

智慧型血糖監測與提醒行動應用程式

二、研究計畫內容(以 10 頁為限)：

(一) 摘要

隨著人口老化，國人慢性病的罹病族群日益遽增，目前國人十大死因中，排名第四名的糖尿病有逐年緩步增加的趨勢。糖尿病及其併發症僅大大影響了國人健康，其所衍生的全民健保醫療負擔亦相當龐大。而近年來機器學習技術蓬勃的發展，也已廣泛的應用於各種醫療資訊分析的議題上。而有鑑於市面上常見的血糖機 APP 大多只有觀察、紀錄之功能，較無針對患者進行各別的資料分析與建議的功能，以及與家屬或醫療院所分享病患血糖控制的現況與趨勢等訊息。；因此本計畫擬開發可供糖尿病患進行血糖管控與自主健康管理的「智慧型血糖監測與提醒行動應用程式」系統。系統中將利用建構個人化的血糖值分析與建議模式，讓使用者能更精準的發現身體狀況的異常，並且將結果以視覺化紅綠燈界面方便使用者用更淺顯易懂的方式得知身體狀況，除了本身狀況也能透過系統通知相關人或醫療院所，讓高風險的使用者能夠盡早向醫療機構尋求協助及治療。

(二) 研究動機與問題

依據國際糖尿病聯盟(International Diabetes Federation, IDF) 2021 年數據顯示，20-79 歲成年人中，全球約有 5.36 億糖尿病者，相較於 2020 年 4.63 億糖尿病者人數增加了約 15%，全球糖尿病盛行率已達 10.5%，平均每 10 人就有 1 人患有糖尿病，國際糖尿病聯盟預估 2045 年將有 7 億人口患有糖尿病(International Diabetes Federation, 2021)。在臺灣，衛生福利部國民健康署¹統計 2017-2020 年 65 歲以上高血糖盛行率達 27.84%，約有 300 萬人為高血糖族群。且依據衛生福利部國民健康署 2013-2016 年「國民營養健康狀況變遷調查」也發現²，18 歲以上國人糖尿病盛行率為 11.8% (男性 13.1%；女性 10.5%)，且每年以 2.5 萬名的速度持續增加。

血糖異常是許多疾病的前兆，單純只有糖尿病，以現在醫學並不難控制，最令人擔憂的是會引起多種併發症。泡在高血糖裡的視網膜細胞會壞死、血管變硬、腎功能受損，日積月累後，眼盲、中風、心臟病、腎衰竭等症狀都會產生(康健雜誌 223 期，陳俊辰)。因此糖尿病及其併發症(包含失明、截肢或洗腎等)，不僅大

¹ 衛福衛生福利部國民健康署: <https://www.hpa.gov.tw/Pages/Detail.aspx?nodeid=641&pid=1231>

² 衛福衛生福利部國民健康署: <https://www.hpa.gov.tw/Pages/Detail.aspx?nodeid=3999&pid=11145>

大影響了國人健康，其所衍生的全民健保醫療負擔亦相當龐大

近年來有越來越多的不同型號的血糖機或血糖數值管理系統陸續上市，但只有少數的血糖機有藍芽傳輸或雲端血糖管理系統(如APP或資訊網頁)，以至於多數糖尿病患者用了不同型號或廠牌的血糖機，仍保持著採用醫院提供的血糖登記手冊，以紙本方式記錄每日量測的血糖值，這對糖尿病患者造成莫大的困擾，尤其是較嚴重且每日需要量測3-6次的第2型糖尿病患者，很容易有漏抄、抄錯或遺失相關血糖紀錄手冊等風險。此外，目前市面上常見的血糖機APP主要只有觀察、紀錄之功能，較無針對患者的資料進行各別的資料分析與建議的功能，以及與家屬或相關醫療院所分享病患血糖控制的現況與趨勢訊息。

