

靜宜大學資料科學暨大數據

分析與應用學系

2022畢業專題

【智慧醫療：麻醉深度之預測】

指導教授：李名鏞教授

專題成員：陳柏宇、鄭家奴、林佳怡、方渝淇、
張智雄、陳芳誼、許俊佑

中 華 民 國 1 1 1 年 5 月

摘要

臺灣麻醉的死亡率遠比其他國高出許多，據報導指出，2002~2008年間，臺灣的麻醉相關致死率平均達十萬分之十二，約是日本的12倍、英國的21倍，更是美國的12~24倍。智慧醫療已經是時代的趨勢，透過收集開刀前的一些臨床資料與相關的臨床處置，將運用大數據分析方法建立模型，以預測麻醉深度與麻醉後的生理訊號。協助麻醉醫師監測病人麻醉後的風險，將病人的麻醉風險降到最低，提升醫療品質。

目次

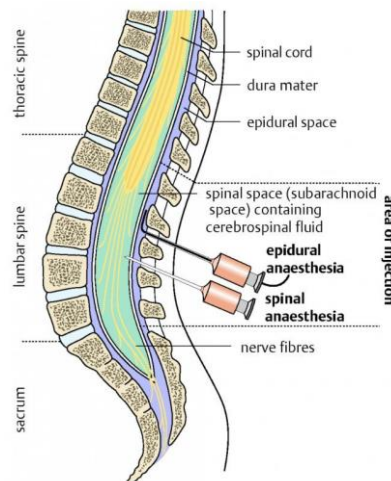
摘要	I
目次	II
第一章 前言	1
第二章 研究方法	4
壹、 使用工具與平台	4
貳、 資料蒐集與名詞解釋	5
參、 特徵工程 (Feature Engineer)	8
肆、 機器學習模型訓練，參數調整與驗證	8
伍、 Overfitting (過度配適之避免)	11
陸、 初步模型評估	11
柒、 評估是否需要進行模型組合(Ensemble)	12
捌、 最終模型評估與交叉驗證	12
第三章 研究步驟	13
壹、 資料清理與前處理	13
貳、 探索性資料分析(EDA)	13
參、 特徵工程	13
肆、 篩選重要變數	14
伍、 建立模型	14
陸、 模型評估	15
第四章 結果	17
壹、 敘述統計與探索性分析	17
貳、 機器學習模型	22
第五章 結論與討論	45
壹、 預測麻醉深度	45
貳、 預測完成半身麻醉10分鐘的收縮壓	45
參、 預測完成半身麻醉10分鐘的舒張壓	45
肆、 預測完成半身麻醉10分鐘的心跳數	45
伍、 預測外科下刀後10分鐘的收縮壓	45
陸、 預測外科下刀後10分鐘的舒張壓	45
柒、 預測外科下刀後10分鐘的心跳數	46
捌、 預測半身麻醉後是否使用到升壓劑	46
玖、 預測半身麻醉後是否使用到止痛藥	46
參考文獻	47

第一章 前言

麻醉通常可以分成兩大類：全身麻醉與區域麻醉；區域麻醉包含**中樞神經阻斷**以及周邊神經阻斷術，主要是利用局部麻醉藥的特性來阻斷神經傳導之過程。局部麻醉藥是經由減少細胞膜上鈉離子通道之通透性，進而抑制神經訊號的傳遞與動作電位的產生，以導致神經相關支配之位置有麻醉的效果。局部麻醉的時效性與麻醉藥物浸泡接觸神經纖維的時間成正比，通常可以阻斷感覺與運動神經的功能，在臨床使用濃度下，不會對神經組織產生永久性的傷害。而臨床上最被廣為使用的中樞神經阻斷技術就是**脊髓麻醉**，也就是通常所謂的**半身麻醉**。

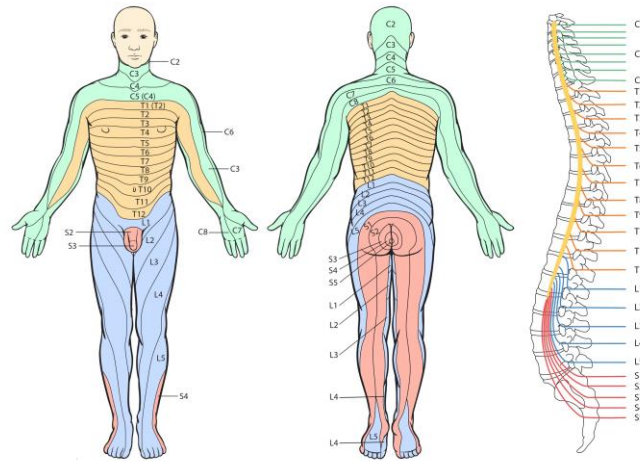
半身麻醉通常使用於手術部位在身體肚臍以下的各式手術，常見用於泌尿科手術、剖腹產手術、腹股溝疝氣修補手術、下肢骨折手術、下肢人工關節置換等等。它的進行方式是在皮膚消毒後，用細長的特殊針具，將特定的幾種局部麻醉藥物，從背後刺入病人脊椎之蜘蛛網膜下腔（subarachnoid space），即腦脊髓液（Cerebrospinal fluid, CSF）存在之處。

局部麻醉藥物會與CSF混合之後，再直接擴散並接觸脊髓旁之神經根，因此所需要的劑量通常較低。利用局部麻醉藥物的效果，可以將負責下肢痛覺與運動的神經一起麻痺，而達到病人下半身暫時無法動彈也沒有感覺的半身麻醉狀態，但區域並不會影響病患之意識。半身麻醉的效果，通常可持續二至四個小時。下圖為半身麻醉施行部位之解剖構造。



在完成半身麻醉之後，麻醉醫護人員會利用麻醉與未麻醉的身體部位對冷熱感覺、疼痛刺激或觸覺的不同感覺，以測試半身麻醉的程度，以確保手術能夠順利的進行。臨床上通常是以**皮節部位**當作半身麻醉程度的表示。通常大部分的下肢手術，需要半身麻醉的程度到達

T10-12的部位，也就是肚臍部位以下皆為麻醉狀態，才可以進行手術，但也有例外，剖腹產手術通常就需要半身麻醉到達T4-T6 的程度，而肛門手術或遠端腳趾手術就只需要到達L4-L5的麻醉程度即可。下圖為皮節之示意圖。



半身麻醉之風險與其麻醉到達皮節程度有相當大的關係，若持續維持在T4 以上的麻醉程度，不只病患會有強烈不適感，常伴隨有噁心嘔吐之發生，也因為抑制交感神經活性，可能會影響呼吸功能，引發心跳下降、心律不整、血壓下降甚至可能會有急救之風險。當然若麻醉程度不夠的話，亦無法進行外科手術。因此如何評估病人情形和外科手術需求，能夠在施行半身麻醉時，給予病人適當麻醉藥物的劑量，以到達合適的麻醉皮節程度，是麻醉科醫師的重要責任。

另外半身麻醉通常是打針一次，同時給予一次性的劑量，臨床上少見有反覆施打之情形，因此，如何能夠在一開始的時候，決定藥物劑量與麻醉程度的關係，更是半身麻醉成功施行的重要關鍵。但臨床上常見不同病人所需要的藥物劑量差異頗大。在臨床上，相同半身麻醉程度時候所需藥物劑量甚至可能會差異到50%！

對於接受麻醉之病人而言，所有麻醉方式均有可能發生副作用或併發症，甚至死亡，例如：面對需要進行救命手術的高齡病患，若本身又有高血壓、糖尿病、心臟相關等疾病時，此時麻醉的風險就如同在大風中走鋼索般，隨時會有掉落之風險。

臨床上，麻醉醫師本身真正能夠控制的因素，只有麻醉藥物的注射總量（體積*濃度）、不同比重麻醉藥物的選擇、打針的位置以及病人的維持姿勢。其中還有有許多不可預測的因素。舉例來說，正常病患CSF 濃度約是1.0003 - 1.0008 g/ml，麻醉醫師可以選擇使用比重較CSF大或比重相同的局部麻醉藥物，作為半身麻醉使用，利用比重不同的藥物在CSF中的

流動傾向不同，來達到不同的麻醉程度，但臨床上我們並無法精確地描述出此類藥物劑量或種類的改變，對於最後病人之半身麻醉程度影響的大小。

有許多研究想要找出真正影響半身麻醉程度的因素，但從1980 年代至今的研究，大多眾說紛紜，比較有共識的部分，除了麻醉醫師可控制的幾個因素外，CSF 的總量與比重也一直被認為是重要影響因素之一，但目前臨床上實際執行半身麻醉時，我們也無法使用這些影響因素，來精確量化並預測半身麻醉的程度。

有鑑於目前並沒有共識的預測模型，臨床上大部分是依照麻醉醫師的臨床經驗來判斷，自然每個人給予的藥量也一定有所差異。過少的劑量，半身麻醉程度不夠，會無法手術；過多的劑量，輕則造成病人不適、術後恢復時間延長、噁心嘔吐機率上升，血壓下降機會上升，重則造成嚴重心律不整，無法呼吸甚至生命危險。

目前麻醉學界對於其相關影響因素雖然已有共識，但各影響因素之權重仍有待討論，而且精確的半身麻醉程度之預測公式模型仍然付之闕如。醫療的將來趨勢逐漸朝向精準化的個人醫療，會針對個體差異給予治療劑量與治療方式的調整。事實上目前的全身麻醉也逐漸朝向精準化的個人醫療，會使用腦電波儀器(BIS, Patient State index 等等)來監測麻醉深度，也有利用ANI(analgesia nociception index)來監測術中的疼痛程度，所以可以更精確調控術中的麻醉藥物使用與麻醉深度。

同樣的，半身麻醉若要朝向更精確的個人化醫療發展，首先勢必要發展經過驗證的半身麻醉程度之預測模型，才能進行個人化劑量的調整。若能成功開發此半身麻醉之預測模型，不但可以因應不同手術需求之半身麻醉程度，給予病患適當之個人化劑量，讓手術順利進行的同時，減少不必要的副作用，例如噁心嘔吐、血壓下降、心跳下降、心律不整等等，亦能增進病人術後恢復速率。

第二章 研究方法

壹、使用工具與平台

一、使用工具：

(一) RStudio

(二) Python

二、套件的使用及介紹

(一) 套件的使用及介紹-R

1. 資料處理－使用套件

(1) Amelia：資料圖形化，查看資料有沒有遺失值

(2) car：將資料分割，以利後續建立特徵

(3) DMwR：資料平衡

2. 特徵工程－使用套件

(1) dummies：將類別變數轉換成虛擬變數

(2) stringr：處理字符串（合併、排序、偵測函式等等）

(3) glmnet：建立Lasso模型，篩選出重要變數

3. 建立模型－使用套件

(1) useful：建立測試集、訓練集

(2) XGBoost：建立XGBoost模型

(3) e1071：建立SVM模型

(4) caret：視覺化混淆矩陣

(5) gbm：建立GBM模型

(6) LightGBM：建立LightGBM模型

(7) pROC：建立Logistic模型

(8) class：建立KNN模型

(9) printr：建立Rule-based classifier模型

(10) C50：建立Rule-based classifier模型

(11) caretEnsemble：交叉驗證

(二) 套件的使用及介紹-Python

1. 資料處理－使用套件
 - (1) sklearn.preprocessing：將變數轉變成索引值
2. 建立模型－使用套件
 - (1) sklearn.model_selection：拆分測試集、訓練集
 - (2) XGBoost.sklearn：建立XGBoost模型
 - (3) catboost：建立CatBoost模型
 - (4) sklearn.metrics：建立混淆矩陣、算準確率並查看預測和實際數值的差異等等
 - (5) sklearn.svm：建立SVM模型
 - (6) sklearn.linear_model：建立Logistic Regression模型
 - (7) sklearn.tree：建立DT模型

貳、資料蒐集與名詞解釋

本專題所有數據資料，由臺北榮總醫院麻醉科提供去識別化資料。

此研究變數介紹：

一、Date：日期

二、Surgery：手術科別，裡面有四個科別:婦產科骨科、泌尿科、大腸直腸外科。

三、ASA：病人麻醉風險等級(1-6)。目前較常用的麻醉風險分類標準是採用「美國麻醉醫師學會」建議的身體狀況分類等級為基準，2001年的定義為：

第一級：正常健康病人。

第二級：有輕微系統疾病但無功能障礙。

第三級：有中等至嚴重程度的系統疾病且導致某些功能障礙。

第四級：有嚴重程度的系統疾病且持續為生命上的威脅及無法行駛功能。

第五級：不管手術治療與否可能活不過24小時的病人。

第六級：即將要捐贈器官的腦死病人。(此案例過於特殊,此研究不納入探討)

四、Emergency：是否為急診手術

五、Age：年紀

六、Height：身高

七、BW：體重

八、BMI：身體質量指數

九、Gender：性別

十、SA_Arrhythmia：下刀後是否有心律不整

十一、Analgesics：半身麻醉完成後，是否使用止痛藥(Fentanyl, Ketamine, NSAID)

十二、Propofol：半身麻醉完成後，是否使用短效靜脈注射麻醉藥(propofol)

十三、Brady_Tx：半身麻醉完成後，是否因心律過慢而使用藥物(atropine, robinu1)

十四、LBW：淨體重計算(去除全部脂肪後的體重一般稱為去脂體重)

十五、HTN：術前是否有高血壓病史

十六、DM：術前是否有糖尿病病史

十七、Arrhythmia：術前是否有心律不整病史

十八、Pregnancy：術前是否懷孕

十九、Ankylosing_Spine：術前是否有僵直性脊椎炎

二十、Spine_OP_Hx：是否曾經接受腰椎手術

二十一、Anal_OP：此次手術是否為痔瘡或肛門瘻管

二十二、Injection Site：施行半身麻醉注射的部位

二十三、Difficult_SA：此次半身麻醉是否為困難半身麻醉

二十四、Patient_Position：半身麻醉注射時病人的姿勢

二十五、In_HR：外科下刀後，第一次量到的心跳數

二十六、In_5_HR：外科下刀5分鐘時，量到的心跳數

二十七、In_10_HR：外科下刀10分鐘時，量到的心跳數

二十八、In_Sat：外科下刀後，第一次量的血氧指數

二十九、IBW：理想體重計算，標準體重反映正常體重較理想和簡單的指標

三十、BSA：身體表面積(為代謝質量比體重更好的指標，因為它較少受到異常的脂肪質量的影響)

