



國立臺灣大學工學院工業工程學研究所

碩士論文

Institute of Industrial Engineering

College of Engineering

National Taiwan University

Master Thesis

運用線性迴歸預測處理器效能

Performance Prediction of Desktop Processors Using

Linear Regression

李侑澄

Yu-Chen Lee

指導教授：洪一薰博士

Advisor: I-Hsuan Hong, Ph.D.

中華民國 111 年 6 月

June 2022

論文口試委員審定書



國立臺灣大學碩士學位論文 口試委員會審定書

論文中文題目：運用線性迴歸預測處理器效能

論文英文題目：Performance Prediction of Desktop
Processors Using Linear Regression

本論文係李侑澄君（學號 P09546002）在國立臺灣大學工業工程學研究所完成之碩士學位論文，於民國 111 年 5 月 30 日承下列考試委員審查通過及口試及格，特此證明

口試委員：

洪一薰
(指導教授)

洪一薰

吳政鴻

吳政鴻

黃奎隆

黃奎隆

陳文智

陳文智

系主任、所長： 洪一薰

洪一薰

致謝

首先誠摯的感謝指導教授洪一薰博士，悉心的教導幫助我更瞭解線性迴歸的理論及分析的優點，不時的討論並指點我正確的方向，使我在這期間獲益匪淺。透過寫作論文的机会，來整理檢視以往在產業的實際應用和分析的流程，並與老師討論時發現了許多自己思考及理解不夠透徹的地方，經過和老師多次的討論、反饋、思辨及研讀後修正，提升了自己在理論及實務結合思考的完整度。學生在論文整理時不僅學習到了更完整的理論分析流程、分析方法背後的涵義及精確的定義，老師對學問的嚴謹、點出問題的精準度更是我輩學習的典範。

其次，在口試期間，承蒙口試委員吳政鴻、黃奎隆及陳文智教授詳細審閱本論文，提供了許多寶貴意見及建議，令學生深感獲益匪淺，使得本論文更加的完整，十分感謝老師們的提點。


最後感謝同組的鑫瓏、盛元及子豪同門師兄在過程中的互相鼓勵，且總能在我迷惘時為我解惑，恭喜我們順利走過這兩年，點點滴滴的協助讓我銘感在心。

研究生 李侑澄

謹誌於 國立臺灣大學工業工程學研究所

中華民國 111 年 6 月

摘要



本研究以桌上型主機板之處理器為主要探討標的，以研發人員的產業領域知識(Domain Knowledge)來找到影響處理器效能的關鍵參數，進行實驗室大量測試數據(Data)的整理，運用簡單的線性迴歸(Linear Regression)方法，讓主機板讀取到使用的處理器及散熱器特性後，來預測處理器在搭配的散熱器下最佳操作頻率及需求電壓，使主機板的處理器在作業系統中能以優化後的效能穩定運作，提供效能需求及內容創作的使用者能增加其使用的應用程式效能、提升使用者的工作效率。此為本研究探討的議題，除了提供效能需求的使用者更佳的體驗，也為產業數據趨勢、線性迴歸的應用和發揮提供一個學術上的案例參考。

關鍵字：線性迴歸、處理器效能、效能優化

Abstract



This research investigates the performance prediction model of desktop processors and identifies the key performance indicators affecting desktop processors performance.

We apply the linear regression to the prediction of desktop processor performance so that a motherboard is able to know the capability of desktop processors and coolers. The proposed method allows us to predict the optimized operating frequency and required voltage while the operating system achieves the optimal performance at the stable status of a motherboard. The results help content creation and performance-hungry users to have a better experience over daily applications and to increase their working efficiency.

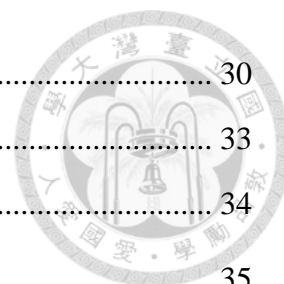
Keywords: linear regression, processor performance, performance optimization

目錄



論文口試委員審定書	i
致謝	ii
摘要	iii
Abstract.....	iv
目錄	v
圖目錄	vii
表目錄	viii
第一章 緒論	1
第二章 處理器效能優化影響參數	5
2.1 處理器頻率	5
2.2 處理器核心電壓	7
2.3 處理器散熱	9
第三章 線性迴歸模型建構與流程	11
3.1 線性迴歸演算法	11
3.2 線性迴歸模型與分析流程	15
3.2.1 問題描述及目標	15
3.2.2 最佳操作頻率預測求解模型	17
3.2.3 最佳核心電壓預測求解模型	18
3.2.4 線性迴歸方法分析流程	20
第四章 預測結果分析與驗證	21
4.1 迴歸預測結果分析	21
4.1.1 預測最佳操作頻率迴歸結果	21
4.1.2 預測最佳核心電壓迴歸結果	23
4.1.3 特徵值關係圖表解析	26
4.2 穩定性驗證預測結果	30

4.2.1 穩定性測試及結果.....	30
4.2.2 研究總結與模型限制.....	33
4.2.3 線性迴歸方法結果比較.....	34
第五章 結論與建議.....	35
參考文獻.....	36





圖目錄

圖 1 電壓調節模組簡易架構(改編自華碩主機板手冊，2022).....	7
圖 2 線性迴歸演算法流程圖	14
圖 3 主機板預測處理器與散熱器特性	15
圖 4 本研究預測處理器效能模型	16
圖 5 處理器核心電壓負載線(loadline).....	19
圖 6 運用線性迴歸方法分析流程	20
圖 7 預測最佳操作頻率相關係數矩陣	22
圖 8 預測最佳操作頻率殘差圖	23
圖 9 預測最佳核心電壓相關係數矩陣	24
圖 10 預測最佳核心電壓殘差圖	26
圖 11 迴歸模型驗證結果與殘差圖(residual plots)	29

表目錄



表 1 處理器超頻上限及餘裕(margin)	6
表 2 處理器核心電壓特性	8
表 3 處理器特定操作頻率及散熱器參數影響	9
表 4 處理器特定操作頻率及核心電壓參數影響	10
表 5 預測最佳操作頻率之相關輸入特徵值意義及說明	17
表 6 預測最佳核心電壓相關輸入特徵值意義及說明	18
表 7 預測最佳操作頻率之相關輸入特徵值訓練集	21
表 8 預測最佳操作頻率模型與結果	22
表 9 預測最佳核心電壓相關輸入特徵值訓練集	24
表 10 預測最佳核心電壓模型與結果	25
表 11 處理器核心數與超執行緒直方圖	26
表 12 處理器核心電壓與操作頻率直方圖	27
表 13 散熱器指數直方圖	27
表 14 處理器體質直方圖	28
表 15 處理器操作頻率與核心電壓趨勢圖	28
表 16 處理器頻率與核心數趨勢圖 vs.穩定輸出電壓.....	29
表 17 效能及穩定性測試平台	30
表 18 散熱器分數範圍(本研究自行測試定義).....	30
表 19 穩定性測試統計表	31
表 20 運用線性迴歸預測處理器效能結果比較	34

第一章 緒論

隨著機器學習的興起，各產業開始蓬勃的發展人工智慧(AI)的應用及企業之間的實作專案，紛紛開始研究如何利用現行已有的大數據來進行企業的數位轉型。而善用大數據已經不是口號，越來越多成功使用大數據的案例，如臉書、eBay、阿里巴巴等等，讓各行各業的企業負責人都開始意識到能否善用大數據，將是影響企業競爭力的重要關鍵之一，將企業轉型成數據驅動(Data-Driven)的企業文化，已是刻不容緩的事。所謂數據驅動，也就是要讓數據來說話，靠分析數據來幫助決策，而不是憑藉感覺來做決定。產業變動瞬息萬變，產生數據的數量(Volume)暴增，數據類型愈來愈多元(Variety)，而且一切講求速度與即時性(Velocity)，要如何才能成為數據驅動的企業，可不是件一蹴可幾的事。

因此在電腦系統的主機板產業也開始發現曙光，本研究在思考是否有機會將研發人員的產業領域知識(Domain knowledge)及數據整理，並僅運用簡單的線性迴歸方法將預測處理器效能演算法導入在主機板，主要希望提供給效能需求的主機板使用者能夠以簡易的方法，在基本輸入/輸出系統(BIOS)或作業系統(OS)界面下能夠選擇是否需要提升處理器的效能，來達到完全的將電腦的效能發揮到極致，特別針對一些需要 AI 運算、影像動畫效能提升。此外，在內容創作的使用者如在轉檔、壓縮等情境加速其應用程式的完成時間，提高其工作效率。綜合上述背景及原因，透過利用大數據及運用線性迴歸方法，為主機板產業及產品注入新的創意，為消費者提供更佳的使用體驗，此為本研究主要動機。

過去十年以來，線性迴歸雖然已被證明在許多領域都很有用，有多種的研究可以從這些學習中受益的技術，我們在這項研究中的主要動機是準確地預測應用未來的處理器性能趨勢，使用線性迴歸模型(Linear Regression)進行分析和訓練網絡，文獻中只有少數研究針對類似的主題，且處理器關鍵技術仍在主要晶片供應商，如 Intel(英特爾)及 AMD(超微半導體)，因此嘗試回顧歷史文獻，並考慮個人電腦(PC)

產業快速迭代的特性及硬體架構，因此本研究主要探討近年來是否有處理器參數分析、數據應用及利用線性迴歸方法進行預測效能的相關知識。

Inal and Küçük (2018)研究提出了用機器學習來分析處理器的特徵值(Features)為輸入(Input)，並運用線性迴歸來針對處理器單核心(Single Core)來預測效能為輸出(Output)，利用正規方程式(Normal Equation)來最佳化預測誤差，並達到70%的準確率是和本研究有高度相關的文獻，探討處理器特徵值對輸出效能的影響，並針對特徵值的數量、正規化(Normalization)來分析對輸出效能準確率的相關性，在實務應用上，處理器效能會因為使用的散熱器而有關鍵性的影響，但此文獻僅單純研究處理器單核心的特性，然而處理器在大部份實際應用及效能最大發揮是在多核心(multi-core)的狀況，雖然在議題研究探討上也有一定的貢獻及參考性，若導入實際應用仍恐有不足之處。

Su et al. (2021) 提到以 Linear regression 和 XGBoost 兩種模型來進行分析，嘗試來預測晶片的效能，以取代以往複雜又昂貴的功能測試方法，其晶片的效能好壞主要以頻率來定義。以 12nm 製程晶片為實驗目標，擷取晶片中會關鍵影響頻率，環形振盪器的相關特徵值做為機器學習的輸入參數，分別在 0.8M-cell 和 1.6M-cell 得到 93.2%及 83.3%的準確率，由實驗結論得到 Linear regression 和 XGBoost 兩種模型的準確率非常接近，且作者最後建議 Linear regression 在此晶片案例訓練時間也較有效率，是更為適合的應用，並此文獻中提到從 Intel 的文件中敘述到晶片有三個主要的參數變異：分別為 Process(製程)、Supply voltage(電壓)和 Temperature(溫度)，以頻率來定義效能好壞判斷，此部份都和本論文要研究分析的特徵值參數有很高的關聯性，但相比於本論文研究除了預測處理器操作頻率及核心電壓兩者組合的最佳狀況，其參數變數有：Voltage(電壓)、Frequency(頻率)、CPU Core(核心數)、Cooler score(散熱器等級)及 HT(執行緒)等，因此在參數的變數及複雜度更高。

Iyer and Pawar (2019) 研究指出，透過不同世代的 CPU 資料的收集、分析與整

理，並運用機器學習和梯度下降法(Gradient Descent)來收斂預測出 CPU 的價格，提供了一個很典型及基本的方法，演示出了資料的應用與機器學習的運用來幫助決策的範例，擷取處理器相關參數：核心數(Cores)、快取(Cache)及核心頻率(Clock Speed)，做為訓練模型的特徵值，說明資料整理的應用、選擇重要具有線性關係的特徵值。雖然此文獻和本研究探討的目標不同，但研究的關鍵零組件處理器和分析的方法都很有相關，也提高了本研究的信心、可信度，由於此文獻的特徵值並不複雜，應可考慮利用正規方程式(Normal Equation)取代梯度下降法來減少運算時的收斂時間。

李侑澄(2014)研究探討 PC 硬體玩家在購買決策之關鍵因素，從台灣的個人電腦(PC)硬體產業對遊戲產業的發展、市場動態及行銷活動逐步剖析，並分析了全球遊戲產業發展、遊戲玩家針對硬體要求的趨勢，使用層級程序分析法(Antalytical Hierarchy Process, AHP)來評估遊戲玩家對硬體功能需求的購買決策，結合相關的學術論文和產業研究報告，針對 146 份問卷來進行統計分析，得到整層級權重的優先順序。從此文獻的研究結論，瞭解到本研究提出的「處理器與散熱」相關議題是消費者在決策購買 PC 時是非常大的關鍵因素，呼應了本研究特別針對處理器(CPU)效能並同時考慮散熱相關功能的分析及探討，是值得在產業投入資源研究的議題。

