

【金融科技知識】補充資料（三）

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」。

(六) 區塊鏈

在區塊鏈中，區塊是區塊鏈數據結構的基本單元。每個區塊由兩個主要部分組成：區塊頭(BlockHeader)和區塊主體(BlockBody)。

1. 區塊頭(BlockHeader)

區塊頭包含描述該區塊的元數據，類似於區塊的「標題」，其主要功能是協助區塊鏈驗證和追蹤區塊。區塊頭的核心元素包括：

- (1)前一個區塊的哈希值：用於鏈接到之前的區塊，形成區塊鏈的連續性和完整性。
- (2)默克爾樹根(MerkleRoot)：包含該區塊內所有交易的哈希摘要，能夠高效驗證區塊中的交易是否未被竄改。
- (3)時間戳(Timestamp)：記錄該區塊被創建的時間。
- (4)難度目標：顯示挖礦時解決該區塊所需的計算難度。
- (5)隨機數(Nonce)：挖礦時為了滿足工作量證明(ProofofWork)找到的數值，用於調整哈希值以符合難度要求。

區塊頭主要用於維持區塊鏈的安全性和連續性，通過哈希鏈結構使得任何對過去區塊的篡改會導致整條鏈上的區塊無效。



2. 區塊主體(BlockBody)

區塊主體包含實際的交易數據，即該區塊中記錄的所有交易。它的主要功能是記錄交易和其他區塊鏈上發生的數據變更。區塊主體主要由以下部分組成：

- (1)交易清單：區塊內記錄的所有交易數據。這些交易經過哈希計算後會生成默克爾樹，用於確認交易的完整性和不可篡改性。
- (2)其他數據：根據具體區塊鏈系統的不同，區塊主體中還可能包含其他額外數據，如智能合約執行結果等。

簡而言之，區塊頭的主要功能是保證鏈的連續性和區塊的完整性，而區塊主體則負責存儲具體的交易數據。

(七) 智能合約

區塊鏈中的智能合約是一種自動執行的程式，根據事先定義好的條件，當條件滿足時，它會自動執行相應的操作。智能合約的條款和條件以程式碼的形式直接寫入區塊鏈中，這使得它具有透明、不可更改和自動執行的特點。

1. 智能合約的主要特點

(1) 自動執行

一旦觸發條件滿足，智能合約就會自動執行操作，無需中間人。

(2) 不可更改

智能合約一旦部署到區塊鏈上，其內容就無法被篡改，保證了交易的可信度和合約的執行力。

(3) 去中心化

智能合約的執行不依賴於任何中心機構，所有參與者都能直接在區塊鏈上驗證執行結果。

(4) 透明性

智能合約的程式碼是公開的，所有參與者都可以查看合約內容，確保了透明性和可信度。

2. 主要應用場景

(1) 金融交易：智能合約可自動處理支付、貸款、保險賠付等操作。

(2) 供應鏈管理：追蹤產品的流轉並自動觸發付款或交貨條件。

(3) 去中心化應用(DApps)：例如去中心化的金融平台(DeFi)或預言機系統(Oracles)。智能合約的出現極大地提高了交易的效率，減少了人工介入的需求，同時降低了成本並增加了安全性。

(八) 公鑰與私鑰

在區塊鏈中，公鑰與私鑰是密碼學中的一對關鍵組合，用於保護資料安全和驗證交易真實性。它們是區塊鏈系統中確保身份驗證和數據加密的基礎。

1. 公鑰(PublicKey)

(1) 用途：公鑰是公開的，可以被任何人查看。它用於接收數位資產(如比特幣)，也是用來驗證數位簽章的工具。

(2) 功能：其他人可以使用你的公鑰來加密信息或資產，這些信息只能用對應的私鑰解密。

(3) 生成方式：公鑰是從私鑰通過單向加密演算法生成的，這樣即使知道公鑰，其他人也無法反推出私鑰。



2. 私鑰(PrivateKey)

- (1)用途：私鑰是個人保管的，必須保持秘密。它用來簽署交易並控制數位資產的使用權限。
- (2)功能：私鑰能用來解密用公鑰加密的信息，並且用來創建數位簽章，證明交易的真實性和所有權。只有持有正確私鑰的人才能發送數位資產或授權交易。
- (3)安全性：私鑰一旦泄露，任何人都可以控制對應的數位資產，因此私鑰必須妥善保管。