三十一、CSF_volume：腦脊髓液量之推測值

三十二、Heavy_Plain：使用Heavy或Plain marcaine作為半身麻醉注射的藥劑。

- 三十三、 LA_dose：半身麻醉注射的藥劑中，marcaine的使用劑量。
- 三十四、 Fentanyl_dose：半身麻醉注射的藥劑中，有無加入fentanyl與marcaine
e 混合的fentanyl劑量。
- 三十五、 Needle_size：施行半身麻醉注射針的大小
- 三十六、 Needle_Bevel：施行半身麻醉注射針的針面方面
- 三十七、 SA_2nd_Dose：施行第二次注射半身麻醉的藥物劑量(若有)
- 三十八、 Experience：施行半身麻醉醫師的年資
- 三十九、 SA_Level_1st：半身麻醉完成後第一次麻醉高度測量
- 四十、 SA_Level_2nd：半身麻醉完成後，第二次麻醉高度測量
- 四十一、 Vasopressor：半身麻醉完成後，是否使用低血壓性休克的血管升壓劑
(ephedrine, Levophed, Epinephrine)
- 四十二、 Vasopressor_dose：半身麻醉完成後，使用升壓劑的總劑量(若有)
- 四十三、 Analgesics_dose：半身麻醉完成後，使用止痛藥的總劑量(若有)
- 四十四、 SA_SBP：進入手術室後，第一次量的收縮壓
- 四十五、 SA_DBP：進入手術室後，第一次量的舒張壓
- 四十六、 SA_5_SBP：完成半身麻醉5分鐘時的收縮壓
- 四十七、 SA_5_DBP：完成半身麻醉5分鐘時的舒張壓
- 四十八、 SA_10_SBP：完成半身麻醉10分鐘時的收縮壓
- 四十九、 SA_10_DBP：完成半身麻醉10分鐘時的舒張壓
- 五十、 SA_HR：進入手術室後第一次量到的心跳數
- 五十一、 SA_5_HR：完成半身麻醉5分鐘時的心跳數
- 五十二、 SA_10_HR：完成半身麻醉10分鐘後時的心跳數
- 五十三、 SA_Sat：進入手術室的第一次血氧指數
- 五十四、 In_SBP：外科下刀後，第一次量到的收縮壓
- 五十五、 In_DBP：外科下刀後，第一次量到的舒張壓
- 五十六、 In_5_SBP：外科下刀5分鐘時，量到的收縮壓
- 五十七、 In_5_DBP：外科下刀5分鐘時，量到的舒張壓
- 五十八、 In_10_SBP：外科下刀10分鐘時，量到的收縮壓

五十九、 In_10_DBP：外科下刀10分鐘時，量到的舒張壓

參、特徵工程 (Feature Engineer)

所謂特徵，在統計中一般稱為解釋變數，而特徵工程簡單說就是將一大堆的解釋變數做一些組合或轉換，得到的新的解釋變數，並用來建立模型，以提升模型的解釋能力。特徵工程的方法，主要包含特徵選擇(Feature Selection)和特徵建構(Feature construction)兩大部分。至於特徵建構，指的是從原始數據中，手動地構建新特徵，透過組合，切割或運算的方式來產生新的特徵，對於各種我們手動或是自動產生出來的特徵，後續仍須透過各種統計或機器學習方法中的特徵選擇，提取出真正重要的特徵，用於後續的模型建立。

肆、機器學習模型訓練，參數調整與驗證

機器學習模型的驗證，可透過幾種不同方法來進行。我們採用的方式是切割出固定比例的訓練資料與驗證資料，透過隨機之方式，將資料切為80%與20%的資料，前者用來進行模型的訓練與驗證(model training and validation)，後者用來進行模型的測試(model testing)。

透過前者訓練出來的模型，可以透過後者來進行驗證。在本研究中預計使用多種機器學習的演算法，以找出在這個問題上表現較佳的數種方法，再看看後續是否有進行模型組合的需要。本研究中將使用的機器學習演算法預計包括：

一、SVM (Support vector machine；支持向量機)

是一種監督式機器學習的分類(classification)演算法，可廣泛地應用於統計分類以及回歸分析。在解決小樣本、非線性及高維模式識別問題中表現出許多特有的優勢，且具有較完備的統計學習理論基礎。我們希望找出一個超平面(hyperplane)，使之將原始的高維度資料中，兩個不同的集合分開。以二維平面來說，我們希望能找出一條線能夠將兩種不同的點分開，而且我們還希望這條線距離這兩個集合的邊界越大越好。

二、XGBoost (Extreme Gradient Boosting)

所謂的 Boosting 就是一種將許多弱學習器集合起來，變成一個更強大的學習器，對於最終預測結果都有更高的準確度。XGBoost (Extreme Gradient Boosting)，是一種擬合殘差的梯度下降算法 Gradient Boosted Tree (GBDT)，每一步學

習是基於之前的錯誤中進步，並會保留原本的模型，並加入新的函數，作為修正上次學習的錯誤，此為集合多個弱學習器。應用方面主要解決監督式學習，可以處理分類也可以解決Regression問題。

三、DT (Decision tree；決策樹)

通過一系列規則對原始資料進行分類的過程。DT是一個預測模型：樹中每個節點表示某個對象，而每個分叉路徑則代表的某個可能的屬性值，而每個葉結點則對應從根節點到該葉節點所經歷的路徑所表示的對象的值。DT僅有單一輸出，若欲有複數輸出，可以建立獨立的DT以處理不同輸出。

四、Regression (迴歸模型)

這是統計領域常用的統計方法之一，可用來解決連續變項或是類別變項的預測問題，透過尋找讓mean square error(MSE)出現最小值時，其個別 feature 所需加權權重的大小的方式建立其模型，在模型建立的過程中也可以進行feature selection，只留下有用的feature來做最終模型的線性組合。

本專題使用兩種迴歸：

邏輯斯迴歸分析(Logistic Regression Analysis)是用來分析與解釋一個名義尺度的反應變數與一個以上的解釋變數間之關係，主要使用反應變數為二元型態之資料。目的是為了要找出類別型態的反應變數和解釋變數之間的關係，因此和簡單線性迴歸分析中最大的差別在於反應變數型態的不同。邏輯斯迴歸在運用上也需符合傳統迴歸分析的一般假設，也就是避免解釋變數之間共線性的問題，以及符合常態分配等的基本假設。因迴歸分析方法中，限制了反應變數為連續型變數，若欲分析的變數非為連續型時則無法使用，此時可選擇邏輯斯迴歸來處理。

線性迴歸(Linear Regression)是利用稱為線性迴歸方程的最小平方函式對一個或多個自變數和因變數之間關係進行建模的一種迴歸分析。這種函式是一個或多個稱為迴歸係數的模型參數的線性組合。只有一個自變數的情況稱為簡單迴歸，大於一個自變數情況的叫做多元迴歸。

五、KNN (K-Nearest Neighbor)

它是監督式學習中分類方法的一種，並可用於分類或迴歸。是一種基於特徵空間中最近訓練實例，對對象進行分類的方法。而訓練樣本是多維特徵空間向量，其中

每個訓練樣本皆帶有一個類別標籤。先在訓練數據集中識別k個觀測值，這些觀測值類似於我們希望分類的新記錄，再使用這些相似的(相鄰)記錄將新記錄分類為一個類。

六、LightGBM(Light Gradient Boosting Machine)

它也是一種常用的機器學習演算法，訓練數據的大小並不會被內存限制。LightGBM提出的主要原因，就是為了解決因多次讀寫訓練數據而所消耗大量時間的問題，其主要想法是利用DT為基底的弱學習器，不斷地迭代訓練並取得最佳的模型。但缺點是少量的資料並不適用此演算法。LightGBM是用最小化損失函數的梯度去建構所有的弱學習器。LightGBM的設計初衷就是提供一個快速高效、低內存占用、高準確度、支持並行和大規模數據處理的數據科學工具。

七、CatBoost

它是Boosting演算法的一種，是一種基於對稱DT (oblivious trees)，為基學習器，實現的參數較少、可使用在類別資料和高準確性的GBDT框架。可以高效並合理地處理類別型特徵，並解決梯度偏差 (Gradient Bias) 及預測偏移 (Prediction shift) 的問題，不但減少overfitting的狀況，還間接提高演算法的準確度和泛化能力。它還有許多優點，像是能夠自動將類別型特徵處理為數值型特徵與缺失值、利用特徵間的相關性，擴大特徵維度等。

八、GBM(Gradient Boosting Machine；梯度提升模型算法)

它是Boosting的其中一種方法。是通過加入新的弱學習器，來努力糾正前面所有弱學習器的殘差，最終這樣多個學習器相加在一起用來進行最終預測，準確率就會比單獨的一個要高。之所以稱為 Gradient，是因為在新增新模型時使用了梯度下降演算法來最小化的損失。

它適用於連續變數與類別變數。不需要對資料進行預處理，也不必插補數值，還可以優化不同的損失函數，並提供幾個超參數調整選項，使函數非常靈活。但它的計算成本偏高，很耗費空間，容易過分強調異常值導致過度擬合，且高靈活性的特質，會致使許多參數相互作用。

九、Rule Based (基於規則)

基於規則通常用於人工智慧應用與研究當中，並用來產生更易於解釋的描述性模型。它的原理是對屬性空間進行直線劃分，並將類指派到每個劃分。如果它允許一條記錄觸發多條規則的話，就可以構造一個更加複雜的決策邊界。此基於類的規則定序方法非常適於處理類分佈不平衡的數據集。而此分類器所產生的規則集，會有互斥規則以及窮舉規則這兩個重要性質。也因為這兩個性質，致使它的模型表達能力幾乎等同於DT。

十、Naive Bayes（樸素貝葉斯）

它是一種十分簡單的分類演算法，思想基礎是：對於給出的待分類項，求解在此項出現的條件下各個類別出現的概率，哪個類別的概率最大，就認為此待分類項屬於哪個類別。對於某些類型的機率模型，在監督式學習的樣本集中能取得非常好的分類效果。在許多實際應用中，樸素貝氏模型參數估計使用最大概似估計方法。換而言之，在不用到貝氏機率或者任何貝葉斯模型的情況下，樸素貝氏模型也能奏效。

伍、Overfitting（過度配適之避免）

為了避免 Overfitting 的出現，我們必須簡化模型的複雜度。一般來說，我們在調整機器學習模型參數的過程中，可以透過 cross validation 的方式來做檢驗，以避免太過複雜的模型，因為這樣的模型顯然在 cross validation 時就會發現它的準確度不同而被淘汰。對於Regression模型來說，可透過調整 Lasso regression的lambda大小的方式，達到正規化的效果，在許多機器學習的模型中，我們也可以透過調整懲罰值的權重，放寬模型對於錯誤的容忍程度，以避免產生過度複雜的模型，對於類神經網路的模型來說，提早讓模型計算終止也是一種避免over-fitting的做法。

陸、初步模型評估

對於目標變數是連續變數，則可以使用Root mean squared error(RMSE)來評估各種不同模型表現的優劣。若是類別型，則分為多分類與二分類兩種，多分類變數使用混淆矩陣及準確率來做評估；二分類則用準確率來評估。

柒、評估是否需要進行模型組合(Ensemble)

依照個別機器學習模型表現的效能優劣，決定是否需要將幾個表現較好的模型進行整合，以近一步提升預測效果。

捌、最終模型評估與交叉驗證

將最終得到的整合模型用於測試資料中，連續型使用交叉驗證評估最終的模型表現效果是否穩定，以及透過交叉驗證的方式尋找最好的結果；多分類變數使用差一類別比率來做評估；而二分類則用敏感度、特異度及交叉驗證來評估。

第三章 研究步驟

壹、資料清理與前處理

資料中若出現缺失值或是其他不合理的數值，會請資料提供者做資料補正。連續變數的離群值使用盒型圖確認，而缺失值部分，需要確認因為人員疏忽沒有登記到或者因為病患不需要做到那方面的檢查。分類變數方面，以次數分配表查看是否有未知的分類出現。

貳、探索性資料分析(EDA)

探索性資料分析主要用來在正規建模或假設檢定作業之外瞭解不同資料之間的相關性，並提供對於資料集變數及其之間相互關係的進一步瞭解。它可以協助識別明顯錯誤，並且更充分地瞭解資料當中的模式、偵測離群值或極端值、尋找變數之間的關係。

透過敘述統計與資料視覺化，對資料進行檢視，我們可以對資料的品質有初步的了解。利用盒鬚圖及散佈圖來檢視資料是否有遺失值或極端值？是否需要進行插補？這些極端值是否來自資料蒐集錯誤？

此外，探索性資料分析尚能幫助我們從資料中看看是否存在甚麼隱含的趨勢，如某些變數間比較存在高度相關性等，對資料有了基本的認識之後，將有利於我們對資料進行清理與前處理的工作，像是重新編碼，或是透過下一階段的特徵工程方法，產生更加複雜的特徵。

參、特徵工程

根據模型，先去探討特徵的合理性，然後對連續型變數進行4種特徵建構(平方、根號、指數、log)，對兩分類跟多分類做one hard encoding跟虛擬變數。最後把資料以兩分類、連續型跟多分類的方式下去做交互作用(分類變數不跟自己做交互作用)。

對於「年齡」，我們並沒有照著WHO規定的方式做分群，而是將它們分成三群，個別是0~65歲為成年人；66~79歲為老年人；80歲以上為高齡。因為這筆資料年輕族群人數很少，與教授和醫師討論後我們決定只分成這三類，並將老年人以及高齡併成一個新的變數。

而「BMI」，我們將其分為六類。分別是18.5以下為過輕；大於等於18.5且小於24為正常範圍；大於等於24且小於27為過重；大於等於27且小於30為輕度肥胖；大於等於30且小於35為中度肥胖；大於等於35為重度肥胖。並將過輕及正常範圍合併；正常範圍及過重合併；過重及輕度肥胖合併；輕度肥胖及中度肥胖合併；中度肥胖及重度肥胖合併；過輕、正常範