回顧過去十年有關資料應用、線性迴歸運用於處理器效能的相關文獻，李侑澄(2014) 研究瞭解到「處理器」效能在「遊戲玩家 PC 硬體購買決策之關鍵因素」整層級權重排序中為第二名，是值得投入研究及分析的主題。在 Inal and Küçük (2018)對處理器效能分析有質性的研究，但僅只針對單核心處理器和有限的輸入特徵的討論，缺乏了加入散熱器及溫度對處理器有重要影響因素的交互關係探討。Iyer and Pawar (2019)由處理器的關鍵特性來預測其對應的價格趨勢，我們瞭解到有效的整理資料、找出有效的特徵值，並輔以線性迴歸、梯度下降法、決策樹(Decision Tree)等機器學習演算法足夠來分析並幫助決策的任務，提供我們很好的方法論參考，而

到了 Su et al. (2021)針對晶片有三個主要的參數變異：分別為 Process(製程)、Supply Voltage(電壓)和 Temperature(溫度)，並且以頻率來定義效能好壞判斷，此部份是和本論文要研究分析的特徵值參數有很高的關聯性，但對比本研究此文獻仍缺乏更多處理器因素、特徵及散熱器等複雜且關鍵的條件。

本研究針對處理器架構，分析和處理實驗室的測試資料，並考慮所有影響的特徵值及變異參數，運用線性迴歸及最小平方法推導正規方程式，提出一個方法預測出當處理器穩定運行下的最佳操作頻率及核心電壓需求，成功達到處理器的效能優化，能夠以最簡單的方式，即能充分發揮其電腦系統中最佳的使用體驗。第二章說明處理器效能優化的關鍵影響參數，解釋並界定這些參數對處理器效能的相關性。在第三章以線性迴歸演算法及文獻得到的相關應用知識，建立一個新的研究方法能夠讓主機板可預測處理器及散熱器特性，針對關鍵影響參數和相關特徵值分析來建構出有效預測處理器效能的模型，建立量化與質性並重之研究架構。第四章會分析運用了線性迴歸方法建構的預測模型、檢驗與衡量模型在預測處理器效能的結果，說明驗證方法以及實際帶來的改善效益，在歸納分析研究資料之後，分享研究成果，最後第五章為結論與建議。

第二章 處理器效能優化影響參數



此章節會先說明處理器效能的關鍵影響參數，從主機板的實務中，要使處理器的效能提升最直接的方法即為提高處理器的核心操作頻率，在 Su et al. (2021)的文獻探討已經敘述了這部份的相關性，然而在實際的操作中必須供應足夠的核心電壓才能讓處理器的頻率有對應的上升，在習知的處理器晶片物理特性，增加電壓即會帶來溫度的副作用要克服及處理，因此本研究所探討的處理器效能優化的定義為：「提升核心頻率並供應足夠且相對低的核心電壓，使處理器在運作的溫度規格內與作業系統下穩定運行，達到處理器頻率、電壓和溫度三個參數的最佳平衡。」因此瞭解了效能優化的定義後，後續章節逐一詳細說明三個重要參數對處理器效能的關係。

2.1 處理器頻率

處理器頻率是衡量處理器效能的一個非常重要的參數，它對處理器效能有很大的影響，頻率愈高、處理器的效能越好。主要原因為頻率愈高，處理器的時脈速度愈高、運算速度越快，可以處理的資料也就愈多。由於處理器運算和執行速度的提升，特別在一些效能需求或內容創作的使用者，有 AI 運算、影片及影音轉檔、專業繪圖，甚至對影像、執行速度高度要求的遊戲玩家都有很高的需求性，因此這種對於提升頻率、超出處理器供應商預設規格的行為又普遍稱之為「超頻」，甚至有許多的電腦組裝廠商出貨時就以超頻的條件提供給效能要求的消費需求使用者，來滿足這些情境的消費使用者工作及應用的需求，提升其工作的效率，也由於這樣龐大的需求及應用，超頻也成為了主機板廠或組裝製造商不斷研究的議題。

瞭解了超頻可帶來這麼多的效率及好處，有強大的需求和應用，主機板在處理器的超頻性能及功能自然就成為最重要投入資源研究的標的，然而要成功讓處理器超頻，並在作業系統下穩定運行也有一定的技術門檻及相關的專業知識，如下先說明部份在超頻時必要瞭解的知識和特性，也會在之後章節線性迴歸建立模型時加入更多的相關性特徵值來分析。



(1) 預設頻率(p0freq, p1freq):

此為處理器供應商在出廠時建議運行的預設頻率，單核心及全核心下的預設頻率我們分別定義為 p0freq 和 p1freq，若運行超出頻率即為超頻。研發人員必須先瞭解處理器在超出預設頻率至超頻上限尚有多少區間的可超頻餘裕稱為(margin)，由於處理器晶片在晶圓生產的物理特性，每顆處理器必然存在不同的餘裕及超頻上限，因此要找出某型號處理器的超頻餘裕，在以往主機板廠商必須投入大量的人力及測試，在開發階段有限的處理器數量進行測試，逐一記錄並進行處理器測試的統計資料，建立成通用規則(Rule-based)，但由於主機板並沒有能力或方法預測並知道不同處理器的餘裕，相對於在市面上銷售且龐大的處理器數量，在建立通用規則上，研發人員就必須考量最大的涵蓋率(Coverage Rate)，來確保提供的超頻功能能夠讓超頻的使用者不會遇到穩定性的問題，如表 1 為建立通用規則的實施實例，以某型號單核心預設頻率(p0freq)為 5.1GHz 的處理器，進行 6 顆相同型號處理器的測試結果，必須以餘裕最低的 CPU4 5.15GHz 為最後規則，來確保都能涵蓋到其它處理器穩定運行。所以要提高通用規則的涵蓋率，只能透過更多的處理器測試及數據。因此若我們有機會通過線性迴歸的方法預測處理器的能力，我們就可以依照不同的處理器預測出來的超頻上限及餘裕設定一個更彈性、更聰明的規則，也能省下研發人員大量的測試及時間。

表 1 處理器超頻上限及餘裕(margin)

編號	CPU1	CPU2	CPU3	CPU4	CPU5	CPU6
頻率	5.25GHzV	5.30GHz	5.20GHz	5.15GHz	5.28GHz	5.21GHz

(2) 核心數(Core):

處理器在設計架構上為了應付不同的使用情境、應用程式在處理上會產生不同的負載(loading)，運算量愈高、執行速度需求愈高的應用程式及情境就會造成處理器高負載的運行，愈高的負載即代表在電源供應上需要更多的消耗能量，意即在處理器上造成的溫度也會自然的上升。因此處理器的設計架構上，就會判斷不同的

應用程式及負載，開啟不同的實體核心數來執行運算，目的就是為了能夠聰明的判斷出應用程式及負載僅開啟適當的實體核心數，達到主機板系統的消耗能量、溫度上的最佳平衡。因此處理器在作業系統下能夠判斷應用程式的需求，在運算量低的應用僅開啟單核心來進行運算；在繪圖、影音運算量大或高畫質需求時即開啟全核心來處理。除了核心數量的判斷，處理器的設計架構也能夠非常聰明的可以在單核心及負載較輕的情境下提升處理器的頻率；多核心及負載重的情境下以較低的處理器頻率來運行，以某型號的處理器，在單核心的運行頻率可以為 5.1GHz，在全核心狀況下可運行在較低的 4.8GHz，來達到效能與消耗能量上的最佳平衡，因此核心數也會加入在線性迴歸模型的特徵參數中來觀測預測模型的影響。

2.2 處理器核心電壓

在前一章節已探討了以往習知的晶片或處理器特性，要提高處理器效能，從硬體的架構上一般即為提升其運作的頻率，隨著頻率的增加，主機板內的供電電壓調節模組(VRM)就需要對處理器的核心電壓(V_{CORE})如圖 1，有相應的上升才能有機會持續提升處理器頻率而得到效能的增加。然而在一般電學的基礎， $P=VI$ (P 為消耗功率， V 為電壓， I 為電流)，瞭解電壓的增加就會造成消耗功率的正比上升，相應就會在處理器上產生熱的問題，若無法即時解決熱造成的溫度問題，電腦系統就會發生穩定性如當機的問題發生，因此核心電壓的設定上對處理器穩定性是非常重要的關鍵，由於需要考慮在不同核心頻率及負載狀況下的電壓設定，因此是本研究最為最困難的地方，如下先說明部份必要瞭解的知識和參數特性。

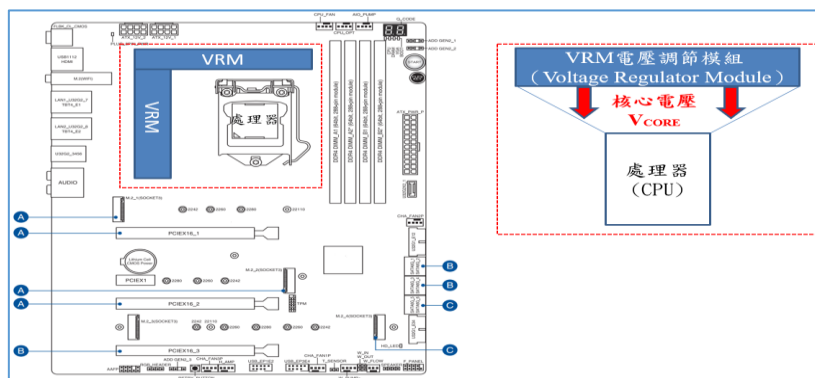


圖 1 電壓調節模組簡易架構(改編自華碩主機板手冊，2022)



以某型號 10 核心(10-Cores)處理器，分別在單核心(1-Core)頻率為 5.1GHz 及全核心(10-Cores)頻率為 4.8GHz 的運作狀況，並考慮在這兩種情況下模擬處理器在輕載和重載對於核心電壓的需求差異來說明。

(1) 核心電壓(p0vid):

處理器在作業系統下以單核心頻率條件讀取到的出廠預設核心電壓，如前章節所說明，由於此時為單核心運作且處理器的負載較輕，因此在此條件由電壓調節模組提供的電壓也不需要太高。

(2) 核心電壓(p0avxvid):

處理器在作業系統下以單核心頻率條件執行 AVX¹指令集讀取到的出廠預設核心電壓，利用執行運用 AVX 指令集的應用程式來模擬加重處理器負載時，記錄處理器此時穩定運行讀取到的核心電壓，由於在單核心頻率下運行，因此在此條件由電壓調節模組提供的電壓需求相對於 p0vid 要來的高。

表 2 處理器核心電壓特性

處理器參數	處理器頻率	處理器負載	處理器核心數
p0vid	高	輕	單
p0avxvid	高	重	單

因此從表 2 中依據不同處理器條件讀取到的核心電壓特性，來預測在特定的頻率下所需要設定的最佳核心電壓，這些核心電壓特性的特徵值除了用來預測判斷處理器的體質，也是預測頻率的重要相關參數。在預測處理器效能中，若這部份的數據資料愈多及有效，則也會對預測最佳核心電壓的精準度有程度上的影響。

¹Intel® 指令集擴充(AVX)是可以在多個資料物件上來運行相同操作時提升效能的指令。在執行 AVX 指令集時會使處理器的浮點運算方面算力增強，同時就會造成處理器負載功耗及熱量上升。



2.3 處理器散熱

處理器在一般運行即會散發出高熱能，在處理器上造成高溫，若無法壓抑住處理器的溫度，處理器就可能因為高溫而產生不穩定或所謂當機的情形，若處理器溫度持續上升就會啟動內部溫度的保護機制(throttling)，一般為 105°C ，若保護機制啟動就會強制讓處理器降速以最低的頻率及低壓來降低溫度。而在超頻的狀況下，由於搭配了核心電壓的上升，為了確保效能及處理器的穩定度，通常會需要在處理器上安裝不同型式的散熱器(cooler)，幫助處理器散去表面的高熱，因此處理器的超頻必須要考量散熱器的能力，不會因為散熱器的散熱能力不足而限制了處理器的超頻能力，如表 3 說明，處理器操作的工作溫度規格及溫度保護機制啟動為 105°C ，從系統三可看到由於使用水冷系統而有較好的散熱能力，所以對比系統二和風扇系統同樣的條件下，仍能正常在處理器的工作溫度規格下穩定運作。

表 3 處理器特定操作頻率及散熱器參數影響

超頻特定操作頻率	系統一	系統二	系統三
核心電壓 V_{CORE}	1.45V	1.47V	1.47V
散熱型式	風扇系統	風扇系統	水冷系統
處理器溫度	95°C	105°C	85°C

在以往的超頻方法由於主機板使用者並不會知道其安裝的散熱器型式及能力，所以除了要找出適合的核心電壓確保處理器消耗功率最低，也要有能力瞭解所使用的散熱器是否足夠壓制處理器的高溫，在實務上若技術能力及經驗不足常常會耗費許多的時間仍無法成功，並且由於市面上的散熱系統產品及型式包羅萬象，這也是為何以往製造商或工程研發人員對於要提升處理器相容性的涵蓋率存在有相當大的困難度，若能找到方法讓主機板除了預測處理器特性、也能預測散熱器能力，就有機會預測處理器的效能。

