

透過語音特徵建構基於堆疊稀疏自編碼器演算法之  
婚姻治療中夫妻互動行為量表自動化評分系統

**Automating Behavior Coding for Distressed Couples  
Interactions Based on Stacked Sparse Autoencoder  
Framework using Speech-acoustic Features**

陳柏軒\*、李祈均\*

**Po-Hsuan Chen and Chi-Chun Lee**

**摘要**

在過去人類行為分析是透過傳統人為觀察方式來記錄。像婚姻治療方面，評分者利用觀看錄影的方式來對一整段夫妻對話中所展現的行為作評分。藉由這樣取得各種行為表達程度的量化，針對此量化分數來更進一步研究夫妻婚姻治療成效，但這種做法非常耗時且會因為評分者的各種主觀因素影響最後的準確性。如果能透過機器學習的方式來自動化處理辨識，將會節省非常多的人工時間和提升客觀性。深度學習(Deep Learning)在目前機器學習上是很熱門的話題。本論文提出以堆疊稀疏自編碼器(Stacked Sparse Autoencoder, SSAE)方式對聲音訊號特徵進行降維，並找出相對關鍵的高階特徵，最後再利用邏輯迴歸分析(Logistic Regression, LR)來辨識。此方法的整體準確率為 75%(丈夫行為平均辨識準確率為 74.9%、太太為 75%)。相對於過去研究的 74.1% (丈夫行為平均準確率 75%，太太為 73.2%) (Black *et al.*, 2013)，提升 0.9%。我們提出的方法在使用更低維度的聲音特徵值中可有效的提升行為辨識準確率。

**關鍵詞：**深度學習，堆疊自編碼器，婚姻治療，人類行為分析，情緒分析

---

\*國立清華大學電機工程學系

Department of Electrical Engineering, National Tsing Hua University

E-mail: theusa20@gmail.com; ccleee@ee.nthu.edu.tw

## Abstract

Traditional way of conducting analyses of human behaviors is through manual observation. For example in couple therapy studies, human raters observe sessions of interaction between distressed couples and manually annotate the behaviors of each spouse using established coding manuals. Clinicians then analyze these annotated behaviors to understand the effectiveness of treatment that each couple receives. However, this manual observation approach is very time consuming, and the subjective nature of the annotation process can result in unreliable annotation. Our work aims at using machine learning approach to automate this process, and by using signal processing technique, we can bring in quantitative evidence of human behavior. Deep learning is the current state-of-art machine learning technique. This paper proposes to use stacked sparse autoencoder (SSAE) to reduce the dimensionality of the acoustic-prosodic features used in order to identify the key higher-level features. Finally, we use logistic regression (LR) to perform classification on recognition of *high* and *low* rating of six different codes. The method achieves an overall accuracy of 75% over 6 codes (husband's average accuracy of 74.9%, wife's average accuracy of 75%), compared to the previously-published study of 74.1% (husband's average accuracy of 75%, wife's average accuracy of 73.2%) (Black *et al.*, 2013), a total improvement of 0.9%. Our proposed method achieves a higher classification rate by using much fewer number of features (10 times less than the previous work (Black *et al.*, 2013)).

**Keywords:** Deep Learning, Stacked Autoencoders, Couple Therapy, Human Behavior Analysis, Emotion Recognition

## 1. 緒論

人與人之間交談互動，常透過語言傳達彼此的想法，並在這交談過程中得知雙方的行為反應。利用人為觀察來分析雙方行為反應，這部分最早常應用在心理學和精神學方面 (O'Brian *et al.*, 1994)。人為行為觀察相當的成功研究在親密關係 (Karney & Bradbury, 1995) (Gonzaga *et al.*, 2007)，即夫妻的行為是影響親密關係程度的因素之一。然而用於人為觀察行為的方式存在一些困難，一方面太消耗時間，另一面也浪費成本。

如果能透過電腦工程的方式來取代人為觀察將大大提升效率，透過低層描述映射高層描述來預測人類行為 (Schuller *et al.*, 2007)，這項研究領域是正在不斷發展的一部分。人類行為信號處理(Behavioral Signal Processing, BSP)目的在幫助連接信號科學和行為處理的方法，建立在傳統的信號處理研究，如語音識別，面手部追蹤等等。相關顯著 BSP 研究已發產於以人為中心的提取音頻，視頻信號，來分析實際上人類行為或是情感方面 (Burkhardt *et al.*, 2009; Devillers & Campbell, 2011)。

