

# 人工智能與金融交易概要

## AI on Financial Trading

台北大學 2018/12/5~6

希奇資本 技術長

董夢雲 博士

# 目 錄

- 一、交易策略概論
- 二、可預測性與效率市場
- 三、時間數列的預測模型
- 四、神經網路的發展
- 五、神經網路的預測應用
- 六、自然語言的交易應用
- 七、程式工具
- 八、線上交易平台的興起
- 九、結論
- 十、參考資料

# 一、交易策略概論

## (一)交易的對象

◆所有的交易都涉及時間的不確定性，亦即，對隨機變數的未來預測。

➤ 以統計的語言，我們是對隨機變數的特定統計量來預測。

◆財務中的避險，基本上是不涉及不確定性。

➤ 避險不存在超額報酬。

➤ 即便是動態避險，都涉及預測。

✓ 動態避險仍然是有高的報酬存在。

◆市場可以成為交易的統計量，大概可以分為三類。

- 期望值：最為直觀的價格買賣。
- 標準差：通過選擇權交易，單純買賣或策略應用。
- 相關性：兩項資產交易，或 option on two assets。

◆一般的 Spot 交易，都是以期望值作為交易的對象。

## (二)市場走勢與策略分類

◆如果從市場走勢的角度來看交易策略，我們可以把交易策略概分為趨勢與盤整兩大類。

### 甲、趨勢策略

◆當市場展現出明顯的上漲或下跌時，我們需要採用的策略。

- 通常又稱為動能(Momentum)策略。
  - ✓ 當市場看漲時，不外就是加碼進場，做多行情。
  - ✓ 當市場看跌時，就採取是退出市場，或積極放空行情。
- 另一相關的策略就是買入之前表現好的股票，放空之前表現差的股票。
  - ✓ 因為好的更好，壞的更壞。
  - ✓ 這也是一種動能的表現，有時也被稱之為順向策略。

## B、盤整策略

◆當市場沒有展現出明顯的上漲或下跌時，亦即，市場出現盤整的現象，我們需要採用的策略。

➤ 通常又稱為均數回復(Mean-Reversion)策略。

✓ 由於在盤整期間，我們可以想像資產價格有一個市場認知的中間價位，當目前價格太高時，投資人會採取賣出方式，以便獲利了結。

✓ 當目前價格太低時，投資人會採取買入方式，以便未來獲利。

➤ 另一相關的策略就是買入之前表現差的股票，放空之前表現好的股票，

✓ 因為好的會轉差，壞的會變好。

✓ 這是一種反轉的表現，有時也被稱之為反向(Contrarian)策略。

### (三)交易策略與對象

	趨勢測略		盤整策略
	Up	Down	
期望值	加碼進場	積極放空	放空高漲 買入低跌
標準差	買入選擇權	賣出選擇權	賣出個股 Strangle 賣出個股 Straddle
相關性	發散測略	收斂測略	賣出組合 Strangle 賣出組合 Straddle

## 二、可預測性與效率市場

### (一)效率性假說

#### ◆ 《效率市場假說》（Efficient Market Hypothesis，EMH）

➤ Eugene Fama(1970, JF)所提出，包含以下幾個要點：

- ✓ 第一，在市場上的每個人都是理性的經濟人，金融市場上每只股票所代表的各家公司都處於這些理性人的嚴格監視之下，他們每天都在進行基本分析，以公司未來的獲利性來評價公司的股票價格，把未來價值折算成今天的現值，並謹慎地在風險與收益之間進行權衡取捨。
- ✓ 第二， 股票的價格反映了這些理性人的供求的平衡，想買的人正好等於想賣的人，即，認為股價被高估的人與認為股價被低估的人正好相等，假如有人發現這兩者不等，即存在套利的可能性的話，他們立即會用買進或賣出股票的辦法使股價迅速變動到能夠使二者相等為止。
- ✓ 第三， 股票的價格也能充分反映該資產的所有可獲得的訊息，即"資訊有效"，當資訊變動時，股票的價格就一定會隨之變動。一個利好消息或利空消息剛剛傳出時，股票的價格就開始異動，當它已經路人皆知時，股票的價格也已經漲或跌到適當的價位了。



## ◆效率市場假說的三種形態

### ➤ 弱式效率市場假說(Weak Form Efficiency)

- ✓ 市場價格已充分反應出所有過去的證券價格資訊，包括股票的成交價、成交量、賣空金額、融資金額等。
- ✓ 推論一：如果弱式效率市場假說成立，則股票價格的技術分析失去作用，基本分析還可能幫助投資者獲得超額利潤。

### ➤ 半強式效率市場假說(Semi-Strong Form Efficiency)

- ✓ 價格已充分反應出所有已公開的有關公司營運前景的資訊，包括成交價、成交量、盈利資料、盈利預測值，公司管理狀況及其它公開披露的財務資訊等。
- ✓ 推論二：如果半強式效率假說成立，則在市場中利用技術分析和基本分析都失去作用，內幕消息可能獲得超額利潤。

### ➤ 強式效率市場假說(Strong Form Efficiency)

- ✓ 價格已充分地反應了所有關於公司營運的資訊，包括已公開的或內部未公開的資訊。
- ✓ 推論三：在強式效率市場中，沒有任何方法能幫助投資者獲得超額利潤，即使有內幕消息者也是一樣。

◆效率市場的資產價格，應呈現出隨機漫步(Random Walk)的走勢。

- Fama(1965, JB), Behavior of stock-market prices, 提出了統計證據。
- Samuelson(1965, Industrial Management Rev.), Proof that properly anticipated prices fluctuate randomly, 給予數學證明。

◆如果市場的資訊，真的被即時且充分反映，那不存在交易獲利的空間。

- 交易不全是訊息驅動的。
- 資訊也不是無污染的。
- 《即時性》是有成本的。

## (二)可預測性

◆股價反應投資人對市場的預期，當訊息產生時，投資人的預期就發生改變。

- 如果這個訊息對公司的營運具體產生的影響，亦即，對公司未來獲利產生實質的變化。
  - ✓ 則股票價格變會出現明顯的變化，市場在解讀這一效果時，股價就會產生趨勢現象。
- 然而，如果沒有足以對公司營運產生具體影響的訊息出現時，市場上交易的訊息往往是一些雜訊，或是基於投資人流動性需求的買賣。
  - ✓ 此時價格的盤整便是一種常態。

◆對於交易的最大考驗就是判斷，究竟市場的價格變動，只是表面的漣漪，還是海嘯的先聲。

- 沒有人能買在最低點，賣在最高點。
- 如果我們能分辨市場所處的區段，反覆的操作適當的策略，那也能夠創造相當的利潤。
- 使用各類的演算法來協助判斷市場所處的位置，人工智慧演算法便是一個可能提供的來源。

◆價格資料本身就是一個時間數列，在學術界這是一個相當重要的研究領域，畢竟預測在商業活動是一項非常重要的事項。

➤ 我們可也先藉由時間數列的模型，來研究資產價格的行為。

# 三、時間數列的預測模型

## (一)資料的可預測性

◆一個  $N$  期的時間數列， $Y_N$ ，是由  $N$  個觀察點資料所構成，表示如下，

$$Y_N = \{y_{N-1}, y_{N-2}, y_{N-3}, \dots, y_2, y_1, y_0\}$$

$$Y_N = \{y_t : 0 \leq t < N, t \in \mathbb{Z}^+\}$$

➤ 如果以資產價格形成一時間數列，我們第一個問題是，甚麼樣的數列具有可預測性。

✓ 一個隨機漫步(Random Walk)所形成的時間數列是不具備可預測性的。

◆我們可以使用布朗運動來形成一個隨機漫步的時間數列如下，

$$dz_t = z_t - z_{t-1} = \sigma \cdot dB_t, \quad z_0 = 0, \quad dB_t = N(0, \sqrt{dt})$$

➤ 令時點  $t$  與時點  $t-1$  的間格為  $dt$ ，時間數列的起點為零。 $\sigma$  表一常數。

➤  $dB_t = N(0, \sqrt{dt})$ ，表常態分配隨機亂數，其平均數為零，標準差為時間間隔的平方根。

◆事實上，我們希望資產價格的時間數列不是隨機漫步，因為我們無法預測之。

➤ 如果資產價格的時間數列呈現盤整或是趨勢現象，我們才有可能由其中獲利。

## A、Hurst Exponent測試

◆我們可以進行 Hurst Exponent 測試，來判斷時間數列所處呈現的狀態。

➤ 我們創造三個時間數列，分別為隨機漫步、盤整、趨勢現象。

$$x_t = \sum_1^t \varepsilon_t + 1000, \varepsilon_t \sim N(0,1)$$

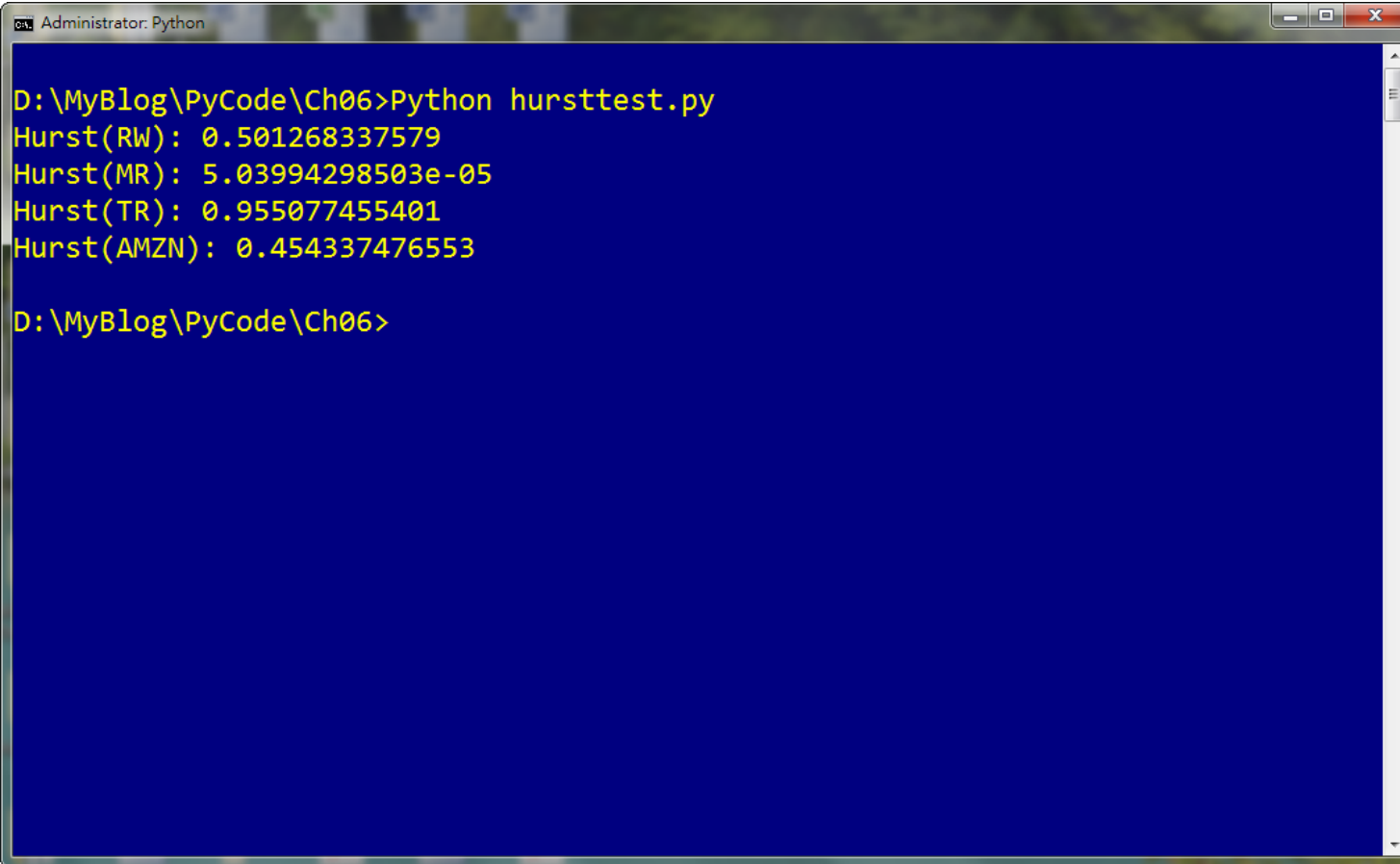
$$y_t = \varepsilon_t + 1000$$

$$z_t = \sum_1^t (\varepsilon_t + 1) + 1000$$

➤ 我們也從 Web 上下載 Amazon 的股價做一對比。

✓ 可以看到隨機漫步數列  $H=0.50126$ 、盤整數列  $H=0.0001$ 、趨勢數列  $H=0.95507$ ，Amazon 股價數列  $H=0.45434$ 。