綜合上述章節，若我們無法找到方法得到這三個重要的參數，以往主機板廠商在效能調整的方式仍得依賴硬體及軟體兩種調教的方式，不管是硬體或軟體的實施方式，其邏輯背後仍依賴工程研發人員建立起來的通用規則來進行調整，由於無法瞭解當下使用的處理器體質特性及散熱器型式，因此在調整的程序不外乎即是通過不斷的試錯(Trial and Error)的調整方式，唯一的不同在於軟體可以透過軟體工程人員加入「自動」的程序，讓軟體偵測到處理器溫度過高或系統穩定性問題後，可重新啟動系統進程序上的調整，最後微調至效能提升又能在作業系統下穩定運行的狀態，然而不管是在硬體或軟體的實現方式，仍然存在無法得知使用者當下使用的處理器體質和可供穩定運行所需要的核心電壓，因此工程研發人員仍必須考量產品銷售後的相容性涵蓋率，如表 4 在核心電壓的設定上必須以最安全、涵蓋率最高的 CPU4 1.48V 為通用規則編寫在主機板硬體的 BIOS ROM(基本輸入輸出系統)或應用軟體的應用程序裡。

表 4 處理器特定操作頻率及核心電壓參數影響

編號	CPU1	CPU2	CPU3	CPU4	CPU5	CPU6
電壓	1.45V	1.47V	1.43V	1.48V	1.44V	1.45V

因此最後總結本章節所敘述，如果我們能找到一種方法：「可以讓主機板預測正在使用的處理器及散熱器特性」，我們是不是就有機會可以依照不同處理器特性、散熱器系統條件下，預測得到本章節探討的三個重要參數，依照當下處理器體質預測出最佳的操作頻率及核心電壓，達成本研究預測處理器效能目的。在線性迴歸演算法的成熟和應用上的推廣，俱備豐富領域專業知識和經驗的工程研發人員，是否可以找到更聰明、有效率的方式解決這些困難？在第三章會進行更多的說明。

第三章 線性迴歸模型建構與流程



回顧第二章節，要能有機會預測處理器效能的目標必須找到方法，讓主機板可預測處理器、體質特性與預測散熱器、瞭解散熱的能力。因此我們在第三章就是希望能找到一個適合的演算法，並建構合適的模型讓主機板可預測處理器和散熱器的特性。

3.1 線性迴歸演算法

參考以往文獻分析、整理和比較後，本研究問題也適用監督式學習的線性迴歸來建立模型，文獻中也探討了線性迴歸在預測處理器或晶片效能上有不錯的結果，因此如下說明本研究線性迴歸模型的分析流程及順序：

(1) 假設函數(hypothesis function):

在線性迴歸模型中，以 x 來表示輸入、以 y 來表示輸出，並需要一個 x 與 y 之間的關係式，假設 $y = f(x)$ 並稱此為目標函數(Target Function)，線性迴歸演算法是預先假設一個函數 $h(x)$ ，目的就是找到一個 $h(x)$ ，使其近似于或擬合 $f(x)$ 。而在不同的線性迴歸模型中，假設函數有不同表達形式，最簡單的一種就是單變量的線性迴歸模型，其形式為：

$$h_{\theta}(x) = \theta_0 + \theta_1 x$$

表示我們要預測一個關於 x 的線性函數 y 。

(2) 代價函數(cost function):

迴歸分析的目標函數或代價函數就是希望能夠從最終的預測模型資料和真實資料的誤差(error)或稱為殘差越小越好，來找參數 θ_0 和 θ_1 ，因此代價函數越小，就代表模型擬合的越好。在很多 x (因素)時，就可以簡化寫成下列方程式來代表(i 代表不同的 θ 和 x 的編號)

$$y = \sum \theta_i * x_i + \theta_0$$

理解了線性迴歸的概念後，我們再連結線性函數 $h_{\theta}(x)=\theta_0 + \theta_1 x$ ，只要解出函數中的 θ_0 和 θ_1 ，就能得到我們想要的假設函數，而要解出 θ_0 和 θ_1 ，就需要用到訓練數據集(training set)，因為機器要進行學習，就要有訓練數據集。在本文研究，輸入特徵值並不是單一的變量，因此會有一個以上的 independent variable (自變數)及一個 dependent variable(應變數)的線性關係，所以為多變數線性迴歸模型，我們瞭解其代價函數(cost function)如下。(其中 J 為代價函數，而 m 是樣本數量)

$$J(\theta_0 \dots \theta_n) = \frac{1}{2m} \left(h_{\theta}(x_i^{(i)}) - y^{(i)} \right)^2$$

(3) 正規方程式(Normal Equation)與梯度下降法(Gradient Descent)

在線性模式要找出代價函數最常見的方法有兩種，分別為由最小平方法推導正規方程式與梯度下降法來找尋讓代價函數達到最小值的最佳參數，然而在本文研究，輸入特徵值並不是單一的變量，因此會有一個以上的自變數及一個應變數線性關係，所以為多變數線性迴歸模型。(其中 $\theta_0 \dots \theta_n$ 為迴歸係數)

$$h_{\theta}(x) = \theta_0 + \theta_1 x_1 + \theta_2 x_2 + \dots + \theta_n x_n$$

梯度下降法和正規方程式都可以求得代價函數，在解線性迴歸時都是常見的方法，使用如下公式來求出 θ 值，若當是複雜的問題時梯度下降法較能解決。總結來說，如果自變數/特徵值數量不多，使用正規方程式可能是相對快速及適合的選擇，反之，則梯度下降法會較適當，Inal and Küçük (2018)。

$$\theta = (X^T X)^{-1} X^T y$$

X 代表特徵向量矩陣

X^T 為矩陣 X 的轉置矩陣

$(X^T X)^{-1}$ 為 X 的轉置矩陣和 X 矩陣相乘後的逆矩陣。

y 是訓練數據中的結果矩陣。

(4) 目標(Goal):

$$\text{minimize } J(\theta_0 \dots \theta_n)$$

目標使達到最小值的最佳參數(θ)，找到一個預測模型讓預測和真實的誤差(error)或殘差(residual)越小越好。並且使用正規方程式可一次算出參數，不需要如梯度下降法的方式需考量學習率及迭代運算，因此希望能以最簡單及快速而有效的方法來解決問題。

(5) 預測模型檢驗:

依照確定的模型建構流程，在預測模型檢驗採用最簡單的方法，以線性迴歸方法找出參數(θ)、建構出本研究的預測模型，並進行迴歸模型表現的衡量與檢驗，我們會觀測以下迴歸指標來進行檢驗:

R 平方(R squared)/ R^2

R 平方意義代表一種衡量迴歸模型表現的指標，從獨立變數 X 可以解釋依變數 Y 變異的比例如下式子，又稱為判定係數(coefficient of determination)。

$$R^2 = 1 - \frac{SS_{res}}{SS_{tot}}$$

其中 SS_{res} 為殘差平方和(residual sum of squares)，殘差平方和越大，表示模型解釋力越低，而 SS_{tot} 為總平方和(total sum of squares)，在衡量迴歸模型的意義上 R 平方愈接近 1 表示模型的預測率愈高。

建構完預測模型後，並進行模型的衡量與檢驗，確保模型的有效性，在下一章節即會說明本研究運用線性迴歸方法預測處理器效能的預測模型，並以衡量指標來檢驗模型的解釋能力及預測率，如圖 2 即為本研究線性迴歸演算法流程圖，也是預測模型的建構流程。

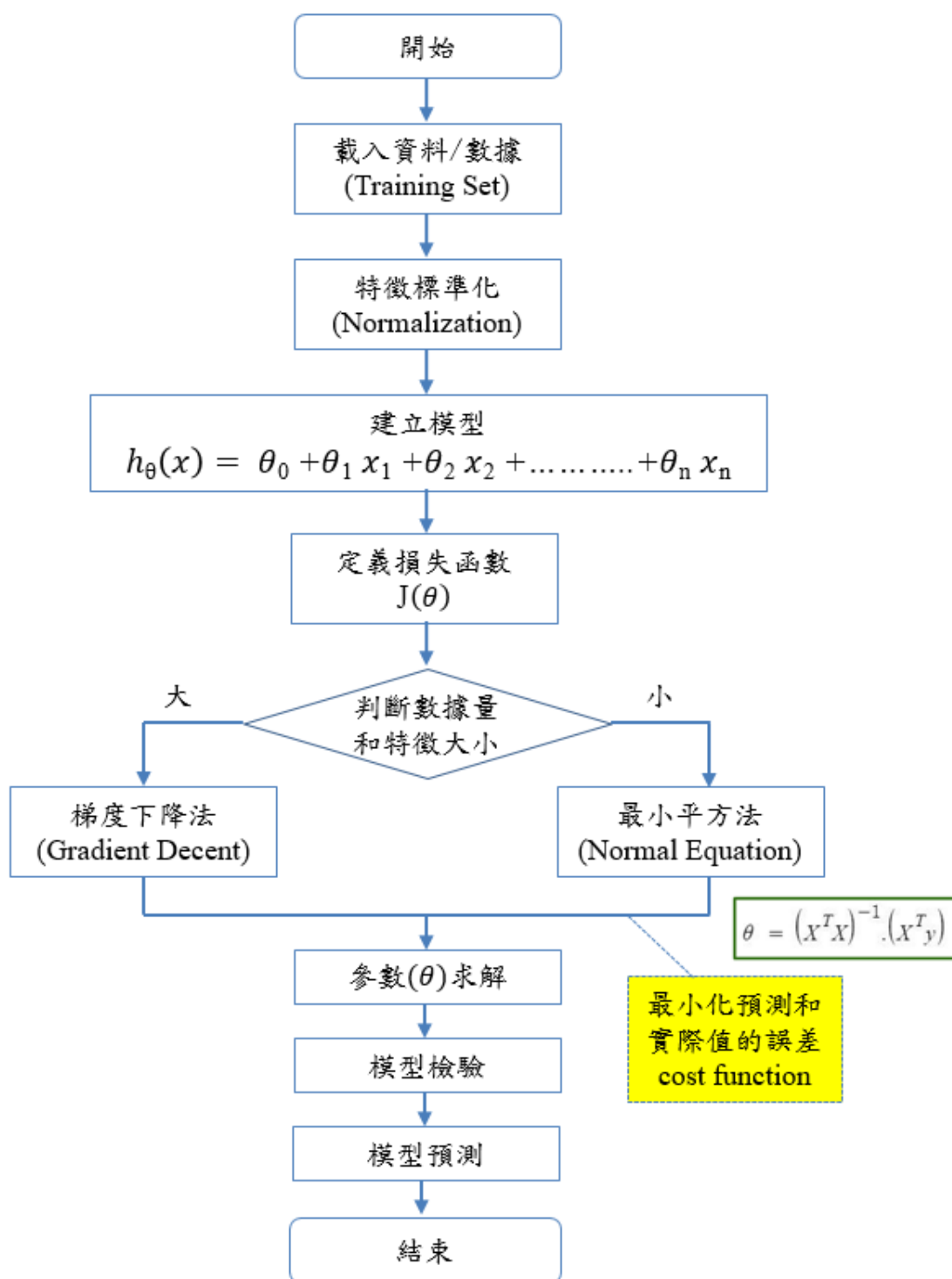


圖 2 線性迴歸演算法流程圖



3.2 線性迴歸模型與分析流程

在此章節我們的目標就是如何利用有效的線性迴歸演算法來讓主機板可預測處理器和散熱器特性，並從以往學者的文獻中已有相關的知識累積和經驗幫助本研究的問題探討，已知利用線性迴歸(Linear Regression)將實驗室即有的資料來進行整理及分析。

3.2.1 問題描述及目標

我們主要運用在實驗室量測到的所有處理器相關電壓、溫度及頻率及暫存器等資料就是我們的訓練資料以及測試資料，利用許多有標記處理器特性的資料庫，這些資料包含每一顆處理器內每顆核心的電壓、溫度、頻率及相關暫存器等相關資訊，找出有效的特徵集合當作輸入，透過線性迴歸的演算法處理資料，輸出則是處理器在相對應頻率所需要的電壓值，這是一個運用線性迴歸演算法做預測的應用，”讓主機板可預測處理器特性” 如圖 3。

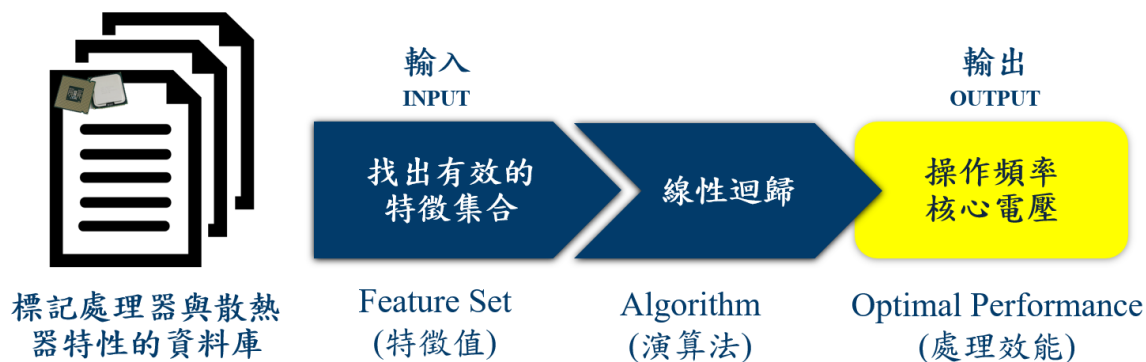


圖 3 主機板預測處理器與散熱器特性

若無法讓處理器得到好的散熱，就會限制處理器效能，也必須進行大量的散熱系統型式資料的統計和參數的整理，建立資料庫”讓主機板可預測散熱器特性”。因此我們必須整理一個關於處理器及散熱器特性的相關測試數據及資料庫，分析並歸納對我們預測處理器效能有用及相關的參數，這是建構良好預測模型最關鍵的任務，有了這樣的資料及數據後，我們接下來就可以來逐一解釋這些參數的意義及重要性，對預測處理器效能的相關性，並套用到線性迴歸演算法，來得到我們想要的預測結果。