## 婚姻治療中夫妻互動行為量表自動化評分系統

本論文利用 BSP 的基本思路應用在婚姻治療資料庫上面 (Christensen *et al.*, 2004)，婚姻治療資料庫會詳細說明在第二章。這個資料庫紀錄了夫妻在一段對話中談述了他們所選擇婚姻中的問題。評分者在根據他們一段話的種種行為根據不同行為量表進行評分(幽默行為、悲傷行為等等)。

延續上篇論文的研究內容來自動化分析夫妻一段對話的行為分數(Black *et al.*, 2013)，一段語音經過預處理，之後作聲音特徵擷取(acoustic feature extraction)，再使用機器學習來作分類辨識，得到最後的準確率。其中，特徵擷取和機器學習的算法都會影響最後的準確率，思考如何改進這些影響因素，對整體準確率的提升是一大重要的課題，也是我們提出這篇論文的因素之一。

在特徵擷取方面，我們沿用三種低階語音特徵(Low Level Descriptors, LLDs)，語韻(prosodic) LLDs、頻譜(spectrum) LLDs 和音質(voice quality) LLDs。切割三種說話者說話區間(speaker domain)，丈夫說話區間、太太說話區間、和不分人說話區間。再來對應各區間提取 20% 語句，經過 7 種統計函數(functionals)，產生 2940 種特徵值。最後我們利用非監督深度學習的做法來降維找出相對關鍵的主要特徵值表現。

深度學習在機器學習領域裡面是最近熱門的話題 (Hinton, 2006)。深度學習可看成是一種資訊的表達方式，利用多層神經網絡，第一層輸入的數據學習之後，產生新的組合輸出，輸出值為第二層的輸入值，再經由學習產生新的輸出值，依此類推重覆把每層的資訊堆疊下去，透過這樣多層學習，可以得到對一個目標值好的特徵表示，相對準確率就能有所提升。至今存在多種深度學習框架如深度神經網路(DNN)、深度信念網路(DBN)和卷積神經網路(CNN)已被應用在語音 (Hinton *et al.*, 2012)、影像辨識 (Smirnov *et al.*, 2014)和手寫識別 (Perwey & Chaturvedi, 2011)等等。

我們利用深度學習中的堆疊稀疏自編碼器(stacked sparse autoencoder, SSAE)，降低特徵值維度，提升特徵值整體相關性，最後利用簡單 LR 辨識行為分數高低。此初期研究結果顯示整體行為平均準確率 75%較之前研究使用 40479 維特徵值結合支持向量器(support vector machine) (Black *et al.*, 2013)提升了 0.9%。

以下簡述各章節的內容。第二章介紹本篇論文所使用的資料庫(database)，第三章介紹我們使用的 SSAE 架構和其演算法，第四章介紹我們提出的系統架構和研究結果，第五章節為結論。

## 2. 婚姻治療資料庫

為了測試我們提出方法的準確率，我們使用和之前論文相同的婚姻治療資料庫(couple therapy database)。以下簡單的介紹的資料庫相關內容：此資料庫的收集是基於研究綜合行為夫婦治療(integrative behavioral couple therapy, IBCT)成效 (Christensen *et al.*, 1995)。資料內容針對 134 對夫妻，每對都長期患有婚姻的問題，如夫妻相處不融洽或是爭執。

治療內容為每對夫妻接受為期一年的治療，研究團隊再讓每對夫妻由太太和丈夫各別選擇一個目前存在嚴重婚姻問題的題目來作為一段 10 分鐘對話，對話中沒有治療師和

研究團隊。透過這 10 分鐘的對話讓夫妻彼此了解雙方之間的問題並且試圖解決當前問題。

每對夫妻皆會進行三個不同階段的對話，治療前、治療中和治療兩年後。透過這三個時間點對話，再經由多位有專業背景的評分者經由兩個行為評分量表，基於社交互動行為評分系統(Social Support Interaction Rating System, SSIRS) (Jones & Christensen, 1998) 和基於夫妻互動行為評分系統(Couples Interaction Rating System, CIRS) (Heavey *et al.*, 2002)進行評分，依據評分結果來了解治療的成效。SSIRS 主要包含 19 種行為準則在四個社交互動分類裡，情感(affectivity)、屈從服從(dominance/submission)、互動表現行為(feature of interaction)和主題評價(topic definition)來作為評分的內容，CIRS 主要包含 13 種行為準則關於夫妻互動問題解決方面，如表 1。

