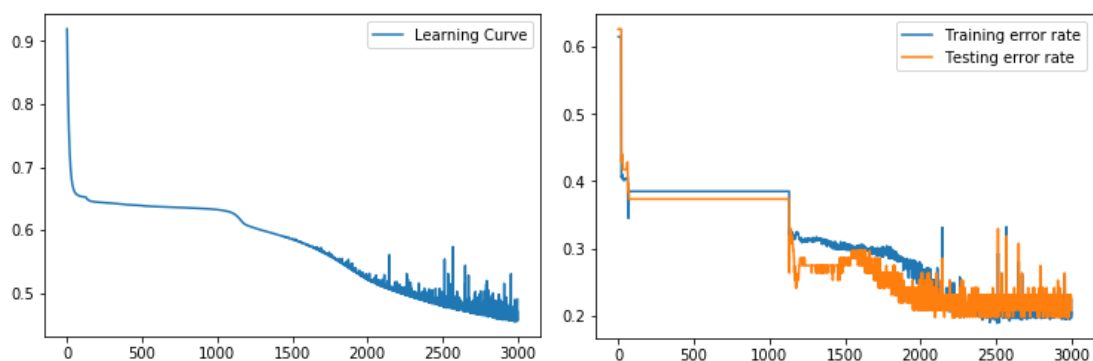


第一題 + 第二題



網路架構

Number of hidden layer		2
類型	Activation	Neurons
DNN	Linear	6
DNN	Sigmoid	3
DNN	Sigmoid	3
DNN	Softmax	2

網路為四層的架構，**橙色層**為 Input Layer，**藍色層**為 Output layer

- **Input layer** 必須為 Linear（即輸入的 Feature 僅經過 w 和 b 計算進入下一層）及 6 個 Neurons（輸入資料有 6 個 features）
- **Output layer** 為 Multi-class classification，所以用 Softmax 作為 Output layer 的 Activation function，並設定為 2 個 Neurons（輸出資料有 2 個 dimension $\rightarrow 01 / 10$ ）
- **Hidden layers** 可以為任何 Activation（在此選用 Sigmoid）

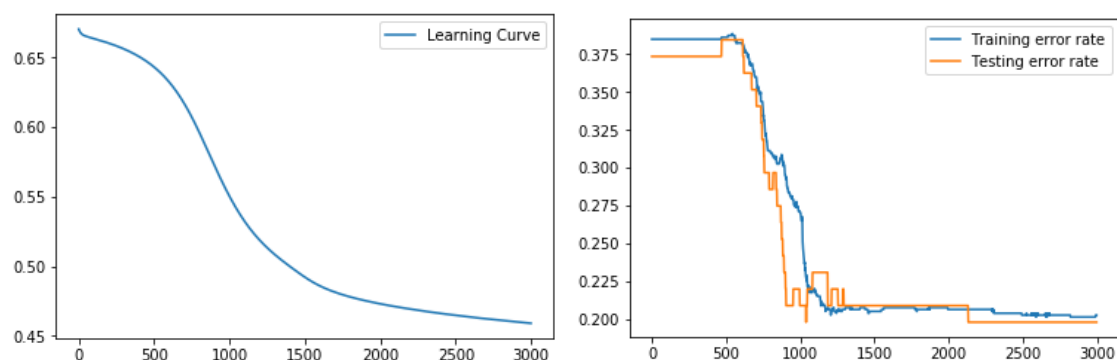
網路參數設定

Loss	Cross entropy
Optimizer	SGD with learning rate 0.01
Epochs	3000
Batch size	20

- **Loss** 由於是 Multi-class classification，所以選用 Cross entropy 作為 Loss
- **Optimizer** 規定使用 SGD 作為 Optimizer
- **Epochs** 大概 3000 個 epochs 就會開始振盪，原因是 Loss 可能已經達到最低點，或卡在 Local minimum
- **Batch size** 每個 minibatch 的 size 為 20

第三題

我認為應該 Normalize 的 Features 有 Fare 及 Age，因為這兩個 Feature 的值域和其他 Features 的相差太多，會造成每個 Feature 的 Gradient 相差甚遠，更新參數不穩定。



以上是 Fare 和 Age 被 Normalized 後，參數更新穩定了許多，網路亦較快收斂。

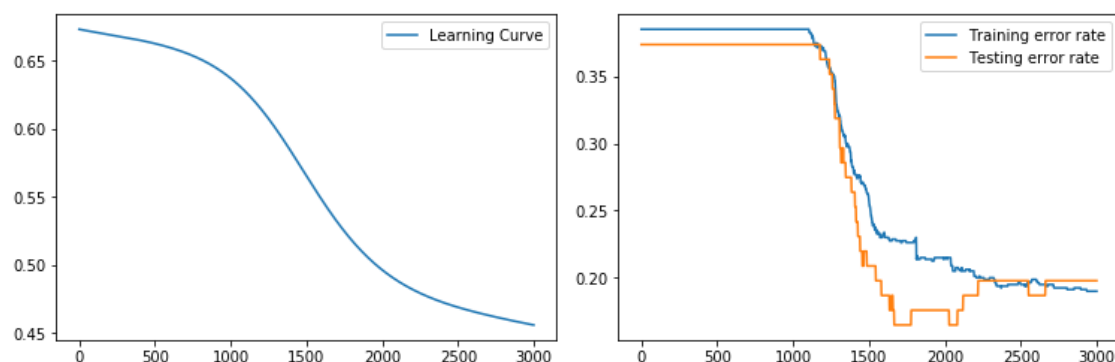
第四題

Index	Survived
Survived	1
Pclass	-0.338481
Sex	-0.543351
Age	0.0105392
SibSp	-0.0353225
Parch	0.0816294
Fare	0.257307

計算 Survived 與其他 Features 的相關係數後，發現 Sex 的絕對數值最大，所以可以判定 Sex 是影響力最高的 Feature。

第五題

Pclass (經濟地位)應該要做 One-hot encoding，原因是每個 Level 之間是互相獨立的，並不包含大小關係。換句話說，當我們只用數值去代表每個 level 時，神經網路會認為這些資料是有序的，但實質上是無序資料，並沒有前後或大小之分。



以上是 Pclass 被 One-hot encoding 後（及 Fare 和 Age 被 Normalized 後）的訓練結果，Learning curve 的 Loss 與第三題沒有太大差異，但 Testing error rate 被降低了不少。

第六題

根據第四題的相關係數表得出以下關係

Feature	生存條件	原因
Sex	當 Sex 為 0，即性別為女性時，生存率越高	在遭遇船難時，女性能優先乘搭救生艇離開
Pclass	當 Pclass 越低，生存率越高	經濟地位低代表其為勞動工作者，體能較佳，生存時間亦越久
Fare	Fare 越高，生存率越高	高花費的遊客大部分為女性，而女性能優先乘搭救生艇離開