



(19)中華民國智慧財產局

(12)發明說明書公告本

(11)證書號數：TW I772287 B

(45)公告日：中華民國 111 (2022) 年 08 月 01 日

(21)申請案號：106102678

(22)申請日：中華民國 106 (2017) 年 01 月 24 日

(51)Int. Cl. : G06Q10/06 (2012.01)

G06Q30/02 (2012.01)

(30)優先權：2016/01/29 中國大陸

201610065638.2

(71)申請人：香港商阿里巴巴集團服務有限公司 (香港地區) ALIBABA GROUP SERVICES LIMITED (HK)

香港

(72)發明人：姜曉燕 (CN)；代斌 (CN)；楊旭 (CN)；褚歲 (CN)；趙耀 (CN)

(74)代理人：林志剛

(56)參考文獻：

TW 201513019A

CN 101986298A

CN 104572734A

CN 105095477A

CN 105243389A

CN 105279288A

US 2015/0088850A1

審查人員：許人偉

申請專利範圍項數：6 項 圖式數：3 共 23 頁

(54)名稱

問題推薦方法及設備

(57)摘要

本發明公開了一種問題推薦方法及設備，該方法包括：在樣本採集週期內，獲取問題並獲取所述問題對應的問題特徵；對所述問題特徵進行處理，處理後的問題特徵在預設的數值區間內；根據各個所述問題及其在所述問題中的第二機率和指定的推薦閾值確定推薦的問題；其中，各個所述問題及其在所述問題中的第二機率是透過所述處理後的問題特徵和第一機率得到的；所述第一機率是透過所述問題特徵得到的。應用本申請的方法，透過對以往問題特徵計算得到需要向使用者推薦的問題，提高了向用戶推薦問題的準確率。

指定代表圖：

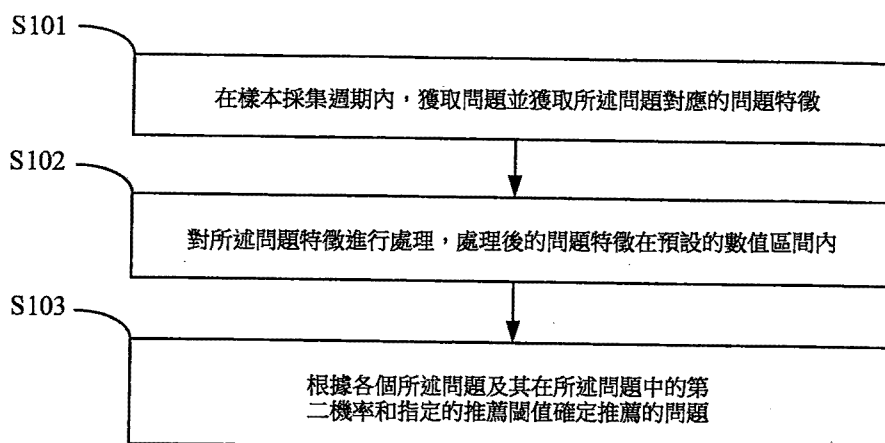


圖 1

I772287

發明摘要

※申請案號：106102678

※申請日：106 年 01 月 24 日

※IPC 分類： G06Q 10/06 (2012.01)
G06Q 30/02 (2012.01)

【發明名稱】(中文/英文)

問題推薦方法及設備

【中文】

本發明公開了一種問題推薦方法及設備，該方法包括：在樣本採集週期內，獲取問題並獲取所述問題對應的問題特徵；對所述問題特徵進行處理，處理後的問題特徵在預設的數值區間內；根據各個所述問題及其在所述問題中的第二機率和指定的推薦閾值確定推薦的問題；其中，各個所述問題及其在所述問題中的第二機率は透過所述處理後的問題特徵和第一機率得到的；所述第一機率は透過所述問題特徵得到的。應用本申請的方法，透過對以往問題特徵計算得到需要向使用者推薦的問題，提高了向用戶推薦問題的準確率。

【英文】

【代表圖】

【本案指定代表圖】：第(1)圖。

【本代表圖之符號簡單說明】：無

【本案若有化學式時，請揭示最能顯示發明特徵的化學式】：無

發明專利說明書

(本說明書格式、順序，請勿任意更動)