**表 1. 32 種人類行為準則包含在兩種行為量表 SSIRS 和 CIRS**

Manual	Codes
<p style="text-align: center;">SSIRS (Social Support Interaction Rating System)</p>	<p>Global positive affect、global negative affect use of humor、sadness、anger/frustration、 belligerence/domineering、contempt/disgust、 tension/anxiety、defensiveness、affection、 satisfaction、solicits partner suggestions、 instrumental support offered、emotional support offered、submissive or dominant、topic a relationship issue、topic a personal issue、 discussion about husband、discussion about wife</p>
<p style="text-align: center;">CIRS (Couples Interaction Rating System)</p>	<p>Acceptance of other、blame、responsibility for self、solicits partner perspective、states external origins、discussion、clearly defines problem、 offers solutions、negotiates、make agreements、pressures for change、 withdraws、avoidance</p>

總共 32 個行為準則，每個行為評分區間為 1 到 9 分。同一對話中，丈夫與妻子會各別被評分。1 為對這項行為所表現的程度最低，9 為對這項行為所表現的程度最高。評分者為 3 到 4 個，透過觀察夫妻 10 分鐘的影片來各別對 32 個行為進行評分。最後總共有 569 個 10 分鐘的會話，117 對夫妻在這個婚姻治療庫裡。

本篇論文延續上一篇論文所使用的 6 種行為來下去作分析，包含認同對方(Acceptance of other)、責備行為(Blame)、夫妻之間正面的互動(Global positive affect)、夫妻之間負面的互動(Global positive affect)、悲傷行為(sadness)、幽默表現行為(humor)，如

表 2。之所以會選擇這 6 種行為，因為和其他 26 種行為評分比起來，這 6 種有較高的評分者認同度(Agreement)，認同度的計算方式為個別評分者的分數和其他評分者評分的平均分數取相關係數(correlation)。其餘行為的認同度介於 0.4 和 0.7 之間，第五章節會比較這 6 種行為預測準確率。

表 2. 對於 6 種行為準則的認同度(agreement)

Code	Agreement
Acceptance of other (acc)	0.751
Blame (bla)	0.788
Global positive affect (pos)	0.740
Global negative affect (neg)	0.798
Sadness (sad)	0.722
Use of humor (hum)	0.755

### 3. 研究方法

在本節，我們首先簡單的介紹自編碼器(Autoencoder)和堆疊稀疏自編碼(Stacked Sparse Autoencoder, SSAE)基本架構以及本篇論文用到的演算法。

#### 3.1 自編碼器(Autoencoder)

深度學習中自編碼器利用非監督學習方式 (Rubanov, 2000)，目標從高維度的輸入特徵值學習到更具代表性的特徵值，最後透過解碼讓輸出值等於輸入值，基本的自編碼器架構如圖 1。

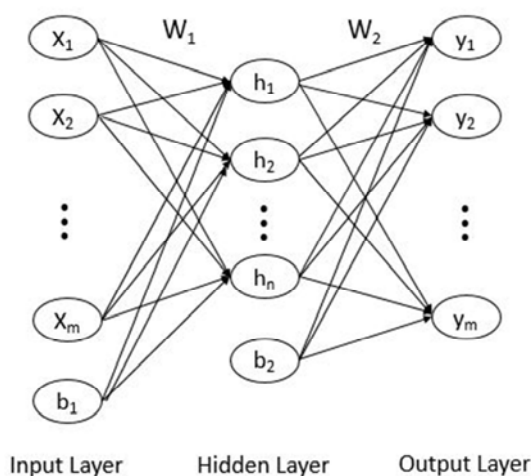


圖 1. 自編碼器

從圖 1，輸入值  $x_i, i = 1, 2, \dots, m, x \in R^m$ ，隱藏層(hidden layer)中的  $h_i, i = 1, 2, \dots, n, h \in R^n$ ，權重矩陣(weight matrix)  $W_1 \in R^{n \times m}$ ，偏移向量(bias vector)  $b_1 \in R^n$ 。由這些因子(factor)構成激活函數(activation function)，如式(1)。

