

**經濟部工業局**

**產業升級創新平台輔導計畫**

**「智慧機械-產業聚落供應鏈數位串流暨AI應用-先期顧問規劃」**

計畫查核點D2

**工具機銷售預測AI建模報告**

契約編號：107年11月13日工知字第10701106401號

計畫名稱：**百德機械體系供應鏈AI應用加值計畫**

計畫期間：自107年09月01 至 108年02月28日止

公司名稱：**百德機械股份有限公司**

製作單位：**財團法人中衛發展中心**

中華民國108年2月

**目 錄**

[一、 文件說明 1](#_Toc2176982)

[二、 前言 1](#_Toc2176983)

[三、 百德工具機AI建模整體流程說明 2](#_Toc2176984)

[(一) 企業需求調查 3](#_Toc2176985)

[(二) 百德銷售資料ETL處理 5](#_Toc2176986)

[(三) 百德銷售預測模型AI建模方法 6](#_Toc2176987)

[四、 百德工具機銷售AI建模參數規劃 7](#_Toc2176988)

[(一) 問題及需求定義 7](#_Toc2176989)

[(二) 數據取得 7](#_Toc2176990)

[(三) 數據分析 7](#_Toc2176991)

[(四) 可視化輔助 7](#_Toc2176992)

[五、 百德工具機AI模型試作 10](#_Toc2176993)

[(一) 建立百德工具機AI模型 10](#_Toc2176994)

[(二) 百德工具機AI模型限制 15](#_Toc2176995)

[六、 結論 16](#_Toc2176996)

**表 目 錄**

[表 1內部銷售預測關鍵因子 8](#_Toc2177007)

[表 2外部銷售預測關鍵因子 9](#_Toc2177008)

[表 3各國PMI與百德工具機銷售訂單數之相關係數表 14](#_Toc2177009)

[表 4百德工具機2015~2018年銷售實績與AI預估比較表 15](#_Toc2177010)

**圖 目 錄**

[圖 1導入流程架構 2](#_Toc2177058)

[圖 2百德AI模式規劃輔導方法 3](#_Toc2177059)

[圖 3百德工具機銷售現況做法及問題 4](#_Toc2177060)

[圖 4資料處理(ETL)流程 5](#_Toc2177061)

[圖 5計畫資料之特徵欄位 6](#_Toc2177062)

[圖 6機器學習步驟流程圖 6](#_Toc2177063)

[圖 7預測建模架構 12](#_Toc2177064)

[圖 8百德主要出口國之製造業採購經理人指數(PMI)趨勢圖 13](#_Toc2177065)

# 文件說明

本報告係根據本計畫D分項工具機銷售預測AI模式規劃之演算試作與AI預測建模規劃內容，進行百德工具機銷售建模參數規劃、演算法試作、建立AI預測模型等，以作為百德未來進行銷售預測AI建置之參考文件。

# 前言

受到大數據浪潮帶動下，產企業紛紛朝轉型及創新目標進行管理改革。面對市場競爭，徒有產品優勢已不足以應付現今的消費市場需求，因此各產業除了必須面臨產品持續創新與精進課題外，消費服務創新亦是各產業現階段刻不容緩且必備的改革重點項目之一。

為了改善產業的管理效率，資訊流的整合及應用逐漸成為不可或缺的管理關鍵要素及技術之一。此外，如何強化資訊透明度亦是必須考量的重點項目之一，故必須藉由系統資訊的輔助方能將龐大的歷史性資訊加以匯流整合並透過資料倉儲方能獲得妥善的資料處理。過去以產品為導向的經營模式將昇華至以服務為導向的發展，由此觀之，產企業將藉由整合供應鏈資訊來有效掌握產銷資訊，從中透過資訊數字化及可視化的應用來瞭解營運的狀態，並透過內部及外部相關營運資訊的量化來掌握市場趨勢。

有鑑於此，為了因應市場變動及需求，本計畫將嘗試透過百德機械資料應用分析，找出管理上潛在的關聯性並建立關鍵營銷商業預測模型。期望透過預測模型的建立，可輔助業者在未來營運策略擬定及決策判斷的參考來源之一。