✓ 長期而言，Amazon 的股價有接近盤整的現象，但隨機性也不低。



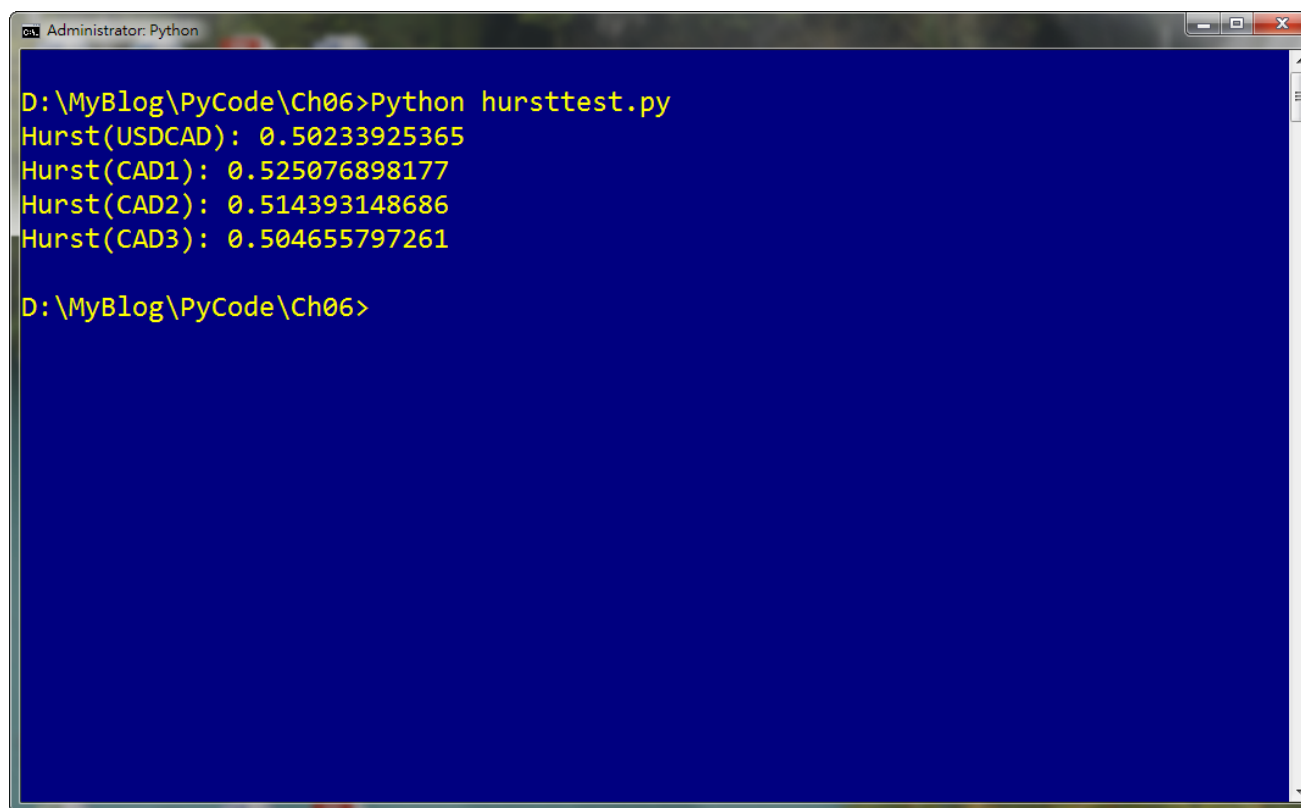
```
Administrator: Python
D:\MyBlog\PyCode\Ch06>Python hursttest.py
Hurst(RW): 0.501268337579
Hurst(MR): 5.03994298503e-05
Hurst(TR): 0.955077455401
Hurst(AMZN): 0.454337476553
D:\MyBlog\PyCode\Ch06>
```



◆我們由 FED 的資料庫下載 CADUSD 的歷史日資料價格,2001/1/1~2016/5/30,進行 Hurst Exponential 測試。

➤ 分別以全部資料,最後 250 天,最後 500 天,最後 1000 天,進行測試,結果如下。

✓ 可以看出來,CADUSD 的時間序列很接近隨機漫步的現象。



```
D:\MyBlog\PyCode\Ch06>Python hursttest.py
Hurst(USDCAD): 0.50233925365
Hurst(CAD1): 0.525076898177
Hurst(CAD2): 0.514393148686
Hurst(CAD3): 0.504655797261

D:\MyBlog\PyCode\Ch06>
```

## (二)穩態資料的處理

◆當我們說一個時間數列是穩態的(Stationary)，是指此時間數列的分配是不隨時變的。

- 亦即，如果我們平移  $Y_0 = \{y_0, y_1, \dots, y_{t-1}, y_t\}$  一段期間， $Y_\alpha = \{y_\alpha, y_{\alpha+1}, y_{\alpha+2}, \dots, y_{\alpha+t-1}, y_{\alpha+t}\}$ ，此數列與原數列有相同的統計分配。

$$Y_0 = \{y_0, y_1, \dots, y_{t-1}, y_t\}$$

$$Y_\alpha = \{y_\alpha, y_{\alpha+1}, y_{\alpha+2}, \dots, y_{\alpha+t-1}, y_{\alpha+t}\}$$

◆穩態的時間數列是我們進行統計分析的主要對象。

- 如果時間數列不具穩態特性，則我們根據歷史資料得到的結果，自然無法用於未來的預測使用。
- ✓ 因為時間數列的統計特性已經改變。
  - ✓ 另外，在進行迴歸分析時，如果隨機變數的歷史資料不具穩態特性，也可能形成假性迴歸的現象。

## ◆穩態的時間數列與交易策略也有重要的關聯。

- 當市場在盤整時，價格數列會具有均數回復的現象。此時，價格數列也就會呈現出穩態的現象。
  - ✓ 我們也可以這麼說，當價格數列呈現出穩態的現象時，均數回復的交易策略才有施展的空間。
- 一個均數回復的價格數列基本上其價格的變動，可以表示為目前價格與平均數之間差距的負比例關係。
  - ✓ 當目前價格高於平均數，則下一期的價格會下降；當目前價格低於平均數，則下一期的價格會上升。

## ◆數學上，我們可以使用 Ornstein-Uhlenbeck 過程，來表示一個均數回復的時間數列，

$$dx_t = \theta(\mu - x_t)dt + \sigma \cdot dB_t$$

- 其中， $\mu$  表時間數列的平均數， $\theta$  表回復均數的速率。
- $dB_t = N(0, \sqrt{dt})$ ，表常態分配隨機亂數，其平均數為零，標準差為時間間隔的平方根。
- $\sigma$  表隨機程序的波動性。

## 甲、ADF測試

◆針對時間數列本身是否具有均數回復的特性，我們可以使用 Augmented Dickey-Fuller (ADF)測試。

➤ 其論點是，一個具有均數回復特性的價格數列，可以用下面的線性模型來描述其價格變動，

$$\Delta y_t = \mu + \beta t + \lambda y_{t-1} + \alpha_1 \Delta y_{t-1} + \alpha_2 \Delta y_{t-2} + \dots + \alpha_k \Delta y_{t-k} + \varepsilon_t$$

➤ ADF 測試是針對  $\lambda$  是否為 0 所進行的。如果  $\lambda$  不為 0，則下一期的變動量，與本期數值有關，則時間數列就不是隨機漫步。

✓ 如果  $\lambda$  為 0，則下一期的變動量，與本期數值無關，則時間數列就是隨機漫步。

➤ 檢定量是將  $\lambda$  的估計值，除以  $\lambda$  的標準差。

✓ 我們預期  $\lambda$  為負值。因此檢定量負值愈大且愈顯著，帶表隨機漫步性愈低。

✓ 這也是我們希望看到的結果，表示時間數列具有較高的可測性。

### (三)共整合的處理

◆實務上，大部分的可交易資產價格都不具備均數回復的性質。

- 如果股價的行為如幾何布朗運動的模式，那麼均數回復的策略對其也就無用武之地了。
  - ✓ 匯率與利率長期有均數回復的性質。
- 雖然單一資產價格的時間數列雖然不具有穩態的特性，但是我們卻可以使用兩個以上的資產，創造出具有穩態特性的組合資產價格時間數列。
  - ✓ 這些資產我們稱其具有共整合的特性。

### 甲、迴歸分析

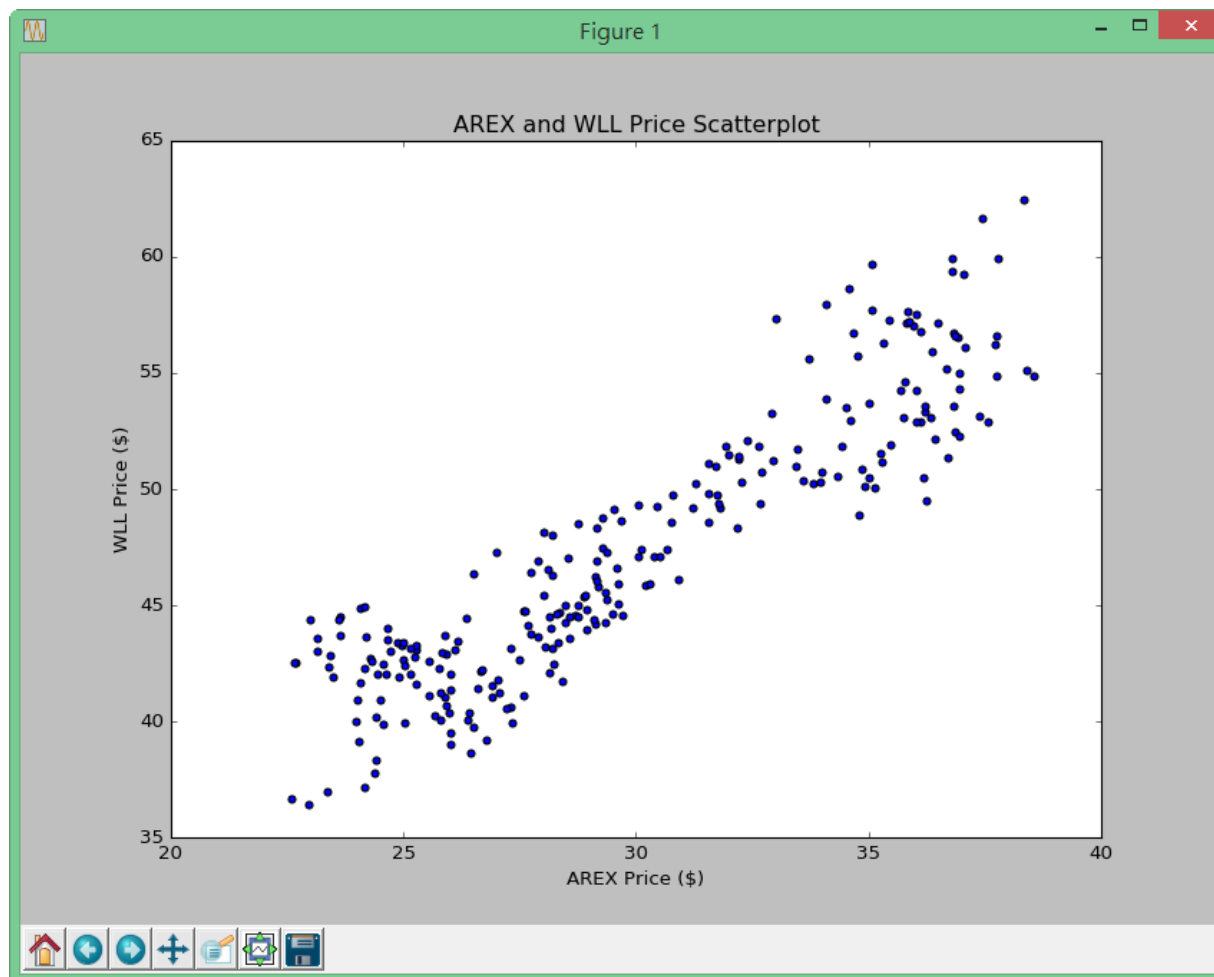
◆讓我們考慮兩個能源類股的股票，其一為股票代碼 AREX 的 Approach Resources 公司，另一為股票代碼 WLL 的 Whiting Petroleum 公司。

