



【金融科技知識】補充資料

顏回 老師提供

一、結構資料、非結構資料與半結構資料

結構資料、非結構資料與半結構資料是根據資料的組織形式進行區分的三種主要類型。每一種資料的結構化程度不同，適用於不同的分析和應用場景。讓我們詳細討論它們的定義、特徵及生活中的例子。

(一) 結構資料(StructuredData)

結構資料是指高度組織化且按照預定格式排列的資料。這些資料通常儲存在關聯式資料庫中（如 SQL 資料庫），有明確的欄位、資料類型（例如整數、字串、日期等）和關聯性，便於查詢、存取與處理。

1. 特徵

- (1) 有固定的資料模型或結構，通常以表格形式呈現。
- (2) 各欄位的資料具有明確定義(例如：客戶姓名、購買日期、價格等)。
- (3) 資料能夠通過查詢語言(如 SQL)輕鬆檢索和分析。
- (4) 適合批量處理和標準化報告。

2. 生活實例

- (1) 銀行交易記錄：每一筆交易都有明確日期、金額、交易類型、帳號等，並存儲在銀行資料庫中。這是典型的結構化資料。
- (2) 公司員工資料：一家公司的 HR 系統，記錄員工的姓名、職位、薪水、入職日期等資訊，這些資訊可以存放在關聯式資料庫中進行管理。
- (3) 商品銷售資料：零售商店的銷售數據，包括商品名稱、售出數量、價格、銷售日期等，通常存儲在數據庫的表格中。

(二) 非結構資料(UnstructuredData)

非結構資料是指沒有明確資料模型或結構的資料。這些資料無法輕易地存放在關聯式資料庫中進行整理和分析，通常由人類生成，如文字、影像、音頻、視頻等。這類資料通常需要使用自然語言處理(NLP)、影像處理等技術來提取有用資訊。

1. 特徵

- (1) 無固定格式，資料隨機且多樣化。
- (2) 很難直接用傳統的 SQL 查詢或統計工具處理。
- (3) 需要更先進分析工具(如人工智慧、機器學習)進行提取、分析和分類。
- (4) 儲存方式靈活，常使用文件存儲系統。

2. 生活實例

- (1) 電子郵件內容：雖然電子郵件有一些基本結構（如發件人、收件人、主題等），但郵件的正文內容可能包含隨意的文字、圖片、附件等，屬於非結構資料。
- (2) 社交媒體貼文：在 Twitter 或 Facebook 上的貼文內容，包含圖片、文字、視頻、表情

符號等，沒有固定的結構。

- (3) 圖片與視頻：生活中的照片、視頻檔案，無法直接用表格形式表示，需要特定的工具來處理和分析。

(三) 半結構資料(Semi-structuredData)

半結構資料介於結構化資料和非結構化資料之間。它具有某種結構，但這個結構不如結構資料那麼嚴格。通常使用標記語言(如 XML、JSON)來表示，其中的元素和屬性提供了資料的一定結構，但可以靈活擴展。

1. 特徵

- (1) 具有部分結構，便於理解與處理。
- (2) 結構靈活，元素和屬性之間的關聯不如結構資料那麼固定。
- (3) 易於在不同系統之間進行資料傳輸與共享。
- (4) 需要一定的解析工具來讀取和處理。

2. 生活實例

- (1) 網頁上的 HTML 文件：雖然網頁有標籤(如<html>、<head>、<body>)來表達結構，但網頁內容是自由的，可能包括文本、圖片、表單等，屬於半結構化資料。
- (2) 電子郵件的元數據：電子郵件標題、時間戳、發件人和收件人等信息，這些資料有一定結構，但不像表格數據那麼嚴格。

(四) 總結

1. 結構資料：適合高效檢索與分析，應用於商務數據、金融系統等。
2. 非結構資料：靈活且多樣，應用於人機互動、社交網絡、媒體內容等。
3. 半結構資料：具有靈活性和組織性，常見於數據傳輸、API 等場景。每種類型的資料在處理上有不同的挑戰和機會，需要根據具體應用選擇合適的技術和工具。

