

《測驗學刊》
第六十八輯第二期 2021 年 6 月 75~99 頁

以支援向量機處理題型符號與文字特徵應用 於微積分試題難度分類

林泓宏¹ 蘇家鈺² 張勝麟³

摘要

本研究主要在建立微積分試題的「題型符號與文字特徵」，透過人工歸納，對各類型的微積分試題擷取試題符號特徵，並轉換為向量表示。接著，對試題特徵向量，分別以主成分分析（Principal Component Analysis, PCA）及線性判別分析（Linear Discriminant Analysis, LDA）做降維處理，找出較符合試題難易分布的特徵空間，最後利用支援向量機對降維後的試題特徵，估計試題的難易度，透過使用支援向量機 RBF 核函數進行「難、中、易」之試題分類。就文獻探討所知，本研究所提出的「題型符號與文字特徵」計算表示形式，為國內外相關研究中創新的特徵集設計。實驗結果顯示：在 5 摺交叉驗證測試下，對單一摺測試集之微積分試題難易度分類，最高可獲取 95% 的正確率，而 5 摺的平均測試正確率也可達 90.19%，基於實驗測試結果遠高於隨機亂猜的 33.33%，而對 3 個類別中，隨機亂猜的 95% 信賴區間上限約在 42.69%，可看出本研究方法的實驗結果大幅高於亂猜達 47.5%，顯示本研究所提出的「題型符號與文字特徵」對於微積分試題難易度分類具有顯著的功效。

關鍵詞：支援向量機、文字特徵、微積分、試題難度分類、題型符號

1. 林泓宏，南臺科技大學資訊工程系助理教授

2. 蘇家鈺，南臺科技大學資訊工程系碩士

3. 張勝麟，南臺科技大學資訊工程系教授

收件日期：2020.03.31；完成修改：2021.05.07；正式接受：2021.05.18

通訊作者：張勝麟；Email：slchang@stust.edu.tw

地址：710301 臺南市永康市南臺街 1 號 南臺科技大學資訊工程系

Difficulty Level Classification of Calculus Exam Questions Using SVMs with Descriptive Features of Symbols and Texts

Horng-Horng Lin¹ Chia-Hong Su² Shing-Lin Chang³

Abstract

A new design of "descriptive symbol and text features" of calculus exam questions has been proposed in this paper. The proposed descriptive features of symbols and texts can be extracted from various calculus questions and are represented by vectors. The high dimensionality of extracted features from test questions is then reduced by principal component analysis (PCA) or by linear discriminant analysis (LDA) for finding a lower dimensional feature space that better fits the difficulty-level distribution of test questions. Subsequently, a support vector machine with radial basis kernel is adopted to categorize calculus questions into three degrees of difficulty, i.e., hard, medium and easy. To the best of our knowledge, the proposed descriptive feature representation of symbols and texts of mathematical questions is a novel design for difficulty level estimation of calculus exam questions and is rarely seen in previous literature. In our experiments of difficulty level classification with 5-fold cross validation (CV), the highest classification accuracy of difficulty level of calculus questions in a test fold is 95%, while the average classification accuracy of 5-fold CV is 90.19%. These results are far higher than the mere 33.33% accuracy of random guess. For the three categories, the upper limit of the 95% confidence interval for random guess is about 42.69%. It can be seen that our result is much higher than the upper limit of random guess by about 47.5%. Validate the significant effectiveness of the proposed descriptive features of symbols and texts of calculus exam questions on automatic difficulty level prediction.

Keywords: calculus, question description symbol, question difficulty level classification, support vector machine, text features

1. Horng-Horng Lin, Assistant Professor, Department of Computer Science and Information Engineering, Southern Taiwan University of Science and Technology
2. Chia-Hong Su, Master, Department of Computer Science and Information Engineering, Southern Taiwan University of Science and Technology
3. Shing-Lin Chang, Professor, Department of Computer Science and Information Engineering, Southern Taiwan University of Science and Technology

Received: 2020.03.31; Revised: 2021.05.07; Accepted: 2021.05.18

Corresponding Author: Shing-Lin Chang; Email: slchang@stust.edu.tw

Address: No. 1, Nantai St., Yongkang Dist., Tainan City 710301, Taiwan

Department of Computer Science and Information Engineering, Southern Taiwan University of Science and Technology

壹、簡介

臺灣的大學入學評量之學科能力測驗（簡稱學測）是用 15 級分制，2020 年出現數學滿級分人數破歷史紀錄達 11%，部分輿論認為題目太簡單、缺乏鑑別度，後續引起大學入學考試中心主任辭職；由於大型考試試題需要保密、不能外洩，所以試題難度設定多以命題者認知為主，且無法廣泛請其他同領域專家來事先評定試題難度，雖然可安排預試生做試題練習、評估，但預試生的人數有限，程度上也不見得能完全隨機挑選，因此很難精準反應出試題對於所有考生的鑑別度。由於現今人工智慧的發展，這類情況可藉由機器學習，分析過去題庫與大考學生的答題結果，來輔助試題難度的判斷，尤其是數學科的試題，評分上少有閱卷者主觀評判之影響，試題難易度皆由學生答題結果反映呈現，更適合發展自動難易度判別的研究。

測驗型態的發展，從早期傳統建立一堆紙本筆試測驗卷，老師直接選出一份考卷或組合新的試題卷，其中自行評估試題難度，再讓學生測驗完畢後做批改登記。到初期電腦的應用，老師可在電腦的記憶體建立題庫存檔、做試題分類，再從中選題，印出考卷，給學生做紙本測驗，接著做批閱。該時期，電腦只讓老師儲存試題資料。再到進階的電腦輔助測驗，老師可建立測驗題庫、試題難易度、出題範圍、可做紙本測驗或在個人電腦讓學生做測驗、並可蒐集學生作答資訊分析。而現在雲端網路平臺的建立，學生可利用個人電腦或平板電腦或手機做到無所不在的學習及雲端同步測驗。因此，老師不只可建立題庫及評量試題的難易度，再根據學生作答能力，即時推薦測驗試題，可讓學生做適性化測驗，也可讓學生利用各種智慧裝置直接做雲端測驗，達到可非定點及定時的測驗，且成績或試題解答也可讓學生及時得知，學生可馬上了解自己的能力的範圍，並求助學習輔導或補救等機制，而學生的作答相關資料也可傳至後端做分析，同時也可給學生適時的推薦學習機制，如多媒體影音教學等，讓學生做適性化學習。然而，這些工作要達成，在試題的建置上除要對試題做結構分類外，還必須建立試題的難易度分析。如此，才能讓學生做適性測驗，進而達到適性學習，也可讓老師及電腦根據學生在學習平臺的測驗資料給出有鑑別度的試卷。

適性化學習的目的在能由學習者經學習或評量後，由系統針對學習者提供較適合的學習內容，以提高學習者的學習效果（Foshee, Elliott, & Atkinson,

