

機器學習課程---Final project

Idea :

目前 AI 還無法做到，但我希望 20 年後 AI 能夠做到捕捉大部分並且紀錄下人們在某些時間內的所有一切，包括情緒、感覺、感知、心境、思維、想法、思想、注意力等等。甚至是能有一種裝置戴上後，透過腦波等技術可以重現，讓人們暫時完全地體會當時有紀錄時的一切。就像是強制做了一場很真實的夢一樣，而這個夢是真實發生過的。若被記錄的當事人同意，也可以選擇進入到別人的過去，沉浸式的體會那個角色的一切。

這類似有時聽到一些舊歌，可以瞬間將人拉回到以前當下聽到這首歌的感受、情境、回憶等等，而當下的可能懵懂無知、青春羞澀的片段會被重新喚起。歌曲像是一種媒介，可以乘載過往的某些面向。然而歌曲所乘載的只是一部分，也很容易被新的感受蓋過去或是漸漸淡化掉。

因人腦記憶體有限，無法記得和紀錄，甚至察覺到當下的所有一切。而且即使記得，也常會有記憶扭曲或重構性記憶的情況，也就是記憶中的感受會隨時間改變、模糊或重新詮釋。因此，在這個 Final project 所想要 AI 達成的，是盡可能的保留那些，並且讓我們能夠重新再體會一次。

應用：

1. 人們可以更了解自己是誰，如何更快樂。也可用於審判案件上。
 - 例如：當你有一個創傷，以往透過心理諮詢想要試圖治癒時，會有上述說的記憶扭曲或重構性記憶的情況。如果可以被記錄下來，有助於釐清並較精準的修復傷口。
2. 有時那些被我們遺棄或遺忘的想法或念頭，可以回過頭一一追朔。
 - 例如：在上完課知道 Final project 大致的主題但還沒看到實際題目時，我只想著 20 年後我希望 AI 會有什麼能力，腦中在不同時間會有很多不同的想法冒出來。但當我在寫這個作業的當下，很多想法也被我遺忘了，因記不了太多；也被我自動遺棄了，因為了要回答問題和比較符合框架。如果能夠被記錄下來，就可以回頭去一個個看，或許我能表達得更豐富和清楚。
 - 例如：或許很多數學上的問題早就被解開了，只是想法常常稍縱即逝，數學家來不及捕捉和意識到。
3. 人類可以更有同理心。
 - 例如當一對情侶因金錢觀不同在吵架，若能戴上裝置的暫時變成對方，

體會對方因經歷一切所產生的反應和想法，就更能理解對方為什麼會這樣想。

4. 可以提升和改善既有的思維。
 - 透過完整被記錄下來的所有想法，可以幫助我們去找出既有思維的路徑，進而去做修正和改善；也能夠觀摩高人的思維，去學習，如此人類或許會更有智慧。
5. 或許能促進世上的文明和科學的進步。
 - 有了紀錄，我們就能知道更細微的東西。例如：戴耳機聽音樂比用喇叭可以聽到更多細小的聲音，感受可能會截然不同。雖然只是很細微的東西，累積起來卻可能會有很不相同的感受。或許世上的文明和科學的進步，關鍵點就在這些細微上。
6. 能夠短暫的回到過去。
 - 在戴上裝置後的那段時間，能夠回到過去快樂的時光。例如：失去親人後，當想念時，可以回到親人還在時的時光。

Toy model :

對於上述目前 AI 無法做到的事情，我認為現在實際可行的第一步是先建構一個用來預測當事人情緒的簡化模型。透過當事人在某一情境下可量測的行為或生理特徵，來分析或預測該時刻所具有的情緒。行為或生理狀態可以使用多種數據表示，例如：心跳速度、說話速度、血壓、血清素或催產素濃度、腦內啡濃度、眼眶濕度等。

