

說明：請各位使用此template進行Report撰寫，如果想要用其他排版模式也請註明題號以及題目內容（請勿擅自更改題號），最後上傳前，請務必轉成PDF檔，並且命名為report.pdf，否則將不予計分。

學號：R12945060 系級：生醫電資所碩二 姓名：羅佳蓉

1. (1%) AUC

- a. (1%) 證明 $AUC = \Pr\{f(x^+) > f(x^-)\}$ ，其中 x^+ 代表實際上positive的一個樣本， x^- 代表實際上negative的一個樣本， f 表示模型預測該樣本positive的機率或程度。

ROC 曲線可以通過改變模型的閾值來得到。對於每一個閾值 t ，模型根據 $f(x) > t$ 來決定樣本是正樣本還是負樣本。

True Positive Rate (TPR) 和 False Positive Rate (FPR) 分別定義為：

$$TPR(t) = P(f(x^+) > t)$$

$$FPR(t) = P(f(x^-) > t)$$

AUC 定義為 ROC 曲線下的面積：

$$AUC = \int_0^1 TPR(FPR^{-1}(u)) du$$

將 AUC 表達為概率形式：

我們可以把 AUC 表示為對所有正樣本和負樣本配對進行平均的概率：

$$AUC = P(f(x^+) > f(x^-))$$

要證明這個等式，我們考慮 ROC 曲線的幾何意義。ROC 曲線的每個點代表不同閾值下的 TPR 和 FPR，而 AUC 是所有可能閾值下 TPR 與 FPR 之間的平均關係。這等價於隨機選擇一個正樣本和一個負樣本，並計算模型對這兩個樣本的預測分數的比較結果。

具體化 AUC 的定義：

考慮所有正樣本和負樣本的配對，AUC 可以表示為：

$$AUC = \frac{1}{N^+ N^-} \sum_{i=1}^{N^+} \sum_{j=1}^{N^-} I(f(x_i^+) > f(x_j^-))$$

其中 N^+ 和 N^- 分別是正樣本和負樣本的數量， $I(\cdot)$ 是指示函數，如果條件成立， $I(\cdot)=1$ ，否則為 0。

上述公式表示我們對所有正樣本和負樣本配對進行遍歷，並計算模型是否成功將正樣本的分數排在負樣本之前。這個平均值即為 AUC，也就是模型在區分正負樣本方面的平均表現。

2. (1%) Surrogate Loss

- a. (0.5%) 實作一個surrogate loss並貼上對應的程式碼。

```
class AUROCSurrogateLoss(nn.Module):
    def __init__(self):
        super(AUROCSurrogateLoss, self).__init__()

    def forward(self, y_pred, y_true):
        y_pred = torch.sigmoid(y_pred)
        pos_pairs = (y_true == 1).nonzero(as_tuple=True)[0]
        neg_pairs = (y_true == 0).nonzero(as_tuple=True)[0]

        if len(pos_pairs) == 0 or len(neg_pairs) == 0:
            return torch.tensor(0.0, device=y_pred.device, requires_grad=True)

        loss = 0
        for p in pos_pairs:
            for n in neg_pairs:
                loss += torch.relu(1 - (y_pred[p] - y_pred[n]))

        return loss / (len(pos_pairs) * len(neg_pairs))
```

- b. (0.5%) 請簡短介紹你使用的loss，並比較使用BCELoss跟Surrogate Loss在Public Leaderboard上的表現。

- BCELoss (Binary Cross Entropy Loss)：用於最小化預測概率與實際標籤之間的距離。適合於預測概率需要準確校準的情況。

 predict.csv Complete · 7d ago	0.59920	<input type="checkbox"/>
--	---------	--------------------------

- 代理損失 (Surrogate Loss)：直接針對AUC進行優化，不是關注單個樣本的預測，而是考慮正負樣本對之間的排名。

 predict.csv Complete · 7d ago	0.60271	<input type="checkbox"/>
--	---------	--------------------------

BCELoss：在需要準確校準概率的模型中表現較好。

代理損失：Surrogate Loss在Leaderboard中的AUC指標上優於BCELoss，表明它在高度不平衡的數據集或關注排序性能的任務中可能是更好的選擇。

3. (1%) Feature importance

- a. 使用Ablation study所判斷出來重要的feature有哪些？

在Ablation Study中，我逐一去除每個特徵，並觀察模型的AUC分數變化。根據AUC變化的結果，以下是對模型預測效果最重要的特徵：

類別特徵

- ◆ family_support_good (家庭支持良好)：該特徵對模型性能有顯著的正面影響，表明家庭支持程度是預測結果的重要因素。
- ◆ firstborn (是否為第一胎)：這個特徵表明是否為第一胎對模型預測有較大影響。
- ◆ edu_M_high (母親教育程度高)：母親教育程度越高，對模型預測的影響越顯著。
- ◆ baby_health (嬰兒健康狀況)：嬰兒健康狀況與預測結果密切相關。
- ◆ Vitamin_tril (懷孕三個月內的維他命攝取)：維他命攝取在懷孕早期對模型的預測效果有較大影響。

連續特徵

- ◆ birth_weight (出生體重)：出生體重是影響模型預測的關鍵因素之一。
- ◆ BW_before_D (懷孕前體重)：母親懷孕前的體重對預測結果具有重要影響。
- ◆ Height_M (母親身高)：母親的身高是模型的重要特徵之一。
- ◆ BW_before_P (生產前體重)：母親在生產前的體重對模型的預測也具有顯著影響。
- ◆ Tril_SO2 (懷孕三個月內的SO2暴露量)：SO2的暴露量在懷孕早期對預測具有較大的影響力。

b. 使用Saliency map所判斷出來重要的feature有哪些？

使用Saliency Map分析過程中，我計算了損失函數對於每個特徵的梯度，並根據梯度的絕對值來衡量特徵的重要性。梯度越大，表示該特徵對損失函數的影響越大，因此對模型的預測越重要。

類別特徵

- ◆ family_support_good (家庭支持良好)
- ◆ firstborn (是否為第一胎)
- ◆ Vitamin_tril (懷孕三個月內的維他命攝取)
- ◆ edu_M_high (母親教育程度高)
- ◆ Vitamin_tri23 (懷孕三個月後的維他命攝取)

這些類別特徵對於模型的預測有顯著影響，尤其是家庭支持、是否為第一胎和維他命攝取的情況，對模型的效果起到了關鍵作用。

連續特徵

- ◆ BW_before_D (懷孕前體重)
- ◆ Tril_NOx (懷孕三個月內的NOx暴露量)
- ◆ BW_before_P (生產前體重)
- ◆ Gestation (妊娠期週數)
- ◆ Tri3_CO (懷孕三個月後的CO暴露量)

- c. 如果一起比較categorical feature, continuous feature，你認為是公平的嗎？

比較類別特徵和連續特徵有時可能不公平

尺度差異：連續特徵的範圍通常較大，而類別特徵通常被編碼為整數或one-hot向量，這可能會影響重要性計算。

表示差異：類別特徵通常通過嵌入表示，其特徵空間與連續特徵不一定具有可比性。