

Sustainability-Oriented Product Representation for Live-Streaming E-commerce via Semantic Normalization and Generative Modeling

Livestreaming e-commerce presents significant challenges for product representation because item information is conveyed through noisy spoken descriptions, informal naming, and inconsistent presentation styles rather than structured catalogs. To address this problem, this study proposes a modular item representation framework that converts unstructured livestream content into stable semantic representations. First, an unsupervised semantic perturbation method is used to extract salient product attributes from livestream transcripts and normalize them into compact semantic anchors. Next, these anchors are used to guide a LoRA-adapted large language model for controllable product description generation, improving attribute grounding while reducing hallucination and excessive promotional bias. Finally, the generated descriptions are encoded into a unified embedding space for similarity search and retrieval-based recommendation. Experiments on real-world livestreaming data show that the proposed method produces more stable and informative item representations across diverse streamer language styles. The results suggest that the framework provides a practical solution for improving product understanding and downstream recommendation in livestreaming e-commerce.

一、研究背景與動機

隨著數位科技與行動通訊的快速發展，直播電商（Live-streaming E-commerce）已逐漸成為新興且重要的商業模式之一。透過即時影音串流與互動式銷售，直播電商突破了傳統電商單向資訊傳遞的限制，使消費者能在觀看商品展示的同時與主播互動，進而提升購物體驗與消費參與度。然而，這種高度即時且互動性強的銷售模式，也使商品資訊的呈現方式產生了根本性的轉變。

在傳統電商平台中，商品資訊通常以結構化方式呈現，例如標準化的商品屬性欄位、規格表與描述文本，使得資訊易於管理與應用於搜尋與推薦系統。然而，在直播電商環境中，商品資訊多半來自主播的口語描述、即時互動對話以及情境式推銷內容，導致資訊呈現高度非結構化、語意不一致且含有大量冗餘與雜訊。不同主播可能對同一商品使用不同描述方式，甚至加入誇飾性語言，使得商品資訊難以被有效整合與理解。

此種資訊不穩定性將直接影響後續的人工智慧應用，例如商品推薦系統、語意搜尋系統與商品分類模型等。當商品屬性無法被準確擷取與建模時，系統將難以進行有效的商品匹配與使用者需求預測，進而降低整體電商服務品質。此外，資訊的不透明與不一致，也可能導致消費者在資訊不足或誤導的情況下做出購買決策，增加錯誤購買與退貨的風險。

從永續發展的觀點來看，此類問題不僅影響商業效率，更涉及資源使用與環境負擔。錯誤購買與退貨行為將增加物流運輸、包裝材料與能源消耗，進而提高碳排放與環境壓力。同時，低品質資訊也會導致系統進行大量無效運算與資料處理，造成數位資源的浪費。因此，如何提升直播電商中商品資訊的品質，使其能被有效轉換為可用的結構化知識，不僅是資訊技術上的重要課題，更具有促進永續消費與數位資源效率的關鍵意義。

基於上述背景，本研究嘗試從語意建模與生成模型的角度出發，探討如何在高度雜訊的直播環境中建立穩定且可信的商品表徵，並進一步提升資訊透明度與系統效率，為永續電商發展提供技術基礎。

二、研究目的

本研究旨在發展一套針對直播電商情境所設計之商品表徵方法，透過語意正規化與生成模型技術，將原本高度非結構化且雜訊豐富的直播內容，轉換為具備一致性與可解釋性的商品資訊表示。透過此方法，期望能有效提升商品資訊的品質，並改善相關應用系統之效能。

在具體目標方面，本研究首先著重於資訊品質的提升。透過語意抽取與標準化技術，將分散且不一致的商品描述轉換為統一的屬性表示，以降低語意歧異與資訊混亂的問題。此舉不僅有助於系統理解商品特性，也能提升消費者獲取資訊的清晰度。

其次，本研究致力於改善生成模型在電商應用中的可靠性。透過引入屬性導向的生成機制，使模型在生成商品描述時能依據已驗證的商品資訊進行內容產生，避免出現不合理推論或誇大描述，進而提升生成內容的可信度與實用性。第三，本研究將建立穩定的商品語意表徵，使其能夠有效應用於推薦系統與搜尋系統中。透過高品質的表徵，可提升系統對商品的理解能力，進而提高推薦準確性與使用者滿意度。

最後，本研究亦從永續發展的角度出發，探討資訊品質提升對於消費行為與資源使用效率的影響。透過改善資訊透明度與推薦效果，可降低錯誤購買與退貨率，進而減少物流與資源浪費，實現更具效率與責任的數位消費模式。

三、研究方法與架構

本研究提出一套整合語意處理與生成模型的商品表徵框架，整體流程可分為三個主要階段，分別為語意正規化、屬性導向生成與商品表徵學習。三者相互配合，形成一個由資料處理到語意建模的完整系統。