二、區塊鏈基本觀念(本教材考題皆來自金融研訓院)

(一) 如何擁有比特幣

1. 用新台幣換取比特幣。
2. 自己當「礦工」，來獲取比特幣。

(二) 區塊鏈與比特幣

1. 「比特幣」是日本人「中本聰」創造的「虛擬貨幣」。
2. 「比特幣」是「區塊鏈」上最易懂與實際的應用，當比特幣被使用的國家與地方越多時，比特幣價值自然增加。

(三) 區塊鏈的由來

1. 「一筆交易=一個節點=記錄在一個區塊中」，即是一本「個人帳本」。
2. 然後再從這一本個人帳本更新交易記錄至所有帳本，並通知所有持有帳本的人。
3. 「一個區塊」代表「一筆交易(一個節點)」，一本「個人帳本(一個區塊)」中有「很多筆交易紀錄」，把一個區塊一個區塊「鏈」起來，就叫做「區塊鏈(一本總帳本)」。
4. 只要「持續開發區塊，更新帳本，並通知所有持有帳本的人」，就可以獲得比特幣，我們稱為「挖礦賺錢」。

(四) 區塊鏈應用-比特幣，特性如下：

- 1.「公開透明(不可被否認)，安全性高(不可被竄改)與去中心化」。
- 2.「每一節點(每一筆交易紀錄)」，進行「工作量證明的計算(用電腦破解數學難題)」，由「最快算出結果的節點來驗證交易」，這就是「取得共識的做法」。
- 3.交易具有「匿名性」，因為每筆交易僅以「錢包地址」作為發送與接收的地址。

(五) 區塊鏈實際交易流程說明

(參考資料來源 <https://www.ithome.com.tw/news/105373>)

1.產生一筆新交易

一筆新交易產生時，經由「數位簽章加密」，然後被廣播到區塊鏈網絡中，並與其它節點進行驗證。

2.各節點將數筆新交易

放進區塊中每個節點會將數筆未驗證的交易收集到區塊中，每個區塊可以包含數千筆交易。

3.決定由誰取得驗證交易

各節點進行「工作量證明的計算」，以「最快算出結果的節點(最先破解數學題者)」者，取得驗證交易，即是「共識機制」。

4.取得驗證權的節點將區塊傳送給所有節點

最快破解數學題的節點，會將自己的區塊傳送(告知)給其它節點。

5.各節點驗證並接上新區塊

其它節點會確認這個區塊的交易是否有效(經由數位簽章+非是重複交易)，接受該區塊，此時「區塊才正式連接成區塊鏈」，並且無法再被竄改。

6.交易驗證完成

所有節點一旦接受該區塊後，先前還沒破解數學題的工作區塊會失效，各節點會重新建立一個區塊，繼續等待下一次「工作量證明的計算」。

三、大數據

(一) 數據本質、特質與金融業應用大數據

1.數據本質：結構化與非結構化。

2.大數據特質(4V)

(1)Volume(大量) (2)Variety(多樣) (3)Velocity(不斷傳輸性、不停產生連續性)

(4)Veracity(真實性、清除不確定性)

3.金融業應用大數據的範例

(1)信用評分與風險評估

金融機構可以通過大數據技術對客戶的信用行為進行更全面的評估。除了傳統的信用記錄，大數據還能整合社交媒體活動、消費習慣、線上行為等多方面資訊，從而更精確地預測客戶的還款能力和風險等級。



(2) 詐欺偵測與預防

大數據能夠通過分析客戶的交易模式、地理位置和歷史數據，快速檢測出異常交易並及時預防詐欺行為。例如，當交易行為偏離客戶正常模式（例如在異常時間、異常地點進行大額消費）時，系統可以立即標記，並進行警示或暫停交易。

