

揭開人工智慧面紗

機器學習開創

產品創新預測的新時代

July 2022

AUTHORS

Colin Ho, Ph.D.

Jiongming Mu

Rich Timpone, Ph.D.

Camille Souchal



**IPSOS
VIEWS**

GAME CHANGERS



人工智慧 (AI) 的應用近年來逐漸普及，語音和臉部辨識等功能皆已內嵌在行動電話、電視、汽車及其他消費性產品中 (如 Amazon Alexa)。演算法使機器人能夠執行任務，幫助我們挑選禮物及找到感興趣的地點。PWC 的一份研究報告預估，AI 將在 2023 年前提高全球經濟產出至少 15 兆美元，此數字遠遠超過目前中國大陸和印度的經濟產值總和。(註1)

在此趨勢下，我們也開始看見 AI 技術能如何同樣為市場研究帶來革命性地改變，創造更快速、更便宜且更好的結果。然而，對 AI 的定義眾多繁雜，常使人難以釐清。加上有些人利用 AI 定義的模糊性從中獲益，導致情況變得更加複雜。(註2)

AI 難以被理解的其中一個原因在於它涵蓋了範圍極廣的多種科技與應用。25 年前，由於當時 AI 覆蓋許多不同領域，包含：(1) 人類知識及智慧的研究與應用、(2) 人工智

慧的研究與開發、(3) 利用人工智慧演算法解決實際問題(註3)，我們意識到，多樣性是造成人們無法真正理解 AI 的主因。自那時起，上述最後一種類別的應用透過各種不同的形式急遽發展，其中機器學習是最為人所知的方法，並已廣泛地運用在我們生活的各個角落。

之後我們將在其他白皮書中另外對如何評估這些方法的準確性進行定義與討論，並探討 AI 的道德倫理問題。本文中，我們聚焦在人工智慧及機器學習 (AI/ML) 的模型建立上，瞭解如何使用消費者語言來準確預測創新產品表現。雖然演算法的性質與品質很重要，但傳統市場研究中重視的資料代表性、品質管控與相關性亦不可或缺，它們能解釋為什麼結合人類智能與人工智慧對於使用人工智慧及機器學習技術預測創新產品的潛力來說如此重要。

什麼是預測分析？ 它與人工智慧/機器學習有何關聯？

預測分析 (Predictive Analytics) 利用統計、機器學習及演算法，基於歷史和/或現有資料對**未來結果**進行預測。除了預測未來，預測分析有時還會用在模型中，瞭解造成特定結果的**驅動因素**，使行銷人員能更有效益地分配資源來創造他們想看見的未來。預測模型可以使用多種不同資料來建立 (如時間序列資料、社群網路資料、調查資料)。進行預測分析前，我們首先要問自己兩個問題：預測模型將如何使用 (例如，我們想預測什麼結果？我們需要瞭解驅動因素嗎？) 以及哪些資料來源可以用來建立模型。

雖然有些人會將預測分析定義為「從經驗 (資料) 中學習的技術」(註4)，但其實預測方法早已存在許久。然而，用來進行預測的市場研究方法卻往往不被視為是使用預測分析，這是因為市場研究和人工智慧使用的術語不太一樣 (例如 forecasting 和 predictive analytics 的差別)，即使是研究中使用 AI/ML 技術的情況下亦是如此。(註5)

平心而論，預測分析所使用的技術已有了長足發展，特別是可用於預測分析上的資料源 (例如銷售資料、社群網路資料、行為資料、影像資料) 以及可使用的演算法 (例如機器學習和深度學習演算法，其部分目標與一般統計方法不同) 數量皆大幅增加(註6)。同時，運算能力的進步也使大規模將預測模型應用在大數據上變得可能。這些技術皆被用在市場研究，且也是益普索許多服務的基礎 (例如，透過社群網路資料來預測市場趨勢、使用影片標籤來預測廣告表現等等)。

本文不對預測分析理論加以說明，而是專注探討「如何利用機器學習來預測產品創新是否能夠獲得成功」。我們透過具體的案例來說明並避免不必要的專業術語，以針對這個主題提供更清楚的解釋。正如我們所見，訓練預測分析模型時，資料性質的選擇極為關鍵，其重要性也將展現在益普索的產品創新成功預測模型如何幫助應對實際遇到的挑戰裡。

訓練人工智慧/機器學習模型以預測產品創新的成功機率

訓練資料對機器學習的重要性

如果我們想教孩子認識手機，我們會將手機的外觀展示給他們看。更好的方法是，提供他們手機和非手機的實際範例，讓他們學會辨識手機獨有的特徵，例如矩形的玻璃螢幕、位於背面的相機鏡頭等。藉由這樣的訓練，孩子可以舉一反三，知道如何辨識所有手機，而不只是認識曾出示給他看過的範例（詳見圖1）。