# 百德工具機AI建模整體流程說明

本計畫之導入流程架構如圖1所示：

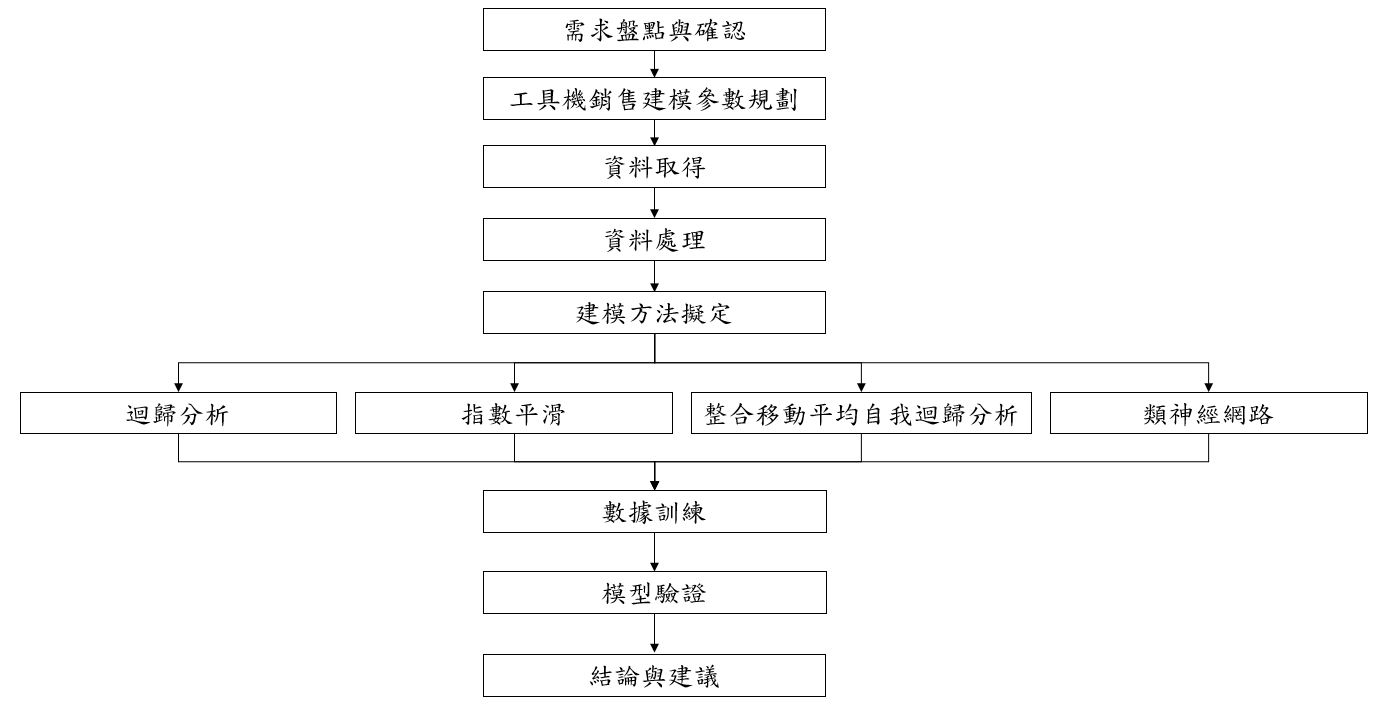


圖 1導入流程架構

本計畫依AI模式規劃三項步驟輔導方法，詳述如下：

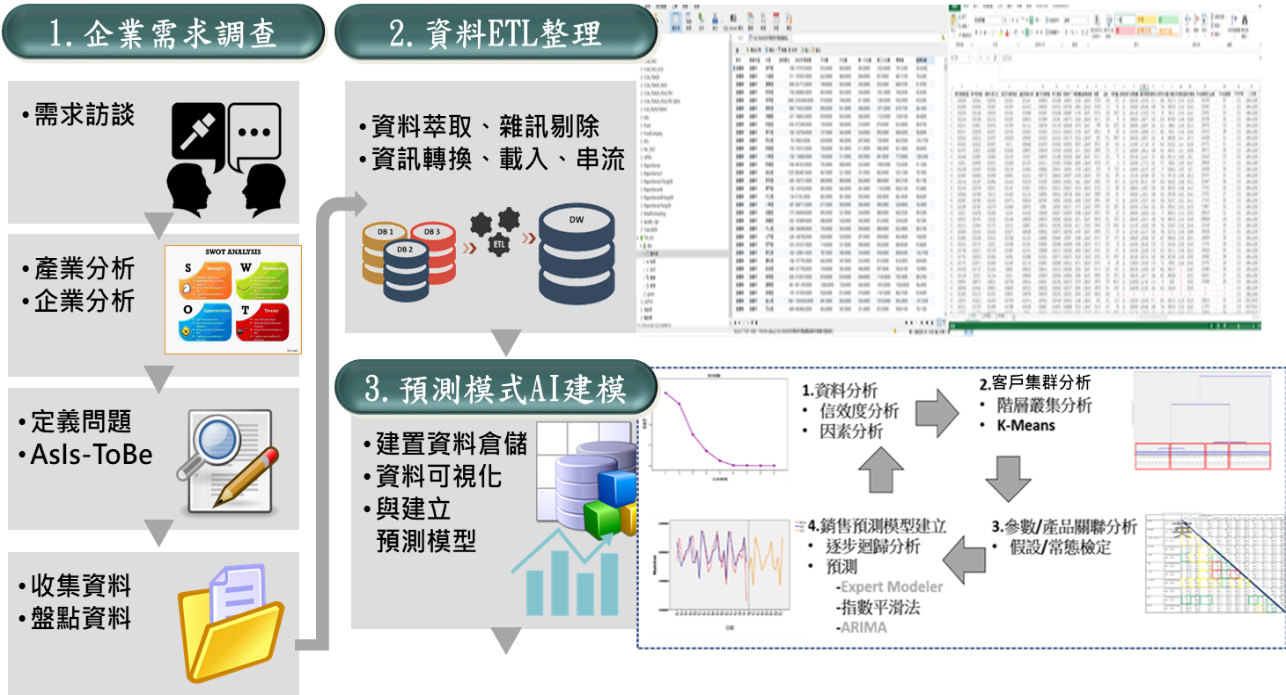


圖 2百德AI模式規劃輔導方法

## 產業現況與企業需求調查

百德機械股份有限公司主要產品為工具機，客戶遍及全球，以歐美為大宗、亞洲及內銷為輔。其出口型產業之特性，受諸多國際因素影響，致使銷售預測評估有其困難之處。

近年來國際局勢瞬息萬變，英國脫歐、川普當選、歐債風暴、美債危機，黑天鵝層出不窮，深深影響產業的穩定性。以歐債風暴為例，從前期客戶取消訂單的損失、至後期急單所衍生的成本，都顯示要在這個競爭激烈的國際市場中生存，就必須學習從資料中嗅出市場先機。如何準確預測銷售數量與機型、有效預知訂單快速變化的問題，將是本次試跑的主要課題。

百德客戶分處五大洲，銷售數量除了受全球景氣影響，各國國內的製造業指數、國民年均所得與產業發展的階段，都可能是影響出口訂單的主因。然國際間總體與個體經濟指標眾多，如何從中挑選適合的指標將是一大考驗。

百德營運模式係採接單式生產與經銷商代理兩種模式。目前銷售預測方式，僅參考之前年度的銷售狀況，配合代理商主觀意見的評估，受限樣本資料數量與主觀判斷，不易精準預測，進而影響到材料庫存成本的壓力。

