

運用機器學習提升國軍歲出 預算用途別判斷效率

李昕恩

國防大學教官

摘要

為解決各部門預算承辦人在預算支用時面臨預算用途別判斷不易之問題，本研究主在探討如何運用人工智慧技術建構國軍歲出預算用途別自動判斷模型，以提升預算用途別判斷效率。本文係以國軍A單位109年至112年預算簽證資料為基礎，運用機器學習技術建立準確率達90%以上的分類模型。此智能分類模型的應用不僅可提升歲出預算用途別判斷的效率，亦帶來下列效益：一、從源頭改善錯誤率，減少預算承辦人用途別誤用情況；二、實現雙贏策略，節省主計人員的輔導成本，同時減少代辦補證案件；三、提升主計職能，使主計人員將審查工作著重在經濟支出與否；四、用途別分類模型可作為預算承辦人編列預算之輔助工具，有助預算編製工作推行。

壹、緒論

依據預算法37條規範，歲出預算應按政事別、計畫或業務別與用途別科目編製，以呈現政府施政狀況與重點，並得以歸屬各計畫所需費用性質。歲出預算用途別在經過行政院逐步

研修後，當前用途別設立目的係以滿足會計資訊之需求，進而發揮成本效益分析及績效評核之功能。在預算表達上，預算之執行控制僅及於第一級科目，第二至第三級則供會計紀錄之用；然為能確切掌握國軍預算執行績效評核及精準成本歸戶等緣由，國防部歷年除依循行政院主計總處針對國軍預算用途別進行調整，亦強化第三級用途別使用範圍之教育訓練工作，體現用途別之於預算執行的重要性。

此外，在帳務中心簽證案件審查與年終巡迴審查（年終駐審）時，預算用途別亦是審查要點之一（如表1、表2），可知歲出預算用途別誤用仍為各單位常見問題。然預算執行要正確且有效，並非預算單位主計部門可獨立達成，還須仰賴各部門預算承辦人員，然而預算承辦人員並非主財官科人員，難以全面知悉經費用途別定義與運用範疇，容易肇生用途別誤用情事，故主計部門預算官便需於請購時費時協助預算承辦人，於結報、簽證時審查用途別是否正確。如遇特殊支出事項，則需橫向諮詢各友軍單位亦或縱向由國防部釋疑。故為減輕預算承辦人員及主財人員在用途別判斷與運用上所耗費之人力，本研究擬以機器學習技術提升歲出預算用途別判斷效率。

表 1 112 年代辦補證態樣

項目	缺失態樣
1	統一發票（收據）應記明事項不全（如廠商統一編號、單價、數量等）
2	各級人員（或驗收人）未簽章
3	表單（冊）內容填列缺漏（誤）
4	統一發票（收據）未註記買受機關或統一編號
5	單價逾 1 萬元且使用年限逾 2 年之財產（或電腦軟體），未檢附財產（軟體）增加單
6	預算用途別與支出事項不符
7	憑單單據份數與實際檢附單據份數未合
8	國內出差旅費列支飛機、高鐵及座（艙）位有分等之船舶者，未檢附票根或購票證明
9	預算科目或用途填列錯誤

資料來源：帳務中心

表 2 112 年年終巡迴審查重點

項目	審查重點
1	工程等採購案件（含小額採購）執行情形
2	原始憑證異常
3	結報事項不當或不經濟支出
4	是否按預算用途別列支經費
5	以當年度之預算之（預）付次年度之費款
6	經費核銷支付款項之要件
7	付款情形延宕且未有適當處理案件

資料來源：國防部

機器學習（Machine Learning, ML）為人工智慧之一環，著重於訓練電腦從資料中學習，並根據經驗改進，以做出最佳決策和預測；當資料越多，準確度則越高。在這數位時代，不論民間企業或是政府部門，從自動駕駛、圖像辨識、語言分析、醫療保健、環境監測、工業自動化到金融服務、財務分析等，都廣泛運用機器學習解決問題及提升作業效率，如金融業將信用卡申請人性別、年齡、年薪、信用記錄、貸款紀錄等各項資料放入機器學習模型

中，即可利用模型判斷該申請人是屬高、中、低風險，據以核發信用卡。

由當前各項應用中，可知機器學習可以發揮最大效益的場域是：當擁有大量資料，但面對的是沒有明確答案、需要一定程度的人為判斷、具有某種潛在規律的問題時。有鑑於國軍預算用途別選用在人為判斷上之不易，遂本研究擬整合國軍雲端資訊系統數據資料與機器學習技術，以提升國軍歲出預算用途別判斷效率。

