

準則式語言模型

本院覽號

05A-1070827

公告日期

智財權狀態

美國US 10,558,926 B2已獲證、美國臨時案已申請

摘要

統計式準則模型（PBA）是一種可解釋的機器學習模型，設計目標在於擷取法則式和統計式機器學習兩者的優點。由於使用近似比對，PBA較易融入不同的應用系統，尤其適合那些需要深度理解的。因此，PBA可以用在文獻探勘，機器閱讀，多輪對談，以及智慧型助教的應用。目前，PBA的美國專利正在審核中。

技術優勢

自然語言處理主要的方法可分法則式模型以及統計或類神經的參數式機器學習模型。法則式模型的缺點在於則必需手工打造每個規則，過於消耗人力。參數式機器學習則難以解釋或做推論。我們提出的PBA則可以類法則式模型進行推論，又可以利用統計N-gram的效果進行全自動或半自動的學習。PBA可以自動分群，並在每群找出一個代表的「準則」。輸入的句子可以和相關的準則比對以決定其最合適的群。當系統判斷錯誤時（譬如：自動分群不合適），使用者可以察覺並調整分群來修正。此外，這些分群的知識可以累積，供其它應用系統借鏡。因此，PBA是一種較容易解釋其決策過程的機器學習。

應用範圍

1. 可以用來學習對談系統的語言及專業知識，並進行多輪問答的理解及省略回填等必需的語言推論；
2. 可以用來學習自然語言指令，並執行相關的代理人任務。
3. 可以學習人工智慧在教育的應用，譬如：自動解題，答案解說，疑難問答，解題暗示，自動出題等等。

創作人

許聞廉



中央研究院
ACADEMIA SINICA