

國立政治大學資訊科學系

碩士論文

基於師生方法學習多層次注意力的

跨領域轉移學習

A Teacher-Student Approach to Cross-domain  
Transfer Learning with Multi-level Attention

指導教授：黃瀚萱 博士

研究生：唐英哲 撰

中華民國 110 年 7 月

## 摘要

本研究應用於跨領域轉移問題上。跨領域轉移問題希望能解決在一個領域資料利用機器學習訓練模型，並將此訓練後的模型應用於其他不同領域的資料。跨領域問題的困難處在於源領域以及目標領域之間的差異，如"快"這個形容詞在跑車產品是好的形容詞，但在電池產品卻是不好的形容詞。在機器學習的問題中，利用已標記資料訓練模型已能達到非常好的效能，但更多情況是沒有足夠的已標記資料訓練模型。基於上述原因，本研究希望可以建立一個既可以解決跨領域轉移問題，又可以解決已標記資料量少的模型。

模型架構可以分為三個部分的多任務學習，分別為監督式學習、師生跨領域轉移注意力模型以及相關度偵測任務。監督式學習使用資料及標籤輸入模型進行學習。師生跨領域轉移模型由教師模型提供學生模型訓練的偽標記資料，學生模型藉由資料層級注意力和領域層級注意力的幫助，為學生模型篩選出適合訓練的偽標記資料。相關度偵測任務用來偵測句子與描述主體之間的關係。

本研究應用於產品意見的情緒立場判斷以及藝人與核能的網路輿情立場判斷問題，實驗結果顯示使用本研究的方法能夠在上述的情緒及輿情立場的分類任務都能達到最好的效能。

## 關鍵字

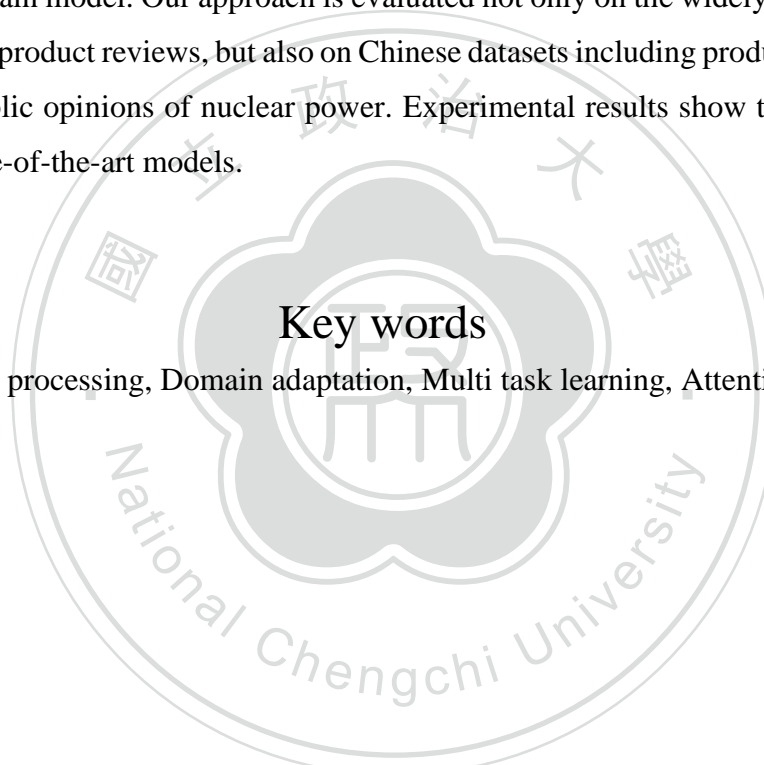
自然語言學習、跨領域轉移問題、多任務學習、注意力機制

## Abstract

The lack of training data forms a challenging issue for applying NLP models in a new domain. Previous work on cross-domain transfer learning aims to exploit the information from the source domains to do prediction for the target domain. To reduce the noises from the out-of-domain data and improve the model's generalization ability, this work proposes a novel teacher-student approach with multi-task learning that transfers the information from source domains to the target domain with sophisticated weights determined by using the attention mechanism at both instance level and domain level. The generalization ability is further enhanced by unsupervised data augmentation. We also introduce a subject detection task for co-training the main model. Our approach is evaluated not only on the widely-adopted English dataset, Amazon product reviews, but also on Chinese datasets including product reviews, artist reviews, and public opinions of nuclear power. Experimental results show that our approach outperforms state-of-the-art models.

### Key words

Natural language processing, Domain adaptation, Multi task learning, Attention mechanism



# 目次

第一章 緒論.....	7
第一節 研究背景.....	7
第二節 研究動機.....	11
第三節 研究目的.....	13
第二章 文獻探討.....	15
第一節 資料增強.....	15
第一項 基礎資料增強.....	15
第二項 非監督式資料增強學習.....	15
第三項 階層式資料增強.....	16
第二節 跨領域轉移問題.....	17
第一項 模型特徵中心法.....	18
第二項 模型損失法.....	19
第三項 資料中心法.....	20
第四項 混合及其他方法.....	21
第三章 研究方法.....	23
第一節 整體模型架構.....	23
第二節 師生跨領域訓練模型架構.....	24
第一項 教師模型.....	24
第二項 學生模型與多層次注意力訓練.....	28
第三節 相關度偵測訓練.....	30
第四節 目標領域模型.....	32
第五節 句子對分類任務.....	34
第四章 資料集、實驗配置與實驗結果.....	35
第一節 資料集.....	35
第二節 實驗配置.....	37
第三節 模型效能.....	40
第四節 模型參數分析.....	48
第五節 模型訓練資料數量分析.....	51
第六節 模型使用情境分析.....	54
第五章 結論.....	56
參考文獻.....	57

## 表次

表 1 負採樣資料範例，以記憶卡為正採樣資料.....	30
表 2 英文實驗四個領域中已標記及未標記資料的數量.....	35
表 3 中文實驗五個領域中已標記及未標記資料的數量.....	36
表 4 使用 BERT 模型作為監督式領域內訓練的基礎模型效能.....	40
表 5 UDA (Unsupervised Data Augmentation)的效能.....	40
表 6 使用 BERT 作監督式跨領域訓練的效能.....	41
表 7 目標領域僅包含未標記資料的設定下與其他方法作比較，結果以 Accuracy (%) 呈現.....	42
表 8 目標領域包含已標記資料及未標記資料的模型效能 (未導入相關度偵測任務).....	43
表 9 加入相關度偵測任務的最終模型效能.....	43
表 10 KKBOX 藝人評論領域作為目標領域實驗.....	45
表 11 核能輿論為目標領域實驗.....	47
表 12 英文資料實驗最佳模型參數.....	49
表 13 中文資料實驗最佳模型參數.....	50
表 14 Books 領域作為目標領域不同資料量效能.....	51
表 15 DVD 領域作為目標領域不同資料量效能.....	51
表 16 Electronic 領域作為目標領域不同資料量效能.....	52
表 17 Kitchen appliances 領域作為目標領域不同資料量效能.....	52
表 18 核能輿論領域作為目標領域訓練資料量分析.....	53

## 圖次

圖 1 監督式領域內訓練.....	8
圖 2 非監督式領域內訓練.....	8
圖 3 半監督式領域內訓練.....	8
圖 4 監督式跨領域轉移訓練.....	9
圖 5 非監督式跨領域轉移訓練.....	9
圖 6 半監督式跨領域轉移訓練.....	9
圖 7 模型多任務學習架構圖.....	23
圖 8 源領域模型架構圖.....	26
圖 9 非監督式資料增強學習示意圖.....	27
圖 10 跨領域模型架構圖.....	29
圖 11 整體模型架構圖.....	33
圖 12 整體模型示意圖.....	34



# 第一章緒論

## 第一節 研究背景

隨著自然語言處理技術 (Nature Language Processing) 的發展，我們可以利用文字以及標記 (答案) 的配對輸入模型訓練，便可以解決許多問題，如分類問題 (情緒分類、商品分類)、序列標記 (詞性標記)、或是問答的問題 (根據問題由模型找出解答)。利用標記過的資料對模型進行監督式訓練，在各種問題上已能達到非常好的效能 [20]。但以上是建立在有大量標記資料的情況，有更多的情況是只有少數的已標記資料，卻有大量的未標記資料。若是要標記這些未標記資料，將會耗費大量的人力及時間成本。而針對擁有大量未標記資料的狀況，可以利用半監督式學習，即利用已標記資料訓練模型，再利用此模型對未標記資料進行偽標記 (Pseudo labeling)，相當於給與這些未標記資料答案，最後使用已標記資料以及偽標記資料一起訓練模型，此即半監督式學習。本研究遇到的情況便是只有少許的已標記資料，卻有大筆的未標記資料，本研究的作法便是基於上述半監督式學習為基礎所發展。

跨領域問題即利用源領域的資料訓練模型，並將此模型用以預測目標領域的資料。在解決同一領域內的問題，已有相當多的研究且能達到好的效果。但跨領域問題的困難之處在於如何消除源領域以及目標領域之間的分歧，舉例來說，Apple 在不同領域可以代表不同意思，在 3C 產品中的 Apple 代表一家電腦公司，Apple 也可以代表的是一種水果。

而同領域的分類問題可以依照資料的標記與否分為監督式訓練 (圖 1)、非監督式訓練 (圖 2) 及半監督式訓練 (圖 3) 三種。跨領域分類問題依據目標領域資料有無標記又可以區分為三種，第一種為目標領域僅有已標記資料的監督式跨領域轉移訓練 (圖 4)，第二種為目標領域僅有未標記資料的非監督式跨領域轉移訓練 (圖 5)，第三種為目標領域既有已標記資料，也有未標記資料的半監督式跨領域轉移訓練 (圖 6)。本研究遇到的跨領域問題為上述的非監督式跨領域轉移訓練和半監督式跨領域轉移訓練，並以領域內的監督式訓練、半監督式訓練及監督式跨領域轉移訓練作為本研究的比較實驗。本研究事先利用已標記資料做監督式訓練，搭配未標記資料一致性訓練的方法，用來訓練源領域資料。對目標領域中的標記資料使用監督式訓練，針對目標領域的未標記資料則利用資料層級以及領域層級的注意力機制，建立目標領域資料與源領域資料之間的連結，找出目標領域資料對各個源領域的注意力，最後做

出預測。相當於另一種監督式訓練與非監督式訓練的結合，解決僅有少量資料及大量未標記資料的跨領域問題。

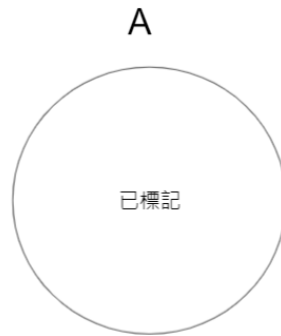


圖 1 監督式領域內訓練



圖 2 非監督式領域內訓練

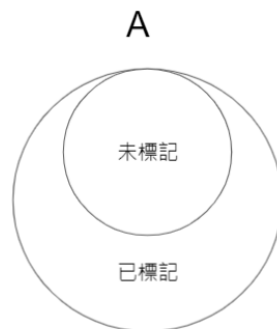


圖 3 半監督式領域內訓練



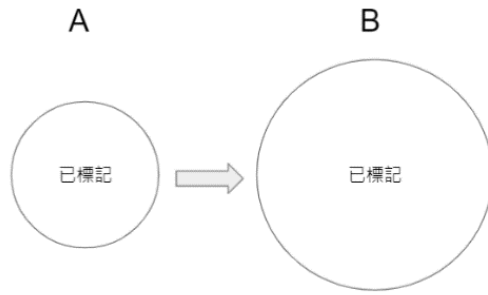


圖 4 監督式跨領域轉移訓練

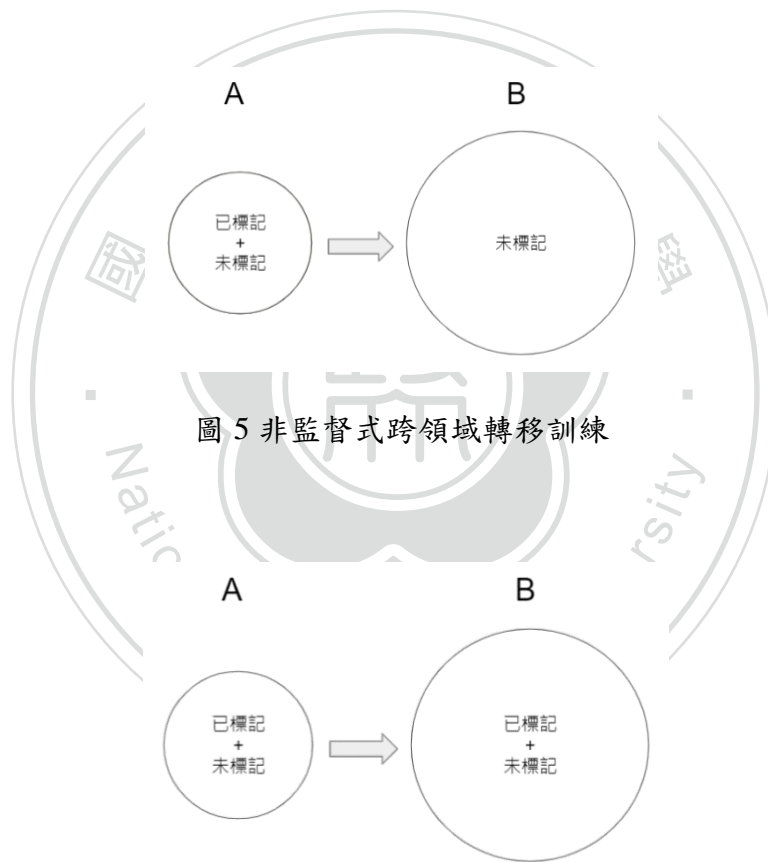


圖 5 非監督式跨領域轉移訓練

圖 6 半監督式跨領域轉移訓練

本研究希望利用 BERT (Bidirectional Encoder Representations from Transformers) 作為模型的 Encoder [11]。BERT 為 Google 所開發的一個用於自然語言處理的模型，可以用來解決單一句子分類問題、句子對分類問題、單一句子標記任務、問答任務。於最近更推出了跨語言的 BERT，可以應用在 109 種多國語言上，在許多典型的任務中取得極佳的效能。BERT 在許多自然語言處理的問題上，都得到較之前作法突破性的成績，故於本篇中使用 BERT 模型作為分類問題的 Encoder。