因此本計畫擬請百德提供2009-2018年共計十年之歷史銷售資料（以下簡稱為本計畫資料），欲配合代理商進行銷售預測問卷，建立銷售預測模型，用以預估百德工具機的銷售數量，並配合歷史資料，推估主要機型的可能數量。

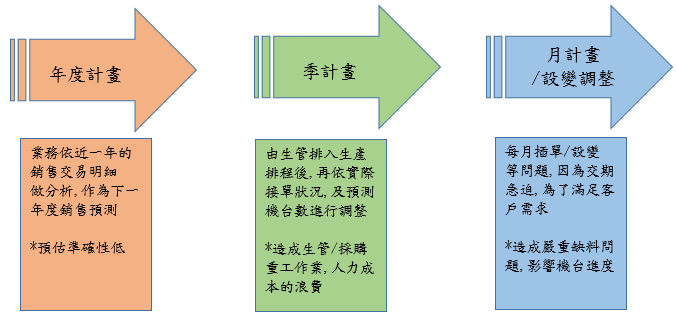


圖 3百德工具機銷售現況做法及問題

本計畫將近年國際局勢變遷之客觀事實，與百德面臨的主觀問題，統整為以下四個命題，並試圖以量化分析降低國際局勢與景氣對於訂單的影響。解決議題如下：

1. 彙整百德歷史銷售訂單之各年度機種/機型/選配件/代理商之出貨狀況。
2. 全球市場區域代理商的經濟變化是否影響百德工具機銷售成效?
3. 百德工具機熱銷產品之各年度/地區代理商之銷售量分析?
4. 百德工具機熱銷產品之選配件需求統計分析?

## 百德銷售資料ETL處理

本計畫先針對計畫資料進行資料品質檢視以及ETL處理。所謂ETL處理係指資料從來源端經過萃取(extract)、轉置(transform)、載入(load)至目的端。接著進行探索性資料分析EDA(Exploratory data analysis)，初步瞭解資料特性，即個別欄位的資料型別、平均數、偏態係數、是否有缺失值、如何填補空值，其資料處理(ETL)流程如圖3所示。