貳、機器學習

一、機器學習介紹

機器學習是一種利用演算法和收集的資料來建立模型的方法，這些模型可以用來分類、預測以做出決策。當模型訓練完成後，它可以用於處理新的資料，並根據過去的經驗來進行預測或分析（李文德，2021）。如果模型的效能可以透過不斷更新和改進，以反映新資料的情況，那麼這個過程就被稱為機器學習。

機器學習（ML）與人工智慧（Artificial Intelligence, AI）二詞看似相同，實則相異。人工智慧是讓機器跟人類一樣有智慧，故機器需藉由「在資料中找規則」的學習方式來變得聰明。Jordan與Mitchell（2015）機器學習可以廣義的分為監督式學習（Supervised Learning）及非監督式學習（Unsupervised Learning）（如圖1）：

(一)監督式學習：是指資料在提供給電腦學習時，會先將資料做標記，意即明確標示X

（特徵值）與Y（目標值），電腦便可藉由 $Y=F(X)$ 模型推導，對未知的資料做目標值預測。監督式學習經常使用於分類或回歸；在分類問題中，目標值Y屬類別變數，如預測貸款違約風險屬性是高度、中度或低度等（如圖2）；針對回歸問題，目標值Y則為連續變數，如預測股價、年薪等。在機器學習中，不同類型的問題也會搭配相應的演算法：迴歸問題多使用線性迴歸及神經網路、分類問題除了使用神經網路，另有決策樹及隨機森林等，而運用何種演算法則需根據具體問題和數據的特點來進行判斷。

(二)非監督式學習：與監督式學習相反，非監督式學習雖然同樣運用 $Y=F(X)$ 模型學習，但投入之資料並未標註目標值Y，意即不提供正確答案，故電腦需自行從所有特徵值X中自行進行分類或分群。典型例子為提供大量貓狗特徵照片給電腦，但不告知電腦照片究竟是貓還是狗，交由電腦自行判讀。搭配使用的演算法則為K平均法及神經網路。