2016; Johnson, Phillips, & Chase, 2009; Kerr, 2016; Koedinger & Aleven, 2007; Lin, Guo, & Lin, 2016; Walkington, 2013; Yarnall, Means, & Wetzel, 2016)，如強調適性教學的教育平臺「因材網」(<http://adaptive-learning.ntcu.edu.tw/>)。由於電腦科技的技術愈來愈進步，相信會有愈來愈多的教育工作者會投入更多適性化學習的研究，建置更多的適性學習平臺及題庫，來提升學習者的學習成就。

在建立雲端題庫時，對試題在整體受試者的能力表現做分類是很重要的。本研究稱為試題的難易度(Difficulty Level)標記，這是需要大量的人力及時間才能完成，目前大致上的作法細分為以下四種：一是由專家去判斷試題的難易度，雖然專家判定的準確率較高，但每次有新試題都須經專家判斷，成本較高，也比較不可行；二是專家判斷與學生施測同時進行，再由專家對二者有偏差的試題做後續的難易度調整，這比作法一更精準，但每個試題需要夠多的施測學生才會反映出該試題的難度，因此若試題多就需較多的學生做測驗，但當有新試題進來如何判讀也會是個問題；三是由學生直接先做測驗，再根據答對率的施測結果做判定，就如同古典測驗理論(Classical Test Theory, CTT)，容易執行(Bichi, 2016)，但若準確率要高，每個試題的學生測驗樣本量要夠大。另外，對新試題還是需要大量學生做測驗才能判定難度；四是學生施測完後，帶入試題反應理論(Item Response Theory, IRT)做判定(余民寧, 2009; Lord, 1980)，例如：Xu 與 Douglas (2006)將該模型建立在電腦化適性測驗上，但這裡需要大規模的學生施測相同的試卷，才能對試卷上的題目做難度判讀，它的好處是試題難度約在-3~3 的連續區間，可對難度做更細緻的分類，缺點是一份試卷不太可能出太多題目，若試題數多時就需分多次測驗。

以上要建立這些試題難易度的機制，不是需要專家判讀、就是需要大量學生做測驗；更麻煩的是當有新試題時，又要重複做這些過程才能獲取難易度，太過於耗時又費工，且對已判讀的舊試題沒有給新試題帶來任何參考資訊，可說是資源的浪費。因此，本研究將對微積分試題做語意分析及難易度分類，產生一個分類模組來判別難易度，如此，對新試題就只需帶入模組做判讀即可，這樣一來可節省很多的人力及具有時效性，此研究方向預期可對試題型態發展與難度評估，建立新的里程碑。

在 Cheng 與 Chang (2016) 的微積分適性化教學平臺實驗中發現，學生做自主學習及測驗，其學期成績有顯著的比沒有上平臺練習的好，這呼應了

林泓宏 蘇家鈺 張勝麟 以支援向量機處理題型符號與文字特徵應用於微積分試題難度分類

很多的論文研究結論，說明自主學習能力愈強的學生，成績會愈好（Dignath & Büttner, 2008; Dignath, Büttner, & Langfeldt, 2008; Fong & Krause, 2014; Hattie, Biggs, & Purdie, 1996; Leidinger & Perels, 2012; Mok, Cheng, Moore, & Kennedy, 2006）。由此可知，設計好的學習教育平臺對自主學習的學生是有幫助的，如「因材網」或「均一教育平臺」（<https://www.junyiacademy.org/>）。然而，若自主學習平臺能導入適性化學習，不但可提升學生的學習效果，也可節省學習時間；而要作到平臺的適性學習或智慧適性診斷，都需要對試題做難易度分析及判斷，這些大量的工作若有專家來執行或有學生來測題，除了要有經費，還要花很多的時間。科技發展的目的在利用機器來取代人類的工作，提高生產或檢測的效率，尤其對需要特定人力識別、判讀的工作，因為人事成本較高與時效性考量，那麼適當的利用機器學習來進行試題難易度判斷，是非常值得探討的問題。

貳、文獻探討

一、試題難度分析

在 CTT 定義試題難度為答對人數的百分比，以 p 表示。也就是計算全體受試者 N 人中，答對該題人數 R 的百分比，其公式為：

$$p = \frac{R}{N} \times 100\% \quad (1)$$

CTT 的判斷方法比較簡易，淺顯易懂（Bichi, 2016）。 p 值愈大，表示該題答對的人愈多，也就是題目愈簡單。另一種根據總分的高低依序排列，分成高分組，為從最高分向下取前 27% 的高分者，再從最低分向上取後 27% 為低分組，再分別計算高分組與低分組在每一個試題的答對人數百分比，分別設為 P_H, P_L ，其計算公式為：

$$P = \frac{P_H + P_L}{2} \quad (2)$$

當 P 愈大，表示試題難度愈低； P 值愈小，表示試題難度愈高，跟式(1)的 p 值相同。然而，這裡的 P 值只是一種順序量尺，表示題目難易的相對位置，卻無法說明各試題之間難度差異的大小，如某測驗前三題的 P 值分別為

測驗學刊，第 68 輯第 2 期

0.6、0.4、0.2，表示第一題難度最低，第二題次之，第三題相對的難。雖然各題間的 P 值差距相同，但並不表示第一題與第二題在難度指數上的差別量等於第二題與第三題在難度指數上的差別量，也就是說他們之間的差距單位並不相等。

針對上述缺點，美國教育測驗服務社（Educational Testing Service, ETS）提出了一種具有等距量尺的 δ 難度指數。以 13 為平均數，4 為標準差，下限為 1，上限為 25 的標準分數。假設 z 為標準化常態分配的標準分數，則：

$$\delta = 13 + 4z \quad (3)$$

由式(3)可看出 δ 值只是 z 值的線性轉換，所以 δ 值愈小，表示試題難度愈低。反之， δ 值愈大，難度愈高，與 CTT 相反。但除了可表示試題難度的相對位置外，而且也可依照常態分配指出不同難度之間的差異量。

另外，蒐集同一份試題卷，對測試者評估每一個試題的資料，再個別分析每個試題難度的試題特徵曲線（Item Characteristic Curve, ICC）。它是由 IRT 獲得，試題反應理論可取得學生能力及試題難易度二變數，並可建立他們之間的關係式，它是反應出對學生的概括性能力之綜合評估，例如：Rasch（1960/1980）提出「Rasch 模型」：

$$p(\theta) = \frac{e^{\theta-b}}{1+e^{\theta-b}} \quad (4)$$

根據這些模型所繪製出的曲線叫做試題特徵曲線，其意義在於描述出「成功解答某一試題的可能性」和「學生能力 θ 」之間的關係。而 b 值是用來決定試題的難易度（Pinar & Oz, 2011），當 $p(\theta)=0.5$ 時 $\theta=b$ ，表示學生能力到達 b 時，會有一半的成功機率答對試題。CTT 與 IRT 的差異在於：(1) CTT 的難度是樣本依賴（Sample Dependent），會導致相同試題在不同受測群樣本的難度值會有所不同。而 IRT 則不受樣本特性的影響，不會因為受測者樣本的不同而有異；(2) CTT 的試題難度是看 p 或 P 值，其值愈大，表示試題難度愈低；值愈小，試題愈困難，這概念與一般想法相反。但 IRT 的試題難度是看 b 值， b 值愈大，表示試題難度愈高； b 值愈小，試題愈簡單，這概念與常理較符合。