## 第二節 研究動機

跨領域問題希望可以利用在源領域訓練的模型，預測目標領域的資料。在訓練資料以及測試資料在同一領域或是相似分佈的情況下所訓練的模型，已有許多研究能達到相當好的效能，但若源領域或是目標領域是不相關的領域或是各自為獨立的領域，要如何消弭源領域以及目標領域的歧異以及由源領域找到適合預測目標領域資料的特徵是非常重要的。

若每遇到一個新的領域便要為其訓練模型並應用，將會耗費大量人力及時間，且若領域資料量不足時，便難以訓練出好的模型，因此本研究希望能夠發展出一個能夠通用於所有立場分類任務的模型，不再侷限於某些特定領域的使用情景。

在 [8] 中，雖能有效的解決跨領域轉移問題，但是適用情景為目標領域僅有未標記資料，但本研究方法希望能同時應用於目標領域既有已標記資料，也擁有未標記資料的情況。在擁有大量標記資料的情況下，許多研究成果都能達到非常好的效果，但現實狀況中，全部資料已被標記的情況卻非常少，更多情況是擁有大量未標記資料，卻只有少量的已標記資料。面對只有少量已標記資料的情況，可以利用資料增強技術來增加標記資料的數量。在 [30] 中提到，利用同義詞替換或是打亂資料中文字順序等方式，可以創造出更多講法不同，但表達意義相同的資料，以此獲得更多已標記資料，解決原本已標記資料不足的問題。

若是要將大量的未標記資料標記完，勢必需要花費大量的時間以及人力成本。而在 [32] 中提出一種能同時訓練已標記資料及未標記資料的架構。此架構希望能利用大量的未標記資料搭配資料增強技術進行一致性訓練，此舉能夠讓模型能學習對不同講法但意思相同的兩個句子做出相同的預測。上述的作法能夠解決僅有少量已標記資料卻有大量未標記資料的問題。文中提及未標記資料較容易取得，但有可能也會使用到別的領域的未標記資料，而跨領域的未標記資料與源領域的資料的分佈可能會有落差，因此文中作法是使用在源領域訓練好的模型對這些未標記資料進行預測，並設定一個臨界值作為篩選偽標記資料的條件，若預測機率高於此臨界值才把此偽標記資料加入訓練集訓練模型。但上述針對跨領域資料，若只使用設定臨界值篩選的作法，將會忽略掉跨領域資料之間關聯程度，或何種領域資料較適合拿來做轉移，領域及領域之間的相關程度為何等問題，且文中沒有特別提及如擁有多個源領域時，該選擇何種領域資料做轉移，因此若喪失這些資訊會使跨領域轉移的效能無法發揮到最大。

本研究希望能發展出利用注意力機制找出適合用來進行跨領域轉移學習的資料，並利用此注意力機制為多任務學習的輔助任務中帶來幫助。基於上述原因，在此歸納出研究動機：

- 本研究希望能建立一個適用於所有分類任務的跨領域模型。
- 本研究希望能建立一個能應用於源領域及目標領域皆含有少量已標記資料及大量未標記資料的跨領域模型。
- 本研究希望利用注意力機制為跨領域任務及其他多任務學習的輔助任務帶來幫助。



### 第三節 研究目的

缺少足夠的訓練資料對跨領域轉移問題帶來很大的困難。以往對於跨領域轉移問題的相關研究關注於利用源領域的資訊去預測目標領域資料。為了減少領域及領域之間的歧異，並增加模型的泛化能力，本文提出了一個全新的師生跨領域模型，使源領域資訊能藉由資料級別注意力機制和領域級別注意力機制轉移至目標領域資料，同時藉由非監督式資料增強學習的方法增加模型的泛化能力。跨領域轉移問題中使用來自源領域的資料來訓練模型，並將此模型用來預測目標領域的資料。源領域和目標領域之間語義表達的不同為跨領域轉移問題的主要障礙，以情感分析任務為例，電影的評論和電子設備的評論有所不同，好電影的形容詞可能不適用於描述電子設備的正面特性，因此可能造成訓練於電影評論情緒分析的模型不適用於預測用戶對咖啡機（電子設備）的情況。近年來已經探索了許多方法，包括用於領域轉換的詞嵌入 [13]，對抗式的轉移學習 [14]，知識型引導的轉移學習 [15]。

本研究提出一個全新的模型架構解決跨領域轉移問題，利用師生跨領域模型將源領域的知識轉移到目標領域。教師模型由所有源領域模型所構成，而教師模型用來為目標領域的未標記資料做預測，並生成偽標記資料供學生模型做訓練。學生模型會和兩種層次的注意力機制合作，分別是資料級別的注意力機制和領域級別的注意力機制，此兩種注意力機制能夠篩選出適合目標領域偽標記資料的源領域特徵供學生模型學習，以減少源領域以及目標領域之間的歧異。此外，我們採用非監督式資料增強學習來提高教師模型的泛化能力，最終學生模型將使用偽標記的領域內資料進行訓練，並受益於由教師模型轉移的知識。除了上述提出的跨領域模型之外，我們也加入了相關度偵測的任務來做多任務學習。相關度偵測任務用來偵測文字評論及描述主體之間的相關度。實驗結果證明加入相關度偵測的多任務學習能夠提升模型的效能。本文利用 Amazon 公司消費者對商品的回覆資料 [2] 及 KKBOX 公司藝人的評論及核能輿論做實驗，其中包含了英文以及中文的資料，本篇會分別對英文以及中文資料做實驗。在英文資料中包含了四個領域，分別是書籍、DVD、電器用品、廚房用品，此資料集也是跨領域轉移問題的相關研究中經常使用到的資料集。中文資料分為 Amazon 公司消費者對商品的回覆資料及 KKBOX 公司藝人的評論及核能輿論。其中 Amazon 商品中文意見資料包含三種領域，分別是嬰兒用品、體育用品、玩具。在進行英文實驗時，會輪流選擇一個領域資料作為目標領域，剩下的領域資料則作為源領域，故英文資料共有四種實驗組合，

以此避免模型只對某些特定領域資料有效，藉由反覆的實驗來驗證模型的穩定性。除了在英文資料做實驗外，本篇也利用中文資料來驗證模型換了語言之後是否能維持相當效能。在中文實驗中，我們分別使用 KKBOX 藝人評論資料以及核能輿論資料作為目標領域，Amazon 三個領域中文資料作為源領域，在中文實驗結果顯示本文所提出的模型能夠有效處理無論是非監督式跨領域轉移問題或是半監督式跨領域轉移問題，都能在這些問題中達到比現行方法更好的效果。本研究提出的模型貢獻如下：

- 本研究提出了一個全新的師生跨領域模型，使目標領域可以使用到來自源領域且不受領域間噪音干擾的知識。
- 本研究提出的模型架構效能，能夠勝過目前解決跨領域轉移問題的相關方法，達到最好的效果。
- 本研究所提出之模型架構能夠與複雜的神經網路結合並應用於其他任務。
- 本研究所提出之模型架構能應用於兩種不同語言的資料。
- 本研究的方法也能應用於多種情境，包含情緒立場及網路輿論的立場判斷。

## 第二章 文獻探討

### 第一節 資料增強

#### 第一項 基礎資料增強

在擁有大量標記資料下的問題，已有許多研究，且已有相當好的成效。針對大量未標記資料的情況下，在 [30] 中介紹各種應用在文字的資料增強技術，希望可以藉由資料增強技術解決已標記資料量過少的問題。本篇提出四種資料增強的方式，包括：同義詞替換、同義詞插入、任意交換文字順序、以及任意刪除文字。

- 同義詞替換

由句子中任意選  $n$  個詞，再把這些選中的詞與其同義詞做替換。

- 同義詞置入

從句子中任意選擇一個詞，並把選中的詞的同義詞任意插入句子中的任意位置，上述作法重複  $n$  次。

- 資料亂序

任意從句子中選擇兩個詞任意交換，並重複此舉  $n$  次。

- 任意刪除

任意刪除句子中的某些詞。

本篇顯示原本的資料經過資料增強過後，再與原資料一起訓練的成果，較沒有使用資料增強方式的效能高。

#### 第二項 非監督式資料增強學習

於 [32] 提出一種能同時學習已標記資料以及未標記資料的模型架構。針對已標記資料進行監督式訓練，針對未標記資料則進行資料增強轉換，再將未標記資料與資料增強後的未標記資料進行一致性訓練，希望模型能從中學習對同一意思，不同講法的句子能有相同的預測值，即模型是否能理解資料換句話說的能力。此篇也提出不同的資料增強技術，分別為反向翻譯以及 TF-IDF 替換：

- 反向翻譯

將資料由源語言翻成目標語言，再由目標語言翻回源語言，藉此獲得不同講法但同個意思的新資料。

- TF-IDF 替換

計算每筆資料中每個詞彙的 TF-IDF 分數，TF-IDF 分數較高的詞彙表示權重較高，較為重要，反之則較低，較不重要。將 TF-IDF 低的辭彙取代為另一個 TF-IDF 低的辭彙，此舉希望能獲得不同講法卻又不會影響整句意思 (因為替換皆為 TF-IDF 較低，較不重要的詞彙) 的新資料。

實驗結果顯示藉由少量的已標記資料進行監督式訓練以及利用未標記資料進行一致性訓練的結果，能達到非常好的效果，在某些任務上甚至超越了只使用已標記資料進行監督式學習的效能。

### 第三項 階層式資料增強

在 [33] 中提出了階層式資料增強的模型架構，模型利用單詞層級的注意力機制以及句子的注意力機制來產生新的資料增強後的訓練資料。單詞以及句子層級的注意力機制會為原本資料中標示出哪些文字資訊是重要的，同時刪除較不重要的資訊，藉由上述方式來得到一個新的訓練資料，解決了原本訓練資料不足的問題。實驗結果顯示，此篇提出之模型架構效能更勝之前提到 [30] 中的基礎資料增強方法，且若階層式資料增強的模型架構與基礎資料增強方法 [30] 結合，能夠達到更高的模型效能。



## 第二節 跨領域轉移問題

上述三個資料增強方法在有些任務上能有效解決擁有少量已標記資料和大量未標記資料的問題，但於本篇希望能應用於跨領域的問題，在源領域及目標領域中都僅有少量的已標記資料及大量未標記資料的情況。

跨領域的轉移學習問題，希望可以訓練某一領域的資料，並將訓練後的模型用來預測目標領域的資料，而跨領域轉移學習的關鍵是如何從源領域中找到適合目標領域的特徵並轉移。跨領域問題在 [17] 中提到解決跨領域問題有三個要素：

- 在源領域上訓練效果的好壞，會影響到目標領域資料的測試效果。
- 在源領域以及目標領域的標記函數差距，此差距應愈小愈好以利於轉移特徵。
- 測量源領域分佈以及目標領域分佈之間距離的方式，若兩個領域分佈間的距離愈小，表示訓練在源領域的模型在目標領域的測試效能愈好。

於 [25] 歸納跨領域問題的方法，主要可以分為三類，在底下一一列出：

- 模型中心法

- 特徵中心法

提取源領域以及目標領域的共同特徵，即這些共同特徵在源領域以及目標領域的分佈表示是相似的，這些特徵不會因領域改變而有很大的差異，而這些特徵非常適合拿來做跨領域的轉移。

- 損失中心法

使用對抗式網絡產生一個不會被特定領域特徵汙染的共同特徵空間並用來預測目標領域，如建立一個分類器學習資料是來自哪個領域，再利用 Gradient reversal layer 給予負的梯度來混淆此分類器，藉此取得不受領域影響的特徵。

- 資料中心法

- 偽標記法

會在每個源領域訓練一個模型，再對目標領域的資料進行預測，作法如：

Co-training、Tri-training。

- 資料篩選法

從源領域資料中選擇最適合目標領域的資料以轉移特徵，用來預測目標領域的資料。

- 預訓練模型法

利用事先預訓練過的語言模型，如：AdaptaBERT、DAPT 等來解決跨領域問題。

- 混合法及其他方法

混合法便是結合模型中心法以及資料中心法來解決跨領域問題。

### 第一項 模型特徵中心法

- Pivot法

在 [3] 中，先找出在源領域以及目標領域中皆有出現，且高頻率的辭彙，這些辭彙稱之為 Pivot。這些選出的 Pivot 在源領域及目標領域中的作用是極為相似的。文章中為每個 Pivot 分別建立一個分類器，分類器的目標為判斷資料中是否出現過此 Pivot，若某資料的預測值為正，表示此筆資料的特徵與此 Pivot 擁有高的相關性。由每個 Pivot 分類器建立由非 Pivot 特徵映射到 Pivot 空間的投影，最終在此投影空間上訓練預測模型。在 [24] 中，也先找出 Pivot 特徵，即在源領域及目標領域中出現頻率極高的特徵。並利用 Bag-of-words 的方式，找出領域特有的特徵以及領域私有的特徵，再藉由 Bipartite Graph 連接領域私有特徵以及領域共有特徵兩個群集，若兩群集的特徵有共同出現在同篇資料中，則為此領域共有特徵以及此領域私有特徵進行連結。藉由此 Bipartite Graph 的建立，可以減少領域私有詞及領域共有詞間的差距，並利用這些特徵來訓練目標領域的分類器。