圍、過重及輕度肥胖合併；正常範圍、過重及輕度肥胖合併；過重、輕度肥胖及重度肥胖合併。經過上述合併後，總共會額外生成8個新變數。

在實務上前一類BMI誤判成後一類BMI是可以接受的，醫師認為半身麻醉的BMI可以用30作為一個切點，所以我們將過輕、正常範圍、過重及輕度肥胖併為一個新變數，我們覺得正常範圍、過重及輕度肥胖的合併跟過重、輕度肥胖及中度肥胖的合併，它們之間的誤差比較可以接受。

以及「手術科別」，我們不拿它跟其他變數做交互作用。「此次是否為困難半身麻醉」，因為舊資料只有一個值所以沒有拿下去做特徵工程（沒有做特徵工程也就是說沒有做交互作用）。「半身麻醉注射部位」、「半身麻醉注射針之大小」、「半身麻醉注射針之針面方面」它們彼此不做交互作用，因為這幾個變數不會與其他變數產生交互作用，所以它們不需要與其他變數去相乘或做交集。

肆、篩選重要變數

在建立完所有特徵之後，我們透過Lasso進行特徵篩選，找出具有影響力(顯著)的變數，使用Lasso時，會有一個參數需要去調，那就是lambda(懲罰值)。不同的lambda會產生不同的收縮效果，所以我們可以利用Cross Validation的手法，驗證在不同lambda值下模型的表現如何，然後取殘差最小的(表現最好)模型，其所對應的lambda算是比較好的值。變數挑選的話，就要觀察在最佳lambda下，哪些變數的係數不為0。

伍、建立模型

我們將訓練模型的資料切割成80%訓練集跟20%測試集，唯有模型九--預測半身麻醉後是否使用到止痛藥的資料切割成60%訓練集跟40%測試集。

預測麻醉深度在建模的時候，有分成連續型以及分類問題兩種模型，而分類的結果比連續型還要好。在這裡呈現5個較佳的模型，分別為XGBoost、SVM、Logistic Regression、DT以及CatBoost，而上述模型當中，由Logistic Regression的模型表現最好。

預測完成半身麻醉10分鐘的收縮壓、預測完成半身麻醉10分鐘的舒張壓、預測完成半身麻醉10分鐘的心跳數，這三個變數有接近40%左右的NA值，因此在預測時把NA值刪掉剩下兩千七百多筆資料。一開始先用XGBoost和SVM進行預測，接著做交叉驗證，再利用Stacking，

先訓練Linear Regression, SVM, GBM三種模型，再分別利用XGBoost、LightGBM將前面的結果進行第二次預測，最後用平均法整合，但是RMSE卻未如預期中降低。

因為麻醉深度可能會影響血壓改變的幅度，因此完成半身麻醉10分鐘的收縮壓，有將半身麻醉完成後的麻醉高度測量篩選出來，因為預測收縮壓需要參考值，因此有將進入手術室後的第一次收縮壓放入。完成半身麻醉10分鐘的收縮壓看交叉驗證的結果即使利用Stacking, RMSE仍大於10許多，代表預測結果極差。

完成半身麻醉10分鐘的舒張壓篩出的變數跟收縮壓差不多，因為是預測舒張壓，所以有將進入手術室第一次舒張壓放入。完成半身麻醉10分鐘的舒張壓預測出的RMSE稍微大於10，模型預測較差，預測不好的原因可能是前面刪了將近40%的樣本數，導致樣本數不足。

預測外科下刀後10分鐘的收縮壓、預測外科下刀後10分鐘的舒張壓、預測外科下刀後10分鐘的心跳數。將預測結果最好的三個分析方法Linear Regression、SVM、GBM，去做Stacking，將個別預測出來的結果再進一步做XGBoost預測，也做了整合模型將Linear Regression、SVM、GBM算出來的預測值取平均，再去計算RMSE，最後做了XGBoost、GBM、LightGBM整合模型。在求出結果最好的模型後，利用交叉驗證法去勘查利用整合模型是否穩定，然後使用交叉驗證讓模型重複執行一百次，再取跑了一百次RMSE取平均的值。預測外科下刀後10分鐘的收縮壓，使用整合模型（線性迴歸+SVM+GBM）得出的RMSE略大於10一點；預測外科下刀後10分鐘的舒張壓，使用GBM得出的RMSE略小於10；而預測外科下刀後10分鐘的心跳數，則是利用Linear Regression，得出的RMSE略大於5。

預測半身麻醉後是否使用到升壓劑、預測什麼情況下會使用到止痛藥。在有無使用升壓劑或止痛藥的人數比為正常情況下會直接建模，但是成效都不是太好，因此考慮進行資料平衡。在此模型裡分別使用了XGBoost、SVM、Rule based、KNN、Naive Bayes、Logistic Regression與CatBoost這7種不同的演算法進行建模，在這裡我只選出其中3個敏感度較佳的進行顯示，可以發現最好的Logistic Regression敏感度接近70%，而交叉驗證後的敏感度與原先的也沒有差很多，但是我們希望敏感度要到90%以上才適合實際應用。

陸、模型評估

透過交叉驗證、RMSE、敏感度、特異度、準確率來對模型做評估，判斷模型的表現。預測麻醉深度是以準確率與差一類別比率做為評估，在臨床上差一類別比率是可以被接受的，所以我們的評估標準採納差一類別比率。預測完成半身麻醉10分鐘的收縮壓、預測完成

半身麻醉10分鐘的舒張壓、預測外科下刀後10分鐘的收縮壓與預測外科下刀後10分鐘的舒張壓，我們使用交叉驗證以及RMSE作為評估，並以RMSE是否在10以內為評估標準。而預測完成半身麻醉10分鐘的心跳數及預測外科下刀後10分鐘的心跳數，也以交叉驗證以及RMSE作為評估，但RMSE以是否在5以內做為評估標準。預測半身麻醉後是否使用到升壓劑、預測半身麻醉後是否使用到止痛藥以敏感度、特異度及交叉驗證作為評估，而二分類類別判斷標準要達九成以上才能被接受。

第四章 結果

壹、敘述統計與探索性分析

一、連續型變數

利用最小值、第一四分位數、中位數和平均數等值來去勘查此資料各個變數的分布狀態，而標準差表達資料間的離散程度，但不同單位不能互相比較其標準差。我們可以觀察到資料蒐集的年紀為二十歲以上的族群。病人的BMI大致落在22至28之間，為正常範圍與過重之間，BMI高低可能會間接的影響身體健康問題，所以可以從此看來大部分病人通常是健康的。進入手術室的第一次收縮壓與進入手術室的第一次舒張壓最大值分別為273和151，表示兩名病人疑似有高血壓的病史。收縮壓、舒張壓、心跳數的標準差有些差異是正常，因為收縮壓的數值範圍較廣，所以標準差會較大，也影響到後續計算RMSE。

(表1)

變數名稱	年紀	身高	體重	身體質量指數
最小值	20.00	122.60	30.00	12.08
第一四分位數	50.00	154.00	57.10	22.68
中位數	66.00	160.20	65.40	25.38
平均數	61.95	160.86	66.73	25.73
第三四分位數	75.00	167.50	74.60	28.34
最大值	101.00	196.00	152.00	55.36
標準差	18.21	9.40	13.54	4.40

(表2)

變數名稱	理想體重計算	身體表面積	淨體重計算
最小值	18.38	1.12	23.11
第一四分位數	47.82	1.58	37.60
中位數	55.15	1.71	44.94
平均數	55.50	1.72	45.97
第三四分位數	63.29	1.85	53.59
最大值	86.04	2.69	85.21
標準差	10.15	0.21	9.96

(表3)

變數名稱	腦脊髓液量之 推測值	半身麻醉注射， marcaine使用劑量	半身麻醉注射， 與marcaine混合 之fentanyl劑量 (若有)
------	---------------	-------------------------	---

最小值	53.64	5.00	10.00
第一四分位數	67.38	12.00	10.00
中位數	70.09	13.00	15.00
平均數	70.37	12.63	17.81
第三四分位數	73.28	13.00	15.00
最大值	85.75	18.00	70.00
標準差	4.11	1.38	11.98

(表4)

可以在下表看到紅色標起來的數值，表示該名病人疑似有高血壓或心臟病的病史。

變數名稱	進入手術室的第一次收縮壓	進入手術室的第一次舒張壓	進入手術室的第一次心跳數	進入手術室的第一次血氧指數
最小值	67.00	27.00	38.00	74.00
第一四分位數	131.00	72.00	65.00	95.00
中位數	148.00	80.00	74.00	97.00
平均數	148.48	81.06	74.80	96.62
第三四分位數	164.00	89.00	83.00	98.00
最大值	273.00	151.00	157.00	100.00
標準差	24.22	13.20	13.67	2.31

(表5)

變數名稱	完成半身麻醉10分鐘的收縮壓	完成半身麻醉10分鐘的舒張壓	完成半身麻醉10分鐘的心跳數
最小值	10.00	30.00	31.00
第一四分位數	115.00	60.00	65.00
中位數	130.00	70.00	73.00
平均數	132.95	70.60	74.55
第三四分位數	149.00	80.00	82.00
最大值	258.00	120.00	175.00
標準差	24.47	13.67	13.93

(表6)

變數名稱	外科下刀的收縮壓	外科下刀的舒張壓	外科下刀的心跳數
最小值	55.00	25.00	35.00
第一四分位數	109.00	59.00	62.00
中位數	123.00	68.00	70.00
平均數	124.98	66.84	72.72

第三四分位數	140.00	75.00	80.00
最大值	245.00	120.00	180.00
標準差	23.54	13.32	14.43

(表7)

變數名稱	外科下刀10分鐘 的收縮壓	外科下科10分鐘 的舒張壓	外科下刀的10分 鐘的心跳數
最小值	60.00	20.00	38.00
第一四分位數	105.00	56.25	60.00
中位數	120.00	63.00	69.00
平均數	120.68	64.60	70.55
第三四分位數	135.00	72.00	79.00
最大值	245.00	115.00	172.00
標準差	21.65	12.06	14.00

類別型變數

類別型變數利用各類占比來看說，每個類別在此變數中佔的比例為多少，而占比多寡也會影響到後續的資料平衡階段。半身麻醉完成後之麻醉高度測量占比較多的為T4-6、T6-8，所以麻醉之後大多是胸腔或腹部的位置會沒有感覺。從二分類的占比來看，除了性別和術前是否高血壓病史之外，其餘的二分類有跟無占比差異較大，可能會是被視為特殊案例而無法去做後續特徵工程的部分，如：此次是否為困難半身麻醉這個變數，No占比為100%，無法使用在特徵工程上。

(表1)

變數名稱	類別與佔比							
半身麻醉完成 後之麻醉高度 測量	Above T4	T4-6	T6-8	T8-10	T10-12	T12-L2	T12-L2	L2-4
	0.93%	33.63%	33.38%	18.38%	12.49%	1.07%	0.00%	0.02%

(表2)

變數名稱	類別與佔比					
手術科別 (OBS婦產 科、Ortho骨科、GU泌 尿科、CRS大腸直腸外 科)	OBS	Ortho	GU	CRS	GS	PS
	6.77%	56.77%	29.80%	3.67%	1.87%	1.12%
半身麻醉注射部位	L1-2	L2-3	L3-4	L3-5	L4-5	L5-S1
	0.02%	2.35%	92.09%	0.07%	5.18%	0.30%

【智慧醫療：麻醉深度之預測】

身體質量指數	過輕	正常	過重	輕度肥胖	中度肥胖	重度肥胖
	2.85%	24.62%	19.27%	14.98%	11.47%	26.81%

(表3)

變數名稱	類別與佔比			
半身麻醉完成後，是否使用止痛藥(Fentanyl, Ketamine, NSAID)	None	Fentanyl	Ketamine	NSAID
	98.75%	1.25%	0.00%	0.00%
半身麻醉施行醫師之年資，分四組(AboveV10, V10-1, R4-3, R2-1)	Above V10	V10-1	R4-3	R2-1
	12.45%	14.43%	24.62%	48.50%
病人麻醉風險等級(1-5)	1	2	3	4
	13.84%	64.43%	21.39%	0.34%

(表4)

變數名稱	類別與佔比		
半身麻醉注射針之大小	23G	25G	27G
	0.05%	91.77%	0.02%
半身麻醉注射針之針面方面	None	Ephedrine	Levophed
	69.70%	29.91%	0.39%
年齡	成年人	老年人	高齡
	45.81%	35.25%	15.94%

(表5)

變數名稱	類別與佔比	
是否為急診手術	No	Yes
	93.34%	6.66%
性別	Male	Female
	48.91%	51.09%
術前是否高血壓病史	No	Yes
	58.78%	41.22%
術前是否糖尿病病史	No	Yes
	82.72%	17.28%
術前是否心律不整病史	None	arrhythmia
	96.30%	3.70%
術前是否懷孕	No	Yes
	94.12%	5.88%
術前是否有僵直性脊椎炎	No	Yes
	99.95%	0.05%
是否曾經接受腰椎手術	No	Yes
	98.50%	1.50%
此次手術是否為痔瘡或肛門瘻管	No	Yes

	98.75%	1.25%
半身麻醉注射，使用Heavy或Plain marcaine	Heavy	Plain
	86.94%	13.06%
此次是否為困難半身麻醉	No	Yes
	100.00%	0.00%

相關係數

由表1和表2可以看出雙變數的相關性，綠字為P值小於0.001，紅字為P值界於0.001到0.05之間，黑字為P值大於0.05，如：外科下刀的心跳數和完成半身麻醉10分鐘的心跳數、外科下刀的心跳數和外科下刀10分鐘的心跳數，皆為相關係數大於0.7且P值小於0.001。並發現不顯著的地方有很多，而這可能會影響到我們後續建模結果會不好。

(表1)