- 兩者面對相似的市場情況，因此很可能具有穩態的配對關係。

◆我們由網路上下載 AREX 與 WLL 這兩支股票 2012/1/1~2013/1/1 的股價，將其時間數列並列於下圖。直觀上，這兩個數列有同步變動的趨勢。



◆下圖是以 AREX 與 WLL 這兩支股票的股價，分列於 X-Y 軸的圖示。直觀上，一條正斜率的帶狀趨勢也可明顯看出，兩者的正相關呼之欲出。



◆配對交易通常針對兩個股票價格，使用線性迴歸模型。

$$y_t = b_0 + b_1x_t + \varepsilon_t$$

- 我們以 WLL 股價為因變數，AREX 股價為自變數，使用 statsmodels 套件進行分析，執行結果摘要如下。

```
Python Win64 - ipython
In [16]: run coint_data.py

=====
                        OLS Regression Results
=====
Dep. Variable:          WLL      R-squared:                0.832
Model:                  OLS      Adj. R-squared:           0.831
Method:                 Least Squares      F-statistic:         1226.
Date:                   Fri, 01 Jul 2016    Prob (F-statistic):    5.84e-98
Time:                   09:48:39            Log-Likelihood:       -569.50
No. Observations:       250              AIC:                  1143.
Df Residuals:           248              BIC:                  1150.
Df Model:                1
Covariance Type:        nonrobust
=====

```

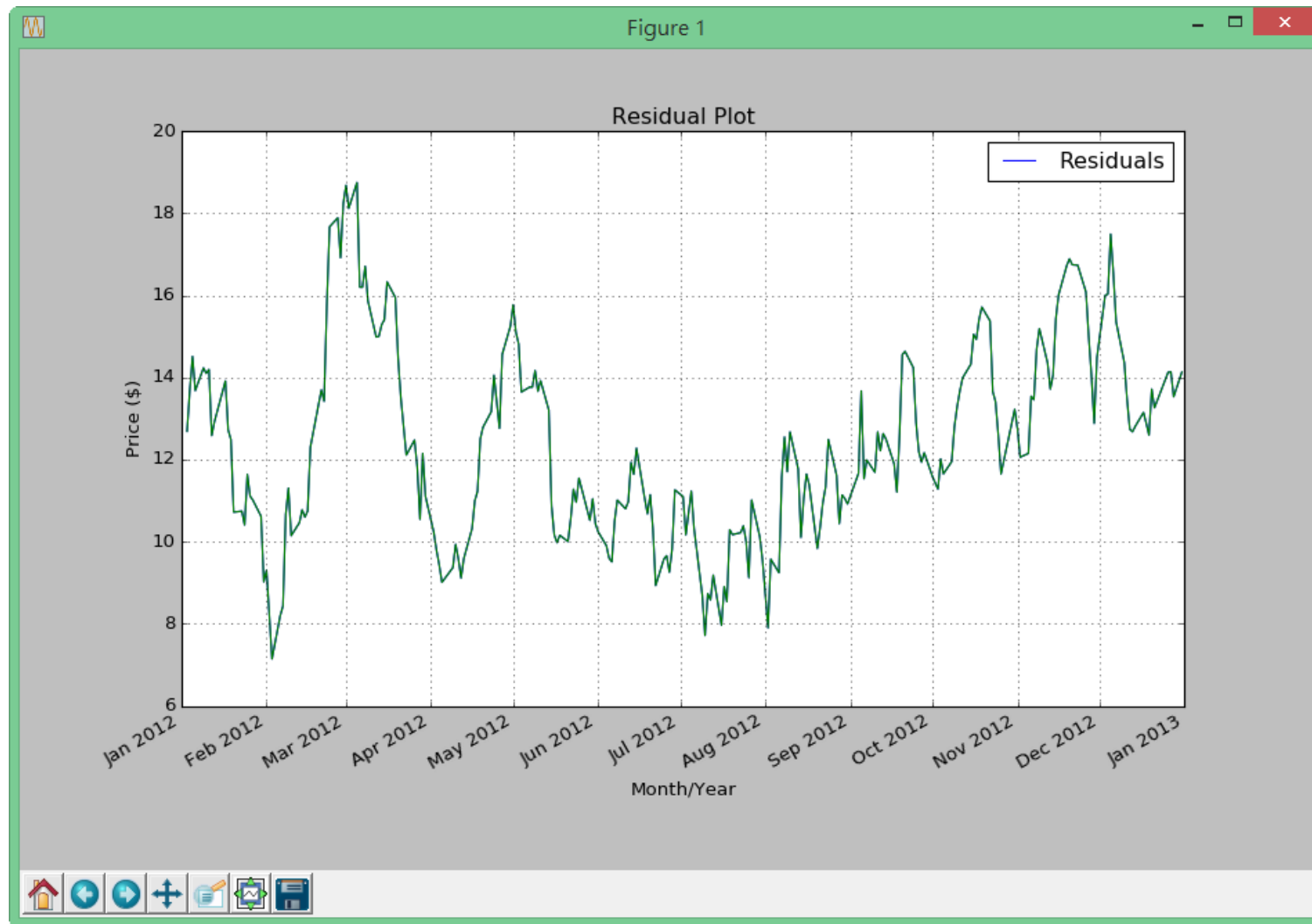
	coef	std err	t	P> t	[95.0% Conf. Int.]
const	12.3599	1.011	12.228	0.000	10.369 14.351
AREX	1.1691	0.033	35.013	0.000	1.103 1.235

```
=====
Omnibus:                 7.364      Durbin-Watson:           0.141
Prob(Omnibus):            0.025      Jarque-Bera (JB):        7.191
Skew:                     0.373      Prob(JB):                0.0274
Kurtosis:                 2.634      Cond. No.                 204.
=====

Warnings:
[1] Standard Errors assume that the covariance matrix of the errors is correctly
specified.
```



- 迴歸殘差圖示如下，似乎呈現穩態的現象。



## 乙、共整合ADF測試

◆單獨執行 ADF 或 Hurst Exponent 測試，只能針對某一特定迴歸 Beta 來檢定。以組成配對交易的 Beta，判定是否穩態。

➤ 共整合的 ADF 測試可以檢測，由兩個時間數列根據最適避險比率形成的線性組合，是否具有穩態的性質。

➤ 我們使用 statsmodels 套件中時間數列的分析工具，可以得到下面的結果。

✓ 檢定的 t 值為-2.9607，對應的 5%的關鍵值為-2.8732。

```
In [18]: run coint_cadf.py
```

```
(-2.9607014636949667, 0.038730957656696367, 0L, 249L,  
{ '1%': -3.4568881317725864, '10%': -2.5729936189738876, '5%': -2.8732185133016057 },  
601.96850620487271)
```

➤ 虛無假設為沒有共整合的性質，因此在 5%的顯著水準下我們拒絕此假設。

✓ AREX 與 WLL 具有共整合的性質。

## (四)ARIMA

◆在傳統的迴歸模型中，我們會假設殘差項在時間上的共變異為零。亦即，我們假設殘差項在時間上是不相關的。

➤ 如果此一前題不成立，則他們便是“自我相關的”或是“序列相關的”。

◆一個變量  $y_t$  的落後期可以表示為  $y_{t-1}$ ，我們以一階差分來表示此變量與其落後一期之間的變動量。

$$\Delta y_t = y_t - y_{t-1}$$

➤ 如果模型存在自我相關的現象，就會導致迴歸的估計不是 BLUE。

✓ 估計結果具備不偏和一致性，但不具備有效性和漸進有效性。

✓ 變異數的估計有偏誤，相關係數絕對值愈大，偏誤愈大，導致自變數的顯著性檢定都不正確。

- ◆模型存在自我相關現象的一個可能原因，可能是由於錯誤設定變量的動態過程。一般的迴歸模型，本質上都是靜態的。

$$y_t = \beta_0 + \beta_1 x_{1t} + \beta_2 x_{2t} + \beta_3 x_{3t} + u_t$$

- 這些模型僅考慮變量之間同時期的關係。

- ◆如果我們假定因變數受到本身前期的影響，或是其他自變數前期的影響，則我們產生的模型變成為一個動態的模型。

$$y_t = \beta_0 + \beta_1 x_{1t} + \beta_2 x_{2t} + \beta_3 x_{3t} + \gamma_0 y_{t-1} + \gamma_1 x_{1t-1} + \gamma_2 x_{2t-1} + \gamma_3 x_{3t-1} + \dots + u_t$$

- 只包含自變數前期變量的模型稱之為分布落後模型(Distributed Lag Model)，
- 包含自變數前期變量與因變數前期變量的模型稱之為自我迴歸分布落後(Autoregressive Distributed Lag Model, ADL)模型。(Chris Brooks, P.151)