$$h(x) = f(W_1x + b_1) \quad (1)$$

其中  $f(x) = 1/(1 + \exp(-z))$  為 sigmoid function。輸出值  $y_i, i = 1, 2, \dots, m, y \in R^m$ ，權重矩陣  $W_2 \in R^{m \times n}$ ，偏移向量  $b_2 \in R^m$ ，自編碼器輸出為式(2)：

$$y = f(W_2h(x) + b_2) \quad (2)$$

為了要求得權重矩陣  $W_1$  和  $W_2$ ，偏移向量  $b_1$  和  $b_2$ ，假設一個樣本集為  $\{(x_1, y_1), (x_2, y_2) \dots (x_m, y_m)\}$ ，有  $m$  組樣本， $x_i$  為樣本輸入特徵值， $y_i$  為對應標籤值，利用代價函數(cost function)，如式(3)。

$$J(W, b) = \left[ \frac{1}{m} \sum_{i=1}^m \left( \frac{1}{2} \|h(x^{(i)}) - y^{(i)}\|^2 \right) \right] + \frac{\lambda}{2} \sum_{l=1}^n \sum_{i=1}^{p_l} \sum_{j=1}^{p_{l+1}} (w_{i,j}^{(l)})^2 \quad (3)$$

式(3)中第一項為均方差項(sum-of-squares error term)，第二項為規則項(regularization term)，其中  $\lambda$  為權重衰減參數(weight decay parameter)， $n$  為自編碼器層數， $p_l$  為第  $l$  層節點數，這項是為了避免訓練過程發生過擬合(overfitting)，之後我們利用反向傳導(back-propagation)演算法和 L-BFGS 優化算法 (Andrew & Gao, 2007)，重複疊代減小  $J(W, b)$  值，最後得到  $W$  和  $b$ 。

而為了讓輸入特徵值更有效的歸類群集並且不同特徵之間的區隔明顯， $J(W, b)$  加入稀疏項(sparsity term)如式(4)，取名為稀疏編碼器(sparse autoencoder) (Obst, 2014)。

$$J_s(W, b) = J(W, b) + \beta \sum_{j=1}^q KL(\rho || \hat{p}_j) \quad (4)$$

其中  $KL = \rho \log \frac{\rho}{\hat{p}_j} + (1 - \rho) \log \frac{1 - \rho}{1 - \hat{p}_j}$ ， $\rho$  為稀疏參數(sparsity parameter)， $\hat{p}_j = \frac{1}{m} \sum_{i=1}^m h_j(x_i)$ ， $\beta$  為控制稀疏項(sparsity term)的參數， $q$  為隱藏層的節點數。

### 3.2 堆疊稀疏自編碼器(Stacked Sparse Autoencoder)

由多個稀疏自編碼器逐層訓練後，堆疊組成的架構為堆疊稀疏自編碼器(Stacked Sparse Autoencoder)，如圖 2，每一層的編碼後輸出為下一層的輸入。從圖 2 可看出，輸入層(Input layer)經由第一個稀疏自編碼器訓練完之後得到第一隱藏層(Hidden layer1)的  $n$  個節點，由這  $n$  個節點在經過第二個稀疏自編碼器訓練得到第二隱藏層(Hidden layer2)的  $p$  個節點，每層的隱藏層節點可視為由上一層產生新的一組特徵，透過這樣逐層訓練可以訓練更多層。

我們實驗採用堆疊稀疏自編碼器(Stacked Sparse Autoencoder, SSAE)，希望透過 SSAE 得到好的特徵表示方式，最後經由分類器產生更好的準確率。

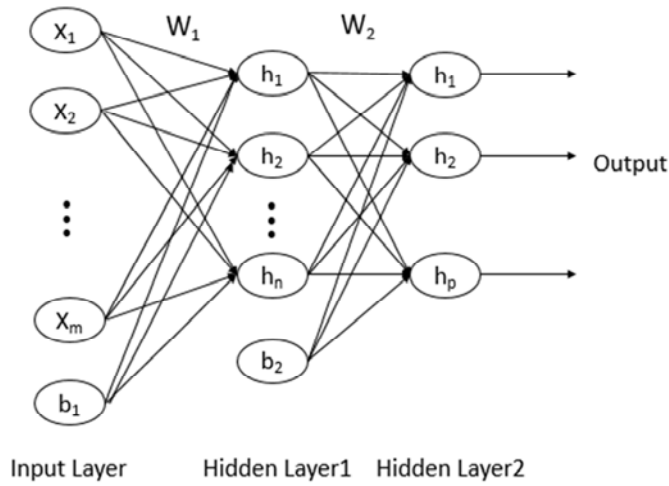


圖 2. 堆疊稀疏自編碼器

### 3.3 實驗架構

我們使用 3 層隱藏層的 SSAE 作為非監督學習的架構，來從低層級特徵(low level feature) 訓練成高層級特徵(high level feature)，然後用 LR 來監督學習作辨識，本實驗第一層稀疏自編碼器架構如圖 3。