- Autoencoder法

在 [6] 使用邊緣化去噪音堆疊的 Autoencoder [29] 來解決跨領域問題。邊緣化去噪音堆疊 Autoencoder 利用線性去噪器來邊緣化噪音的方式，解決以往去噪音堆疊 Autoencoder 需要花費大量成本訓練以及無法擴展至高維度特性的問題。文中應用邊緣化去噪音堆疊的 Autoencoder 在源領域以及目標領域中學習特徵以訓練模型。

## 第二項模型損失法

[14] 中建立一個特徵提取器提取特徵供模型使用，特徵會應用在兩處，分別是標記預測器以及領域判斷器。標記預測器用來判斷資料的標記，而領域判斷器則用來判斷資料來自哪個領域。其中在特徵提取器以及領域判斷器中會加入一層梯度反轉層，此梯度反轉層作用為讓梯度乘上負一變成負數，使模型進行反向梯度更新時，混淆領域判斷器，最終使源領域特徵分佈及目標領域的特徵分佈相似，創造出共享特徵。而 [34] 解決多源頭的跨領域問題，也應用到類似的概念，在文中所使用的模型架構也使用到梯度反轉層來混淆領域判斷器，獲得共同特徵。

[10] 中提出一個 Weighting Scheme based Unsupervised Domain Adaptation 的架構。此架構分為四個部分，分別是私有特徵的 Extractor、公有特徵的 Extractor、分類器以及領域判別器。

此篇應用到對抗式網絡的概念，利用領域判別器對輸入資料預測來源為哪個領域，藉此讓私有特徵與公有特徵的區分更顯著，並以此領域判別器的損失作為對抗，以期公有特徵的 Extractor 獲得不被私有領域影響的特徵。而判別器也能讓分別在源領域訓練的模型獲得一個相對於目標資料的注意力或權重，讓模型能在預測目標領域資料時能利用此權重選擇適合轉移的領域。而分類器用來預測資料的分類，在此篇為資料的情緒分類。私有特徵的 Extractor 的作用是保留與各自領域相關的特徵，並把這些與各自領域相關的特徵給予領域判別器作資料來源為何領域的預測。對於未標記資料處理，是由源領域所訓練的模型來對其進行偽標記，以此賦與這些未標記資料一個標籤。在偽標記的處理上，文章中希望能獲得預測後有相當信心水準的資料，但又不希望數量太多而影響模型效能，因此提出了選擇相當信心水準的資料的條件。(1) 模型對未標記資料預測出來的機率必須大於某個臨界值。(2) 在訓練過程中會逐步降低此臨界值。(3) 在訓練過程中，產生出的偽標記資料數量小於某個自定數量或是臨界值，便會停止訓練。藉由對抗式生成網絡的概念，賦與每個源領域相對於目標領域資料的權重，以此進行預測，並搭配著設定臨界值條件的偽標記方法來獲得偽標記資料，最終得到損失並更新模型，此概念即為此篇的架構。

### 第三項 資料中心法

- 偽標記法

於 [27] 中，以 Tri-training 的概念為基礎，建立跨領域的模型。Tri-training 可以應用於訓練含有已標記資料以及未標記資料中，將已標記資料分為三份，並利用這三份資料分別獨立訓練三個模型。再利用訓練完的模型來對未標記資料進行預測，若兩個模型對未標記資料的預測值皆相同，表示此筆資料具有高的訓練價值，便將此資料連同此資料的預測值放入第三個模型進行訓練。整個訓練過程在沒有任何模型進行訓練的時候停止。

而此篇認為若用 Tri-training 的方式來訓練跨領域模型，將會花費大量的成本，因為必須訓練三個模型。文章中提出改進的方式為將訓練三個模型改成在 LSTM (長短期記憶模型) [19] 後外接三個輸出層，將 Tri-training 的概念由三個模型改為應用在三個輸出層，若兩個輸出層對未標記資料的預測值相同，便將此資料及預測值加入第三個輸出層進行訓練，如此一來可以減少大量的成本。此作法相當於在利用源領域資料訓練模型，並利用對目標領域未標記資料的預測值作為是否加入第三個模型的依據。

- 資料篩選法

於 [17] 面對多個源領域的問題，則建立一個能夠動態挑選源領域的選擇器。文章中也認為測量源領域資料以及目標領域資料之間距離的方式十分重要，因為在跨領域轉移學習的問題中，如何挑選一個與目標領域最相似、最適合用來轉移的源領域資料是跨領域轉移學習的關鍵，因此如何挑選一個最適合模型的測量方式是不可忽略的重點。

在文中方法則將源領域資料與目標領域資料之間的距離加入預測標記的損失函數中，一同更新模型。

挑選源領域選擇器的設計，則是任意挑選一個源領域供模型訓練，由模型訓練中所得到的損失作為此領域選擇器的回饋，選擇器會根據此回饋及環境因素做出動作，即挑選源領域供目標領域資料使用，此作法與強化學習 (Reinforcement Learning) 的概念十分相似，訓練模型基於環境以及回饋來做出動作，做完動作後得到回饋，此回饋再供下一個動作使用。

#### 第四項 混合及其他方法

在 [9] 中面對目標領域只有未標記資料的情況下，文章提出了 Projection-based 和 Self-training 的方式來解決未標記資料的跨領域問題。

Projection-based 作法是在源領域以及目標領域分別訓練一個投影矩陣，使在向量空間中，兩個不同已標記資料，如：正評、負評資料間的距離能夠分得越開越好，以利模型學習，較不會出現誤判的情況。

Self-training 即先利用源領域已標記資料訓練模型，再利用此訓練好的模型預測目標領域中的未標記資料，給予未標記資料一個偽標記。訓練過程中，若模型對未標記資料的預測信心程度大於某個自訂的臨界值，便將此偽標記資料加入目標領域的訓練資料中訓練。

目標領域的模型會利用 Projection-based 的方式，利用 Self-training 所產生偽標記資料，訓練一個投影矩陣來分開向量空間中，不同標記資料的距離，最終為資料做出預測，並與偽標記計算損失以更新模型。

在 [18] 中利用計算每個目標領域資料與各源領域之間的距離，即計算點到集合的距離，以此得到各源領域對目標領域資料的權重。同時，進行針對已標記資料的監督式學習，並利用 Maximum mean discrepancy [16] 的方式計算領域間的距離，期待能減少源領域以及目標領域之間的距離。

[1] 中，提出四種解決跨領域問題的方法，第一種為將所有源領域資料全部用來訓練一個分類器。第二種方法為第一種方法的延伸，即把訓練特徵縮減到那些只會在目標領域中出現的特徵。第三種方法為分別在多個源領域各自訓練一個分類器，再利用這些分類器為目標領域資料進行預測與多數決投票或是有加權的多數決投票。最後一種方法為，針對目標領域中的擁有部分已標記資料以及大量未標記資料的情況下，先利用已標記資料訓練模型，再利用訓練好的模型為未標記資料預測標記，最後將原本已標記資料及模型預測後的偽標記資料結合訓練模型，但未應用到源領域資訊。

於 [21] 中使用 memory network [28] 的方式來解決跨領域問題。作者認為使用 LSTM、GRU [7] 等記憶模型的記憶機制容量仍無法有效解決長期記憶表現不好的問題，因此使用 memory network 的方式來儲存記憶，即每筆資料都用一個記憶體儲存。整體模型架構可以分為兩個部分，第一部分使用 memory network 的方式訓練用來預測資料標記的模型，另一部分則同樣使用 memory network 的方式來訓練用來判斷資料來源的領域判別器，特別的是在領域判別器的架構中也導入了對抗

式學習的概念，使用了梯度反轉層混淆領域判別器，由此得到供所有領域使用的共同特徵。

[31] 中建立一個情緒用詞的圖形，圖形的頂點代表情緒用詞，而圖形的邊則表示情緒立場的關係，此關係則是利用出現在文中的頻率所計算的，如在同一領域中”好”跟”快”皆在形容某一產品，而在另外領域中”好”為一個正向的形容詞，因此能推論出”快”也是一個正向的形容詞。文中的另一重點則是計算領域之間的相似度，相似度的計算藉由皆出現在不同領域的辭彙，利用 JS divergence [23]，計算此辭彙在不同領域中分佈的相似度。最後結合情緒用詞圖形以及領域間的相似度來解決跨領域問題。

在 [22] 中提出階層式的注意力轉移網路，此模型架構由 P-net 和 NP-net 所組成。P-net 的功能用來萃取領域之間公有的特徵，而 NP-net 的功能為萃取領域的獨有特徵。在跨領域轉移問題中，階層式的注意力轉移網路可以由源領域轉移對單詞和句子級別情感上的注意力至目標領域資料，藉由此注意力找到最適合目標領域的源領域特徵做轉移。

在 [8]，此篇旨在解決多對一的跨問題，非一對一的跨領域問題，即擁有多個源領域及一個目標領域，且在目標領域中僅有未標記資料。因此該選擇哪個源領域的資料做轉移十分重要。

文中的方法先對每個源領域以 Self-training 的方式訓練模型，先分別利用源領域已標記資料訓練源領域模型，再將這些源領域模型用來預測源領域中的未標記資料，即給這些未標記資料一個偽標記，最後再將原有的已標記資料與偽標記資料結合重新訓練源領域模型，每個源領域所訓練的模型皆為獨立且互不干擾的。對目標領域的未標記資料則利用原先在源領域訓練好的模型進行偽標記並以多數決的方式選擇最終答案，如三個源領域模型的預測值分別為正評、正評及負評，則正評為最終多數決結果，此偽標記則視為資料的真實答案。

文中對於目標領域所建的模型是利用 Relatedness map learning 以及 Instance-based Domain-Attention 的方式來預測目標領域的未標記資料。訓練 Relatedness map learning 希望能找出每筆源領域的資料對於預測目標領域資料的貢獻，找出適合目標領域資料的源領域資料。訓練 Instance-based Domain-Attention 希望可以依據不同的目標領域資料調整對每個源領域的注意力或權重，根據這些源領域權重，模型可以知道如何選擇適合目標領域的源領域做轉移。

## 第三章 研究方法

### 第一節 整體模型架構

本文所提出之模型架構可以包括三個部分，第一部分為使用監督式學習訓練目標領域內已標記資料。第二部分為師生跨領域模型，利用非監督式資料增強學習來分別為源領域內資料做訓練，各源領域所訓練的模型構成教師模型，為目標領域的未標記資料做偽標記，學生模型利用資料級別的注意力機制以及領域級別的注意力機制來為這些偽標記資料做跨領域學習。第三部分為相關度偵測任務，此任務希望藉由偵測文句與描述目標主體的相關程度來與主任務一起訓練，藉由多任務學習來提升整體模型的效能。整體之多任務模型架構如圖7所示。

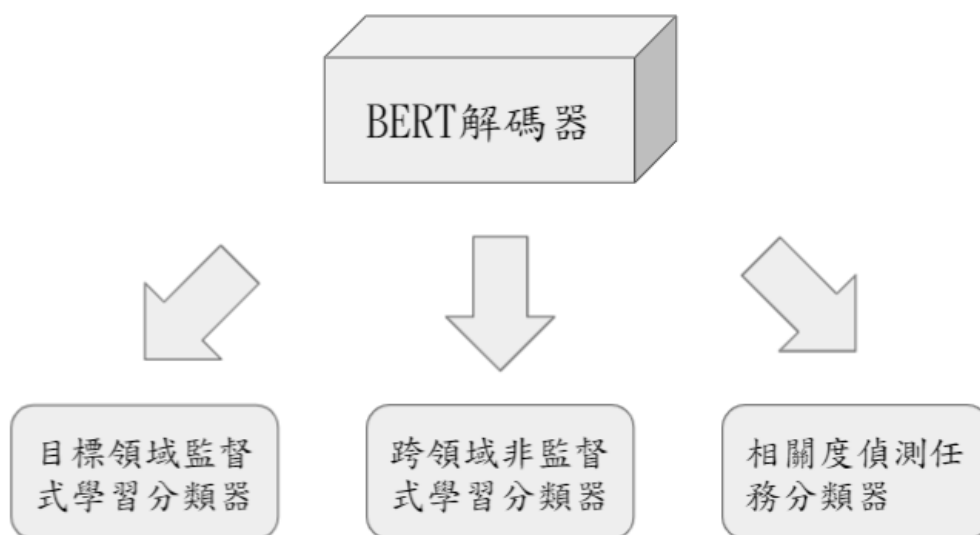


圖 7 模型多任務學習架構圖

## 第二節 師生跨領域訓練模型架構

本研究的跨領域訓練模型架構包括兩個階段。首先，準備  $k$  個源領域模型，並使用來自  $k$  個源領域的已標記資料進行訓練。為了提高這些源領域模型的泛化能力，我們引入了來自  $k$  個源領域的其他未標記資料，作為非監督式資料增強學習的材料 [32]，這些源領域的模型構成教師模型，提供知識讓學生模型學習。其次，我們訓練學生模型，該模型將成為用來預測目標領域資料的最終模型。

與傳統的集成模型預測方法不同， $k$  個源領域模型不會直接應用於最終預測。這  $k$  個領域源模型會結合成教師模型用於為目標領域內的未標記資料生成偽標記供學生模型學習，我們的學生模型將會利用該偽標記進行訓練。而學生模型進一步引入了多層次的注意力機制來進行跨領域間的知識轉移。本研究的師生跨領域最終模型將以一種複雜的方式從源領域資料和目標領域資料中獲取知識，從而減少了領域及領域之間的歧異。以下將詳細介紹上述的兩階段模型架構：