【發明名稱】(中文/英文)

問題推薦方法及設備

【技術領域】

本申請關於通信技術領域，特別關於一種問題推薦方法，本申請同時還關於一種問題推薦設備。

【先前技術】

隨著雙十一、雙十二等電商節日的興起，越來越多人開始網上購物，但“節日經濟”對電商也造成了雙重衝擊即銷量和客服的壓力暴增。企業的客服一般分為人工客服和自助客服，客服壓力的增加促使企業使用更多的人工客服，投入的成本也隨之越來越多。

因此，需要自助客服系統具有更大的處理能力，才能滿足客服需求。自助客服系統，能夠自動處理解決用戶的問題。而自助客服的系統中待處理資料量的增加，使現有的方法已經不能夠處理全量資料。現有的演算法隨著問題增加，其計算效率下降。並且大部分特徵是稀疏的，而現有技術適用於處理稠密特徵，這樣，系統中問題特徵量增加的同時，對用戶問題的預測精度就會下降。另外現有技術中的模型單一，效果受限。所以，隨著資訊的不斷爆炸，現在的機器學習模型已經不能滿足需求。

因此，本領域技術人員極待解決的技術問題就是如何透過對以往問題特徵計算得到需要向使用者推薦的問題，提高向用戶推薦問題的準確率，進而在自助客服節點解決用戶問題，減少使用者進入人工客服，降低人工客服成本。