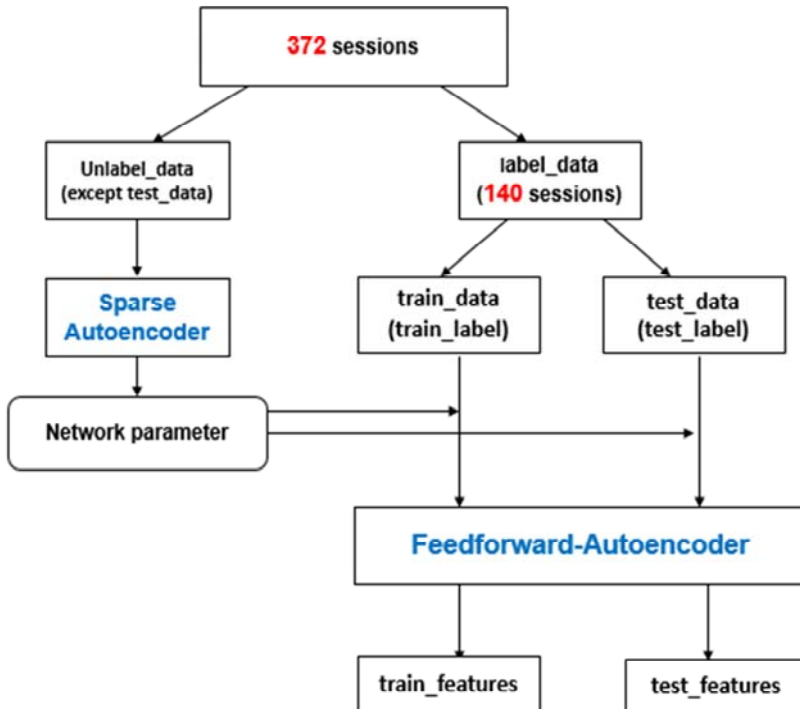


圖 3. 實驗架構

一段語音經過預處理，降低雜訊影響，才不會影響之後的特徵擷取，而這部分預處理在上篇論文已經被處理過了 (Black *et al.*, 2013)。本篇論文改變特徵擷取方法，這部分下一章會介紹。如圖 3，正規化後的特徵值，一種行為包含 372 筆 10 分鐘會話(session)，分為有標籤數據(labeled data)和沒有標籤數據(unlabeled data)，沒有標籤數據利用稀疏自編碼器來訓練網絡參數，訓練好後再把 140 筆有標籤數據分為訓練資料和測試資料，輸入自訓練好的網絡參數，產生新的一組特徵。新的一組特徵為下一層輸入值，重複利用圖 3 架構可以產生更多層。我們希望新的特徵值對於行為分數將有更好的表示，下面章節會證明之。

## 4. 實驗設計和結果

### 4.1 特徵值

如圖 4，利用原本 LLDs，在三種對話區間裡(speaker domain)，丈夫時間區間(husband、H)、太太時間區間(wife、W)和不分人時間區間(full、F)所說句子，切割成以 20% 句子為一個時間區間，切割完後合成一個行向量，行向量的特徵值，再經由如表 3 所列的 7 種 functionals 處理過後，產生最後 2940 個特徵值。在輸入 SSAE 以前，我們把這些特徵值正規化在 0 和 1 的區間。詳細的特徵值內容可參考 (Black *et al.*, 2013)。

