應用語言模型於線上招募詐欺偵測

余銘忠 1

國立高雄科技大學 企業管理系 副教授

洪英皓 2

國立高雄科技大學 企業管理系碩士 研究生

一、 摘要

為了改善但本研究發現過去較少學者研究招募詐欺偵測的議題,且大多皆以特徵工程與機器學習方法的組合進行偵測。因此本研究藉此提出兩種深度學習模型架構(BERT-DNN, GRU-DNN)來驗證深度學習在線上招募詐欺偵測中的表現,除此之外,也透過詞向量(word2vec, fastText)與主題模型來提升模型偵測文字資料時的表現。

本研究以 EMSCAD, 英文維基百科作為模型訓練與測試資料集。在研究結果中,發現本研究的深度學習模型比過去研究採用的機器學習更為優異,其中 GRU-DNN 更為突出。因此未來若採用深度學習的方法進行偵測或延伸,將能有效降低受害機率

關鍵詞：GRU，BERT，詐欺偵測。

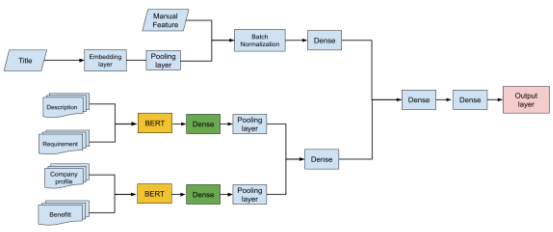
二、 動機與目的

Vidros et al.於 2017 年從 workable 收集資料並將其開源給學術界存取,但本研究亦發現過去較少研究採用深度學習方法,因此本研究以兩種深度學習方法進行設計並驗證。除此之外,本研究調查了在招募詐欺領域中之詞向量模型的影響力,也利用 LDA 與現有特徵工程方法來挖掘出潛在的特徵。

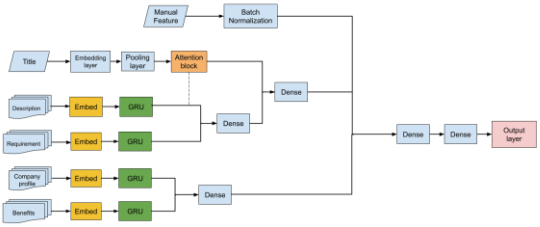
三、 相關文獻與研究方法

本研究依照資料屬性將其歸成三種類別(文字,名目,二元),每種類別採用其對應的預處理與特徵工程進行清洗與轉換。最後則是利用處理後的資料訓練兩種模型,藉此找出能有效判讀詐欺招募文的參數。

(一) BERT-DNN 模型架構



(二) GRU-DNN 模型架構



四、 研究結果與討論

(一) 最適參數與過往研究模型之表現比較

	Accuracy	F-score	precision
GRU-DNN	0.9733	0.7713	0.649
BERT-DNN	0.9436	0.5976	0.4384
Random Forest	0.8269	0.4098	0.282

本研究藉由一系列的實驗設計來訓練出最適參數之模型。在實驗結果中發現相較於 BERT-DNN,GRU-DNN 更能有效地偵測不平衡資料集的少數類別,並且兩種模型在各評分準則下皆比 Vidros et al. (2017)提出之模型還要好。

五、 結論

本研究發現在人工智慧領域中,過往較少研究在招募詐欺偵測議題中發展,並且大多皆採用傳統的機器學習模型與特徵工程來偵測。故本研究以現今當紅的深度學習方法進行偵測並驗證。在研究結果中，證實了深度學習模型表現較為優異，其中 GRU-DNN 與 word2vec 的 CBOW 模型組合更適合被應用在此領域中。

但本研究亦發現此資料集涵蓋了大量的錯字、缺失值，且由於模型較為複雜,需要大量的運算時間與記憶體空間才能偵測。因此未來若要發展深度學習之模型,可採用更細緻的斷詞方法來處理錯字問題;在運算時間與記憶體空間方面,則是採用可平行化運算的模型(CNN, Transformer)與 Network compression 技術來更進一步優化線上招募詐欺偵測系統。