【發明內容】

本發明提供了一種問題推薦方法，用於提高向用戶推薦問題的準確率。所述方法包括以下步驟：

在樣本採集週期內，獲取問題並獲取所述問題對應的問題特徵；

對所述問題特徵進行處理，處理後的問題特徵在預設的數值區間內；

根據各個所述問題及其在所述問題中的第二機率和指定的推薦閾值確定推薦的問題；

其中，各個所述問題及其在所述問題中的第二機率是透過所述處理後的問題特徵和第一機率得到的；所述第一機率是透過所述問題特徵得到的。

較佳地，所述問題特徵包括數值型特徵和文本型特徵，所述數值型特徵為連續的，所述文本型特徵為不連續的。

較佳地，所述獲取問題，具體包括：

在特徵獲取週期內獲取所述問題；

若存在特徵獲取週期內未獲取的問題，則所述未獲取

的問題的值為空；

若不存在特徵獲取週期內未獲取的問題，則以獲取的問題作為所述問題。

較佳地，所述獲取所述問題對應的問題特徵，具體包括：

在特徵獲取週期內獲取問題特徵；

若存在特徵獲取週期內未獲取的問題特徵，且該問題特徵為所述數值型特徵時，則以獲取的與問題對應的問題特徵的數值的均值作為問題特徵；

若存在特徵獲取週期內未獲取的問題特徵，且該問題特徵為所述文本型問題特徵時，則以獲取的與問題對應的問題特徵中出現頻率最高的作為問題特徵；

若不存在特徵獲取週期內未獲取的問題特徵，則以獲取的問題特徵作為所述問題特徵。

較佳地，對所述問題特徵進行處理，具體包括：

若所述問題特徵為數值型問題特徵，則對問題特徵進行歸一化處理；

若所述問題特徵為文本型問題特徵，則對問題特徵進行向量化處理，向量化處理後的問題特徵為數值型問題特徵。

較佳地，所述第二機率是透過對所述處理後的問題特徵和所述第一機率進行深度神經網路 DNN 計算得到的。

相應的本申請還提出了一種問題推薦設備，所述設備包括：

獲取模組：在樣本採集週期內，獲取問題並獲取所述問題對應的問題特徵；

處理模組：對所述問題特徵進行處理，處理後的問題特徵在指定的數值區間內；

確定模組：根據各個所述問題及其在所述問題中的第二機率和指定的推薦閾值確定推薦的問題；

其中，各個所述問題及其在所述問題中的第二機率は透過所述處理後的問題特徵和第一機率得到的；所述第一機率は透過所述問題特徵得到的。

較佳地，所述問題特徵包括數值型特徵和文本型特徵，所述數值型特徵為連續的，所述文本型特徵為不連續的。

較佳地，所述獲取模組具體用於：

在特徵獲取週期內獲取所述問題；

若存在特徵獲取週期內未獲取的問題，則所述未獲取的問題的值為空；

若不存在特徵獲取週期內未獲取的問題，則以獲取的問題作為所述問題。

較佳地，所述獲取模組具體用於：

在特徵獲取週期內獲取問題特徵；

若存在特徵獲取週期內未獲取的問題特徵，且該問題特徵為所述數值型特徵時，則以獲取的與問題對應的問題特徵的數值的均值作為問題特徵；

若存在特徵獲取週期內未獲取的問題特徵，且該問題

特徵為所述文本型問題特徵時，則以獲取的與問題對應的問題特徵中出現頻率最高的作為問題特徵；

若不存在特徵獲取週期內未獲取的問題特徵，則以獲取的問題特徵作為所述問題特徵。

較佳地，所述處理模組具體用於：

若所述問題特徵為數值型問題特徵，則對問題特徵進行歸一化處理；

若所述問題特徵為文本型問題特徵，則對問題特徵進行向量化處理，向量化處理後的問題特徵為數值型問題特徵。

較佳地，所述第二機率是透過對所述處理後的問題特徵和所述第一機率進行深度神經網路 DNN 計算得到的。

由此可見，透過應用本申請的技術方案，對問題特徵進行處理和分類計算，然後對得到的結果進行進一步的深度計算，即可準確地輸出各個所述問題和第二機率。本申請提高了向用戶推薦問題的準確率。本申請可處理稠密性的問題特徵值，同時適合處理大規模的稀疏資料，提高了預測精度。

【圖式簡單說明】

圖 1 為本申請提出的一種問題推薦方法的流程示意圖；

圖 2 為本申請具體實施例所提出的一種 DNN 模型示意圖；

圖 3 為本申請提出的一種問題推薦設備的結構示意圖。

【實施方式】

有鑒於現有技術中的問題，本發明提出了問題推薦方法，該方法應用於問題推薦系統中，結合機器學習模型和深度神經網路模型 DNN 進行模型訓練。該系統能夠根據歷史記錄為使用者推薦其所需的問題，並且擅長處理稀疏型和稠密性的問題特徵，可用於提高向用戶推薦問題的準確率。

如圖 1 所示，為本申請提出的一種核驗資訊處理方法流程示意圖，包括以下步驟：

S101，在樣本採集週期內，獲取問題並獲取所述問題對應的問題特徵。

本申請旨在向需要問題推薦服務的使用者推薦問題。使用者可以是該需要問題推薦的使用者和其他用戶。本發明的系統中存在使用者的歷史記錄，在歷史記錄中包含各個問題以及對應的問題特徵。問題推薦系統通常包括收集層、處理層、儲存層和輸出層。收集層負責收集其他設備發送的問題及問題特徵。處理層利用收集到的問題及問題特徵進行模型訓練。儲存層負責資料儲存，其中儲存了使用者的歷史記錄。輸出層進行問題及問題特徵的輸出。本申請中的問題推薦系統可以在伺服器上實現，較佳採用分散式伺服器。並且本申請可以使用一個伺服器，也可以使用

用多個伺服器組成的集群。

所述問題特徵包括數值型特徵和文本型特徵，所述數值型特徵為連續的，比如，數值型特徵為使用過某個應用軟體的次數，數值 9 代表使用過 9 次，所述文本型特徵為不連續的，比如，文本型特徵為發票狀態，對應未開發票和已開發票。因為歷史記錄中的問題和問題特徵都具有一定的時效性，所以設置樣本採集週期，以採集一段時期之內問題及問題特徵，比如一周內、一個月內。其他設備向系統發送問題及問題特徵時，由於不同設備的 IP 不同，一些問題、問題特徵的獲取時間較長，可能導致系統在很長時間內不能獲取完畢。

