

# 基於深度學習方法之佛教引用句推薦系統

釋廣慧<sup>1</sup>、王昱鈞<sup>2,\*</sup>、洪振洲<sup>3</sup>

## 摘要

在寫作之中，適度引用名言佳句是常用之撰寫技巧，能增進文章說服力，讓文章更加優美令人信服。而佛教典籍之中有非常多深具哲理與啟發之名句，許多創作之中亦時常引用佛教經典之語句以闡述其文之要旨，然對於現代人來說，要在寫作時引用佛教相關之引用句實非易事。因此能於文章寫作時依據當前內容自動推薦合適的佛教引用句即為一重要的需求與研究課題，本論文即針對佛教引用句推薦問題，提出基於長短期記憶 (Long Short-Term Memory, LSTM) 與基於轉換器的雙向編碼器表徵技術 (Bidirectional Encoder Representations from Transformers, BERT) 兩種深度學習方法之佛教引用句推薦系統，可自動從文章內容進行分析推薦合適的佛教引用句。我們建置佛教引用句資料集，藉以訓練深度學習模型，實驗結果顯示系統之推薦準確率可高達 0.9148，其能有效進行寫作時之佛教引用句推薦。

**關鍵詞：**佛教引用句、推薦系統、深度學習、長短期記憶、BERT

---

投稿日期：2022 年 9 月 28 日；通過日期：2022 年 11 月 14 日。

<sup>1</sup> 法鼓文理學院博士生。

<sup>2</sup> 法鼓文理學院助理教授。

<sup>3</sup> 法鼓文理學院教授。

\* 通訊作者：王昱鈞，Email: ycwang@dila.edu.tw

## 壹、前言

在寫作之中，適度引用名言佳句是常用之撰寫技巧，除了能添飾文采外，亦能增進文章說服力，讓創作之文章更加優美令人信服。此類名言佳句，多來自傳統諺語、名人之語錄，以及許多哲學思想與宗教典籍之名句，是人類千百年來思想與人生經驗的結晶。然而，要在文章寫作之中適當引用合適的名言佳句，對於一般人來說並非十分容易的事，有時於寫作時非常渴望在文章某處引用名言佳句，但不一定當下即能想出使用相關合適的佳句加強文章的力度。名言佳句之數量浩繁，要記誦並熟知這些名言佳句的意涵與適於引用的時機並不容易。因此，若能有一針對寫作之引用句推薦系統，以文章內容推薦合適的引用句，將可解決此寫作之難題。

在選擇引用句時，除了傳統諺語與知名人士之名言之外，來自宗教典籍如佛教經典或基督教之《聖經》等文本中之經典名句亦是常用之引用對象。然而，這類宗教典籍之引用句，多與其宗教意涵有高度關連，除非對於此類宗教典籍與其教義十分熟稔，要適度引用此類引用句對一般人來說更加困難。

特別於東方文化之中，起源自印度的佛教對於整個東亞文明的發展具有相當深遠之影響。自東漢佛教開始傳入中土以來，佛教經典不斷持續輸入，引發了千年來佛經翻譯之工作，大量的佛經被翻譯為漢文，成為今日的佛教大藏經。而佛教的教義與思想也隨之融入中華文明之中，並拓展至整個東亞漢字文化圈，成為東亞文化重要成分。佛教的思想能給人一種啟發作用，是一種精神性的文化資源，能為人的精神世界提供資糧。佛教典籍之中有非常多深具哲理與啟發之名句，古來許多文人的創作之中亦皆時常引用佛教經典之語句以闡述其文之要旨，然對於現代人來說，要在寫作時引用如此精奧之佛教相關之引用句實非易事。因此，能於文章寫作時依據當前內容自動推薦合適的佛教引用句即成為一重要之需求，如能實現此一佛教引用句之推薦方法，除了能解決寫作上之難處之外，更能將這類佛教引用句中所包含之佛教積極、樂觀、正信、進步及多采多姿的一面呈現出來，使佛教在今天人們建設精神家園中發揮更多的積極作用。

本論文即針對佛教引用句推薦問題，提出一個自動從文章內容進行分析推薦合適的佛教引用句方法與系統。我們採用現時於自然語言處

理領域已廣泛導入之深度學習方法，使用長短期記憶（Long Short-Term Memory, LSTM）與基於轉換器的雙向編碼器表徵技術（Bidirectional Encoder Representations from Transformers, BERT）之語言表徵模型以建構佛教引用句推薦系統。

## 貳、文獻回顧

對於佛教引用句推薦問題，現前雖無相關針對佛教領域之自動引用推薦之研究，然過往對於引用推薦與基於內容之推薦任務已有許多研究成果，其與佛教引用句推薦所欲解決之目標接近且相關，以下分別詳述相關之研究。

引用推薦基本上可視為一種基於文本內容之推薦任務，傳統之基本內容之推薦任務最初主要是用於產品推薦，其根據產品之描述文本內容，以及使用者的偏好與歷程紀錄，透過內容分析以計算排序適合的推薦項目（Pazzani & Billsus, 2007）。文本內容之推薦任務中與引用句推薦最相關之研究問題為基於文本內容的引用參考推薦（content-based citation recommendation; Strohman, Croft, & Jensen, 2007），其根據文本的內容推薦適於插入文中之學術論文參考文獻項目。He、Pei、Kifer、Mitra 與 Giles（2010）提出利用非參數式之機率模型用以計算文本內容與引用文獻內容之相似度以產生引用文獻推薦。Huang 等人（2012）將文本內容與參考文獻引用之間視為一種特殊的翻譯問題，其將欲進行引用的文本內容與參考文獻之內容視為兩種語言，利用 IBM Model 機器翻譯模型進行訓練建構引用文獻推薦系統。Huang、Wu、Liang、Mitra 與 Giles（2015）則利用神經機率模型以學習出欲進行引用的文本內容與參考文獻之內容的向量表徵，藉此計算文本與參考文獻之匹配機率以輸出推薦結果。此類基於文本內容的推薦方法，其多可利用許多結構資訊，如論文標題、摘要內容，論文間引用參考關係等，且其文本長度一般都較長。而本論文欲解決之引用句推薦問題，引用句一般皆很短，且除了引用句前後文字內容外並無其它資訊，故更具挑戰性。

另一相近之研究問題為成語慣用句推薦（idiom recommendation），Liu、Liu、Shan 與 Wang（2018）其利用神經網路學習文本內容與成語之

分布表徵，以計算其相似度進行排序推薦成語。王皓儀（2019）以中文成語例句語料庫來訓練 Word2Vec 詞嵌入模型以獲得中文詞向量，接著以中文文本內容訓練 LSTM 與雙向 LSTM 模型進行分類，以輸出適合用於該文本之中文成語。Liu、Pang 與 Liu（2019）將成語推薦任務視為機器翻譯問題，採用基於注意力機制之編碼器－解碼器架構，訓練一可將輸入文本內容翻譯為成語之翻譯模型。

Tan、Wan 與 Xiao（2015）最先定義提出於寫作上的引用推薦任務，其提出一基於 Learning-to-Rank 框架之引用推薦方法，其基於 16 種以人工設計之特徵，將引用推薦視為一般任務，其自古騰堡計畫（Project Gutenberg）中收集 3,158 句引用句及 64,323 篇文章做為資料集。Tan、Wan 與 Xiao（2016）與 Tan、Wan、Liu 與 Xiao（2018）則導入神經網路模型於引用推薦任務，其使用 LSTM 神經網路以學習出引用句與文本內容之向量分布表徵，以此學習出之向量計算引用句與相關文本之匹配。Ahn、Lee、Jeon、Ha 與 Lee（2016）則結合四種不同的模型，包括基於統計之層級調整（granularity adjustment）、隨機森林、卷積神經網路（convolutional neural network, CNN）與 LSTM，依此建立 hybrid 混合模型以進行的引用推薦，其亦自古騰堡計畫電子書與線上部落格收集近 400 項包含引用句與文本以建構資料集。

除了用於寫作之引用推薦任務外，許多研究亦關注於對話時的引用推薦。Lee、Ahn、Lee、Ha 與 Lee（2016）提出 LSTM-CNN 整合模型用於自 Twitter 連續對話串中推薦引用。Wang、Li、Zeng、Zhang 與 Wong（2020）使用編碼器－解碼器框架自對話歷程與目前查詢中學習生成引用句內容以作為回應。Wang、Zeng 與 Wong（2021）則採用語義編碼方法，將對話歷程以 Transformer（Vaswani et al., 2018）與 GRU（Cho et al., 2014）進行編碼，並將引用內容以 Transformer 進行編碼，再經由編碼結果進行引用匹配之計算。

近期相關之引用推薦研究，多採用基於神經網路學習文本之表徵以進行分類或翻譯，作為推薦之結果輸出，在許多推薦資料集上皆有相當好之效果。於佛教引用句推薦任務上，我們亦借鑑於引用推薦與成語推薦之作法，採用深度學習之循環神經網路進行分類，並導入預訓練語言模型以解決過往推薦語料集規模較小之問題。