模型內容：

- 資料來源 $\mathbf{x}=(x_1, x_2, \dots, x_n)$ 為收集或採用的 n 個數值，其中每個 x_i 代表一項可量測的數值。
- 目標標籤 \mathbf{y} 為當事人所標示真實擁有的情緒。若有該情緒就用 1 表示，沒有就用 0 表示，形成一個 multi-hot encoding 的向量。
- 因情緒不是互斥的，有可能是既開心又興奮。因此此模型為多標籤分類模型 (multi-label classification)。
- 由於每筆資料都有明確的答案(情緒標籤)可供模型學習。(至少在不會產生太多複雜情緒下的情況可以知道。)，所以採用監督式學習方法訓練模型。
- 在模型的最後一層的 activation function 採用 sigmoid，將每個情緒對應的輸出值限制在 0 和 1 之間 (每個情緒各自獨立輸出)。得到的數值可視為模型對於該情緒的信心程度或情緒的程度。
- 因每個標籤為獨立的二元判定，因此 Loss function 採用 binary cross entropy，作為模型在訓練時更新參數的評斷標準。

舉例：

假設我的情緒集合為 {開心, 難過, 緊張}。現收集 3 個數據，我在某時刻的當下心跳的速度為 120, 眨眼的速度為 150, 眼眶的濕度為 0.7, 則資料來源 $x=[120, 150, 0.7]$ ；若我當下實際是感到難過和緊張的，則目標標籤 $y=[0,1,1]$ 。而經由模型預測出來的值若是 $y^*=[0.1,0.9,0.8]$ ，表示開心程度為 0.1, 難過程度為 0.9, 緊張程度為 0.8。

Code implementation of Toy model :

◆ Data :

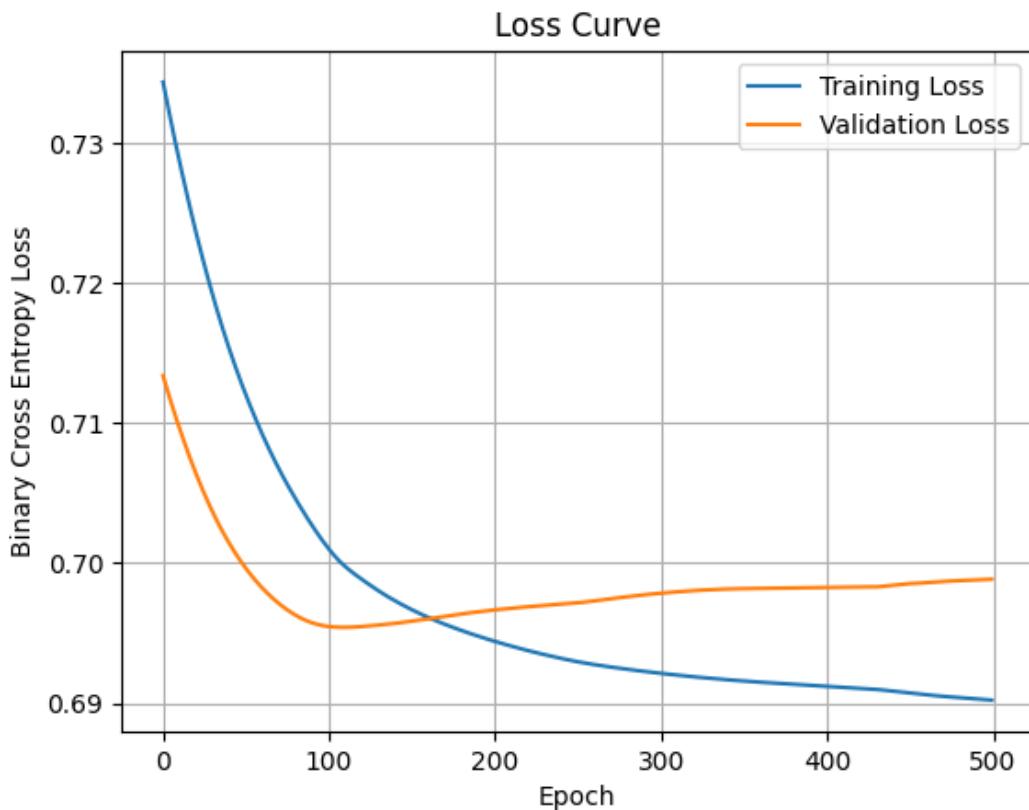
用程式隨機產生 2000 筆資料，分成 Training data (70%), Validation data (15%)，Test data (15%)。且手動輸入 4 筆資料，分別為 $x=[120,150,0.7]$, $y=[0,1,1]$ 、 $x=[70,40,0.9]$, $y=[0,1,0]$ 、 $x=[60,30,0.1]$, $y=[1,0,0]$ 、 $x=[100,130,0.2]$, $y=[1,0,1]$ 。將手動輸入的第一筆資料加入 Test data 用以與預測的結果比較；其餘手動輸入的三筆資料加入 Training data 。