◆金融市場中的模型需要放入落後項的可能原因為，

- 因變量的慣性作用：正相關，Trend。
- 過度反應：負相關，Mean Reversion。

## 四、神經網路的發展

### (一)神經網路的處理

#### 甲、神經元

##### ◆神經元生理結構

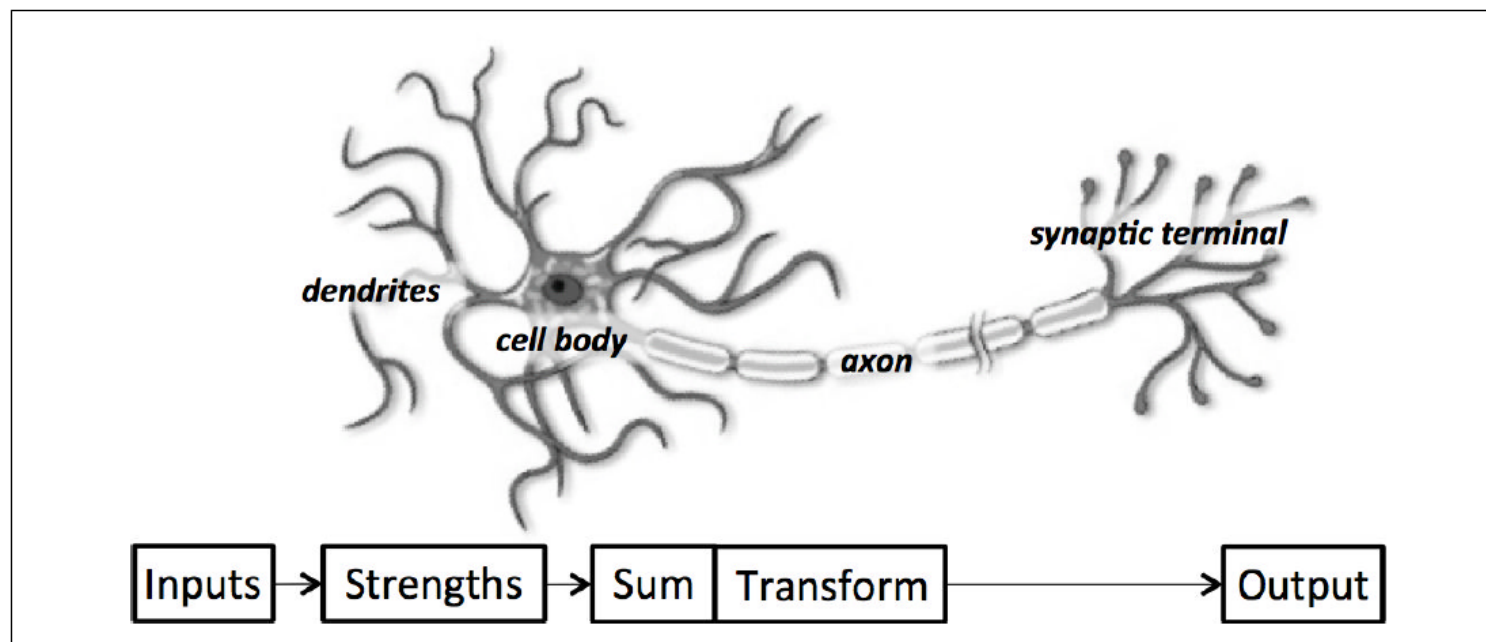


Figure 1-6. A functional description of a biological neuron's structure

## ◆人工神經網路數學型式

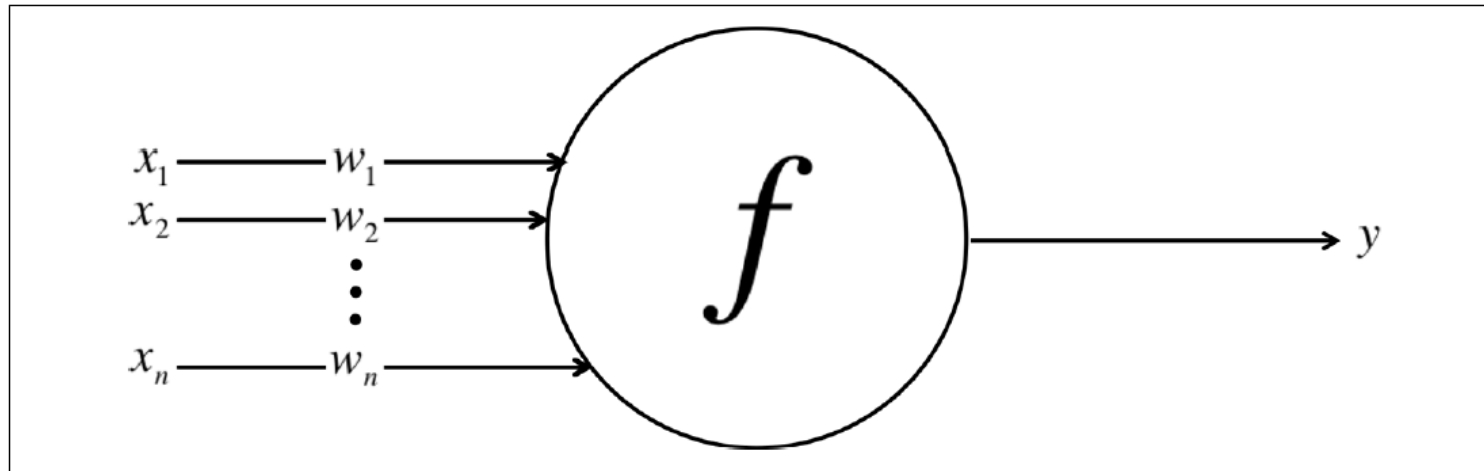


Figure 1-7. Schematic for a neuron in an artificial neural net

◆ 期末考試感知器表示

$$h(\mathbf{x}, \boldsymbol{\theta}) = \begin{cases} -1, & 3x_1 + 4x_2 - 24 < 0 \\ 1, & 3x_1 + 4x_2 - 24 \geq 0 \end{cases}$$

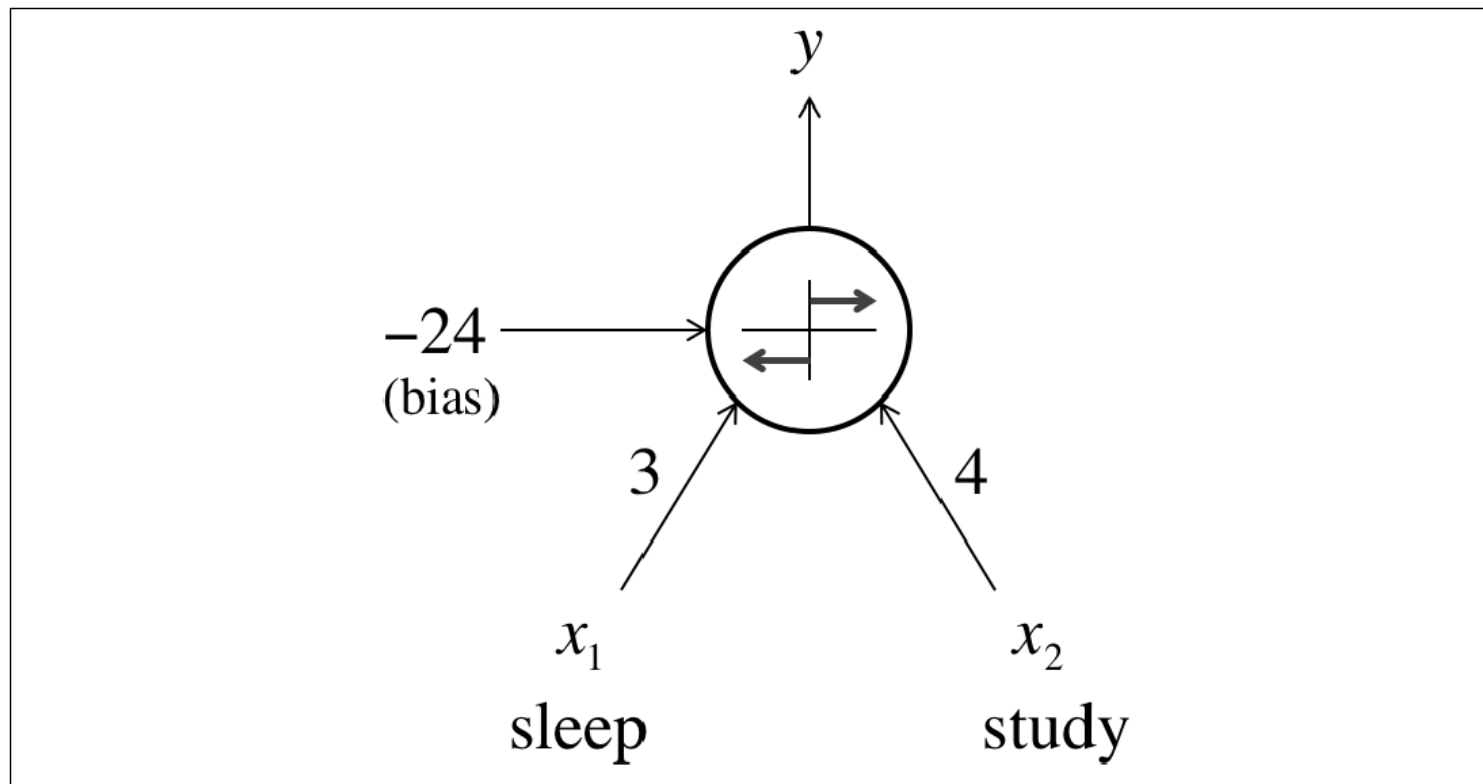


Figure 1-8. Expressing our exam performance perceptron as a neuron



## 乙、FFN

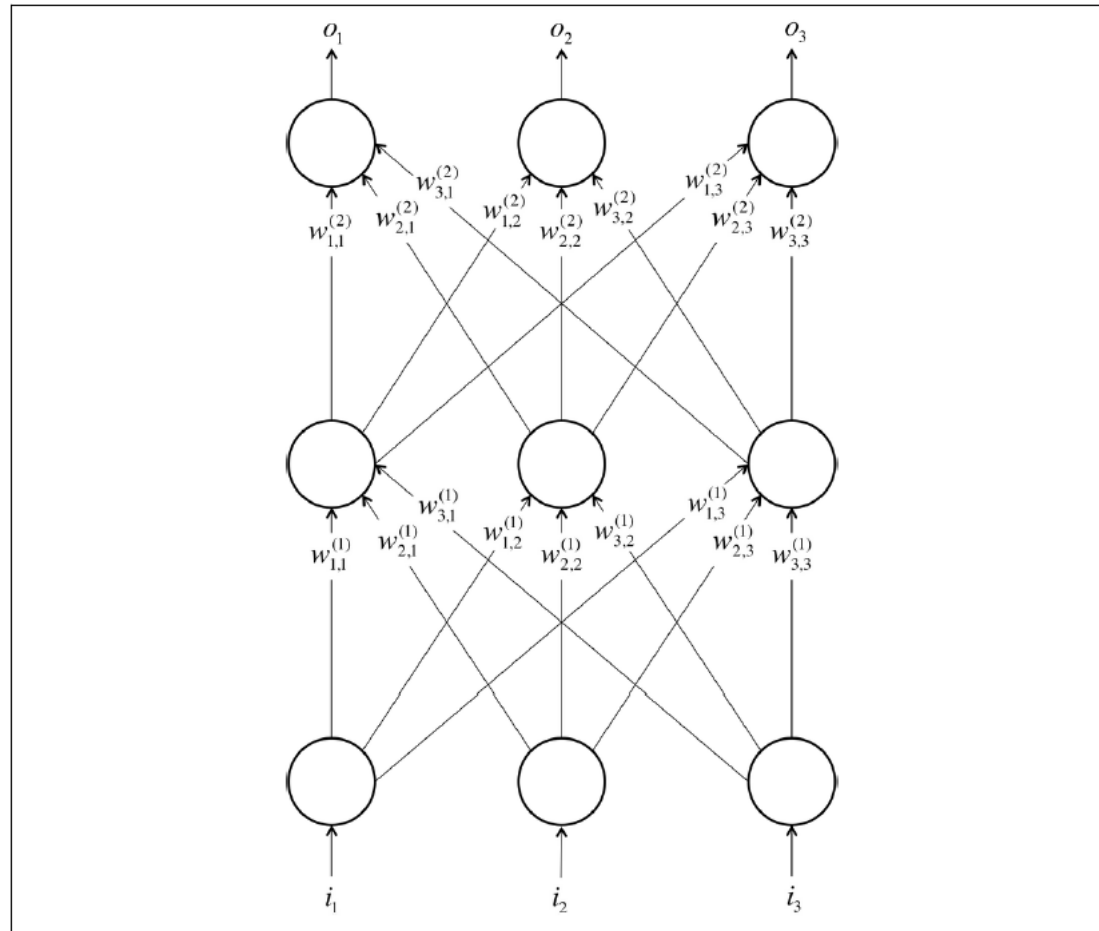


Figure 1-9. A simple example of a feed-forward neural network with three layers (input, one hidden, and output) and three neurons per layer