為了更高效地對資料進行處理，在本申請的較佳實施例中，設定特徵獲取週期，在特徵獲取週期內獲取所述問題。若存在特徵獲取週期內未獲取的問題，則所述未獲取的問題的值為空，若不存在特徵獲取週期內未獲取的問題，則以獲取的問題作為所述問題。在特徵獲取週期內獲取問題特徵，若存在特徵獲取週期內未獲取的問題特徵，且該問題特徵為所述數值型特徵時，則以獲取的與問題對應的問題特徵的數值的均值作為問題特徵，若存在特徵獲取週期內未獲取的問題特徵，且該問題特徵為所述文本型問題特徵時，則以獲取的與問題對應的問題特徵中出現頻率最高的作為問題特徵，若不存在特徵獲取週期內未獲取的問題特徵，則以獲取的問題特徵作為所述問題特徵。

在獲取問題及其對應的問題特徵之後，推薦系統透過

對問題特徵進行篩選，以刪除一些特徵，比如刪除所有使用者都相同的問題特徵、容易超出特徵獲取週期的問題特徵和與經營業務無關的問題特徵。經過篩選得到的特徵可以為後續建立分類模型作準備。

S102，對所述問題特徵進行處理，處理後的問題特徵在預設的數值區間內。

在獲取問題和對應的問題特徵後，問題推薦系統對問題特徵進行處理。若所述問題特徵為數值型問題特徵，則對問題特徵進行歸一化處理以使處理後的問題特徵在指定的數值區間內；若所述問題特徵為文本型問題特徵，則對問題特徵進行向量化處理，以使處理後的問題特徵為數值型問題特徵並在指定的數值區間內。

在本申請的較佳實施例中，若問題特徵是數值型特徵，可以採用百分位分箱演算法進行歸一化處理，使所有問題特徵在經過處理後都處在指定的數值區間內。在進行百分位分箱演算法處理時，將原始數值歸納到 100 個箱中，然後對箱進行編碼，比如 0.01，0.02.....1.00。經過處理的數值型問題特徵處於 0 至 1 的數值區間內。

由於文本型問題特徵是以文字的形式呈現的，無法參與計算。所以，需要對文本型問題特徵進行向量化處理，將問題特徵由文本型特徵轉化為數值型特徵。可以採用 one hot 編碼處理文本型特徵，計算各個特徵的頻率，從而按頻率給出 one hot 編碼。舉例來說，文本型特徵為發票狀態，對應未開發票和已開發票，經過向量化處理後，

得到數值型特徵 0 和 1，處於 0 至 1 的數值區間內。

問題特徵經過處理後，處於指定的數值區間內，以便於參與後續計算。需要說明的是，本申請需要得到處於指定數值區域內的問題特徵，故進行上述百分位分箱演算法和向量化處理方法僅為本申請較佳實施例提出的示例，本申請的保護範圍並不限於此，以上僅為本申請較佳實施例提出的示例，在此基礎上還可以選擇其他方式來進行計算，以使本申請適用於更多的應用領域，這些改進都屬於本發明的保護範圍。

S103，根據各個所述問題及其在所述問題中的第二機率和指定的推薦閾值確定推薦的問題。

在獲取問題和對應的問題特徵後，本申請還需要對問題特徵進行簡單分類模型計算，可以採用決策樹演算法，從而得到第一機率。在進行決策樹計算時，首先需要兩輪採樣過程。第一輪對問題特徵進行隨機採樣，得到決策樹可處理的問題特徵。第二輪進行重要特徵採樣，根據所述可處理的問題特徵計算權重。

在當今資料介面越來越多的情況下，資料集的原始變數、衍生變數會越來越多，因此資訊值 IV (Information Value) 在實際資料應用中十分重要。資訊值 IV 用來表示每一個變數對目標變數來說有多少“資訊”的量，從而使得特徵選擇變得簡單快速。

在進行特徵選擇時，總是在將特徵的重要程度量化之後再進行選擇，而如何量化特徵，就成了各種方法間

最大的不同。在資訊增益中，重要性的衡量標準就是看特徵能夠為分類系統帶來多少資訊，帶來的資訊越多，該特徵越重要。因此對於一個特徵而言，資訊增益為系統存在該特徵和不存在該特徵時候的資訊量各是多少，兩者的差值就是這個特徵給系統帶來的資訊量，即資訊增益 IG (Information Gain)。