IRT 的方法大概只能在一開始建立試題難度時，做大量測驗即可獲得相關資訊，若要在適性化平臺上做即時的試題難易度調整或新試題的難度判

林泓宏 蘇家鈺 張勝麟 以支援向量機處理題型符號與文字特徵應用於微積分試題難度分類

斷，則是不可行的。因為在這個過程中，每個學生所測驗的題目及題數都不盡相同，此時無法再帶入 Rasch 模型。

二、自然語言處理

為了讓電腦理解微積分試題上面的語詞及符號的能力，自然語言處理及影像識別的應用就很重要了，必須先教會電腦區分文字、數學式及符號，再進行對應的特徵擷取。在這過程中，電腦要學會文字的「斷詞」和數學式的「表達意思」，在 1980 年代末期，統計機器學習開始應用到自然語言處理，不再列出語言的所有規則讓電腦來執行，而是建立人工智慧演算法，讓電腦從訓練集中自行尋找出資料所含的模式與趨勢。目前，常用的中文自然語言處理系統包含有：語言技術平臺 (<https://ltp.ai/>)、語言雲 (<http://www.ltp-cloud.com/>)、Stanford NLP (<https://stanfordnlp.github.io/CoreNLP/>)、National Language Toolkit (NLTK) (<https://www.nltk.org/>)、NLPIR-ICTCLAS 漢語分詞系統 (<http://ictclas.nlpir.org/>)、結巴中文分詞 (<https://github.com/fxsjy/jieba>)、漢語言處理包 (<https://github.com/hankcs/HanLP>)，以及 Word2Vec (<https://code.google.com/archive/p/word2vec/>) 等，皆是針對自然語言分析提出不同的計算模型與應用平臺。但上述這些方法都是對中／英文詞意做分詞，卻沒有針對數學式做處理。就研究者的理解，利用人工智慧做試題難度判別有 Chang 與 Cheng (2017)，其中利用學生作答的正確率做試題難度的動態修正，而用自然語言理解數學問題的則有 Facebook AI 研究院 Lample 與 Charton (2020)，主要探討如何利用深度學習解微分方程及不定積分，採用樹的形式來定義數學符號，以求出特定結構之方程式解，屬於特殊求解應用。上述這些研究與本文方法自行定義數學符號結構，進行機器學習以判定一般化微積分試題難度，有顯著的分析方法與應用標的差異。

三、支援向量機相關研究

目前有很多演算法應用於人工智慧的分類上，如感知器 (Perceptron)、邏輯斯迴歸 (Logistic Regression)、k 最鄰近分類器 (k-Nearest Neighbors, KNN)、決策樹分類器 (Decision Tree Classifier) 等。其中，常被用到的是支援向量機 (Support Vector Machines, SVMs)，可在任意樣本數量、分布下，保有最大分隔疆界的線性學習模型，較能避免過度擬合 (Overfitting)，並保證極值為全域的最佳解，且可利用核函數 (Kernel

測驗學刊，第 68 輯第 2 期

Function) 來擴充處理非線性問題。其中，最被廣泛應用的核心函數是 Radial Basis Function (RBF) Kernel，或稱高斯核 (Gaussian Kernel)。本研究將嘗試利用 SVMs 演算法結合 RBF 核函數對微積分試題的難度判斷上。

在人工智慧領域中，SVMs 已有很廣泛的應用，例如：Dey、Chowdhury 與 Ghosh (2017) 用 SVMs 做臉部影像辨識；Putel 與 Jhaveri (2015) 用 SVMs 來做網路節點安全的偵測，對節點正常或丟包的分類；Le、Porwal、Hdden 與 Dentith (2012) 利用 SVMs 演算法及改良法做遙控影像資料的岩性分類研究；楊智為 (2007) 將 SVMs 結合多重貝氏網路來與專家診斷學生的錯誤類型做比較，發現可提升辨識效果；但利用人工智慧於教育應用上的資料相對較少，尤其是應用在數學式的分析上又更少了。本研究將首次利用 SVMs，結合上述所提出的語意符號特徵做數學式特徵分析，應用於微積分試題的難度估計。

參、研究方法

一、資料來源

本研究利用臺灣某科技大學建置的微積分電腦適性化學習平臺，希望能提升學生的微積分學習效果，也能減輕老師的教學負擔。該校集合有數學背景且教微積分相關課程的老師成立「微積分電腦適性化教學」社群，共有 10 位專業老師參與，花了一學年時間完成試題的認知屬性及難度判定，可見試題難度的專家判斷是很費時及耗人力的工作。該平臺利用電腦科技的技術經過幾年的努力，完成整個微積分學習、測驗、適性學習及學習履歷等功能的平臺，將微積分內容分為 14 個認知單元，每個認知單元約包含 7~8 個認知屬性，亦即該試題的主要操作能力，需人為判定的特徵，此認知屬性斷定容易，一般由出題老師就可界定。

如附錄所示為單元 3-2 及部分 3-1 試題的認知屬性與試題範例；平臺中的全部試題已超過 1500 題，平均每個單元約有 50 題老師測驗試題及學生練習的適性化試題約 60 題，其中適性化試題均由測驗試題做類似設計，題型不變，且每個試題都由專業老師界定它的認知屬性及難易度，其中難易度分為難、中、易三個等級，學生在做自主學習時，平臺會依學生的認知能力給予適時適性的練習考題，最後會給學生學習輔導及建議資料，不但可增加學生

的學習成效，也可減輕老師的輔導工作，在 Chang 與 Cheng (2017) 中有利用該平臺。因此，這些資料讓本實驗的樣本具有高度可用性。

二、微積分試題的特徵性質

在做自動試題難度估計之前，需先分析微積分試題的一些結構、差異性及表示方式等，針對微積分試題的特徵性質，茲說明重要特點如下。

(一) 題型探討

對微積分試題而言，同樣的運算方式或解法，但不同的出題方式，對學生的作答及難度感覺可能會有差異。為了研究試題的難易度特徵，做了題型研究的實驗，在一份 12 題的試卷中，分甲、乙二班，皆由同一老師上課的一年級新生，分別施測同問題但不同形式的測驗卷，其中有 3 題為定錨試題（均為選擇題），分別為附錄中試題 3-1-3、3-1-7 及 3-1-43，利用兩母體比例差雙尾檢定，考慮 $\alpha=0.05$ ，發現這三題答對率如表 1 所示，並無顯著差異（ $p\text{-值} > \alpha$ ），表示二班無顯著差異。

表 1 中的非定錨試題 3-1-12（詳見附錄）是選擇題，為甲班試題；試題 3-1-12-1 是填充題，為乙班試題，答對率就有顯著差異（ $p\text{-值} < \alpha$ ）。一般而言，選擇題本身的選項就有猜測及誘答的功能，以難度來講應該是會比填充題簡單。因此，為研究的一致性及其原本就建立有難易度的試題樣本，本研究使用的樣本以選擇題為主。