機器學習技術已被廣泛用於檢測人體狀況(Abbas et al., 2018)，分析疾病的相關因素(Mahyoub, et al., 2018)及診斷高血壓、糖尿病等各種疾病的研究中。通過機器學習的技術可以發現目標變數和特徵變數之間的隱藏關係，進而建構模式來利用特徵變數進行分類或預測的目的。基本上，性別、年齡及身體質量指數(Body Mass Index；BMI)是與糖尿病控制有關的重要人口統計風險因子。糖尿病患者在不同性別、年齡及 BMI 等風險因子與患病時間及量測時間的組合下，會有不同的異常標準³，若能以相關的資訊使用機器學習進行相關的分析與建模，將能夠提供個別化的糖尿病控制的相關資訊。

本計畫的目標是開發可供糖尿病患者進行血糖管控與自主健康管理的「智慧型血糖監測與提醒行動應用程式」系統。可以提供給糖尿病患者，當量測完血糖數值後，直接用透過血糖機的藍芽傳送數據到所開發的 APP，包含可註記量測的時間與日期，提供給每個已登入帳號的糖尿病患者或家屬，可以讀取相關資料，並將系統自動生成的血糖數值連續性資料的統計圖表(含其他數據的健康儀表板)，直接呈現在 APP 上。並且透過機器學習技術，建構個人化的血糖值分析與建議模式。而為了與 APP 內的架構以及圖形化的顯示配合，本計畫將以三個廣泛使用的決策樹技術，包含 Classification and Regression Trees (CART)、C4.5、C5.0 三種演算法，來進行分析模式的建立。

(三) 文獻回顧與探討

3.1 糖尿病定義

糖尿病 (Diabetes mellitus) 主要原因是血糖長期在於高於標準值上，但主要

³ <https://blog.health2sync.com/individual-diabetes-glucose-management-goals/>

的因素是因為胰臟分泌胰島素不足，或者胰島素受到阻礙，而葡萄糖無法進入細胞，所引起的高血糖，如不進行診斷或治療，會引發許多的併發症，嚴重的話會引起心血管疾病、中風、慢性腎臟疾病、糖尿病足和視網膜病變等之併發症。(林雅鈴，2016)

3.2 糖尿病控制的重要風險因子

與糖尿病控制有關的風險因子眾多，下面介紹性別、年齡與 BMI 等三個比較重要且易於與 APP 結合的人口統計風險因子

(一) 性別:

在性別方面，大多的研究顯示男性患有糖尿病之比例高於女性患有糖尿病。美國疾病管制與預防中心（CDC）指出，在美國，男性在 2019 年患有糖尿病人口中有 15.4%，而女性在 2015 年患有糖尿病人口中有 14.1%。在美國人口中，年齡組群之糖尿病患病率為（45~64、65~74 及 75 歲以上）男性比女性有高的糖尿病患病率，而且男性糖尿病患病率有上升的趨勢（Centers for Disease Control and Prevention，2020）。除了大多研究表示，糖尿病患病率大多是男性高於女性，但在土耳其地區發現女性患有糖尿病比率是高於為男性，糖尿病患病率為 11.4%，男性糖尿病患病率為 9.7%，女性糖尿病患病率為 12.7%。在土耳其之所以女性糖尿病患病率會高於男性，是因為在中年和老年女性的肥胖率遠高於男性（Ilhan et al.，2013）。Esayas et al.（2013）學者指出，在南非洲的許多國家中，女性糖尿病患病率比例會高於男性，可能原因是女性肥胖率和體重都高於為男性。因此性別與相關的因子的結合會是討論糖尿病風險的重要考量。

(二) 年齡:

在年齡方面，研究顯示成年人患有糖尿病之比例低於老年人患有糖尿病，但糖尿病的比率有年輕化的趨勢。目前許多研究證實，不管糖尿病前期或是第 2 型糖尿病的發病率之年齡在於 65 歲以上，且有更嚴重的併發症風險（Selvin et al.，2006）。美國疾病管制與預防中心（CDC）指出，在美國，18~44 歲患有糖尿病患病率約為 4.8%，45~64 歲患有糖尿病患病率約為 18.9%，大於 65 歲患有糖尿病患病率約 29.2。年齡越年長，患有糖尿病機會也越高，這大多研究已證實年齡越大，患有糖尿病機率也越大，因為年齡越年長是患有糖尿病的重要決定因素之一，血糖濃度上升是隨著年齡增長的（Choi et al.，2001）。