因為資訊值 IV 和資訊增益 IG 都可以表示問題特徵對應的權重，所以，所述權重為資訊值 IV 和/或資訊增益 IG，根據權重選出重要特徵，然後根據所述重要特徵建立分類模型。然後，透過分類模型對篩選得到的問題特徵進行分析得到第一機率。將各個問題特徵經過決策樹計算後得到的相應機率作為所述第一機率。

需要說明的是，本申請進行以上基於決策樹演算法得到的各個問題特徵對應機率值的過程僅為本申請較佳實施例提出的示例，除此之外還可以選擇比如邏輯回歸計算等的其他計算方法，以使本申請適用於更多的應用領域，這些改進都屬於本發明的保護範圍。

在得到處理後的問題特徵和第一機率後，對這兩者進行計算，得到各個所述問題及其在所述問題中的第二機率。在較佳的實施例中，所述各個所述問題及其在所述問題中的第二機率透過 DNN 計算得到。問題推薦系統中的深度神經網路 DNN (Deep Neural Network)，包括輸入節點和計算節點。DNN 計算包括以下步驟：(1) 輸入節點獲取所述處理後的問題特徵和第一機率。(2) 計算節

點透過全連接層、啟動函數 ReLu 和多分類損失函數 softmax loss 對處理後的問題特徵和第一機率進行計算，得到第二機率。

在具體的應用場景中，操作步驟如圖 2 所示：

a.由輸入層獲取所述處理後的問題特徵和第一機率。

在 DNN 訓練前，可以使用決策樹對資料進行初步分類，同時可以透過第一機率控制深度神經網路 DNN 中網路節點所占的權重。

b.由中間層即計算層進行問題推薦，計算層透過全連接層、啟動函數 ReLu 和多分類損失函數 softmax loss 對所述處理後的問題特徵和第一機率進行計算，得到各個問題特徵對應的問題及第二機率。

利用啟動函數 ReLu 進行計算，使一部分網路中神經元的輸出為 0，因而造就了網路的稀疏性，並且減少了參數的相互依存關係，緩解了過擬合問題的發生。同時，使計算節點的計算量較小，有利於提高系統推薦問題的效率。另外，DNN 訓練可以使用 GPU，能夠針對矩陣計算進行加速，進一步提高計算速度。除啟動函數 ReLu 之外，還可以使用 sigmoid 層進行計算。

c.輸出層輸出各個所述問題及其對應的第二機率。

需要說明的是，本申請是由第一機率和處理後得到的數值型問題特徵，得到第二機率，本申請提出的計算方式是 DNN 計算，本申請的保護範圍並不限於此，以上僅為較佳實施例提出的示例，在此基礎上還可以選擇其他方式

來進行計算，以使本申請適用於更多的應用領域，這些改進都屬於本發明的保護範圍。

在本申請的較佳實施例中，問題推薦系統根據各個所述問題及其在所述問題中的第二機率和指定的推薦閾值，確定推薦的問題。然後根據閾值得到閾值內的問題特徵，進而以該問題特徵對應的問題作為推薦的問題。比如在閾值內得到六個問題的問題特徵，則系統推薦這六個問題。本發明在針對各個使用者對應的歷史記錄中的問題和問題特徵進行計算，進而確定待推薦的問題之後，當使用者訪問問題推薦系統的時候直接調用相應的結果。透過本申請中的問題推薦系統，使用者能夠直接獲取與之相關性非常高的問題。

