透過學習器評比後的訓練集誤判率及測試集誤判率，觀察模擬結果，我們發現不同的機器學習模型在訓練和測試階段呈現出一致的趨勢。LDA 和 QDA 在兩個階段的誤判率都相對較高，顯示對於這組資料，這兩個模型的配適能力可能受到一定的挑戰。KNN 在訓練集上呈現相對低的誤判率，但在測試集上的誤判率相對提升，顯示可能存在過度配適的現象。ANN 的表現也呈現出較高的誤判率，這可能是因為模型過於複雜或者需要進一步的參數調整。

總體而言，這些結果提示我們在使用這些模型時需要謹慎選擇，並且可能需要進一步的調優和優化。

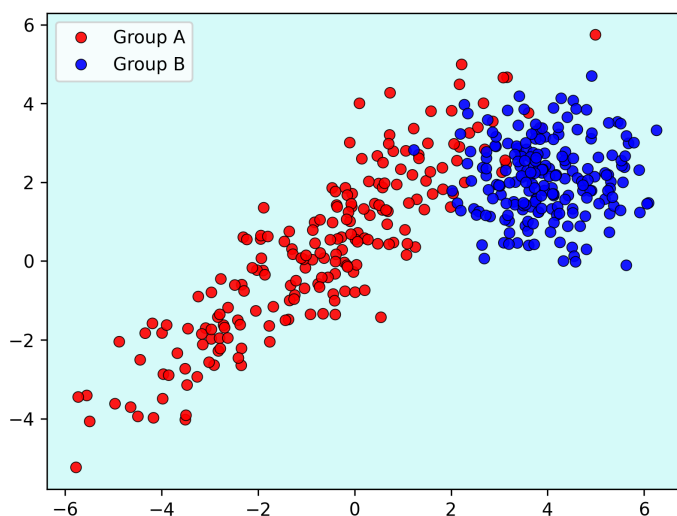


圖 4: 資料 6 散佈圖

表 4: 資料 6 不同學習器評比結果

學習器	LDA	QDA	KNN(5)	KNN(15)	ANN(10)	ANN(20)
誤判率						
training error	0.0416	0.0378	0.0444	0.0459	0.0403	0.0409
testing error	0.0435	0.0395	0.0521	0.0504	0.0486	0.0477

1.2 三群資料

1.2.1 資料 7

資料 6 是生成三組多元常態分佈樣本，第一組的平均值為 $(0.5, -0.2)$ ，包含 300 個樣本，共變異數矩陣為 $\begin{bmatrix} 2 & 0.3 \\ 0.3 & 1 \end{bmatrix}$ ，第二個群組的平均值為 $(2, 2)$ ，包含 200 個樣本，共變異數矩陣為 $\begin{bmatrix} 1 & 0.2 \\ 0 & 1 \end{bmatrix}$ ，而第三個群組的平均值為 $(-1, 2)$ ，包含 100 個樣本，共變異數矩陣為 $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ ，圖 5 為其散佈圖，表 5 為透過學習器評比後的訓練集誤判率及測試集誤判率，觀察模擬結果，LDA 和 QDA 呈現相對較高的誤判率，這可能顯示這兩種方法並不是最理想的選擇，或者資料的特性不夠符合這兩個模型的假設。KNN 在訓練集上顯示出相對較低的誤判率，但在測試集上的誤判率有所上升。這可能表明模型在訓練時過度配適，導致在新資料上配適能力較差。ANN 在訓練和測試集上均呈現相對較高的誤判率。這可能顯示模型需要更多的優化，或者考慮使用更複雜的網路結構提升性能。