相關係數	完成半身麻醉 10分鐘的 收縮壓	完成半身麻醉 10分鐘的 舒張壓	完成半身麻醉 10分鐘的 心跳數	外科下刀 10分鐘的 收縮壓	外科下刀 10分鐘的 舒張壓
年紀	0.336	0.012	-0.097	0.224	-0.041
身高	-0.051	0.119	-0.038	0.020	0.169
體重	-0.036	0.078	-0.010	-0.030	0.070
身體質量指數	-0.011	0.007	0.015	-0.054	-0.039
理想體重計算	-0.033	0.129	-0.050	0.042	0.180
身體表面積	-0.044	0.097	-0.021	-0.019	0.102
淨體重計算	-0.008	0.134	-0.058	0.047	0.161
腦脊髓液量之推值	-0.050	0.119	-0.038	0.020	0.169
半身麻醉注射，marcaine使用劑量	-0.390	0.081	-0.021	0.003	0.065
半身麻醉注射，與marcaine混合之fentanyl劑量(若有)	0.056	0.019	-0.029	0.042	-0.013
進入手術室的第一次收縮壓	0.603	0.295	-0.002	0.450	0.206
進入手術室的第一次舒張壓	0.327	0.477	0.206	0.245	0.388
進入手術室的第一次心跳數	-0.035	0.083	0.767	-0.072	0.023
進入手術室的第一次血氧指數	-0.135	-0.044	-0.053	-0.088	0.013
外科下刀的收縮壓	0.605	0.402	-0.043	0.731	0.504
外科下刀的舒張壓	0.361	0.567	0.111	0.517	0.685
外科下刀的心跳數	0.02	0.157	0.796	-0.031	0.106

(表2)

相關係數	外科下刀10分鐘的心跳數	半身麻醉完成後之麻醉高度測量
年紀	-0.109	0.053
身高	-0.009	0.156
體重	-0.021	-0.067
身體質量指數	-0.019	-0.190
理想體重計算	-0.021	0.171
身體表面積	-0.022	-0.018
淨體重計算	-0.047	0.091
腦脊髓液量之推值	-0.009	0.156
半身麻醉注射，marcaine 使用劑量	-0.044	0.015
半身麻醉注射，與marcaine混合 之fentanyl劑量(若有)	-0.033	-0.012
進入手術室的第一次收縮壓	-0.070	-0.041
進入手術室的第一次舒張壓	0.146	-0.031
進入手術室的第一次心跳數	0.702	-0.064
進入手術室的第一次血氧指數	-0.047	-0.026
外科下刀的收縮壓	0.013	0.136
外科下刀的舒張壓	0.158	0.143
外科下刀的心跳數	0.884	-0.002

貳、機器學習模型

一、模型一：預測麻醉深度(SA_Level_1st)

(一) 作法

本部分，對於預測變數SA_Level_1st實施LabelEncode(Above T4:0, L2-4:1, T1 0-12:2, T12-L2:3, T4-6:4, T6-8:5, T8-10:6)，並利用分類模型來進行預測。

(二) 模型建立

模型部分，分別為Logistic Regression、DT、SVM、XGBoost、Catboost等分類模型，利用27個特徵訓練，並於最後預測各分類麻醉深度。模型以正確率以及差一個等級的正確率評估模型的好壞。

(三) 各模型結果

一、SVM預測結果混淆矩陣

預測 真實	Above T4	L2-4	T10-12	T12-L2	T4-6	T6-8	T8-10
Above T4	0	0	0	0	0	0	0
L2-4	0	0	0	0	0	0	0
T10-12	0	0	11	0	1	1	2
T12-L2	0	0	0	0	0	0	0
T4-6	4	0	30	3	182	136	47
T6-8	4	1	79	7	112	156	112
T8-10	0	0	0	0	0	0	0

準確率：38.61%

差一個類別比例：80.52%

二、Logistic Regression預測結果混淆矩陣

預測 真實	Above T4	L2-4	T10-12	T12-L2	T4-6	T6-8	T8-10
Above T4	0	0	0	0	0	0	0
L2-4	0	0	0	0	0	0	0
T10-12	1	0	3	2	3	3	4
T12-L2	0	0	0	0	0	0	0
T4-6	3	0	28	1	180	114	48
T6-8	4	1	75	7	112	172	100
T8-10	0	0	4	0	0	4	9

準確率：41.45%

差一個類別比例：79.38%

三、DT預測結果混淆矩陣

預測 真實	Above T4	L2-4	T10-12	T12-L2	T4-6	T6-8	T8-10
Above T4	0	0	0	0	0	0	0
L2-4	0	0	0	0	0	0	0
T10-12	1	0	5	2	2	7	3
T12-L2	0	0	0	0	0	0	0
T4-6	4	0	33	4	173	129	52
T6-8	3	1	72	4	120	157	106
T8-10	0	0	0	0	0	0	0

準確率：38.15%

差一個類別比例：79.27%

四、XGBoost預測結果混淆矩陣

預測 真實	Above T4	L2-4	T10-12	T12-L2	T4-6	T6-8	T8-10
Above T4	0	0	0	0	0	0	0

L2-4	0	0	0	0	0	0	0
T10-12	2	0	8	2	12	7	8
T12-L2	0	0	2	0	0	0	1
T4-6	4	1	28	2	146	104	51
T6-8	2	0	53	5	124	158	84
T8-10	0	0	19	1	13	24	17

準確率：37.47%

差一個類別比例：76.42%

五、Catboost預測結果混淆矩陣

預測 真實	Above T4	L2-4	T10-12	T12-L2	T4-6	T6-8	T8-10
Above T4	0	0	0	0	0	0	0
L2-4	0	0	0	0	0	0	0
T10-12	3	0	5	3	4	3	5
T12-L2	0	0	0	0	0	0	0
T4-6	5	0	32	2	172	130	52
T6-8	0	1	69	3	117	153	99
T8-10	0	0	4	2	2	7	5

準確率：38.15%

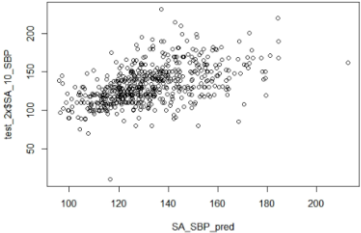
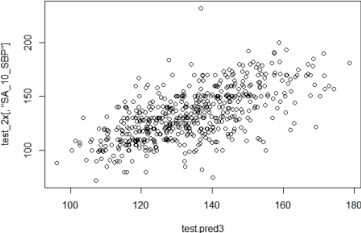
差一個類別比例：78.92%

總結：在於各項分類模型中，對於預測T4-6、T6-8兩類別是敏感的，對於任何可能的類別，基本上被分類於這兩類當中，而兩類別在於各模型當中的準確率皆為在7成以及8成左右，其中以Logistic Regression模型準確率最高，XGB差一個類別比率最小。

二、模型二：完成半身麻醉10分鐘的收縮壓

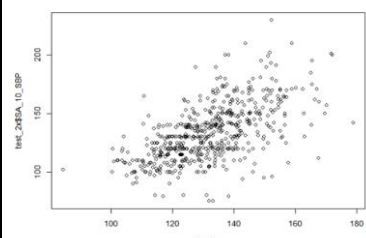
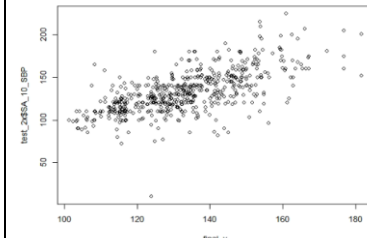
完成半身麻醉10分鐘的收縮壓有1592筆NA值，因此在預測時將NA值刪除，資料最後只剩2794筆資料。一開始我們先使用XGBoost和SVM進行預測，模型預測結果光看RMSE非常不好。而在真實值和預測值的圖形中，看到有些值偏離圖形，且形狀有點胖，且真實值跟預測值差距大於10的數量幾乎佔了一半以上。

	XGBoost	SVM
篩選變數	[1]手術科別為泌尿科 [2]年齡和理想體重的交互作用 [3]年齡和腦脊髓液量之推測值的根號之交互作用 [4]年齡的平方和理想體重的平方之交互作用 [5] BMI取log和術前沒有僵直性脊椎炎 [6]半身麻醉注射針之大小為23G或25G和年齡和理想體重之交互作用	

	[7]半身麻醉注射針之大小為23G或25G和年齡的平方和理想體重之交互作用 [8]半身麻醉注射針之大小為23G或25G或27G和年齡和理想體重之交互作用 [9]半身麻醉注射針之大小為23G或25G或27G和年齡和腦脊髓液量之推測值的平方之交互作用 [10]半身麻醉注射針之針面方面Lateral或Head或Feet和年齡和理想體重之交互作用 [11]半身麻醉注射針之針面方面Lateral或Head或Feet和年齡和腦脊髓液量之推測值的平方之交互作用 [12]此次手術不是痔瘡或肛門瘻管和年齡和理想體重之交互作用 [13]半身麻醉完成後之麻醉高度測量 [14]進入手術室的第一次收縮壓	
RMSE	22.2256	18.2201
預測值-真實值 >10	58.86%	55.28%
圖形(預測值和真實值)		

因為我們對結果不是很滿意，因此我們利用Stacking試圖將RMSE降低。先訓練Linear Regression、SVM、GBM三種模型，在分別利用XGBoost、LightGBM將前面三種模型的預測方法在進行第二次的預測。結果看似有好一些，但是卻仍然是不理想狀態

完成半身麻醉10分鐘的收縮壓	初始模型 1. Linear Regression 2. Svm 3. Gbm 利用XGBoost進行整合	初始模型 1. Linear Regression 2. Svm 3. Gbm 利用LightGBM進行整合
篩選變數	[1]手術科別為泌尿科 [2]年齡和理想體重的交互作用 [3]年齡和腦脊髓液量之推測值的根號之交互作用 [4]年齡的平方和理想體重的平方之交互作用 [5] BMI取log和術前沒有僵直性脊椎炎 [6]半身麻醉注射針之大小為23G或25G和年齡和理想體重之交互作用 [7]半身麻醉注射針之大小為23G或25G和年齡的平方和理想體重之交互作用 [8]半身麻醉注射針之大小為23G或25G或27G和年齡和理想體重之交互作用 [9]半身麻醉注射針之大小為23G或25G或27G和年齡和腦脊髓液量之推測值的平方之交互作用	

	[10] 半身麻醉注射針之針面方面Lateral或Head或Feet和年齡和理想體重之交互作用 [11] 半身麻醉注射針之針面方面Lateral或Head或Feet和年齡和腦脊椎液量之推測值的平方之交互作用 [12] 此次手術不是痔瘡或肛門瘻管和年齡和理想體重之交互作用 [13] 半身麻醉完成後之麻醉高度測量 [14] 進入手術室的第一次收縮壓	
RMSE	19.4504	18.77162
預測值-真實值 >10	56.88%	49.91%
圖形		

為了避免我們是因為運氣問題預測不好，因此我們做了交叉驗證，將模型做了100遍，最後將RMSE進行平均，可以看似種模型的結果都不太好，所以這四種模型都不太推薦，而預測始終不好的原因可能是因為樣本數過少的緣故。

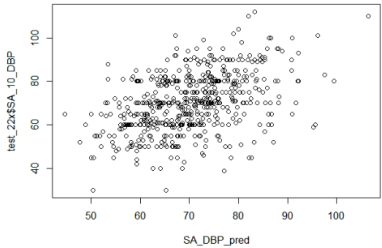
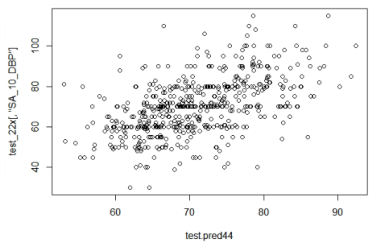
交叉驗證結果

XGBoost	SVM	初始模型 1. Linear Regression 2. SVM 3. GBM 利用XGBoost進行第二次預測	初始模型 1. Linear Regression 2. SVM 3. GBM 利用LightGBM進行第二次預測
21.30275	19.29558	19.4293	19.16953

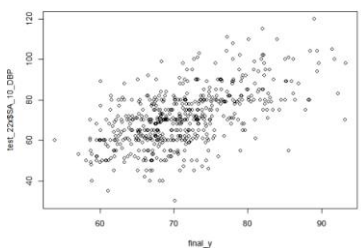
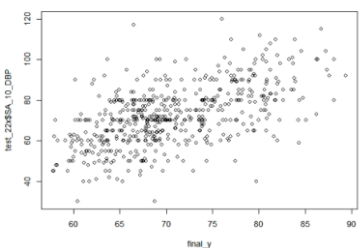
三、 模型三：完成半身麻醉10分鐘的舒張壓

完成半身麻醉10分鐘的舒張壓有1592比NA值，因此在預測時將NA值刪除，資料最後只剩2794筆資料。預測結果看起來勉強還算可行，然而預測值跟真實值差10的比例有40%還是略嫌偏高，且圖形稍嫌離散。

	XGBoost	SVM
篩選變數	[1] 手術科別為骨科 [2] 病人麻醉風險等級為2和性別為男性和理想體重平方之交互作用 [3] 年齡為65歲以上和性別為女性和BMI的平方之交互作用 [4] BMI介於18.5和30之間和術前有高血壓病史和理想體重平方之交互作用	

	[5]性別為女性和半身麻醉注射，使用Heavy marcaine和BMI的平方之交互作用 [6]半身麻醉注射針之大小為25G和年齡在65歲以下和年齡的平方之交互作用 [7]半身麻醉注射針之大小為23G或25G和年齡在65歲以下和年齡的平方之交互作用 [8]半身麻醉注射針之大小為23G或25G和不是急診手術和理想體重的平方之交互作用 [9]半身麻醉完成後之麻醉高度測量 [10]進入手術室的第一次收縮壓 [11]進入手術室的第一次舒張壓	
RMSE	12.4608	11.9113
預測值-真實值 >10	39.71%	38.28%
圖形(預測值和真實值)		

完成半身麻醉10分鐘的舒張壓	初始模型 1.Linear Regression 2.SVM 3.GBM 利用XGBoost進行整合	初始模型 1.Linear Regression 2.SVM 3.GBM 利用LightGBM進行整合
篩選變數	[1]手術科別為骨科 [2]病人麻醉風險等級為2和性別為男性和理想體重平方之交互作用 [3]年齡為65歲以上和性別為女性和BMI的平方之交互作用 [4] BMI介於18.5和30之間和術前有高血壓病史和理想體重平方之交互作用 [5]性別為女性和半身麻醉注射，使用Heavy marcaine和BMI的平方之交互作用 [6]半身麻醉注射針之大小為25G和年齡在65歲以下和年齡的平方之交互作用 [7]半身麻醉注射針之大小為23G或25G和年齡在65歲以下和年齡的平方之交互作用 [8]半身麻醉注射針之大小為23G或25G和不是急診手術和理想體重的平方之交互作用 [9]半身麻醉完成後之麻醉高度測量 [10]進入手術室的第一次收縮壓 [11]進入手術室的第一次舒張壓	
RMSE	11.37679	12.31675