## 參、佛教引用句推薦方法

佛教引用句推薦其目標即希望透過目前已撰寫之文章內容，由方法自動推薦出合適的引用句。例如有一現前之文章內容為：「人生苦短如夢，稍縱即逝，若不謹慎持握，可能就突然過去了。生命脆弱如絲，一呼一吸之間瞬息消滅，如果不能善力發揮，真是可惜人生一場。…，世間的一切永珍，就像睡夢、水泡、朝露、閃電那樣的短促無常，有什麼好貪婪、執著、眷戀的？」，在文章之中之「…」之處即為希望加入佛教引用句之位置。在此例中，填上「一切有為法，如夢幻泡影」此一佛教引用句最為適合，因為人生中之人、事、物是「一切有為法」，因執著而鬱悶難當，能觀照這一切「如夢幻泡影」，心就能輕安自在。方法的輸入即為所欲使用引用句之處「…」的前文與後文，而推薦方法根據所輸入之前文與後文，於佛教引用句集合之中選出最適合於此處之引用句。

我們將佛教引用句推薦問題定義如下：

$$c = \arg \max_{q_i \in Q} \text{Prob}(T_{\text{previous}}, T_{\text{afterward}}, q_i) \quad (1)$$

其中  $Q$  為佛教引用句之集合， $q_i$  代表  $Q$  中的第  $i$  個引用句， $T_{\text{previous}}$  為欲引用處之前文， $T_{\text{afterward}}$  為欲引用處之後文， $\text{Prob}()$  函式則為模型所計算輸出之機率，而  $c$  則為模型所輸出之最適合的佛教引用句之索引值。

針對佛教引用句推薦問題，我們分別使用 LSTM 與 BERT 兩種深度學習方法以建構推薦模型，此外並以傳統機器學習方法之支援向量機（support vector machine, SVM）模型作為基線方法以進行比較。在進行機器學習模型進行引用句推薦之前，輸入之資料需先經過前處理，以下分節詳述前處理之方式與各推薦模型之細節。

### 一、資料前處理

佛教引用句推薦是以文章之前後文作為輸入，其基本上皆為現代中文之資料。由於不同的模型對於中文資料之輸入表示方式不同，對於中文資

料我們採用兩種方式進行前處理，對於 LSTM 與 SVM 基線方法其中文資料會先進行中文分詞，對於 BERT 方法則將輸入字串依中文字元進行記號化 (tokenization)。

在中文分詞方法上，我們採用 Jieba 中文分詞工具 (<https://github.com/fxsjy/jieba>)，並使用 Jieba GitHub 所提供之包含繁體中文之大型詞庫 ([https://github.com/fxsjy/jieba/blob/master/extra\\_dict/dict.txt.big](https://github.com/fxsjy/jieba/blob/master/extra_dict/dict.txt.big)) 作為主要之分詞詞典。Jieba 之分詞方法為將輸入之中文語句字串中的每一個字元與其後之連續字元之組合於詞典中進行搜尋，如有成詞者即記錄其所有可能之組合，整句之可能詞組組合搜尋完整後，即可建構出有向無環圖 (directed acyclic graph)，而後再以動態規畫方法找出成詞機率最高的路徑，即為最終之分詞結果。Jieba 在現代中文分詞上已為常用之工具，其分詞正確率亦有相當之效果，且其易於與各種程式語言與平臺整合，故十分適合用於本推薦方法之中文分詞工作。若輸入之中文文章內容如下：

別人認不認真與我不相干，所以自己要想成就首先要肯定三藏十二部是諸佛菩薩教導我的。我要發心作佛弟子，我要發願接受佛菩薩的教誨，別人學不學與我不相干。

經過 Jieba 分詞後，可得到如下結果：

別人認不認真與我不相干，所以自己要想成就首先要肯定三藏十二部是諸佛菩薩教導我的。我要發心作佛弟子，我要發願接受佛菩薩的教誨，別人學不學與我不相干。

對於 BERT 模型，其對於中文之處理方式並不進行分詞，而是直接以單一字元做為輸入之單位。我們使用中文 BERT 提供之 BERT Tokenizer 進行輸入中文文章語句之處理，同樣之輸入內容經過 BERT Tokenizer 處理後，其輸出結果如下：

[CLS] 別人認不認真與我不相干，所以自己要想成就首先要肯定三藏十二部是諸佛菩薩教導我的。我要發心作佛弟子，我要發願接受佛菩薩的教誨，別人學不學與我不相干。 [SEP]

其中開頭與結尾之 [CLS] 與 [SEP] 為 BERT 中所保留之特殊 token，由 BERT Tokenizer 自行添加，開頭之 [CLS] 用以作為 BERT 進行整句分類問題之輸出位置，而 [SEP] 則是用於切分兩個以上的句子使用。

## 二、基於機器學習之推薦方法

### (一) LSTM

LSTM 為循環神經網路 (recurrent neural network, RNN) 之一種，其架構適於閱理序列之數據。傳統的前饋式神經網路，其對於輸入之資料，皆將其當作各自獨立的網路輸入，輸入資料的先後順序對於神經網路的計算上沒有任何影響。然而由於語言之表達與語法結構皆與詞彙之順序有高度關聯，而循環神經網路則是設計用以處理此類序列之問題，其架構上 RNN 可以將狀態在自身的網路中循環傳遞，將前一次神經網路的狀態傳遞至本次的神經網路，從而讓神經網路可以學習到前後關聯之訊息。然而傳統的 RNN 在隨著時間序列不斷增長時，在模型訓練上的遭遇梯度消失的問題，前面較久遠的資訊對於當前的結果幾乎不存在影響，意即傳統 RNN 很難學到較久遠以前的資訊與當前輸出的關係，造成 RNN 無法解決長期依賴問題。而 LSTM 即設計用以解決此問題，透過在其神經網路中加入記憶空間，並利用運算閘以決定是否要儲存或刪除內部記憶空間之資訊，而運算閘亦設計為權重值參數來決定其開啟與否，從而讓運算閘也能夠一同被訓練，LSTM 在訓練過句中會自行學習隨著序列輸入，哪些訊息需要被記憶，哪些訊息不重要應該要遺忘，從而解決 RNN 無法處理長期依賴之問題。

在佛教引用句推薦之方法上，我們採用基於 LSTM 的循環神經網路模型其示意圖如圖 1 所示。

輸入每一個 token 會於嵌入層 (embedding layer) 轉換為向量表徵，並填充至最大輸入長度，而後分別作為 LSTM 循環神經網路層之輸入。在引用句推薦之模型設計上，於循環神經網路層，我們分別採用兩種設計，其一為僅單一 LSTM 循環神經網路所構成，另一則是疊加兩層 LSTM 循環神經網路，建構雙向 LSTM (Bidirectional LSTM, BiLSTM)，其中一

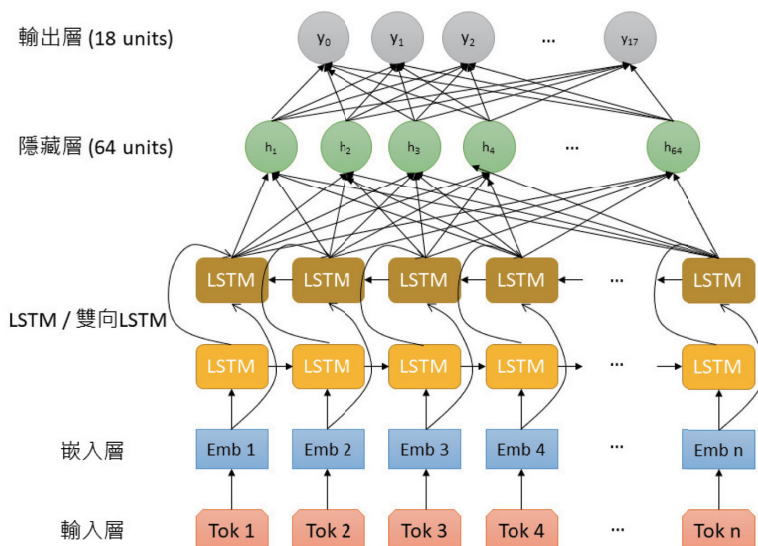


圖 1 LSTM 佛教引用句推薦方法模型架構

資料來源：作者自製。

層為從前至後傳遞，另一層則為從後至前傳遞，這兩層皆以嵌入層作為時間序列之輸入。前向之 LSTM 層可學習到從前文而來的訊息，而後向之 LSTM 層則可學習到從後文來的訊息，可有效學習雙向的語意依賴，雙向 LSTM 層各自之輸出結果將進行拼接 (concatenation) 為一個輸出向量。LSTM 層之輸出再連接至一具 64 個神經元之全連接層，最後再連至與引用句集合數量相當之神經元數量之全連接層，作為模型最終之輸出。