(三) 身體質量指數（BMI）

肥胖患病率在世界各地快速的增加（James，2008），第 2 型糖尿病和肥胖

對於長期的健康有很重要的危險因素之風險 (Holly et al, 2010)，依據世界衛生組織 (2003) 表示，在 2014 年有超過三分之一的成年人為超重，而有超過十分之一的成年人為肥胖，超重和肥胖或久坐不動的人是主要增加第 2 型糖尿病的風險之一 (Abdullah et al, 2010)，但也有可能是因為遺傳因素，遺傳也會造成糖尿病之風險。在美國人口中，有肥胖者和 BMI 超過正常值的人約有三分之二患有糖尿病之風險，肥胖的糖尿病患病率為 54% (James, 2008)。

3.3 機器學習在糖尿病之應用

現今，機器學習形式的 AI 已大量運用在各行業，醫療行業也存於此浪潮當中，而血糖疾病患病人數隨著逐年成長，已有許多文獻進行廣泛的研究。例如 (Shon et al, 2019) 研究中使用機器學習和類神經網路技術，對糖尿病數據進行預測；(Alic et al, 2017) 研究中使用貝葉斯網路對糖尿病及心血管疾病進行分類；(Divya et al, 2019) 資料說明如何透過類神經網路和模糊神經網路作為混合模型來預測是否為糖尿病患者，並針對模型之準確性，敏感性和特異性進行了評估，也都獲得良好的預測結果。

(四) 研究方法與步驟

4.1 研究方法

本計畫擬採用決策樹產生規則式預測模式，方便與 APP 結合。決策樹演算法是屬於監督式學習法的一種，藉由分類已知的事物來建立樹狀式的結構，以從中歸納並萃取規則，並進行未知樣本的預測。決策樹的層級架構，可以分析不同層級的變因對目標變數之影響，因此隨著不同資料，採用不同的演算法，得到的樹狀結構自然不同。決策樹產生的基本演算法是利用貪婪演算法 (greedy algorithm)，它是一種由上而下 (top-down) 的方法，用遞迴 (recursive) 和各個擊破 (divide-and-conquer) 來建立樹狀結構。由於決策樹演算法無法處理連續性的數值。因此數值屬性必須先轉換成類別型態，方能進行決策樹分析。

使用決策樹進行資料分析的過程包含三大步驟：(1) 資料來源與處理；(2) 資料彙整與轉換以及 (3) 建立決策樹。其中，(1) 資料來源與處理及 (2) 資料彙整與轉換都稱為資料前置處理。資料前置處理完畢之後，接下來便開始進入決策樹建構過程。將依據每次所選到的屬性，反覆地將樣本分隔開來，直到同一個分支子集合內的資料都屬於同一個類別，或是沒有屬性可分為止。目前常見的決策樹演算法

包括：分類與迴歸樹(classification and regression trees, CART)、卡方自動交互檢測(Chi-squared automatic interaction detection, CHAID)、C4.5/C5.0 等一系列方法。

CART 演算法的分支變數的準則是 Gini 係數，在每個分枝節點進行資料分割，並建立一個二分式的決策樹，其計算出每種門檻值下的 Gini 值，並選取獲得以最小的 Gini 值的門檻值將資料分成兩個子集合。CART 演算法在每一個可能的分枝變數中會選擇具有最大化純度的分枝水準作為候選分枝依據，再經由比較所有候選分枝變數中具有最大純度作為節點的分枝。當利用訓練資料完成決策樹的建構，CART 會利用成本複雜性的修剪方法，修剪其不必要的分枝。