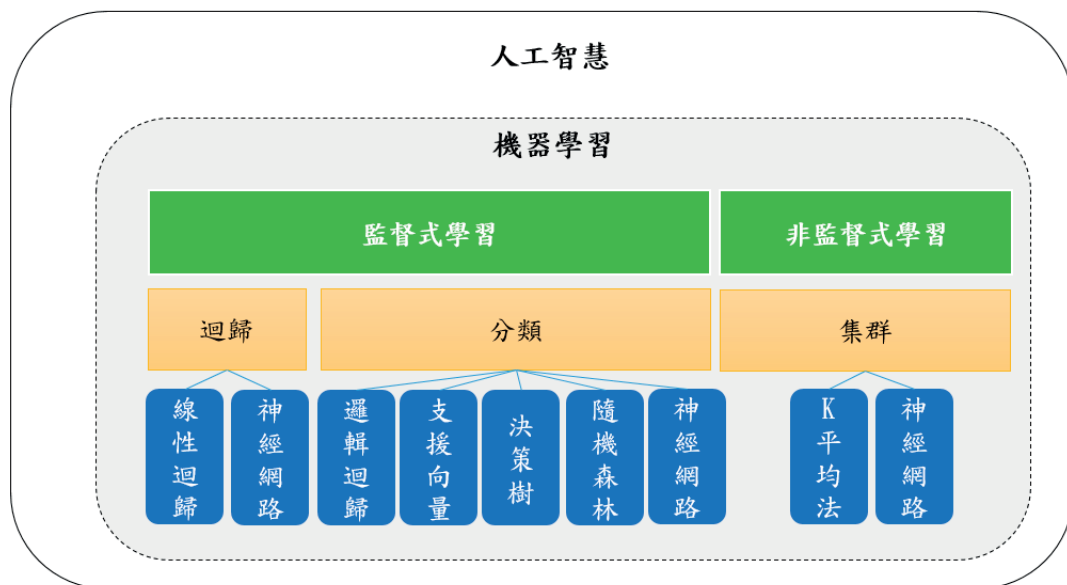


圖 1 機器學習簡介

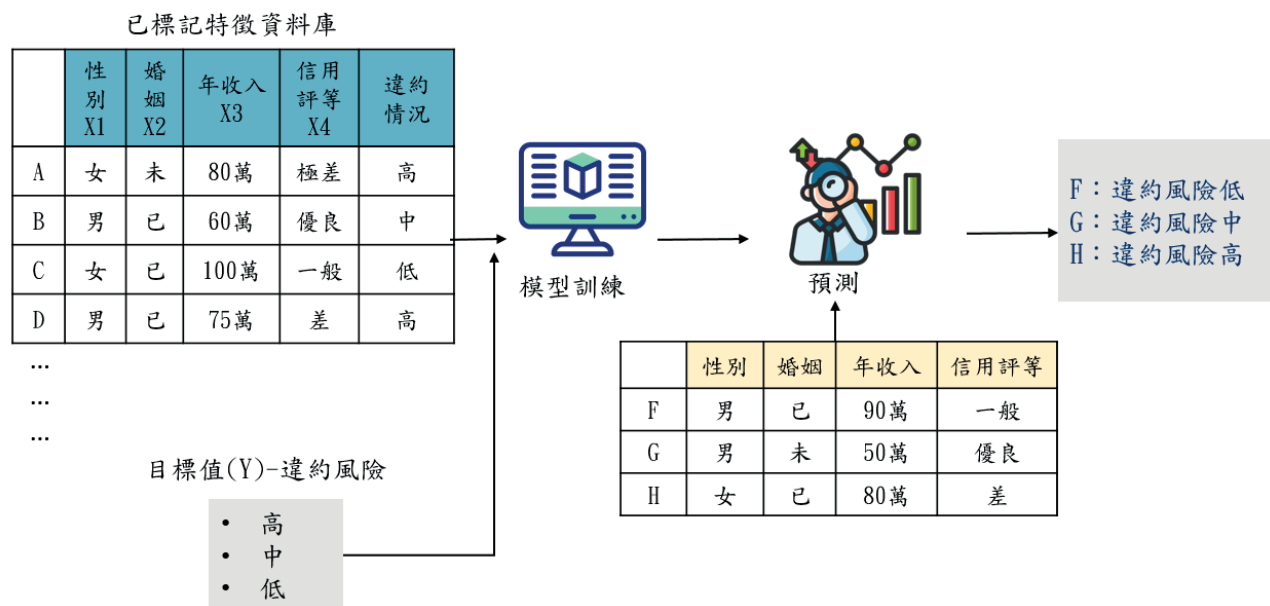


圖 2 監督式學習案例—債務違約預測示意圖

二、機器學習在財務領域的應用

近年在財務領域中，已有許多運用機器學習輔助預測的例子。孫嘉明（2022）大多數公司在進行雜項支出或非例行性支出時，各部門請購人員會在ERP系統中繕打支付憑單，此時請購人員需依支付項目自行鍵入會計科目。但由於支付類型龐雜，會計科目亦因交易用途而有差異，即便會計部門針對各部門人員進行教育訓練，仍可能發生疏漏與錯誤。為提升會計作業行政效率，Liu等（2021）與臺灣半導體公司進行產學合作，運用機器學習進行會計科目判斷—各部門人員只需透過填寫交易明細¹，模型便會透過明細分析出合適的會計科目，且模型正確率高達90%以上。在建置模型後，財會部門僅需針對未使用系統推薦的會計科目做重點審查，有效節省人力與時間。

林佳伶（2023）運用機器學習模型預測臺灣上市公司的經營績效，其中績效因素涵蓋財務及ESG²相關之非財務指標。研究經由測試不同的演算法，得知支持向量迴歸模型與類神經網路模型能精準預測經營績效，進一步顯示ESG因素在預測及解釋公司經營績效具有重要角色。此外，在判別公司信用風險方面，投資人通常藉由會計師對當期財報提出之關鍵查核事項，據以進行投資決策。故會計師提出之關鍵查核事項是否能有效判別公司信用風險高低，實為一重要問題。因此，卓佳慶、王泓達、陳育仁與段宇娟（2021）以2016年至2018年我國上市櫃公司為樣本，結合文字探勘方式解析每一篇關鍵查核事項的文字資訊，再以機器學習方法中的遺傳基因演算法優化的支援向量機模式，得出關鍵查核事項確實可有效區分公司信用風險高低，有助投資人執行決策。

