شبکه‌های عصبی مصنوعی
۰۱-۷۱۳-۱۱-۱۳
بخش ششم

ART1



دانشگاه شهید بهشتی
دانشکده‌ی علوم و مهندسی کامپیوتر
بهار ۱۳۹۵
احمد محمودی ازناوه

فهرست مطالب

- نظریه تشدید وفقی (ART1)
  - پایداری در برابر انعطاف‌پذیری
  - ساختار شبکه
  - شیوه‌ی آموزش
  - چند مثال
- ویرایش‌های مختلف ART





پیش‌گفتار

- این نظریه بر اساس نحوه‌ی پردازش اطلاعات در مغز توسط Stephen Grossberg و Gail Carpenter مطرح شد.
- این نظریه مدلی از شبکه‌های عصبی را توصیف می‌کند که به هر دو شیوه‌ی بی‌نظارت و بانظارت آموزش می‌بینند و برای بازشناخت الگو و پیش‌بینی کاربرد دارد.



Carpenter, G.A. and S. Grossberg, A massively parallel architecture for a self-organizing neural pattern recognition machine. Comput. Vision Graph. Image Process., 1987. 37(1): p. 54-115.

Adaptive Resonance Theory

Stability-Plasticity Dilemma

- این سیستم «پایداری» و «انعطاف‌پذیری» را همزمان در نظر می‌گیرد.
- Stability:
  - رفتار سیستم در برابر رویدادهای غیرمرتبط پایدار است.
- Plasticity:
  - سیستم در برابر رویدادهای بزرگ خود را تطبیق می‌دهد.
- در این صورت همانند رفتار بیولوژیکی انسان سیستم همواره در حال آموزش است و از دانش قبلی خود استفاده می‌کند.





Adaptive Resonance Theory انعطاف‌پذیری

- ART1 یک شبکه‌ی بی‌نظارت است، البته به صورت بانظارت هم تعمیم یافته است.
- برای ورودی‌های جدید اگر خوشه‌ی (Cluster) مناسب یافت نشود که این ورودی را در خود جای دهد، گروه جدیدی ایجاد می‌گردد.
- این کار با استفاده از پارامتری به نام احتیاط(هوشیاری) (vigilance) انجام می‌شود. در واقع با این پارامتر تعادلی بین پایداری و انعطاف‌پذیری ایجاد می‌شود.
- در این شبکه به جای استفاده از اختلاف نمونه‌ها (absolute difference)، شباهت نسبی (relative similarity) مورد استفاده قرار می‌گیرد.

پارامتر احتیاط (هوشیاری)

- پارامتر احتیاط که با ρ نشان داده می‌شود، وظیفه‌ی مراقبت از شبکه را بر عهده دارد.
- هر چقدر ρ بیشتر باشد، گروه‌ها کوچک‌تر و تعداد آن‌ها بیشتر خواهد بود.
- شبکه‌ی ART به گونه‌ای برای ورودی‌های جدید که امکان طبقه‌بندی شدن در گروه‌های جاری را ندارند، گروه جدید ایجاد می‌کند.
- ART1 تنها از بردارهای ورودی **دودویی** استفاده می‌کند.





- N الگو داریم که هر یک اندازه‌ی m دارند.