$ \text{預測值}-\text{真實值} >10$	34.88%	38.99%
圖形(預測值和真實值)		

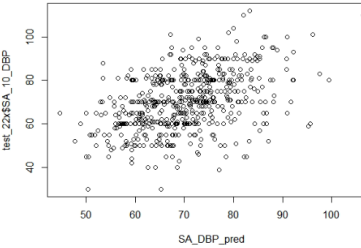
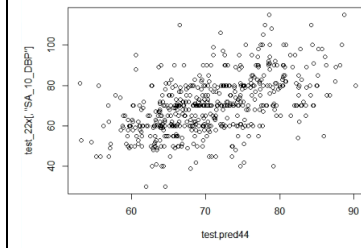
交叉驗證結果

XGBoost	SVM	初始模型 1. Linear Regression 2. SVM 3. GBM 利用XGBoost 進行第二次預測	初始模型 1. Linear Regression 2. SVM 3. GBM 利用LightGBM進行第二 次預測
12.90861	11.78874	11.77794	11.68829

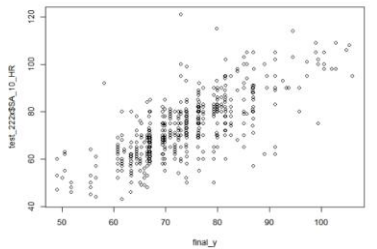
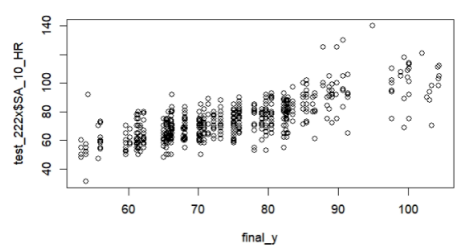
交叉驗證的結果出來，Stacking第二次使用LightGBM結果有降低一些，並且RMSE結果是最低的，因此我們推薦使用先訓練Linear Regression、SVM、GBM三種模型，再分別利用LightGBM將前面三種模型的預測方法在進行第二次的預測。

四、模型四：完成半身麻醉10分鐘的心跳數

為了避免離群值過高影響到模型預測，因此將完成半身麻醉10分鐘的舒張壓離群值為175刪掉，完成半身麻醉10分鐘的舒張壓有1592筆NA值，因此在預測時將NA值刪除，資料最後只剩2799筆資料。而篩出的變數只有一個，可能原因是因為多數變數與預測變數的相關係數連一成都不到，導致篩選不出來。

	XGBoost	SVM
篩選變數	進入手術室的第一次心跳數	
RMSE	12.4608	11.91131
$ \text{預測值}-\text{真實值} >10$	39.71%	38.28%
圖形(預測值和真實值)		

完成半身麻醉10分鐘的心跳數	初始模型 1. Linear Regression 2. SVM	初始模型 1. Linear Regression 2. SVM
----------------	--	--

	3. GBM 利用XGBoost進行整合	3. GBM 利用LightGBM進行整合
篩選變數	進入手術室的第一次心跳數	
RMSE	8.707026	9.24635
$ \text{預測值}-\text{真實值} >10$	20.17%	21.78%
圖形(預測值和真實值)		

交叉驗證結果

XGBoost	SVM	初始模型 1. Linear Regression 2. SVM 3. GBM 利用XGBoost 進行第二次預測	初始模型 1. Linear Regression 2. SVM 3. GBM 利用LightGBM進行第二次 預測
9.05734	8.984841	9.043543	8.827359

交叉驗證的結果出來，Stacking第二次使用LightGBM結果有降低一些，並且RMSE結果是最低的，因此我們推薦使用先訓練Linear Regression、SVM、GBM三種模型，再分別利用LightGBM將前面三種模型的預測方法在進行第二次的預測。

五、 模型五：外科下刀後10分鐘的收縮壓

放入的解釋變數有半身麻醉完成後之麻醉高度測量、進入手術室的第一次收縮壓、進入手術室的第一次舒張壓、進入手術室的第一次心跳數、進入手術室的第一次血氧指數、外科下刀的收縮壓、外科下刀的舒張壓、外科下刀的心跳數、外科下刀的血氧指數和特徵工程。

Lasso變數篩選結果：

- [1]手術科別中的泌尿科(GU)
- [2]進入手術室的第一次收縮壓
- [3]外科下刀的收縮壓

反應變數跟解釋變數的相關性

	手術科別中的 泌尿科(GU)	進入手術室的 第一次收縮壓	外科下刀的收縮壓
--	-------------------	------------------	----------

外科下刀10分鐘 的收縮壓	0.1606	0.4502	0.7312
------------------	--------	--------	--------

表(1)

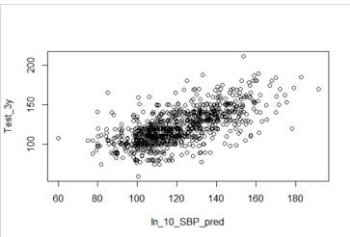
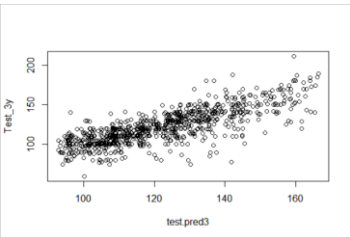
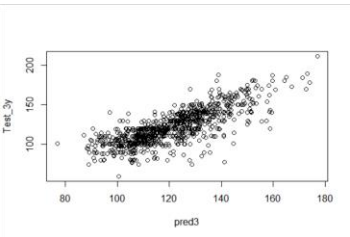
可以從表(1)知道進入手術室的第一次收縮壓(解釋變數)和外科下刀的收縮壓(解釋變數)與外科下刀後10分鐘的收縮壓(反應變數)相關性較高，分別為0.4502 和0.7312，意思為進入手術室的第一次收縮壓(解釋變數)和外科下刀的收縮壓(解釋變數)為影響預測外科下刀後10分鐘的收縮壓(反應變數)的重要因子。

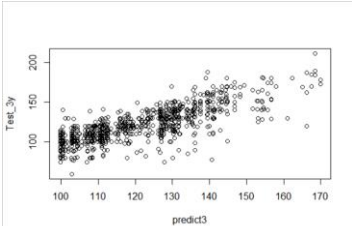
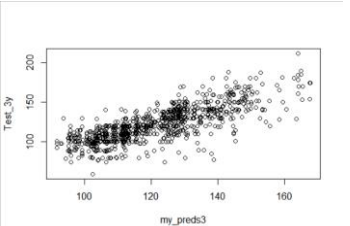
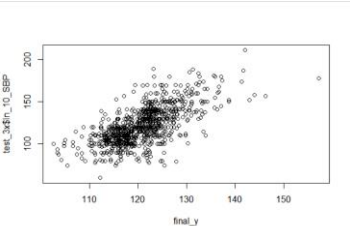
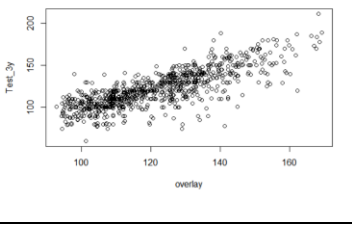
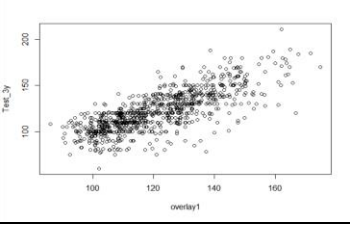
將Lasso篩選出來的變數放入XGBoost、SVM、Linear Regression、GBM、LightGBM、Stacking(Linear Regression、SVM、GBM=> XGBoost)、整合模型(Linear Regression+SVM+GBM)和整合模型(XGBoost+GBM+LightGBM)進行建模與預測，並計算預測值與真實值的差異(均方根誤差RMSE)與計算真實值減預測值絕對值大於10占總比數(測試集)的多少%。

註：1. Stacking(Linear Regression、SVM、GBM=>XGBoost)：Linear Regression、SVM和GBM個別預測出來的結果，再進一步的利用XGBoost去做預測，最後將預測出來的值平均，去計算RMSE。

2. 整合模型(Linear Regression+SVM+GBM)：將Linear Regression、SVM和GBM算出來的預測值平均再去計算RMSE。

3. 整合模型(XGBoost+GBM+LightGBM)：將XGBoost、GBM和LightGBM算出來的預測值平均再去計算RMSE。

XGBoost	SVM	Linear Regression
1. RMSE : 17.306 2. 真實值-預測值 >10的個數/測試集數量*100% : 49.26% 3. 真實值與預測值的散佈圖： 	1. RMSE : 13.4109 2. 真實值-預測值 >10的個數/測試集數量*100% : 39.45% 3. 真實值與預測值的散佈圖： 	1. RMSE : 13.4609 2. 真實值-預測值 >10的個數/測試集數量*100% : 40.25% 3. 真實值與預測值的散佈圖： 
GBM	LightGBM	Stacking(Linear Regression、SVM、GBM=> XGBoost)
1. RMSE : 13.4933	1. RMSE : 13.6553	1. RMSE : 18.1167

<p>2. 真實值-預測值 >10的個數/測試集數量*100%：40.71%</p> <p>3. 真實值與預測值的散佈圖：</p> 	<p>2. 真實值-預測值 >10的個數/測試集數量*100%：40.25%</p> <p>3. 真實值與預測值的散佈圖：</p> 	<p>2. 真實值-預測值 >10的個數/測試集數量*100%：54.96%</p> <p>3. 真實值與預測值的散佈圖：</p> 
<p>整合模型 (Linear Regression+SVM+GBM)</p>	<p>整合模型 (XGBoost+GBM+LightGBM)</p>	<p>補充說明： 五個分析(XGBoost、SVM、Linear Regression、GBM和LightGBM)分法預測出來最好的三個分析(Linear Regression、SVM和GBM)去做Stacking跟整合模型來去勘查建模效果是否有比較好，也想看說整合模型(XGBoost、GBM和Lighgbm)會是如何？而做了此分析。</p>
<p>1. RMSE：13.3121</p> <p>2. 真實值-預測值 >10的個數/測試集數量*100%：39%</p> <p>3. 真實值與預測值的散佈圖：</p> 	<p>1. RMSE：13.9403</p> <p>2. 真實值-預測值 >10的個數/測試集數量*100%：42.42%</p> <p>3. 真實值與預測值的散佈圖：</p> 	

表(2)

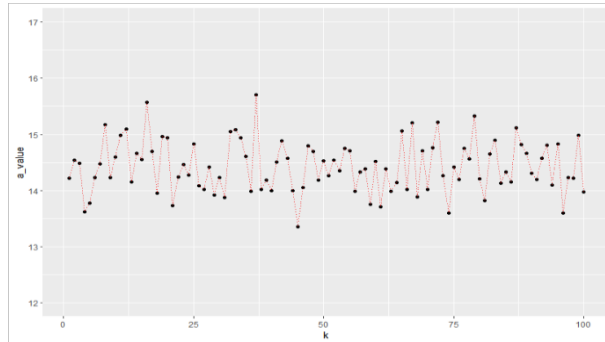
由表(2)的結果可以看出這八種模型以整合模型 (Linear Regression+SVM+GBM)為最佳分析方法，RMSE為13.3121，且真實值減預測值絕對值大於10占總比數(測試集)的39%，也能從散佈圖中看出真實值與預測值之間的關係，此分析方法做出的散佈圖，點更傾向45度角延伸，表示預測出來的值跟真實值很接近，但大部分的點較為鬆散。

將最佳模型(整合模型(Linear Regression+SVM+GBM))做交叉驗證，放入自製的for迴圈當中，讓程式反覆執行一百次，並得出一百個結果(RMSE的值)，然後取平均，勘查模型是否穩定並比較沒有使用交叉驗證和有使用交叉驗證預測出來的結果。

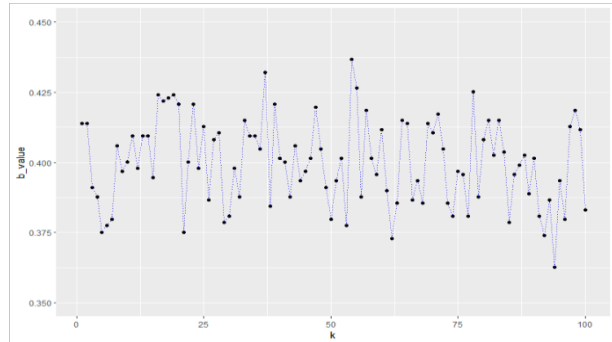
模型疊加(Linear Regression+SVM+GBM)

	RMSE	真實值-預測值 >10的個數/測試集數量*100%
外科下刀的10分鐘的收縮壓	13.3121	39%
外科下刀的10分鐘的收縮壓(交叉驗證)	14.4248	40%

表(3)



圖一、一百次的RMSE結果圖



圖二、一百次的真實值減預測值大於十占比結果圖

從表(3)可以看出沒有做交叉驗證得來的比有做交叉驗證來的好，也可以從圖一圖二得知，做了一百的交叉驗證中有些預測出來的值與真實值相差較大，所以平均之後拉低了RMSE的值與提高了真實值減掉預測值絕對值大於十的占比，並可以從圖一看出利用整合模型(Linear Regression+SVM+GBM)預測一百次的RMSE大致落在13.5到15之間，可以發現在第六十次循環之後，RMSE相較比較穩定，所以預測外科下刀後10分鐘的收縮壓以整合模型(Linear Regression+SVM+GBM)為最佳分析方法。

六、 模型六：外科下刀後10分鐘的舒張壓

放入的解釋變數有半身麻醉完成後之麻醉高度測量、進入手術室的第一次收縮壓、進入手術室的第一次舒張壓、進入手術室的第一次心跳數、進入手術室的第一次血氧指數、外科下刀的收縮壓、外科下刀的舒張壓、外科下刀的心跳數、外科下刀的血氧指數和特徵工程。