總體而言，這些結果凸顯了模型性能受到多方面因素影響的複雜性，包括模型的複雜度、資料的特性以及參數的選擇。

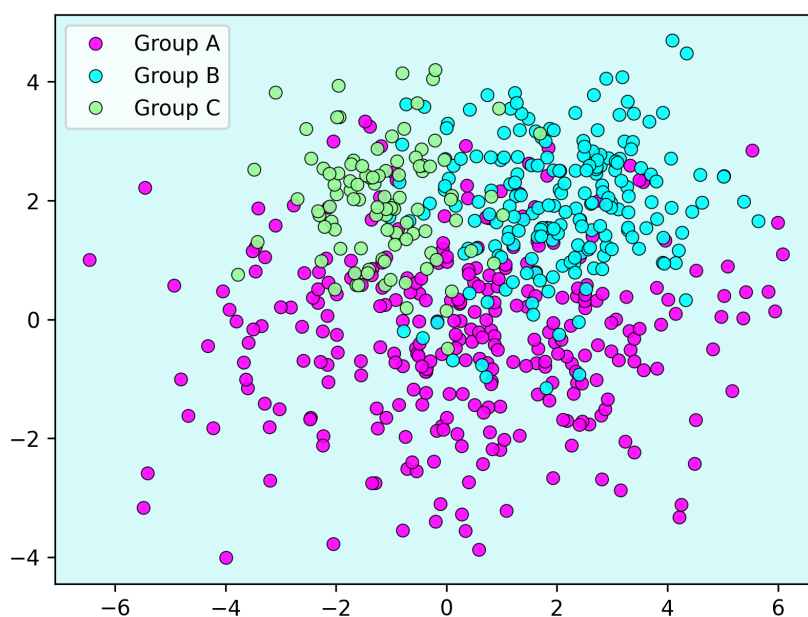


圖 5: 資料 7 散佈圖

表 5: 資料 7 不同學習器評比結果

學習器 誤判率	LDA	QDA	KNN(5)	KNN(15)	ANN(10)	ANN(20)
training error	0.1463	0.1458	0.1246	0.1350	0.1421	0.1418
testing error	0.1433	0.1442	0.1665	0.1422	0.1418	0.1511

1.2.2 資料 8

資料 8 是生成三組多元常態分佈樣本，第一組的平均值為 $(0.5, -0.2)$ ，包含 300 個樣本，共變異數矩陣為 $\begin{bmatrix} 5 & 0.3 \\ 0.3 & 2 \end{bmatrix}$ ，第二個群組的平均值為 $(2, 2)$ ，包含 200 個樣本，共變異數矩陣為 $\begin{bmatrix} 2 & 0.2 \\ 0 & 1 \end{bmatrix}$ ，而第三個群組的平均值為 $(-1, 2)$ ，包含 100 個樣本，共變異數矩陣為 $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ ，圖 6 為其散佈圖，表 6 為透過學習器評比後的訓練集誤判率及測試集誤判率，觀察模擬結果，首先，LDA 的誤判率相對較高，顯示其在訓練和測試階段都沒有很好地配適資料。其次，QDA 相對於 LDA 表現稍好，但仍有進一步改進的空間。接著，KNN 在鄰近點數量增加時呈現出誤判率下降的趨勢，但在某些情況下可能產生過配適。最後，ANN 顯示出相對低的誤判率，表現較好且在訓練和測試集上均能保持穩定。

總體而言，這些實驗結果強調了模型性能的依賴性，並凸顯了不同機器學習方法的優勢和劣勢。

表 6: 資料 8 不同學習器評比結果

學習器 誤判率	LDA	QDA	KNN(5)	KNN(15)	ANN(10)	ANN(20)
training error	0.2241	0.1861	0.1669	0.1881	0.1835	0.1819
testing error	0.2272	0.1942	0.2233	0.2116	0.1957	0.1930