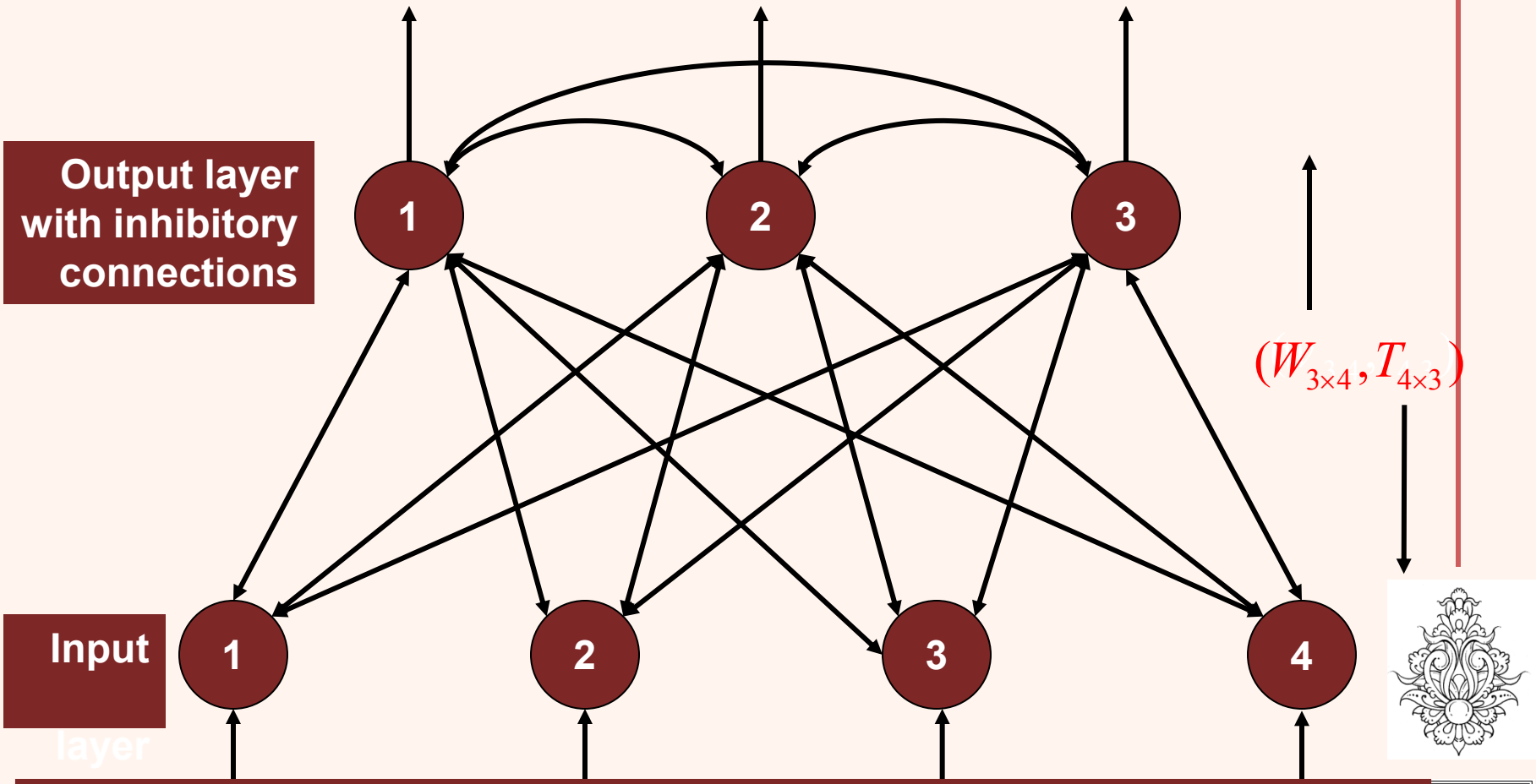
$$X_{n\ (m\times 1)} \qquad n = 1, 2, 3, ..., N$$

- **هدف:**
  - خوشه‌بندی نمودن بردارهای X به گونه‌ای که هر دسته دارای یک مرکز باشد.
  - به این مرکز Prototype (P) یا همان شاخص دسته گویند.
- **فرضیات**
  - بردارهای X_n دارای اعضای صفر و یک هستند.

The ART-1 Network

ساختار شبکه

نزدیک‌ترین فاصله میان ورودی و شاخص مشخص می‌دارد ورودی به کدام گروه تعلق دارد



دو گروه وزن وجود دارد :

Bottom-up weight : مشخص می‌کند بهترین کاندید برای خروجی ورودی جاری کدام است.

Top Down Weight: شاخص یا همان prototype بهینه گروه انتخاب شده را معین می‌کند.