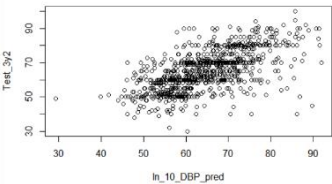
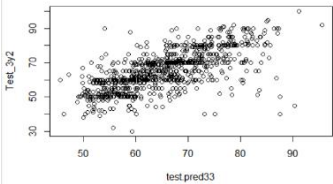
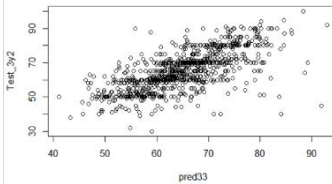
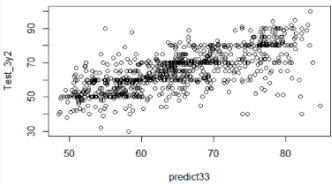
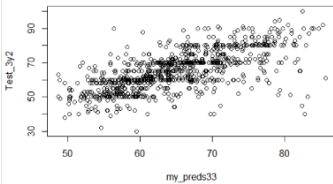
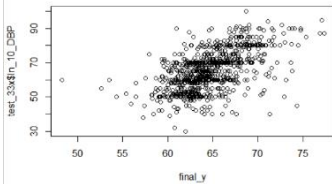
Lasso變數篩選結果：

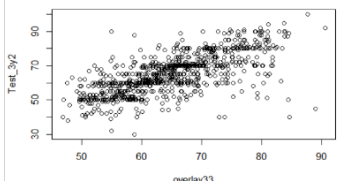
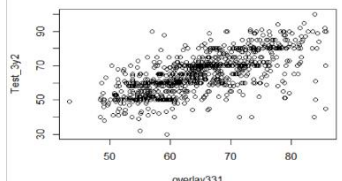
- [1] 女性x年齡取平方 x體重取平方
- [2] 女性x年齡取平方 x身體質量指數取平方
- [3] 病人麻醉風險等級3x術前有心律不整病史x半身麻醉注射，與marcaine混合之fentanyl劑量取平方
- [4] 病人麻醉風險等級3x術前有心律不整病史x半身麻醉注射，與marcaine混合之fentanyl劑量取指數
- [5] 半身麻醉注射針大小為25Gx 女性x身體質量指數取平方
- [6] 半身麻醉注射針大小23G, 25G, 27G聯集x女性x身體質量指數取平方
- [7] 半身麻醉注射針之針面方面為Lateral, Head聯集x不是急診手術x理想體重計算取平方

- [8]半身麻醉完成後之麻醉高度測量
- [9]進入手術室的第一次舒張壓
- [10]外科下刀的收縮壓
- [11]外科下刀的舒張壓

這十一個變數有去做跟反應變數的相關係數，可以發現進入手術室的第一次舒張壓(解釋變數)、外科下刀的收縮壓和外科下刀的舒張壓(解釋變數)與外科下刀後10分鐘的舒張壓(反應變數)相關性較高，分別為0.3885、0.5045 和0.6851，意思為進入手術室的第一次舒張壓(解釋變數)、外科下刀的收縮壓和外科下刀的舒張壓(解釋變數)為影響預測外科下刀後10分鐘的舒張壓(反應變數)的重要因子。

將Lasso篩選出來的變數放入XGBoost、SVM、Linear Regression、GBM、LightGBM、Stacking(Linear Regression、SVM、GBM=>XGBoost)、整合模型(Linear Regression+SVM+GBM)和整合模型(XGBoost+GBM+LightGBM)進行建模與預測，並計算預測值與真實值的差異(均方根誤差RMSE)與計算真實值減預測值絕對值大於10占總比數(測試集)的多少%。

XGBoost	SVM	Linear Regression
1. RMSE : 9.2769 2. 真實值-預測值 >10的個數/測試集數量*100% : 22.35% 3. 真實值與預測值的散佈圖 : 	1. RMSE : 8.4493 2. 真實值-預測值 >10的個數/測試集數量*100% : 18.13% 3. 真實值與預測值的散佈圖 : 	1. RMSE : 8.3785 2. 真實值-預測值 >10的個數/測試集數量*100% : 19.04% 3. 真實值與預測值的散佈圖 : 
GBM	LightGBM	Stacking(Linear Regression、SVM、GBM=>XGBoost)
1. RMSE : 8.2416 2. 真實值-預測值 >10的個數/測試集數量*100% : 18.36% 3. 真實值與預測值的散佈圖 : 	1. RMSE : 8.4563 2. 真實值-預測值 >10的個數/測試集數量*100% : 19.27% 3. 真實值與預測值的散佈圖 : 	1. RMSE : 9.9871 2. 真實值-預測值 >10的個數/測試集數量*100% : 32.84% 3. 真實值與預測值的散佈圖 : 
整合模型	整合模型	補充說明:

(Linear Regression+SVM+GBM)	(XGBoost+GBM+LightGBM)	五個分析(XGBoost、SVM、Linear Regression、GBM和LightGBM)分法預測出來最好的三個分析(Linear Regression、SVM和GBM)去做Stacking跟整合模型來去勘查建模效果是否有比較好，也想看說整合模型(XGBoost、GBM和LightGBM)會是如何?而做了此分析。
1. RMSE : 8.2658 2. 真實值-預測值 >10的個數/測試集數量*100% : 17.9% 3. 真實值與預測值的散佈圖 : 	1. RMSE : 8.434 2. 真實值-預測值 >10的個數/測試集數量*100% : 19.27% 3. 真實值與預測值的散佈圖 : 	

表(2)

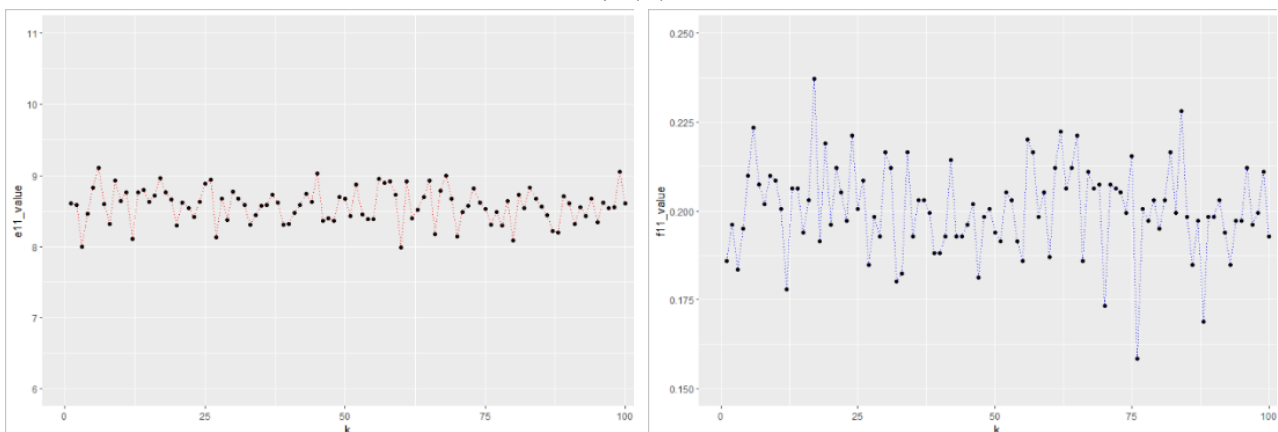
由表(2)的結果可以看出這八種模型以GBM為最佳分析方法，RMSE為8.2416，且真實值減預測值絕對值大於10占總比數(測試集)的18.36%，也能從散佈圖中看出真實值與預測值之間的關係，此分析方法做出的散佈圖，點更傾向45度角延伸，表示預測出來的值跟真實值很接近，但還是有部分的點較為鬆散，X軸與Y軸範圍沒有差很大，所以真實值與預測值的差異比預測外科下刀後10分鐘的收縮壓來的小。

將最佳模型(GBM模型)做交叉驗證，放入自製的for迴圈當中，讓程式反覆執行一百次，並得出一百個結果(RMSE的值)，然後取平均，勘查模型是否穩定並比較沒有使用交叉驗證和有使用交叉驗證預測出來的結果。

GBM

	RMSE	真實值-預測值 >10的個數/測試集數量*100%
外科下刀的10分鐘的舒張壓	8.2416	18.36%
外科下刀的10分鐘的舒張壓(交叉驗證)	8.5774	20.05%

表(3)



圖一、一百次的RMSE結果圖

圖二、一百次的真實值減預測值大於十占比結果圖

從表(3)可以看出沒有做交叉驗證得來的比有做交叉驗證來的好，也可以從圖一圖二得知，做了一百的交叉驗證中有些預測出來的值與真實值相差較大，所以平均之後拉低了RMSE的值與提高了真實值減掉預測值絕對值大於十的占比，並可以從圖一看出利用GBM模型預測一百次的RMSE大致落在8.5左右，只有少數預測出來的值比較突兀，所以利用GBM模型預測是穩定的，所以預測外科下刀後10分鐘的舒張壓以GBM為最佳分析方法。

七、模型七：外科下刀後10分鐘的心跳數

放入的解釋變數有半身麻醉完成後之麻醉高度測量、進入手術室的第一次收縮壓、進入手術室的第一次舒張壓、進入手術室的第一次心跳數、進入手術室的第一次血氧指數、外科下刀的收縮壓、外科下刀的舒張壓、外科下刀的心跳數、外科下刀的血氧指數和特徵工程。

Lasso變數篩選結果：

- [1]非急診手術x術前沒有懷孕 x BMI取根號
- [2]術前沒有懷孕 x此次手術非為痔瘡或肛門瘻管xBMI取根號
- [3]術前沒有懷孕 x半身麻醉使用Heavy藥劑x BMI取根號
- [4]術前沒有懷孕 x半身麻醉使用Heavy藥劑x BMI取log
- [5]進入手術室的第一次心跳數
- [6]外科下刀的心跳數

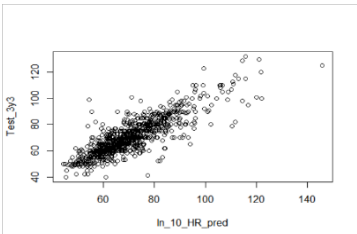
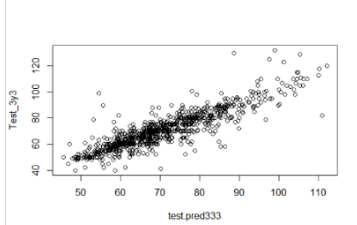
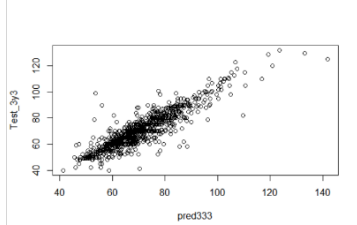
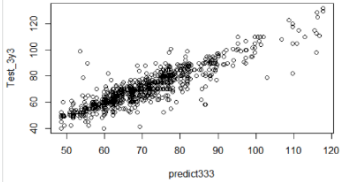
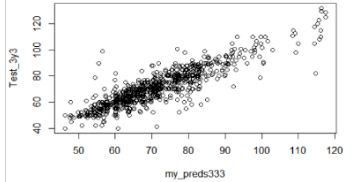
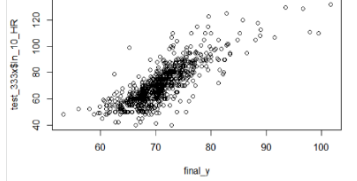
反應變數跟解釋變數的相關性

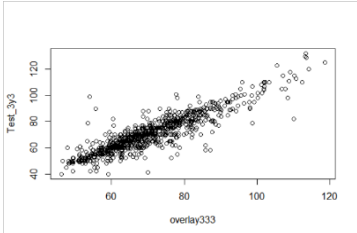
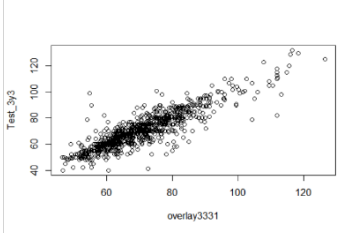
	非急診手術x術前沒有懷孕 x BMI取根號	術前沒有懷孕 x此次手術非為痔瘡或肛門瘻管xBMI取根號	術前沒有懷孕 x半身麻醉使用Heavy藥劑x BMI取根號	術前沒有懷孕 x半身麻醉使用Heavy藥劑x BMI取log	進入手術室的第一次心跳數	外科下刀的心跳數
外科下刀後10分鐘的心跳數	-0.2732	-0.2223	-0.2346	-0.2359	0.7019	0.8844

表(1)

可以從表(1)知道進入手術室的第一次心跳數(解釋變數)和外科下刀的心跳數(解釋變數)與外科下刀後10分鐘的心跳數(反應變數)相關性較高，分別為0.7019和0.8844，意思為進入手術室的第一次心跳數(解釋變數)和外科下刀的心跳數(解釋變數)為影響預測外科下刀後10分鐘的心跳數(反應變數)的重要因子。

將Lasso篩選出來的變數放入XGBoost、SVM、Linear Regression、GBM、LightGBM、Stacking(Linear Regression、SVM、GBM=>XGBoost)、整合模型(Linear Regression+SVM+GBM)和整合模型(XGBoost+GBM+LightGBM)進行建模與預測，並計算預測值與真實值的差異(均方根誤差RMSE)與計算真實值減預測值絕對值大於10占總比數(測試集)的多少%。

XGBoost	SVM	Linear Regression
1. RMSE : 7.7847 2. 真實值-預測值 >10的個數/測試集數量*100% : 15.51% 3. 真實值與預測值的散佈圖 : 	1. RMSE : 6.7422 2. 真實值-預測值 >10的個數/測試集數量*100% : 9.24% 3. 真實值與預測值的散佈圖 : 	1. RMSE : 6.3418 2. 真實值-預測值 >10的個數/測試集數量*100% : 8.67% 3. 真實值與預測值的散佈圖 : 
GBM	LightGBM	Stacking(Linear Regression、SVM、GBM=> XGBoost)
1. RMSE : 6.4499 2. 真實值-預測值 >10的個數/測試集數量*100% : 9.12% 3. 真實值與預測值的散佈圖 : 	1. RMSE : 6.5418 2. 真實值-預測值 >10的個數/測試集數量*100% : 9.81% 3. 真實值與預測值的散佈圖 : 	1. RMSE : 10.5404 2. 真實值-預測值 >10的個數/測試集數量*100% : 29.99% 3. 真實值與預測值的散佈圖 : 
整合模型 (Linear Regression+ SVM+GBM)	整合模型 (XGBoost+GBM+LightGBM)	補充說明: 五個分析(XGBoost、SVM、Linear Regression、GBM和LightGBM)