¹ 該研究特徵值X包含：交易原因（文字資料）、成本資料（分類型變數）、交易金額（連續型變數）、廠商、廠房/單位、採購人員；目標值Y為會計科目（分類型變數）

² ESG分別為環境保護（Environmental）、社會責任（Social）以及公司治理（Governance），代表的是企業社會責任。現今眾多企業和投資者已將ESG評分視為評估企業永續經營性的重要指標，對於投資決策也有相當程度的影響。

有鑑於此，本研究為解決國軍歲出預算用途別選用在人為判斷上之不易，擬建構機器學習模型進行國軍用途別預測，以降低用途別誤用情事，並提升主財人員及經費承辦人預算執行作業效率。

參、模型建置

一、決定目標

建立國軍預算用途別判斷模型，提升主計人員及經費承辦人預算執行效率。

二、蒐集資料

蒐整國軍A單位109年至112年預算簽證資料，共計5萬786筆。

三、變數分析與資料清洗

(一)變數分析：模型目標為判斷預算用途別，故目標值Y設定為三級用途別編號，另從簽證資料中擇取可能影響用途別判斷的因素為特徵值：X1預算科目代號、X2支出事項及X3金額。其中，Y用途別編號及X1預算科目代號屬名目變數；X2支出事項為文字，後續將把文字轉換為編碼，模型才得以判讀；X3金額則為連續變數（表3）。

(二)資料清洗：原始資料計5萬786筆係涵蓋所有簽證類別（AA、AC、AB、BB、BC、BD），為避免資料重複產生偏誤，故刪除BB與BD及其他具缺漏值之資料共545筆，正式使用資料集總計5萬241筆。以一級用途別劃分，20業務費總金額及簽證筆數占比最高（分別為63.30%、85.22%），10人事費次之（分別為16.04%、10.36%），總計資料分析如表4、圖3、圖4。

表 3 變數列表

	變數名稱	變數涵義	類別	舉例
特徵值	X1	預算科目代號	國軍預算科目表之歲出預算科目	名目變數
	X2	支出事項	預算支出事由	文字資料
	X3	金額	預算簽證金額	連續變數
目標值	Y	用途別編號（三級）	國軍預算科目表之歲出用途別	名目變數

表 4 資料統計分析

用途別	金額（元）	金額百分比	筆數	筆數百分比
10 人事費	627,482,881	16.04%	5207	10.36%
20 業務費	2,476,183,315	63.30%	42817	85.22%
30 設備及投資	736,339,317	18.82%	1803	3.59%
40 獎補助費	71,758,772	1.83%	414	0.82%
總計	3,911,764,285	100%	50241	100%

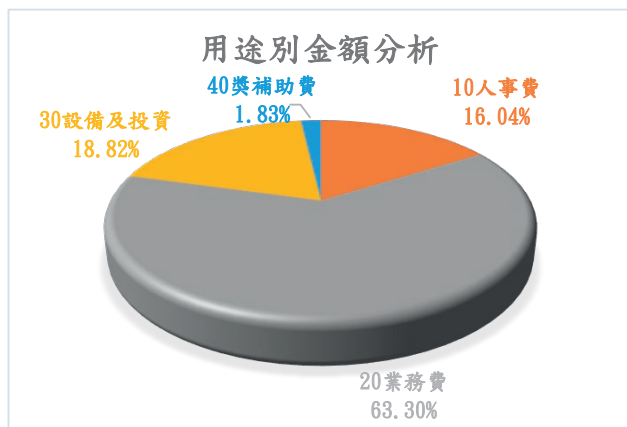


圖3 用途別金額分析

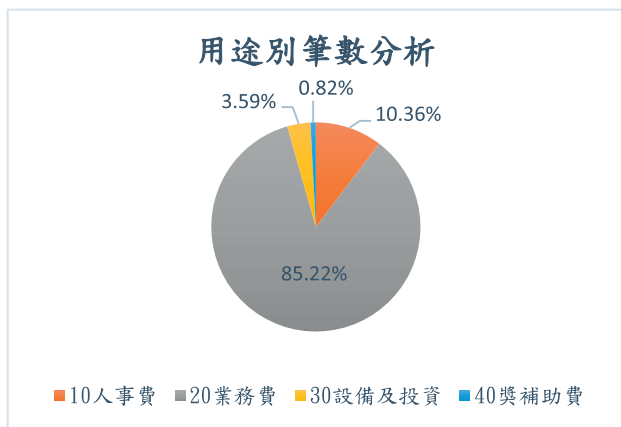


圖4 用途別筆數分析

四、模型選擇

本研究使用資料集已明確標註X及Y，故屬監督式學習；其中目標值Y用途別編號為名目變數，遂本研究採用分類型機器學習模型。分類型之機器學習模型包含決策數、隨機森林、神經網絡等等，考量本文目的、研究構想及變數設定與Liu et al (2021) 一文相近，將參酌其運用隨機森林模型，主要原因如下：