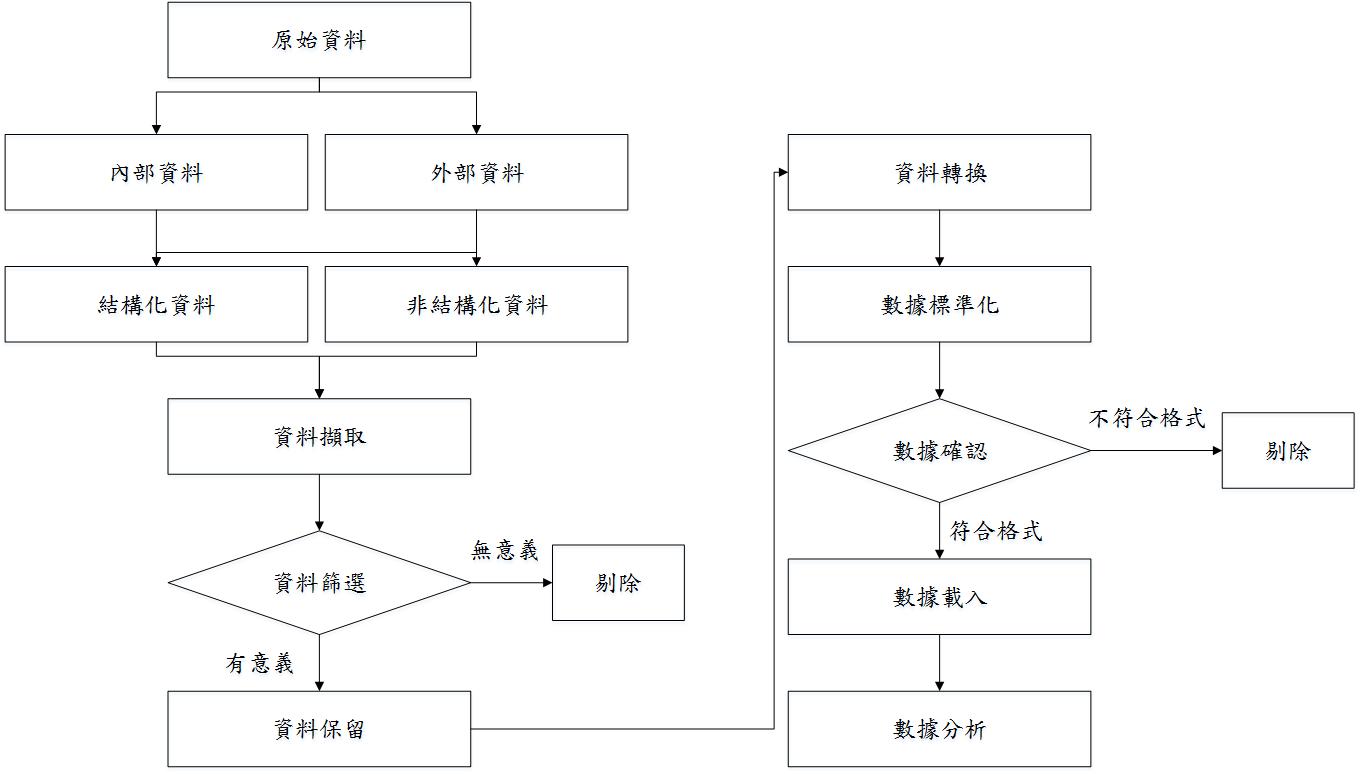


圖 4資料處理(ETL)流程

計畫資料經前述步驟處理後，由原始資料之特徵欄位(如圖4)進行缺值填補、欄位分割及特徵篩選。其中數值型欄位包括入庫數量、出貨數量及訂單編號衍生之訂單數量共三欄；類別型資料衍生出控制器、主軸、刀庫等選配件欄位；並依需求目標挑選出貨日期，作為時間序列的主要依據。最後針對各項指標進行資料標準化作業，以及類別型資料進行資料攤平處理。

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 客戶代號 | 客戶說明 | 料號 | 料品名稱 | 料品規格 | 明細備註 |

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 訂單單號 | 訂單日期 | 出貨單號 | 出貨數量 | 出貨單確認日期 | 製令單號 | 開工日期 | 完工日期 | 入庫數量 |

圖 5計畫資料之特徵欄位

## 百德銷售預測模型AI建模方法

本計畫使用Python程式語言Scikit-learn套件進行分析，Scikit-learn依照目的將機器學習演算法分為四大類：分類(classification)、群集(clustering)、迴歸(regression)與降維(dimensionality reduction)，機器學習步驟流程如圖4所示。

首先確認資料數量(大於50筆)，以確保模型穩定性與預測能力，第二步確認預測目標為類別型資料或數值型資料。預測數值型資料，可視資料量及複雜程度以決定是否降維，之後進行迴歸分析預測；若預測目的為類別型資料，則視資料標籤的有無，決定進行分類分析或集群分析。