圖 4 說明本研究預測處理器效能的模型，也藉助線性迴歸演算法(Learning Algorithm)，從訓練數據集(Training Set)中，來不斷改進預測/假設函數(hypothesis)的精準度，並且有了預測函數後，就可以將新的問題抽象為變量 x (通常稱為 feature)，代入預測函數，就可以得到預測的結果 y 。

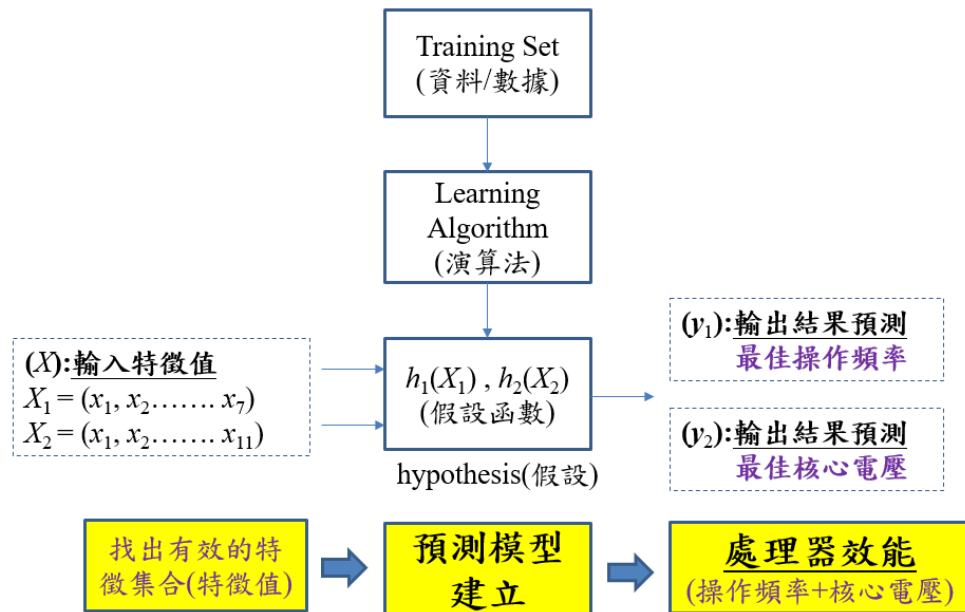


圖 4 本研究預測處理器效能模型

要達到處理器效能的優化，除了單純提高操作頻率，也必須考量頻率提升後其搭配的電壓能儘可能的低，來減少處理器功耗和溫度的影響。而處理器供應商出貨時都有其預設的電壓建議，但為確保涵蓋率通常在設定上也都是偏高及保守，本研究方法必須先將處理器在使用的散熱器和運作溫度規格條件下，將超頻後的最佳操作頻率預測出來之後，再針對不同的頻率設定來預測出處理器儘可能低且又能穩定運作的最佳核心電壓設定。因此為多目標的預測先找出 y_1 最佳操作頻率，再以此最佳頻率的狀況下預測出足夠穩定的 y_2 最佳核心電壓設定，雖然為兩個目標但仍存在密切的關連性。因此預測處理器的效能必要條件為：最佳操作頻率及能穩定運作的核心電壓。要找到這兩項必要條件，就必須分析及找出關鍵影響的參數、特徵值諸如處理器體質相關特性、散熱器能力，由研發人員的領域知識，如下會分別針對處理器和散熱器的特性來說明模型裡所挑選的參數及特徵值意義，以瞭解預測處理器最佳操作頻率和核心電壓的相關性。



3.2.2 最佳操作頻率預測求解模型

參照 2.1 及 2.2 章節關於處理器頻率及核心電壓參數的說明，表 5 為預測處理器最佳操作頻率相關使用的特徵值，其中核心數量(core)與超執行緒(HT)是處理器本身規格特性，不需要透過測試即可瞭解和頻率的高度相關性，愈少的核心數及超執行緒數量對處理器負載需求愈低，因此處理器能運行的頻率也會較高，因此本研究在預測處理器頻率時，還會觀測其它可能影響的參數，如處理器體質相關的出廠預設的頻率、核心電壓，甚至散熱器能力和分數對處理器頻率的相關性。

表 5 預測最佳操作頻率之相關輸入特徵值意義及說明

	特徵值	特徵值意義及說明
x_1	$p0freq/p0vid$	處理器體質(頻率與電壓關係式)
x_2	$p0freq/p0avxvid$	處理器體質(頻率與電壓關係式)
x_3	$p0freq$	單核心出廠預設頻率
x_4	$p1freq$	全核心出廠預設頻率
x_5	$cooler$	測試使用的散熱器能力及分數(散熱器分數)
x_6	$core$	處理器核心數量
x_7	HT	超執行緒 (HT, Hyper-Threading) 數量

並以上表特徵值建立預測處理器最佳操作頻率之訓練集(Training Set)，由實驗室建立的測試資料作為演算法的訓練集，使演算法從此資料中學習和建構、訓練模型如下式，並且能夠使用在未來資料上進行預測。

[最佳操作頻率的求解模型]:

$$h_{\theta}(x) = \theta_0 + \theta_1 * \frac{p0freq}{p0vid} + \theta_2 * \frac{p0freq}{p0avxvid} + \theta_3 * p0freq + \theta_4 * p1freq \\ + \theta_5 * cooler + \theta_6 * cores + \theta_7 * HT$$



3.2.3 最佳核心電壓預測求解模型

表 6 為預測處理器可以穩定運行的最佳核心電壓相關使用的特徵值，在第二章已瞭解，頻率(Frequency)的設定和核心電壓有最高的相關性必須加入關鍵特徵，也同時觀測其它處理器相關體質、特性和散熱器能力的影響性。

表 6 預測最佳核心電壓相關輸入特徵值意義及說明

	特徵值	特徵值意義及說明
x_1	<i>Frequency</i>	處理器測試設定頻率(依不同頻段)
x_2	<i>p0vid</i>	單核心頻率下讀取到的出廠預設核心電壓
x_3	<i>Freq.-p0freq</i>	測試設定頻率與單核心出廠頻率比值差異
x_4	<i>p0freq/p0vid</i>	處理器體質(頻率與電壓關係式)
x_5	<i>p0freq/p0avxvid</i>	處理器體質(頻率與電壓關係式)
x_6	<i>p0freq</i>	單核心出廠預設頻率
x_7	<i>p1freq</i>	全核心出廠預設頻率
x_8	<i>cooler</i>	測試時所使用的散熱器能力及分數(散熱器分數)
x_9	<i>cores</i>	處理器核心數量
x_{10}	<i>HT</i>	超執行緒(HT, Hyper-Threading)數量
x_{11}	<i>loadline</i>	處理器核心電壓負載線(loadline)設定

並以上表特徵值建立預測處理器最佳核心電壓之訓練集(Training Set)，由實驗室建立的測試資料作為演算法的訓練集，使演算法從此資料中學習和建構、訓練模型如下式，並且能夠使用在未來資料上進行預測。

[最佳核心電壓的求解模型]:

$$\begin{aligned}
 h_{\theta}(x) = & \theta_0 + \theta_1 * Frequency + \theta_2 * p0vid + \theta_3 * Freq.-p0freq \\
 & + \theta_4 * \frac{p0freq}{p0vid} + \theta_5 * \frac{p0freq}{p0avxvid} + \theta_6 * p0freq + \theta_7 * p1freq \\
 & + \theta_8 * cooler + \theta_9 * cores + \theta_{10} * HT + \theta_{11} * loadline
 \end{aligned}$$



(1) 處理器體質(SP, Silicon Parameter):

從表 6 參數中產生以下關係式: $\left(\frac{p0freq}{p0vid}\right)$ 或 $\left(\frac{p0freq}{p0avxvid}\right)$

從關係式瞭解，在處理器預設的出廠頻率下，若讀取到所需運作核心電壓愈低，我們可以將此判定為體質相對優異的處理器，並所造成的熱能也會較低，本研究會利用這個相關性來研判處理器的體質使用。

(2) 散熱器分數(cooler):

當使用的散熱器型式的能力愈好，代表能帶走的熱能愈多，處理器的溫度也就愈低，相對於處理器的穩定運行有關鍵的影響，因此工程研發人員會測試不同的處理器及搭配各種型式的散熱器，透過記錄處理器的核心數目(cores)、執行緒(HT)、溫度、核心電壓(Vcore)和消耗的功率(瓦數)來設定散熱器分數，歸類出不同的散熱器分數(Score)和能力，以便做為預測處理器最佳操作頻率的重要參數。

(3) 超執行緒技術(HT):

使每個核心執行多執行緒，能夠平行處理的工作愈多，處理器會向每個實體核心公開兩個執行內容，猶如兩個「邏輯核心」，可以處理不同的軟體執行緒，相較於單一執行緒核心處理工作的效率更高，改善了過去核心必須等待其他工作完成的閒置時間及處理器的輸送量。

(4) 核心電壓負載線(loadline):

透過調整核心電壓範圍來控制處理器負載線設定，選擇較高的電壓(負載線較為平緩)獲得更佳的系統效能，選擇較低的電壓(負載較陡峭)獲得更佳的電源效率。設定值依主機板及電壓調變模組而有不同的調整等級，負載線設定越高則效能越好，提高了電壓準位值與超頻能力，但副作用則較為耗電如圖 5。

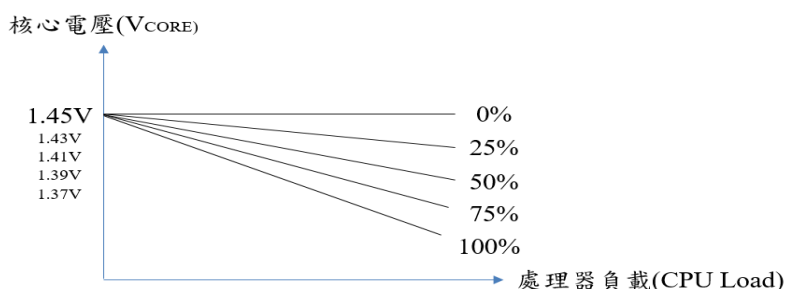


圖 5 處理器核心電壓負載線(loadline)



3.2.4 線性迴歸方法分析流程

所以我們從概念及研究動機發起，將以往領域知識收集到的數據加以統計整理及分析，從中觀察探索與擷取可以當作演算法重要及有效的特徵值，依據本章節選定的有效演算法及分析程序建立模型，然後將模型導入驗證機台進行實作及模型的驗證與分析，最後搭配實際機台測試的統計資料及驗證結果匯整，以比對模型預測的狀況，若預測結果導入機台結果不理想，則返回檢視模型的微調，直至最終結果符合預期的目標及準確度，至此確認本研究整個分析流程如圖 6。

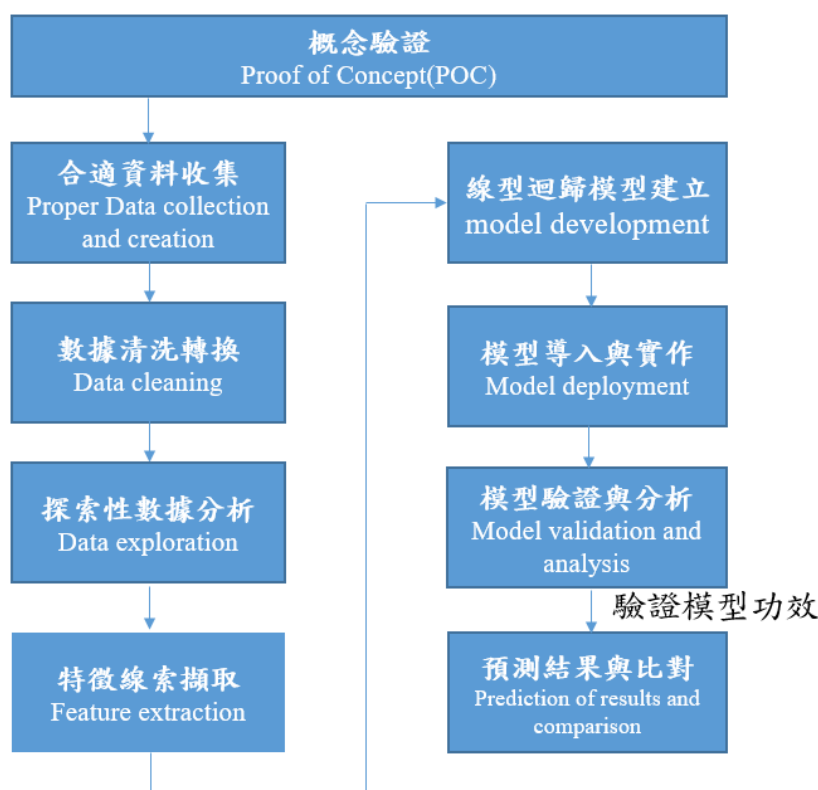


圖 6 運用線性迴歸方法分析流程

要能有效預測得到處理器效能，其中最困難的部分是如何透過工程研發人員的領域知識，從收集的測試資料找出有效的特徵集合，建構一套適合的預測模型，希望透過這樣的架構我們來建立讓主機板可預測處理器和散熱器特性的機制，能夠精準預測並建議目前使用的電腦系統能夠穩定運作的最佳效能設定，幫助所需要的工作效率和任務，因此在下一章節即會進到預測結果及驗證階段的方法、流程說明，並利用本章節說明的迴歸模型衡量與檢驗指標，來解譯本研究工作流程及運用線性迴歸方法是否達成本研究預期的目標和結果。