(3) 個性化金融產品推薦

金融機構可以根據客戶的歷史行為、消費模式和生活習慣，透過大數據分析來量身定制金融產品和服務。這不僅能提高客戶滿意度，還有助於提升交叉銷售和增加收入來源。

(4) 市場趨勢預測與投資決策

大數據可以整合全球經濟指標、股票市場數據、公司財報、社交媒體反饋等信息，從而幫助投資者和金融機構更準確地分析市場趨勢，進行智能投資決策。例如，透過社交媒體上對於某些公司的輿情分析，可以預測股價的波動。

(二) 大數據的應用層次、方向

1. 大數據分析的「應用層次」

(1) Foundational(基礎分析)

基礎分析是最簡單的數據分析層次，用於總結和描述現有數據，幫助理解過去發生的事情。例如：零售業使用基礎分析來匯總月銷售額、產品銷量等，了解不同產品的表現如何。

(2) Prescriptive(規範分析)

規範分析深入探究數據，找出生成某事件或趨勢的原因。它試圖回答「為什麼會發生」的問題。例如：在醫療領域，規範分析可以用來了解某些疾病在特定地區爆發的原因，分析環境、生活習慣等多個因素之間的相關性。

(3) Advanced, Predictive(進階預測)

進階預測是利用統計模型、機器學習等方法來預測未來可能發生的情況，旨在回答「將來會發生什麼」的問題。例如：金融機構利用進階預測分析客戶信用風險，通過歷史交易行為預測客戶未來違約的可能性。

2. 大數據分析的「應用方向」

(1) 客戶關係管理

大數據在客戶關係管理中的應用，主要是通過分析客戶的行為數據來了解他們的需求和偏好，從而提供個性化的產品和服務，提升客戶滿意度與忠誠度。例如：電子商務平台分析客戶的購買歷史、瀏覽行為和產品評價，提供精準的產品推薦，幫助提升銷售轉化率和增強客戶體驗。

(2) 建立 360 度客戶視覺圖

「建立 360 度客戶視覺圖」是指整合來自各種資料來源的客戶資訊，形成一個全面、全方位的客戶檔案，從而更深入地了解客戶的需求、偏好與行為。這種視覺圖像是客戶關係管理（CRM）中的重要工具，幫助企業更好地與客戶互動，提供個性化的服務與產品。



(3) 風險管理

大數據在風險管理中的應用有助於識別潛在風險，並通過預測和數據分析，減少財務損失或避免風險事件的發生。例如：銀行使用大數據分析客戶的信用紀錄、消費行為和社會活動來評估其還款風險，從而決定是否發放貸款，並提前預防可能的信用違約。

(4) 績效管理

大數據可以幫助企業監控和評估內部流程及員工表現，根據實時數據進行調整與優化，從而提升整體運營效率。例如：企業利用大數據工具監控生產線上的機器運行狀況和員工生產效率，識別低效環節，並及時調整資源分配，以提升生產績效。

3. 大數據分析的「資料方向」

(1) 傳統結構化的資料分析

大數據傳統結構化資料分析是指對高度組織化的、規範排列的資料進行處理和分析的過程。這類資料通常以行列格式存儲在關聯式資料庫中，資料具有固定的結構，如數字、文本、日期等類型的欄位。傳統的分析方法主要依賴於統計工具和查詢語言（如 SQL）來挖掘有用的資訊。例如：銀行使用結構化資料分析，來分析客戶的歷史交易數據。透過 SQL 查詢，它可以篩選出過去一年內的所有交易，並進行聚合計算，以了解某一段時間內的總交易金額、交易頻率等。

(2) 非結構化的資料分析

大數據非結構化資料分析是指對沒有固定結構或格式的資料進行處理和分析的過程。這類資料通常來源廣泛，包含文字、圖片、音頻、視頻等，無法直接用傳統的行列格式來表示，因此需要使用專門的技術和工具來處理。例如：社交媒體上的文字、圖片和影片屬於非結構化資料。企業可以通過自然語言處理（NLP）來分析這些社交媒體內容，了解用戶對某產品的情感態度，進行品牌監控或客戶反饋分析。