### 丙、閾値(Thresh-Hold)

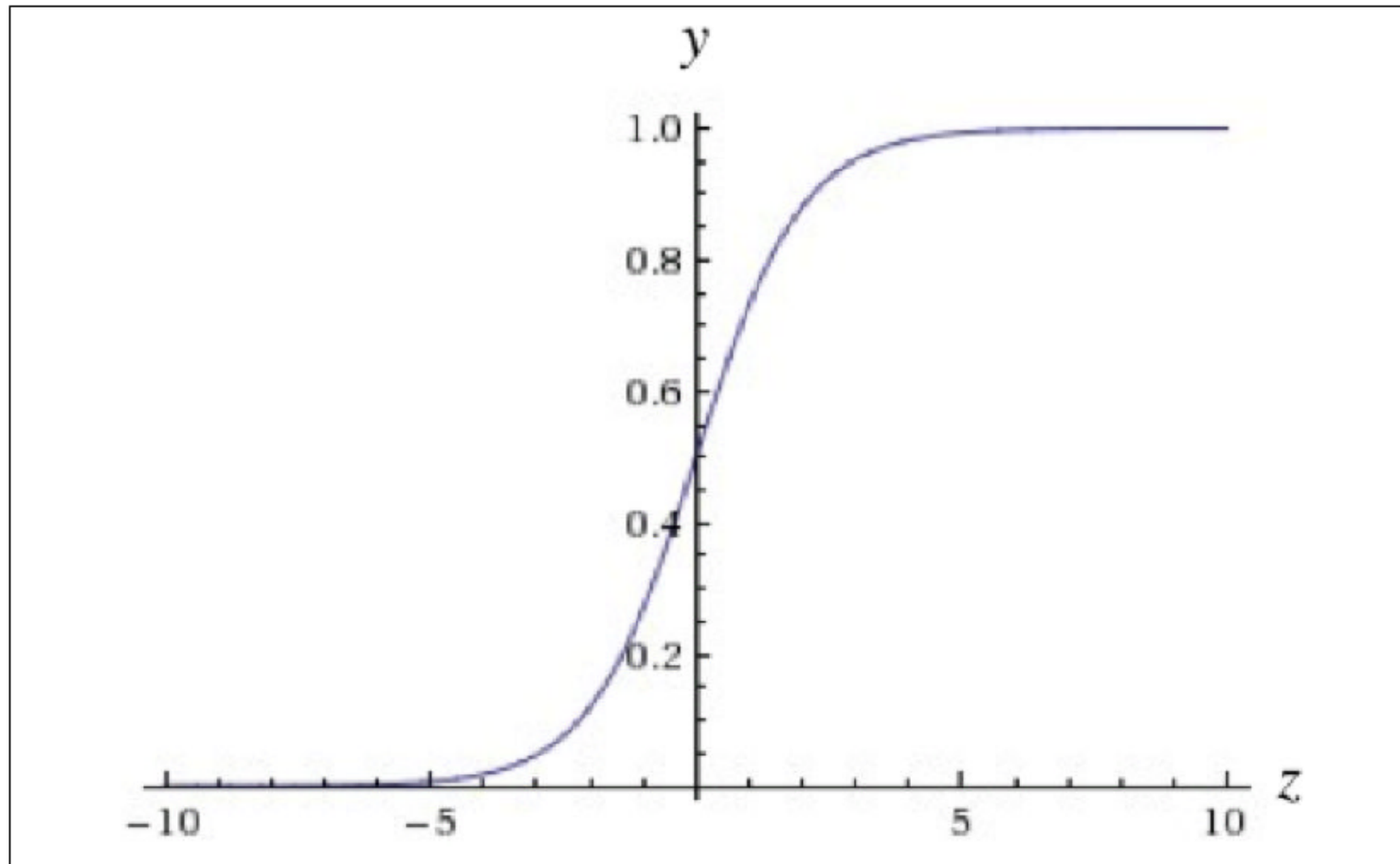


Figure 1-11. The output of a sigmoid neuron as  $z$  varies

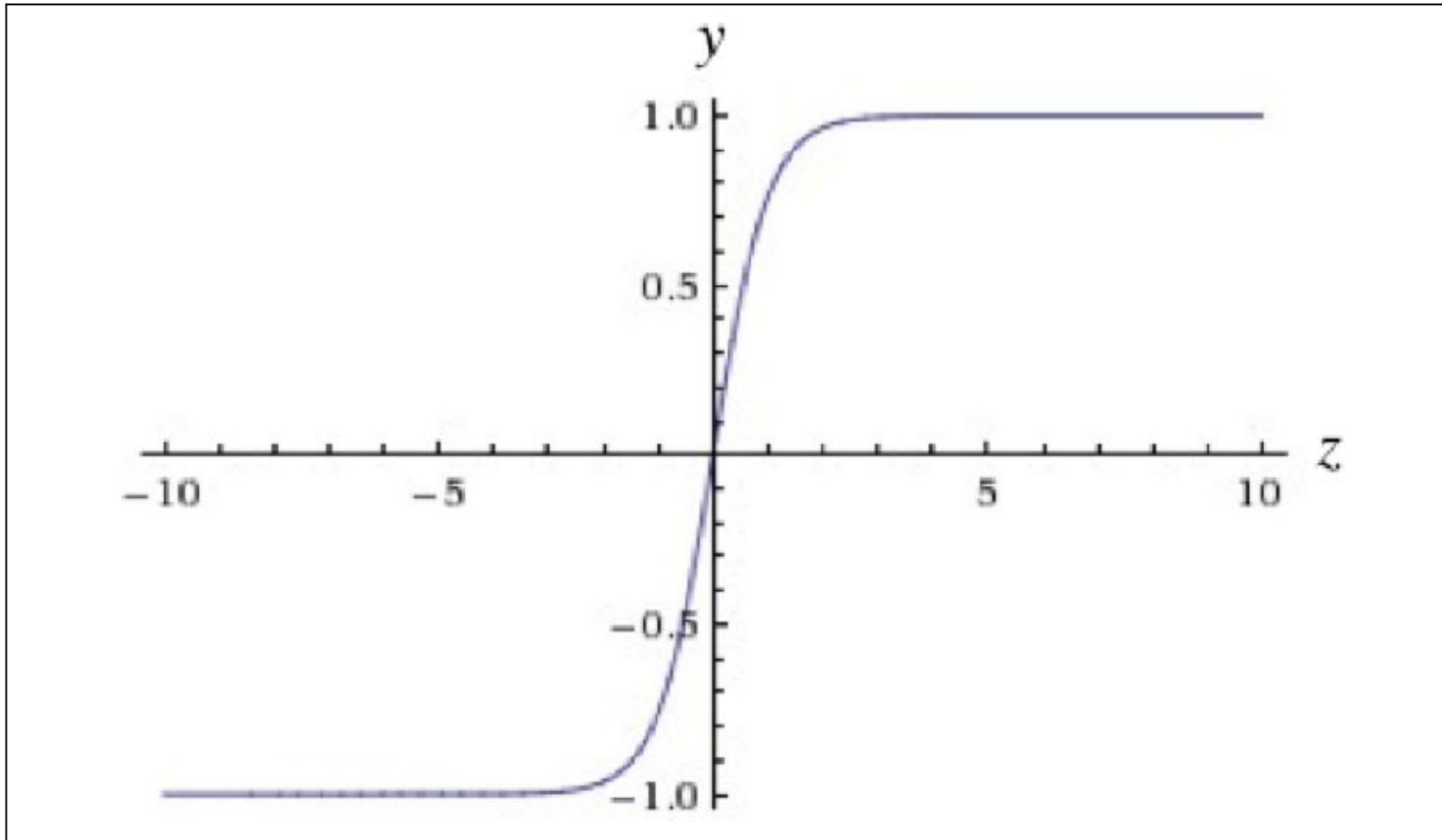


Figure 1-12. The output of a tanh neuron as  $z$  varies

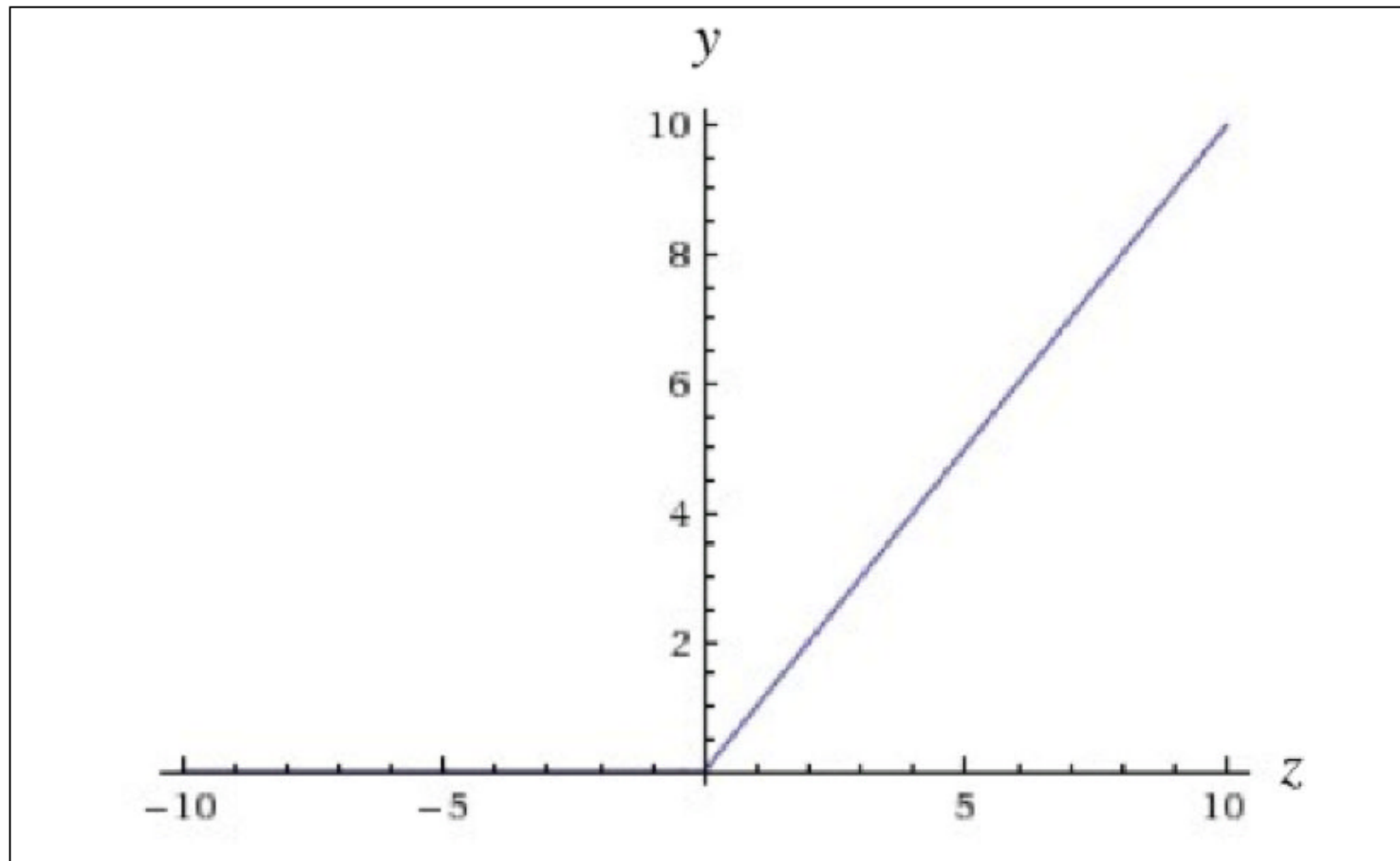


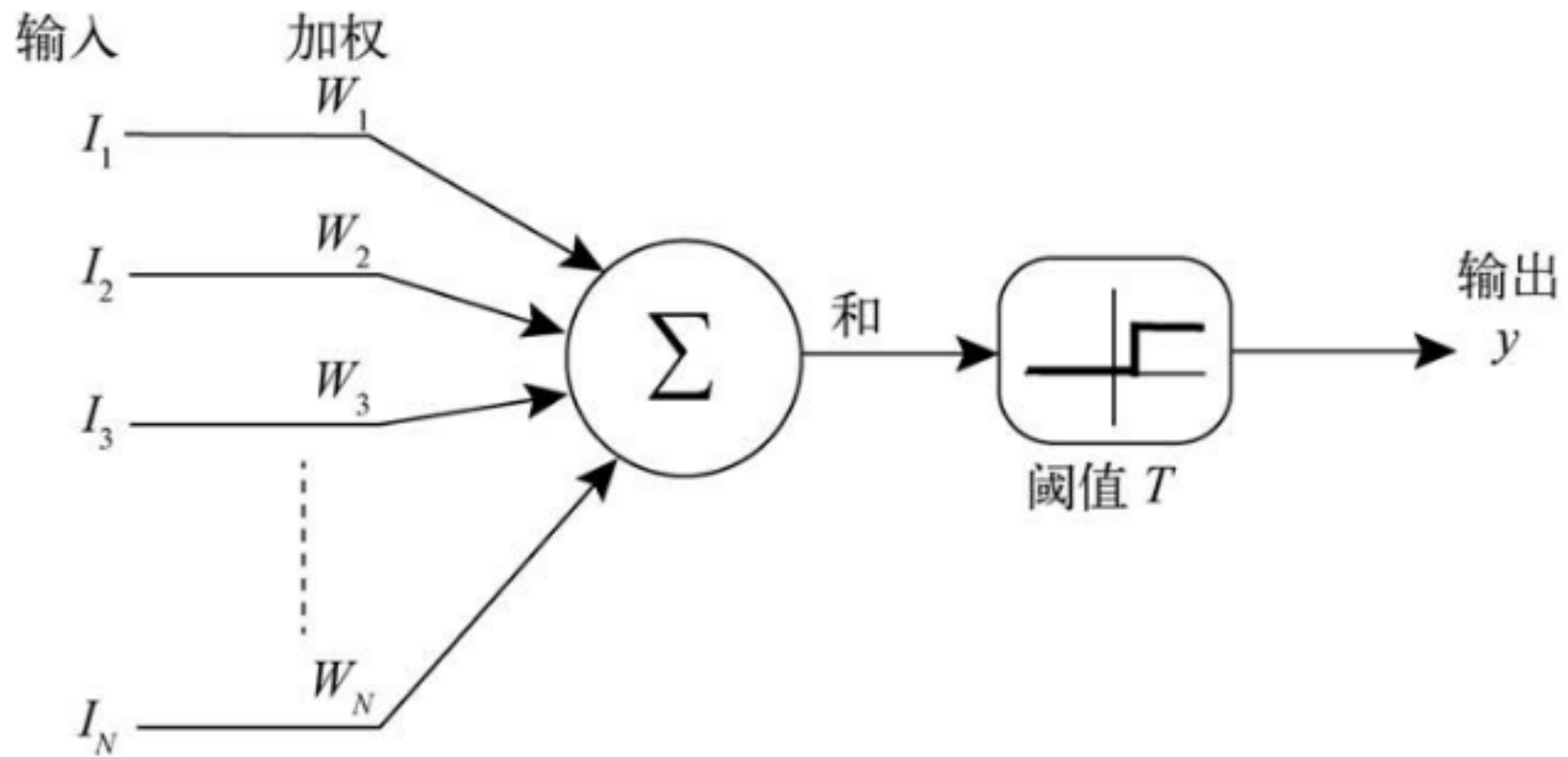
Figure 1-13. The output of a ReLU neuron as  $z$  varies

## (二)神經網路的原型

◆最早的神經網路研究可以追溯到 1943 年電腦發明之前。當時，沃倫·麥卡洛克（Warren McCulloch）和沃爾特·匹茲（Walter Pitts）二人提出了一個單個神經元的計算模型，如圖 1-7 所示。

- 在這個模型中，左邊的  $I_1, I_2, \dots, I_N$  為輸入單元，可以從其他神經元接受輸出，然後將這些信號經過加權（ $W_1, W_2, \dots, W_N$ ）傳遞給當前的神經元並完成匯總。
- 如果匯總的輸入資訊強度超過了一定的閾值（ $T$ ），則該神經元就會發放一個信號  $y$  給其他神經元或者直接輸出到外界。
- 該模型後來被稱為麥卡洛克-匹茲模型，可以說它是第一個真實神經元細胞的模型。

圖 1-7 麥卡洛克和匹茲的神經元模型



◆ 1957 年，弗蘭克·羅森布拉特（Frank Rosenblatt）對麥卡洛克-匹茲模型進行了擴充，即在麥卡洛克-匹茲神經元上加入了學習演算法，擴充的模型有一個響亮的名字：感知機 (Perception)。

➤ 感知機可以根據模型的輸出  $y$  與我們希望模型的輸出  $y^*$  之間的誤差，調整權重  $W_1, W_2, \dots, W_N$  來完成學習。

- ✓ 我們可以形象地把感知機模型理解為一個裝滿了大大小小水龍頭（ $W_1, W_2, \dots, W_N$ ）的水管網路，學習演算法可以調節這些水龍頭來控制最終輸出的水流，並讓它達到我們想要的流量，這就是學習的過程。
- ✓ 感知機就好像一個可以學習的小孩，無論什麼問題，只要明確了我們想要的輸入和輸出之間的關係，都能通過學習得以解決。

◆1969 年，人工智慧界的權威人士馬文・閔斯基給連接學派帶來了致命一擊。