- (一) 隨機森林係傳統機器學習模型中，分類功能強大的模型。
- (二) 隨機森林是從決策樹集成而來，在模型解釋上相對容易。
- (三) 隨機森林可容易看出各個變數的重要性。

五、進行訓練

運用python程式及5萬241筆簽證資料建立歲出預算用途別預測模型。此處針對訓練過程

的重點進行說明：

- (一) 模型訓練概念：以資料集內的「支出事項」為主要變數，再搭配「預算科目」其「金額」去學習用途別。其中5萬241筆資料將隨機被分為80%的訓練集，20%的測試集，避免模型出現過度擬合狀況，影響模型預測準確度。
- (二) 關鍵詞庫匯入：本研究變數一支出事項為文字資料，是攸關用途別使用的關鍵，故文字須能轉換為電腦理解形式，模型才能具有顯著的預測能力。本研究參考眾多文獻後，使用基礎且實用之jieba（斷詞）³套件進行句子分詞，再運用One-Hot Encoding（熱編碼）技術⁴將文字轉化為數值。但因國軍預算執行使用詞彙不會出現在jieba內建資料庫中，為使斷詞及熱編碼更加精確，故本研究彙整現有簽證資料之用途別關鍵詞彙，並加入原始資料庫中（表5節錄部分關鍵詞庫）。

³ jieba（斷詞）係透過內建文字資料庫來判斷句子的斷點，如將「專任教師超課鐘點費」自動斷成「專任教師」、「超課鐘點費」。

⁴ One-Hot Encoding技術是將所有可出現的元素用1及0組成的陣列來表示，假定可能出現元素有「超課」、「兼課」、「鐘點費」、「教師」，如有一個句子為「專任教師超課鐘點費」，此時便能用One-Hot陣列[1,0,1,1]表示這個句子。

表 5 國軍 A 單位簽證資料之用途別關鍵詞彙（節錄）

用途別 編號	三級用途別	關鍵詞庫
101510	軍人待遇	加給；教召薪餉；空勤；資訊
101525	教師待遇	超課；鐘點費；專任；教師；兼課；超兼；超（兼）課；教官
102015	兼課鐘點費	兼課；鐘點費；兼任；超（兼）課；教師；教官；外籍教師；薪資； 軍人特質；ROTC 大學儲備軍官訓練團
103505	主副食及服裝	勤召；伙食費；服裝代金；教召；主副食費
200305	教育費	學分費；公餘進修；補助；語文檢定；短期交流教育費
202410	規費	水源；保護保育；下水道使用；空氣汙染；空污費；規費；垃圾袋；水費； 專利維護；燃料使用；服務費；防制；許可費；資料；申請

(三) 模型準確率：本研究以資料全集（5萬241筆）為機器學習資料，但為確認關鍵詞庫對於模型訓練之重要性，將模型區分為「斷詞時使用原始內建資料庫」及「斷詞時使用自定義關鍵詞庫」（意即在模型中引入前述之用途別關鍵詞彙），結果顯示：當使用自定義關鍵詞庫時，模型準確率87%，遠高於使用jieba套件原始資料庫，可知關鍵詞庫是模型訓練的重要資訊（表6）。

六、模型調整

然而，為使模型準確率達到90%以上，以

彰顯機器學習對於國軍歲出預算用途別判別成效，本研究試將資料集進行區分⁵，並同樣以有無使用自定義關鍵字庫來測試準確率（表7）：

(一) 10用途別：在模型僅需判斷10人事費項下三級用途時，因支出事項相較20業務費單純，故在有定義用途別關鍵字庫時，準確率高達98.5%。

(二) 20用途別：業務費支用筆數多，且支出用途範圍廣且繁雜，在使用自定義關鍵字庫進行預測時，與放入全部用途資料集相比，準確率從87%變成88%，並無明顯提升。

表 6 原始模型準確率

資料集範圍	使用原始資料庫	自定義關鍵字庫
資料全集	71.3%	87%

表 7 模型調整準確率

資料集範圍	使用原始資料庫	自定義關鍵字庫
全部用途	71.3%	87%
10用途	87%	98.5%
20用途	73.6%	88%
排除 205105、205115 用途	78%	93.3%