訓練一個能預測產品創新成功機率的人工智慧/機器學習模型（以下稱為 AI/ML 模型）在概念上與教導孩子辨識手機相仿。我們需要同時提供產品創新成功及失敗的實際案例，AI/ML 模型才能學習辨識產品創新成功需具備什麼樣的特徵。在獲得足夠的範例之後，這個經過訓練的模型就可以開始對它從未看過的創新產品進行成功機率預測了。訓練是所有預測模型的基礎，因此訓練資料的品質和相關性將決定預測模型的好壞。

提到預測產品創新的成功機率時，大多數人會先想到將已被驗證過的產品概念作為訓練

如果想要 AI 學會辨識人臉，我們會提供 AI 模型一些臉部範例。雖然這種模式可以套用於許多地方，卻不適用在建立產品創新預測模型上。 ”

資料。這有可能是人們受到大眾媒體傳遞的資訊影響。舉例來說，通常我們會提供 AI 會 AI 雖然這種模式可以套用在許多地方，卻不適用在建立產品創新預測模型上。與其他領域不同，我們想預測的事物並非固定不變（例如，下棋的規則和棋子本身不會改變），或者是發生變化的速度足夠慢時，訓練資料的更新會比較準確。然而產品創新預測，就是要預測**現今尚不存在的事物**。

正如我們接下來將談到的，這並不代表預測分析無法被拿來應用，而是在使用時，所選擇的訓練資料極為關鍵。然而在這樣的情況下，許多可用資料源並無法應對動態環境。考量到訓練資料對於建立預測模型的重要性，我們將對此做進一步說明。

圖1 從過去經驗中學習如何幫助預測未來，特別是在未來與過去相似的情況下



過去未必能幫助預測未來

在教導孩子辨識物品的例子中，我們預設現在與未來手機的外觀相仿。但當新推出手機的外觀與過去完全不同時（例如 Apple Watch 這種沒有數字鍵盤、佩戴在手腕上的手機），以過去辨別未來將毫無幫助（詳見圖2）。

訓練資料與未來新品的相關性對於利用 AI/ML 模型預測產品創新成功機率來說至關重要。想像一下，我們將已測試過的产品概念作為訓練資料建立一個 AI/ML 模型，以預測全新個人清潔產品取得成功的可能性。為能更具體說明，接下來我們將以 1800 年代後期推出並獲得成功的美國香皂品牌 Ivory 為例。

Ivory 香皂的其中一個特點在於它能漂浮在水上，這對於泡澡的人來說極具吸引力，因為如此一來就不需要在浴缸底部四處摸索找尋肥皂。然而，由於淋浴現在更為普遍，Ivory 香皂的這個特點已不再有優勢。

簡單來說即是在產品創新領域，**過去的產品特色不一定能用來預測未來能取得成功的產品概念**。以往成功的產品及其優勢，在未來可能不再是人們重視的要素。隨著時間推移，市場上相繼推出的新產品改變了人們對產品的期望。想想看手機在過去十年間的變化，我們期待手機應具備的功能也隨之改變。當未來的變化性高時，通常不建議使用過去甚至是現下的產品概念來訓練預測模型。

圖2 過去經驗不一定能幫助預測變化性高的未來

將手機的外觀
展示給孩子看

無法辨識出一款
外觀完全不同的
未來手機



瞬息萬變的世界

不只有產品會改變並導致消費者期待發生變化，整個世界也可能會隨時間變遷，而使消費者重視的產品特色變得不同或轉而找尋新的產品好處。我們只需觀察氣候變遷、新冠疫情及科技發展，就能看到我們的世界如何在各個不同領域裡出現巨大改變，進而改變消費者的偏好。

近年來，永續產品變得越來越受到歡迎。有鑑於永續概念的發展還很新，若使用過去五年內測試過的产品概念來訓練模型以預測永續產品是否會成功，將會面臨到樣本不足的問題。新冠疫情則是另一個改變世界的重大事件，人們重新思考並調整健康的重要性，特別是免疫力方面。雖然在五年前，免疫力也同樣是受歡迎的一個產品好處，但疫情使免疫力的重要程度大幅提升。因此，以過去五年的產品概念為基礎訓練出來的預測模型，很有可能會低估具備提升免疫力的產品好處現今對消費者的重要性。最後，想想在這個科技進步的數位時代，快消品（CPG）如何日新月異。舉例來說，現在有透過在網路上回答一系列問題來進行客製化的個人化維他命（例如 Care/of），還有使用擴增實境技術來打造刷牙體驗的兒童牙刷（例如高露潔 Magik 魔法泡泡互動式兒童牙刷）。而在