## (二) BERT

近年來在自然語言處理 (Natural Language Processing, NLP) 之技術上，基於大量語料預訓練之語言表徵學習模型得到廣泛之應用，在各種 NLP 問題上導入此類預訓練模型皆有顯著之進步。其中最為突破性且得到廣泛應用的，即為 Google 所提出之 BERT。BERT 為 Google 於 2018 年所提出之語言表徵模型 (Devlin, Chang, Lee, & Toutanova, 2018)，其基於 Transformer 模型中之編碼器架構，為一種深度雙向之非監督式語言表徵之方法。BERT 主要之目的為給定一些詞彙的前提下，評估出現的下一個詞彙的機率分布。由於 BERT 的訓練並不需要標注語料，可以自極大量之



未標注語料中學習詞彙的搭配表達，加上多層自注意力機制的深度學習模型之能力，使得 BERT 除了單純的詞彙表現之外，更能夠學到語法結構與一定之語義訊息，在許多 NLP 之問題上，如語言理解、自動問答等，皆有優於先前方法之表現。

BERT 為基於 Transformer 之深度雙向語言表徵模型，其基本上是利用 Transformer 架構建構出多層雙向的編碼器神經網路，Transformer 為 Google 於 2017 年提出之基於自注意力機制（self-attention）的深層模型（Vaswani et al., 2018），其模型架構如圖 2 所示。

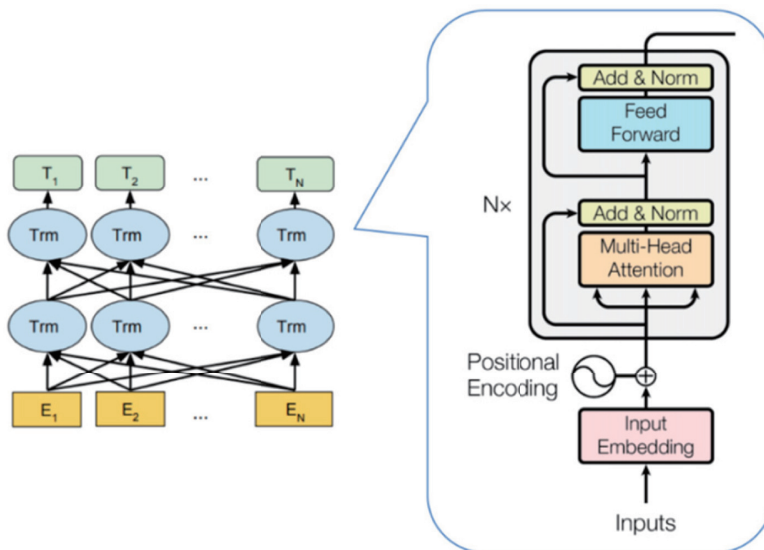


圖 2 BERT 模型架構

資料來源：美團技術團隊（2020）。

BERT 其輸入為 token 所構成之語句，依其所處理之語言不同有不同之 Tokenization 策略，以中文來說，一般皆是以單一漢字為單位。其輸入除了 token 外，亦包含其位置編碼與區段編碼之訊息，用以表示該 token 於句中的位置，以及其屬於何句之用。輸入之語句皆會經過多層的 Transformer 網路，其每一個 Transformer 單元皆以多層多頭自注意力機制層建構而成，其能有效觀察出整個語句中的前後所有字詞之資訊以決定其輸出。

在 BERT 模型之訓練上，其僅需未標注之大量文本，其主要以兩項任務做為訓練標的：「遮罩語言模型」(Masked Language Model, MLM)與「下一句預測」(Next Sentence Prediction, NSP)，此兩項任務巧妙地利用語言的特性，可利用大量收集來的無標注文本，產生自標注 (self-annotated) 文本以進行訓練。MLM 任務其將輸入之語句中隨機將其中之 15% 的 token 遮掩，要求模型預測被遮掩之位置的 token 實際為何。此 MLM 訓練任務讓模型得以學會上下文前後詞組搭配之關係，理解語言於同一句中使用之詞語組合。NSP 任務則是輸入兩個句子，而模型則需進行二元分類任務，判斷該兩句是否在實際的語言使用上為連續的前後兩個句子。NSP 任務可以讓模型學會句子與句子間的語意連續之脈絡，理解語言於對話前後脈絡下的語意連貫。藉由如上 MLM 與 NSP 任務，便能訓練出得以進行語言表徵的 BERT 語言模型。

完成 BERT 語言模型之預訓練後，BERT 即能有效表示輸入之該語言的語句之訊息，在實際方法之建構上，會將訓練好之 BERT 語言模型利用遷移學習 (Transfer Learning) 之方式，將 BERT 語言模型再次透過所欲解決之問題之相關語料進行微調 (fine-tuning) 訓練，以得到最終之模型。在佛教引用句推薦方法上，我們將已完成預訓練之中文 BERT 語言模型，利用佛教引用句語料進行分類任務之微調訓練，將其遷移以適用於佛教引用句推薦，圖 3 為 BERT 微調訓練之架構。

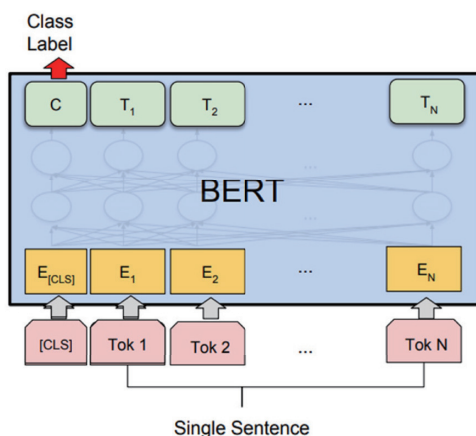


圖 3 BERT 之分類任務微調訓練

資料來源：Devlin 等人 (2018)。

### （三）基線方法：SVM

在基線方法的選擇上，我們採用於機器學習領域常見且分類效果較佳之 SVM 模型（Vapnik, 1999）。SVM 是一組用於分類、回歸和異常值檢測的監督學習方法。SVM 的分類方法是在類別資料間找出最佳的決策邊界，針對一訓練資料集  $\{(x_i, y_i)\}$ ，其中  $x_i \in R^n$  為輸入資料，而  $y_i \in \{-1, +1\}$ ， $i = 1, \dots, l$  為分類之標籤， $l$  為訓練資料集之大小，而 SVM 則欲找出一超平面  $f(x_i) = (\omega \cdot x_i) + b$ ，即求解如下之方程式之全域最佳解：

$$\min_{\omega, b, \xi} \frac{1}{2} \|\omega\|^2 + C \sum_{i=1}^l \xi_i \quad \text{s.t.} \quad y_i((\omega \cdot x_i) + b) \geq 1 - \xi_i, \xi_i \geq 0 \quad (2)$$

對於原始之輸入資料，其經由非線性對應轉換至高維度空間，而此超平面則於此轉換之高維度空間中尋找，為降低計算複雜度，於計算時導入核函數（kernel function）方法（Aizerman, 1964），核函數可使得於高維度空間中估量兩個向量之間之相似度之計算，得以於原先之低維度空間之計算得出，藉由降低計算之複雜度。找出超平面後，對於未知之資料進行預測時，即可以其於超平面之相對位置，得出分類之結果。

## 肆、實驗與討論

### 一、訓練資料集建置

為建置佛教引用句推薦之訓練資料集，首先參考淨行居士編著之《佛學名句類編》一書，針對在寫作與一般社會大眾認知上較廣為熟知之引用句，共擇定 18 句，表 1 為該 18 句佛教引用句之列表。

針對擇定之 18 句佛教引用句，以自網路上以該佛教引用句作為搜尋關鍵字，從各類資料網站中查詢出包含該佛教引用句之網頁文章，再透過網路爬蟲技術擷取該文章，表 2 為資料來源之網站列表及自其中所擷取之文章數目。

表 1 擇定佛教引用句列表

| 標籤編號 | 佛教引用句         | 資料數量  |
|------|---------------|-------|
| 0    | 一切有為法，如夢幻泡影   | 590   |
| 1    | 凡所有相，皆是虛妄     | 775   |
| 2    | 若真修道人，不見世間過   | 329   |
| 3    | 應無所住，而生其心     | 923   |
| 4    | 過去心不可得，現在心不可得 | 325   |
| 5    | 色不異空，空不異色     | 600   |
| 6    | 不生不滅，不垢不淨     | 445   |
| 7    | 不取於相，如如不動     | 328   |
| 8    | 如來者，無所從來      | 164   |
| 9    | 無有定法，如來可說     | 163   |
| 10   | 無有恐怖，遠離顛倒夢想   | 236   |
| 11   | 法尚應捨，何況非法     | 290   |
| 12   | 本來無一物，何處惹塵埃   | 518   |
| 13   | 一念悟時，眾生是佛     | 65    |
| 14   | 何期自性，本自清淨     | 182   |
| 15   | 不是風動，仁者心動     | 25    |
| 16   | 如來是真語者、實語者    | 164   |
| 17   | 是法平等，無有高下     | 243   |
| 數量總計 |               | 6,365 |