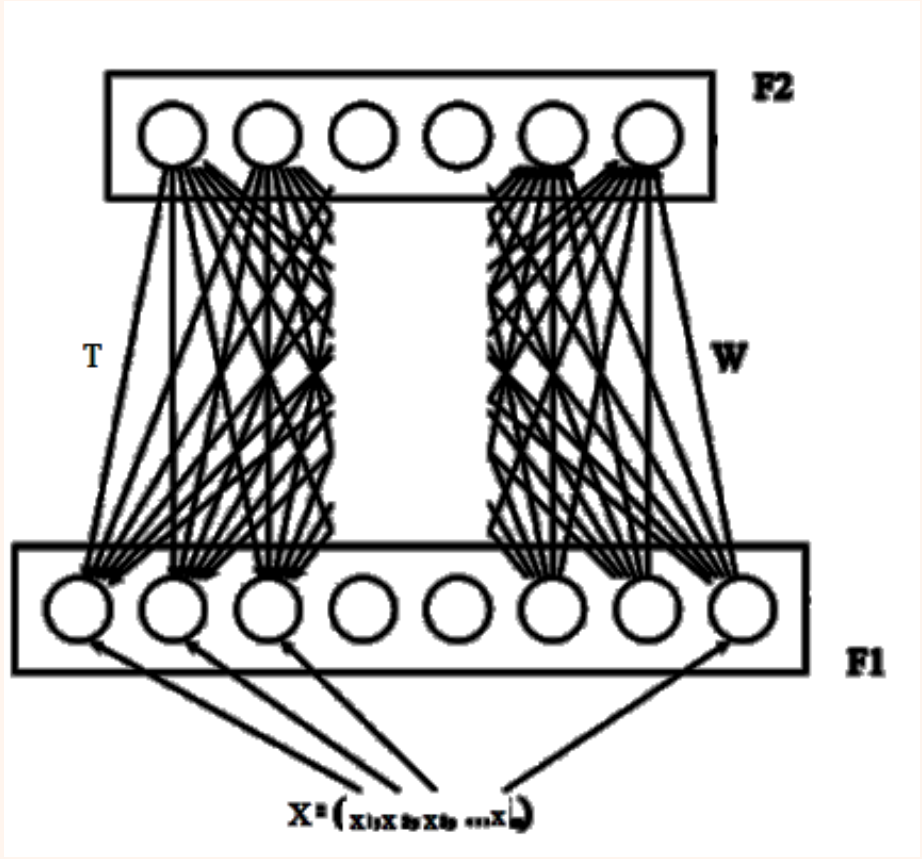
ART1

- یک الگو ممکن است چندین بار به شبکه اعمال شود.
- برخی شبکه‌ها با کاهش نرخ آموزش به پایداری می‌رسند، در این صورت به شبکه اجازه‌ی یادگیری داده‌های جدید داده نمی‌شود.
- توانایی شبکه در یادگیری داده‌های جدید در هر مرحله Plasticity نامیده می‌شود.



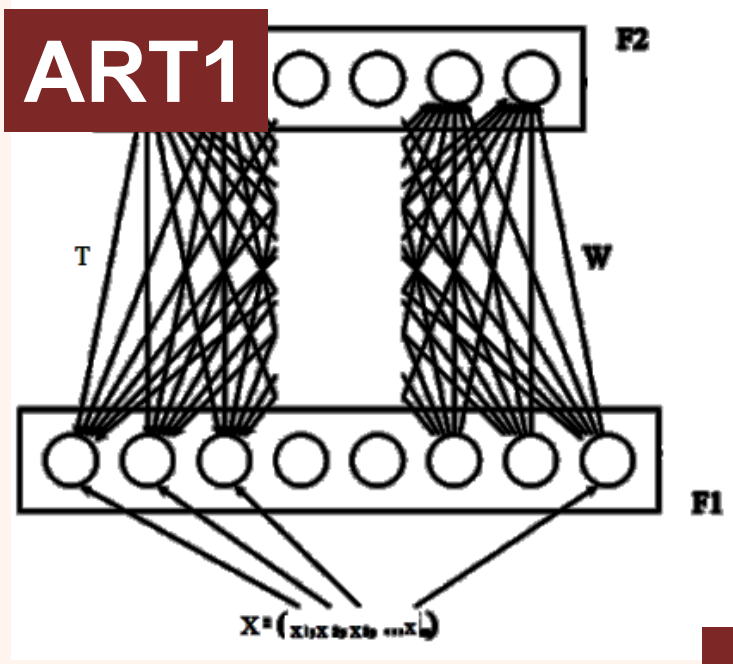


- تعداد یک‌های موجود در بردار X_n $\|X_n\|$
- تعداد یک های مشترک در دو بردار $X.Y = X \cap Y$



Basic ART Structure

ART1



- اگر وزن‌های اتصالی از واحدهای ورودی به واحد iام خروجی را مانند زیر تعریف کنیم:

$$W_i = \begin{bmatrix} w_{i1} \\ . \\ . \\ w_{iN} \end{bmatrix}$$

Weight Matrix Feed Forward

- پس از این‌که وزن‌ها اعمال شد، در خروجی واحد برنده انتخاب شده، یکی برنده و بقیه صفر می‌شوند.
- اگر چند واحد مساوی بود چپ‌ترین(طبق ساختار شکل) برنده است.

ART1

T: Weight Matrix Feed Backward

- پس از انتخاب برنده، یک ماتریس وزن‌دهی رو به عقب وجود دارد.
- اعضای T با توجه به ورودی که باعث برنده شدن شده است تنظیم می‌شوند.
- در انتها مقادیر وزن‌های ماتریس T به عنوان شاخص گروه مشخص می‌گردد.
- تمامی این فرآیند یک Iteration محسوب می‌شود.