過去五年所累積的快消品產品中，卻沒有任何運用數位科技的個人化或擴增實境產品的例子。

我們可以從新冠疫情初期看到快速變化的世界對 AI 模型的實際影響。人們的行為發生劇烈改變，從健康到購物行為的既有 AI/ML 模型因而失效了。在疫情爆發前訓練好的模型已無法在這個變化後的世界進行準確預測。^(註7)

本文重點在於，我們無法以現有的產品概念來預測世界或產品的未來發展。使用過去幾年間甚至時下的產品概念作為訓練範例，可能會導致預測模型完全忽略新的產品好處（例如個人化）或者是高估/低估因環境背景變遷而改變的現有好處的重要性（例如免疫力）。以產品概念為基礎的預測模型可能會在前一兩年內表現優異，但很快就會隨著市場的快速變化而過時。**這種模型只對那些與市場上已有產品相似的產品有效**，而不適用在完全推翻現況的創新產品概念。數位化、疫情和氣候變遷無不體現了這個世界以及人們的需求可以多麼迅速地轉變。

消費者看的是整體產品概念，而非個別元素

消費者針對產品創新的整體概念做出評價，而非產品的個別元素（例如洞察、好處、可信度、價格）。以獨立產品特點及好處為基礎的預測模型，較難準確地掌握一個產品真正吸引消費者的關鍵因素。創新產品概念中的各個元素時常相互影響，進而改變消費者對整體產品的評價。也就是說，消費者對新產品的反應並非只是個別元素的總和，而需更進一步考慮到每個元素之間的交互作用。

舉例來說，基於概念的預測模型可能會判斷當產品描述包含「酪梨」或「辣椒醬」的詞彙，將在食品概念中獲得更高的分數。如果兩者都存在於同一個產品時，過於簡單的預測模型可能會告訴你這個產品概念將能獲得極為正面的消費者回饋！但在現實生活中，如果出現寫著「酪梨辣椒果昔」的產品，消費者更有可能會給予負面評價！消費者看的是整體產品吸不吸引人，而不只是其個別組成元素。

謹慎地在產品創新測試中使用預測分析

評估學習效果的準確度

孩子在初次學習辨識手機時，若遇到先前沒看過的手機便很有可能會出錯，AI/ML 模型也是如此。預測模型的分析結果並非百分之百準確。如果要使用預測模型，我們就要做出取捨。例如接受在預測未來可能結果上的準確度降低，但是獲得結果的速度較快且成本較低。評估 AI/ML 模型的準確度對於判斷該模型是否適用於特定產品概念來說極其重要。有鑑於預測模型並非完美無缺，我們建議在早期概念發想階段就使用產品創新測試預測模型，幫助品牌在眾多產品概念中進行篩選，以進一步做更嚴格的产品測試。

AI/ML 模型會使用一組未被用來訓練模型、但其市場結果我們已經知道的範例來進行驗證。這代表我們能運用這個 AI/ML 模型來評估創新產品概念，預測模型從未見過

這些產品，但我們已經知道這些產品是否能在市場上取得成功。若模型的預測分析符合實際結果，預測準確度即為 100%。在實務上，大多數產品創新預測模型的準確度都落在 60% 至 80% 的區間。對準確度的可接受範圍因人而異，一般會權衡做出錯誤決策的潛在風險以及使用預測模型能帶來的好處。

需特別注意的是，對於產品成功或不成功的定義往往較為主觀（例如與期望結果、市場平均表現或競爭對手相比等）。無論採用何種評估指標，使用者都應該要確實瞭解該預測模型的準確度、它的計算方式以及其預測結果是基於何種指標。

評估 AI/ML 模型的準確度
對於判斷該模型
是否適用於特定產品概念
來說極其重要。 ”

一個模型只能對一種產品概念進行預測

回到前面孩童辨識手機的例子，孩子只學會辨識手機，但無法分辨其他物品（例如咖啡機、冰箱），我們所學習的很多東西都是具體明確的。AI/ML 也一樣，它目前的發展主要集中在弱人工智慧（或稱狹義人工智慧）領域，其演算法僅能用來處理單一任務。因此，像是圖像辨識、國際象棋及語音辨識等 AI 應用，都是使用不同範例來訓練並分開進行驗證。即使是使用了相同演算法的類似任務（例如 AI 學習如何製作林布蘭和莫內的圖像）亦是如此。對於訓練機器學習模型時的限制同樣適用在產品創新預測上。