1. RMSE：6.3749 2. 真實值-預測值 >10的個數/測試集數量*100%：8.55% 3. 真實值與預測值的散佈圖： 	1. RMSE：6.6317 2. 真實值-預測值 >10的個數/測試集數量*100%：10.49% 3. 真實值與預測值的散佈圖： 	htGBM)分法預測出來最好的三個分析(Linear Regression、SVM和GBM)去做Stacking跟整合模型來去勘查建模效果是否有比較好，也想看說整合模型(XGBoost、GBM和Lighgbm)會是如何?而做了此分析。
---	--	---

表(2)

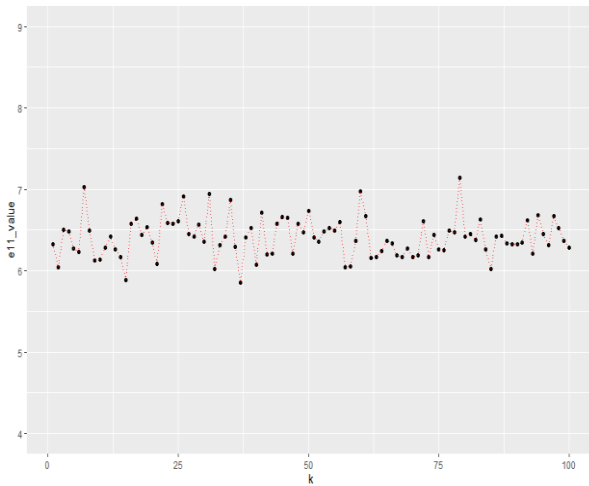
其實這八種模型預測出來的結果都差不多，但重點是真實值減預測值絕對值差異的值會影響到RMSE，由表(2)的結果可以看出這八種模型以Linear Regression分析為最佳分析方法，RMSE為6.3418，且真實值減預測值絕對值大於10占總比數(測試集)的8.67%，也能從散佈圖中看出真實值與預測值之間的關係，此模型預測的效果只有些許的值與真實值有落差，所以點幾乎都是45度角延伸。

將最佳模型(Linear Regression)做交叉驗證，放入自製的for迴圈當中，讓程式反覆執行一百次，並得出一百個結果(RMSE的值)，然後取平均，勘查模型是否穩定並比較沒有使用交叉驗證和有使用交叉驗證預測出來的結果。

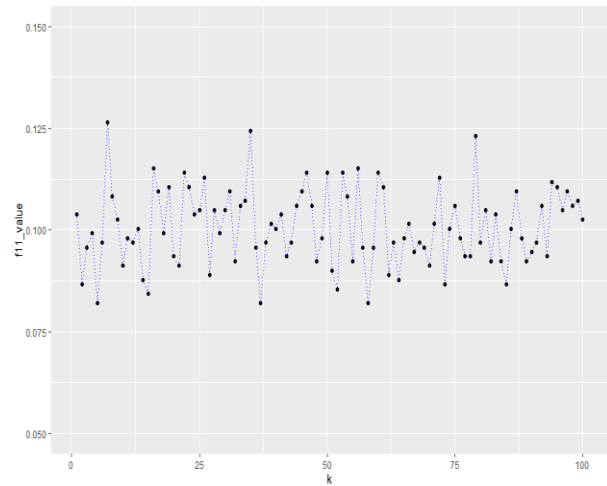
Linear Regression

	RMSE	真實值-預測值 >10的個數/測試集數量*100%
外科下刀的10分鐘的心跳數	6.3418	8.67%
外科下刀的10分鐘的心跳數(交叉驗證)	6.4079	9.96%

表(3)



圖一、一百次的RMSE結果圖



圖二、一百次的真實值減預測值大於十占比結果圖

從表(3)可以看出沒有做交叉驗證得來的比有做交叉驗證來的好，也可以從圖一圖二得知，做了一百的交叉驗證中有些預測出來的值與真實值相差較大，所以平均之後拉低了RMSE的值與提高了真實值減掉預測值大於十的占比，並可以從圖一看出利用Linear Regression建立模型去預測是穩定的，RMSE大致落在6到6.7之間，所以預測外科下刀後10分鐘的心跳數以Linear Regression為最佳分析分法。

八、 模型八：預測半身麻醉後是否會使用到升壓劑

特徵篩選

特徵工程結束後，我們生成20幾萬筆的變數，將這麼龐大的數據量丟入模型中，會消耗大量的電腦效能，而且也容易造成過度配飾的問題，因此需要透過特徵篩選，只留下對反應變數(dependent variable)造成顯著影響的自變數(independent variable)進行建模。

透過Lasso Regression將20幾萬個變數進行篩選，總共篩選出了33個變數，其中又以「手術科別」、「身體質量指數」、「是否有懷孕」、「半身麻醉完成後之麻醉高度測量」、「半身麻醉注射針之針面方面」、「血壓」，這6個因子對此模型有一定的影響力。

資料平衡

透過隨機抽樣將原始資料的80%作為訓練集，其餘的20%作為測試集，而沒有使用升壓劑的人數比有使用升壓劑的人數多了2.5倍，因為兩類的人數差距沒有到4倍以上，正常情況下會直接進行建模，但是模型預測結果都不是很好，因此使用R語言的SMOTE函數將數量較多的類別採用Under-Sampling的概念降低筆數，讓每一類的人數差距減少，避免模型在訓練或預測時數量多的類別容易被預測到，導致模型敏感度降低。

醫療行為上會使用到升壓劑，代表術中病人有危急情況，因此我們希望有使用升壓劑的準確度能達到9成以上，所以主要觀察敏感度的結果來判斷模型好壞，資料平衡後分別使用7種演算法(XGB、SVM、Rule Based、KNN、Naive Bayes、Logistic Regression、Catboost)進行建模，我們從7個模型當中選出3個敏感度結果最好的顯示如下：

SVM

預測結果：

	實際 No	實際 Yes
預測 No	432 (49.26%)	83 (9.46%)
預測 Yes	182 (20.75%)	180 (20.52%)

敏感度＝實際有使用升壓劑且預測為有使用升壓劑的人數／實際有使用升壓劑的總人數

特異度＝實際沒有使用升壓劑且預測為沒有使用升壓劑的人數／實際沒有使用升壓劑的總人數

$$\text{敏感度} = 180 / (180 + 83) = 68.44\%$$

$$\text{特異度} = 432 / (432 + 182) = 70.36\%$$

(B)Rule Based

預測結果：

	實際 No	實際 Yes
預測 No	424 (48.35%)	90 (10.26%)
預測 Yes	190 (21.66%)	173 (19.73%)

$$\text{敏感度} = 173 / (90 + 173) = 65.78\%$$

$$\text{特異度} = 424 / (424 + 190) = 69.06\%$$

(C) Logistic Regression

預測結果：

	實際 No	實際 Yes
預測 No	422 (48.12%)	70 (7.98%)
預測 Yes	192 (21.89%)	193 (22.01%)

$$\text{敏感度} = 193 / (70 + 193) = 73.38\%$$

$$\text{特異度} = 422 / (422 + 192) = 68.73\%$$

交叉驗證

為了確認模型不是受到極端數據的影響，造成預測結果的敏感度不佳，因此我們重複100次的隨機抽樣進行模型訓練，將100次的敏感度結果取平均與原先模型敏感度相比差異有多少。

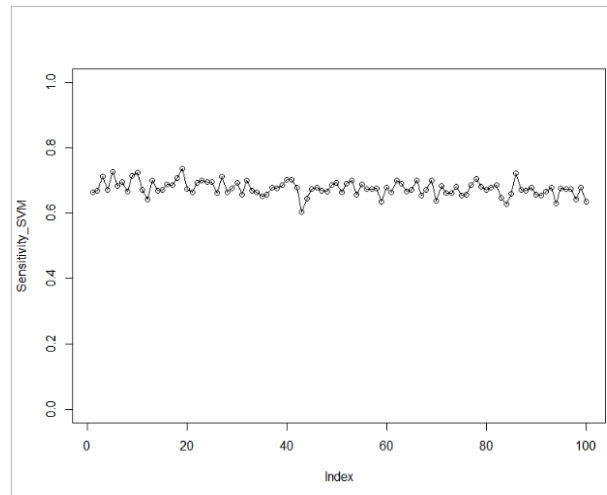
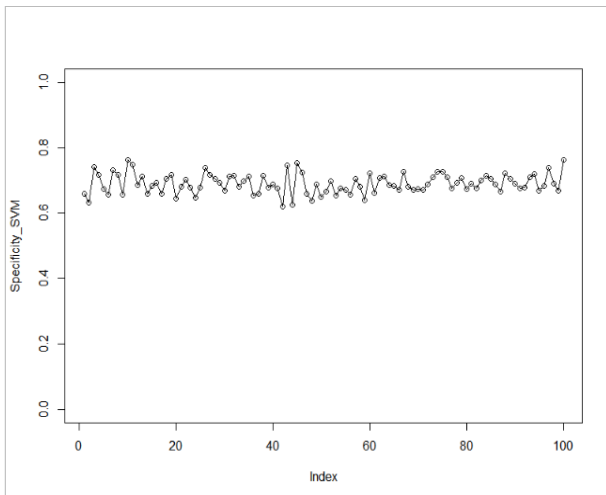
(A)SVM

預測結果(原始)：

敏感度：68.44%，特異度：70.36%

預測結果(交叉驗證後)：

敏感度：69.03%，特異度：67.60%



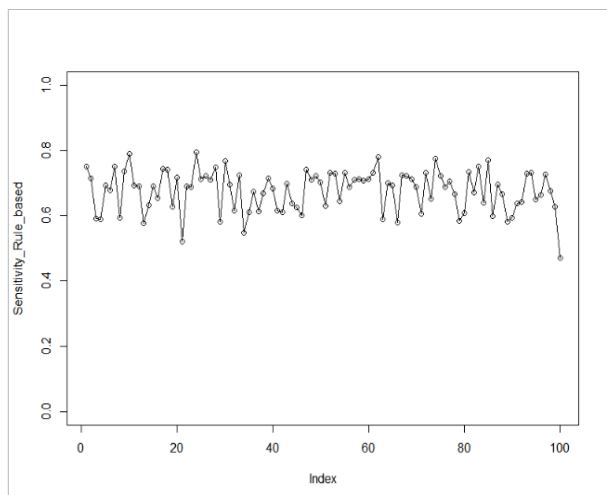
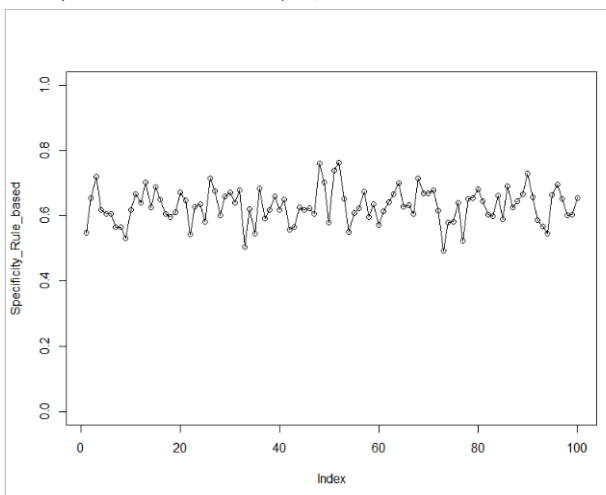
(B)Rule_Based

預測結果(原始)：

敏感度：65.78%，特異度：69.06%

預測結果(交叉驗證後)：

敏感度：63.60%，特異度：67.72%



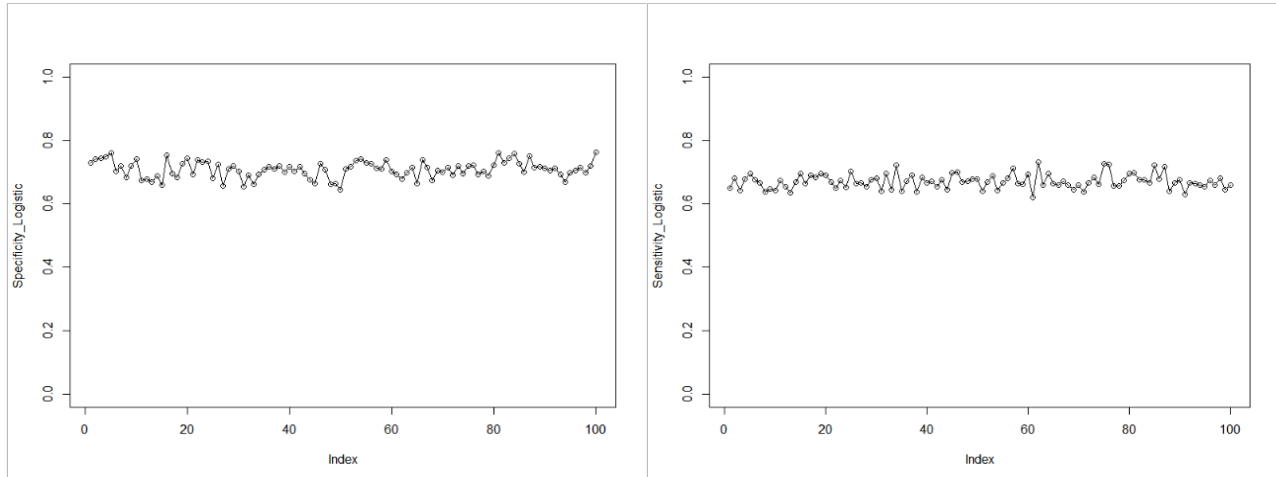
(C) Logistic Regression

預測結果(原始)：

敏感度：73.38%，特異度：68.73%

預測結果(交叉驗證後)：

敏感度：70.84%，特異度：67.04%



圖中可以看出每一次的敏感度以及特異度結果，其中又以Logistic Regression的敏感度結果最為穩定，但是敏感度平均落在7成左右，還沒有辦法實際應用到醫療行為上。

九、 模型九：預測半身麻醉後是否會使用到止痛藥

特徵篩選

我們透過Lasso Regression進行特徵篩選，總共篩選出了5個變數，結果如下：

- [1]手術科別為婦產科
- [2]病人麻醉風險為1與術前有懷孕與身體質量指數之交互作用
- [3]半身麻醉注射部位為L4-5或L5-S1與術前有懷孕與年齡之交互作用
- [4]半身麻醉注射部位為L4-5或L5-S1與術前有懷孕與年齡的平方之交互作用
- [5]半身麻醉注射部位為L2-3與術前有懷孕與理想體重的平方之交互作用