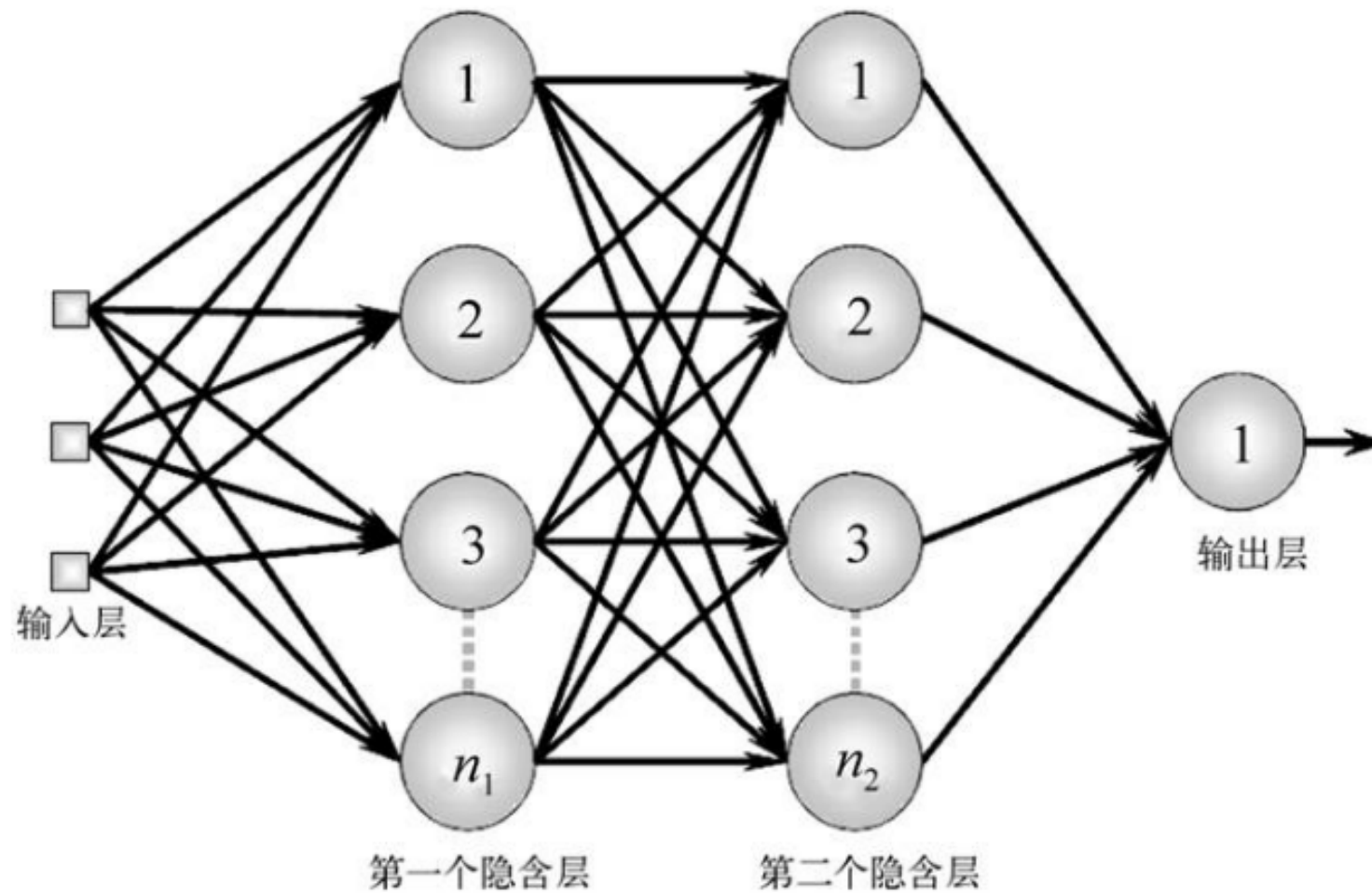
- 他通過理論分析指出，感知機並不像它的創立者羅森布拉特宣稱的那樣可以學習任何問題。
  - ✓ 連一個最簡單的問題：判斷一個兩位的二進位數字是否僅包含 0 或者 1 (即所謂的 XOR 問題) 都無法完成。
- 這一打擊是致命的，本來就不是很熱的神經網路研究差點就被閔斯基這一棒子打死了。



◆1974 年，人工智慧連接學派的救世主傑夫·辛頓（Geoffrey Hinton）終於出現了。

- 他曾至少兩次挽回連接學派的敗局，1974 年是第一次。
- 辛頓的出發點很簡單——“多則不同”。
- ✓ 只要把多個感知機連接成一個分層的網路，那麼，它就可以圓滿地解決閼斯基的問題。
- 如圖 1-8 所示，多個感知機連接成為一個四層的網路，最左面為輸入層，最右面為輸出層，中間的那些神經元位於隱含層，右側的神經元接受左側神經元的輸出。

圖 1-8 多層感知機



◆那麼多個神經元，可能有幾百甚至上千個參數需要調節，我們如何對這樣複雜的網路進行訓練呢？

- 辛頓等人發現，採用幾年前亞瑟·布賴森（Arthur Bryson）等人提出來的反向傳播演算法（Back Propagation Algorithm，簡稱 BP 演算法）就可以有效解決多層網路的訓練問題。

◆以水流管道為例來說明，當網路執行決策的時候，水從左側的輸入節點往右流，直到輸出節點將水吐出。

- 而在訓練階段，我們則需要從右往左來一層層地調節各個水龍頭，要使水流量達到要求，我們只要讓每一層的調節只對它右面一層的節點負責就可以了，這就是反向傳播演算法。
- 事實證明，多層神經網路裝備上反向傳播演算法之後，可以解決很多複雜的識別和預測等問題。

### (三)深度學習

◆21 世紀的第二個十年，最惹人注目的人工智慧研究，要數深度學習（Deep Learning）。

- 2011 年，谷歌 X 實驗室的研究人員從 YouTube 視頻中抽取出 1000 萬張靜態圖片，把它餵給“谷歌大腦”，一個採用了深度學習技術的大型神經網路模型，在這些圖片中尋找重複出現的模式。
  - ✓ 三天後，這台超級“大腦”在沒有人類的幫助下，居然自己從這些圖片中發現了“貓”。
- 2012 年 11 月，微軟在中國的一次活動中，展示了他們新研製的一個全自動的同聲翻譯系統——採用了深度學習技術的計算系統。
  - ✓ 演講者用英文演講，這台機器能即時地完成語音辨識、機器翻譯和中文的語音合成，也就是利用深度學習完成了同聲傳譯。
- 2013 年 1 月，百度公司成立了百度研究院，其中，深度學習研究所是該研究院旗下的第一個研究所。
- .....