### 第一項 教師模型

資料增強是通過增加訓練資料的多樣性來提高模型泛化能力。在自然語言處理中，關於資料增強的概念是為每筆已標記資料生成大量相似語意的新資料，並且可以將標籤直接轉移到此筆新資料。對於原本已具有標籤  $y$  的資料  $x$ ，可以將相同的標籤  $y$  分配給由  $x$  生成的新資料  $\hat{x}$ ，因為  $\hat{x}$  在語義上與  $x$  是相等的。藉由這種方式，可以透過學習具有相同含義但不同表達方式資料來訓練模型，使模型更具有泛化能力。多種資料增強方法，例如反向翻譯和基礎資料增強 [30]，已被證明能夠有效增加訓練資料的多樣性。在本研究選擇使用反向翻譯作為資料增強的方式而非使用基礎資料增強 [30]，原因是因為基礎資料增強雖能為原本的資料創造新的相同意思卻不同講法的資料，但相較於反向翻譯，反向翻譯能為原本資料創造更多元且不同變化的新資料，使模型學習到更豐富的”換句話說”能力。本研究的教師模型由使用非監督式資料增強學習各別為源領域資料訓練而得的模型集結而成。與傳統的資料增強的方法不同，非監督式資料增強學習不會將標籤從原資料直接轉移至其資料增強後的資料，而是利用資料增強後的資料進行一致性訓練的任務來共同訓練模型，在該任務中，期待模型能夠對在原本資料與其資料增強後資料進行一致的預測。換句話說，在主任務中，以監督式學習的方式訓練已標記資料，而在輔助任務中用未標記的資料及其經過資料增強後的資料來進行一致性訓練。如此一來便可以使用到大量的未標記資料。在



主任務中使用交叉熵損失來計算損失，在輔助任務中使用 KL-divergence 來計算損失，而最後的損失則是兩個損失搭配權重後的總和。在此訓練中，本篇使用反向翻譯來作為資料增強的方式。圖 8 說明了多個源領域的非監督式資料增強訓練，而圖 9 則以藝人資料為例，目標任務用來判斷回覆對藝人的評價為正面或負面。輔助任務中，針對未標記資料則以反向翻譯作為資料增強的方式，並判斷模型對未經資料增強前的資料以及經資料增強後資料的預測是否一致。本篇採用流行的預訓練編碼器 BERT 作為我們將文字轉為向量的基本模型。在 BERT 之後進一步添加了其他分類器以預測標籤。監督式訓練和一致性訓練中使用的編碼器和分類器是共享的。在這裡，我們將比較使用 BERT 做監督式學習和 BERT 做非監督式資料增強學習的模型效能，然後選擇效能較高的模型作為源領域的最終模型。



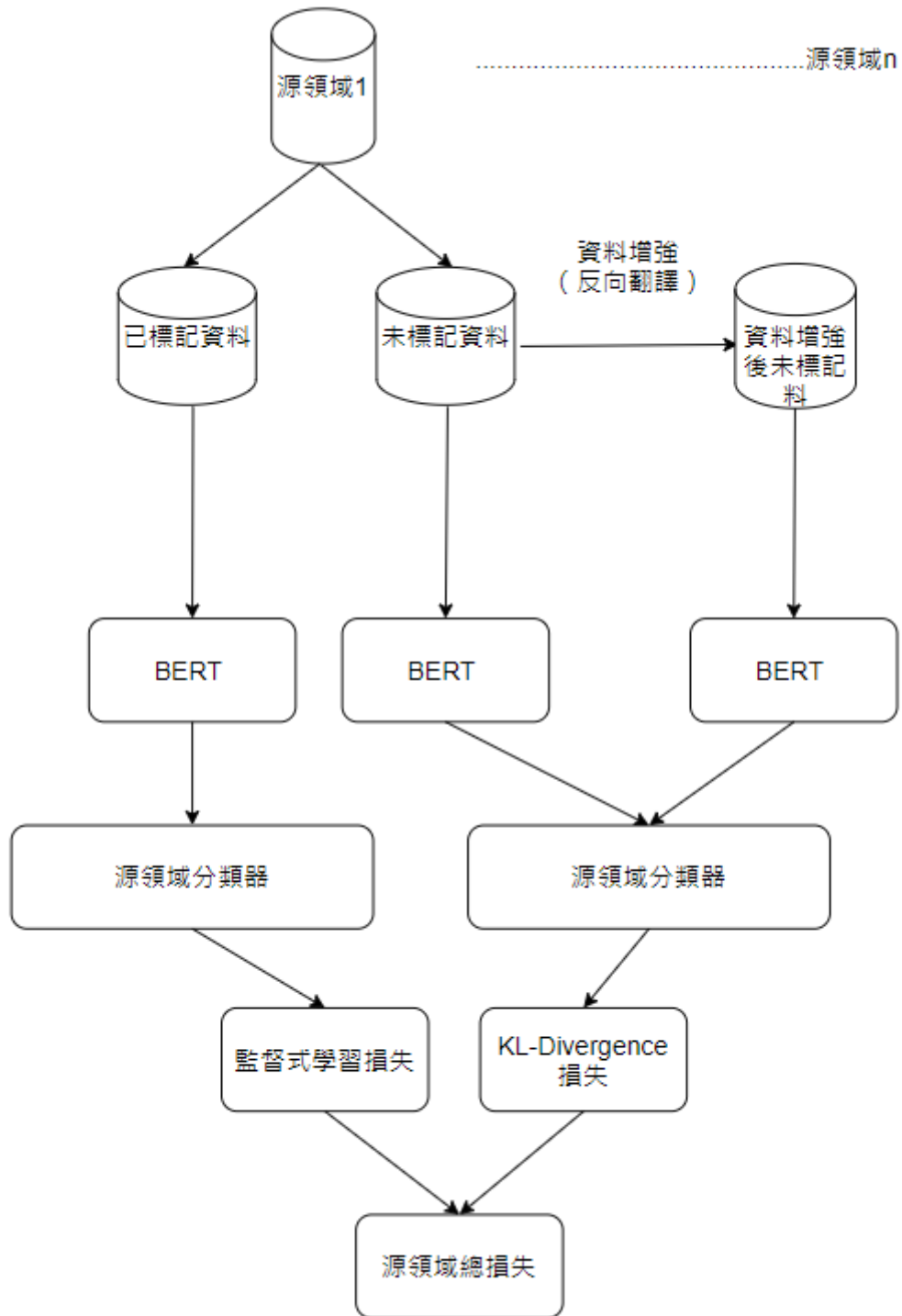


圖 8 源領域模型架構圖

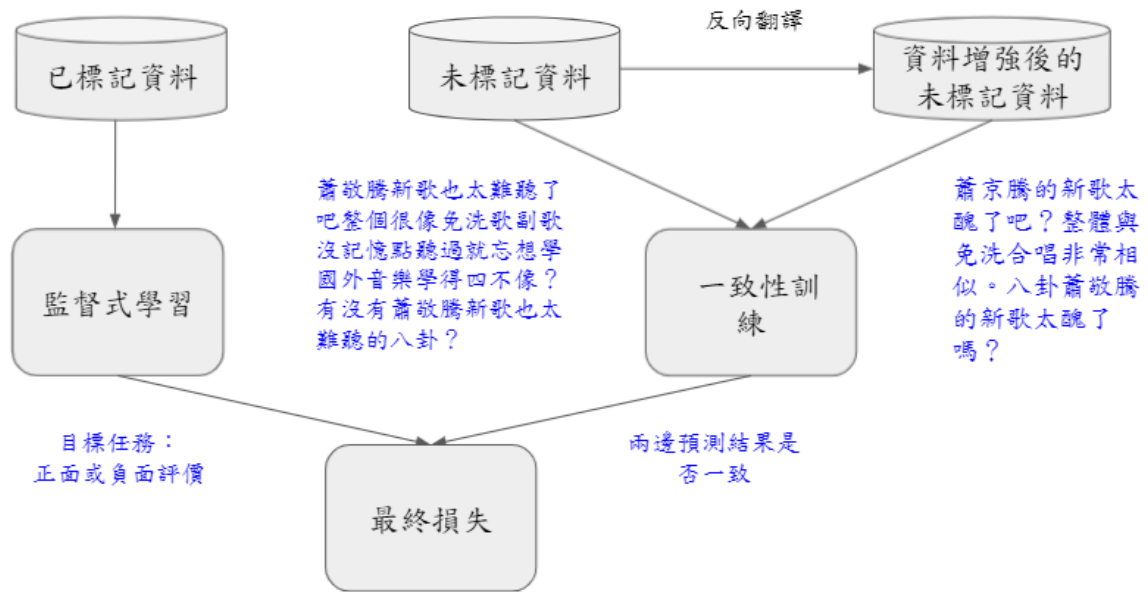


圖 9 非監督式資料增強學習示意圖



## 第二項 學生模型與多層次注意力訓練

不同的源領域模型對目標領域資料的貢獻程度不同。我們將  $k$  個源領域模型結合成教師模型對目標領域內未標記資料進行預測，再使用所有預測值進行多數決投票，舉例來說若三個源領域對目標領域未標記資料進行預測得到三個預測值，分別是負評、負評、正評，則最終多數決投票結果為負評。若四個源領域對目標領域未標記資料進行預測得到四個預測值，分別是負評、負評、正評、正評的平手情況，則最終多數決投票隨機選擇正評或是負評作為結果。而  $k$  個源領域模型產生的  $k$  個預測值構成教師模型，教師模型最終會為目標領域的每筆未標記資料產生一個偽標籤，我們使用此機制組合了來自多個源領域的知識。不同的領域資料對模型的貢獻程度也有所不同，換句話說，由教師模型所提供的知識並非都適合學生模型做學習。基於這些觀察，本研究引入了多層次的注意力機制至學生模型，替學生模型篩選由教師模型所得到的知識並學習。多層次的注意力機制利用了來自資料級別和領域級別的訊息。在跨領域學習部分，資料級別注意力機制參考了 [8] 中的 Relatedness map learning。定義如下式：

$$\psi_i(x_T, x_i^L) = \exp(x_T^T, x_i^L) / \sum_{x' \in S_i^L} \exp(x_T^T, x') \quad (3.1)$$

$x_T$  代表目標領域的未標記資料， $x_i^L$  代表源領域的已標記資料， $S_i^L$  表示源領域已標記資料集。而領域層級注意力機制參考了 [8] 中的 Instance-based Domain-

Attention，定義如下：

$$\theta(x_T, S_i) = \exp(x_T^T, \phi_i) \quad (3.2)$$

$x_T$  代表目標領域的未標記資料， $S_i$  代表源領域， $\phi_i$  代表源領域的向量表達。在此我們使用 Xavier initialization 來訓練源領域的向量，並藉由模型損失來更新源領域向量。

資料層級注意力機制希望可以找出適合目標領域未標記資料的源領域資料特徵做轉移，而領域層級注意力機制希望可以找出適合目標領域未標記資料的源領域做轉移。

最終兩個級別的注意力會結合成為學生模型，用來預測目標領域未標記資料。學生模型定義如下：

$$\hat{y}(x_T) = \sigma\left(\sum_{i=1}^N \sum_{x_i^L \in S_i^L} y(x_i^L) \psi_i(x_T, x_i^L) \theta(x_T, S_i)\right) \quad (3.3)$$

$\hat{y}(x_T)$  代表目標領域未標記資料的預測值， $\sigma$  代表 sigmoid 函數， $N$  代表源領域數量， $y(x_i^L)$  代表源領域已標記資料的標籤， $x_T$  代表目標領域未標記資料， $x_i^L$  代表源領域的已標記資料， $S_i^L$  表示源領域已標記資料集， $S_i$  代表源領域向量。再利用預測值  $\hat{y}(x_T)$  以及偽標籤來計算 entropy loss 更新模型。跨領域模型架構如圖 10。