⁵ 考量資料中20業務費用途及10人事費用途為主要，故僅針對此兩類做模型測試。

(三)刪除205105（消耗品）及205115（非消耗品）：經審視20業務費用用途的資料集，觀察出三級用途別205105及205115之支出事項過於多元，故難以運用關鍵詞庫定義；且考量經費承辦人能容易判斷物品費之使用，故此處將205105及205115資料排除，發現此模型相較於全部用途模型，在有定義關鍵字庫時，準確率達93.3%，參照相關文獻，模型準確率若能達到90%，已能證明模型效用。

七、進行預測

模型調整完成後，本研究選定排除消耗品及非消耗品之模型為測試模型。於此，使用者可輸入「預算科目」、「支出事項」、及「金額」，模型即可提供最適用途別（圖5、圖6）。示意圖如下：使用者輸入預算科目代號「120103」，支出事項「辦理演講鐘點費」，金額「4000」，模型立即給予用途別203615（講座鐘點費）建議（圖7）。

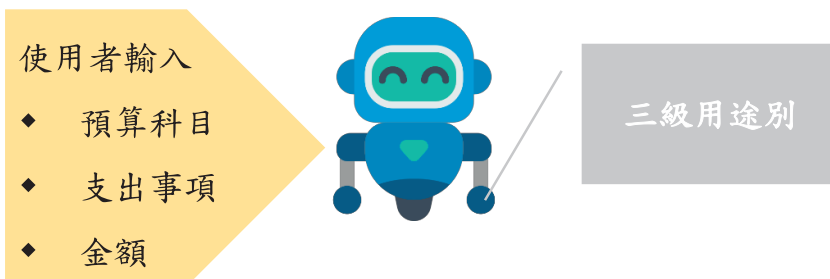


圖5 本研究模型預測示意圖（一）

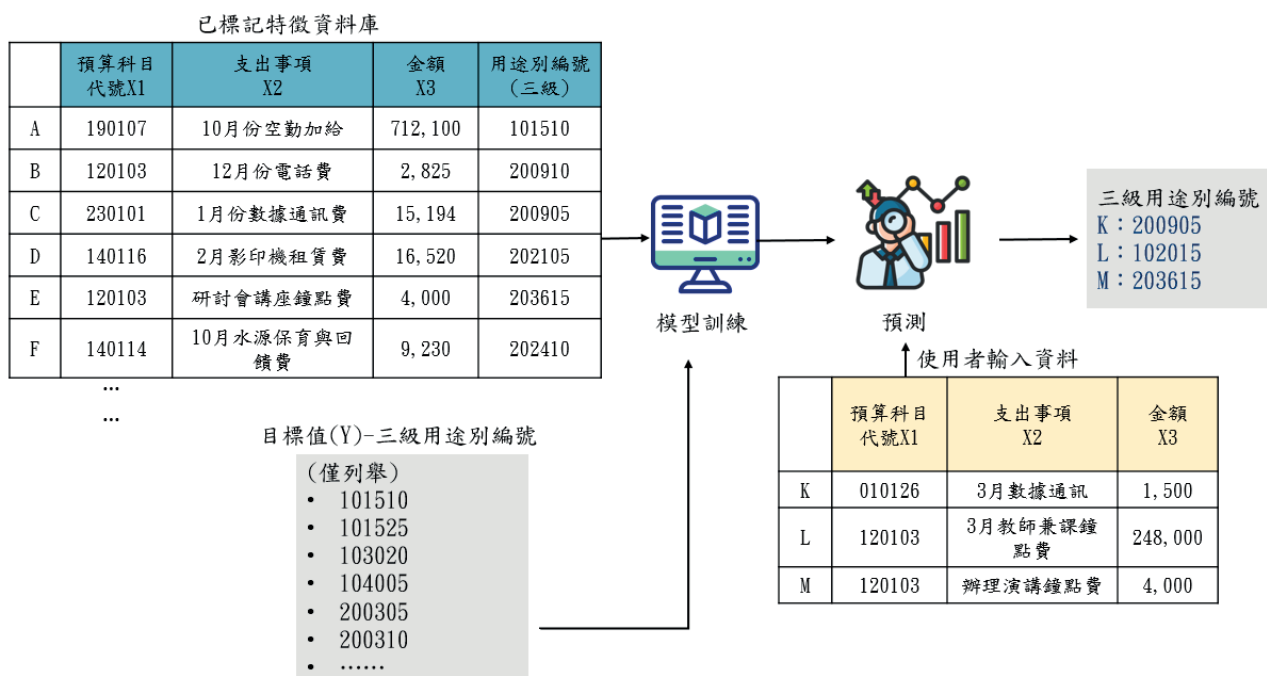


圖6 本研究模型預測示意圖（二）

請輸入預算科目代號：120103
 請輸入支出事項：辦理演講鐘點費
 請輸入金額：4000
 預測的用途別是：203615
 Accuracy: 0.9330456226880395

圖 7 模型實際預測畫面

肆、結論與建議

一、為解決各部門預算承辦人與主財人員在預算支用時，面臨預算用途別判斷之不易，本研究運用國軍A單位109年至112年預算簽證資料進行機器學習，建立起準確率達90%以上之分類模型，可有效提升歲出預算用途別判斷效率。透過此智能分類模型，各部門預算承辦人可輸入預算科目、支出事項及金額，據以獲得用途別科目建議，可帶來幾項成效：

- (一)從源頭改善錯誤率：減少預算承辦人在預算執行時辨別用途別的窒礙與誤判情事。
- (二)雙贏策略：降低主計人員協助預算承辦人案件審查之時間成本，並減少國軍平時及年終駐審的代辦補證案件。
- (三)提升主計職能：在當前主財人力負荷大的環境下，主計職能轉而發揮在經濟支出與否的審查工作。
- (四)有助預算編製工作：預算編製與預算執行實為預算的生命週期，對預算承辦人而言，無論是執行或是編列，用途別分類模型都能發揮輔助作用。

二、為使機器學習模型可落實於實務工作，本研究提出幾項建議：

- (一)模型運用於預算動支階段：本研究為確保資料正確性，係以預算簽證資料進行