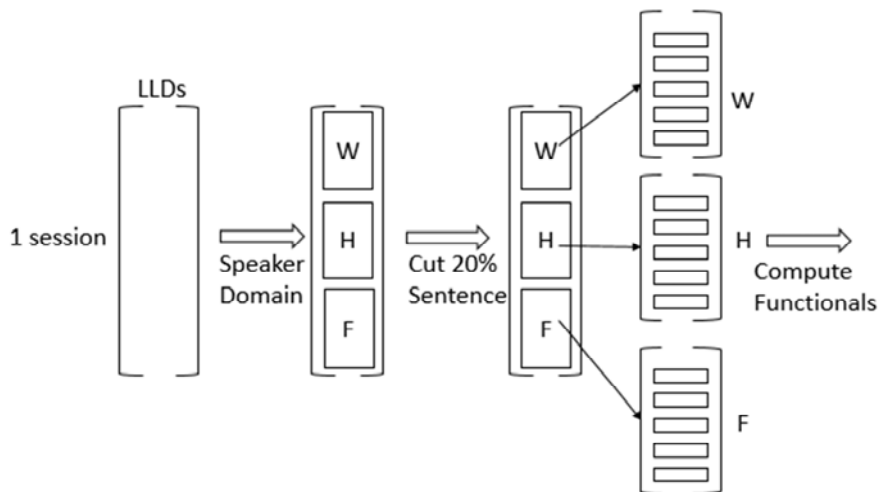


圖 4. 實驗特徵提取架構



表 3. 28 種特徵值和 7 種 functionals

LLDs	Functionals
1. MFCC[0-14]	1. Mean
2. MFB[0-7]	2. Median
3. F0normlog	3. Standard deviation
4. VAD(speech/no speech)	4. Skewness
5. Intensity	5. Kurtosis
6. Jitter	6. Max position
7. Jitter of Jitter	7. Min position
8. Shimmer	

## 4.2 資料

由原本資料庫 569 筆對話、117 對夫妻，經由上篇論文預處理過後(Black *et al.*, 2013)，產生最後的 372 筆對話、104 對夫妻。在 372 筆對話裡面丈夫和太太都會被評分到，對應在 6 種行為準則，我們選擇前 20% 的分數和後 20% 的分數的對話當作實驗的辨識，共 140 筆對話；兩種標籤值 0 和 1，1 為對應到高分，0 為對應到低分。而在這些取出來被預測的對話裡，夫妻數介於 68 到 77 對，利用這些行為對應到夫妻對數來作交叉驗證，1 對夫妻作驗證，其餘對數作訓練，重複循環 6 種行為對應到的夫妻對數來作驗證。

## 4.3 實驗設定

在這實驗裡，我們用 SSAE 來作為非監督學習，LR 來監督學習預測，留一對夫妻法則 (leave-one-couple-out) 的方式來作交叉驗證。一開始先用貪婪訓練算法 (greedy layerwise) 逐層預學習 (pre-training)，訓練完參數初始值輸入至 SSAE，SSAE 有五個因子會影響最後的表現，分別是隱藏層節點 (hidden units)、計算損失函數 (cost function) 的疊代次數和三個超參數 (hyper-parameters) 為  $\lambda$ 、 $\rho$ 、 $\beta$ ， $\lambda$  為權重衰減參數 (weight decay parameter)， $\rho$  為稀疏參數 (sparsity parameter)， $\beta$  為控制稀疏項 (sparsity term) 的參數，這些參數在第三章有介紹過。我們先用 1 層隱藏層來測試準確率，如表 4。透過改變不同的隱藏層節點數，根據準確率來決定我們下一層所使用的隱藏層節點數。

如表 4 可得知，隱藏層數目為 300 的時候，丈夫和太太被評分的 6 種行為平均準確率為最高，使用的疊代次數為 15 次， $\rho = 0.1$ ， $\lambda = 0.002$ ， $\beta = 2$ 。

接下來測試二層隱藏層的準確率，第一層隱藏數已經決定好了，我們測試的第二層隱藏層節點數，如表 5。從表中得知，第二層隱藏層的節點數為 200 的時候，準確率為最高，使用的疊代次數為 15 次， $\rho = 0.1$ ， $\lambda = 0.0001$ ， $\beta = 1$ 。