Initialize phase

مقداردهی اولیه

- ماتریس T: تمامی اعضایش برابر با یک قرار داده می‌شوند.
- ماتریس W: تمامی اعضایش 1/(m+1) قرار داده می‌شوند.
- m (اندازه‌ی بردار X)
- N (تعداد بردار X)





Recognition phase

Attentional Subsystem

- بردار ورودی X به شبکه اعمال می‌شود و از طریق ماتریس W به لایه‌ی خروجی هدایت می‌شود. (ماتریس T غیرفعال)
  - در خروجی واحد برنده انتخاب می‌شود.





Comparison phase

Orienting Subsystem

$$S = \frac{\sum_{i=1}^{m} t_{ij} . X_i}{\sum_{i=1}^{m} X_i}$$

- مقدار S برای واحد برنده‌ی j محاسبه می‌شود.
- پارامتر ρ مشخص می‌کند آیا ورودی متعلق به گروهی موجود در نظر گرفته خواهد شد یا باید گروهی جدید برایش ایجاد گردد.
- آیا S>ρ
  - بلی، ورودی به گروه j تعلق دارد (بردار شاخص t_j اصلاح می‌شود)

fast learning

$$t_{i,j}(k+1) = t_{i,j}(k) X_i \qquad W_{i,j^*}(k+1) = \frac{t_{i.j^*}(k) X_i}{0.5 + \left\| T^T{}_{j^*} X \right\|}$$

  - خیر، ورودی به گروه جدیدی تعلق دارد، گروه جدیدی اضافه خواهد شد.

Learning rule:

$$t_{i,j}(k+1) = (1-\beta) t_{i,j}(k) + \beta t_{i,j}(k) X_i$$

نکات

- تمام ورودی‌ها در طی تکرارهای متوالی اعمال می‌شوند.
- برای مجموعه‌ای محدود، از الگوهای آموزشی الگوریتم به پایداری می‌رسد.
- پس از رسیدن به پایداری هر الگوی ورودی مستقیماً به بردار شاخص خود دسترسی دارد.



A. Initialize each top-down weight $t_{l,j}$ (0) = 1;

B. Initialize bottom-up weight $w_{j,l}$ (0) = 1/m+1 ;

C. **While** the network has not stabilized, **do**

1. Present a randomly chosen pattern $x = (x_1,\ldots,x_n)$ for learning

2. Let the active set A contain all nodes; calculate

$y_j = w_{j,1}\ x_1 + \ldots + w_{j,n}\ x_n$ for each node $j \quad A$;

3. **Repeat**

a) Let j^* be a node in A with largest y_j, with ties being broken arbitrarily;

b) Compute $s^* = (s^*_1,\ldots,s^*_n)$ where $s^*_l = t_{l,j^*}\ x_l$;

c) Compare similarity between s^* and x with the given vigilance parameter ρ :

if $\dfrac{\sum_{l=1}^{n} s^*_l}{\sum_{l=1}^{n} x_l} \leq \rho$ **then** remove j^* from set A

else associate x with node j^* and update weights:

w_{j^*l} (new) = $\dfrac{t_{l,j^*}(old)x_l}{0.5 + \sum_{l=1}^{n} t_{l,j^*}(old)x_l}$ $\qquad$ t_{l,j^*} (new) = $t_{l,j^*}(old)x_l$

Until A is empty or x has been associated with some node j

4. If A is empty, then create new node whose weight vector coincides with current input pattern x;

end-while

مثال

- فرض کنیم چهار بردار ورودی همانند زیر داریم که می‌خواهیم در سه گروه قرار بگیرند:

$$X_1 = [1,\ 0,\ 1,\ 0]^T$$
$$X_2 = [1,\ 1,\ 0,\ 0]^T$$
$$X_3 = [0,\ 1,\ 1,\ 0]^T$$
$$X_4 = [1,\ 1,\ 1,\ 0]^T$$

$m=4$ $N=4$

- برای وزن‌های اولیه خواهیم داشت:

$$t_{ij\,(0)} = 1$$

$$w_{ij\,(0)} = \frac{1}{1+m} = \frac{1}{5} = 0.2$$

vigilance parameter ρ = 0.6

ادامه‌ی مثال

- فرض می‌کنیم هر گروه یک شاخص دارد برای ابتدای کار شاخص را برابر با ۱ قرار می‌دهیم:

$$T = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}_{4\times 3} \qquad W = \begin{bmatrix} 0.2 & 0.2 & 0.2 \\ 0.2 & 0.2 & 0.2 \\ 0.2 & 0.2 & 0.2 \\ 0.2 & 0.2 & 0.2 \end{bmatrix}_{4\times 3}$$