資料來源：作者自製。

表 2 資料來源網站列表

| 網站名稱      | 網址  | 擷取數量  |
|-----------|---|-------|
| Google 台灣 | <a href="https://www.google.com.tw">https://www.google.com.tw</a>                             | 2,719 |
| 法鼓全集網頁版   | <a href="http://ccdd.omtf.pl/pc.htm">http://ccdd.omtf.pl/pc.htm</a>                           | 237   |
| 般若文海      | <a href="https://book.bfn.org/">https://book.bfn.org/</a>                                     | 1,325 |
| 中台禪寺      | <a href="https://ctworld.org.tw/index.htm">https://ctworld.org.tw/index.htm</a>               | 481   |
| 聯合新聞網     | <a href="https://udndata.com/ndapp/Index?cp=udn">https://udndata.com/ndapp/Index?cp=udn</a>   | 332   |
| 佛光山資訊網    | <a href="https://www.fgs.org.tw/">https://www.fgs.org.tw/</a>                                 | 206   |
| 香光資訊網     | <a href="http://www.gaya.org.tw/query/index.html">http://www.gaya.org.tw/query/index.html</a> | 127   |
| 台灣學佛網     | <a href="http://big5.xuefo.net/">http://big5.xuefo.net/</a>                                   | 257   |
| 數量總計      |   | 5,684 |

資料來源：作者自製。

透過程式從抓取的文章中搜尋有佛教引用句之位置點，擷取引用句前後約各 100 個字，前後文長度大約 200 個字，產生佛教引用句與其上下文之配對作為訓練資料集之項目。

## 二、評估方法

為評估佛教引用句推薦結果之有效性，我們採用 Top- $k$  準確率（Top- $k$  Accuracy）做為評估方法，Top- $k$  Accuracy 之定義如下：

$$\text{Top-}k \text{ Accuracy} = \frac{|\text{correct classifications in the first } k \text{ outputs}|}{|\text{all first } k \text{ outputs}|} \quad (3)$$

其中分母為推薦方法針對輸入所產生之前  $k$  筆推薦項目之數量，分子則為推薦結果的前  $k$  筆項目中符合正確分類的數量。對於推薦問題來說，我們皆希望推薦方法能於推薦的第一個項目即符合我們所希望的答案，其即為 Top-1 Accuracy 所估算出準確率。但於實際應用中，有時不一定需要如此嚴苛的評判標準，方法所推薦之項目之中的前數筆資料若有包含答案通常亦可接受，故可適度放大  $k$  的數量以評估方法實際應用的有效性。於本論文中分別採用 Top-1、Top-3 與 Top-5 Accuracy 做為評估標準。

## 三、實驗設定

為有效評估推薦效果，進行佛教引用句推薦實驗時採行十折交叉驗證（10-fold cross-validation）方法，將整個佛教引用句資料集經隨機打亂順序後，切分為十等分，依序取其中一等分作為實驗之測試資料，其餘九等分合併作為實驗之訓練資料，以訓練機器學習之佛教引用句推薦系統，藉以消弭因資料切分之不同所可能造成之結果偏差。

於基線 SVM 模型中，我們採用於分類問題中表現較佳之 Gaussian Radial Basis Function（RBF）函數作為分類器之核函數。於 LSTM 模型之實驗上需預先設定相關之超參數（hyperparameter），於本論文實驗中所設定之固定參數如表 3 所示。

此外，於 LSTM 與 BiLSTM 模型中，於嵌入層之詞向量維度與 LSTM 單元之輸出維度之兩項參數由於對於分類結果具有較大之影響，故額外進行實驗以測試不同之參數組合之結果，以擇定最佳之參數值。表 4 為 LSTM 模型對於不同之詞向量維度與 LSTM 輸出維度實驗之正確率結果，

表 5 則為 BiLSTM 模型對於不同之詞向量維度與 LSTM 輸出維度實驗之正確率結果。

從實驗結果可知隨著詞向量維度與輸出維度提高，模型正確率皆與之上升，而於詞向量維度達到 512 與輸出維度達到 512 時，於 LSTM 與 BiLSTM 模型皆達到最高之正確率，若再繼續增加詞向量維度或輸出維度，則模型之正確率便開始下降，顯示模型可能因為參數增加而導致過擬

表 3 LSTM 模型固定參數

| 固定參數             | 設定值                       |
|------------------|---------------------------|
| Optimizer        | Adam                      |
| Loss Function    | Categorical cross-entropy |
| Vocabulary Size  | 4600                      |
| Max Input Length | 200                       |
| Batch Size       | 32                        |
| Epoch            | 20                        |

資料來源：作者自製。

表 4 LSTM 模型參數實驗結果 (Top-1 Accuracy)

| 詞嵌入維度 | 輸出維度   |        |        |        |        |        |
|-------|--------|--------|--------|--------|--------|--------|
|       | 32     | 64     | 128    | 256    | 512    | 768    |
| 32    | 0.7055 | 0.7126 | 0.7365 | 0.7316 | 0.7588 | 0.7286 |
| 64    | 0.7200 | 0.7557 | 0.7605 | 0.7717 | 0.7818 | 0.7852 |
| 128   | 0.7379 | 0.7494 | 0.7624 | 0.7786 | 0.7925 | 0.8031 |
| 256   | 0.7278 | 0.7712 | 0.7822 | 0.7786 | 0.7898 | 0.7970 |
| 512   | 0.7629 | 0.7884 | 0.7892 | 0.7906 | 0.8187 | 0.8175 |

資料來源：作者自製。

表 5 BiLSTM 模型參數實驗結果 (Top-1 Accuracy)

| 詞嵌入維度 | 輸出維度   |        |        |        |        |        |
|-------|--------|--------|--------|--------|--------|--------|
|       | 32     | 64     | 128    | 256    | 512    | 768    |
| 32    | 0.7063 | 0.7344 | 0.7351 | 0.7338 | 0.7426 | 0.7347 |
| 64    | 0.7167 | 0.7530 | 0.7555 | 0.7596 | 0.7624 | 0.7671 |
| 128   | 0.7071 | 0.7519 | 0.7608 | 0.7588 | 0.7764 | 0.7958 |
| 256   | 0.7115 | 0.7440 | 0.7723 | 0.7838 | 0.8016 | 0.8124 |
| 512   | 0.6723 | 0.7519 | 0.7868 | 0.7844 | 0.8168 | 0.8097 |

資料來源：作者自製。

合 (overfitting)，至使於實際資料之表現變差。而 LSTM 模型與 BiLSTM 模型相較，其表現略佳，增加雙向 LSTM 對於佛教引用句推薦之問題上並未有較突出之表現。故於 LSTM / BiLSTM 模型，最終擇定採用 LSTM 模型，並將詞向量維度與 LSTM 輸出維度皆設定為 512。

針對 BERT 模型，其固定之參數設定如表 6 所示，模型參數之設定依據深度學習運算設備之記憶體限制與資料集之輸入長度而定。

表 6 BERT 模型參數

| 參數         | 設定值 |
|------------|-----|
| Batch Size | 10  |
| Epoch      | 5   |
| Max Length | 300 |

資料來源：作者自製。

#### 四、實驗結果

表 7 為基線 SVM 模型、LSTM 模型與 BERT 模型之實驗結果。從實驗數據可知，LSTM 模型與 BERT 模型整體平均優於基線 SVM 模型，LSTM 模型於 Top-1 準確率為 81.87%，而 Top-3 準確率已達 90.13%，已具相當良好之準確程度。而 BERT 模型相較其它兩個模型達到最佳之效果，其於 Top-1 準確率即已達到 91.48%，其 Top-5 準確率更高達 97.97%，顯見其對於佛教引用句之推薦已有相當好之效果。

表 7 佛教引用句推薦模型實驗結果

| Approach       | Top-1 Accuracy | Top-3 Accuracy | Top-5 Accuracy |
|----------------|----------------|----------------|----------------|
| Baseline (SVM) | 0.7943         | 0.8948         | 0.9422         |
| LSTM           | 0.8187         | 0.9013         | 0.9340         |
| BERT           | 0.9148         | 0.9656         | 0.9797         |

資料來源：作者自製。

註：SVM：支援向量機 (support vector machine)；LSTM：長短期記憶 (Long Short-Term Memory)；BERT：基於轉換器的雙向編碼器表徵技術 (Bidirectional Encoder Representations from Transformers)。

## 五、錯誤分析

從實驗數據可知基於深度學習之 LSTM 與 BERT 方法之佛教引用句推薦模型皆已有相當好之效果，然其仍有部分無法正確預測之情況，對於無法正確推薦之錯誤例，以下進行詳細之分析。為觀察深度學習模型之實際表現，我們以先前進行十折交叉驗證之實驗中隨機挑選其中一個組合進行錯誤項目之觀察，LSTM 與 BERT 方法皆針對同一組合之結果進行分析。