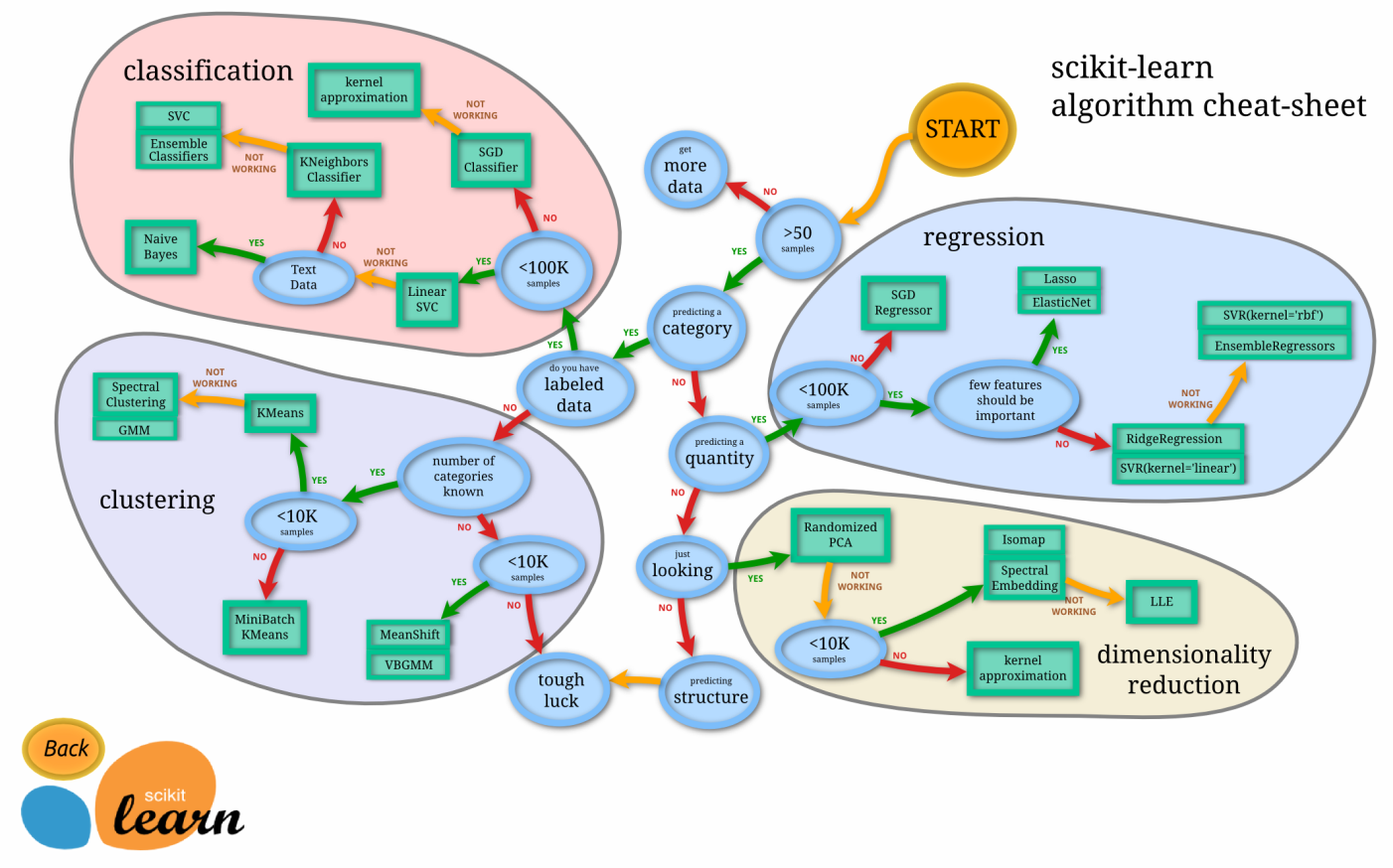


圖 6機器學習步驟流程圖

本計畫需求目標為銷售預測，預測欄位係屬連續型數值資料，故挑選迴歸相關類別之演算法。其次，因機型、選配件欄位經統計方法進行初步判別，與銷售訂單數量皆無顯著相關，故本計畫再依據時間序列的資訊，衍生相關欄位。

# 百德工具機銷售AI建模參數規劃

為即時掌握企業營運動態及市場變動，數據的應用將扮演著極為重要的角色。傳統企業管理主要是以人為主觀認知進行商業決策，但基於人為的不確定因素，難免會造成結果的偏頗及失準，故為了提升企業決策並降低決策風險，善用數據的輔助將能適時提升管理效益並能藉此達到管理創新的目的。企業對於數據的導入及應用大致可分四階段：

## 問題及需求定義

為釐清企業面臨問題，導入前期首要必須具備清楚的問題定義及改善目標，方能進一步擬定策略來因應並採取相對應的對策。

## 數據取得

為了實現大數據應用及資訊價值串流，數據將是不可或缺的關鍵。因此，數據的取得及處理亦是後續導入及推進的重點。

## 數據分析

除了問題盤點外，導入後期的數據分析也是影響決策判斷重要關鍵之一，包括數據的判讀及演算模型的選定。

## 可視化輔助

為了達到即時化的決策輔助，數據可視化的呈現，亦是後續AI建模及企業管理決策判斷的關鍵。因此，善用數據分析軟體輔助產業可視化管理資訊亦能提升管理績效。

為了提升後續預測模型準確性，並更貼近市場脈動，除了必須考量企業內部關鍵管理因子外，外部市場關鍵指標也是不容忽視的。有鑑於本計畫業者百德機械目前係以接單式生產為主並以外銷為大宗，在銷售決策仍以人工經驗法則進行判斷，故造成其在銷售預測準確度低落，因而導致管理成本管控成效不佳窘境。為了改善業者在管理預測的準確度，本計畫主要是協助業者透過供應鏈協同來達到即時資訊串流及回饋，並將數據反映至相關製程及訂單交期進行整合性管控，進而強化整體預測效益。因此，針對百德機械銷售預測建模，本計畫透過內部團隊與外部專家共同盤點與研議下，分別針對內部及外部影響銷售因子彙總如表1及表2所示：