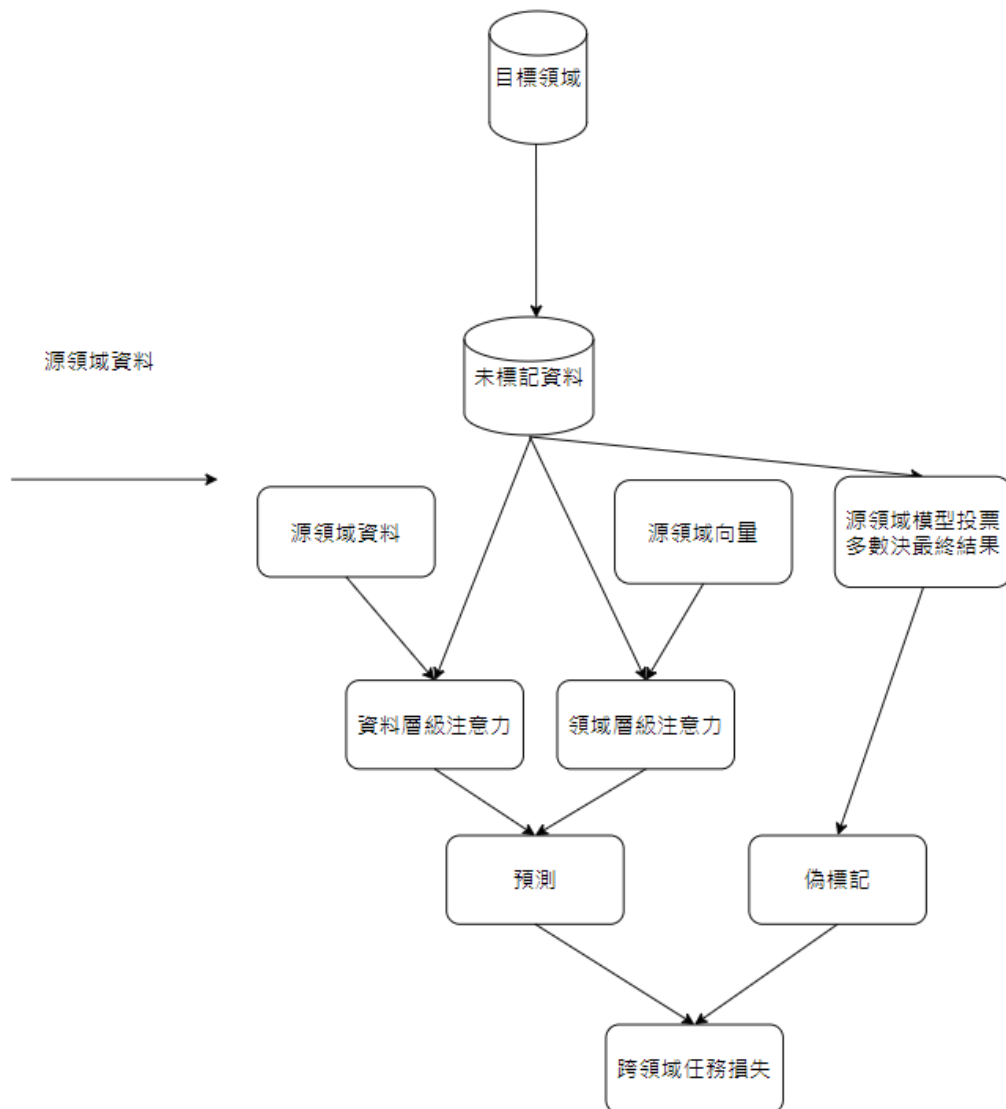


圖 10 跨領域模型架構圖

### 第三節 相關度偵測訓練

在 [26] 中認為多任務學習在某些情況中的效能能夠更勝單純訓練單一任務的效能。為了增加本文所提之跨領域模型的效能，因此我們新增了一個相關度偵測的任務作為輔助任務一同訓練，其目的為偵測相關度。我們希望藉由找出文字以及目標描述主體之間的相關程度來幫助模型訓練。在相關度偵測任務，為了增加相關度訓練任務的負採樣資料，以 Amazon 產品回覆 [2] 為例，我們將原本的顧客回覆搭配與顧客回覆中不相關的目標產品作為負採樣資料。而負採樣資料包含兩種，第一種為顧客回覆配上不一致但仍為同一個領域的目標產品，第二種為顧客回覆配上不一致且不同領域的目標產品。表 1 舉出以記憶卡為正採樣資料，負採樣資料的範例。

訓練資料	文字描述	目標主體
正採樣資料	非常建議這張記憶卡，體積小巧，傳輸速度快	記憶卡
同領域負採樣資料	非常建議這張記憶卡，體積小巧，傳輸速度快	電視
不同領域負採樣資料	非常建議這張記憶卡，體積小巧，傳輸速度快	百科全書

表 1 負採樣資料範例，以記憶卡為正採樣資料

在相關度偵測任務中，本研究提出兩種有效使用負採樣資料的機制，以下將詳細說明此兩種機制。

- 注意力機制

在相關度偵測的任務中利用注意力機制作為篩選負採樣資料的方式。此處使用的注意力機制與上述跨領域轉移學習的注意力機制有所不同。此處計算正採樣資料及負採樣資料之間的 Cosine similarity 來作為注意力，本研究針對正採樣與負採樣資料 768 維向量的每個維度分別計算 Cosine similarity。Cosine similarity 越高者，表示正採樣資料以及負採樣資料之間的相似度越高，越難讓模型區分兩者之間的差異，越適合拿來作為相關度偵測任務的訓練資料，因此給予此筆負採樣資料越高的注意力。注意力機制的定義如式 3.4 及式 3.5 所示：

$$Attention(x_P, x_N) = \frac{(x_P \cdot x_N)}{|x_P||x_N|} \quad (3.4)$$

$$\hat{x}_N = Attention(x_P, x_N)x_N \quad (3.5)$$

$x_P$  代表正採樣資料經過 BERT 轉換後的向量， $x_N$  代表負採樣資料經過 BERT 轉換後的向量， $\hat{x}_N$  代表經過 Attention 轉換後的負採樣資料向量。

- 半衰期機制

在訓練相關度偵測任務的過程中，我們使用半衰期機制來控制上述兩種負採樣資料數量間的比例，半衰期機制的定義如下：

$$\lambda \times |X_{in}| + (1 - \lambda) \times |X_{out}| \quad (3.6)$$

$|X_{in}|$  代表同領域負採樣資料的數量， $|X_{out}|$  代表不同領域負採樣資料的數量，而  $\lambda$  代表控制同領域及不同領域負採樣資料數量比例的參數，定義如下：

$$\lambda = 0.5^{\frac{t}{k}} \quad (3.7)$$

$t$  代表模型訓練的疊代數， $k$  代表半衰期。在剛開始訓練時，我們將  $\lambda$  設為 1， $\lambda$  會隨著式 3.7 降低。換句話說一剛開始所有負採樣資料皆為同領域的負採樣資料，隨著訓練疊代數逐漸增加，同領域負採樣資料會逐漸減少，而不同領域的負採樣資料會逐漸增加。

## 第四節 目標領域模型

根據目標領域資料標記有無可以分為兩種訓練模型，若目標領域內僅有未標記資料則使用本研究之師生跨領域模型，若目標領域含有已標記資料及未標記資料，針對目標領域中的已標記資料，本研究利用傳統的監督式學習來訓練模型，即賦予模型資料與標籤進行訓練，而針對目標領域中的未標記資料則採用師生跨領域模型來進行非監督式的轉移學習，學生模型會與監督式學習一同訓練。最終監督式學習模型會與師生跨領域模型以及相關度偵測任務模型進行多任務訓練，三者的損失會結合，並一同更新三者的模型。在此，我們使用 Adam optimizer 來優化梯度。在監督式學習模型、跨領域模型及相關度偵測模型中，我們使用 BERT 作為解碼器，使文字換為向量，此解碼器會在三個模型之間共享，整體模型架構及示意圖呈現於圖 11 以及圖 12。整體模型的損失定義如式 3.8。

$$Loss = sup + weight 1 * transfer + weight 2 * relevant \quad (3.8)$$

Loss 代表整體模型的損失，sup 代表監督式學習損失，transfer 代表師生跨領域模型的損失，relevant 代表相關度偵測任務的損失。其中 weight1 用來控制跨領域模型損失的權重，weight2 用來控制相關度偵測任務損失的權重。最終整體模型損失會用來更新監督式學習任務模型、跨領域任務模型以及相關度偵測任務模型。



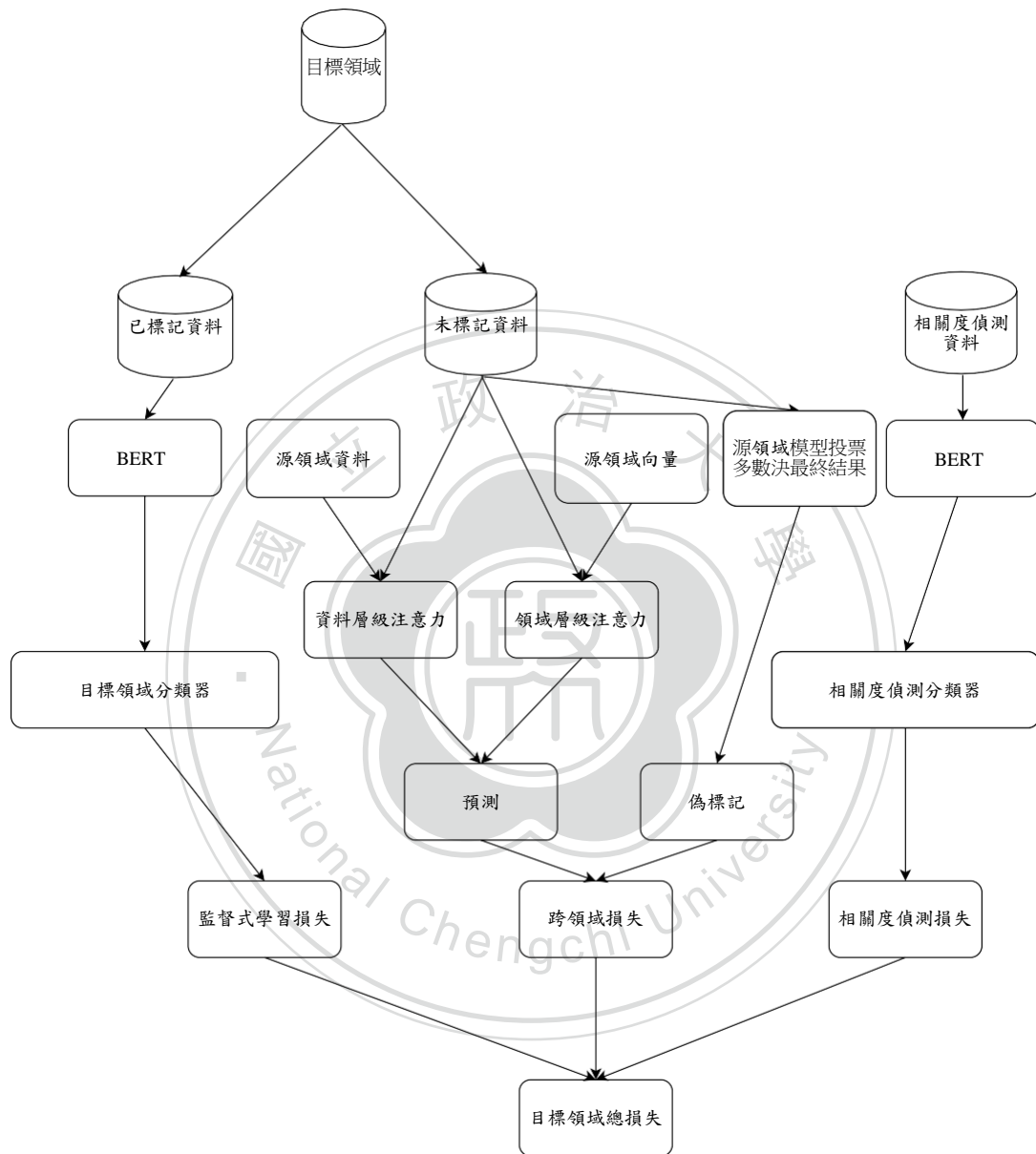


圖 11 整體模型架構圖

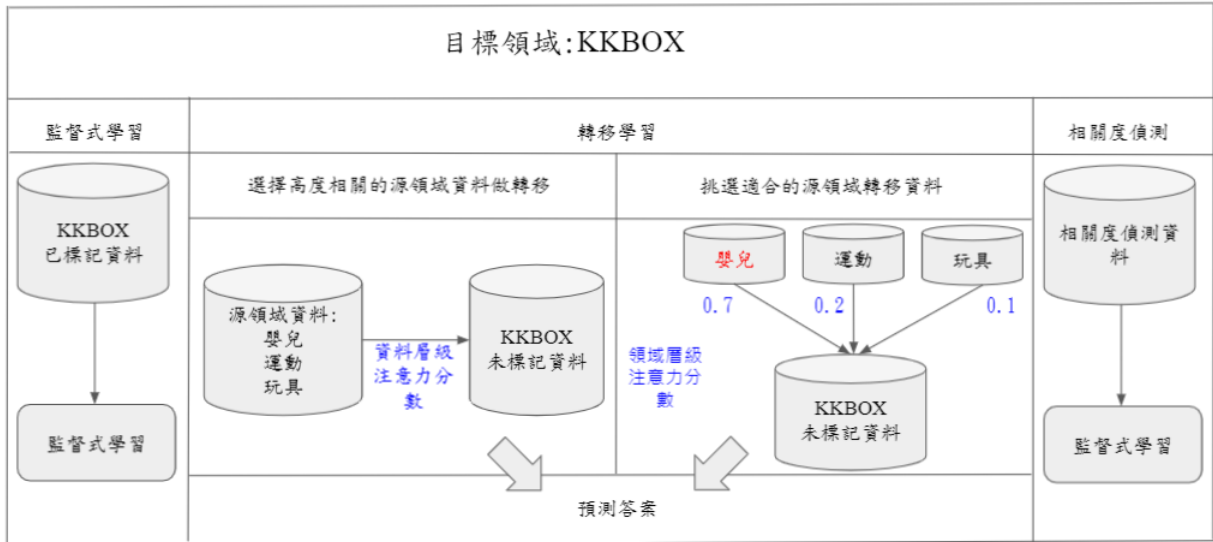


圖 12 整體模型示意圖

### 第五節 句子對分類任務

在上述的監督式學習、跨領域學習及相關度偵測任務中，本研究視為句子對的分類任務，即一筆資料含有一段文字及這段文字所描述的主體。在監督式學習以及跨領域學習中，使用文字搭配描述主體作為訓練資料，正評、負評為答案。在相關度偵測的任務中一樣使用文字搭配描述主體作為訓練資料，有關、無關則為答案。

## 第四章 資料集、實驗配置與實驗結果

### 第一節 資料集

- 英文資料集

英文資料的實驗是參考 [2] 使用 Amazon 公司的商品客戶回覆資料，共有四種領域商品的資料集包括：書籍、DVD、電器用品、廚房用具。每筆資料包含商品名稱、消費者對此商品的意見回饋，以及對商品的滿意程度。滿意程度由 1 到 5 分，分數愈高則愈滿意，此實驗參考 [2] 的配置，若分數大於或等於四分，便標記為正評 (positive)，若分數小於或是等於兩分，則標記為負評 (negative)，資料若為三分則省略。四個領域中已標記以及未標記資料的數量如表 2 所呈現。