LSTM 模型對於實驗資料集之 18 句佛教引用句進行推薦之預測結果，其錯誤數量分別如表 8 所示，表中之錯誤數為於該標籤類別之測試資料中，模型未能正確分類為與測試資料中之標籤相符之數量。其數據以圖表方式呈現如圖 4，另以混淆矩陣 (Confusion Matrix) 方式呈現於表 9，顯示所有標籤之錯誤分布。

觀察其於 18 句各句之預測結果之錯誤情形，共可歸納為兩種主要原因。其一為 18 句中之部分佛教引用句其意涵相近，在文章引用時其前後文描述十分接近，造成模型不易區分此類引用句，因而產生分類錯誤。於表 8 中，錯誤率較高的引用句前幾項之中有標籤 0、標籤 1、標籤 3、標籤 7、標籤 11、標籤 12。其實際之引用句原文條列如下：

- (一) 標籤 0：一切有為法，如夢幻泡影
- (二) 標籤 1：凡所有相，皆是虛妄
- (三) 標籤 3：應無所住，而生其心
- (四) 標籤 7：不取於相，如如不動

表 8 LSTM 模型預測錯誤數量

| 標籤    | 0    | 1    | 2    | 3    | 4    | 5    | 6    | 7    | 8    |
|-------|------|------|------|------|------|------|------|------|------|
| 總數    | 590  | 775  | 329  | 923  | 325  | 600  | 445  | 328  | 164  |
| 錯誤數   | 107  | 158  | 40   | 157  | 10   | 48   | 25   | 80   | 22   |
| 錯誤率 % | 18.1 | 20.4 | 12.2 | 17.0 | 3.1  | 8.0  | 5.6  | 24.4 | 13.4 |
| 標籤    | 9    | 10   | 11   | 12   | 13   | 14   | 15   | 16   | 17   |
| 總數    | 163  | 236  | 290  | 518  | 65   | 182  | 25   | 164  | 243  |
| 錯誤數   | 44   | 26   | 114  | 111  | 30   | 44   | 9    | 11   | 68   |
| 錯誤率 % | 27.0 | 11.0 | 39.3 | 21.4 | 46.2 | 24.2 | 36.7 | 6.7  | 28.0 |

資料來源：作者自製。



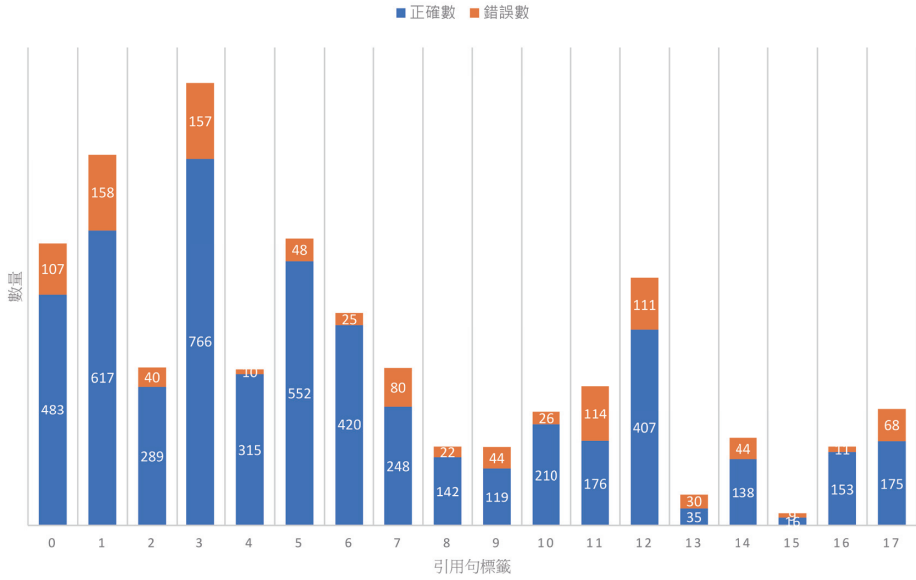


圖 4 LSTM 模型預測錯誤分析圖

資料來源：作者自製。

表 9 LSTM 模型預測結果之混淆矩陣

| Actual | Predicted |     |     |     |     |     |     |     |     |     |     |     |     |    |     |    |     |     |
|--------|-----------|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|----|-----|----|-----|-----|
|        | 0         | 1   | 2   | 3   | 4   | 5   | 6   | 7   | 8   | 9   | 10  | 11  | 12  | 13 | 14  | 15 | 16  | 17  |
| 0      | 483       | 35  | 3   | 20  | 2   | 5   | 1   | 11  | 3   | 1   | 4   | 8   | 12  | 0  | 0   | 0  | 0   | 2   |
| 1      | 24        | 617 | 4   | 44  | 3   | 9   | 2   | 15  | 4   | 0   | 4   | 18  | 16  | 0  | 2   | 0  | 0   | 13  |
| 2      | 3         | 8   | 289 | 11  | 0   | 0   | 2   | 1   | 0   | 0   | 0   | 3   | 9   | 2  | 0   | 0  | 1   | 0   |
| 3      | 11        | 39  | 7   | 766 | 0   | 9   | 0   | 9   | 1   | 3   | 5   | 23  | 23  | 3  | 10  | 1  | 0   | 13  |
| 4      | 0         | 2   | 0   | 2   | 315 | 0   | 1   | 0   | 0   | 0   | 0   | 0   | 3   | 0  | 0   | 0  | 1   | 1   |
| 5      | 6         | 11  | 1   | 7   | 0   | 552 | 6   | 2   | 0   | 0   | 3   | 1   | 8   | 0  | 0   | 0  | 0   | 3   |
| 6      | 1         | 1   | 0   | 2   | 0   | 6   | 420 | 0   | 0   | 1   | 2   | 2   | 5   | 0  | 0   | 1  | 0   | 4   |
| 7      | 17        | 24  | 3   | 15  | 0   | 0   | 1   | 248 | 0   | 2   | 0   | 11  | 2   | 1  | 1   | 0  | 0   | 3   |
| 8      | 2         | 4   | 0   | 3   | 1   | 1   | 0   | 1   | 142 | 2   | 0   | 2   | 1   | 2  | 0   | 0  | 3   | 0   |
| 9      | 1         | 3   | 1   | 6   | 0   | 0   | 3   | 1   | 0   | 119 | 0   | 17  | 0   | 1  | 1   | 1  | 2   | 7   |
| 10     | 3         | 4   | 0   | 5   | 0   | 3   | 4   | 1   | 0   | 0   | 210 | 1   | 4   | 0  | 1   | 0  | 0   | 0   |
| 11     | 11        | 32  | 4   | 20  | 2   | 6   | 2   | 4   | 1   | 5   | 0   | 176 | 14  | 3  | 2   | 0  | 1   | 7   |
| 12     | 5         | 18  | 9   | 34  | 0   | 12  | 1   | 3   | 0   | 1   | 2   | 12  | 407 | 1  | 9   | 1  | 0   | 3   |
| 13     | 3         | 4   | 1   | 5   | 0   | 0   | 1   | 0   | 1   | 4   | 0   | 0   | 5   | 35 | 3   | 0  | 0   | 3   |
| 14     | 2         | 1   | 0   | 20  | 0   | 2   | 0   | 2   | 0   | 0   | 1   | 1   | 13  | 1  | 138 | 0  | 0   | 1   |
| 15     | 0         | 1   | 1   | 2   | 0   | 1   | 0   | 1   | 0   | 2   | 0   | 1   | 0   | 0  | 0   | 16 | 0   | 0   |
| 16     | 1         | 1   | 1   | 3   | 0   | 0   | 2   | 0   | 1   | 0   | 0   | 1   | 0   | 0  | 1   | 0  | 153 | 0   |
| 17     | 1         | 11  | 2   | 7   | 0   | 4   | 2   | 6   | 3   | 7   | 2   | 11  | 3   | 6  | 1   | 2  | 0   | 175 |

資料來源：作者自製。

(五) 標籤 11：法尚應捨，何況非法

(六) 標籤 12：本來無一物，何處惹塵埃

此六句佛教引用句，皆與佛教之禪宗思想有關，多出自於《金剛般若波羅蜜經》之內容。標籤 0「一切有為法，如夢幻泡影」，其中之一切有為法是世出世間一切諸法包括起心動念的心，心也是有為法。夢如幻、泡、影三個字，是形容夢之不實。而標籤 1「凡所有相，皆是虛妄」，凡所有相，就是前句之一切有為法，其皆為虛妄的意思，是暫時的、臨時的，如夢幻、泡、影。標籤 3「應無所住，而生其心」，應無所住之「無住」妙旨，全在「不取於相，如如不動」上，因為凡所有相皆是虛妄！此句是把為何「不取於相，如如不動」的真相說出來。標籤 7「不取於相，如如不動」，即表示我們今天見色聞聲，六根接觸六塵境界就執著、就著相，外不著相，內不動心，其心即為清淨心。標籤 11「法尚應捨，何況非法」，其強調就算是如來所說法，亦不可執著，不可著相於此，如同標籤 3「應無所住，而生其心」。標籤 12「本來無一物，何處惹塵埃」，本來無一物就是無念，心有所住，就是妄念，六塵競起，如能一念不生，覺性現前，便是標籤 3 的「應無所住，而生其心」。從佛學義理進行深入之探究，此五句所談及之概念彼此互為表裡，相互關連，因此對模型區分來說確實較為困難。