表 1內部銷售預測關鍵因子

| **分類** | **指標項目** |
| --- | --- |
| 主軸 | 主軸碼 |
| 主軸名稱 |
| 代理商名稱 | 地區 |
| 國別 |
| 經銷商 |
| 客戶代號 |
| 全球客戶名稱 | 地區 |
| 國別 |
| 經銷商 |
| 客戶代號 |
| 刀數 | 刀數碼 |
| 刀數名稱 |
| 控制器 | 控制器碼 |
| 控制器名稱 |
| 機種機型 | 機碼 |
| 機種機型名稱 |
| 製造及訂單 | 出貨數量 |
| 刀數碼 |
| 訂單單號 |
| 訂單號 |
| 訂單日期 |
| 訂單序號 |
| 供應型態 |
| 機碼 |
| 結案日期 |
| 客戶 |
| 客戶名稱 |
| 控制器碼 |
| 料號 |
| 料品規格 |
| 料品名稱 |
| 退回數量 |
| 完工日期 |
| 完工數量 |
| 完工允諾日 |
| 序號 |
| 預計交貨日期 |
| 原始交期 |
| 製令單號 |
| 製令結案日 |
| 製令開工日 |
| 製令完工日 |
| 製造數量 |
| 主軸碼 |
| 機種 |
| 銷售機台分析 | 年 |
| 月份 |
| 數量 |

表 2外部銷售預測關鍵因子

| **分類** | **指標項目** |
| --- | --- |
| 景氣指標及燈號 | 製造業銷售量指數 |
| 工業生產指數 |
| 股價指數 |
| 產業外銷動態 | 臺灣機械出口值年增率 |
| 臺灣工具機出口值年增率 |
| 製造業採購經理人指數  (電力暨機械設備產業) | 原物料價格 |

# 百德工具機AI模型試作

針對數據建模及演算法應用上，近年來類神經網路演算法已廣泛運用在各領域問題解決課題中。有別於傳統的演算法，類神經網路演算可透過不斷學習及修正，解決相關線性及非線性議題，達到預測及辨識等效益。

然而，不論內部或外在環境都存在著各種影響銷售的因素，為了能提升銷售預測的準確性並降低影響因素所造成的不確定性，透過歷史數據及演算法的整合應用，將可找到最適切的預測模型，使企業透過模型的評估及應用，更有效掌握供應鏈供需平衡。因此，本計畫在銷售預測模型的建立上將聚焦於以督式學習為主軸，分別會使用四種據時間序列之演算法進行評估，包括迴歸分析、指數平滑法、ARIMA以及類神經網路等。

## 建立百德工具機AI模型

本計畫在銷售預測建模過程中除了會將業者內部營業數據納入分析外，亦會結合開放資料(Open Data)的運用進行更全面的資料分析，過程中所有取得的資料，不論是結構化資料或是非結構化資料均會透過數據的標準化，使其數據具備一致性。此外，亦會進行雜訊或缺漏值處理，以提升數據分析的可行性；待數據處理完成後，接著透過數據分析軟體(如：SPSS或Matlab等)進行銷售預測模型的建立；再接著，透過數據的測試及驗證來比較不同預測模型的預測準確值；最後，將分析數據與結果予以可視化圖表呈現，根據圖表的數據的呈現，嘗試找到資料的規則性，並從中萃取出關鍵特徵以發展更精準且適切的行銷策略。

### 資料取得步驟

由於本計畫的數據主要區分兩大部分，分別有內部營運資料及外部開放資料。首先，無論內部或外部資料的部分將透過SQL資料庫來進行介接與匯出 (表1及表2)。

### 資料處理(ETL)

完成資料取得步驟後，接著將進行資料處理階段，此部分可區分三大階段：

1. 階段一主要是將所取得的原始資料進行整理。透過人為篩選來擷取有用的資料予以保存，而無意義的資料則給予剔除。
2. 階段二則將擷取後的資料進行轉換，例如數據的標準化等。
3. 階段三則將轉換好的數據匯入分析軟體中等待後續的分析。

### 預測建模

待完成資料處理後，則將ETL後的數據匯入數據分析軟體中，並針對預測目標擬定建模參數。本計畫所設定之建模參數如下：

1. **依變數**

各機種年度銷售數量

1. **自變數**

* 內部指標：八大產銷分類指標項目，包含主軸、代理商名稱、全球客戶名稱、刀數、控制器、機種機型、製造及訂單以及銷售機台分析等(表1)。
* 外部指標：三大經濟指標項目，包含景氣指標及燈號、產業外銷動態以及製造業採購經理人指數(電力暨機械設備產業)(表2)。

數據將劃分為兩大部分，分別為測試及驗證。其中，測試階段將擷取70%的數據、驗證的部分則擷取30%數據進行分析；爾後，所得之四種預測建模結果將會相互進行交叉驗證與比較；模型的衡量指標將採取RMSE、MAPE與R^2進行準確度的評估；待完成預測建模並經交叉比對及驗證找出最適模型後，初期會就既有的分析數據及結果導入Microsoft Power BI軟體來建立可視化營運分析儀表板。本模型待後續經業者驗證確認其可行性與效益性後，則會規劃介接資料庫，透過模型的運算及可視化系統的輔助，可有效將靜態營運資料加值為動態互動資料，進而達到數據即時性的分析效益。本計畫預測建模架構如圖2所示。