在第一階段，語意正規化主要針對直播逐字稿中的商品資訊進行處理。由於直播內容通常包含大量口語化表達、情緒性語句與重複描述，本研究透過語意分析與關鍵資訊擷取方法，篩選出具有代表性的商品屬性，並將其轉換為統一格式的語意表示。此過程的核心在於從高雜訊資料中提取高價值資訊，以建立穩定的語意基礎。

在第二階段，屬性導向生成則利用大型語言模型進行商品描述生成。與傳統自由生成不同，本研究將已篩選之商品屬性作為條件輸入，限制模型的生成範圍，使其內容能夠與實際商品資訊保持一致。透過此種方式，可有效降低模型產生錯誤資訊的機率，並提升描述的品質與可信度。

在第三階段，商品表徵學習將生成的商品描述轉換為語意向量，並進一步結合商品影像資訊，建立多模態商品表示。透過整合文本與視覺特徵，可使商品表徵更完整且具備更強的語意表達能力，進而提升其在推薦與搜尋任務中的表現。此外，本研究亦設計一套評估機制，從準確性、完整性與語意一致性等多個面向，全面評估商品表徵與生成內容的品質，以確保所提出方法在實務應用上的有效性。

四、研究創新與貢獻

本研究在方法設計與應用價值上具有多項創新與貢獻。首先，本研究針對直播電商中非結構化與高雜訊資料的特性，提出一套完整的語意正規化流程，能有效提升商品資訊的一致性與可用性，填補現有研究多集中於結構化資料的不足。

其次，本研究導入屬性導向生成機制，將生成模型的自由生成轉變為受控生成，使其能夠依據已驗證資訊產生內容，顯著降低生成錯誤與不實描述的問題，提升生成模型在電商應

用中的可靠性。第三，本研究整合多模態資訊進行商品表徵學習，使商品表示不僅包含語意資訊，也能反映視覺特徵，提升模型對商品的理解能力與應用範圍。

此外，本研究提出兼顧「資訊正確性」與「行銷說服力」的評估指標，使商品描述品質的評估更具全面性，能同時反映技術與實務需求。最重要的是，本研究將人工智慧技術與永續發展議題結合，強調資訊品質對於消費行為與資源效率的影響，拓展資訊管理研究在社會與環境層面的應用價值。

五、與永續發展之關聯性

本研究與永續發展目標具有高度關聯，特別是在數位經濟快速發展的背景下，資訊品質與資源效率之間的關係日益重要。首先，在責任消費與生產方面，本研究透過提升商品資訊的準確性與透明度，使消費者能基於完整且可靠的資訊進行決策，降低因資訊不對稱所導致的錯誤購買與退貨行為。此舉有助於減少物流運輸、包裝浪費與能源消耗，進而降低環境負擔。其次，在產業創新與數位基礎建設方面，本研究透過人工智慧技術提升電商系統的資訊處理能力與推薦效率，促進智慧電商的發展，並提升整體產業競爭力。

此外，本研究亦符合綠色人工智慧（Green AI）的概念。透過提升資訊品質與模型效率，可降低不必要的計算與資料處理需求，減少能源消耗與碳排放，使人工智慧技術的發展更加符合永續原則。整體而言，本研究不僅著重於技術創新，更強調其在促進永續消費、降低資源浪費與提升系統效率上的長期影響。

六、預期成果

本研究預期能建立一套高品質且具實務應用價值的商品表徵方法，使直播電商中的非結構化資料能夠被有效轉換為穩定且具語意一致性的商品資訊。在技術層面，本研究可提升商品屬性擷取的準確性與生成內容的可信度，並改善推薦與搜尋系統的效能。在應用層面，本研究可協助電商平台提升資訊品質與使用者體驗，並降低錯誤推薦與資源浪費。此外，本研究亦預期能產出具學術價值之研究成果，並為未來相關研究提供方法與評估基準。

七、未來發展與應用

本研究可進一步應用於實際直播電商平台，進行大規模驗證與系統整合，以評估其在真實環境中的效益。此外，可擴展至跨語言與跨平台之商品表徵研究，以提升方法的泛用性。在技術發展方面，未來可進一步整合語音、影像與文本等多模態資料，建立更完整的商品理解模型。同時，亦可結合永續指標（如碳足跡或環保屬性），發展具永續導向的推薦系統，引導消費者做出更環境友善的選擇。長期而言，本研究可作為智慧電商與永續數位經濟的重要基礎，促進科技發展與環境保護之間的平衡。

八、研究對數位永續與綠色人工智慧之貢獻

隨著人工智慧技術在各產業中的廣泛應用，其所帶來的能源消耗與計算成本問題亦逐漸受到關注。近年來，「綠色人工智慧」（Green AI）逐漸成為重要研究方向，強調在提升模型效能的同時，也需兼顧能源效率與環境影響。然而，目前多數研究主要著重於模型架構優化或計算資源節省，較少從「資訊品質」的角度切入。