表 4. 1st hidden unit 分析丈夫和太太對應到 6 種行為的準確率，粗體字為較高的準確

1 <sup>st</sup> hidden unit	Rated Spouse	Acc (%)	Bla (%)	Pos (%)	Neg (%)	Sad (%)	Hum (%)	Avg (%)
100	Husband	67.9	76.4	65.7	78.6	52.9	61.4	67.2
	Wife	70	73.6	65	74.3	58.6	59.3	66.8
200	Husband	72.9	76.4	71.4	82.1	57.1	67.1	71.2
	Wife	71.4	82.9	65.7	77.1	64.9	57.9	70
300	Husband	77.1	77.9	72.1	82.9	58.6	67.1	<b>72.6</b>
	Wife	75.7	82.1	71.4	78.6	58.6	63.6	<b>71.7</b>
500	Husband	70	78.6	68.6	82.9	55	62.1	69.5
	Wife	74.3	82.1	69.3	80.7	58.6	62.9	71.3
1000	Husband	75	77.9	69.3	84.3	58.6	65.7	71.8
	Wife	72.1	79.3	69.3	80	53.6	62.9	69.5
Previous method(Black <i>et al.</i> , 2013)	Husband	78.6	72.9	72.1	84.3	60	71.4	73.2
	Wife	77.9	84.3	74.3	80	66.4	67.1	75

表 5. 2nd hidden unit 分析丈夫和太太對應到 6 種 code 的準確率，粗體字為較高的準確率

1 <sup>st</sup> Hidden Layer	2 <sup>nd</sup> Hidden Layer	Rated Spouse	Acc (%)	Bla (%)	Pos (%)	Neg (%)	Sad (%)	Hum (%)	Avg (%)
300	100	Husband	75	78.6	68.6	83.6	57.9	67.9	71.9
		Wife	71.4	80.7	72.9	77.1	58.6	62.9	70.6
	200	Husband	77.1	77.1	71.4	83.6	57.9	69.3	<b>72.7</b>
		Wife	72.1	82.1	72	77.1	62.1	65.7	<b>71.9</b>
	300	Husband	73.6	76.4	72.1	84.3	58.6	67.1	72
		Wife	72.9	80.7	71.4	76.4	55	70	71.3
Previous method (Black <i>et al.</i> , 2013)	Husband	78.6	72.9	72.1	84.3	60	71.4	73.2	
	Wife	77.9	84.3	74.3	80	66.4	67.1	75	

最後一層的節點數我們設為 150，如表 6，得到最後的準確率。使用的疊代次數為 20 次， $\rho = 0.1$ ， $\lambda = 0.0001$ ， $\beta = 1$ 。

**表 6. 3rd hidden unit 分析丈夫和太太對應到 6 種 code 的準確率和之前研究準確率比較**

1 <sup>st</sup> Hidden Layer	2 <sup>nd</sup> Hidden Layer	3 <sup>rd</sup> Hidden Layer	Rated Spouse	Acc (%)	Bla (%)	Pos (%)	Neg (%)	Sad (%)	Hum (%)	Avg (%)
300	200	150	Husband	80	78.6	73.6	84.3	59.3	73.6	74.9
			Wife	80	83.6	72.9	81.4	65	67.9	75
Previous method (Black <i>et al.</i> , 2013)			Husband	78.6	72.9	72.1	84.3	60	71.4	73.2
			Wife	77.9	84.3	74.3	80	66.4	67.1	75

#### 4.4 實驗結果比較

我們所使用三種不同層的稀疏自編碼器和之前的論文整體平均準確率結果如表 7。

**表 7. 整體平均正確率對於四種不同方法**

Method		Avg(%)
Previous (Black <i>et al.</i> , 2013)		74.1
SSAE	One Layer	72.2
	Two Layers	72.3
	Three Layers	75.0

由表 7 中可得知 3 層的 SSAE 較之前研究提高 0.9%。之前研究使用 40479 個特徵值來作預測，而我們使用 2940 個特徵值，理論上看來較多的特徵值相對於準確率會較高，但透過深度學習的方式，降低數據的維度，找出相對關鍵的特徵，對於準確率的提升是有幫助的，從表 7 中看來雖然訓練 1 層和 2 層準確率表現沒有比較好，在使用 3 層之後就有好的表現，此論點由此可證。

#### 5. 結論

現今存在越來越多資料庫，如何快速且準確預測資料，是近來研究的熱門議題。在這篇論文中，我們提出堆疊稀疏自編碼器改變特徵提取的方法和以為主體架構，來比較和之前研究的準確率，目的在藉由降低特徵數量，提升訊息的含量的方式，找到相對關鍵的特徵，來達到更好的準確率並減少訓練時間。最後結果也證明了利用非監督學習來訓練出新的一組特徵值，經由監督學習作分類，準確率較之前研究來的好，提出新的方法整體平均為 75% 高於舊的研究 74.1%，提升 0.9%。