表 1 定錨試題（前 3 題）與非定錨試題（後 2 題）的答對率

答對率	試題 3-1-3	試題 3-1-7	試題 3-1-43	試題 3-1-12	試題 3-1-15
甲班	0.65	0.40	0.78	0.72	0.82
乙班	0.71	0.40	0.71	0.52	0.79
$p\text{-值}$.48	.97	.35	.02	.71

(二) 選擇項是否要列入特徵

選擇題的每個選項是否要列入題目的特徵資料集，要看試題本身的結構。對某些題目而言很直覺是一定要的，如試題 3-2-1，選項本身就是問題主體的一部分。但如試題 3-1-15，除一個正確答案外，其他選項均為誘導選

測驗學刊，第 68 輯第 2 期

項，放在甲班，另變化試題為：

若 $f(x) = (2x^3 - 1)^4$ ，則 $f'(1) =$ (1) 4 (2) 1 (3) 864 (4) 24

放在乙班，如表 1 所示，發現答對率無顯著差異 (p -值 $= 0.71 > \alpha$)，所以這裡就不會把選擇項列入特徵，而如試題 3-1-3 的「不存在」或特殊符號，像 ∞ 、 $-\infty$ 等的選項，就會列入特徵集。

(三) 能區分變數位置不同解法有差異

微積分數學式的擷取特徵比文字擷取難很多，如試題 3-2-3，把 $x \cdot y$ 對調，困難度就提升很多，這裡所謂的特徵擷取是要能區分顯函數及隱函數的差異。

(四) 指數函數與二層指數的概念

比較試題 3-2-24、3-2-27，若對二題目均了解的學生會說 3-2-24 較難，3-2-27 應該是簡單的，因屬冪函數的類型，但實際測驗發現難度卻很高，主要是學生很容易因不清楚變數及常數而誤解，因此在特徵擷取上，這類的函數會加特徵「二層指數」的結構來強化指數現象。

(五) 導數、導函數及需用計算機求近似值

因求導數的試題一般都是先求導函數再取值，但在難度的標示上，求導數要加一層帶值運算，但這影響難度不大，除非答案還需一些運算知識，如試題 3-2-14，若學生不知 $\ln e = 1$ ，則可能會微分，但會選錯答案。因此，把求導數跟導函數分開為不一樣的特徵。而有些題目須用工程計算機求近似值的選題，學生有些不會按計算機或根本對那函數不理解而無法計算出答案，所以這類題目會加標示「需用計算機」的特徵，如試題 3-2-50 會比試題 3-2-60 多增加這個特徵。

(六) 認知屬性

每個試題都會經專家認定給該試題在該單元的認知屬性，這是以試題的認知操作能力來區分。原則上每個試題至少有一個在解題時需要的能力，但

林泓宏 蘇家鈺 張勝麟

以支援向量機處理題型符號與文字特徵應用於微積分試題難度分類

有些題目可能需要若干的能力，而加入這些特徵是否對難易度的判讀有幫助，則將實驗之。

（七）合成函數

本研究將會區分二個合成函數的組合運算與三層的合成函數，例如： $f(x)=e^{2x}\sin(x^3)$ 為二個合成函數做乘法運算組成，但 $f(x)=e^{2\sin(x^3)}$ 是屬三層的合成函數，在難度上是比前項函數難微分。

三、題型符號與文字特徵分類

就各試題的文字及數學式的特徵分類是很重要之工作，因為分的好愈能反映出實驗效果。本研究參考 Word 方程式工具及 LaTeX 數學式語法的設計，把微積分試題的式子分為「結構符號及對應文字」、「函數符號及對應文字」、「運算符號及對應文字」三類，分別列舉在表 2、表 3、表 4，這三個表形成整體的「一般化微積分題型符號特徵描述集合」（Feature Description Set for Typical Calculus Question Types）；對於每一題微積分試題，無論是選擇題還是填充題，都可以透過「題型符號、文字特徵」比對，將該題目中是否出現某一符號或對應文字特徵進行統計，以計算出該題型難易度分類的特徵向量。

表 2 結構符號及對應文字

連續、連續值	導數、 $\left.\frac{dy}{dx}\right _{x=a}$	微分近似值	反函數導數、 $(f^{-1})'(a)$	向量、 \vec{u}	矩陣、 $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$
不定積分、 $\int f(x)dx$	面積	弧長	旋轉體體積	側表面積	極值
切線斜率	切線、切線方程式	極座標	極方程式	二次曲面	極限值、 $\lim_{x \rightarrow a} f(x)$
定積分、 $\int_a^b f(x)$	二重積分、 $\iint_a^b f(x)$	三重積分、 $\iiint_a^b f(x)$	封閉積分、 $\oint_a^b fx$	偏微分、 $\frac{\partial y}{\partial x}$

測驗學刊，第 68 輯第 2 期

表 3 函數符號及對應文字

常數函數、 $f(x)=c$	線性函數、 $f(x)=ax+b$	二次函數、 $f(x)=ax^2+bx+c$	高次函數、 $f(x)=a_nx^n+\dots+a_1x+a_0$	冪函數 (開平方)、 $f(x)=\sqrt{x}$	冪函數 (開 n 方)、 $f(x)=\sqrt[n]{x}$
連續函數	導函數、 $\frac{dy}{dx}$	微分、 dy, dx	反函數的導函 數、 $\frac{dx}{dy}$ 、 $(f^{-1})'(x)$	有理函數、 $f(x)=\frac{g(x)}{h(x)}$	冪函數、 $f(x)=x^a$
隱函數、 $xy^2=x^3+y^3$	二階導函數、 $y'', f''(x)$	高階導函數、 $y^{(n)}, f^{(n)}(x)$	反函數、 $f^{-1}(x)$	參數式、 $\begin{cases} x=x(t) \\ y=y(t) \end{cases}$	合成函數、 $(f \circ g)(x)$
$\log_a f(x)$ 、 對數函數	$\ln f(x)$ 、 自然對數函數	條件函數-2 $f(x)=\begin{cases} g(x) \\ h(x) \end{cases}$	條件函數-3 $f(x)=\begin{cases} g(x) \\ h(x) \\ I(x) \end{cases}$	e^x 、 自然指數函數	a^x 、 指數函數
$\sin \theta$	$\cos \theta$	$\tan \theta$	$\cot \theta$	$\sec \theta$	$\csc \theta$
$\sin^{-1} \theta$	$\cos^{-1} \theta$	$\tan^{-1} \theta$	$\cot^{-1} \theta$	$\text{Sec}^{-1} \theta$	$\text{csc}^{-1} \theta$
$\sinh \theta$	$\cosh \theta$	$\tanh \theta$	$\coth \theta$	$\text{sech} \theta$	$\text{csch} \theta$
$\sinh^{-1} \theta$	$\cosh^{-1} \theta$	$\tanh^{-1} \theta$	$\coth^{-1} \theta$	$\text{sech}^{-1} \theta$	$\text{csch}^{-1} \theta$

表 4 運算符號及對應文字

\cup, \cap	$+, -, \times, \div$	\forall	$\infty, -\infty$	$\geq, \leq, >, <$	$\sqrt{\quad}$	$\sqrt[3]{\quad}$	$\sqrt[n]{\quad}$	\wedge, \vee
\subset, \subseteq	\exists 、 存在	∇ 、 梯度	$=, \neq$!、唯一	\approx, \sim	\in	\nexists 、 不存在