(3) 數位通路的資料分析

大數據的數位通路資料分析指的是通過數位渠道（如網站、移動應用、電子郵件、社交媒體等）收集和分析客戶互動、行為及交易數據，以提供更好的用戶體驗和商業決策。這些數據可以包括點擊行為、瀏覽時間、購物車操作、點擊的廣告等。例如：一家電商網站利用數位通路數據分析來了解用戶的瀏覽行為和購物車操作，通過分析哪些產品被頻繁瀏覽但未購買，來優化推薦算法或推出針對性的折扣活動，提升購買轉化率。

(三) 大數據的應用情境與情境行銷的建立

1. 大數據分析的「應用情境說明」

(1) 客戶視圖呈現

客戶視圖呈現是利用大數據技術將客戶的各種行為和屬性數據整合成一個全方位的視覺化檔案。這通常涉及來自不同來源的資料，包括購買歷史、社交媒體互動、瀏覽記錄等，為企業提供一個完整的、動態更新的客戶畫像。例如：銀行可以通過客戶視圖呈現，掌握某個客戶的借貸歷史、投資偏好以及消費習慣，從而提供量身定制的理財產品建議。



(2) 客戶區隔建立

客戶區隔建立是通過大數據分析，將客戶根據不同的屬性或行為特徵分成若干組。這些區隔可以基於年齡、購買行為、收入、興趣等，讓企業可以針對不同的群體採取差異化的行銷策略。例如：電商公司利用客戶區隔建立，把消費者分成高頻購物者、價格敏感型客戶和新品愛好者，針對不同群體推出不同的促銷活動或推薦商品，提升銷售轉化率。

2. 「情境式行銷」的建立

(1) 大數據的情境式行銷 (Contextual Marketing) 是指根據消費者在特定情境下的需求和行為，實時提供個性化的行銷信息和產品推薦。這種行銷策略充分利用大數據技術，分析顧客的行為、環境和情境，從而針對性地滿足他們的需求。

(2) 例如：餐飲業，當顧客走進某個咖啡店時，透過位置服務，商家可以立即發送優惠券或推薦特定的飲品，吸引顧客進店消費。例如：電商平台，在顧客瀏覽某類商品時，平台可以根據其過去的購買歷史或收藏商品，實時推送相關的優惠或類似產品推薦。

(3) 總結，情境式行銷利用大數據提供的即時和個性化資訊，能夠大幅提升行銷效果和客戶滿意度，促進銷售並增強品牌忠誠度。這種策略使企業能夠在合適的時間以合適的方式接觸顧客，實現更高的轉化率。

(四) 大數據的社群資料分析

大數據的社群資料分析 (Social Data Analytics) 是指通過收集和分析社交媒體平台上的數據，以了解用戶行為、情感和趨勢。這些分析幫助企業識別受眾需求、優化行銷策略，並增強與客戶的互動。

1. 特徵

(1) 多樣的資料來源

社群資料分析可以來自多種社交媒體平台 (如 Facebook、Twitter、Instagram、TikTok 等) 上用戶生成的內容，包括貼文、評論、分享、標籤和訊息等。

(2) 情感分析

利用自然語言處理技術，分析用戶對品牌或產品的情感反應，判斷正面、負面或中性的情感，以便調整行銷策略。

(3) 趨勢識別

通過分析熱門話題、標籤和趨勢，企業可以發現消費者的最新需求，迅速調整產品或行銷活動。

2. 例子

(1) 品牌監控：企業可以利用社群資料分析來追蹤用戶對某一新產品的反饋，及時了解消費者的看法，並根據反饋調整產品特性或行銷策略。

(2) 行銷活動效果評估：在推出行銷活動後，企業可以分析社交媒體上的討論和互動，評估活動的成功與否，並調整未來的行銷計劃。

3. 總結

社群資料分析幫助企業從大量社交媒體數據中提取有價值的見解，提升品牌知名度、加強客戶關係，並在市場競爭中獲得優勢。這種分析不僅能夠提供即時的反饋，還能幫助企業預測未來的消費趨勢。