領域	正評	負評	未標記
Books	1000	1000	6000
DVD	1000	1000	3,4741
Electronics	1000	1000	1,3153
Kitchen appliances	1000	1000	1,6785

表 2 英文實驗四個領域中已標記及未標記資料的數量

- 中文資料集

除了 Amazon 英文資料集之外，本篇也使用中文資料做實驗，包括 Amazon 公司的商品客戶回覆資料、KKBOX 公司的藝人評論資料以及核能輿論資料。與英文資料集不同的是中文資料集中擁有三個標籤，分別是正評、中立以及負評，分類的難度更高。Amazon 中文資料包含三個領域，分別為嬰兒用品、運動用品、及玩具。表 3 顯示此資料集中三個領域、KKBOX 藝人評論資料以及核能輿論的資料筆數。

領域	正評	中立	負評	未標記
嬰兒用品	3565	474	1207	0
運動用品	1945	212	192	0
玩具	4191	474	630	0
KKBOX	409	560	92	5,3617
核能輿論	1662	1649	1518	1,6772

表 3 中文實驗五個領域中已標記及未標記資料的數量



## 第二節 實驗配置

- 英文資料集

我們利用 Amazon 產品的顧客回覆英文資料集 [2] 作為模型驗證的資料集，此資料集也是在跨領域問題的相關論文中常使用的資料集。為了模擬多來源跨領域轉移問題的實驗配置，我們參考了 [8]。在實驗中，我們會將資料集中的一個領域作為目標領域，而剩下的領域則作為源領域，每個領域都會輪流作為源領域以及目標領域。在每個領域中，我們將 2000 筆已標記資料中的 1600 筆作為訓練集，剩下的 400 筆作為測試集。1600 筆的訓練集中含有 800 筆正評資料以及 800 筆負評資料，而 400 筆的測試集中含有 200 筆的正評資料以及 200 筆的負評資料，此實驗配置能確保所有標籤數量是平衡的。每個領域中也提供充足數量的未標記資料。

在本文中，我們利用上述 Amazon 顧客回覆資料進行以下實驗，驗證本文提出的模型架構，實驗分別為：

- 監督式領域內訓練

我們使用 BERT 模型分別對四個領域的已標記資料做訓練。

- 半監督式領域內訓練

本實驗參考了 Google 所提出的 Unsupervised data augmentation [32]，用此方法分別對四個領域的領域內已標記資料和未標記資料做實驗。

- 監督式跨領域訓練

使用 BERT 模型做跨領域轉移問題的實驗，將其中一個領域作為目標領域，剩餘的領域作為源領域。此實驗使用所有源領域資料訓練一個模型，並應用於目標領域，使用的資料皆為已標記資料。

- 非監督式跨領域訓練 (未導入相關度偵測任務)

在非監督式跨領域訓練中，將使用本研究提出之師生跨領域模型做實驗。利用源領域中的已標記資料以及未標記資料做訓練，特別的是目標領域中只使用未標記資料作訓練。如同上述實驗配置，其中一個領域作為目標領域，剩餘的領域作為源領域，每個領域皆會輪流作為目標領

域以及源領域。

– 半監督式跨領域訓練 (未導入相關度偵測任務)

本研究所提出之師生跨領域模型並搭配僅訓練目標領域已標記資料的監督式學習的多任務模型做此實驗。利用所有領域中的已標記資料以及未標記資料做跨領域轉移學習，其中一個領域作為目標領域，剩餘的領域作為源領域，每個領域皆會輪流作為目標領域以及源領域。

– 半監督式跨領域訓練 (導入相關度偵測任務)

最後，此實驗應用本研究所提出之師生跨領域模型並搭配僅訓練目標領域已標記資料的監督式學習，並導入相關度偵測任務的整體模型，整體模型利用了每個領域中所有的已標記資料以及未標記資料。如同上述實驗配置，其中一個領域作為目標領域，剩餘的領域作為源領域，每個領域皆有機會作為目標領域以及源領域。在導入相關度偵測的訓練中，除了參照式 3.6 的方式調整相關度偵測任務的訓練資料，另外也將式 3.6 中的  $\lambda$  乘在不同領域的負採樣資料上但仍設為 1，即一開始所有負採樣資料皆為不同領域的負採樣資料，隨著訓練疊代數逐漸增加，不同領域負採樣資料會逐漸減少，而同領域的負採樣資料數量會逐漸增加。

• 中文資料集

在中文資料集的實驗中，使用 Amazon 公司三個領域產品意見作為源領域，KKBOX 公司的藝人評論以及核能輿論分別作為目標領域進行跨領域轉移學習。針對中文資料也作了以下的實驗：

– 監督式領域內訓練

使用 BERT 模型分別對 KKBOX 藝人評論領域以及核能輿論領域的已標記資料做訓練。

– 半監督式領域內訓練

本實驗參考了 Google 所提出的 Unsupervised data augmentation [32]，用此方法分別對 KKBOX 藝人評論領域以及核能輿論領域的已標記資料及未標記資料做訓練。

– 監督式跨領域訓練

使用 BERT 模型做跨領域轉移問題的實驗，將其中一個領域作為目標領域，剩餘的領域作為源領域。此實驗使用所有源領域資料訓練一個模型，並應用於目標領域，使用的資料皆為已標記資料。實驗利用 Amazon 公司三個產品領域資料訓練一個模型並應用於 KKBOX 藝人評論領域及核能輿論領域。

– 非監督式跨領域訓練 (未導入相關度偵測任務)

此實驗利用本研究之師生跨領域模型做實驗，模型利用所有源領域中的已標記資料以及目標領域的未標記資料做跨領域轉移學習，其中 KKBOX 藝人評論及核能輿論領域分別作為目標領域，Amazon 中文資料的三個領域作為源領域。

– 半監督式跨領域訓練 (未導入相關度偵測任務)

此實驗利用本研究所提出之師生跨領域模型並搭配僅訓練目標領域已標記資料的監督式學習的多任務模型做訓練。模型利用所有源領域中的已標記資料以及目標領域的已標記資料及未標記資料做跨領域轉移學習，其中 KKBOX 藝人評論及核能輿論領域分別作為目標領域，Amazon 中文資料的三個領域作為源領域。

– 半監督式跨領域訓練 (導入相關度偵測任務)

最後，和英文的實驗相同，此實驗利用本研究所提出之師生跨領域模型並搭配僅訓練目標領域已標記資料的監督式學習，並導入相關度偵測任務的整體模型做訓練。所有源領域中的已標記資料以及目標領域的已標記資料及未標記資料皆會使用。如同上述實驗配置，KKBOX 藝人評論領域及核能輿論領域分別作為目標領域，Amazon 中文資料的三個領域作為源領域。在導入相關度偵測的訓練中，則使用了式 3.4 以及式 3.5 的注意力機制篩選適合的負採樣資料作訓練。

### 第三節 模型效能

針對上個章節的英文及中文實驗配置，於本章將逐一列出實驗結果與實驗分析。

- 英文實驗結果

- 監督式領域內訓練

我們使用 BERT 模型分別對四個領域的已標記資料做訓練的結果呈現於表 4。結果顯示模型在四個領域的效能不論是 Accuracy 或是 Macro-F1 皆能達到 89.50% 以上，代表本篇使用 BERT 作為基礎模型是十分有效且合適的。

領域	Accuracy	Macro-F1
Books	91.25%	91.25%
DVD	89.50%	89.50%
Electronics	94.25%	94.25%
Kitchen appliances	93.75%	93.75%

表 4 使用 BERT 模型作為監督式領域內訓練的基礎模型效能

- 半監督式領域內訓練

使用 Google 所提出的 Unsupervised data augmentation [32] 方法的實驗結果如表 5。實驗結果顯示使用 Unsupervised data augmentation 方法，一同訓練已標記資料和未標記資料的模型效能能夠超越僅用已標記資料所訓練出的模型效能(表 4)。

領域	Accuracy	Macro-F1
Books	91.50%	91.49%
DVD	90.25%	90.25%
Electronics	93.75%	93.75%
Kitchen appliances	94.25%	94.24%

表 5 UDA (Unsupervised Data Augmentation)的效能



– 監督式跨領域訓練

我們使用 BERT 模型做監督式跨領域轉移問題的實驗結果如表 6 呈現。此實驗結果與表 4 與表 5 比較，可以看出效能較低，此實驗可以驗證在源領域及目標領域的歧異性會降低模型的預測效能。

目標領域	源領域	Accuracy	Macro-F1
Books	DVD, Electronics, Kitchen appliances	90.00%	90.00%
DVD	Books, Electronics, Kitchen appliances	88.25%	88.25%
Electronics	Books, DVD, Kitchen appliances	92.75%	92.74%
Kitchen appliances	Books, DVD, Electronics	93.25%	93.25%

表 6 使用 BERT 作監督式跨領域訓練的效能

– 非監督式跨領域訓練 (未導入相關度偵測任務)

本研究的師生跨領域方法將與在同資料集且同實驗配置上的不同方法做比較，特別的是此實驗為配合其他研究方法之實驗配置，並不會使用”描述產品”的資訊至模型訓練。以下將逐一列出比較方法：

\* uni-MS [1][34][18]

使用所有源領域資料共同訓練一個模型，再使用此模型對目標領域資料進行預測。此方法可以視為最基本的跨領域轉移學習方法。

\* SCL: Structural Correspondence Learning [3][2]

SCL 中先找出 pivots，再找出能將 non-pivots 特徵投影到 pivots 空間的投影，最終模型會用此投影空間進行訓練。實驗結果引用 [31]。

\* SFA: Spectral Feature Alignment [24]

SFA 建立 bipartite graph 連結了領域獨立特徵以及領域相依特徵，並將這些特徵做聚合，最終利用原本的特徵以及聚合後的特徵訓練模型。實驗結果引用 [31]。

\* SST: Sensitive Sentiment Thesaurus [5][4]

SST 方法利用所有源領域特徵建立一個情緒辭庫，並用此詞庫來訓練跨領域情緒分類問題。實驗結果引用[5]。

\* SDAMS: Sentiment Domain Adaptation with Multiple Sources [31]

SDAMS 建立情緒詞彙圖以找出領域之間的相似度來解決跨領域問題。情緒詞彙圖中涵蓋所有情緒詞彙之間的連結，而領域之間的相似度是藉由比較同一詞彙在不同領域的分布之間的相似度而來。

\* AMN: End-to-End Adversarial Memory Network [21]

利用 memory network [28] 的方式訓練兩個任務，第一為情緒分類任務，第二為領域分類任務完成跨領域轉移問題。該模型也能利用注意力機制自動選 Pivots。實驗結果引用 [12]。

\* Attention: Multi-source domain adaptation for unsupervised domain adaptation [8]

利用 Relatedness map learning 和 Instance-based Domain-Attention 挑選適合目標領域資料的源領域資料做轉移。

針對目標領域僅有未標記資料預測的結果呈現於表 7，結果顯示本研究的師生跨領域模型效能較其他所有方法來得高(第 9 欄)。其中較為特別的是，在此實驗中不需使用訓練到最好的源領域模型即可以達到非常好的模型效能，顯示本研究的方法是有效果的。

目標領域	uni-MS	SCL	SFA	SST	SDAMS	AMN	Attention	師生
Books	80.00	74.57	75.98	76.32	78.29	79.75	83.50	87.25
DVD	76.00	76.30	78.48	78.77	79.13	79.83	80.50	86.75
Electronics	74.75	78.93	78.08	83.63	84.18	80.92	80.00	90.25
Kitchen appliances	85.25	82.07	82.10	85.18	86.29	85.00	86.00	90.50

表 7 目標領域僅包含未標記資料的設定下與其他方法作比較，結果以 Accuracy (%) 呈現

– 半監督式跨領域訓練 (未導入相關度偵測任務)

半監督式跨領域訓練結果呈現於表 8。純粹訓練在目標領域未標記資料的效能表 7 與訓練在目標領域已標記資料以及未標記資料效能表 8 做比較，可以發現後者的效果更勝前者，換句話說本研究的師生

跨領域模型不單單只能解決目標領域只有未標記資料的情境，若遇到目標領域中同時含有已標記資料以及未標記資料的情況，本研究的模型依舊能夠有很高的效能。

目標領域	Accuracy	Macro-F1
Books	89.75%	89.74%
DVD	90.50%	90.50%
Electronics	94.50%	94.50%
Kitchen appliances	94.25%	94.25%

表 8 目標領域包含已標記資料及未標記資料的模型效能 (未導入相關度偵測任務)

– 半監督式跨領域訓練 (導入相關度偵測任務)

導入相關度偵測任務至本研究的師生跨領域模型訓練結果呈現於表 9。結果顯示導入相關度偵測的跨領域模型的效能除了 Books 領域外，其餘三個領域皆能夠超越表 5，Google 所提出的 Unsupervised data augmentation [32] 方法。導入相關度偵測任務的跨領域模型在所有實驗組合皆能超越表 8，未導入相關度偵測任務的跨領域模型，證實了本研究的最終模型，師生跨領域模型與相關度偵測任務一同訓練的多任務學習是有效果的，且與上述所有實驗結果比較效果是最高的。

目標領域	Accuracy	Macro-F1
Books	91.00%	91.00%
DVD	91.50%	91.49%
Electronics	94.50%	94.50%
Kitchen appliances	94.50%	94.50%