為達到以上技術目的，本申請還提出了一種問題推薦設備，如圖 3 所示，所述設備包括：

獲取模組 310：在樣本採集週期內，獲取問題並獲取所述問題對應的問題特徵；

處理模組 320：對所述問題特徵進行處理，處理後的問題特徵在指定的數值區間內；

確定模組 330：根據各個所述問題及其在所述問題中的第二機率和指定的推薦閾值確定推薦的問題；

其中，各個所述問題及其在所述問題中的第二機率は透過所述處理後的問題特徵和第一機率得到的；所述第一機率は透過所述問題特徵得到的。

在較佳的實施例中，所述問題特徵包括數值型特徵和

文本型特徵，所述數值型特徵為連續的，所述文本型特徵為不連續的。

在較佳的實施例中，所述獲取模組具體用於：

在特徵獲取週期內獲取所述問題；

若存在特徵獲取週期內未獲取的問題，則所述未獲取的問題的值為空；

若不存在特徵獲取週期內未獲取的問題，則以獲取的問題作為所述問題。

在較佳的實施例中，所述獲取模組具體用於：

在特徵獲取週期內獲取問題特徵；

若存在特徵獲取週期內未獲取的問題特徵，且該問題特徵為所述數值型特徵時，則以獲取的與問題對應的問題特徵的數值的均值作為問題特徵；

若存在特徵獲取週期內未獲取的問題特徵，且該問題特徵為所述文本型問題特徵時，則以獲取的與問題對應的問題特徵中出現頻率最高的作為問題特徵；

若不存在特徵獲取週期內未獲取的問題特徵，則以獲取的問題特徵作為所述問題特徵。

在較佳的實施例中，所述處理模組具體用於：

若所述問題特徵為數值型問題特徵，則對問題特徵進行歸一化處理；

若所述問題特徵為文本型問題特徵，則對問題特徵進行向量化處理，向量化處理後的問題特徵為數值型問題特徵。

在較佳的實施例中，所述第二機率是透過對所述處理後的問題特徵和所述第一機率進行深度神經網路 DNN 計算得到的。

透過以上的實施方式的描述，本領域的技術人員可以清楚地瞭解到本申請可以透過硬體實現，也可以借助軟體加必要的通用硬體平臺的方式來實現。基於這樣的理解，本申請的技術方案可以以軟體產品的形式體現出來，該軟體產品可以儲存在一個非揮發性儲存媒體（可以是 CD-ROM，隨身碟，移動硬碟等）中，包括若干指令用以使得一台電腦設備（可以是個人電腦，伺服器，或者網路設備等）執行本申請各個實施場景所述的方法。

本領域技術人員可以理解附圖只是一個較佳實施場景的示意圖，附圖中的模組或流程並不一定是實施本申請所必須的。

本領域技術人員可以理解實施場景中的裝置中的模組可以按照實施場景描述進行分佈於實施場景的裝置中，也可以進行相應變化位於不同於本實施場景的一個或多個裝置中。上述實施場景的模組可以合併為一個模組，也可以進一步拆分成多個子模組。

上述本申請序號僅僅為了描述，不代表實施場景的優劣。

以上公開的僅為本申請的幾個具體實施場景，但是，本申請並非局限於此，任何本領域的技術人員能思及的變化都應落入本申請的保護範圍。

【符號說明】

310：獲取模組

320：處理模組

330：確定模組

申請專利範圍

1. 一種問題推薦方法，其特徵在於，該方法包括：

在樣本採集週期內，獲取問題並獲取該問題對應的問題特徵；

對該問題特徵進行處理，處理後的問題特徵在預設的數值區間內；

根據各個該問題及其在該問題中的第二機率和指定的推薦閾值確定推薦的問題；

其中，各個該問題及其在該問題中的第二機率是透過該處理後的問題特徵和第一機率得到的；該第一機率是透過該問題特徵得到的，該問題特徵包括數值型特徵和文本型特徵，該數值型特徵為連續的，該文本型特徵為不連續的；該第二機率是透過對該處理後的問題特徵和該第一機率進行深度神經網路 DNN 計算得到的；