C4.5/C5.0 決策樹演算法是由其前身 ID3 學習法改進而來。ID3 為一決策樹歸納技術，在構建決策樹過程中，ID3 以資訊獲利(Information Gain)為基準，選擇最佳的屬性當成決策樹的節點，使得所導致的決策樹為一最簡單(或接近最簡單)的決策樹。C4.5 演算法以資訊增益比作為決定分枝變數的準則，且為多元分枝決策樹，C4.5 演算法最常用於處理類別型資料，若遇連續型資料則需事先將其轉化成類別變數。相較於其他分類演算法的預測準確性、複雜度和訓練時間，C4.5 決策樹演算法提供了較佳的準確性及資料解釋能力。由於遺漏值會在建立決策樹的過程中被忽略與取代，因此遺漏值不影響資訊增益比的計算。C5.0 決策樹演算法是 C4.5 決策樹演算法的進階版，C5.0 決策樹演算法增加了交互驗證(cross-validation)與訓練資料重覆抽測的機制(boosting)，與 C4.5 決策樹演算法相比，不僅決策樹的結果更準、計算速度更快、且需佔用的記憶體資源也較少。(鄒佳玲,2016)。

4.2 研究步驟

本計畫的 APP 的系統進行流程圖，如圖 1 所示。整個建構流程可以分為前置作業與上線使用與測試兩個部分。在前置作業，包含有 APP 開發、血糖判斷模式的訓練與建立，以及資料庫建置等三個。其中血糖判斷模式的建構，會與合作的醫師，先以現有的糖尿病患者的紀錄資料，在去識別化與通過 IRB 的情況下，使用醫師建議的合適風險因子，利用決策樹技術來生成樹的規則模式，並且經與醫師討論後做為本計畫採用之模式。