表 9 加入相關度偵測任務的最終模型效能

• 中文實驗結果

以 Amazon 公司的嬰兒用品、運動用品、玩具三個領域作為源領域，KKBOX 藝人評論及核能輿論分別做為目標領域的三個實驗結果呈現於表 10 及表 11。

– KKBOX 藝人評論作為目標領域

\* KKBOX 藝人評論領域做監督式領域內訓練

在使用 BERT 為 KKBOX 領域做領域內訓練中，由於為三分類問題且訓練資料較少的原因，因此模型效能 (Macro-F1) 僅為 75.57%。

\* KKBOX 藝人評論領域做半監督式領域內訓練

使用 Google 所提出的 Unsupervised data augmentation [32] 方法的實驗結果較直接使用 KKBOX 藝人評論做監督式領域內訓練效能低。

\* KKBOX 藝人評論領域為目標領域做監督式跨領域訓練

使用 BERT 以 Amazon 三個領域產品意見資料訓練模型，以 KKBOX 藝人評論領域作為目標領域的模型效能非常低，顯示出 Amazon 產品意見領域與 KKBOX 藝人評論領域相差非常大，非常不適合轉移。

\* KKBOX 藝人評論領域為目標領域做非監督式跨領域訓練 (未導入相關度偵測任務)

本研究的師生跨領域模型應用在非監督式跨領域訓練的效果並不是很好，歸咎原因，因為目標領域僅有未標記資料的非監督式跨領域訓練本就較為困難，且訓練資料為雜訊較多的網路輿情資料，而源領域產品資料與目標領域藝人評論資料差距太大，難以進行轉移，且為三分類任務，故使用本研究之師生跨領域模型的效果較為不好。

\* KKBOX 藝人評論領域為目標領域做半監督式跨領域訓練 (未導入相關度偵測任務)

而使用本研究所提出之師生跨領域模型，利用源領域的資訊，同時訓練目標領域的已標記資料以及未標記資料，與監督式領域內訓練結果相比，模型效能 (Macro-F1) 提升了近 2 個百分比，可以顯示本文之跨領域模型確實可以提升模型預測的效能。

\* KKBOX 藝人評論領域為目標領域做半監督式跨領域訓練 (導入相關度偵測任務)

最終，如同英文的實驗，我們導入相關度偵測的任務至師生跨領域模型一同進行多任務訓練，表 10 顯示相較於未導入相關度偵測任務的跨領域模型，導入相關度偵測任務的跨領域模型的效能更為提升，也驗證了導入相關度偵測任務的跨領域模型能增加效能的想法。

方法	Accuracy	Macro-F1
監督式領域內訓練	72.77%	75.57%
半監督式領域內訓練	69.48%	72.60%
監督式跨領域訓練	39.44%	30.96%
非監督式跨領域訓練	26.76%	25.70%
半監督式跨領域訓練 (無相關度偵測任務)	74.65%	77.26%
加入相關度偵測任務的最終模型效能	74.65%	77.65%

表 10 KKBOX 藝人評論領域作為目標領域實驗

#### — 核能輿論作為目標領域

\* 核能輿論領域做領域內訓練

使用 BERT 為核能輿論做領域內訓練中，由於為三分類問題且皆為雜訊較高的網路及新聞輿論資料，因此模型效能並非相當高。

\* 核能輿論領域做半監督式領域內訓練

使用 Google 所提出的 Unsupervised data augmentation [32] 方法的實驗結果較直接使用核能輿論做監督式領域內訓練效能低。

\* 核能輿論領域為目標領域做監督式跨領域訓練

使用 BERT 以 Amazon 三個領域產品意見資料訓練模型，以核能輿論領域作為目標領域的模型效能非常低，顯示出 Amazon 產品意見領域與核能輿論領域相差非常大，非常不適合轉移。

- \* 核能輿論領域為目標領域做非監督式跨領域訓練 (未導入相關度偵測任務)

本研究的師生跨領域模型應用在非監督式跨領域訓練的效果並不是很好，因為目標領域僅有未標記資料的非監督式跨領域訓練本就較為困難，和 KKBOX 藝人評論資料相同，訓練資料為雜訊較多的網路輿情資料，且源領域產品資料與目標領域核能輿論資料差距太大，難以進行轉移，並且為三分類任務，故使用本研究之師生跨領域模型的效果較為不好。

- \* 核能輿論領域為目標領域做半監督式跨領域訓練 (未導入相關度偵測任務)

而使用本研究所提出之師生跨領域模型並搭配目標領域已標記資料的監督式學習，利用源領域的資訊，同時訓練目標領域的已標記資料以及未標記資料，結果顯示效果 (Macro-F1)提升了近 1.2 個百分比，顯示本研究之跨領域模型可以提升效能。

- \* 核能輿論領域為目標領域做半監督式跨領域訓練 (導入相關度偵測任務)

導入相關度偵測的任務與師生跨領域模型並搭配目標領域已標記資料的監督式學習一同進行多任務訓練的整體模型，表 11 顯示，導入相關度偵測任務的跨領域模型的效能僅在 Macro-F1 值上超越用 BERT 模型做監督式領域內訓練的效能，且效能反而比未導入相關度偵測任務的跨領域模型來的低。在核能輿論實驗中由於沒有明確的描述目標主體，如人名或是產品，僅以 ”核能立場” 作為描述目標，故相關度偵測任務也難以使模型學習到資訊，使最終的模型效能反而比未導入相關度偵測任務來的低。

方法	Accuracy	Macro-F1
監督式領域內訓練	60.83%	60.54%
半監督式領域內訓練	59.07%	59.06%
監督式跨領域訓練	38.45%	38.17%
非監督式跨領域訓練	30.57%	30.11%
半監督式跨領域訓練(無相關度偵測任務)	61.66%	61.70%
加入相關度偵測任務的最終模型效能	60.83%	61.05%

表 11 核能輿論為目標領域實驗



## 第四節 模型參數分析

此章節提供模型能夠達到最佳效能的參數，在此會分英文資料集實驗和中文資料集實驗做分析。

### • 英文資料集實驗

表 12 歸納出在英文實驗中，使用本研究方法能夠達到最佳實驗結果的模型參數，並歸納出以下結論：

- 學習率在  $1e-5$  或  $2e-5$  的模型效能是最好的，此學習率範圍也十分接近 [11] 中建議使用 Adam optimizer 訓練 BERT 模型的學習率。
- 跨領域學習的權重與相關度偵測任務的權重在本研究中不需要太大即能為整體模型架構帶來效能的提升。
- 由於訓練資料中未標記資料為多數，因此在轉移學習任務中設置較大的訓練批次數量。
- 在相關度偵測的任務中，由於使用式 3.4 及式 3.5 的注意力機制式篩選負採樣資料的訓練效果不佳，因此改為使用式 3.6 及式 3.7 的半衰期機制訓練。在 Books、DVD 和 Kitchen appliances 這三個領域實驗結果顯示，使用半衰期機制訓練相關度偵測任務，若一開始所有的負採樣訓練資料皆為同領域的負採樣訓練資料，即  $\lambda$  是乘在同領域的負採樣訓練資料數量上的，模型的效果能達到最佳。
- 使用半衰期機制訓練相關度偵測任務，依據實驗結果顯示半衰期設為 2 (接近2)或 4 能為模型訓練帶來最好的效果。



模型參數	Books	DVD	Electronic	Kitchen appliances
學習率	1e-5	2e-5	2e-5	2e-5
跨領域學習的權重	0.2	0.2	0.2	0.4
相關度偵測任務的權重	0.2	0.2	0.2	0.8
監督式學習批次訓練數量	4	4	4	4
轉移學習批次訓練數量	16	16	16	16
相關度偵測任務學習批次訓練數量	4	4	4	4
一開始訓練時同領域負採樣資料比例 (相關度偵測)	1	1	0	1
半衰期 (相關度偵測)	1.82	2	4	2

表 12 英文資料實驗最佳模型參數

• 中文資料集實驗

表 13 歸納出在中文實驗中，使用本研究方法能夠達到最佳實驗結果的模型參數，並歸納出以下結論：

- 不論是 KKBOX 領域作為目標領域或是核能輿論作為目標領域，學習率在 2e-5 時的模型效能是最好的，此學習率也如 [11] 中建議使用 Adam optimizer 訓練 BERT 模型的學習率。
- 和英文實驗相似，跨領域學習的權重與相關度偵測任務的權重在本研究中不需要太大，即能為整體模型架構提升效能。
- 和英文實驗相似，由於訓練資料中未標記資料為多數，因此在跨領域轉移學習任務中設置較大的訓練批次數量。
- 在相關度偵測的任務中，使用注意力機制 (式 3.4 及式 3.5) 篩選負採樣資料，此機制能讓模型自動篩選出適合用來訓練的負採樣資料，實驗結果也顯示此機制確實能有提升的效果。

模型參數	KKBOX	核能輿論
學習率	2e-5	2e-5
跨領域學習的權重	0.1	0.1
相關度偵測任務的權重	0.1	0.3
監督式學習批次訓練數量	4	4
轉移學習批次訓練數量	16	16
相關度偵測任務學習批次訓練數量	4	4

表 13 中文資料實驗最佳模型參數



## 第五節 模型訓練資料數量分析

本節進行模型的訓練資料數量對於模型效能的分析。在英文資料集的實驗中使用了1600筆目標領域內已標記資料，本研究好奇訓練資料數量對於模型效能的影響，故分別輸入1200、800、400筆目標領域已標記資料(各標籤數量照原資料集標籤比例分配)到本研究的最終模型(半監督式師生跨領域模型及相關度偵測模型進行多任務訓練)進行訓練，在目標領域已標記資料量為0時，僅利用非監督式師生跨領域模型及相關度偵測模型的多任務模型進行實驗。英文資料實驗選取四個領域中的其中一個作為目標領域，剩餘的三個領域則作為源領域。另外此實驗以BERT模型對目標領域內的已標記資料進行領域內的監督式訓練效能作為比較值(已標記資料量為0時不討論)。以下將詳細列出不同實驗配置中，不同訓練資料的模型效能。

數量	Proposed Accuracy	Proposed Macro-F1	BERT Accuracy	BERT Macro-F1
1600	91.00%	91.00%	91.25%	91.25%
1200	89.25%	89.19%	91.00%	90.99%
800	89.75%	89.73%	90.00%	90.00%
400	88.00%	87.99%	89.25%	89.24%
0	87.50%	87.50%	None	None

表 14 Books 領域作為目標領域不同資料量效能

數量	Proposed Accuracy	Proposed Macro-F1	BERT Accuracy	BERT Macro-F1
1600	91.50%	91.49%	89.50%	89.50%
1200	89.25%	89.24%	89.00%	89.00%
800	89.25%	89.24%	87.75%	87.75%
400	88.25%	88.24%	89.00%	89.00%
0	86.75%	86.74%	None	None

表 15 DVD 領域作為目標領域不同資料量效能

數量	Proposed	Proposed	BERT	BERT
	Accuracy	Macro-F1	Accuracy	Macro-F1
1600	94.50%	94.50%	94.25%	94.25%
1200	93.00%	93.00%	92.75%	92.75%
800	91.00%	90.99%	92.50%	92.50%
400	90.50%	90.50%	92.25%	92.25%
0	88.25%	88.25%	None	None

表 16 Electronic 領域作為目標領域不同資料量效能

數量	Proposed	Proposed	BERT	BERT
	Accuracy	Macro-F1	Accuracy	Macro-F1
1600	94.50%	94.50%	93.75%	93.75%
1200	91.25%	91.24%	93.50%	93.48%
800	91.50%	91.48%	91.75%	91.75%
400	92.75%	92.74%	91.25%	91.25%
0	91.00%	90.99%	None	None

表 17 Kitchen appliances 領域作為目標領域不同資料量效能

在表 14、表 15、表 16、表 17 中，Proposed 為本研究之跨領域導入相關度偵測的模型，BERT 代表使用 BERT 模型對已標記資料進行領域內訓練的效能。單論本研究之模型而言，模型效能隨著目標領域已標記資料量的下降，模型效能呈現下降的趨勢，因此本研究的模型在目標領域輸入越多已標記資料訓練，則模型效能會越高。與 BERT 監督式訓練比較的結果顯示，除了在 Books 領域，本研究在所有不同資料量配置實驗效果皆略低於使用 BERT 監督式訓練之外，在另外三個領域中，本研究在越多已標記資料訓練的情況下，模型效能高於使用 BERT 做監督式訓練，也驗證了本研究之師生跨領域搭配相關度偵測模型能帶來更好的模型效能。

中文資料實驗由於 KKBOX 藝人評論的目標領域已標記資料量較少，故選擇核能輿論領域做目標領域訓練資料量與模型效能分析，並以 Amazon 中文產品回覆資

料的三個領域作為源領域。表 18 為中文資料集的實驗分析，Proposed 為本研究半監督式師生跨領域搭配相關度偵測模型的效能，在目標領域已標記資料量為 0 時，僅利用非監督式師生跨領域模型及相關度偵測模型的多任務模型進行實驗。BERT 為使用 BERT 模型為已標記資料進行領域內監督式訓練的模型效能 (已標記資料量為 0 時不討論)。單論本研究之模型效能，核能輿論領域作為目標領域訓練資料量分析和英文實驗相同，隨著核能輿論領域中的已標記資料量降低，模型效能也隨之降低。與 BERT 監督式訓練效能比較的結果在三種不同已標記資料配置，本研究之模型大部分較 BERT 做領域內監督式訓練高，顯示不只在英文實驗本研究之模型是有效的，換到中文資料本研究模型依然能提升效能。

數量	Proposed	Proposed	BERT	BERT
	Accuracy	Macro-F1	Accuracy	Macro-F1
3864	60.83%	61.05%	60.83%	60.54%
2000	56.06%	55.95%	57.10%	57.13%
1000	51.40%	51.18%	49.84%	49.48%
0	32.54%	32.53%	None	None