四、試題轉化為向量模式

對定義完成的符號特徵，將利用來決定各試題含有哪些特徵，再將特徵轉換成數碼，其中對文字敘述要做斷詞的動作，對方程式及符號要做結構分解及判斷。為提升模型效果，特徵資料處理是很重要的工作，透過資料檢視與特徵描述整理，所得到乾淨的資料集將使模型學習到更有意義的特徵，而不是在無關的雜訊上過度擬合。如刪除所有不相關的字元，在試題 3-2-24 中，需刪除「若」、「則」、「之值為何」及標點符號「，」、「？」，就完成了前處理的工作。再將每個試題得到的特徵集轉化為電腦能理解的向量模式，每個特徵都視為一個獨立維度，如同 One-Hot 編碼，這種方法稱為詞

林泓宏 蘇家鈺 張勝麟 以支援向量機處理題型符號與文字特徵應用於微積分試題難度分類

袋模型，因為這種表示方法完全忽視了句子中的詞彙順序，但這在微積分試題上比較不會發生誤解，如試題 3-1-15 及 3-1-15-1 這二題，在試題的詞意上是一樣的，不會因式子次序顛倒而被誤答。因此，One-Hot 編碼 (Alkharusi, 2012; Cohen, Cohen, West, & Aiken, 2013) 是目前研究用的方法，以試題 3-2-9 及 3-2-10 的編碼說明在表 5，每題第一個元素為它的難易度，如試題 3-2-9 對應的向量為 [3201100010...]，其中難度類別為 3；試題 3-2-10 對應的向量為 [2110011211...]，難度類別為 2。

表 5 試題轉換為向量模式

出現頻率	難易度	高次 多項式	導數	二階導 函數	隱函數	自然指 數函數	自然對 數函數	合成 函數	＋、－	×
試題 3-2-9	3	2	0	1	1	0	0	0	1	0
試題 3-2-10	2	1	1	0	0	1	1	2	1	1

五、降維方法

為了將詞袋模型的向量進行簡化與後續視覺化顯示，在訓練試題難易度判別分類器前，可將試題特徵向量做降維處理，以找出合適的試題難易分布特徵空間。降維方法有 PCA (Principal Component Analysis) (Abdi & Williams, 2010)，用於非監督式學習，其中追求的是在降維之後能保持最大化數據的內在信息，並通過衡量在投影方向上的數據方差之大小來衡量該方向的重要性；另一種常見的線性降維方法是 LDA (Linear Discriminant Analysis) (Rao, 1948)，是一種監督式的線性降維算法，跟 PCA 保持數據信息不同，LDA 是希望降維後不同類別的數據點儘可能區分開來。

本研究將分別利用 PCA 及 LDA 做降維計算，及用 SVM 加 RBF 核函數做分類器，並利用 5-Fold 交叉驗證法的方式做重複的測試及評估，其中也對 SVM 的懲罰因子 C 及核函數參數 γ 求優化值。

肆、實驗結果

本研究採用先前介紹的微積分試題庫做自動難易度判別，其中單元 3-1 及 3-2 為樣本資料集，共有 102 題，規劃 A 組有 84 個特徵，B 組多加入 16 個認知屬性，共有 100 個特徵。

一、SVM 模型參數選擇

為了找出 SVM 做題型難度分類時的最佳模型參數設定，實驗方法對 SVM 的兩個主要參數：正確率調控參數 C 與 RBF 核函數參數 γ ，透過如圖 1 所示的參數網格搜尋法進行參數選擇計算。其中， $C \in \{10^{-2}, 10^{-1}, \dots, 10^{10}\}$ 與 $\gamma \in \{10^{-9}, 10^{-8}, \dots, 10^3\}$ 被選為參數測試區間，透過 A 組中 102 筆資料做 5 摺交叉驗證（5-Fold Cross Validation, CV），以測試集平均正確率最高為最佳解。若只利用 SVM 求得的最佳參數為 $C=1000$ 與 $\gamma=0.001$ ，最佳平均正確率（與專家建立的難易度做比較）為 46.49%，為提升正確率有必要考慮降維模式。當以 PCA + SVM 做為訓練方法，在 PCA 降維至 $N=26$ 維時，可獲得測試集平均正確率最高為 55.81%，此時 SVM $C=1000000$ 與 RBF $\gamma=0.001$ 。而若以 LDA + SVM 做訓練，其中將 LDA 降至 $N=2$ 維，可獲得測試集平均正確率最高為 88.19%，此時 SVM $C=100$ 與 RBF $\gamma=0.0001$ 。因此，將以上述選定模型的參數做以下的實驗。

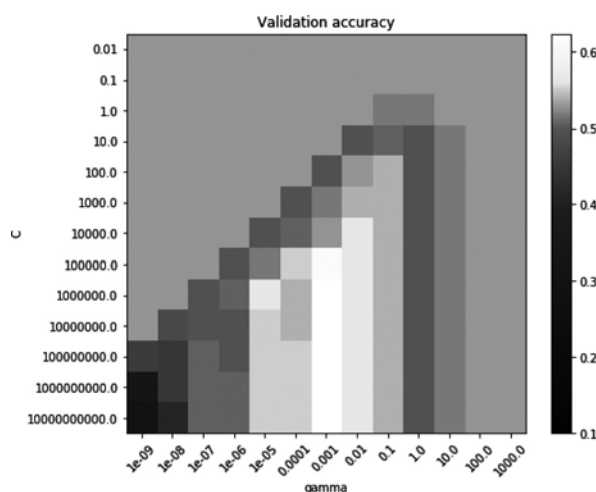


圖 1 SVM 模型參數選擇之交叉驗證正確率（Validation accuracies）圖

二、PCA + SVM 及 LDA + SVM 比較

首先將對 A 組 102 筆資料，84 個特徵做實驗，其中試題的易、中、難度分別為 25、54、23 題，比例約為 12.5%、52%、23.5%，隨機分成 5 摺，其中

林泓宏 蘇家鈺 張勝麟 以支援向量機處理題型符號與文字特徵應用於微積分試題難度分類

4 份作為訓練資料，1 份作為測試，並進行交叉驗證。在參數的選擇上，以各模型中的最佳值為實驗用參數，即 PCA + SVM 用降維至 $N=26$ ，SVM $C=1000000$ 與 RBF $\gamma=0.001$ ，而 LDA + SVM 用降維至 $N=2$ ，SVM $C=100$ 與 RBF $\gamma=0.001$ 。

