Essence of a neural net approach to discerning ML patterns

1. First period (mid 1980s)
   First used in databases to `reduce’ the number of dimensions in complex multi-dimensional situations

2. Second period (early 1990s)
   Combined with artificial neural networks (ANN) to `learn’ new patterns of activity, independently of human agent intervention
   [an alternative to the then-predominant non-adaptive structured query language (SQL) approach]

3. Third period (late 1990s-present)
   Combined with fuzzy logic in applications to carbon-based (ie human feature) recognition systems and fraud detection
ANN is based on Parallel Distributed Processing (PDP)

Recurring input-output instances are ‘associated’ or connected

Second (output) vector is transposed

Weight (or covariance) matrix

Cross product formed

Accumulated Memory $A$

$$A = x y^T$$
With training, data memories `Accumulate`:

**Learning to identify recurrent activity patterns (eg regular cash movements):**

\[ M_1 = \begin{array}{ccc} 0 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{array} \]

\[ M_2 = \begin{array}{ccc} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \end{array} \]

\[ A = M_1 + M_2 = \begin{array}{ccc} 0 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 1 & 0 \end{array} \]

**Learning to identify odd activity patterns (eg suspect cash movements):**

\[ M_1 = \begin{array}{ccc} 0 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 1 & 0 \end{array} \]

\[ M_2 = \begin{array}{ccc} 0 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 1 & 0 \end{array} \]

\[ A = M_1 + M_2 = \begin{array}{ccc} 0 & 1 & 0 \\ 1 & 2 & 1 \\ 0 & 2 & 0 \end{array} \]
So after training, responses will be:

If it looks like this:

Then **non-suspect**

If it looks like this:

Then **suspect**

---

Compared to regular template $R$, **1 square** is different.

Compared to suspect template $S$, **3 squares** are different!

`Deviation` is **least** for the regular pattern $R$
So this activity is not suspicious
POSTSCRIPT:
Abnormally high or irregular levels of `noise’ in communication channels (chatter & money flows), blurring the template, did indeed feature in the days leading up to 9/11.
Important points to note re using artificial neural networks (ANN) as an anti-ML pattern recognition technique:

1. ANNs can readily adapt (or morph) to new patterns
2. ANNs rely almost exclusively on data on which to base learning (no data; no rules)
3. ANNs may be slow to learn, depending on the learning algorithm used
4. ANNs may over-learn (detecting spurious patterns), or inappropriately generalise
5. ANNs may be programmed to embody a `memory decay’ term, to speed up processing.

For further details refer to the associated commentary to part 6 of this interactive CD.