我們再進一步仔細觀察此五句佛教引用句其中各句之推薦結果之錯誤分布。標籤 0「一切有為法，如夢幻泡影」其中分類錯誤之分析如圖 5 所示，由圖中可以看出原本應為標籤 0 者，模型將其錯誤分類至其它五句，即標籤 1、3、7、11、12 的比例相對較多。而圖 6 為標籤 1「凡所有相，皆是虛妄」的分類錯誤之分析，從圖中亦可看出，其分類之錯誤分布於其它五句標籤 0、3、7、11、12 的比例亦相對較高。

第二種預測結果之錯誤主要因為訓練資料量較少，使得模型未能足夠學習進行有效預測。於表 8 中之錯誤數量中，其中標籤 13 與標籤 15 即屬於此類型。由於建置訓練資料集時所能收集之數量與其它引用句相較少很多，標籤 13 僅 65 筆，標籤 15 更僅有 25 筆，此造成對於這兩句引用句的預測結果錯誤率亦較高。以標籤 13 為例，細部觀察其錯誤之分布情形如圖 7 所示，可見其錯誤預測為其它引用句標籤之分布較為分散，未特定集中於某些引用句標籤。

標籤0：一切有為法，如夢幻泡影  
錯誤例分布

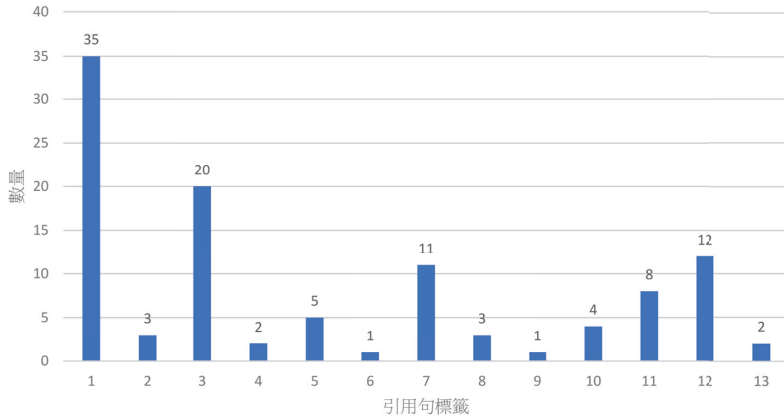


圖 5 標籤 0 「一切有為法，如夢幻泡影」之錯誤情形

資料來源：作者自製。

標籤1：凡所有相，皆是虛妄  
錯誤例分布

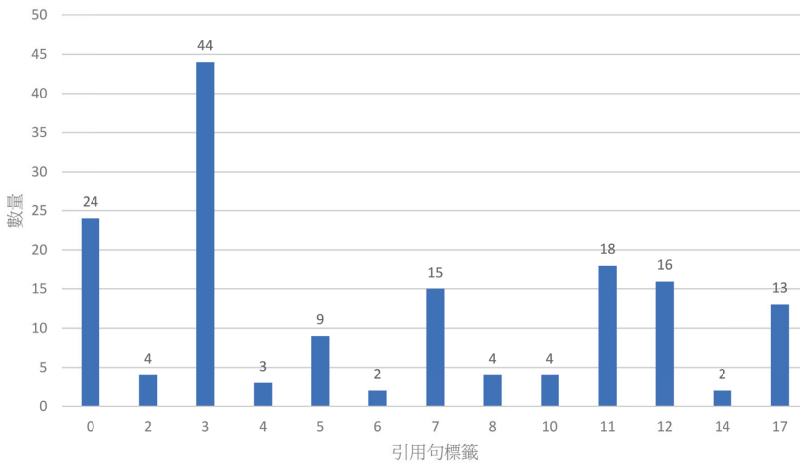


圖 6 標籤 1 「凡所有相，皆是虛妄」之錯誤情形

資料來源：作者自製。

從實驗結果可知，BERT 模型之表現整體大幅優於 LSTM 模型之效果。對於 BERT 模型之錯誤例之數量，亦同樣統計於表 10，圖 8 為錯誤例之分析圖表。

標籤13：一念悟時，眾生是佛

錯誤例分布

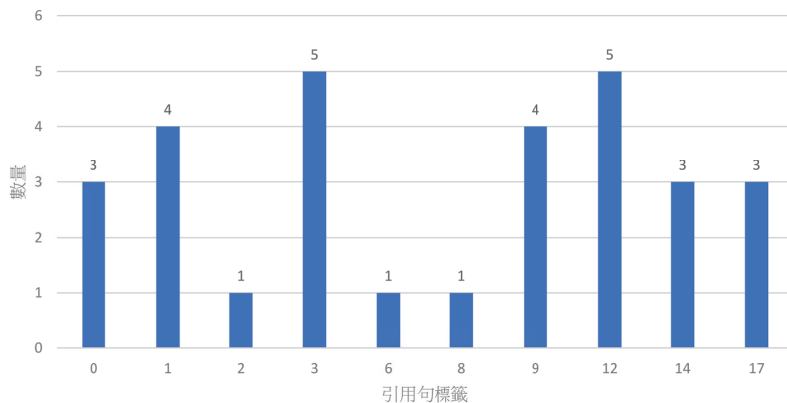


圖 7 標籤 13 「一念悟時，眾生是佛」之錯誤情形

資料來源：作者自製。

表 10 BERT 模型預測錯誤數量

| 標籤    | 0    | 1   | 2    | 3   | 4    | 5   | 6   | 7    | 8    |
|-------|------|-----|------|-----|------|-----|-----|------|------|
| 總數    | 590  | 775 | 329  | 923 | 325  | 600 | 445 | 328  | 164  |
| 錯誤數   | 79   | 68  | 26   | 91  | 1    | 25  | 12  | 45   | 5    |
| 錯誤率 % | 13.4 | 8.8 | 7.9  | 9.9 | 0.3  | 4.2 | 2.7 | 13.7 | 3.0  |
| 標籤    | 9    | 10  | 11   | 12  | 13   | 14  | 15  | 16   | 17   |
| 總數    | 163  | 236 | 290  | 518 | 65   | 182 | 25  | 164  | 243  |
| 錯誤數   | 17   | 10  | 50   | 49  | 9    | 15  | 0   | 4    | 36   |
| 錯誤率 % | 10.4 | 4.2 | 17.2 | 9.5 | 13.8 | 8.2 | 0.0 | 2.4  | 14.8 |

資料來源：作者自製。

從 BERT 模型之錯誤結果分析亦可看出，對於第一種錯誤類型，即佛教引用句意涵相近之情況，BERT 模型於標籤 0、1、3、7、11 該五句引用句其錯誤率仍較其它標籤來得高，但與 LSTM 模型相較，其錯誤數量皆有明顯減少。對於第二種錯誤類型，即訓練資料不足之情況，BERT 則有大幅的改善，標籤 13 之錯誤率從 LSTM 的 46% 降至 13.8%，而標籤 15 之錯誤率則降為 0%。我們將 LSTM 模型與 BERT 模型於各引用句標籤之錯誤類型並列於表 11。