第四章 預測結果分析與驗證



我們在第三章節確定了本研究的演算法架構及分析程序，本章節主要來進行預測結果的分析說明與驗證。

4.1 迴歸預測結果分析

此章節會分別針對最佳操作頻率和核心電壓的部份，從數據集開始整理出相關係數矩陣來觀測參數的相關性，並以線性迴歸方式建構預測模型，利用 R 平方來判斷、衡量迴歸模型的預測率，最後以殘差圖、直方圖及迴歸線圖來輔助說明本研究各重要參數的結果和意義。

4.1.1 預測最佳操作頻率迴歸結果

如表 7 為示意部份預測處理器最佳頻率之訓練集(Training Set)，以線性迴歸演算法從此資料中學習和建構、訓練模型並進行預測。

表 7 預測最佳操作頻率之相關輸入特徵值訓練集

操作頻率	p0freq/p0vid	p0freq/p0avxvid	p0freq	p1freq	cooler	cores	HT
4550	3.32103321	3.32103321	4500	4400	132	4	2
4800	3.32103321	3.32103321	4500	4400	154.5	4	2
4700	3.372434018	3.299856528	4600	4300	161	6	1
4700	3.372434018	3.299856528	4600	4300	161	6	1
4650	3.385826772	3.32046332	4300	4100	163	6	1
4700	3.422619048	3.372434018	4600	4300	164	6	1
4200	3.433476395	3.375527426	4000	4000	95.7	4	1
4500	3.433476395	3.375527426	4000	4000	132	4	1
4600	3.433476395	3.375527426	4000	4000	151	4	1
4200	3.467741935	3.37254902	4300	4100	95.7	6	1

- (1) 相關係數矩陣(Correlation Matrix): 首先如圖 7 相關係數矩陣列出本研究在預測最佳操作頻率及核心電壓的所有的特徵值其互相影響的相關性，衡量變量之間相關程度的係數，相關係數若越接近於±1，兩個係數之間的相關性越強。在預測最佳操作頻率的特徵參數影響性分析上，除了處理器核心(cores)及超執行緒數量(HT)為架構規格上顯而易見的特性，若要能夠提高預測準確度，可以看到處理器體質相關(p0freq/p0vid、p0freq/p0avxvid)為 0.34 及 0.35、出廠預設頻率

(p0freq、p1freq)為 0.47 及 0.32，散熱器分數(cooler)為 0.42，都是不可忽視的重要的相關性參數。

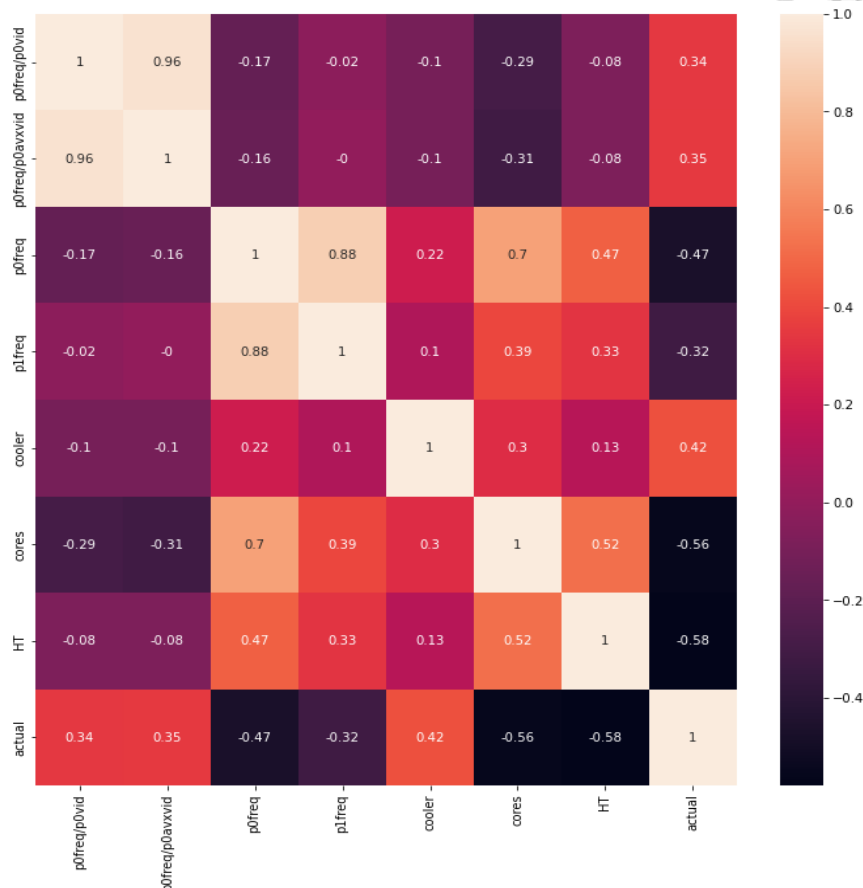


圖 7 預測最佳操作頻率相關係數矩陣

(2) 預測模型與結果: 依第三章確定的模型建構流程，我們建構預測最佳操作頻率的模型如表 8，並得到 R 平方為 0.84，參考 3.1 章節關於迴歸模型衡量與檢驗指標，此迴歸模型對預測結果的可解釋程度約為 84%，模型擬合度尚可接受。

表 8 預測最佳操作頻率模型與結果

迴歸結果	
R 的倍數	0.92
R 平方	0.84
調整的 R 平方	0.84
標準誤	95.38
觀察值個數	1736



預測最佳操作頻率模型		
截距	2594.07	θ_0
$p0freq/p0vid$	14.41	θ_1
$p0freq/p0avxvid$	498.26	θ_2
$p0freq$	-0.05	θ_3
$p1freq$	-0.03	θ_4
$cooler$	8.57	θ_5
$cores$	-45.64	θ_6
HT	-191.42	θ_7

(3) 殘差圖分析: 如圖 8 為預測最佳操作頻率的殘差圖，觀測都還很平均以 0 為橫軸的直線上下隨機散佈，因此迴歸直線對各個觀測值的擬合情況也符合預期。

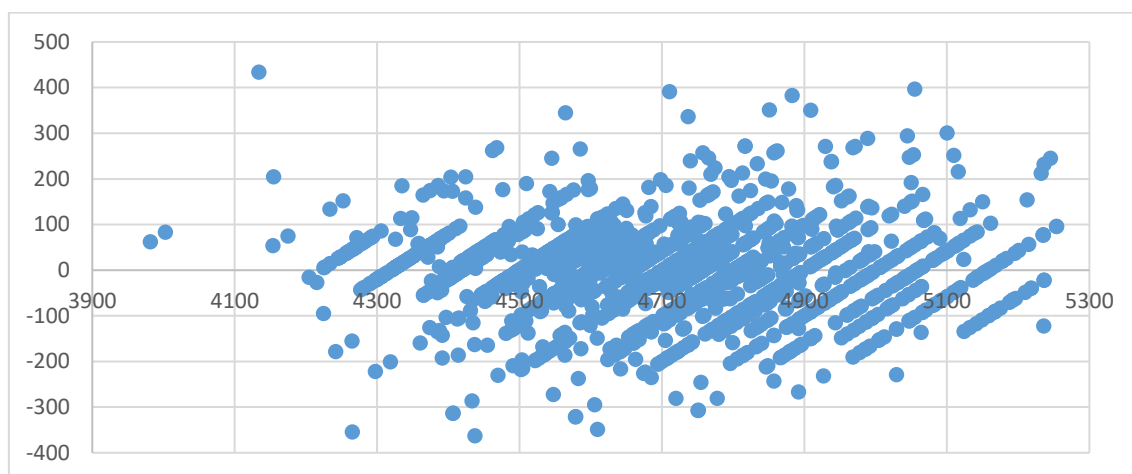


圖 8 預測最佳操作頻率殘差圖

4.1.2 預測最佳核心電壓迴歸結果

預測最佳核心電壓的模型，在工程研發人員及製造商的領域知識中，要預測出特定頻率下能穩定運行的最佳核心電壓，是預測處理器效能最困難與關鍵的部份，如第二章所討論，核定電壓的設定需要有足夠、有效的測試資料與數據，才能建立有把握及高涵蓋率的核心電壓設定基礎。如表 9 部份預測處理器最佳核心電壓之訓練集(Training Set)，以線性迴歸演算法從此資料中學習和建構、訓練模型並進行預測。

表 9 預測最佳核心電壓相關輸入特徵值訓練集

核心電壓	Frequency	p0vid	Freq-p0freq	p0freq/p0vid	p0freq/p0avxvid	p0freq	p1freq	cooler	cores	HT	loadline
1476.35	5100	1528	-200	3.46859	3.41275	5300	4900	134	10	2	4
1082.24	4300	1528	-1000	3.46859	3.41275	5300	4900	176	10	2	4
1107.24	4400	1528	-900	3.46859	3.41275	5300	4900	176	10	2	4
1162.24	4500	1528	-800	3.46859	3.41275	5300	4900	168	10	2	4
1207.24	4600	1528	-700	3.46859	3.41275	5300	4900	170	10	2	4
1232.24	4700	1528	-600	3.46859	3.41275	5300	4900	171	10	2	4
1267.24	4800	1528	-500	3.46859	3.41275	5300	4900	174	10	2	4
1322.24	4900	1528	-400	3.46859	3.41275	5300	4900	174	10	2	4
1377.24	5000	1528	-300	3.46859	3.41275	5300	4900	175	10	2	4
1452.24	5100	1528	-200	3.46859	3.41275	5300	4900	174	10	2	4

(1) 相關係數矩陣(Correlation Matrix): 如圖 9 在最佳核心電壓的特徵參數影響性分析上, 頻率(Frequency)為 0.89 證明了其必然的重要相關性, 而要提高預測準確度, 處理器核心電壓負載線(loadline)為 0.23 似乎也是需要考量的參數。

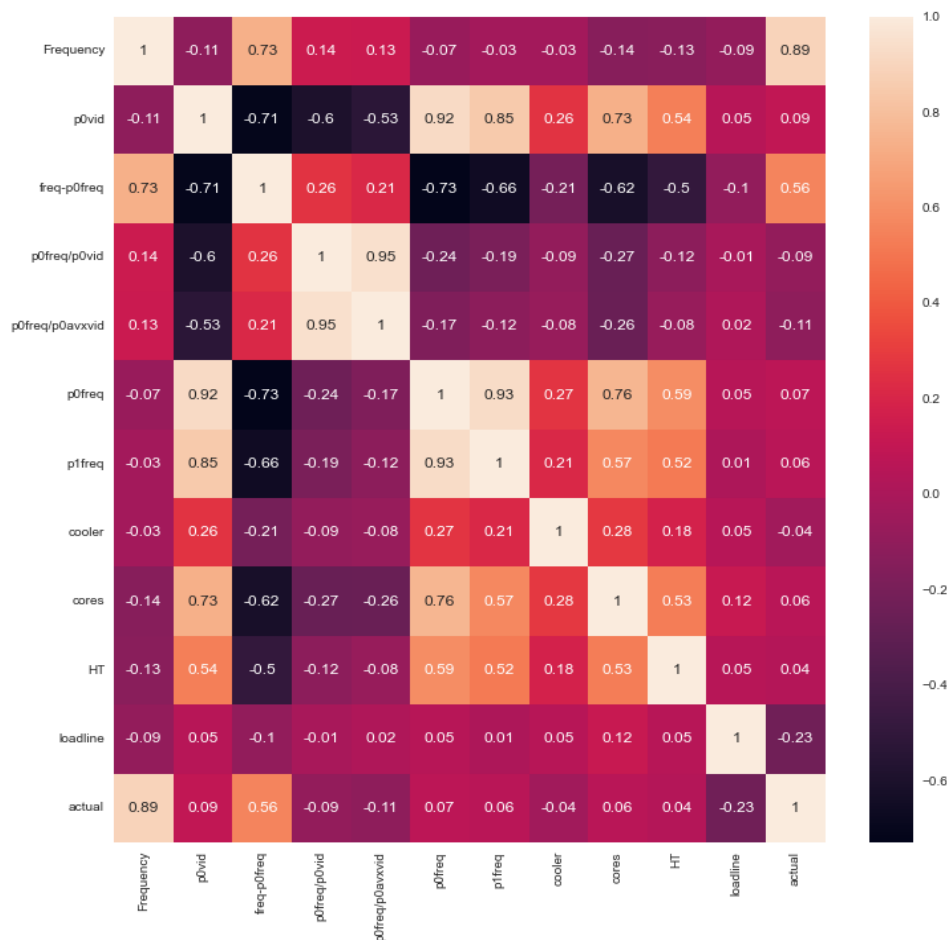


圖 9 預測最佳核心電壓相關係數矩陣

- (2) 預測模型與結果: 本研究收集了 37442 筆的測試數據, 建構出預測模型如表 10, 並得到 R 平方為 0.92, 此回歸模型對預測結果的可解釋程度約為 92%, 模型擬合度非常良好。