3. 公鑰與私鑰的運作方式：

- (1)A 用 B 公鑰加密，B 用自己的私鑰解密。
- (2)A 用 B 的私鑰簽章，B 用 A 的公鑰驗章。

(九) 區塊鏈中比特幣交易具有匿名性

這種匿名性實際上是一種偽名性，而非完全的隱匿。這意味著，在比特幣交易中，交易雙方的身份並不直接顯示，而是以公鑰(或比特幣地址)的形式進行標識。這使得比特幣交易不會直接揭示用戶的真實身份。

1. 身份隱匿：

比特幣交易中，沒有涉及用戶的真實姓名、地址或其他個人資料。每個用戶僅以一個或多個比特幣地址(由公鑰生成的字母和數字組合)來進行交易，從而保護了真實身份。

2. 公開透明：

儘管用戶身份是隱藏的，但比特幣的交易紀錄是完全公開且可追溯的。區塊鏈上每筆交易的詳情(如交易金額、發送方和接收方的地址等)都會被記錄並向所有人公開。雖然無法直接知道交易背後的身份，但交易記錄是永久保留且不可篡改的。

3. 偽名性：

然交易地址沒有與真實身份直接相關，但如果用戶通過某些途徑將自己比特幣地址與真實身份連接起來，例如在交易所進行身份驗證，則可以追溯其交易活動。因此，比特幣的匿名性是有限的，交易地址一旦與真實身份連接，相關交易行為也會被揭露。

比特幣交易的匿名性指的是用戶的身份並不在交易中顯示，僅使用地址來進行交易，從而隱匿了個人的真實身份。然而，這並不意味著交易完全不可追蹤，因為所有交易記錄都是公開的，而且一旦地址與真實身份建立聯繫，交易活動仍然可以被追溯。

(十) 區塊鏈的資訊解決方法與組成元件

區塊鏈的資訊解決方案透過去中心化、透明化和安全的技術，為許多行業提供可靠的數據管理、交易處理和智能合約自動執行等功能。這些解決方案通常由多個組合元件構成，這些元件共同合作以實現區塊鏈技術的應用。

1. 區塊鏈的主要資訊解決方案

(1) 分散式資料庫

區塊鏈本質上是一種去中心化資料庫，數據不是存儲在單一的中心化伺服器中，而是分散存儲在多個節點上，提升了資料安全性和容錯能力。

(2) 數位資產與加密貨幣



區塊鏈可以支持數位資產(如比特幣、以太坊等加密貨幣)的創建和交易，使得資產的所有權和交易過程透明化且不可篡改。

(3)智能合約

區塊鏈上可以執行智能合約，這些合約根據預先設定好的條件，自動執行交易和協議，無需第三方介入，從而提高效率並降低成本。

(4)供應鏈管理

區塊鏈可以追蹤產品在供應鏈中的流動，記錄每個階段的操作，確保產品的來源、品質和交付過程透明可靠。

(5)身份認證與管理

區塊鏈可以用來建立去中心化的身份認證系統，用戶的身份信息可以分佈式儲存，提升隱私性和安全性。

2.組合元件

區塊鏈解決方案的實現需要依賴多個核心元件，這些元件共同構成了區塊鏈的基礎架構：

(1)共識機制

決定區塊鏈中節點如何就區塊中的交易達成一致。常見的共識機制包括工作量證明(Proof of Work, PoW)和權益證明(Proof of Stake, PoS)。這些機制確保區塊鏈的數據一致性和安全性。

(2)密碼學技術

- a.公鑰與私鑰：用於加密交易和簽署數位資產，保障交易的安全性和用戶的身份隱匿。
- b.哈希函數：用來生成交易和區塊的唯一識別碼，防止篡改。

(3)P2P 網絡(點對點網絡)

區塊鏈依賴於 P2P 網絡來進行數據的傳輸和同步。每個節點相互連接，協作維護區塊鏈的完整性。

(4)分佈式賬本技術(DLT, Distributed Ledger Technology)

區塊鏈本身就是一種 DLT。它通過分佈式節點共同維護一個不可篡改的交易記錄。

(5)智能合約平台

例如以太坊，它允許開發者在區塊鏈上部署自動執行的智能合約，支持去中心化應用(DApps)的開發和運行。

(6)區塊鏈節點(Node)

節點是區塊鏈網絡中的計算設備，它們負責維護、驗證和記錄區塊鏈中的交易數據。節點可以分為全節點和輕節點，全節點會儲存整個區塊鏈的副本，輕節點則只儲存關鍵數據。

3.結論

區塊鏈的資訊解決方案基於去中心化的分佈式賬本、加密技術和共識機制，組合元件包括共識機制、密碼學技術、P2P 網絡、智能合約平台等。這些元件協同工作，支持不同場景的應用，從金融到供應鏈，提供安全、透明且自動化的解決方案。

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，找出發生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。