本研究由於是對男女句子取出特徵值，再利用深度學習和機器學習去分析，準確率能有所提升，但是上升並不明顯還有改進空間。在未來，不管是透過改變非監督學習的演算法，或者是改變特徵擷取，只要產生出好的特徵值對資料的表示，準確率勢必會有更大的突破，把訓練好的模型套用在其他資料庫來講，也大大縮短人為預測所消耗的時間和成本。

## 參考文獻

- Andrew, G., & Gao, J. (2007). Scalable training of L 1-regularized log-linear Models. In *Proceedings of the 24th international conference on Machine learning. ACM*, 33-40.
- Black, M., Katsamanis, A., Baucom, B., Lee, C., Lammert, A., Christensen, A., Georgiou, P., & Narayanan, S. (2013). Toward automating a human behavioral coding system for married couples' interactions using speech acoustic features. *Speech Communication*, 55(1), 1-21.
- Burkhardt, F., Polzehl, T., Stegmann, J., Metze, F., & Huber, R. (2009). Detecting real life anger. In *Proc. IEEE Int'l Conf. Acous., Speech, and Signal Processing*, 4761-4764.
- Christensen, A., Atkins, D.C., Yi, J., Baucom, D.H., & George, W.H. (2004). Couple and individual adjustment for 2 years following a randomized clinical trial comparing traditional versus integrative behavioral couple therapy. *J. Consult. Clin. Psychol.*, 72, 176-191.
- Christensen, A., Jacobson, N.S., & Babcock, J.C. (1995). Integrative behavioral couple therapy. In: Jacobsen, N.S., Gurman, A.S. (Eds.), *Clinical Handbook of Marital Therapy, second ed. Guilford Press, New York*, 31-64.
- Devillers, L., & Campbell, N. (2011). Special issue of computer speech and language on affective speech in real-life interactions. *Comput. Speech Lang.*, 25, 1-3.
- Gonzaga, G.C., Campos, B., & Bradbury, T. (2007). Similarity, convergence, and relationship satisfaction in dating and married couples. *J. Personal. Soc. Psychol.*, 93, 34-48.
- Heavey, C., Gill, D., & Christensen, A. (2002). Couples interaction rating system 2 (CIRS2)., *University of California, Los Angeles. Los Angeles, CA, USA*.
- Hinton, G. (2006). Reducing the Dimensionality of Data with Neural Networks. *Science*, 313(5786), 504-507.
- Hinton, G., Deng, L., Yu, D., Dahl, G., Mohamed, A., Jaitly, N., et al. (2012). Deep Neural Networks for Acoustic Modeling in Speech Recognition: The Shared Views of Four Research Groups. *IEEE Signal Process. Mag.*, 29(6), 82-97.
- Jones, J., & Christensen, A. (1998). Couples interaction study: Social support interaction rating system. *University of California, Los Angeles. Los Angeles, CA, USA*.
- Karney, B.R., & Bradbury, T.N. (1995). The longitudinal course of marital quality and stability: A review of theory, methods, and research. *Psychol. Bull.*, 118, 3-34.

- O'Brian, M., John, R.S., Margolin, G., & Erel, O. (1994). Reliability and diagnostic efficacy of parent's reports regarding children's exposure to marital aggression. *Violence and Victims*, 9(1), 45-62.
- Obst, O. (2014). Distributed machine learning and sparse representations. *Neurocomputing*, 124, 1.
- Perwej, Y., & Chaturvedi, A. (2011). Machine recognition of Hand written Characters using neural networks. *International Journal of Computer Applications*, 14(2), 6-9.
- Rubanov, N. (2000). The layer-wise method and the backpropagation hybrid approach to learning a feedforward neural network. *IEEE Trans. Neural Netw.*, 11(2), 295-305.
- Schuller, B., Batliner, A., Seppi, D., Steidl, S., Vogt, T., Wagner, J., et al. (2007). The relevance of feature type for automatic classification of emotional user states: Low level descriptors and functionals. In *Proc. Interspeech*, Antwerp, Belgium, 2253-2256.
- Smirnov, E., Timoshenko, D., & Andrianov, S. (2014). Comparison of Regularization Methods for ImageNet Classification with Deep Convolutional Neural Networks. *AASRI Procedia*, 6, 89-94.