| | Training data | Validation data | Test data |
|------|---------------|-----------------|-----------|
| 資料數量 | 1403 | 300 | 301 |

◆ Hypothesis :

- i. $f: R^3 \rightarrow R^3$ (輸入 3 個測量數值，輸出 3 維向量)
- ii. number of hidden layer : 1
- iii. number of neuron in hidden layer : 8
- iv. activation function : Hidden layer 使用 ReLU ; Output layer 使用 Sigmoid
- v. loss function : binary cross entropy

◆ Result :



```
== Test Example Prediction ==
Input x = tensor([[120.0000, 150.0000,    0.7000]])
True y = tensor([[0., 1., 1.]])
Pred y* = tensor([[0.9200, 0.0589, 0.6357]])
```

由程式結果得知，模型預測放在 Testing data 中的第一筆手動資料 $x=[120, 150, 0.7]$ 的結果不太理想；亦從 Loss 曲線可知，即使跑過完整資料 500 次 (Epoch=500)，Training Loss 仍無法下降至接近 0，而是趨近於 $0.693=\log(2)=-\log(0.5)$ ，表示模型的策略是將預測值會接近 0.5，也就是無法區分 0 和 1，等同於隨機亂猜。

因資料集合中的 2000 筆資料 (x,y) 是隨機生成的，表示 x 與 y 之間不存在任何規律讓模型學習。雖有在 Training data 中額外加入了 3 筆手動資料，但因比例太低，尚不足以讓模型學習到相對正確的判斷。因此，這部分與理論的預測是一致的。

備註：

以下內容為我一開始思考 Final project 的部分思路，內容應該有很多模糊且不正確、混亂的地方，且肯定和原始想法剛出現的內容有些差異，體現了記憶扭曲的現象。我認為能夠保留這些不完整甚至不對的部分，也是重要的紀錄。同時也呼應了我這個 Final project 的主題，因此將相對原始的想法紀錄在這。

想法：

當一件事情或是一個外來的刺激發生時，每個人都會有不同的感受和情緒，那些可能是錯縱複雜的，可能參雜了一些開心、一點釋然、大量惆悵等等，或是複雜到無法用言語形容，甚至一瞬間可能會有成千上萬種感受。然而，隨著時間的過去，我們的心情、情緒、經歷和環境等等，都會影響著記憶中當時感受到的一切。因此當我們在回憶起這些時，往往會失真。(這稱為記憶扭曲或重構性記憶，也就是記憶中的感受會隨時間改變、模糊或重新詮釋。)

但那些本能反應和情緒是最真實且直接的，或許也是能夠代表一個人內在靈魂的象徵。(我認為人到這世上，重要的任務之一是知道自己到底是誰。這是這幾年經歷些事情後，我有所體悟認為的。)

而接在反應和情緒的後面我認為可能是想法，這與人生所經歷的一切有關，是在本能反應後所演化出的。像是小嬰兒肚子餓了就哭，這是當下的本能反應；或許長大經歷哭了就被打，可能就會故意哭的更大聲，或就不哭了。這就是演化出來的，是一種想法。

當人生累積夠多的想法，想法與想法之間的跳躍，我認為就形成了思維。而每個人都有不同的思維，但等年紀變大後，基本上大概就固定了。(就像我現在大概是理工腦，寫下這些也是藉由已經固定的思維去想，甚至連這句話也是，甚至連我說甚至連這句話也是，…，無窮下去)。