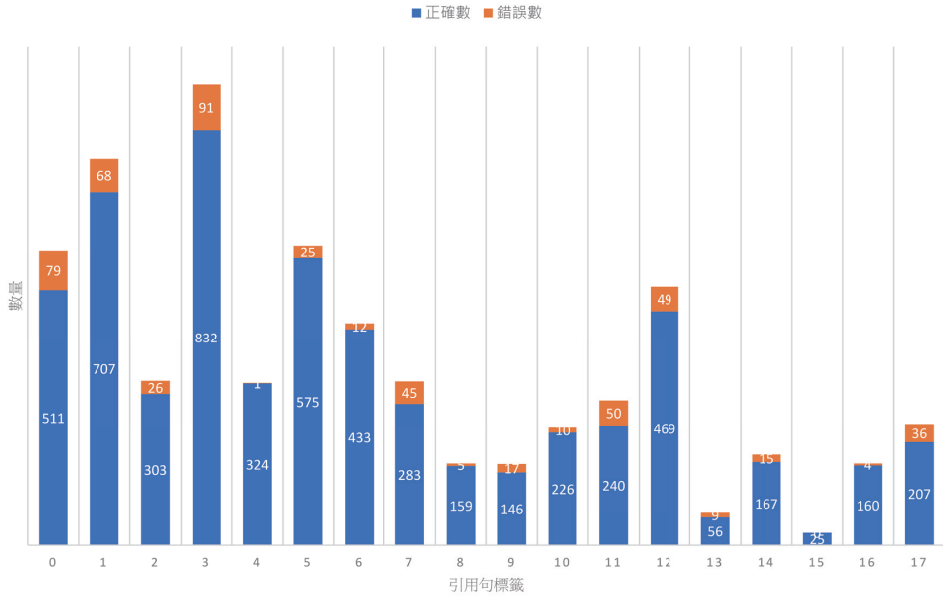


圖 8 BERT 模型預測錯誤分析圖

資料來源：作者自製。

表 11 LSTM 模型與 BERT 模型錯誤比較

| 標籤    | 0    | 1    | 2    | 3    | 4    | 5    | 6    | 7    | 8    |
|-------|------|------|------|------|------|------|------|------|------|
| 總數    | 590  | 775  | 329  | 923  | 325  | 600  | 445  | 328  | 164  |
| 錯誤數   |      |      |      |      |      |      |      |      |      |
| LSTM  | 107  | 158  | 40   | 157  | 10   | 48   | 25   | 80   | 22   |
| BERT  | 79   | 68   | 26   | 91   | 1    | 25   | 12   | 45   | 5    |
| 錯誤率 % |      |      |      |      |      |      |      |      |      |
| LSTM  | 18.1 | 20.4 | 12.2 | 17.0 | 3.1  | 8.0  | 5.6  | 24.4 | 13.4 |
| BERT  | 13.4 | 8.8  | 7.9  | 9.9  | 0.3  | 4.2  | 2.7  | 13.7 | 3.0  |
| 標籤    | 9    | 10   | 11   | 12   | 13   | 14   | 15   | 16   | 17   |
| 總數    | 163  | 236  | 290  | 518  | 65   | 182  | 25   | 164  | 243  |
| 錯誤數   |      |      |      |      |      |      |      |      |      |
| LSTM  | 44   | 26   | 114  | 111  | 30   | 44   | 9    | 11   | 68   |
| BERT  | 17   | 10   | 50   | 49   | 9    | 15   | 0    | 4    | 36   |
| 錯誤率 % |      |      |      |      |      |      |      |      |      |
| LSTM  | 27.0 | 11.0 | 39.3 | 21.4 | 46.2 | 24.2 | 36.7 | 6.7  | 28.0 |
| BERT  | 10.4 | 4.2  | 17.2 | 9.5  | 13.8 | 8.2  | 0.0  | 2.4  | 14.8 |

資料來源：作者自製。

從表中可看出，BERT 對於每一個佛教引用句之錯誤數量相較 LSTM 模型皆有一定程度之減少，且對於訓練資料顯著不足之標籤 13 與 15 兩句之錯誤率改善最為明顯。結果顯示 BERT 透過大量中文語料進行預訓練學習語言之表達方式，再利用佛教引用句訓練資料集進行遷移學習，能有效克服訓練資料過少不平均之難題，提升佛教引用句之推薦表現。

## 伍、結論

在寫作之中，適度引用名言佳句是重要的撰寫技巧，而佛教典籍之中有非常多深具哲理與啟發之名句，然對於現代人來說，要在寫作時引用佛教相關之引用句實非易事。因此能於文章寫作時依據當前內容自動推薦合適的佛教引用句即為一重要的需求與研究課題，本論文即針對佛教引用句推薦問題，提出基於 LSTM 與 BERT 兩種深度學習方法之佛教引用句推薦系統，可自動從文章內容進行分析推薦合適的佛教引用句。我們自文獻中常見之佛教引用句之中，擇定 18 句，並透過網路爬蟲建置包括引用句之上下文內容之訓練資料集。實驗結果顯示 LSTM 模型之 Top-1 準確率達到 0.8187，而 BERT 模型之 Top-1 準確率高達 0.9148，顯見我們所提出之佛教引用句推薦系統之有效性。針對模型推薦結果進行錯誤分析，主要可分為兩種主要錯誤類型，其一為許多佛教引用句彼此意涵於義理上十分近似，致使深度學習模型較不易區分；另一則為部分佛教引用句所能收集到的訓練資料筆數過少，使得深度學習模型不足以有效進行訓練，而利用大量語言進行預訓練的 BERT 模型較能有效克服訓練資料較少的問題。

佛經語言是非常珍貴且豐富的語料庫，雖目前僅擇定一部分之佛教引用句，然實驗結果已證明本方法之有效性，未來可再增補更多佛教引用句非佛甚或非佛教領域之引用句以擴大資料集數量，拓展應用層面以實現更廣泛的推薦方法，且可進行寫作的支援系統與互文性（intertextuality）之相關研究，期能於各類引用句推薦及佛學研究與推展上創造更多實用價值。

## 參考文獻

- 王皓儀 (2019)。深度學習應用於中文成語與同義成語寫作之推薦 (未出版之碩士論文)。國立臺北科技大學電機工程系, 臺北。
- 美團技術團隊 (2020)。美團點評 2019 評技術年貨：算法篇。取自 [https://s3plus.meituan.net/v1/mss\\_e63d09aec75b41879dcb3069234793ac/file/%E7%AE%97%E6%B3%95%E7%AF%87.pdf](https://s3plus.meituan.net/v1/mss_e63d09aec75b41879dcb3069234793ac/file/%E7%AE%97%E6%B3%95%E7%AF%87.pdf)
- Ahn, Y., Lee, H., Jeon, H., Ha, S., & Lee, S.-g. (2016). *Quote recommendation for dialogs and writings*. Paper presented at the CBRecSys 2016, Boston, MA.
- Aizerman, M. A. (1964). Theoretical foundations of the potential function method in pattern recognition learning. *Automation and remote control*, 25, 821-837.
- Cho, K., Van Merriënboer, B., Gulcehre, C., Bahdanau, D., Bougares, F., Schwenk, H., & Bengio, Y. (2014). Learning phrase representations using RNN encoder-decoder for statistical machine translation. In A. Moschitti, B. Pang, W. Daelemans (Eds.), *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)* (pp. 1724-1734). Doha, Qatar: Association for Computational Linguistics. doi:10.3115/v1/D14-1179
- Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2018). BERT: Pre-training of deep bidirectional transformers for language understanding. *arXiv*, 1810, 04805.
- He, Q., Pei, J., Kifer, D., Mitra, P., & Giles, L. (2010). Context-aware citation recommendation. In *WWW'10: Proceedings of the 19th International Conference on World Wide Web* (pp. 421-430). New York, NY: Association for Computing Machinery. doi:10.1145/1772690.1772734
- Huang, W., Kataria, S., Caragea, C., Mitra, P., Giles, C. L., & Rokach, L. (2012). Recommending citations: Translating papers into references. In *CIKM'12: Proceedings of the 21st ACM International Conference on Information and Knowledge Management* (pp. 1910-1914). New York, NY: Association for Computing Machinery. doi:10.1145/2396761.2398542
- Huang, W., Wu, Z., Liang, C., Mitra, P., & Giles, C. L. (2015). A neural probabilistic model for context based citation recommendation. In *Proceedings of the Twenty-Ninth AAAI Conference on Artificial Intelligence*

- (pp. 2404-2410). Palo Alto, CA: AAAI Press.
- Lee, H., Ahn, Y., Lee, H., Ha, S., & Lee, S.-g. (2016). Quote recommendation in dialogue using deep neural network. In *SIGIR '16: Proceedings of the 39th International ACM SIGIR Conference on Research and Development in Information Retrieval* (pp. 957-960). New York, NY: Association for Computing Machinery. doi:10.1145/2911451.2914734
- Liu, Y., Liu, B., Shan, L., & Wang, X. (2018). Modelling context with neural networks for recommending idioms in essay writing. *Neurocomputing*, 275, 2287-2293. doi:10.1016/j.neucom.2017.11.005
- Liu, Y., Pang, B., & Liu, B. (2019). Neural-based Chinese idiom recommendation for enhancing elegance in essay writing. In A. Korhonen, D. Traum, & L. Màrquez (Eds.), *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics* (pp. 5522-5526). Florence, Italy: Association for Computational Linguistics. doi:10.18653/v1/P19-1552
- Pazzani, M. J., & Billsus, D. (2007). Content-based recommendation systems. In P. Brusilovsky, A. Kobsa, & W. Nejdl (Eds.), *The adaptive web: Methods and strategies of web personalization* (pp. 325-341). Berlin, Germany: Springer. doi:10.1007/978-3-540-72079-9\_10
- Strohman, T., Croft, W. B., & Jensen, D. (2007). Recommending citations for academic papers. In *SIGIR '07: Proceedings of the 30th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval* (pp. 705-706). New York, NY: Association for Computing Machinery. doi:10.1145/1277741.1277868
- Tan, J., Wan, X., Liu, H., & Xiao, J. (2018). QuoteRec: Toward quote recommendation for writing. *ACM Transactions on Information Systems*, 36(3), 1-36. doi:10.1145/3183370
- Tan, J., Wan, X., & Xiao, J. (2015). Learning to recommend quotes for writing. *Proceedings of the AAAI Conference on Artificial Intelligence*, 29(1), 2453-2459. doi:10.1609/aaai.v29i1.9530
- Tan, J., Wan, X., & Xiao, J. (2016). A neural network approach to quote recommendation in writings. In *CIKM'16: Proceedings of the 25th ACM International on Conference on Information and Knowledge Management* (pp. 65-74). New York, NY: Association for Computing Machinery. doi:10.1145/2983323.2983788