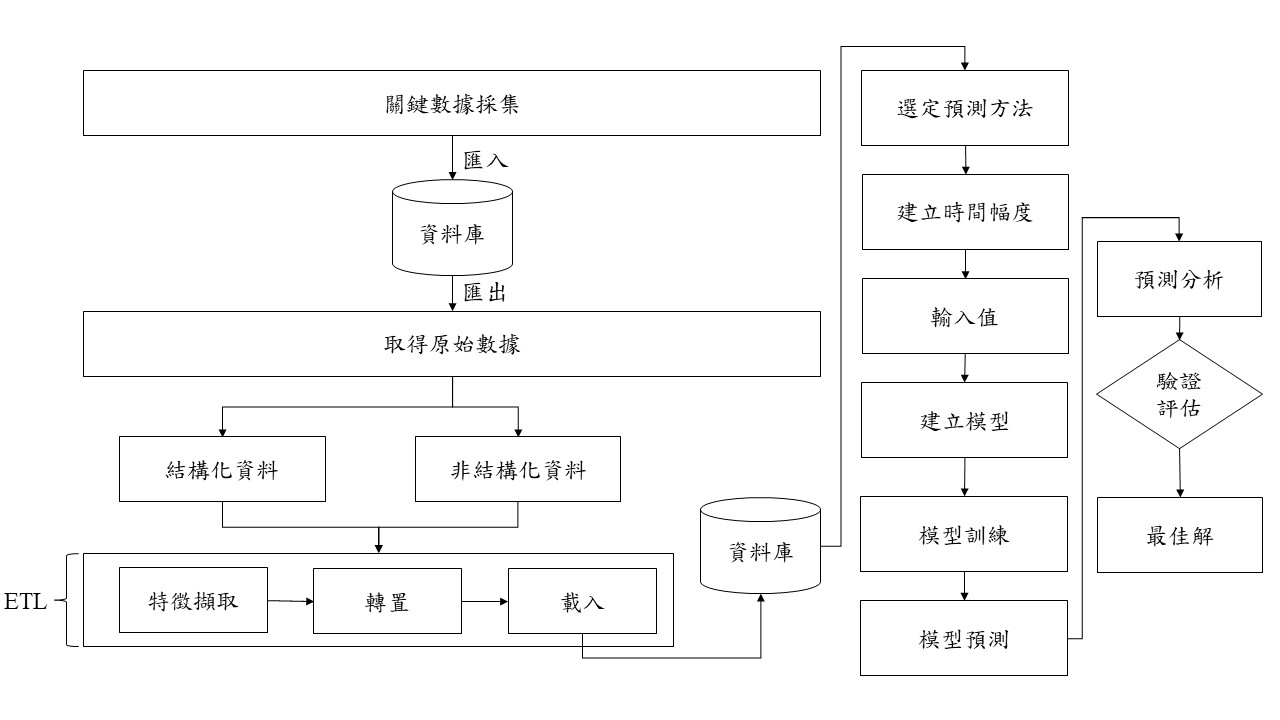


圖 7預測建模架構

迴歸模型的產出為數值型資料，屬於監督式學習，是透過輸入資料，將預測或挖掘特徵能力內化於電腦程式之中的一種方法。但電腦只能學習到資料本身的資訊，因此資料品質與特徵選取就已經決定電腦在AI學習的上限，模型本身只能逼近該上限。

由百德歷史銷售資料，可將銷售資訊連結年份、月份，將資料整理為時間序列特性的資料，特徵欄位的主體為時間與訂單數量兩項，其餘特徵欄位主要作為分類使用。計畫資料中能預測銷售訂單數量的欄位不足，須導入外部參數，以提升模型的預測能力。

考量百德客戶主要為出口，銷售訂單量除百德自身能掌握的因素外，尚受到國內外景氣所影響。因此嘗試納入百德眾多代理商中，訂單數量前五名：義大利、英國、土耳其、日本、瑞士，以及台灣、中國、美國共八國的製造業採購經理人指數（Purchasing Managers Index，以下簡稱PMI），作為反映國內外景氣的欄位。PMI除了反映製造業狀況、就業及物價表現，也是對出口很有預測力的前瞻性指標，對國家經濟活動的監測與預測具有重要作用。百德主要出口國之製造業採購經理人指數(PMI)趨勢圖如下圖：