機器學習，然如期望發揮分類模型效益，此分類模型應運用在預算請購階段，而非預算簽證階段。

- (二)關鍵詞庫建立與維護：從前章模型測試可知，當自變數有文字類型資料，在進行文字分析時，考量國軍預算執行支出事項與一般使用詞彙不同，故在機器學習過程中需導入自行建立的關鍵字資料庫。而當預算執行數據漸增，關鍵詞庫更應保持更新，以維模型預測效能。

三、本研究首次提出機器學習運用於主財業務構想，故僅以國軍A單位進行測試，且該單位支出事項範圍較明確。後續如能取得各預算支用單位資料，則可考慮將「採購單位」（如：人事、情報、作戰等科室）納入，使模型可以據該單位的支出紀錄給予適當用途別之建議，預期可大幅提升預算簽證之正確性。

參考文獻

中文部分

1. Liu, S. Z., Wang, C. K., Lee, C. C., & Huang, Y. S. (2021)。機器學習與會計科目判斷—臺灣半導體公司使用人工智慧提升行政效率個案對會計教育之啟示。中華會計學刊，17（2），245-268。

2. 張智揮（2023）。主計季刊。國軍預算科目及用途別科目研析。主計季刊第64卷第2期，18-30。
3. 孫嘉明（2022）。數位轉型趨勢下之內部稽核數據分析與人工智慧應－2021年全球內稽年會活動分享。內部稽核，（116），4-9。
4. 林佳伶（2023）。機器學習模型在財務與非財務性ESG指標對臺灣上市公司經營績效之探討與應用。
5. 卓佳慶、王泓達、陳育仁、段宇娟（2021）。以關鍵查核事項偵測公司信用風險－機器學習之應用。會計審計論叢，11（2），85-111。

英文部分

1. Goodell, J. W., Kumar, S., Lim, W. M., & Pattnaik, D. (2021). Artificial intelligence and machine learning in finance: Identifying foundations, themes, and research clusters from bibliometric analysis. *Journal of Behavioral and Experimental Finance*, 32, 100577.
2. Jordan, M. I., & Mitchell, T. M. (2015). Machine learning: Trends, perspectives, and prospects. *Science*, 349 (6245), 255-260.