◆2017 年 5 月 23 至 27 日在烏鎮圍棋峰會上，最新的強化版 AlphaGo，和世界第一棋士柯潔比試，取得 3 比零全勝的戰績。

➤ 此次 AlphaGo 利用 Google TPU 執行，加上快速進化的深度學習法。

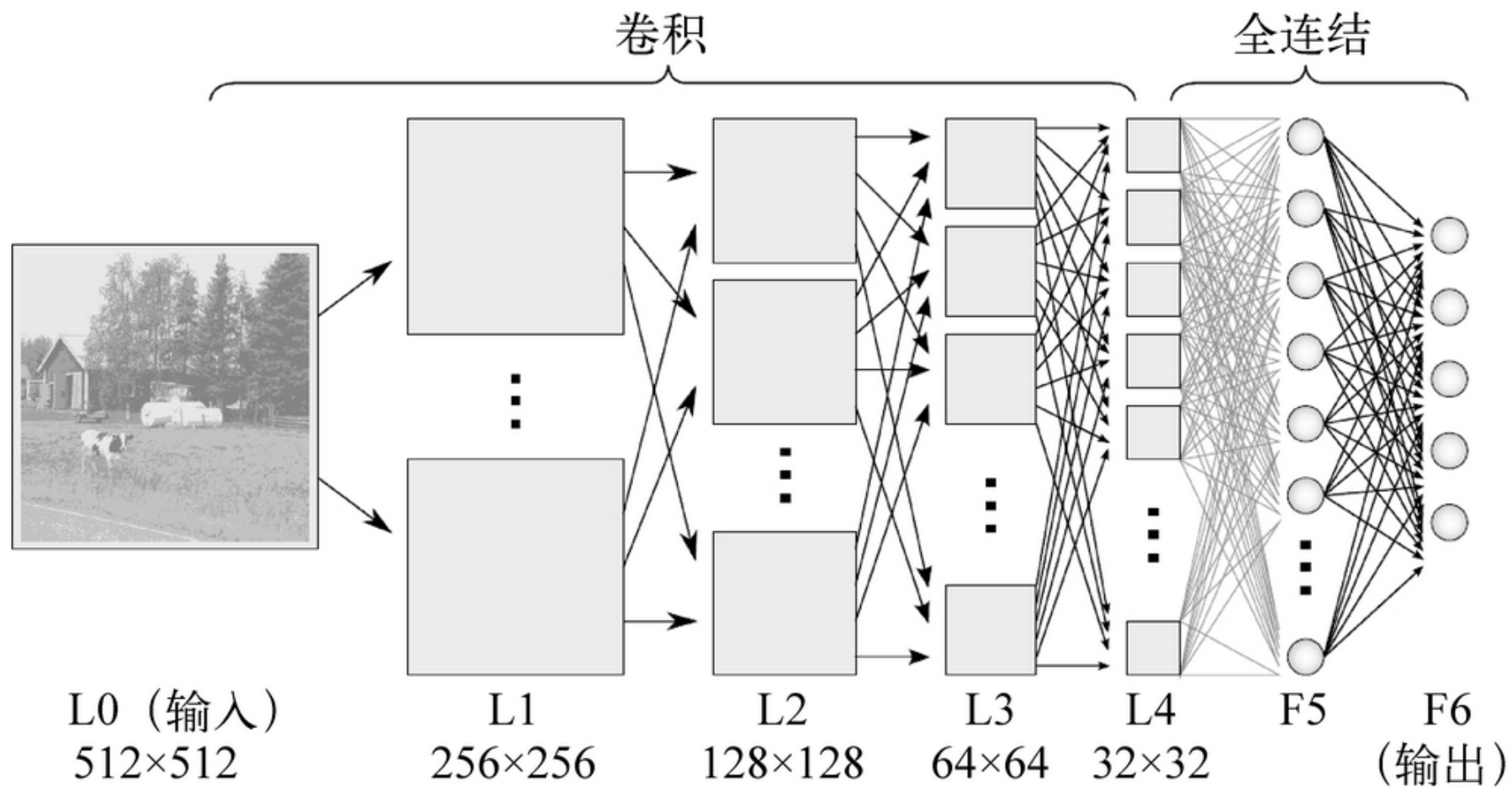
✓ 在與柯潔的比賽結束後，中國圍棋協會授予 AlphaGo 職業圍棋九段的稱號。

◆全球頂尖的電腦、互聯網公司都不約而同地對深度學習表現出了極大的興趣。深度學習為何？

➤ 深度學習仍然是一種神經網路模型，只不過這種神經網路具備了更多層次的隱含層節點，同時配備了更先進的學習技術，如圖 5 所示。

◆當我們將超大規模的訓練資料餵給深度學習模型的時候，這些具備深層次結構的神經網路仿佛搖身一變，成為了擁有感知和學習能力的大腦，表現出了遠遠好於傳統神經網路的學習和泛化的能力。

圖 5 一個深度神經網路模型



## ◆深度學習神經網路其實早在 20 世紀 80 年代就出現了。

➤ 當時的深度網路並沒有表現出任何超凡能力。

✓ 因為當時的資料資源遠沒有現在豐富，而深度學習網路恰恰需要大量的資料以提高它的訓練實例數量。

## ◆到了 2000 年，當大多數科學家已經對深度學習失去興趣的時候，傑夫·辛頓帶領他的學生繼續在這個冷門的領域裡堅持耕耘。

➤ 2009 年，辛頓小組獲得了意外的成功。

✓ 他們的深度學習神經網路在語音辨識應用中取得了重大的突破，轉換精度已經突破了世界紀錄，錯誤率比以前減少了 25%。

✓ 辛頓小組的研究讓語音辨識領域縮短了至少 10 年的時間。

➤ 蘋果公司把他們的研究成果應用到了 Siri 語音辨識系統上，使得 iPhone 5 全球熱賣。

## ◆為什麼把網路的深度提高，配合上大資料的訓練就能使得網路性能有如此大的改善呢？

- 因為人腦恰恰就是這樣一種多層次的深度神經網路。
- 已有的證據表明，人腦處理視覺資訊就是經過多層加工完成的。
  - ✓ 深度學習實際上只不過是對大腦的一種模擬。
- 模式識別問題長久以來是人工智慧發展的一個主要瓶頸。深度學習技術似乎已經突破了這個瓶頸。
  - ✓ 有人甚至認為，深度學習神經網路已經可以達到 2 歲小孩的識別能力。

## ◆有理由相信，深度學習會將人工智慧引入全新的發展局面。



## 五、神經網路的預測應用

◆與簡單的分類或迴歸問題不同，時間數列問題增加了順序的複雜性，或觀察之間的時間相依性。

- 增加資料特別處理的難度，當我們要進行模型的配湊(Fitting)與評估(Evaluating)。
- 而其時間結構可能有助於模型化，如果額外的結構，如趨勢與季節性能被用來提升模型化能力。

◆傳統上，通常以線性方法如 ARIMA 來進行時間數列的預測，

- 因為它被廣為所知，而且在許多問題是有效的。
- 然而，它也有許多的限制。(民用版)
  - ✓ 需要資料完整。
  - ✓ 聚焦線性關係。
  - ✓ 聚焦固定的時間相依性。
  - ✓ 單一變數。
  - ✓ 單步預測。