資料平衡

因為有使用止痛藥的人數只有55位，為了提高模型的穩定性，透過隨機抽樣將原始資料的40%作為測試集，藉此收集更多有使用止痛藥的個案，之後使用R語言的SMOTE函數將數量少的類別採用Over-Sampling的概念放大，數量較多的則採用Under-Sampling的概念減少，讓

每一類的人數差距減少達到資料平衡的效果，避免模型訓練或預測時數量多的類別容易被預測到。

醫療行為上會使用到止痛藥，代表病人麻醉深度不足在手術過程中甦醒過來，因此我們希望有使用止痛藥的準確度能達到9成以上，所以主要觀察敏感度的結果來判斷模型好壞，將資料平衡後的訓練集分別使用7種演算法(XGB、SVM、Rule Based、KNN、Naive Bayes、Logistic Regression、Catboost)進行建模，我們從7個模型當中選出3個預測結果最好的顯示如下：

(A)XGBoost

預測結果：

	實際 No	實際 Yes
預測 No	1637 (93.33%)	7 (0.40%)
預測 Yes	97 (5.53%)	13 (0.74%)

敏感度＝實際有使用止痛藥且預測為有使用止痛藥的人數／實際有使用止痛藥的總人數

特異度＝實際沒有使用止痛藥且預測為沒有使用止痛藥的人數／實際沒有使用止痛藥的總人數

敏感度＝ $13 / (7 + 13) = 65.00\%$

特異度＝ $193 / (70 + 193) = 94.41\%$

(B)SVM

預測結果：

	實際 No	實際 Yes
預測 No	1637 (93.33%)	5 (0.29%)
預測 Yes	97 (5.53%)	15 (0.86%)

敏感度＝ $193 / (70 + 193) = 75.00\%$

特異度＝ $193 / (70 + 193) = 94.41\%$

(C) Logistic Regression

預測結果：

	實際 No	實際 Yes
預測 No	1637 (93.33%)	6 (0.29%)
預測 Yes	97 (5.53%)	14 (0.80%)

敏感度＝ $193 / (70 + 193) = 70.00\%$

特異度 = $193 / (70 + 193) = 94.41\%$

交叉驗證

為了確認模型不是受到極端數據的影響，造成預測結果的敏感度不佳，因此我們重複100次的隨機抽樣進行模型訓練，將100次的敏感度結果取平均與原先模型敏感度相比差異有多少。

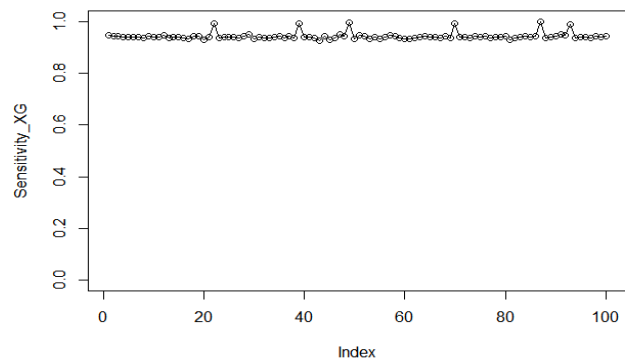
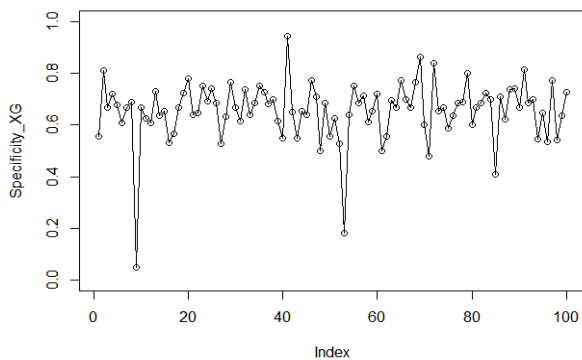
(A) XGBoost

預測結果(原始)：

敏感度：65.00%，特異度：94.41%

預測結果(交叉驗證後)：

敏感度：65.51%，特異度：94.25%



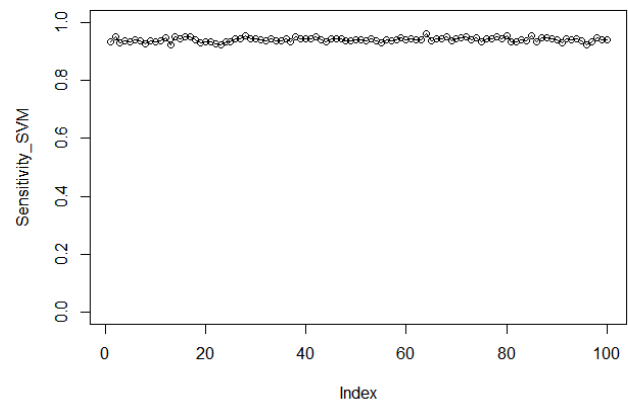
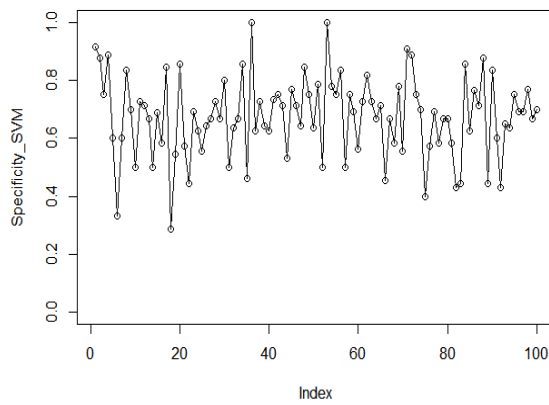
(B) SVM

預測結果(原始)：

敏感度：75.00%，特異度：94.41%

預測結果(交叉驗證後)：

敏感度：67.46%，特異度：93.98%



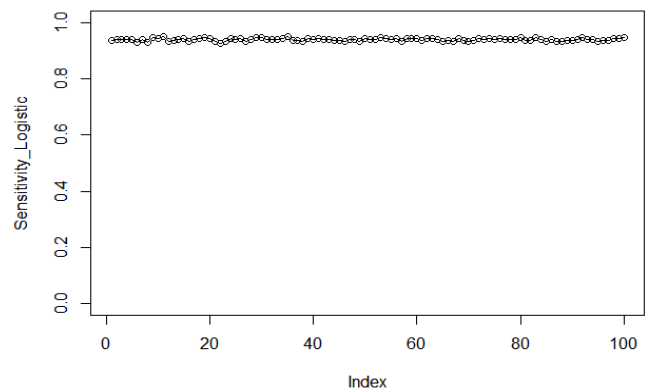
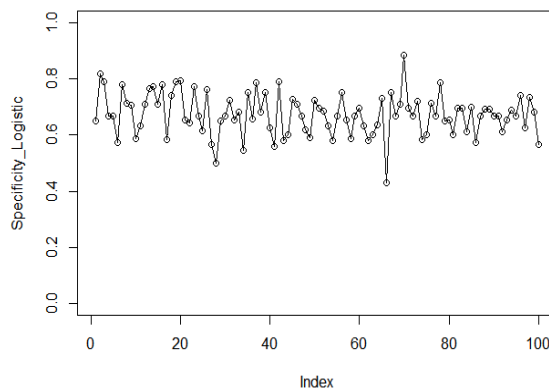
(C) Logistic Regression

預測結果(原始)：

敏感度：70.00%，特異度：94.41%

預測結果(交叉驗證後)：

敏感度：67.30%，特異度：93.91%



圖中可以看出每一次的敏感度以及特異度結果，其中又以Logistic Regression的敏感度結果最為穩定，但是敏感度平均落在7成左右，且每次的敏感度差異過大，沒有辦法實際應用到醫療行為上。

第五章 結論與討論

使用麻醉藥的劑量或施打的部位，都會對手術及術後結果有一定的影響，所以我們想研究透過麻醉施行過程資料，來去預測會對病患造成怎樣的後果。我們用醫院病患基本的生理資料與半身麻醉施行過程資料，預測9個主題：

壹、 預測麻醉深度

以「半身麻醉完成後之麻醉高度測量」當作應變數，並使用選出的27個特徵當作自變數。在多個模型中，以Logistic Regression模型準確率最高，XGB差一個類別比率最小。

貳、 預測完成半身麻醉10分鐘的收縮壓

使用選出的14個特徵當作自變數，一開始模型預測結果光看RMSE非常不好，雖利用Stacking試圖降低，也透過交叉驗證將RMSE進行平均，但還是降不太下來，所以此訓練的模型都不太推薦。

參、 預測完成半身麻醉10分鐘的舒張壓

使用選出的11個特徵當作自變數，我們推薦使用Regression、SVM、GBM三種模型，其中又以LightGBM的RMSE結果是最低的，為11.68829。

肆、 預測完成半身麻醉10分鐘的心跳數

以篩出的「進入手術室的第一次心跳數」為自變數，一樣推薦使用Regression、SVM、GBM三種模型，其中又以LightGBM的RMSE結果是最低的，為8.827359。

伍、 預測外科下刀後10分鐘的收縮壓

使用選出的3個特徵當作自變數，預測外科下刀後10分鐘的收縮壓以整合模型(Linear Regression+SVM+GBM)為最佳分析方法，RMSE為13.3121。

陸、 預測外科下刀後10分鐘的舒張壓

使用選出的11個特徵當作自變數，預測外科下刀後10分鐘的舒張壓以GBM為最佳分析方法，rmse為8.24。

柒、 預測外科下刀後10分鐘的心跳數

使用選出的6個特徵當作自變數，預測外科下刀後10分鐘的心跳數以Linear Regression為最佳分析方法，RMSE為6.3418。

捌、 預測半身麻醉後是否使用到升壓劑

使用選出的6個特徵當作自變數，其中以SVM、Rule Based跟Logistic Regression結果最好，進行交叉驗證後以Logistic Regression的敏感度結果最為穩定，敏感度平均落在7成左右。

玖、 預測半身麻醉後是否使用到止痛藥

使用選出的5個特徵當作自變數，其中以XGBoost、SVM跟Logistic Regression結果最好，進行交叉驗證後以SVM的敏感度結果較好，敏感度平均落在7成左右，但是每次的敏感度結果差異大。

我們在建立模型的過程中，發現有許多因素會直接或間接地影響預測結果，進而導致模型結果不佳。像是資料中的極端值、特徵工程中的交互作用以及模型中的參數等等。其中遇到的困難為變數相關性低，導致篩選不到重要的變數，以及可能缺少個案的實驗室檢驗資料，例如：血液、尿液、腦波、生化數據和藥物使用，與個人飲食習慣，如：過量飲酒、咖啡、茶類等。

參考文獻

一、英文文獻：

1. Greene NM: Distribution of local anesthetic solutions within the subarachnoid space. *Anesth Analg* 64:715, 1985.
2. Stienstra R, Greene NM. Factors affecting the subarachnoid spread of local anesthetic solutions. *Reg Anesth* 1991;16:1-6.
3. Stienstra R, Veering BT: Intrathecal drug spread: Is it controllable? *Reg Anesth Pain Med* 23:347, 1998
4. Cameron AE, Arnold RW, Ghoris MW, et al: Spinal analgesia using bupivacaine 0.5% plain: Variation in the extent of the block with patient age. *Anaesthesia* 36:318, 1981.
5. Pitkanen M, Haapaniemi L, Tuominen M, et al: Influence of age on spinal anesthesia with isobaric 0.5% bupivacaine. *Br J Anaesth* 56:279, 1984.
6. McClure JH, Brown DT, Wildsmith JAW: Effect of injected volume and speed of injection on the spread of spinal anesthesia with isobaric amethocaine. *Br J Anaesth* 54:917, 1982.
7. Holman SJ, Robinson RA, Beardsley D, et al: Hyperbaric dye solution distribution characteristics after pencil-point needle injection in a spinal cord model. *Anesthesiology* 86:966, 1997.
8. Urmei WF, Stanton J, Bassin P, et al: The direction of the Whitacre needle aperture affects the extent and duration of isobaric spinal anesthesia. *Anesth Analg* 84:337, 1997.
9. Qing-he Zhou, Bo Zhu, Chang-Na Wei, et al: abdominal girth and vertebral column length can adjust spinal anesthesia for lower limb surgery, a prospective, observational study. *BMC Anesthesiology* (2016) 16:22
10. Chi-Hang Kuok, Chung-Hsin Huang, Pei-Shan Tsai, et al: Preoperative measurement of maternal abdominal circumference relates the initial sensory block level

vel of spinal anesthesia for cesarean section: An observational study. Taiwanese Journal of Obstetrics & Gynecology 55 (2016) 810-814

11. Alex Tiong Heng Sia, Kok Hian Tan, Ban Leong Sng, et al: Hyperbaric Versus Plain Bupivacaine for Spinal Anesthesia for Cesarean Delivery. Anesth Analg 2015;120:132-40

二、 中文文獻：

1. 黃惠鈴(2014-04-01)：〈為什麼台灣麻醉致死率高？〉《天下雜誌》，544期。
2. 林喜音(2018-05-30)：〈全身麻醉危險？麻醉醫師：其實真正的危險在於…〉《元氣網》。
3. 《麻醉科-光田醫療社團法人光田綜合醫院》(2021/8/17)：〈全身麻醉與半身麻醉那一個風險高？麻醉風險如何計算？〉。
4. 《程式人生》(2018-12-31)：〈KNN演算法(有監督學習演算法)〉。
5. 《光田醫療社團法人光田綜合醫院》：〈關於全身麻醉〉。
6. 《維基百科》：〈LightGBM〉。
7. 《iT邦幫忙》：〈[Day 17] 輕量化的梯度提升機 - LightGBM〉
8. 《今天頭條》(2020-03-08)：〈「ML」深入理解CatBoost〉。
9. Blessy_Zhu (2019-06-11)：〈機器學習（十二）-基於規則的分類器〉《台部落》。
10. 《維基百科》：〈基於規則的系統〉。
11. 江奕(2016/02)：〈邏輯斯迴歸分析〉《eNews》，第11期。
12. 《百科知識》：〈線性迴歸〉。