表 10 預測最佳核心電壓模型與結果

迴歸結果	
R 的倍數	0.96
R 平方	0.92
調整的 R 平方	0.92
標準誤	41.5
觀察值個數	37442

預測最佳核心電壓模型		
截距	6470.87	θ_0
<i>Frequency</i>	0.61	θ_1
<i>p0vid</i>	-4.82	θ_2
<i>Freq-p0freq</i>	-0.19	θ_3
<i>p0freq/p0vid</i>	-1766.56	θ_4
<i>p0freq/p0avxvid</i>	-140.81	θ_5
<i>p0freq</i>	1.02	θ_6
<i>p1freq</i>	0.04	θ_7
<i>cooler</i>	-0.48	θ_8
<i>cores</i>	13.50	θ_9
<i>HT</i>	36.70	θ_{10}
<i>loadline</i>	-21.50	θ_{11}

- (3) 殘差圖分析: 圖 10 為預測最佳核心電壓的殘差圖, 觀測電壓的分佈平均以 0 為橫軸的直線上下隨機散佈, 因此迴歸直線對各個觀測值的擬合情況也符合預期。

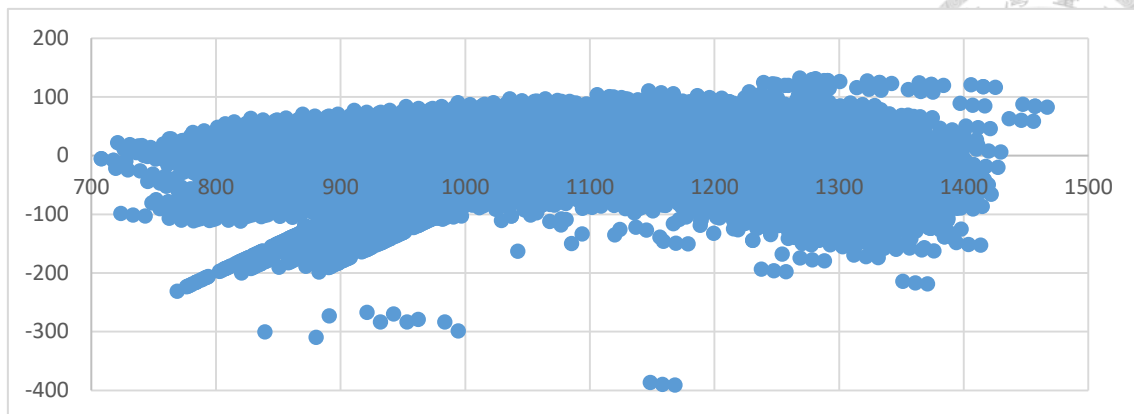


圖 10 預測最佳核心電壓殘差圖

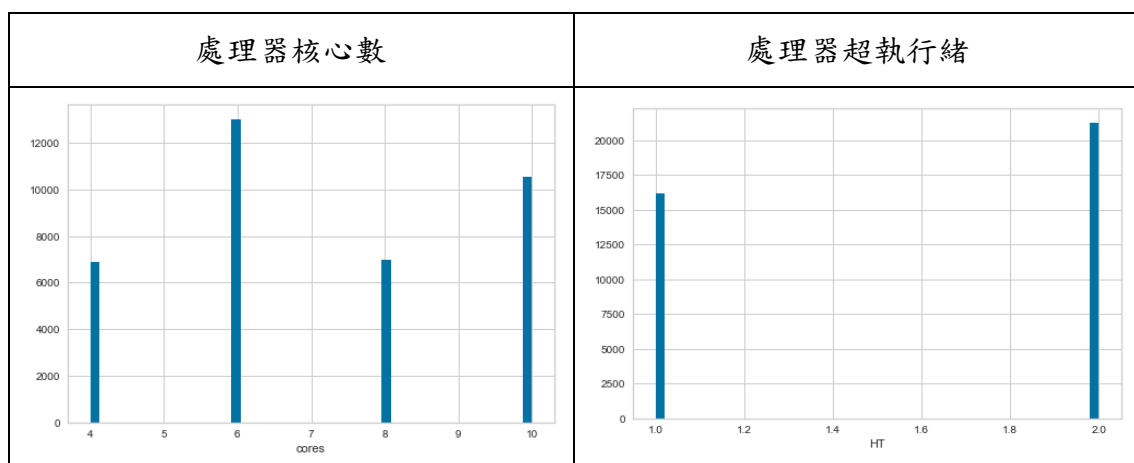
4.1.3 特徵值關係圖表解析

由於核心電壓為預測處理器效能的最重要關鍵因子，所以本研究在進行核心電壓測試數據的建立非常嚴謹，後續會特別針對此部份的相關特徵選擇及特性說明，仍會依照第三章確定的模型建構流程，以 37442 筆的核心電壓測試資料來進行訓練，針對目前用來預測最佳核心電壓的 11 個特徵值(features)來做成直方圖及特徵關係的圖表，並隨後針對部份圖表做相關補充說明。

(1)「處理器核心數與超執行緒直方圖」對應表：

表 11 說明本研究處理器型號的核心數與超執行緒在 37442 筆中資料數量的分佈，分析的資料庫包含了 4、6、8 與 10 核心的處理器，在目前處理器架構的超執行緒數量，依不同規格的處理器而有 1 個或 2 個數目的方式來分配處理應用程式。

表 11 處理器核心數與超執行緒直方圖

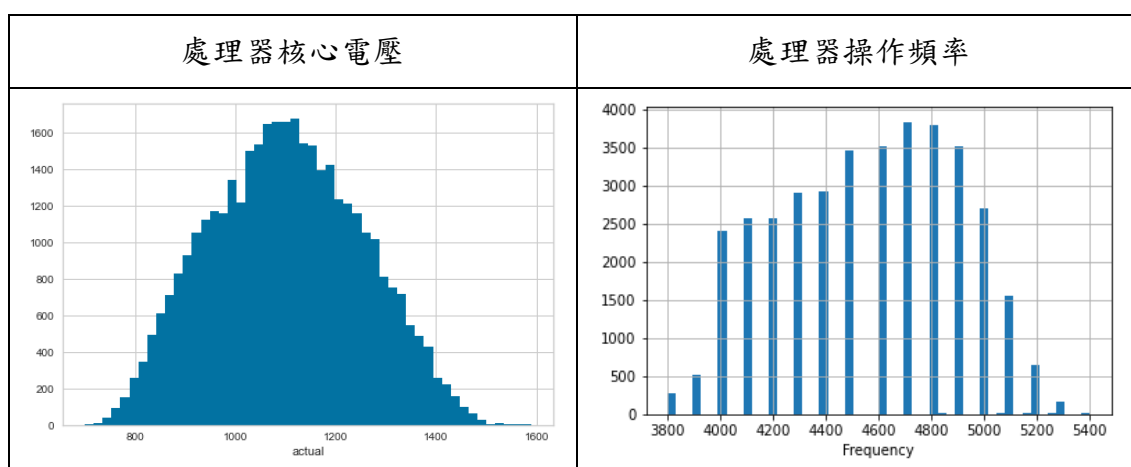




(2)「處理器核心電壓與操作頻率直方圖」對應表：

表 12 說明在不同處理器型號核心電壓與操作頻率的資料分佈，資料呈現高斯分佈(Gaussian distribution)，表示不同的處理器型號、頻率，能穩定運作的核心電壓也會不同，也呈現出分析的處理器資料中以 4500MHz~4800MHz 頻率的處理器為最多數，而能穩定運作的電壓範圍為 1.0V ~ 1.3V 為主，當然以處理器特性若要得到較高的效能及頻率，電壓仍必須隨之正比上升才能使系統穩定運作。

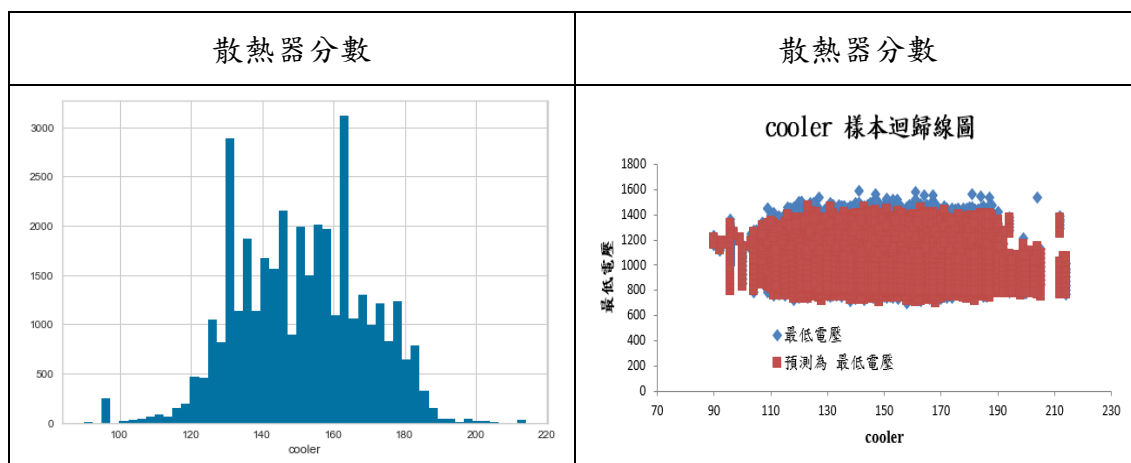
表 12 處理器核心電壓與操作頻率直方圖



(3)「散熱器指數直方圖」對應表：

表 13 說明本研究資料中處理器散熱器特性及效能的分佈，並加入由迴歸線圖解釋實際和預測的狀況，資料中包含不同型式的風扇及水冷散熱系統，在本研究中定義 170 分以上的散熱系統有非常優異的散熱能力，如下散熱器分數純粹為本研究定義出來可供方便分類及最終效能的判斷參考依據。

表 13 散熱器指數直方圖

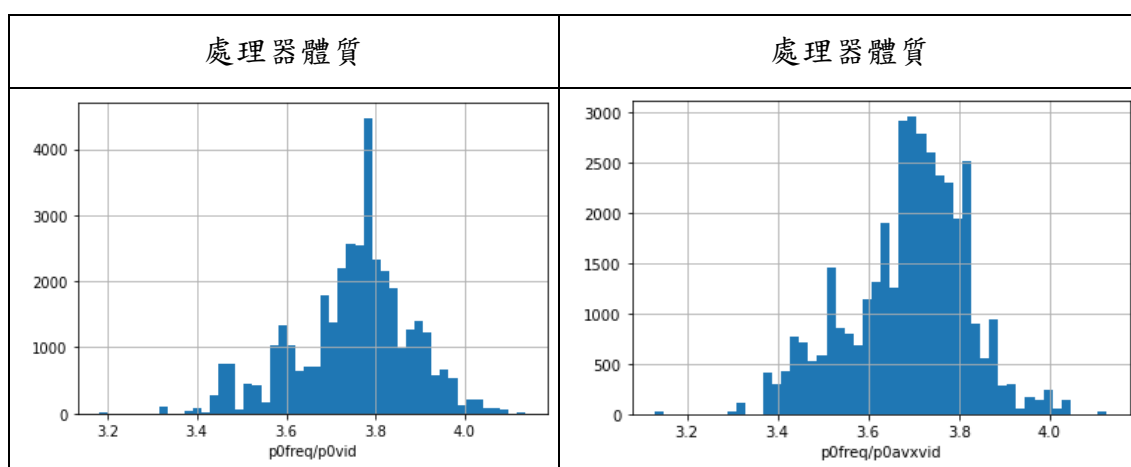




(4)「處理器體質直方圖」對應表：

在本研究的目標中，若能用線性迴歸的流程建構快速判斷處理器效能及體質優劣的方法，除了可以幫助研發工程人員快速判斷篩選，也可供後續演算法預測出處理器效能的關鍵因素。如表 14 中數值 3.7~3.8 意義體質的處理器為多數，如同晶圓特性的優異呈現高斯分佈，愈往左代表處理器體質較差、愈往右則此處理器體質愈優異。