在上線使用部分，則將根據有無量測需求以及有無授權者，來進行 APP 使用流程的規劃、畫面設計以及後續的修改。

而根據流程，本計畫預計發展的 APP 可以分成三部分—血糖分析、個人

化通知、及歷史資料查詢，主要的功能如圖 2 所示。以下分別說明之。

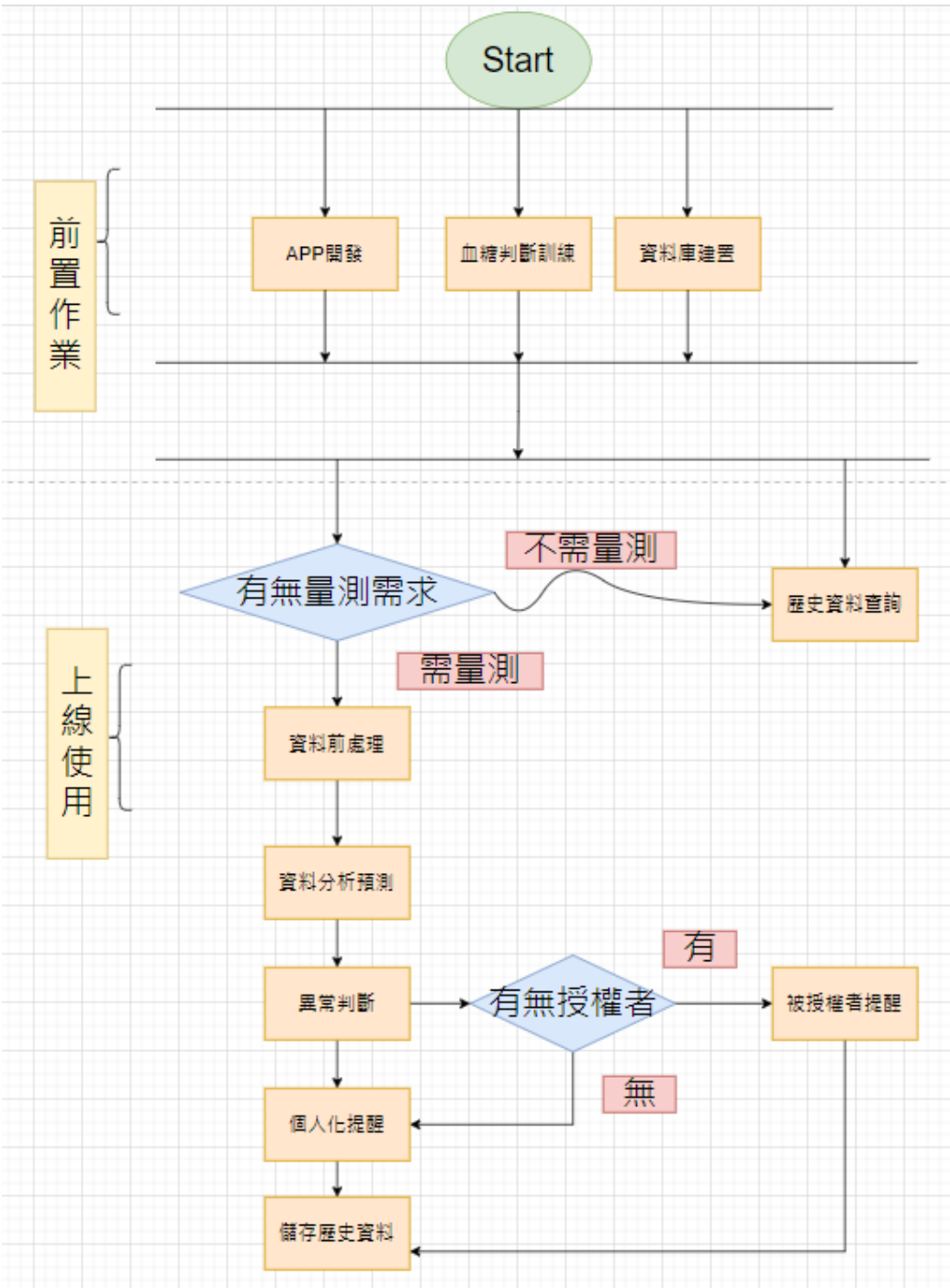


圖 1、系統進行流程圖

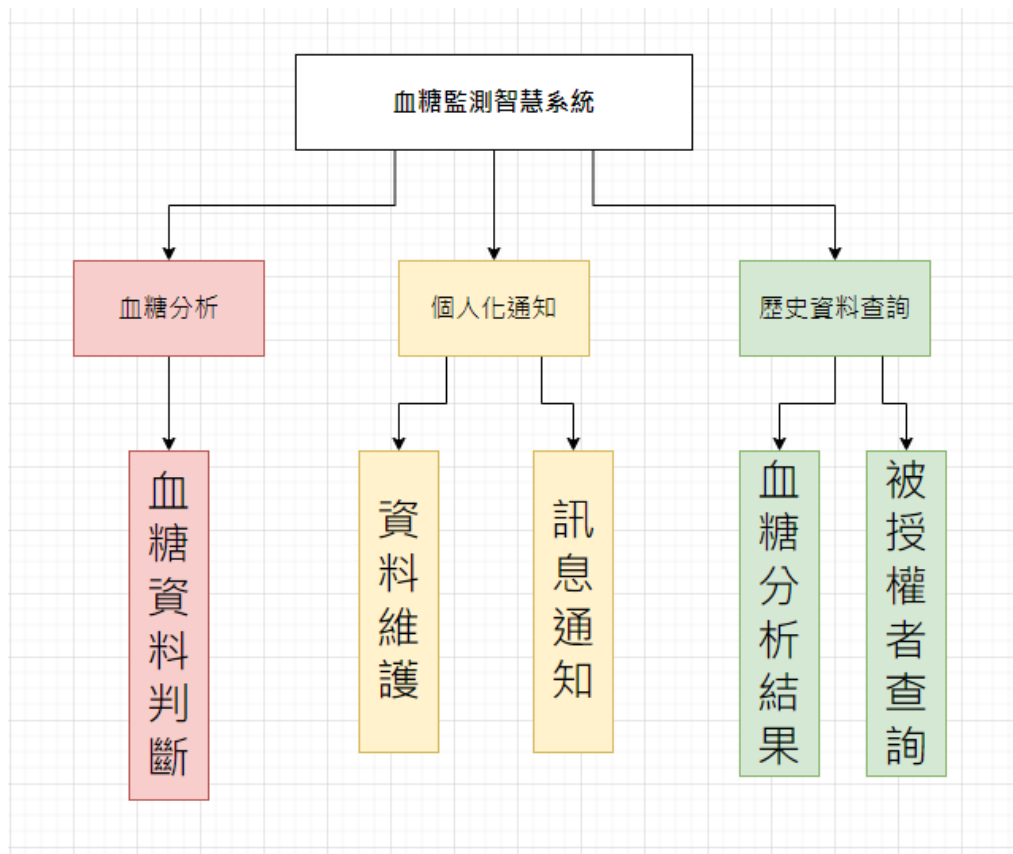


圖 1、「血糖監測智慧系統」功能架構圖

(一) 血糖分析

這功能主要是進行血糖判斷分析的設計。透過決策樹模式以現有的資料進行訓練已產生規則。並且會使用上述所提三種決策樹的結果做比較，選擇成效較好之並與醫生做討論，將所得模型規則寫入，再依規則產生紅黃綠等三種警示行為。進行血糖判斷時，首先會從資料庫獲取使用者個人相關資訊，然後對所擷取出來的資料進行前處理；最後再將經過資料前處理後的資料，放入訓練後的模型，就能預測是否處在高風險的狀態。也會做異常分析，若是血糖維持穩定突然高過由使用者設定異常值，會做高風險的提醒。

(二) 個人化提醒

主要是當判讀到使用者的使用異常時，依據使用者個人提醒的設置，提供適當的提醒，以提高使用者避免遭受危害。系統首先會判斷使用者是否有準時測量，若有異常使用一些預設的提醒方式，如跳出通知或是對設置的聯絡人進行提醒；若是分析完發生異常，會給予紅燈並提醒，反之則為綠燈，黃燈則為警示需要多加留意。系統會依據使用者自行設定個人的喜好，發出相對

應的提醒。

然後，系統會於資料庫中，記錄使用者量測及個人相關資訊，隨著使用者對於系統的使用，系統會不斷累積使用者因所產生的資料，有了這些資料，再透過資料分析，即能對於不同使用者，作出最適合使用者個人狀態判斷的方式。如當體重減少的使用者被偵測血糖上升、或當年紀增加的使用者被偵測到血糖值超出年紀所適合的血糖值時，當按下分析按鈕系統能判斷為紅燈並即時做出提醒。