其中，在該深度神經網路 DNN 訓練前，透過第一機率控制深度神經網路 DNN 中網路節點所占的權重，

其中，該獲取該問題對應的問題特徵，具體包括：

在特徵獲取週期內獲取問題特徵；

若存在特徵獲取週期內未獲取的問題特徵，且該問題特徵為該數值型特徵時，則以獲取的與問題對應的問題特徵的數值的均值作為問題特徵；

若存在特徵獲取週期內未獲取的問題特徵，且該問題特徵為該文本型問題特徵時，則以獲取的與問題對應的問題特徵中出現頻率最高的作為問題特徵；

若不存在特徵獲取週期內未獲取的問題特徵，則以獲取的問題特徵作為該問題特徵。

2.如申請專利範圍第 1 項所述的方法，其中，該獲取問題，具體包括：

在特徵獲取週期內獲取該問題；

若存在特徵獲取週期內未獲取的問題，則所述未獲取的問題的值為空；

若不存在特徵獲取週期內未獲取的問題，則以獲取的問題作為該問題。

3.如申請專利範圍第 1 項所述的方法，其中，對該問題特徵進行處理，具體包括：

若該問題特徵為數值型問題特徵，則對問題特徵進行歸一化處理；

若該問題特徵為文本型問題特徵，則對問題特徵進行向量化處理，向量化處理後的問題特徵為數值型問題特徵。

4.一種問題推薦設備，其特徵在於，該設備包括：

獲取模組：在樣本採集週期內，獲取問題並獲取該問題對應的問題特徵；

處理模組：對該問題特徵進行處理，處理後的問題特徵在指定的數值區間內；

確定模組：根據各個該問題及其在該問題中的第二機率和指定的推薦閾值確定推薦的問題；

其中，各個該問題及其在該問題中的第二機率是透過

該處理後的問題特徵和第一機率得到的；該第一機率是透過該問題特徵得到的，該問題特徵包括數值型特徵和文本型特徵，該數值型特徵為連續的，該文本型特徵為不連續的，該第二機率是透過對該處理後的問題特徵和該第一機率進行深度神經網路 DNN 計算得到的；