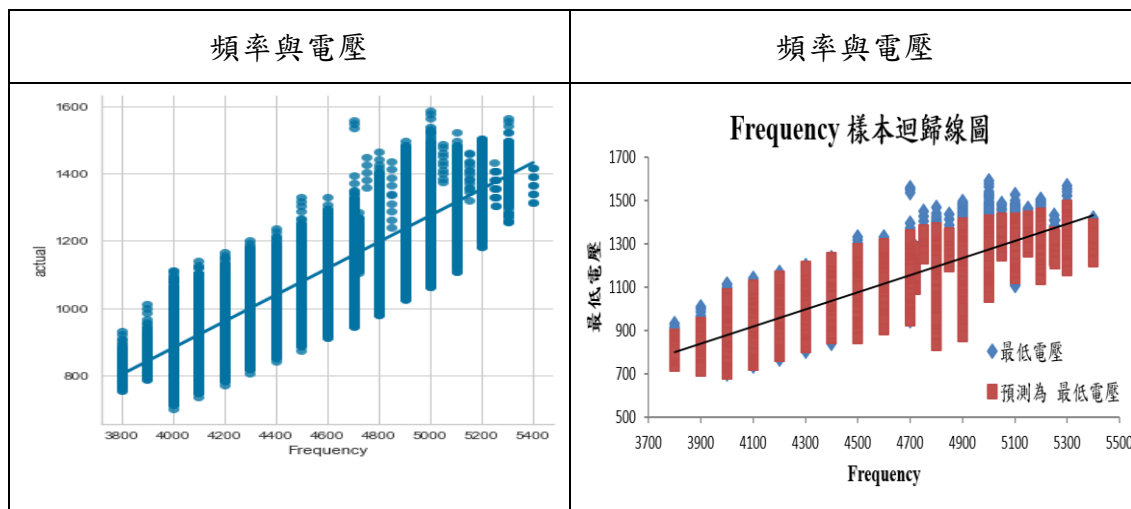
表 14 處理器體質直方圖



(5)「處理器操作頻率趨勢圖」對應表：

從表 15 說明操作頻率與核心電壓的趨勢圖，如同先前章節描述，要提升頻率得到好的效能則必須相對提供較高的電壓，因此從下表趨勢瞭解頻率和電壓為正比，並呈現線性的趨勢，也是再驗證本研究以線性迴歸來進行分析的佐證。

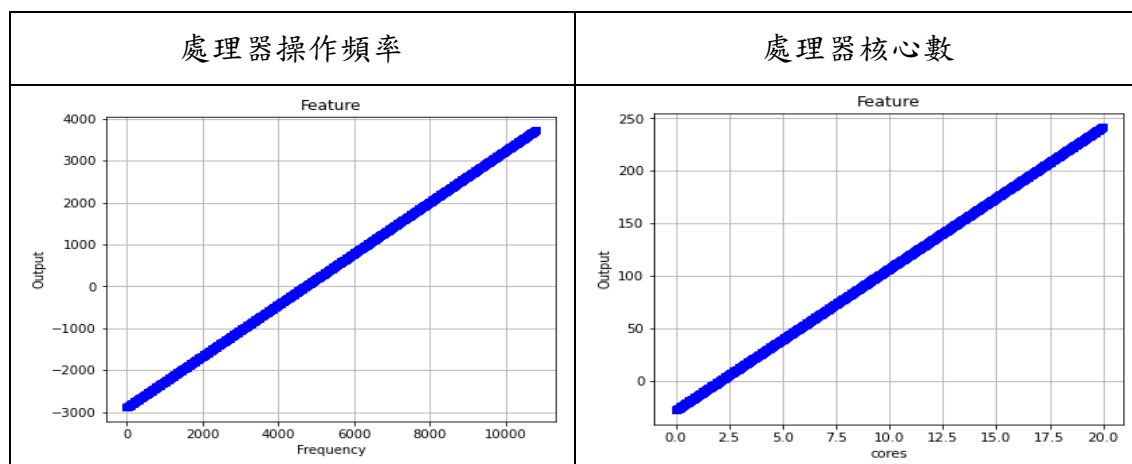
表 15 處理器操作頻率與核心電壓趨勢圖



(6)「處理器頻率與核心數趨勢圖 vs. 穩定輸出電壓」對應表：

從表 16 關係圖可供工程研發人員快速瞭解本文研究的相關特徵關連性，由於機器或電腦系統僅會真實呈現分析資料的關係和曲線，至於各個特徵關連及是否符合研究的特性，仍必須仰賴工程研發人員的領域知識來進行有效性的判斷。

表 16 處理器頻率與核心數趨勢圖 vs. 穩定輸出電壓



(7)「殘差圖(residual plots)」：

如圖 11 為本文研究預測結果與驗證以“殘差圖”為最後分析圖形來總結，所有描繪的點都以 0 為橫軸的直線上下隨機散佈，因此迴歸直線對各個觀測值的擬合情況是良好的，說明了變數 x 與 y 之間有顯著的線性相關性， R 平方(R^2) 數值為 0.925 非常接近 1.0，代表此迴歸模式能夠解釋全體變異量的比例愈大、此模式愈有解釋能力，證明使用的模型與這些量測數據有很高符合性。

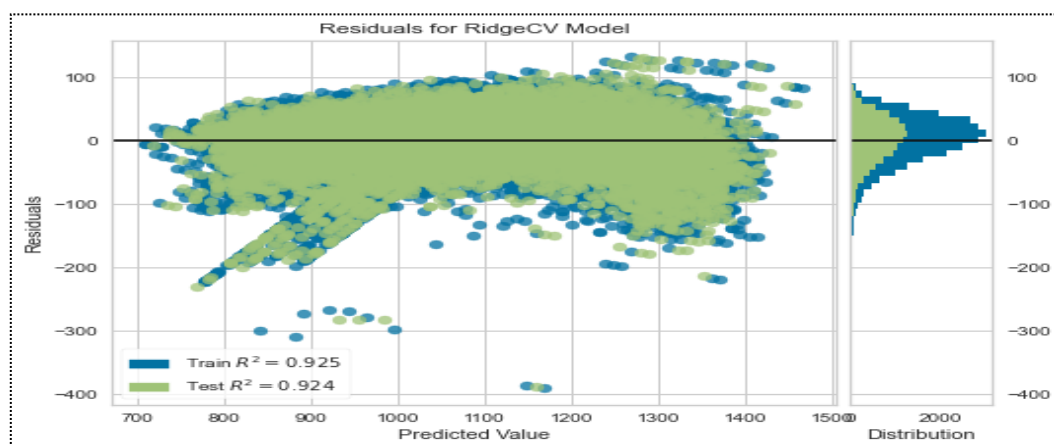


圖 11 迴歸模型驗證結果與殘差圖(residual plots)



4.2 穩定性驗證預測結果

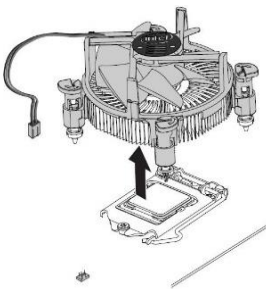
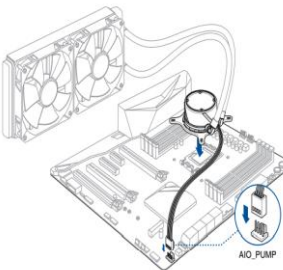

進行迴歸分析後預測準確度的可靠度及穩定性驗證，以大量不同的測試平台、設備及條件來進行作業系統下執行壓力測試應用軟體 Real Bench 1 小時的測試如表 17，並以一個月的時間每天記錄各測試平台的測試結果，並將統計表提供給工程研發人員參考供演算法及參數的調整。

4.2.1 穩定性測試及結果

表 17 效能及穩定性測試平台

主機板	支援線性迴歸演算法-相容的各型號主機板
處理器	Intel i5/i7/i9 K 系列各型號處理器
散熱器	水冷型-各型號/廠牌散熱器(定義歸類如表 18)
電源供應器	750W 含以上各型號/廠牌電源供應器
測試程式	RealBench ² 測試程式
測試時間	每天進行各燒機平台的測試及記錄，維持一個月

表 18 散熱器分數範圍(本研究自行測試定義)

散熱器 分數	~130pts	150 ~ 160pts	170pts+
散熱器 型式	<u>風扇</u> 散熱器	<u>水冷</u> 散熱器	高階水冷或散熱能力非常優異的 <u>水路式</u> 散熱器
參考外觀			
			(聯力水道板 官方配置圖，2022)

²RealBench，<https://rog.asus.com/articles/guides/guide-realbench-hwbot-edition-v2-43/>

由於壓力測試條件及環境較為嚴苛，研發人員選擇較一般普及常用的水冷散熱器為主要測試設備，其分數為 150-160pts 左右，以減少機台變數，此僅為本研究研發人員設計給測試人員的測試條件及方法。

表 19 穩定性測試統計表

處理器編號	水冷散熱器型號	電源供應器型號(瓦數)	超頻%	核心數	結果
QWJS-0010	ROG AYUO 120 RGB	ANTEC SIGNITURE 1000W	28%	6	Pass
QWJS-0005	ROG STRIX LC 120 RGB	ASUS ROG Thor 1200W	25%	6	Pass
QX27-0002	ROG STRIX HIRYU 120	Corsair RM850X 850W	51%	8	Pass
QWJT-0007	CRYORIG A-Series	IN WIN IRP-COM1500 1500W	28%	6	Pass
QWX4-0014	TUF LC 240	Corsair RM850X 850W	38%	8	Pass
QWX6-0023	CRYORIG A-Series	IN WIN IRP-COM1500 1500W	38%	8	Pass
QWJS-0010	ROG AYUO 120 RGB	ANTEC SIGNITURE 1000W	28%	6	當機
QWJS-0010	ROG AYUO 120 RGB	ANTEC SIGNITURE 1000W	28%	6	Pass
QWX6-0024	ROG STRIX LC 120 RGB	ASUS ROG Thor 1200W	38%	8	Pass
QWX6-0009	TUF LC 240	Corsair HX1200I 1200W	38%	8	Pass
QWX6-0022	ASUS ROG Ryujin 360	ASUS ROG Thor 850W	38%	8	Pass
QX27-0002	ROG STRIX LC 120 RGB	ASUS ROG Thor 1200W	51%	8	Pass
QWX4-0004	M13EG 一體式水冷	Corsair RM850I	38%	8	Pass
QWJS-0024	ASUS ROG Ryuo 240	ASUS ROG Thor 1200W	25%	6	Pass
QWX6-0019	ASUS ROG Ryuo 240	ASUS ROG Thor 1200W	38%	8	Pass
QWJS-0026	ASUS ROG Ryuo 240	EnerMax PRO87+ 600W	28%	6	Pass
QWX6-0017	ROG STRIX LC 240 RGB	Corsair HX850I 850W	41%	8	Pass
QX27-0004	M13EG 一體式水冷	Corsair RM850I 850W	51%	8	Pass
QXX6-0016	IDCOOLING 240	FSP HDP1000M 1000W	38%	8	Pass
QWJS-0021	IDCOOLING 240	FSP HDP1000M 1000W	25%	6	Pass
QWJS-0004	ASUS ROG STRIX LC-120	Super Flower 1000W	28%	6	Pass
QWX6-0037	ASUS ROG STRIX LC-120	Super Flower 1000W	38%	8	Pass
QWJS-0028	ASUS ROG STRIX LC-120	Super Flower 1000W	28%	6	Pass
QWJS-0027	ASUS TUF LC-120	FSP AURUM 1200W	28%	6	Pass
QWX6-0039	ASUS TUF LC-120	FSP AURUM 1200W	38%	8	Pass
QWJS-0012	ASUS ROG STRIX LC-240	Silverstone DA850	28%	6	Pass
QWJS-0025	ROG RYUJIN 240	ROG THOR 850W PLATINUM	28%	6	Pass



QWX6-0040	ROG RYUJIN 240	ROG THOR 850W PLATINUM	41%	8	Pass
QWX4-0016	TUF LC 120 RGB	ANDYSON AD-1200W	38%	8	Pass
QWX6-0038	TUF LC 120 RGB	ANDYSON AD-1200W	38%	8	Pass
QX29-0004	ASUS ROG STRIX LC-240	Silverstone DA850	51%	8	Pass
QWX6-0013	ASUS TUF LC-120	COOLMASTER 1250W	38%	8	Pass
QWJS-0015	ROG STRIX LC-120 RGB	TUF650W-GAMING	28%	6	Pass
QWX6-0011	NH-U14S	ROSEWILL HIVE-850 850W	38%	8	Pass
QWJS-0018	ROG STRIX LC 240 RGB	Corsair AX850 850W	28%	6	Pass
QWJS-0022	ROG RYUJIN 240	ROSEWILL 1000W	28%	6	Pass
QX29-0003	ROG STRIX LC-120 RGB	TUF650W-GAMING	51%	8	Pass
QWJS-0014	ASUS ROG STRIX LC-240	Corsair HX1000I 1000W	25%	6	Pass
QWX6-0031	ASUS ROG STRIX LC-240	Corsair HX1000I 1000W	41%	8	Pass
QX29-0002	ROG STRIX LC 360 RGB	FOCUS 650 GOLD 650W	51%	8	Pass
QWX6-0032	ROG STRIX LC 240 RGB	Corsair CMPSU 750W	34%	8	Pass
QWJS-0014	ASUS ROG STRIX LC-240	Corsair HX1000I 1000W	28%	6	Pass
QWX6-0031	ASUS ROG STRIX LC-240	Corsair HX1000I 1000W	38%	8	Pass
QWX4-0015	ASUS ROG STRIX LC-240	TT 1200W	38%	8	Pass
QX29-0002	ASUS ROG STRIX LC-240	Corsair CMPSU 750W	51%	8	Pass