表 18 核能輿論領域作為目標領域訓練資料量分析

## 第六節 模型使用情境分析

在本節中會對於上述實驗中的方法，以及本研究的方法進行比較及詳細說明使用情境。

- 監督式領域內訓練

針對只有單一領域且領域內只有已標記資料

- 半監督式領域內訓練

針對只有單一領域但領域內含有已標記資料及未標記資料。在第四章的英文實驗中，使用 Google 所提出 Unsupervised data augmentation 方法的模型效能，能超越使用領域內已標記資料的監督式學習，因此若同領域內不但擁有已標記資料和未標記資料，則適合使用半監督式領域內訓練。

- 監督式跨領域訓練

監督式跨領域訓練中，源領域以及目標領域皆僅使用已標記資料。根據第四章的實驗結果，監督式跨領域訓練效能會較監督式領域內訓練以及半監督式領域內訓練還差，原因是因為源領域及目標領域的差異，導致所訓練出來的模型不適合應用於目標領域資料，且並沒有特殊機制為目標領域資料篩選出源領域資料作特徵轉移。

- 非監督式跨領域訓練 (本研究方法)

非監督式跨領域轉移使用情景為目標領域僅有未記資料的跨領域訓練。使用此訓練方法的效果會較使用監督式領域內訓練、半監督式領域內訓練、監督式跨領域訓練差，因為目標領域僅為使用未標記資料的非監督式訓練。

- 半監督式跨領域訓練 (本研究方法)

半監督式跨領域訓練適用的情景為目標領域擁有已標記資料及未標記資料，此訓練方法也能較上述所有訓練方法(監督式領域內訓練、半監督式領域內訓練、監督式跨領域訓練、非監督式跨領域訓練)的效能高。因為模型結合了監督式訓練及非監督式跨領域訓練，藉由跨領域知識的幫助，以及額外的相關度偵測任務，使整體模型效能能達到最高。

綜合上述分析，使用本研究的半監督式師生跨領域模型搭配相關度偵測任務的多任務訓練模型，能夠在所有無論領域內訓練或是跨領域訓練中，達到最好的效果。



## 第五章 結論

本篇提出了全新的師生教學式的跨領域轉移學習架構，有效地解決了不論是非監督式的跨領域轉移問題或是半監督式的跨領域轉移問題。模型架構中的教師模型使用了非監督式資料增強學習，不但可以為學生模型創造更多低噪音的偽標記資料並學習，學生模型藉由與多層次的注意力學習機制合作，能順利轉移適合目標領域的源領域特徵。模型同時導入了相關度偵測的多任務學習，藉由讓模型學習描述文字與目標主體之間的相關度，使模型效能有所提升。實驗結果顯示本研究的模型架構超越了所有現存的方法，達到此跨領域轉移問題資料集的最佳表現。除此之外，本研究的方法也能應用於多種情境，就上述實驗來說，除了可以應用於英文資料集判斷顧客對於產品意見回覆的情緒立場，更可以應用於中文資料集輿論的立場判斷上，驗證了模型在不同分類任務仍能維持穩定。

實驗結果顯示，本篇提出的跨領域多任務訓練模型，不但解決了跨領域問題中的兩個主要問題，分別是從源領域挑選適合用來轉移至目標領域的資料以及減少源領域以及目標領域的歧異性，也能在有限數量已標記資料的情況下，使用本文之模型進行跨領域轉移學習及半監督式學習模型訓練不論是源領域或是目標領域的已標記資料及未標記資料，達到比僅使用已標記資料的監督式學習所訓練出的模型更高的效能。相關度偵測任務的導入也驗證了描述文字與目標描述主體之間的相關度能讓模型學習更多用來判斷立場的資訊，同時也證明了多任務學習確實能更勝單一任務學習的效能。

在本文提出的英文實驗中，全部的領域皆會輪流作為源領域以及目標領域，由此反覆驗證本文的模型能夠利用不同領域作為源領域以及目標領域進行訓練，仍能有效的提高效能，顯示了本文提出之跨領域模型的穩定度非常高。除了使用英文資料集做實驗之外，我們更使用了中文的資料集進行實驗，實驗結果也顯示本研究模型能有效的解決跨領域轉移問題以及少量已標記資料的問題，也驗證了本文模型能應用於兩種不同語言的資料集上。



## 參考文獻

- [1] Anthony Aue and Michael Gamon. “Customizing Sentiment Classifiers to New Domains: A Case Study”. In: Jan. 2005.
- [2] John Blitzer, Mark Dredze, and Fernando Pereira. “Biographies, Bollywood, Boomboxes and Blenders: Domain Adaptation for Sentiment Classification”. In: *Proceedings of the 45th Annual Meeting of the Association of Computational Linguistics*. Prague, Czech Republic: Association for Computational Linguistics, June 2007, pp. 440–447. URL: <https://www.aclweb.org/anthology/P07-1056>.
- [3] John Blitzer, Ryan McDonald, and Fernando Pereira. “Domain Adaptation with Structural Correspondence Learning”. In: *Proceedings of the 2006 Conference on Empirical Methods in Natural Language Processing*. Sydney, Australia: Association for Computational Linguistics, July 2006, pp. 120–128. URL: <https://www.aclweb.org/anthology/W06-1615>.
- [4] Danushka Bollegala, David Weir, and John Carroll. “Cross-Domain Sentiment Classification Using a Sentiment Sensitive Thesaurus”. In: *IEEE Transactions on Knowledge and Data Engineering* 25.8 (2013), pp. 1719–1731. DOI:10.1109/TKDE.2012.103.
- [5] Danushka Bollegala, David Weir, and John Carroll. “Using Multiple Sources to Construct a Sentiment Sensitive Thesaurus for Cross-Domain Sentiment Classification”. In: *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies*. Portland, Oregon, USA: Association for Computational Linguistics, June 2011, pp. 132–141. URL: <https://www.aclweb.org/anthology/P11-1014>.
- [6] Minmin Chen et al. “Marginalized Denoising Autoencoders for Domain Adaptation”. In: *CoRR* abs/1206.4683 (2012). arXiv: 1206.4683. URL: <http://arxiv.org/abs/1206.4683>.
- [7] Junyoung Chung et al. “Empirical Evaluation of Gated Recurrent Neural Networks on Sequence Modeling”. In: *CoRR* abs/1412.3555 (2014). arXiv: 1412.3555. URL: <http://arxiv.org/abs/1412.3555>.
- [8] Xia Cui and Danushka Bollegala. *Multi-source Attention for Unsupervised Domain Adaptation*. 2020. arXiv: 2004.06608 [cs.CL].
- [9] Xia Cui and Danushka Bollegala. “Self-Adaptation for Unsupervised Domain Adap-

- tation”. In: *RANLP*. 2019.
- [10] Yong Dai et al. *Adversarial Training Based Multi-Source Unsupervised Domain Adaptation for Sentiment Analysis*. 2020. arXiv: 2006.05602 [cs.CL].
- [11] Jacob Devlin et al. “BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding”. In: *CoRR* abs/1810.04805 (2018). arXiv: 1810.04805. URL: <http://arxiv.org/abs/1810.04805>.
- [12] X. Ding et al. “Learning Multi-Domain Adversarial Neural Networks for Text Classification”. In: *IEEE Access* 7 (2019), pp. 40323–40332. DOI: 10.1109/ACCESS.2019.2904858.
- [13] Hady Elsahar and Matthias Gallé. “To Annotate or Not? Predicting Performance Drop under Domain Shift”. In: *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)*. Hong Kong, China: Association for Computational Linguistics, Nov. 2019, pp. 2163–2173. DOI: 10.18653/v1/D19-1222. URL: <https://www.aclweb.org/anthology/D19-1222>.
- [14] Yaroslav Ganin et al. “Domain-Adversarial Training of Neural Networks”. In: *Journal of Machine Learning Research* 17.59 (2016), pp. 1–35. URL: <http://jmlr.org/papers/v17/15-239.html>.
- [15] Deepanway Ghosal et al. “KinGDOM: Knowledge-Guided DOMain Adaptation for Sentiment Analysis”. In: *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*. Online: Association for Computational Linguistics, July 2020, pp. 3198–3210. DOI: 10.18653/v1/2020.acl-main.292. URL: <https://www.aclweb.org/anthology/2020.acl-main.292>.
- [16] Arthur Gretton et al. “A Kernel Two-Sample Test”. In: *Journal of Machine Learning Research* 13.25 (2012), pp. 723–773. URL: <http://jmlr.org/papers/v13/gretton12a.html>.
- [17] Han Guo, Ramakanth Pasunuru, and Mohit Bansal. *Multi-Source Domain Adaptation for Text Classification via DistanceNet-Bandits*. 2020. arXiv: 2001.04362 [cs.CL].
- [18] Jiang Guo, Darsh J. Shah, and Regina Barzilay. “Multi-Source Domain Adaptation with Mixture of Experts”. In: *CoRR* abs/1809.02256 (2018). arXiv: 1809.02256. URL: <http://arxiv.org/abs/1809.02256>.

- [19] Sepp Hochreiter and Jürgen Schmidhuber. “Long Short-Term Memory”. In: *Neural Computation* 9.8 (1997), pp. 1735–1780. DOI: 10.1162/neco.1997.9.8.1735. eprint: <https://doi.org/10.1162/neco.1997.9.8.1735>. URL: <https://doi.org/10.1162/neco.1997.9.8.1735>.
- [20] Yann LeCun, Yoshua Bengio, and Geoffrey Hinton. “Deep Learning”. In: *Nature* 521.7553 (2015), pp. 436–444. DOI: 10.1038/nature14539. URL: <https://doi.org/10.1038/nature14539>.
- [21] Zheng Li et al. “End-to-End Adversarial Memory Network for Cross-domain Sentiment Classification”. In: *Proceedings of the Twenty-Sixth International Joint Conference on Artificial Intelligence, IJCAI-17*. 2017, pp. 2237–2243. DOI: 10.24963/ijcai.2017/311. URL: <https://doi.org/10.24963/ijcai.2017/311>.
- [22] Zheng Li et al. “Hierarchical Attention Transfer Network for Cross-domain Sentiment Classification”. In: Jan. 2018.
- [23] [https://en.wikipedia.org/wiki/Jensen%E2%80%93Shannon\\_divergence](https://en.wikipedia.org/wiki/Jensen%E2%80%93Shannon_divergence)
- [24] Sinno Jialin Pan et al. “Cross-Domain Sentiment Classification via Spectral Feature Alignment”. In: *Proceedings of the 19th International Conference on World Wide Web*. WWW ’10. Raleigh, North Carolina, USA: Association for Computing Machinery, 2010, pp. 751–760. ISBN: 9781605587998. DOI: 10.1145/1772690.1772767. URL: <https://doi.org/10.1145/1772690.1772767>.
- [25] Alan Ramponi and Barbara Plank. *Neural Unsupervised Domain Adaptation in NLP—A Survey*. 2020. arXiv: 2006.00632 [cs.CL].
- [26] Sebastian Ruder. “An Overview of Multi-Task Learning in Deep Neural Networks”. In: *CoRR* abs/1706.05098 (2017). arXiv: 1706.05098. URL: <http://arxiv.org/abs/1706.05098>.
- [27] Sebastian Ruder and Barbara Plank. “Strong Baselines for Neural Semi-supervised Learning under Domain Shift”. In: *CoRR* abs/1804.09530 (2018). arXiv: 1804.09530. URL: <http://arxiv.org/abs/1804.09530>.
- [28] Sainbayar Sukhbaatar et al. “Weakly Supervised Memory Networks”. In: *CoRR* abs/1503.08895 (2015). arXiv: 1503.08895. URL: <http://arxiv.org/abs/1503.08895>.

- [29] Pascal Vincent et al. “Extracting and Composing Robust Features with Denoising Autoencoders”. In: *Proceedings of the 25th International Conference on Machine Learning*. ICML '08. Helsinki, Finland: Association for Computing Machinery, 2008, pp. 1096–1103. ISBN: 9781605582054. DOI: 10.1145/1390156.1390294. URL: <https://doi.org/10.1145/1390156.1390294>.
- [30] Jason W. Wei and Kai Zou. “EDA: Easy Data Augmentation Techniques for Boosting Performance on Text Classification Tasks”. In: *CoRR* abs/1901.11196 (2019). arXiv: 1901.11196. URL: <http://arxiv.org/abs/1901.11196>.
- [31] Fangzhao Wu and Yongfeng Huang. “Sentiment Domain Adaptation with Multiple Sources”. In: *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*. Berlin, Germany: Association for Computational Linguistics, Aug. 2016, pp. 301–310. DOI: 10.18653/v1/P16-1029. URL: <https://www.aclweb.org/anthology/P16-1029>.
- [32] Qizhe Xie et al. “Unsupervised Data Augmentation”. In: *CoRR* abs/1904.12848 (2019). arXiv: 1904.12848. URL: <http://arxiv.org/abs/1904.12848>.
- [33] Shujuan Yu et al. “Hierarchical Data Augmentation and the Application in Text Classification”. In: *IEEE Access* PP (Dec. 2019), pp. 1–1. DOI: 10.1109/ACCESS.2019.2960263.
- [34] Han Zhao et al. “Adversarial Multiple Source Domain Adaptation”. In: *Advances in Neural Information Processing Systems*. Ed. by S. Bengio et al. Vol. 31. Curran Associates, Inc., 2018, pp. 8559–8570. URL: <https://proceedings.neurips.cc/paper/2018/file/717d8b3d60d9eea997b35b02b6a4e867-Paper.pdf>.