為了要比較降維方法在試題難易度的優劣，本實驗分別做 PCA + SVM 及 LDA + SVM 的 5 摺交叉驗證法來比較。表 6 為利用上述 PCA + SVM 的最佳參數，執行 25 次 5 摺交叉驗證後，選取 Test 平均正確率最高的那組來呈現。其中，在訓練集及測試集的平均正確率分別為 53.92% 及 55.81%，都有達到 5 成以上，但與表 7 的 LDA + SVM 比較，發現表現都比較差。對 LDA + SVM 在訓練集及測試集的平均正確率分別為 89.70% 及 88.19%，圖 2 及圖 3 分別為表 7 中第四摺的訓練集及測試集分類圖。但若參數選用跟 PCA + SVM 相同，即 SVM $C=1000000$ 與 RBF $\gamma=0.001$ 時，訓練集及測試集的平均正確率分別為 91.40% 及 45.71%，測試集變得很差，可見參數的選擇對 LDA + SVM 很重要。另外，對 PCA + SVM 與 LDA + SVM 的最佳參數下，做 5 fold CV 隨機挑選，各執行 25 次，分別得到測試集的平均正確率為 51.42% 與 89.19%，而標準差為 0.046 與 0.045，用二母體平均數差的 t 檢定發現 p 值 ≈ 0 （雙尾），達顯著差異。表示整體來看，LDA + SVM 在測試集的正確率有比 PCA + SVM 顯著高。

表 6 PCA + SVM 5-Fold 交叉驗證法

PCA (N=26) + SVM，未加認知屬性；80% training set						
正確率	1	2	3	4	5	平均
Test	50%	50%	61.90%	60%	57.14%	55.81%
Training	54.88%	53.66%	54.32%	54.88%	51.85%	53.92%
PCA (N=26) + SVM，加入認知屬性；80% training set						
正確率	1	2	3	4	5	平均
Test	50%	61.90%	60%	52.38%	65%	57.87%
Training	56.10%	54.32%	59.76%	51.85%	52.44%	54.89%

註：此 5 摺交叉驗證 Test 平均正確率高於 Training 平均正確率，可能導因於這 5 摺資料隨機分組時，測試集題目相對較少、變異較大之故；整體 25 次「5 摺交叉驗證」的 Test 正確率平均為 51.42%，皆在 5 成以上。

測驗學刊，第 68 輯第 2 期

表 7 LDA + SVM 5-Fold 交叉驗證法

LDA (N=2) + SVM，未加認知屬性：80% training set						
正確率	1	2	3	4	5	平均
Test	85%	90.48%	85%	90.48%	90%	88.19%
Training	91.46%	83.95%	87.80%	95.06%	90.24%	89.70%
LDA (N=2) + SVM，加入認知屬性：80% training set						
正確率	1	2	3	4	5	平均
Test	90.48%	85%	90%	90.48%	95%	90.19%
Training	93.83%	92.68%	91.46%	92.59%	95.12%	93.14%

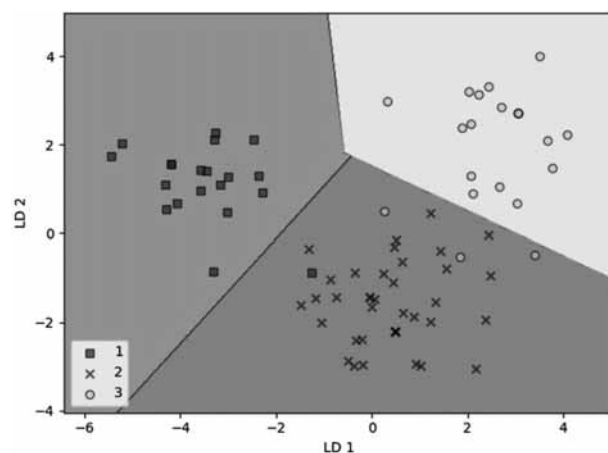


圖 2 表 7 中第四摺的訓練集分類圖

註：1、2、3 分別為試題易、中、難度。

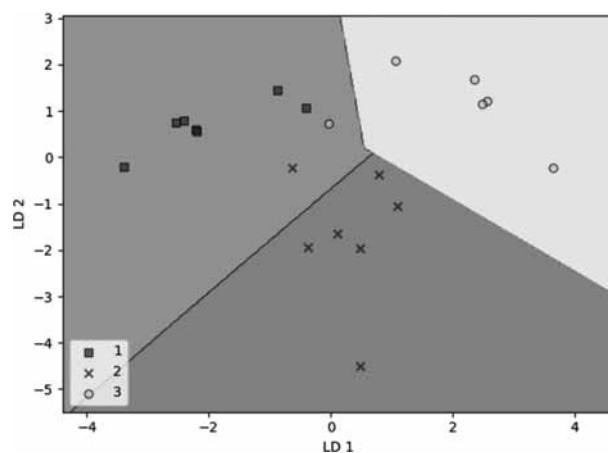


圖 3 表 7 中第四摺的測試集分類圖

註：1、2、3 分別為試題易、中、難度。

林泓宏 蘇家鈺 張勝麟 以支援向量機處理題型符號與文字特徵應用於微積分試題難度分類

三、比較加入認知屬性的特徵

在以 B 組 102 筆資料，加入認知屬性後共 100 個特徵做實驗，從表 6-7 後三格可看出，加入認知屬性後 PCA + SVM 及 LDA + SVM 的測試集及訓練集都有提升，但約只增加 1~3% 的正確率，所以認知屬性的特徵對難易度的判斷是有些微幫助。另外，在 PCA + SVM 的方法中，發現當降至 $N=32$ 時，同樣的參數可得測試集的平均正確率為 57.14%，但有某交叉驗證的正確率達 71.43%，預測錯誤率為 $ERR = 0.29$ ，對 PCA + SVM 算不錯的結果，其對應的測試集混淆矩陣（confusion matrix）如表 8 所示。

表 8 PCA + SVM 有認知屬性的混淆矩陣

		預測值		
測試題數		易	中	難
	易	1	2	0
實際值	中	1	12	0
	難	0	3	2

四、比較特徵有加入「需計算機」及「二層指數」

此處的特徵比較實驗，乃是針對 PCA + SVM 且有加入認知屬性特徵的部分做討論；在此實驗中，移除「需計算機」及「二層指數」二項特徵，測試結果如表 9 所示，其特徵維度 98 維，做 PCA 時降到 12 維。與表 6 的「PCA（ $N=26$ ）+ SVM，加入認知屬性」做比較，表 6 比表 9 的測試集增加 6.44% 的正確率，由此可看出加入這二項的特徵是需要的。

表 9 PCA + SVM 移除「需計算機」及「二層指數」特徵的 5-Fold 交叉驗證法

PCA ($N=12$) + SVM，加入認知屬性；80% training set						
正確率	1	2	3	4	5	平均
Test	42.86%	42.86%	47.62%	61.90%	61.90%	51.43%
Training	60.49%	55.56%	54.32%	51.85%	54.32%	55.31%

伍、結論

本研究提出對微積分試題的特徵結構分析，涵蓋符號與對應文字的關聯，透過「結構符號及對應文字」、「函數符號及對應文字」、「運算符號及對應文字」等三大類的特徵集設計，可將此特徵集應用在各類、一般化的微積分試題上，計算出適合做為試題難易度分類的特徵向量表示形式。

透過一般化「題型符號與文字特徵」，使用支援向量機之 **RBF** 核函數進行「難、中、易」試題分類，實驗結果顯示，**LDA + SVM** 在測試的正確率顯著高於 **PCA + SVM**。