本研究從另一個面向探討人工智慧的永續性，即透過提升輸入與輸出資訊的品質，減少不必要的計算與資料處理需求。在直播電商情境中，若商品資訊本身充滿雜訊與不一致性，系統將需要進行更多次的運算與修正，才能達到可用的效果。相對地，若能在前端即建立穩定且高品質的商品表徵，將可大幅降低後續模型處理的複雜度與計算成本。

此外，本研究透過屬性導向生成機制，使模型生成內容更具精準性與一致性，避免產生大量冗餘或錯誤資訊。這不僅提升了生成效率，也降低了因資訊錯誤而導致的重複運算與系統負擔。從整體系統角度來看，此種「以品質提升取代計算堆疊」的策略，符合綠色人工智慧所強調的高效率與低資源消耗原則。因此，本研究不僅在應用層面改善電商系統效能，也在方法論上提供一種新的永續思維，即透過優化資訊本身，達到降低能源消耗與提升系統效率的雙重目標。

九、對消費行為與社會影響之探討

直播電商的興起，不僅改變了商品銷售模式，也對消費者行為產生深遠影響。由於直播環境強調即時互動與情境帶動，消費者在觀看過程中容易受到情緒與氛圍影響，進而產生衝動購買行為。在資訊不完整或描述誇大的情況下，此類決策更容易偏離實際需求，導致購買後不滿意甚至退貨的情形。

本研究透過提升商品資訊的準確性與一致性，使消費者能在更清楚且透明的資訊基礎上做出決策，有助於降低衝動消費與資訊誤導的影響。當商品描述能夠真實反映其特性時，消費者對商品的預期將更接近實際體驗，進而降低退貨率與資源浪費。此外，本研究亦有助於建立更健康的數位消費環境。在資訊透明且可靠的情境下，消費者將更傾向進行理性比較與選擇，而非依賴情緒或誇飾性語言做出決策。長期而言，此種轉變有助於推動「負責任消費」的形成，並促進電商產業朝向更永續的發展方向。從社會層面來看，本研究亦可提升消費者對平台與資訊的信任度，進而強化數位經濟的穩定性與長期發展潛力。

十、研究限制與風險評估

儘管本研究提出一套完整的商品表徵方法，仍存在若干限制與潛在風險需進一步考量。首先，在資料來源方面，直播內容具有高度變異性，不同平台、主播風格與商品類型皆可能影響語意表達方式。因此，本研究方法在不同場景中的泛用性仍需透過更多實驗驗證。其次，在生成模型應用方面，雖然透過屬性導向機制可降低錯誤生成的機率，但仍無法完全避免模型在特殊情境下產生不符合預期的內容。因此，未來仍需進一步強化模型的可控性與穩定性。

此外，在多模態表徵方面，影像與文本之間的語意對齊仍具挑戰性，特別是在商品展示方式不一致或影像品質不佳的情況下，可能影響表徵效果。在永續層面上，本研究主要從資訊品質與系統效率進行分析，尚未直接量化其對碳排放或資源消耗的具體影響。因此，未來可進一步結合實際數據進行更深入的環境效益評估。透過對上述限制的分析，本研究可作為後續研究改進與延伸的基礎。

十一、研究時程規劃

本研究預計分為四個階段進行，確保研究進度與成果品質。第一階段為資料蒐集與前處理，將收集直播電商相關資料，並進行語音轉文字與資料清理，以建立可用的研究資料集。接著，第二階段為語意正規化與屬性抽取模型開發，設計並實作商品屬性擷取與語意標準化方法，建立初步商品表示。而第三階段為生成模型與商品表徵學習，導入大型語言模型進行屬性導向生成，並建立多模態商品表徵。最後，第四階段為系統整合與評估分析，透過多項指標評估方法效能，並分析其在推薦系統與永續應用上的潛在影響。透過分階段規劃，可確保研究能有系統地推進，並逐步驗證各模組的有效性。

十二、結論

本研究針對直播電商中商品資訊高度非結構化與語意不穩定的問題，提出一套結合語意正規化、生成模型與多模態學習的商品表徵方法。透過提升商品資訊的準確性與一致性，本研究不僅改善了電商系統的效能，也為消費者提供更透明且可靠的資訊基礎。

更重要的是，本研究將資訊技術與永續發展議題結合，從資訊品質的角度出發，探討其對於消費行為與資源使用效率的影響。透過降低錯誤資訊、提升推薦準確性與減少不必要的資源浪費，本研究展現人工智慧在促進永續消費與綠色數位經濟中的潛在價值，隨著直播電商持續成長，本研究所提出的方法可進一步擴展至更大規模的應用場景，並結合更多永續指標，推動電商產業朝向更高效率與更具責任的發展方向。