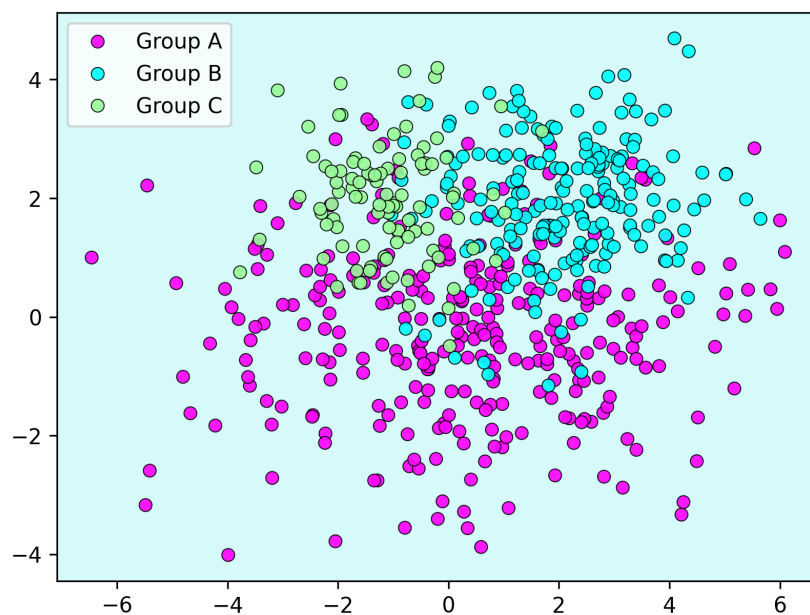


圖 6: 資料 8 散佈圖

1.3 綜合評比

為了更好去比較不同的資料與不同的學習器之間的關聯，以下將前面比較過的所有結果放在一起做比較。

表 7 為六個資料的訓練集在六個學習器模擬中誤判率的比較表格。

表 7: 訓練集誤判率比較表格

Training Error \ DATA	DATA					
	Data1	Data4	Data5	Data6	Data7	Data8
LDA					Worst	Worst
QDA	Worst	Best	Best	Best		
KNN(K=5)					Best	Best
KNN(K=15)		Worst	Worst	Worst		
ANN(10)						
ANN(20)	Best					

圖 7 為六個資料的訓練集在六個學習器模擬中誤判率的折線圖。

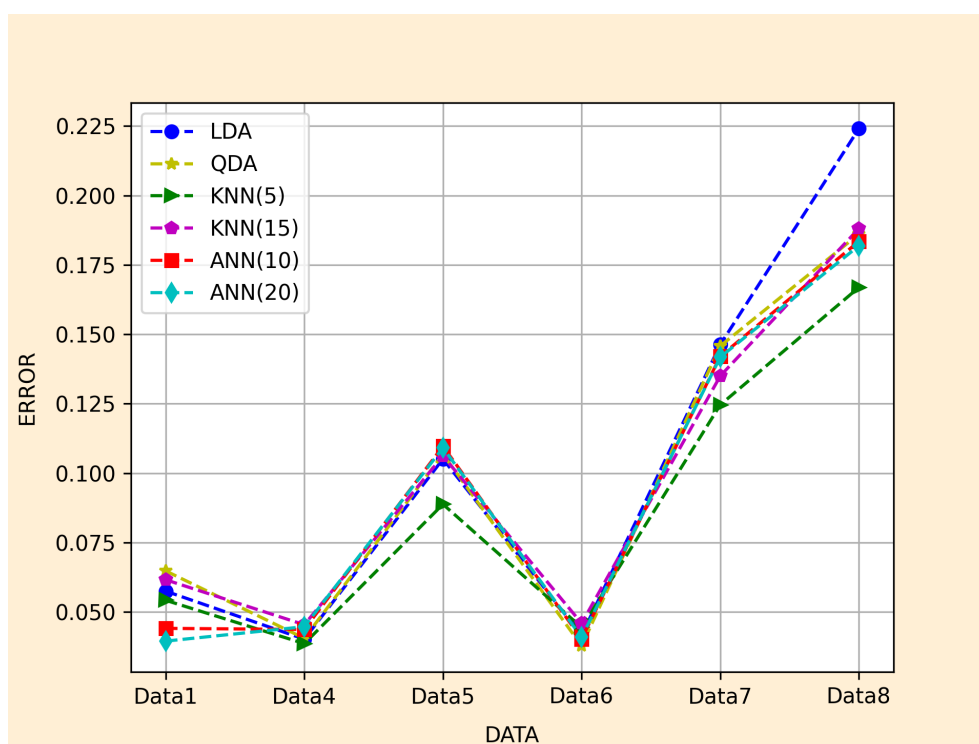


圖 7: 訓練集誤判率折線圖

表 8 為六個資料的測試集在六個學習器模擬中誤判率的比較表格。

表 8: 測試集誤判率比較表格

Testing Error \ DATA	DATA					
	Data1	Data4	Data5	Data6	Data7	Data8
LDA			Best			Worst
QDA				Best		
KNN(K=5)	Worst	Worst		Worst	Worst	
KNN(K=15)						
ANN(10)	Best	Best	Worst		Best	
ANN(20)						Best