(三) 歷史資料查詢

歷史資料查詢主要是依賴資料維護中寫入的血糖測量資料、時間等資料，透過 APP 頁面的方式，讓使用者能隨時隨地去查詢個人的身體健康狀況，並透過相關的統計圖與儀表板進行呈現，若有異常能及時尋求協助。被授權人也能獲得使用者資訊，若是使用者有異常行為，也能及時獲得資訊。

(五) 預期結果

本計畫期望能達到下列的成果：

1. 可提供病患一個簡便且容易操作的血糖數值記錄方式，亦可使病患的朋友、家屬或是相關的醫療院所輕易掌握該病患血糖控制的現況與趨勢
2. 以實際的血糖資料集來建構能與 APP 結合的智慧型血糖控制資料分析模式。
3. 對於有血糖相關問題的國人提供更好的監測、預防方法。
4. 能將所得出之規則實際與血糖機廠商配合，將 APP 推廣至實際使用面，進而獲取更多資料以優化學習之模型。
5. 研究成果申請專利。

(六) 參考文獻

1. 林雅鈴，「以糖化血色素與空腹血糖值預測糖尿病之發生情形：以台灣中老年糖尿病高危險群者為例」，亞洲大學健康產業管理系碩士班碩士論文，2016。
2. 鄒佳玲，「應用資料探勘技術建立退貨預測模式-以某面板公司為例」，健行科技大學工業管理系碩士班碩士論文，2016。
3. Abbas, R., Hussain, A. J., Al-Jumeily, D., Baker, T., & Khattak, A. (2018). Classification of Foetal Distress and Hypoxia Using Machine Learning Approaches. *Int. Conf. Intell. Comput.*, 767-776.