其中，在該深度神經網路 DNN 訓練前，透過第一機率控制深度神經網路 DNN 中網路節點所占的權重，

其中，該獲取模組具體用於：

在特徵獲取週期內獲取問題特徵；

若存在特徵獲取週期內未獲取的問題特徵，且該問題特徵為該數值型特徵時，則以獲取的與問題對應的問題特徵的數值的均值作為問題特徵；

若存在特徵獲取週期內未獲取的問題特徵，且該問題特徵為該文本型問題特徵時，則以獲取的與問題對應的問題特徵中出現頻率最高的作為問題特徵；

若不存在特徵獲取週期內未獲取的問題特徵，則以獲取的問題特徵作為該問題特徵。

5.如申請專利範圍第 4 項所述的問題推薦設備，其中，該獲取模組具體用於：

在特徵獲取週期內獲取該問題；

若存在特徵獲取週期內未獲取的問題，則該未獲取的問題的值為空；

若不存在特徵獲取週期內未獲取的問題，則以獲取的問題作為該問題。

6.如申請專利範圍第 4 項所述的問題推薦設備，其中，該處理模組具體用於：

若該問題特徵為數值型問題特徵，則對問題特徵進行歸一化處理；

若該問題特徵為文本型問題特徵，則對問題特徵進行向量化處理，向量化處理後的問題特徵為數值型問題特徵。

圖 式

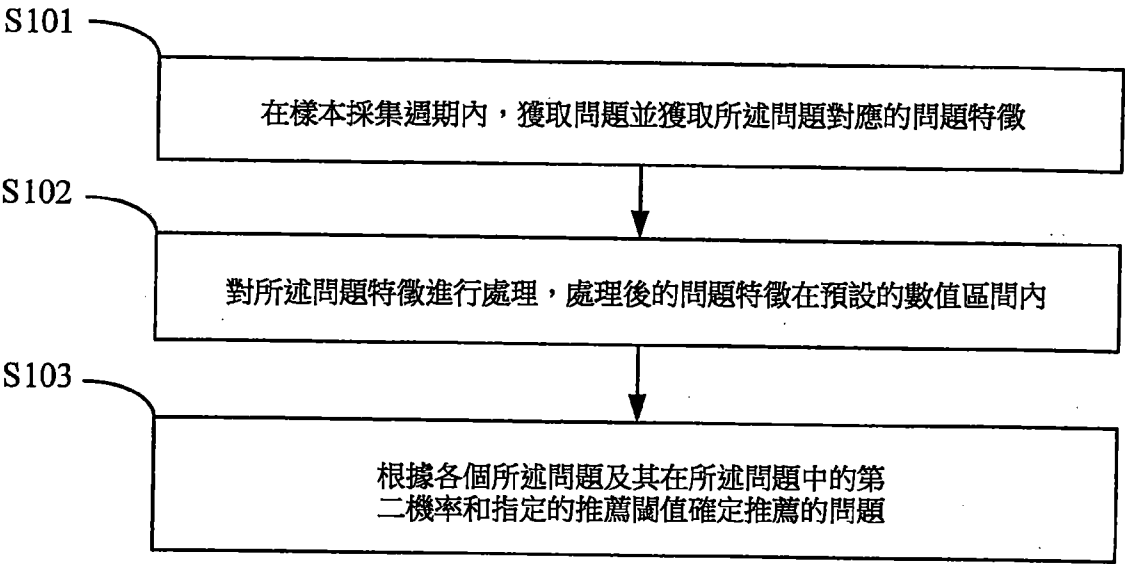


圖 1

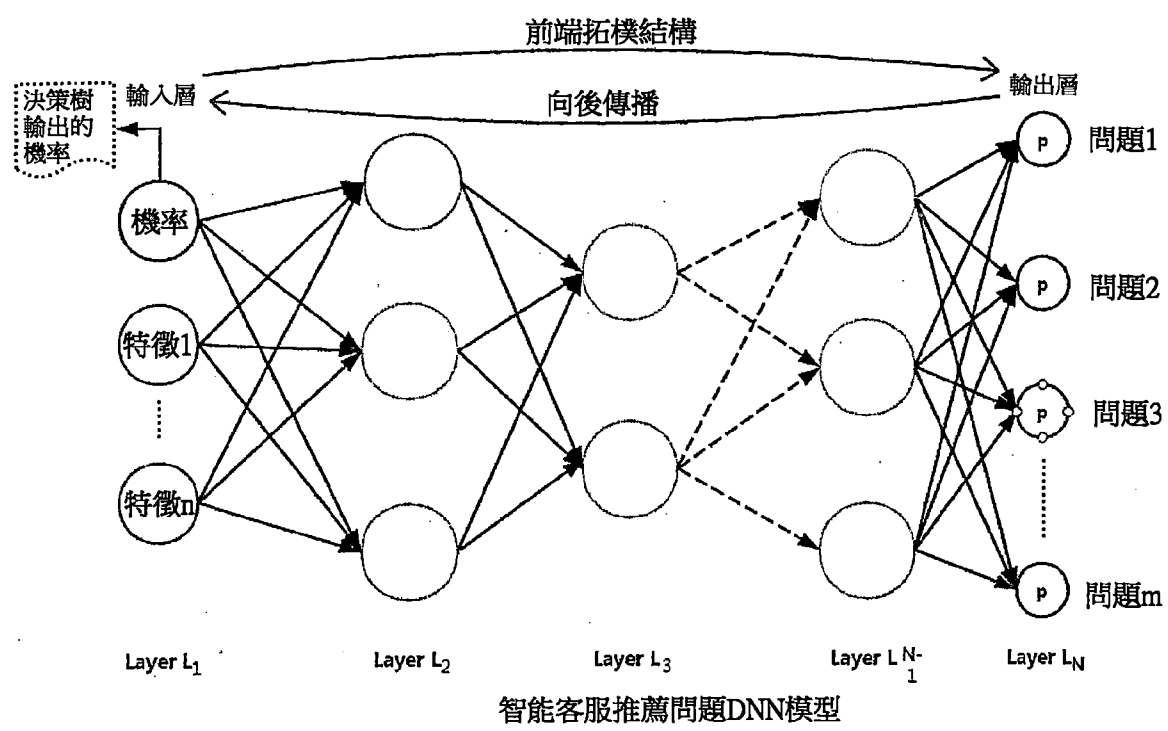


圖 2

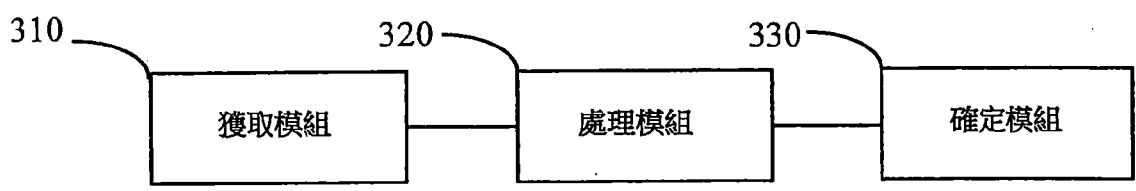


圖 3