圖 8 為六個資料的測試集在六個學習器模擬中誤判率的折線圖。

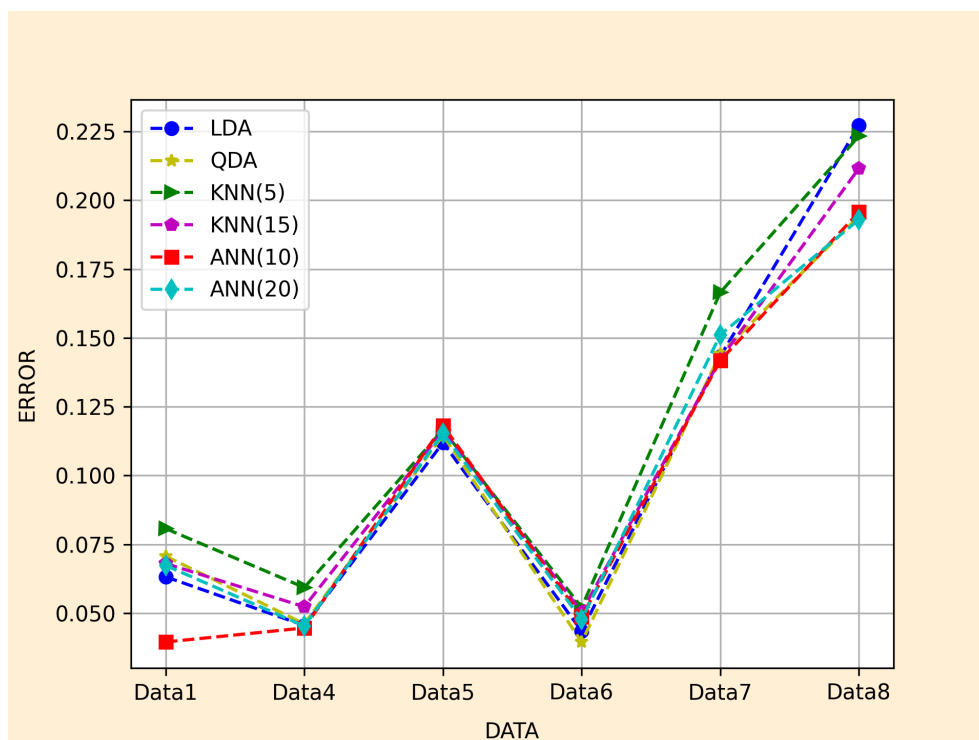


圖 8: 測試集誤判率折線圖

根據以上的表格與折線圖，可初步判斷出 LDA 與 KNN 相較於其他的學習器表現相對較不好，而 QDA 及 ANN 會是比较好的選擇，當然可能因為資料的不同造成不同的結果，但以本章所評比的資料結果就是如此，讀者也可以自行嘗試其他更多種類型的資料做比較。

2 結語

這份蒙地卡羅模擬實驗的文章深入探討了不同機器學習方法在多組資料上的表現比較。通過觀察散佈圖和評比結果，我們得知在特定資料集上，LDA、QDA、KNN 和 ANN 這四種學習器的表現各有優劣。在某些情況下，ANN(20) 表現最佳，而在其他情況下，則可能是其他學習器表現較佳。這凸顯了模型性能受到多方面因素影響的複雜性，包括模型的複雜度、資料的特性以及參數的選擇。因此，在實際應用中，需要根據具體情況選擇最適合的學習器。這份實驗為我們提供了寶貴的見解，幫助我們更好地理解不同機器學習方法的適用範圍，並為未來的相關研究和應用提供了重要的參考依據。