4. Centers for Disease Control and Prevention. (2020), <https://www.cdc.gov/diabetes/data/statisticsreport/diagnosedundiagnosed-diabetes.html>
5. Ilhan Satman, Beyhan Omer, Yildiz Tutuncu, Sibel Kalaca, Selda Gedik, Nevin Dincag, Kubilay Karsidag, Sema Genc, Aysegul Telci, Bulent Canbaz, Fulya Turker, Temel Yilmaz, Bekir Cakir, Jaakko Tuomilehto.(2013). Twelve-year trends in the prevalence and risk factors of diabetes and prediabetes in Turkish adults. *Eur J Epidemiol*, 28:169–180.
6. Esayas Haregot Hilawe, Hiroshi Yatsuya, Leo Kawaguchi, Atsuko Aoyama.(2013).Differences by sex in the prevalence of diabetes mellitus, impaired fasting glycaemia and impaired glucose tolerance in sub-Saharan Africa: a systematic review and meta-analysis. *Bull World Health Organ*,91:671–682D. doi: <http://dx.doi.org/10.2471/BLT.12.113415>
7. Selvin E, Coresh J, Brancati FL.(2006). The burden and treatment of diabetes in elderly individuals in the U.S. *Diabetes* , 29:2415– 2419.
8. B. C. K. Choi, F. Shi.(2001). Risk factors for diabetes mellitus by age and sex: results of the National Population Health Survey. *Diabetologia*,44: 1221-1231.
9. James WP.(2008).The epidemiology of obesity: the size of the problem. *J Intern Med*, 263:336–352.
10. Holly Kramer, Guichan Cao, Lara Dugas, Amy Luke, Richard Cooper, Ramon Durazo-Arvizu.(2010).Increasing BMI and waist circumference and prevalence of obesity among adults with Type 2 diabetes: the National Health and Nutrition Examination Surveys. *Journal of Diabetes and Its Complications*,24;(2010);368–374.
11. J. Shon, R. Kumudha and K. Deepa, "Associated Machine Learning Techniques Based On Diabetes Based Predictions" in *Proc. ICICCS*, (2019) , pp.1445-1450.
12. B. Alic, L. Gurbeta and A. Badnjevic, "Machine Learning Techniques for Classification of Diabetes and Cardiovascular Discases," in *Proc. MECO,Bar, Montenegro*, (2017).
13. S. Divya, and P. Asha, "Earlier Diagnosis and Survey of Diabetes Mellitus Using Machine Learning Techniques," in *Proc. ICISC, Coimbatore, India*,2019, pp.37-38.

(七)需要指導教授指導內容

在修習資訊系統專案時，接觸了一些有關醫療的概念，在日常生活中、新聞上或是各種影片當中，看到了許多關於血糖機的使用方式或是利用血糖機去測試一項食物適不適合容易有高血糖的人去食用，然後將資料記錄在App中，或是只能透過紙本記錄，但也都僅僅只是記錄，血糖機的使用不限於老年人，在各個年齡層

都有，血糖機App對於他們有重要的存在意義不能只有紀錄，有鑒於此決定要讓血糖機App不單單只是只有紀錄的功能，而是利用他實現更多的功用，把使用價值發揮到最大。我修習過多個數理相關的課程，如微積分、統計學及離散數學。同樣修習過多個程式設計的課程掌握良好撰寫程式基礎，像是Web前端設計、後端程式設計、進階web程式設計、C語言、python等等。雖然經過相關的學習，但關於本計畫的知識還是有許多有待加強的地方，因此本計畫需要指導教授指導的內容可以分為下列幾部分，

1. 協助與醫師討論醫療資訊議題的介紹，以及醫療資料整理的方式。
2. 協助與醫師討論血糖預測相關議題以及重要風險因子變數的討論
3. 機器學習分類技術之原理以及其應用方法的介紹。
4. 協助建構使用者模型預測之模式。
5. 協助建立實驗流程。
6. 協助與醫師一起評估模式結果
7. 協助與醫師一起評估APP的應用結果。