從文獻探討可知，本研究所提出的「題型符號與文字特徵」表示形式，為國內外相關研究中創新的特徵集設計。在 5 摺交叉驗證測試下，對單一摺測試集之微積分試題難易度分類，透過 **LDA + SVM** 最高可獲取 95% 的正確率，而 5 摺的平均測試正確率也可達 90.19%，實驗測試結果遠高於隨機亂猜的 3.33%。若根據 Müller-Putz、Scherer、Brunner、Leeb 與 Pfurtscheller (2008) 文中對 3 個類別，以本文 102 題的平均每類別樣本數為 34 個計算，利用內插法得到亂猜的 95% 信賴區間上限約在 42.69%，可看出本研究方法的結果高於亂猜上限 47.5%。

進一步分析，在選到最佳參數下，**LDA + SVM** 的正確率都有達到 85% 以上，但 **PCA + SVM** 的平均正確率卻都低於六成，可能原因包含因樣本不多（僅 102 題）而特徵維度達 100 維，導致原始特徵空間中的題目點太分散，儘管採用了降低特徵維度的技術，但仍可能造成分類器過度擬合（over-fitting）的結果。

本研究日後改進的方向，除從增加測試集題數外，也可進一步探討自動化題型特徵集生成方法，如研究有回饋機制的深度學習方法來生成特徵集與建構分類器，以提高預測準度。另外，在當題型有圖形時，需專家給出對應的符號或文字做特徵編碼，這是本方法的限制，未來可朝圖形自動判讀方向發展。

謝誌

本論文由南臺科技大學之「教育部高教深耕子計畫：高階智慧生醫聯合

林泓宏 蘇家鈺 張勝麟 以支援向量機處理題型符號與文字特徵應用於微積分試題難度分類

研究中心」及科技部計畫（MOST 108-2511-H-218-003-MY2）補助，特此致謝。同時，感謝審查委員的細心指正與提供寶貴的修改建議。

參考文獻

中文部分

- 余民寧（2009）。*試題反應理論（IRT）及其應用*。臺北市：心理。
- 楊智為（2007）。以 SVM 結合多重貝氏網路在教育測驗上的應用（未出版之碩士論文）。國立臺中教育大學，臺中市。

英文部分

- Abdi, H., & Williams, L. J. (2010). Principal component analysis. *Wiley Interdisciplinary Reviews: Computational Statistics*, 2(4), 433-459.
- Alkharusi, H. (2012). Categorical variables in regression analysis: A comparison of dummy and effect coding. *International Journal of Education*, 4(2), 202-210. doi:10.5296/ije.v4i2.1962
- Bichi, A. A. (2016). Classical test theory: An introduction to linear modeling approach to test and item analysis. *International Journal for Social Studies*, 2(9), 27-33.
- Chang, S.-L., & Cheng, S.-C. (2017). Computer adaptive learning platform for calculus. In T.-T. Wu, R. Gennari, Y.-M. Huang, H. Xie, & Y. Cao (Eds.), *Emerging technologies for education* (pp. 153-162). New York, NY: Springer. doi:10.1007/978-3-319-52836-6_18
- Cheng, S.-C., & Chang, S.-L. (2016). *Assessment of the difficulty of items in computer adaptive learning platform for calculus*. Paper presented at 2016 International Symposium on Novel and Sustainable Technology, Tainan, Taiwan.
- Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (2013). *Applied multiple regression/correlation analysis for the behavioral sciences*. London, UK: Routledge.
- Dey, A., Chowdhury, S., & Ghosh, M. (2017). *Face recognition using ensemble support vector machine*. 2017 Third International Conference on Research in Computational Intelligence and Communication Networks (ICRCICN), Kolkata, India. doi:10.1109/ICRCICN.2017.8234479
- Dignath, C., & Büttner, G. (2008). Components of fostering self-regulated learning among students: A meta-analysis on intervention studies at primary and secondary school

林泓宏 蘇家鈺 張勝麟 以支援向量機處理題型符號與文字特徵應用於微積分試題難度分類

- level. *Metacognition and Learning*, 3(3), 231-264. doi:10.1007/s11409-008-9029-x
- Dignath, C., Büttner, G., & Langfeldt, H. P. (2008). How can primary school students learn self-regulated learning strategies most effectively? A meta-analysis on self-regulation training programmes. *Educational Research Review*, 3(2), 101-129. doi:10.1016/j.edurev.2008.02.003
- Fong, C. J., & Krause, J. M. (2014). Lost confidence and potential: A mixed methods study of underachieving college students' sources of self-efficacy. *Social Psychology of Education*, 17(2), 249-268. doi:10.1007/s11218-013-9239-1
- Foshee, C. M., Elliott, S. N., & Atkinson, R. K. (2016). Technology-enhanced learning in college mathematics remediation. *British Journal of Educational Technology*, 47(5), 893-905. doi:10.1111/bjet.12285
- Hattie, J., Biggs, J., & Purdie, N. (1996). Effects of learning skills interventions on student learning: A meta-analysis. *Review of Educational Research*, 66(2), 99-136.
- Johnson, B. G., Phillips, F., & Chase, L. G. (2009). An intelligent tutoring system for the accounting cycle: Enhancing textbook homework with artificial intelligence. *Journal of Accounting Education*, 27(1), 30-39. doi:10.1016/j.jaccedu.2009.05.001
- Kerr, P. (2016). Adaptive learning. *ELT Journal*, 70(1), 88-93. doi:10.1093/elt/ccv055
- Koedinger, K. R., & Aleven, V. (2007). Exploring the assistance dilemma in experiments with cognitive tutors. *Educational Psychology Review*, 19(3), 239-264. doi:10.1007/s10648-007-9049-0
- Lample, G., & Charton, F. (2020). *Deep learning for symbolic mathematics*. Paper presented at 2020 Eighth International Conference on Learning Representations, Addis Ababa, Ethiopia.
- Le, Y., Porwal, A., Hdden, E. J., & Dentith, M. (2012). Towards automatic lithological classification from remote sensing data using support vector machines. *Computers & Geosciences*, 45, 229-239. doi:10.1016/j.cageo.2011.11.019
- Leidinger, M., & Perels, F. (2012). Training self-regulated learning in the classroom: Development and evaluation of learning materials to train self-regulated learning during regular mathematics lessons at primary school. *Education Research International*, 1-14. doi:10.1155/2012/735790
- Lin, C.-C., Guo, K.-H., & Lin, Y.-C. (2016). A simple and effective remedial learning system with a fuzzy expert system. *Journal of Computer Assisted Learning*, 32(6), 647-662.
- Lord, F. M. (1980). *Applications of item response theory to practical testing problems*. Hillsdale, NJ: Lawrence Erlbaum Associates.