「你只能預測你用來訓練模型的內容」的原則對於產品研究中的 AI/ML 應用具重大影響。如果我們建立的模型是用來預測飲料的成功，就不該用它來預測家庭清潔用品。而當 AI/ML 模型是用以預測飲料在美國市場是否能夠成功，我們就不應該用它來

對中國市場的飲料進行預測。這是因為每個產品類別對於成功的預測因子都不一樣，一個國家消費者在意的因素，對另一個國家的消費者來說可能並不重要。

AI/ML 預測模型的使用者需要知道模型使用了哪些訓練範例，以及模型的預測因子是否符合自身所需。瞭解這些能幫助使用者評估模型可應用的範圍。當用途改變或使用在不同情境時，訓練好的模型需要重新再被訓練一次。我們不能期待專為執行單一任務的模型可以在不同背景脈絡下仍表現優異，人工智慧和機器學習並非萬能。

雖然我們也可以建立更為通用的 AI/ML 模型來根據品類和地理差異進行預測，但處理這種異質性問題所需的相關資料量往往遠超過實際可用的數量。使用數量不足或較不相關的資料，將降低產品概念預測結果的準確度。



圖像辨識、國際象棋及語音辨識等 AI 應用，都是使用不同的範例來訓練並分開進行驗證。即使是使用了相同演算法的類似任務亦是如此。 ”

益普索的產品創新預測模型

從前面針對基於產品概念的模型的討論中，不難猜出我們並不把對產品概念的描述作為預測的基礎。相反地，我們使用**消費者對產品概念的逐字回應**來進行預測。基本上，消費者會被展示一個新的產品概念，然後被要求回答一個開放式問題以瞭解他們看見該產品概念時，**最先浮現在腦海中的想法**。這些針對產品概念的即時反應將被我們用在益普索的預測模型中。

我們使用消費者的反應來訓練模型並對新產品是否能成功進行預測。我們相信，消費者的直覺反應是對創新產品最有利且穩

定的預測因子。消費者對於新產品的反應不太可能隨時間改變。十年後，消費者依然會在喜歡新產品時回以正面評價，不喜歡時給予負面評價，或是在沒有興趣時表現出無動於衷的態度。以消費者的即時回應來建立 AI/ML 預測模型，有一點要注意：我們需要蒐集到消費者的即時反應，才能對新產品概念的未來表現進行預測。

訓練資料

截至 2022 年 2 月為止，我們已蒐集到來自超過 60 個國家、7 個主要品類（人類食品、飲料、醫療保健、居家清潔、個人護理、美妝及寵物用品）約 400 萬個消費者的即時回應。這個資料庫提供我們訓練預測模型所需的資料，並確保預測結果的品質穩定。資料庫中的資料會隨著新的研究出現而自動增加，使我們能定期更新模型。

我們使用兩組資料來建立預測模型，分別是 (1) 消費者對新產品概念的第一反應以及 (2) 針對相關性、昂貴程度及差異化 (RED) 這三個關鍵指標，消費者會選擇新產品還是原有產品。^(註8)基本上，消費者對新產品概念的即時反應將被用來預測他們的選擇。

消費者的第一反應是從他們觀看新產品概念後立即回覆的一個開放性問題所獲得，且是未經過處理或潤飾的原始資料。舉例

來說，針對食品概念的回應可能會包含正面回饋如「美味」或「令人垂涎三尺」，或者也有可能會出現「噁心」、「不是我的菜」等負面評價。這些回應被用來對消費者是否會在 RED 這三個關鍵指標上選擇新產品進行預測。人們如何回應總能透露出他們對新產品的正面或負面態度，因此這些被用來訓練模型的範例將保持與未來的相關性，且預測模型可以預測創新產品未來是否成功而不會過時。

我們理解有些企業較偏好不需要消費者參與的預測模型，但出於前面所討論過的顧慮，仍需要一定限度的消費者參與。只需要一個開放式問題，就能夠快速且有效地篩選出具備潛力的產品創新。雖然此方法有其限制，我們仍會持續跟進評估，以瞭解產品概念是否能增進模型在產品開發後期階段的預測能力，而這將是建立在消費者即時回應的基礎上更進一步的優化。

人們如何回應總能透露出他們對新產品的正面或負面態度。 ”