- Vapnik, V. N. (1999). *The nature of statistical learning theory* (2nd ed.). New York, NY: Springer. doi:10.1007/978-1-4757-3264-1
- Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., ... Polosukhin, I. (2018). Attention is all you need. In I. Guyon, U. Von Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, & R. Garnett (Eds.), *Advances in neural information processing systems 30* (pp. 5999-6009). Red Hook, NY: Curran Associates, Inc.
- Wang, L., Li, J., Zeng, X., Zhang, H., & Wong, K.-F. (2020). Continuity of topic, interaction, and query: Learning to quote in online conversations. In B. Webber, T. Cohn, Y. He, & Y. Liu (Eds.), *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)* (pp. 6640-6650). Association for Computational Linguistics. doi:10.18653/v1/2020.emnlp-main.538
- Wang, L., Zeng, X., & Wong, K.-F. (2021). Quotation recommendation and interpretation based on transformation from queries to quotations. In C. Zong, F. Xia, W. Li, & R. Navigli (Eds.), *Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing (Volume 2: Short Papers)* (pp. 754-758). Association for Computational Linguistics. doi:10.18653/v1/2021.acl-short.95

# A Buddhist Quote Recommendation System Based on Deep Learning Approaches

Guanghai Shi<sup>1</sup>, Yu-Chun Wang<sup>2,\*</sup>, Jen-Jou Hung<sup>3</sup>

## Abstract

In writing, citing famous sayings is a common writing technique, which can enhance the persuasiveness of the article and make the article more convincing. Buddhist quotes are one the vital sources of maxims. These Buddhist quotes are full of wisdom and often enlighten people. However, for modern people, knowing how to cite a suitable inspiring Buddhist quote are not an easy job. Therefore, a recommendation system which is able to automatically recommend suitable Buddhist quotes according to the content while writing becomes an urgent demand. This paper proposes a Buddhist quote recommendation system based on two deep learning approaches, such as LSTM and BERT. We compile a data set for Buddhist quote recommendations and then train a deep learning model. The experimental results show that the accuracy of our system achieves 0.9148, which demonstrates our system effectively recommends suitable Buddhist quotes when writing.

**Keywords:** Buddhist quotes, recommendation system, deep learning, LSTM, BERT.

---

Manuscript received: September 28, 2022; Accepted: November 14, 2022

<sup>1</sup> PhD Student, Dharma Drum Institute of Liberal Arts.

<sup>2</sup> Assistant Professor, Dharma Drum Institute of Liberal Arts.

<sup>3</sup> Professor, Dharma Drum Institute of Liberal Arts.

\* Email: ycwang@dila.edu.tw

## Extended Abstract

In writing, moderately quoting famous quotes is a common writing technique. It can not only embellish the literary talent but also enhance the persuasiveness of the article. It can make the created article more beautiful and convincing. Such famous quotes are mostly taken from traditional proverbs, quotations from famous people, or from many philosophical thoughts and religious classics. They are the crystallization of human thought and life experience for thousands of years. However, it is not very easy for ordinary people to properly quote appropriate famous sentences when writing articles. If there is a quote recommendation, which can recommend suitable quotes based on the content of the article, it will solve this writing problem.

In addition to traditional proverbs and famous sayings, meaningful sayings from religious texts such as Buddhist canons or the Christian Bible are often used in the selection of quotes. However, most of the quotes from such religious texts are strongly related to their religious connotations. Unless they are very familiar with such religious texts and their teachings, it is more difficult for ordinary people to use such quotes properly. This paper aims at recommending Buddhist quotes and proposes a method that automatically analyzes and recommends appropriate Buddhist quotes from the content of an article. We adopt the deep learning approaches that have been widely introduced in natural language processing, using the language of Long Short-Term Memory (LSTM) and the Bidirectional Encoder Representations from Transformers (BERT) model to construct the Buddhist quote recommendation system.

The recommendation of Buddhist quotes is based on the texts of the article, which are basically written in modern Chinese. For the LSTM-based model, we use the Jieba word segmentation tool with a large-scale lexicon as the word segmentation dictionary. Then, each input token will be converted into a vector representation in the embedding layer, filled up to the maximum input length, and then to LSTM recurrent neural network layer. In the design of the model for quote recommendation, we adopt two designs for the recurrent neural network layer, one is composed of only a single LSTM recurrent neural network, and the other is superimposed with two layers of LSTM to construct a Bidirectional LSTM (BiLSTM), one layer is passed from front to back, and the other layer is passed from back to front. Both layers use the embedded layer as the input. The forward LSTM layer can learn the information from the previous text, while the backward LSTM layer can learn the information from the subsequent text,

which can effectively learn the two-way semantic dependence, and the output results of the two LSTM layers will be concatenated is an output vector. The output of the LSTM layer is then connected to a fully connected layer with 64 neurons and finally connected to a fully connected layer with the number of neurons equivalent to the number of reference sentences set as the final output of the model.

For the BERT model, Chinese texts are not performed word segmentation but directly use a single character as the input unit. BERT is a deep language model based on Transformer architecture which constructs a multi-layer bidirectional encoder neural network. The input sentences will pass through the multi-layer Transformer network constructed from the attention mechanism, it can effectively observe the information of all words in the entire sentence to determine its representation as output. Based on the pre-training from the large-scale training corpora, BERT can effectively represent the information of the input sentence of the language. In the construction of the actual Buddhist quote recommendation model, we fine-tune the pre-trained Chinese BERT language model to the classification task using the Buddhist quotes dataset and transfer it to apply to the recommendation of Buddhist quotes.

In order to build a training data set for Buddhist quote recommendation, we first select a total of 18 quotes that are widely known in writing and general public cognition. Then, using the Buddhist quote as the search keyword, query the webpage articles containing the Buddhist quote from various well-known websites, and then use the web crawler to collect the article. From the article, the texts about 100 characters before and after the quote, about 200 characters in length, are extracted, and the pairing of Buddhist quotes and their context are generated as items of the training data set.

From the experiment results, the LSTM model and the BERT model are on average better than the baseline SVM model on average. The top-1 accuracy of the LSTM model is 81.87%, while the top-3 accuracy has reached 90.13%. Compared with the other two models, the BERT model achieves the best results. Its top-1 accuracy has reached 91.48%, and its top-5 accuracy is as high as 97.97%. The results show the effectiveness of our proposed approach for recommending appropriate Chinese Buddhist quotes.

Although the accuracy of our proposed method is very high, there is still some kind of error. The errors can be divided into two main types. One is that many Buddhist quotes are very similar in meaning, which makes it difficult for

the models to discriminate. Of the 18 quotes in our dataset, five quotes tend to be more error-prone than others. These include some famous quotes from the Diamond Sutra related to the school of Zen Buddhism, such as “All things contrived are like dream, illusion, bubble, and shadow” (一切有為法，如夢幻泡影), “Everything with form is unreal” (凡所有相，皆是虛妄), and “They should develop a mind which does not abide in anything” (應無所住，而生其心). The main topic of these quotes is about not clinging to unreal material forms. A deeper study of Buddhist doctrine reveals that the concepts discussed in these five quotations are interrelated, making it difficult to distinguish the models from the surrounding contexts.

The second case is that some quotes in our data set have a small amount of training data because there are not enough contexts crawled from the web. Two of the 18 quotes in our dataset have only 65 and 25 instances, respectively. Due to the insufficient amount of training data for the two quotes, the deep learning approaches, especially the LSTM-based models, cannot effectively learn a suitable classifier. The BERT-based approach, which relies on pre-training from large-scale corpora, can effectively overcome this shortage problem of less training data.

Chinese Buddhist scriptures are very precious and rich corpora that contain the essence of human wisdom. Our proposed Chinese Buddhist quote recommendation method can effectively recommend appropriate Buddhist quotes based on the context during writing. Although the size of the current data set is limited, the experiment results have proven the effectiveness of our proposed methods. In the future, we will increase the number of quotes and add more Buddhist and other religious quotes to broaden the application of our recommendation method. The quote recommendation approach can also be extended to the writing support system and has advantages for the study of intertextuality. Digital techniques can support the research of Buddhist studies and promote the application of religious scriptures to create more practical value in the current era.