<https://chatgpt.com/share/cc1759b0-397d-4fbf-b60a-1e1f315d7e1a>

輸入的資料:

feature:

element的排名?

生成的patch是否為正確修復(plausible或是correct)?

此element是甚麼類型的錯誤(如:value error、type error)?

(可能要找相關論文defects4j錯誤分類)

label:

使用的FL?

使用的APR?

應該是指一個組合，input資料是甚麼FL+APR作為一個label

所以假設有3種FL，4種APR，則設定的input就是這些，然後將他們設定的label可能就是1~12。

如果是FL和APR的組合，那label設定應該是用one-hot方式標記?

(結果應該是一個排名，例如用哪個組合是最好的機率是多少，依序排名下來)。

(所以可能是以排名模型為主，而不用分類模型)

(分類模型代表輸入的資料要用哪個是最好的預測結果)

(排名模型表示這個輸入的資料可以依序使用這些組合能夠達到最好的結果)

實際實作:

可能的方法:

將element有生成出patch的結果加上feature以及label。

選擇一種模型進行訓練(要再考慮是使用甚麼模型進行)。

80%資料訓練、10%驗證、10%測試。

評估訓練結果方法:

loss function做修正(如softmax常使用在multi-lable classfication)，這部分還不太熟悉，不確定要用什麼樣的基準修正，可能需要之後更加理解後再修改。

用輸出的結果->FL與APR的組合效果是否效果最好作為評估，最好的判定以排名(即FL)與patch是否正確(即APR，如果此APR無法生成patch則視為無效)做基準，排名越前面且patch正確，效果則越好。

(不確定這樣能不能，也不確定這樣的假設是否有效)

Label應該是用learning to rank得出排名模型

像是xi輸出得yi，yi這就是他的label

這就算是label。

所以應該要再去看看learning to rank的詳細理解。

因為如果是以圖片辨識，貓、狗等等，是監督式學習的label

但是learning to rank因為模型不同，所以模型不一樣，當初在討論label時候才會不懂。

所以單純是使用的模型不一樣才會造成label認知上的誤差。