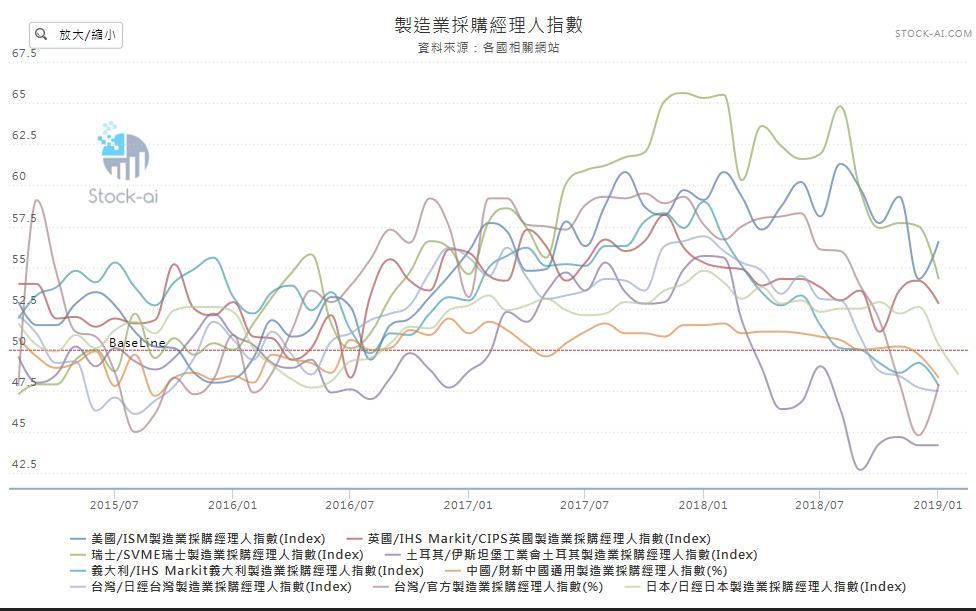


圖 8百德主要出口國之製造業採購經理人指數(PMI)趨勢圖

模型採用計畫資料中2015至2018年資料，首先針對年份、月份等類別型資料進行攤平（One-hot encoding），再將各國PMI進行標準化，接著透過機器學習方式建立廻歸分析模型，方程式如下：

銷售訂單總量 = 13.6606 \* ITA +33.6175 \* SWI -19.4924 \* CHI -18.9444 \* TUR +10.4387 \* TAI2 +19.2215 \* JPN +17.535 \* 2015 +18.3218 \* 2016 -30.3826\*MON1-9.7936 \* MON4 +20.2867 \* MON11 -22.992 \* MON12+38.0896

藉由程式的結構可以觀察到，我們可利用各國上月PMI指標，來預測本月銷售訂單數量，另外因為2015、2016年的銷售量優於其他年度，以及在每年的1月2月的淡季、12月的旺季，模型也都做權重調整。各國PMI與百德工具機銷售訂單數之相關係數表、百德工具機2015~2018年每月銷售實績與AI預估比較如下表：

表 3各國PMI與百德工具機銷售訂單數之相關係數表

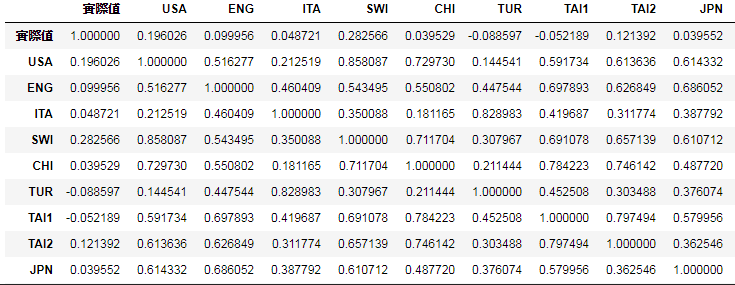
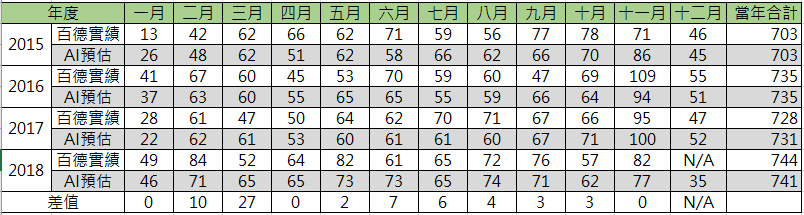


表 4百德工具機2015~2018年銷售實績與AI預估比較表



## 百德工具機AI模型限制

由於本模型是以量化資料為基礎，透過演算法建立預測能力的模型。為確保模型具有穩定且良好預測能力，最佳狀況必須大量數據資料作為輔助。本模型穩定性受樣本量數量限制，以及特殊事件和外部因子皆未能列入考量，建議若需實際建置，特徵欄位可加入領域知識所建議之欄位，並將重大事件影響納入考量。

# 結論

工具機產業是我國產業發展的重要基石之一，為了提升國家經濟發展水平與競爭力，善用產企業資源來扶植產業邁向轉型及升級已是國家與產企業們近年來積極投入改革的重點項目，亦是當前大環境下必須因應且欲解決的當務之急課題。

受到工業4.0帶動下，數位化管理轉型與智慧製造儼然已成為產企業必須具備且不可或缺的管理要素。為此，如何妥善應用企業資源並發展創新且差異化之商業模式使管理績效優化將是現階段及未來企業營運成敗的決勝關鍵。此外，為了滿足快速變動的消費市場以及消費者期待，預測手法及創新管理思維導入亦是當前許多產企業不斷追求突破藉此提高產業附加價值的成功契機之一。有鑑於此，本計畫現階段配合業者需求及資料的整合初步建立關鍵預測模型，期許透過數據輔助及學習，能加以落實管理的有效性及效率性，協助業者創造更卓越的管理績效。