準確性和特定性

我們的預測模型是根據三個 RED 指標進行驗證。也就是說，我們將模型的預測結果與在過去的概念測試中，受訪者的實際選擇進行比較。舉例來說，如果在新產品和現有解決方案之間做選擇時，有 100 名受訪者選擇了新產品，而我們的模型預測則有 75 名受訪者選擇新產品，25 名選擇現有解決方案，那麼準確率就是 75%。根據我們迄今為止建立的模型，對 RED 指標的預測準確率大概落在 70-75% 的範圍內。由於早期篩選階段的目標是確保勝出的產品概念能進入下一階段進行更嚴格的測試，這樣的準確度是可以被接受的，它能帶來更高的敏捷度及更大的成本效益。

需特別注意的是，若使用預測模型來預測不同品類/國家的產品概念，將降低預測準確度。舉例來說，美國食品模型的預測準確度，單指該模型對新美國食品概念的預測能力。如需建立範圍更廣的預測模型，例如美國以外的國家，則建立一個全球性的食品預測模型可能會更好，如此一來便能做更廣泛的應用。全球食品模型的準確度也許會略低於僅限美國市場的模式，但它能夠對多個不同國家進行預測。

結語

我們試著以簡單易懂的語言解釋如何開發 AI/ML 模型並將其應用於預測產品創新的成功。我們也明確指出，預測分析結果在很大程度上取決於訓練資料與未來的相關性。為能更聰明地運用這些模型，需要瞭解其中將面臨到的取舍（例如準確度與成本之間的平衡），以及由於可用資料不足而導致模型的侷限性。

將 AI/ML 模型用於產品創新的預測，可以使產品創新測試的速度更快、成本更低，將專案時程從數週或數月縮短至數小時或數天。但就像其他研究或分析方法一樣，我們需要對其進行深入瞭解以確保正確使用和應用。因此，我們專注在產品創新測試中結合人工智能和機器學習的增強智能（Augmented Intelligence）以實現最大效益。

機器學習的時代已經來臨，讓我們介紹如何在市場研究中適當地應用它。如果您曾經投資過或曾閱讀投資公司的招股書，您可能會注意到「過去表現不代表未來結果」的提醒。在使用過去產品的表現來預測未來產品的成功時亦是如此。建立一個預測模型來決定今天尚不存在的產品成功與否，與建立一個 AI/ML 模型來下棋、開車或推薦您可能喜歡的電影截然不同。

參考資料

1. Horgan, J. (2020). Will Artificial Intelligence Ever Live Up to Its Hype? Replication problems plague the field of AI, and the goal of general intelligence remains as elusive as ever. Scientific American online <https://www.scientificamerican.com/article/will-artificial-intelligence-ever-live-up-to-its-hype/>
2. Olson, P. (2019). "Nearly Half of All 'AI Startups' Are Cashing in on Hype." Forbes <https://www.forbes.com/sites/parmyolson/2019/03/04/nearly-half-of-all-ai-startups-are-cashing-in-on-hype/#5635ecf1d022>
3. Taber, C.S. and Timpone, R.J. (1996). Computational Modeling. Sage University Paper Series on Quantitative Applications in the Social Science, series no. 07-113. Thousand Oaks, CA: Sage
4. Siegel, E. (2013). Predictive Analytics. Hoboken, NJ: John Wiley & Sons.
5. Venugopal, V., & Baets, W. (1994). Neural Networks and Statistical Techniques in Marketing Research: A Conceptual Comparison. Marketing Intelligence and Planning, 12 (7), pp 30-38
6. Agrawal, A, Gans, J., and Goldfarb, A. (2018). Prediction Machines. Boston, MA: Harvard Business Review Press
7. Heaven, WD. (2020). "Our weird behavior during the pandemic is messing with AI models," MIT Technology Review <https://www.technologyreview.com/2020/05/11/1001563/covid-pandemic-broken-ai-machine-learning-amazon-retail-fraud-humans-in-the-loop>
8. Ho, C., Mu, J., Brown, A. (2019). Speed Dating with Innovations: What's Your Innovation Pick-Up Line? Ipsos POV. <https://www.ipsos.com/sites/default/files/ct/publication/documents/2019-06/innovation-speed-dating-ipsos-views.pdf>

揭開人工智慧面紗

機器學習開創產品創新預測的新時代

AUTHORS

Colin Ho, Ph.D.

Chief Research Officer, Innovation and Market Strategy and Understanding, Ipsos

Jiongming Mu

Global Innovation Solutions Lead, Ipsos

Rich Timpone, Ph.D.

Managing Director, Global Science Organization, Ipsos

Camille Souchal VP,

Innovation Technologist, Ipsos

The **Ipsos Views** white papers are produced by the **Ipsos Knowledge Centre**.

www.ipsos.com

@Ipsos

GAME CHANGERS