- Mok, M. M. C., Cheng, Y. C., Moore, P. J., & Kennedy, K. J. (2006). The development and validation of the Self-Learning Scales (SLS). *Journal of Applied Measurement*, 7(4), 418-449.
- Müller-Putz, G., Scherer, R., Brunner, C., Leeb, R., & Pfurtscheller, G. (2008). A closer look on BCI results. *International Journal of Bioelektromagnetism*, 10, 52-55.
- Pinar, R., & Oz, H. (2011). Validity and reliability of the Philadelphia Geriatric Center Morale Scale among Turkish elderly people. *Quality of Life Research*, 20, 9-18. doi: 10.1007/s11136-010-9723-4
- Putel, N. J., & Jlavier, R. H. (2015). Detecting packet dropping misbehaving nodes using support vector machine (SVM) in MANET. *International Journal of Computer Applications*, 122(4), 26-32. doi:10.5120/21689-4794
- Rao, R. C. (1948). The utilization of multiple measurements in problems of biological classification. *Journal of the Royal Statistical Society, Series B*, 10(2), 159-203.
- Rasch, G. (1960/1980). *Probabilistic models for some intelligence and attainment tests*. (Copenhagen, Danish Institute for Educational Research), expanded edition (1980) with foreword and afterword. Chicago, IL: The University of Chicago Press.
- Walkington, C. A. (2013). Using adaptive learning technologies to personalize instruction to student interests: The impact of relevant contexts on performance and learning outcomes. *Journal of Educational Psychology*, 105(4), 932-945. doi:10.1037/a0031882
- Xu, X., & Douglas, J. (2006). Computerized adaptive testing under nonparametric IRT models. *Psychometrika*, 71, 121-137.
- Yarnall, L., Means, B., & Wetzel, T. (2016). *Lessons learned from early implementations of adaptive courseware*. Menlo Park, CA: SRI International.

林泓宏 蘇家鈺 張勝麟 以支援向量機處理題型符號與文字特徵應用於微積分試題難度分類

附錄

單元 3-2 導數運算的認知操作（包含試題範例）

認知屬性	操作能力
1	會做 e 的極限操作
2	會做對數函數的導數
3	會做指數函數的導數
4	會做三角函數的極限操作
5	會做三角函數的導數
6	會做反三角函數的導數
7	會做雙曲與反雙曲函數的導數
8	了解高階導數及運算

試題 3-2-1. 若 $f(x) = x^2 |x|$ ，則下列何者錯誤

$$(1) f'(x) = \begin{cases} 3x^2, & x \geq 0 \\ -3x^2, & x < 0 \end{cases} \quad (2) f''(x) = \begin{cases} 6x, & x \geq 0 \\ -6x, & x < 0 \end{cases}$$

$$(3) f'''(x) = \begin{cases} 6, & x \geq 0 \\ -6, & x < 0 \end{cases} \quad (4) f^{(4)}(x) = 0, x \neq 0$$

答案：3 認知屬性：8 難易度：2

試題 3-2-3. 若 $y = \sin(x^3 + 2)$ 則 $y'(2)$ 之值為

$$(1) \sin 10 \quad (2) 12 \sin 10 \quad (3) -12 \cos 10 \quad (4) 12 \cos 10$$

答案：4 認知屬性：5 難易度：1

試題 3-2-9. 若 $x^3 - y^3 = 1$ ，則 y'' 之值為

$$(1) -\frac{2x}{y^5} \quad (2) \frac{2x}{y^5} \quad (3) -\frac{x^2}{y^2} \quad (4) \frac{x^2}{y^2}$$

答案：1 認知屬性：8 難易度：3

測驗學刊，第 68 輯第 2 期

試題 3-2-10. 若 $f(x) = e^{2x} \ln(x^3 + 1)$ ，則 $f'(1)$ 之值為

- (1) $(2 \ln 2 + 3)e^2$ (2) $\left(2 \ln 2 + \frac{3}{2}\right)e^2$ (3) e^2 (4) $\left(2 \ln 2 + \frac{1}{2}\right)e^2$

答案：2 認知屬性：2、3 難易度：2

試題 3-2-14. 若 $f(x) = \frac{\ln x}{x}$ 則 $f'(e) =$

- (1) 0 (2) 1 (3) $\frac{1}{e}$ (4) $\frac{1}{e^2}$

答案：1 認知屬性：2 難易度：1

試題 3-2-24. 若 $y = x^e + e^x + e^e$ ，則 $y'(1)$ 之值為何？

- (1) e (2) $2e$ (3) $1 + e + e^e$ (4) $1 + e$

答案：2 認知屬性：3 難易度：1

試題 3-2-27. 若 $y = x^{a^a}$, $a > 0$ 則 $\frac{dy}{dx} =$

- (1) $x^{a^a} \cdot \ln x$ (2) $x^{a^a} \cdot \ln a$ (3) $a^a x^{a^a - 1} \cdot \ln x$ (4) $a^a x^{a^a - 1}$

答案：4 認知屬性：3 難易度：3

試題 3-2-50. 若 $f(x) = 5^x + x^5$ ，則 $f'(1)$ 約為

- (1) 10 (2) 6 (3) 13.0472 (4) 6.6094

答案：3 認知屬性：3 難易度：2

試題 3-2-60. 若 $f(x) = 5^x + x^5$ ，則 $f'(x) =$

- (1) $x5^{x-1} + 5x^4$ (2) $5^x + 5x^4$ (3) $5^x \ln 5 + 5x^4$ (4) $5x^4$

答案：3 認知屬性：3 難易度：1

林泓宏 蘇家鈺 張勝麟 以支援向量機處理題型符號與文字特徵應用於微積分試題難度分類

試題 3-1-3. 若 $f(x) = (x-1)\sin(x^3+2)$ 則 $f'(1)$ 之值為

- (1) 0 (2) $\sin 3$ (3) $2\sin 3$ (4) 不存在

試題 3-1-7. 若 $(x^2 - y^2)^5 = x + 3y$ ，則在 $(1, 0)$ 的切線方程式為

- (1) $y = -3(x-1)$ (2) $x = 3(y-1)$ (3) $y = 3(x-1)$ (4) $x = -3(y-1)$

試題 3-1-12. 若 $f'(a) = 3$ ，則 $\lim_{h \rightarrow 0} \frac{f(a+2h) - f(a)}{h} =$

- (1) 3 (2) 6 (3) -6 (4) 不存在

試題 3-1-12-1. 若 $f'(a) = 3$ ，則

$$\lim_{h \rightarrow 0} \frac{f(a+2h) - f(a)}{h} = \underline{\hspace{2cm}}$$

試題 3-1-15. 若 $f(x) = (2x^3 - 1)^4$ ，則 $f'(x) =$

- (1) $4(2x^3 - 1)^3$ (2) $24(2x^3 - 1)^3$ (3) $4(6x^2)^3$ (4) $24x^2(2x^3 - 1)^3$

試題 3-1-15-1. 求 $f'(x)$ ，若 $f(x) = (2x^3 - 1)^4$ 。

- (1) $4(2x^3 - 1)^3$ (2) $24(2x^3 - 1)^3$ (3) $4(6x^2)^3$ (4) $24x^2(2x^3 - 1)^3$

試題 3-1-43. 求 $y = \sqrt[3]{6x+1}$ 的導函數為

- (1) $\frac{4}{6x+1}$ (2) $\frac{2}{\sqrt[3]{(6x+1)^2}}$ (3) $\frac{4}{\sqrt[3]{6x+1}}$ (4) $\frac{2}{\sqrt[3]{6x+1}}$

測驗學刊，第 68 輯第 2 期