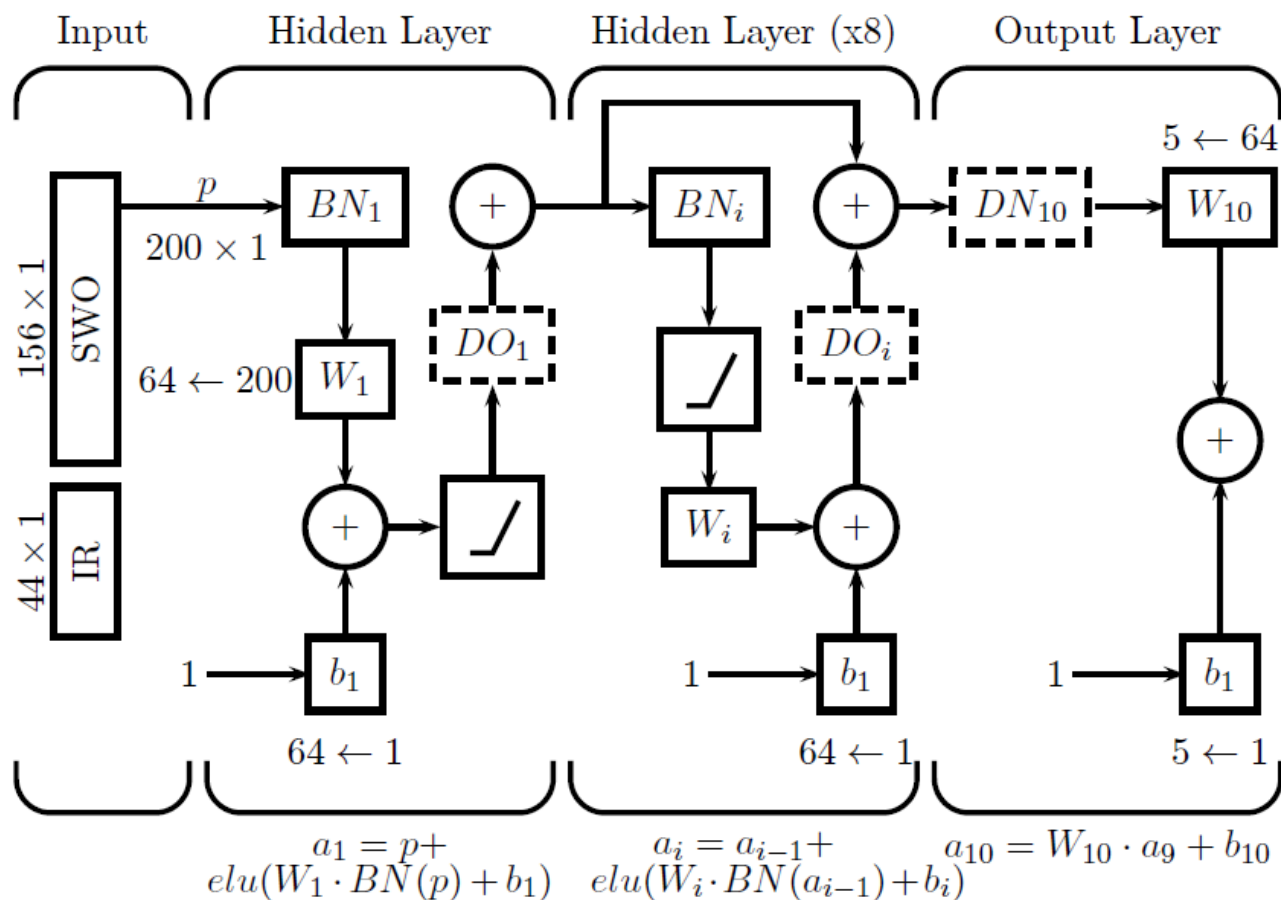
◆機器學習，在一些方向具有優勢。

- 多輸入變數。
- 複雜非線性關係。
- 資料缺失。

◆傳統上，為求成效良好，需要具備訊號處理背景的行業專家(Domain Expert)或實務工作者手工提取特徵。

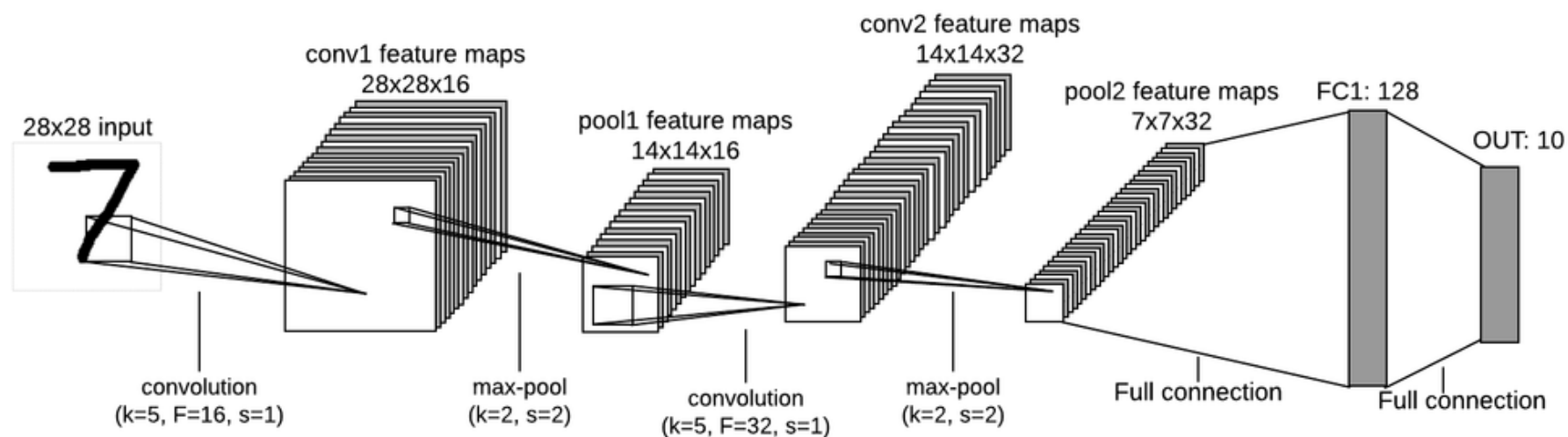
# (一)多層感知機(MLP)的處理

## ◆Classical



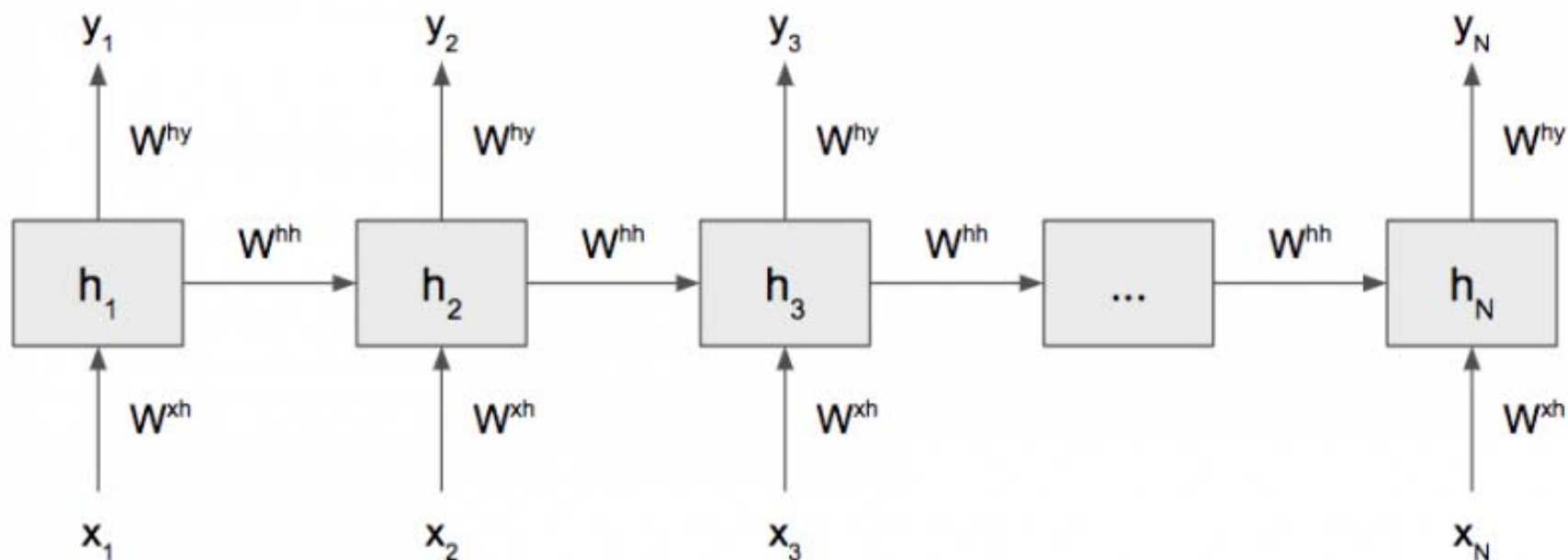
## (二)捲積神經網路(Convolutional NN)的處理

### ◆Convolutional Kernel , Visual Application 。



### (三)循環神經網路(Recurrent NN)的處理

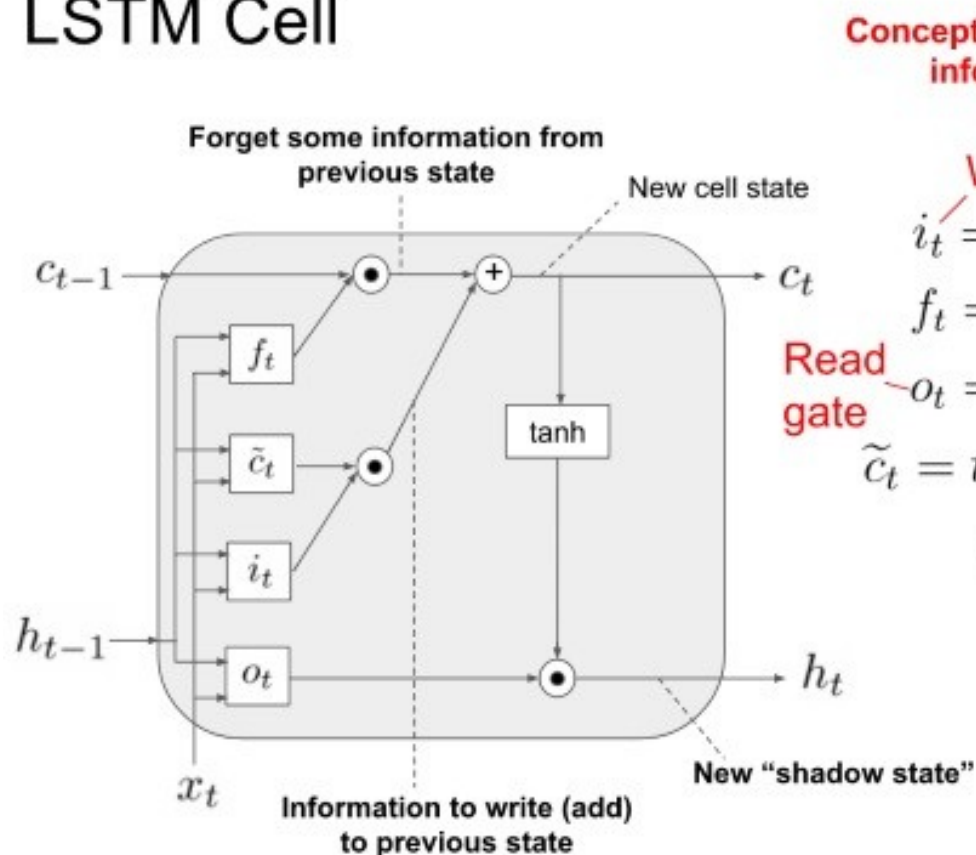
#### ◆Dependence



## (四)長短期記憶模型(Long-Short Term Model)的處理

### ◆Long Term and Short Term Memory

#### LSTM Cell



Conceptually, we lose information

We use shadow state to calculate gates

LSTM equations

Write gate

$$i_t = \sigma(W_i h_{t-1} + U_i x_t + b_i) \quad \text{Input gate}$$

$$f_t = \sigma(W_f h_{t-1} + U_f x_t + b_f) \quad \text{Forget gate}$$

$$o_t = \sigma(W_o h_{t-1} + U_o x_t + b_o) \quad \text{Output gate}$$

$$\tilde{c}_t = \tanh(W h_{t-1} + U x_t + b) \quad \text{Memory cell candidate}$$

$$c_t = f_t \circ c_{t-1} + i_t \circ \tilde{c}_t \quad \text{Memory cell}$$

$$h_t = o_t \circ \tanh(c_t) \quad \text{Shadow state}$$

$$y_t = h_t \quad \text{Cell Output}$$

Read occurs after writing



142

◆Fischer, Thomas & Christopher , "Deep learning with long short-term memory networks for financial market predictions" , FAU Discussion Papers in Economics , Nov, 2017 ◦

# 六、自然語言的交易應用

## (一)市場訊息的傳遞過程

### ◆市場的訊息反應

- 意見的擴散(文字)：中性字眼—>意見字眼
- 耳語的形成(文字)：意見字眼—>情緒字眼
- 交易的產生(數字)：交易量—>變動量
- 價格的變化(數字)：交易價—>變動量

### ◆Fed 的看法(Statement)。



# Text of November FOMC statement

Information received since the Federal Open Market Committee met in September indicates that the labor market has continued to strengthen and that economic activity has been rising at a strong rate. Job gains have been strong, on average, in recent months, and the unemployment rate has declined. Household spending has continued to grow strongly, while growth of business fixed investment has moderated from its rapid pace earlier in the year. On a 12-month basis, both overall inflation and inflation for items other than food and energy remain near 2 percent. Indicators of longer-term inflation expectations are little changed, on balance.

Consistent with its statutory mandate, the Committee seeks to foster maximum employment and price stability. The Committee expects that further gradual increases in the target range for the federal funds rate will be consistent with sustained expansion of economic activity, strong labor market conditions, and inflation near the Committee's symmetric 2 percent objective over the medium term. Risks to the economic outlook appear roughly balanced.

In view of realized and expected labor market conditions and inflation, the Committee decided to maintain the target range for the federal funds rate at 2 to 2-1/4 percent.

In determining the timing and size of future adjustments to the target range for the federal funds rate, the Committee will assess realized and expected economic conditions relative to its maximum employment objective and its symmetric 2 percent inflation objective. This assessment will take into account a wide range of information, including measures of labor market conditions, indicators of inflation pressures and inflation expectations, and readings on financial and international developments.

## (二)NLP的應用過程

◆社群媒體的討論

◆語彙的判讀(斷詞)

◆詞頻的計算

◆正反面情緒

◆買、賣意見的討論

◆交易行為的產生

# 七、程式工具

## (一)傳統工具

### ◆MultiCharts，

➤ <http://www.multicharts.com.tw/>

### ◆TradeStation，

➤ <https://www.tradestation.com/>

### ◆精業系統，各券商系統，CMoney。

### ◆技術分析與財報資料分析為主。

## (二)一般量化交易工具

### ◆Excel VBA ,

- 與 Excel Spread Sheet 資料整合。
- 共通的介面。

### ◆Matlab ,

- 多個數學工具箱。
- 神經網路工具箱。
- 矩陣資料導向。

### (三)高頻交易工具

#### ◆C++/GPU/Linux System

- 效率為主要考量。
- 專業機構使用。

## (四)AI交易策略工具

### ◆Python

- AI(NN) Framework
  - ✓ TeansorFlow(Google)
  - ✓ MXNet(Amazon)
  - ✓ CNTK(Microsoft)
  - ✓ Theano(Univ. of Montreal)
  - ✓ H2O(MIT)
- Quantitative Trading Platform

### ◆C# 、Java 、C++

- Open Source(SourceForge.net)
- Quantitative Trading Platform

## 八、線上交易平台的興起

### ◆Quantopian ,

- <https://www.quantopian.com/> , Python 平台 , 提供撰寫策略的便利性

### ◆bincentive ,

- <https://www.bincntive.com/> , 交易平台 , 區塊鏈交易紀錄公開 , 安全機制可以跟單

## 九、結論

- ◆AI 應用方興未艾。
- ◆Deep Learning 重中之重。
- ◆數據處理已經成熟。
- ◆文本處理(NLP)快速成長。
- ◆速度是關鍵。
- ◆GPU 平行運算必備工具。
- ◆程式能力是必要的。



## 十、參考資料

- ◆零起點 TensorFlow 與量化交易，何海群，電子工業出版社，2018。
- ◆高頻交易，Irene Aldridge，談效俊等譯，機械工業出版社，2015。
- ◆GPU 高性能編程 CUDA 實戰，Jason Sanders & Edward Kandrot，聶雪軍等譯，機械工業出版社，2011。
- ◆C++語言。
- ◆C#語言。
- ◆Python 語言。
- ◆VBA 語言。

金融研訓院 特約講師  
證券暨投資分析人員合格(CSIA)  
希奇資本 技術長(CTO)

董 夢 雲 財務博士

Mobil: (Taiwan)0988-065-751 (China)1508-919-2872

EMail: [dongmy@ms5.hinet.net](mailto:dongmy@ms5.hinet.net)

Line ID/WeChat ID: andydong3137

## 專長

GPU 平行運算與財務工程，C#、.Net Framework、CUDA、OpenCL、C、C++。

外匯與利率結構商品評價實務，股權與債權及衍生商品評價實務。

風險管理理論與實務，資本配置與額度規劃。

## 經歷

中國信託商業銀行交易室研發科主管

凱基證券風險管理部主管兼亞洲區風險管理主管

中華開發金控、工業銀行風險管理處處長

永豐金控、商業銀行風險管理處處長

永豐商業銀行結構商品開發部副總經理

## 著作

金融選擇權：市場、評價與策略，第二版，1997，新陸書局。

財務工程與 Excel VBA 的應用：選擇權評價理論之實作，2005，證券暨期貨發展基金會。

## 翻譯

衍生性金融商品與內部稽核，2003，金融研訓院。