經過測試人員每日將穩定性壓力測試報表及結果提供給工程研發人員，以便立即進行相關參數及演算法的調整，從測試報表結果中可發現處理器效能會依不同的處理器體質及搭配的散熱器能力如上表 19，提升至 25% ~ 51% 不等，研發人員可從表中瞭解測試平台發生當機故障(Fail)的時間及設備條件，以便立即調整可於隔日進行驗證。研發人員以此測試方式及流程來建立演算法導入的可靠度，並瞭解預測的準確率，在 45 個測試平台中有 1 平台發生當機故障，即 2.22% fail rate，研發人員可再調整相關參數和演算法模型持續提升預測的準確率及可靠度。



4.2.2 研究總結與模型限制

歸納本研究與以往利用線性迴歸來預測處理器、晶片效能文獻的差異及方法，提出分析處理器在不同的核心數、負載狀況擷取的頻率($p0freq$, $p1freq$)、電壓($p0vid$, $p0avxvid$)、核心數($cores$)、超執行緒(HT)等處理器特性與散熱器($cooler$)關鍵影響處理器效能參數，透過這些特徵參數的優化及測試數據，利用成熟的線性迴歸演算法並先以線性的模型來建立多目標的預測模型，分別在最佳的操作頻率及核心電壓得到 92%與 84%不錯的準確率，有機會在提高了操作頻率後，透過模型預測建議出基於此操作頻率下最佳核心壓的設定，結合得到處理器的效能優化。

然而在真實的應用，特徵參數如頻率、電壓、溫度、散熱能力等不會是完全的線性關係，基於本研究多變數線性迴歸模型，由於實際場域上的數據並不是完全線性的分佈，所以在穩定性的驗證仍存在 2.22%的 fail rate，代表主機板使用者拿到處理器於 BIOS or OS 下選擇啟動效能優化的功能後仍會有機率發生當機的狀況，因此本研究曾嘗試在一些特徵參數直接加上平方項/二次項(Quadratic)或是三次項(Cubic)的非線性曲線來擬合數據，但經過多次的驗證，對於再提升準確率及減少故障率非常有限，因此為了在實際應用上的需求，僅能將操作頻率及核心電壓的預測模型，切分成數個處理器頻率的分段處理，再採取決策樹分類方法來減少故障率。此部份也是本研究目前僅以線性迴歸方法的限制，以決策樹的分類方式主要目的為調整模型參數使得穩定性測試的故障率為 0%的目標，較難以有邏輯、學術性的方法整理，此為本研究未來可再持續研究及精進的地方。



4.2.3 線性迴歸方法結果比較

主機板與電腦系統效能的提升，在第 2 章節已探討過以往由通用規則建立的硬體或軟體方式來實現，表 20 加入本研究運用線性迴歸預測處理器效能的比較和結果，工程研發人員以上述說明的驗證方法，以大量機台的燒機(Aging)測試方式，以確保主機板使用者在提升效能成功機率及作業系統下運行穩定度都得到良好的可靠度。

表 20 運用線性迴歸預測處理器效能結果比較

效能調整	硬體方式	軟體方式	線性迴歸方法
優點	◆ 不需要在 OS 下安裝軟體	◆ 可編寫自動調整程序	◆ 可依據處理器體質及散熱器特性並進行”動態”調整
			◆ 可實施在 BIOS 或系統作業軟體方式 ◆ 簡單、易用，使用操作直覺不需俱備調教知識
缺點	◆ 需增加實體按鈕或外接設備	◆ 需要在 OS 下另外安裝軟體	◆ 需花時間建立和收集大量資料
	◆ 操作上需具備一定的效能調教領域知識	◆ 無法判斷處理器體質及散熱器	◆ 需具備效能調教與演算法知識
	◆ 無法判斷處理器體質及散熱器	◆ 調教時間長	
成功機率	低	中	高
穩定度	低	中	高

因此本研究到此確定達成了我們在第三章所設定任務：

「讓主機板可預測處理器和散熱器特性」，運用線性迴歸預測處理器效能，也找到一個新的方法來改善以往硬體及軟體方式的缺點，帶給主機板使用者的處理器效能優化上有較快及成功機率也更高的方式。

第五章 結論與建議



本研究結論，有別以往運用線性迴歸或機器學習的相關文獻，都僅單純針對單晶片或處理器有限條件下的預設效能、價格的預測。本研究以工程研發人員的領域知識、測試資料來數據趨動，並輔以線性迴歸的演算法建立一個有效的方法：

「讓主機板可預測處理器及散熱器特性」，針對不同的處理器及散熱器特性、體質後，預測出處理器在最佳操作頻率及核心電壓兩者重要的參數組合，使主機板使用者得到處理器在優化後的效能、並在作業系統下能夠穩定運行。歸納出工程研發人員的領域知識是最絕對的關鍵因素，透過跨領域的學習數據的應用整理，增加關鍵的輸入變數 x 及選擇適用的演算建立適用的預測模型，最後得到良好可靠度的輸出預測結果 y ，分別在預測最佳操作頻率得到 84%、在關鍵的最佳核心電壓為 92% 的準確率。僅利用線性迴歸演算法及正規方程式的優化流程，在投入規模性的測試驗證後，確定就能夠針對本研究欲達成的目標及預測有非常好的結果，改善以往硬體及軟體效能優化的缺點，除了在產業樹立一個數據趨動的案例，也提供了在運用線性迴歸的參考應用和學術上演算法實作的研究方法探討。

最後，在本研究多變數線性迴歸模型的限制，由於實際場域上的數據並不是完全線性的分佈，若要再提高預測模型的準確率及減少實際導入的穩定性運作故障率，嘗試在一些特徵參數直接加上平方項/二次項 (Quadratic) 或是三次項 (Cubic) 的非線性曲線來擬合數據，對於再提升準確率及減少故障率非常有限，因此為了在實際應用上的需求，目前較有效方法僅能基於原預測模型以不同的處理器頻率範圍來分段，並以決策樹演算法來分類後，持續調整預測模型，來儘可能達到減少穩定性測試的故障率，較難以有邏輯、學術性的方法整理，此為本研究建議及提供未來學者參考並能夠繼續研究、探討的方向。

參考文獻



Andrew Ng, Instructor Founder, DeepLearning.AI & Co-founder, Coursera

<https://www.coursera.org/learn/machine-learning>

Borkar, S., Karnik, T., Narendra, S., Tschanz, J., Keshavarzi, A., & De, V. (2003, June).

Parameter variations and impact on circuits and microarchitecture. In *Proceedings of the 40th annual Design Automation Conference* (pp. 338-342).

Borkar, S. (2005). Designing reliable systems from unreliable components: the challenges of transistor variability and degradation. *Ieee Micro*, 25(6), 10-16.

Cazorla, F. J., Ramirez, A., Valero, M., & Fernández, E. (2004, December). Dynamically controlled resource allocation in SMT processors. In *37th International Symposium on Microarchitecture (MICRO-37'04)* (pp. 171-182). IEEE.

Chang, K. W., Huang, C. Y., Mu, S. P., Huang, J. M., Chen, S. H., & Chao, M. C. T. (2018, August). DVFS binning using machine-learning techniques. In *2018 IEEE International Test Conference in Asia (ITC-Asia)* (pp. 31-36). IEEE.

Chen, T., & Guestrin, C. (2016, August). Xgboost: A scalable tree boosting system. In *Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining* (pp. 785-794).

Choi, S., & Yeung, D. (2006, June). Learning-based SMT processor resource distribution via hill-climbing. In *33rd International Symposium on Computer Architecture (ISCA'06)* (pp. 239-251). IEEE.

Humeau, S., Wijaya, T. K., Vasirani, M., & Aberer, K. (2013, October). Electricity load forecasting for residential customers: Exploiting aggregation and correlation between households. In *2013 Sustainable internet and ICT for sustainability (SustainIT)* (pp. 1-6). IEEE.

Inal, G., & Küçük, G. (2018, November). Application of machine learning techniques on prediction of future processor performance. In *2018 Sixth International Symposium on Computing and Networking Workshops (CANDARW)* (pp. 190-195). IEEE.

Intel® Core™ 處理器產品。

<https://www.intel.com.tw/content/www/tw/zh/products/details/processors/core.html>

Ipek, E., Supinski, B. R. D., Schulz, M., & McKee, S. A. (2005, August). An approach to performance prediction for parallel applications. In *European Conference on Parallel Processing* (pp. 196-205). Springer, Berlin, Heidelberg.

Iyer, S. G., & Pawar, A. D. (2019, November). Machine learning model for predicting price of processors using multivariate linear regression. In *2019 International Conference on Smart Systems and Inventive Technology (ICSSIT)* (pp. 52-56). IEEE.

Jiménez, D. A., & Lin, C. (2001, January). Dynamic branch prediction with perceptrons. In *Proceedings HPCA Seventh International Symposium on High-Performance Computer Architecture* (pp. 197-206). IEEE.

Joseph, P. J., Vaswani, K., & Thazhuthaveetil, M. J. (2006, February). Construction and use of linear regression models for processor performance analysis. In *The Twelfth International Symposium on High-Performance Computer Architecture, 2006.* (pp. 99-108). IEEE.

Karkhanis, T. S., & Smith, J. E. (2004, June). A first-order superscalar processor model. In *Proceedings. 31st Annual International Symposium on Computer Architecture, 2004.* (pp. 338-349). IEEE.

List of Intel processors.

https://en.wikipedia.org/wiki/List_of_Intel_processors

Lo, J. L., Parekh, S. S., Eggers, S. J., Levy, H. M., & Tullsen, D. M. (1999). Software-directed register deallocation for simultaneous multithreaded processors. *IEEE Transactions on Parallel and Distributed Systems*, 10(9), 922-933.

Machine Learning week 1: Cost Function, Gradient Descent and Univariate Linear Regression

https://medium.com/@lachlanmiller_52885/machinelearning-week-1-cost-function-gradient-descent-andunivariate-linear-regression-8f5fe69815fd

Maxwell, P., Hartanto, I., & Bentz, L. (2000, October). Comparing functional and structural tests. In *Proceedings International Test Conference 2000 (IEEE Cat. No. 00CH37159)* (pp. 400-407). IEEE.

Monreal, T., González, A., Valero, M., González, J., & Viñals, V. (2000). Dynamic register renaming through virtual-physical registers. *Journal of Instruction Level Parallelism*, 2, 4-16.

Phan, T. D. (2018, December). Housing price prediction using machine learning algorithms: The case of Melbourne city, Australia. In *2018 International Conference on Machine Learning and Data Engineering (iCMLDE)* (pp. 35-42). IEEE.

Sadi, M., Tehranipoor, M., Wang, X., & Winemberg, L. (2015, May). Speed binning using machine learning and on-chip slack sensors. In *Proceedings of the 25th edition on Great Lakes Symposium on VLSI* (pp. 155-160).

Sarangi, S. R., Greskamp, B., Teodorescu, R., Nakano, J., Tiwari, A., & Torrellas, J. (2008). VARIUS: A model of process variation and resulting timing errors for microarchitects. *IEEE Transactions on Semiconductor Manufacturing*, 21(1), 3-13.

Sharkey, J., Ponomarev, D., & Kanad Ghose, M. S. I. M. (2005). A Flexible Multithreaded Architectural Simulation Environment. *CS-TR 05-DP01, State Univeristy of New York at Binghamton*.

Shi, Q., Wang, X., Winemberg, L., & Tehranipoor, M. M. (2016). On-chip sensor selection for effective speed-binning. *Analog Integrated Circuits and Signal Processing*, 88(2), 369-382.

Su, M. Y., Lin, W. C., Kuo, Y. T., Li, C. M., Fang, E. J. W., & Hsueh, S. S. Y. (2021, April).

Chip Performance Prediction Using Machine Learning Techniques. In *2021 International Symposium on VLSI Design, Automation and Test (VLSI-DAT)* (pp. 1-4). IEEE.

The Secret Of CPU HYPERTHREADING In Depth

[The Secret Of CPU HYPERTHREADING In Depth \(ourtechroom.com\)](https://ourtechroom.com)

Varma, A., Sarma, A., Doshi, S., & Nair, R. (2018, April). House price prediction using machine learning and neural networks. In *2018 second international conference on inventive communication and computational technologies (ICICCT)* (pp. 1936-1939). IEEE.

Viktorovich, P. A., Aleksandrovich, P. V., Leopoldovich, K. I., & Vasilevna, P. I. (2018, August). Predicting Sales Prices of the Houses Using Regression Methods of Machine Learning. In *2018 3rd Russian-Pacific Conference on Computer Technology and Applications (RPC)* (pp. 1-5). IEEE.

Wang, H., Koren, I., & Krishna, C. M. (2008, October). An adaptive resource partitioning algorithm for SMT processors. In *Proceedings of the 17th international conference on Parallel architectures and compilation techniques* (pp. 230-239).

Wang, X., Tehranipoor, M., & Datta, R. (2008, November). Path-RO: A novel on-chip critical path delay measurement under process variations. In *2008 IEEE/ACM International Conference on Computer-Aided Design* (pp. 640-646). IEEE.

What is Machine Learning?

<https://deeptai.org/machine-learning-glossary-and-terms/machine-learning>

李侑澄. (2014). 運用層級分析法探討 [遊戲玩家 PC 硬體購買決策之關鍵因素].