- پس از مرحله‌ی مقداردهی اولیه، ورودی‌ها را به سیستم اعمال می‌کنیم:





ادامه‌ی مثال

$$X_1 \rightarrow W^T X_1 = \begin{bmatrix} 0.4 \\ 0.4 \\ 0.4 \end{bmatrix}$$

همه با هم برابرند
اولی انتخاب می‌شود

$$j^* = 1$$

- تست پارامتر vigilance

$$S = \frac{X_1^T . T_1}{\|X_1\|} = \frac{2}{2} = 1 > \rho$$



$$T_{j^*}(1) = T_1(1) = T_1(0).X_1$$

$$= [1\ 0\ 1\ 0]^T$$

ادامه‌ی مثال

$$W_{ij*}(1) = \frac{t_{ij*}(0)X_{1i}}{0.5 + X_1^T . T_{j*}(0)} = \frac{[1\ 0\ 1\ 0]}{2.5} = [0.4\ 0\ 0.4\ 0]$$

$$T = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 1 \\ 0 & 1 & 1 \end{bmatrix}_{4\times 3} \qquad W = \begin{bmatrix} 0.4 & 0.2 & 0.2 \\ 0 & 0.2 & 0.2 \\ 0.4 & 0.2 & 0.2 \\ 0 & 0.2 & 0.2 \end{bmatrix}_{4\times 3}$$

اتمام یک iteration

ادامه‌ی مثال

- X_2 اعمال می‌گردد.

$$X_2 \rightarrow W^T X_2 = \begin{bmatrix} 0.4 \\ 0.4 \\ 0.4 \end{bmatrix}$$

$$j^* = 1$$

- تست ρ

$$S = \frac{X_2^T T_1}{\|X_2\|} = \frac{1}{2} = 0.5 < \rho$$



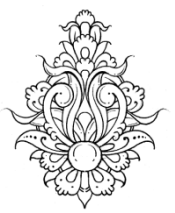
- از ۲ به عنوان برنده استفاده می‌کنیم:

$$S = \frac{X_2^T . T_2}{\|X_2\|} = \frac{2}{2} = 1 > \rho$$



$$T_2(2) = T_2(1) \; . X_2$$

$$= [1\ 1\ 0\ 0]^T$$





ادامه‌ی مثال

$$T_2(2) = T_2(1) \wedge X_2$$

$$= [1\ 1\ 0\ 0]^T$$

$$W_2 = \frac{T_2(1).X_2}{0.5 + X_2^T.T_2(1)} = \frac{[1\ 1\ 0\ 0]}{2.5} = [0.4\ 0.4\ 0\ 0]$$

$$T = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \\ 0 & 0 & 1 \end{bmatrix}_{4\times 3} \qquad W = \begin{bmatrix} 0.4 & 0.4 & 0.2 \\ 0 & 0.4 & 0.2 \\ 0.4 & 0 & 0.2 \\ 0 & 0 & 0.2 \end{bmatrix}_{4\times 3}$$

اتمام دومین iteration





ادامه‌ی مثال

- اعمال X_3

$$X_3 \rightarrow W^T X_3 = \begin{bmatrix} 0.4 \\ 0.4 \\ 0.4 \end{bmatrix}$$

$$j^* = 1$$

- تست ρ

$$S = \frac{X_3^T . T_1(2)}{\|X_3\|} = \frac{1}{2} = 0.5 < \rho$$



- از ۲ به عنوان برنده استفاده می کنیم:

$$S = \frac{X_3^T . T_{2(2)}}{\|X_3\|} = \frac{1}{2} = 0.5 < \rho$$



شبکه عصبی

ادامه‌ی مثال

- از ۳ به عنوان برنده استفاده می کنیم:

$$S = \frac{X_3^T T_{3(2)}}{\|X_3\|} = \frac{2}{2} = 1 > \rho$$



$$T_3(3) = T_3(2) \,.\, X_3 \qquad W_3 = [0 \quad 0.4 \quad 0.4 \quad 0]$$

$$= [0\ 1\ 1\ 0]^{\mathrm{T}}$$

$$T = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}_{4\times 3} \qquad W = \begin{bmatrix} 0.4 & 0.4 & 0 \\ 0 & 0.4 & 0.4 \\ 0.4 & 0 & 0.4 \\ 0 & 0 & 0 \end{bmatrix}_{4\times 3}$$





اتمام سومین iteration

شبکه عصبی

ادامه‌ی مثال

- X_4 اعمال می‌گردد:

$$X_4 \rightarrow W^T X_4 = \begin{bmatrix} 0.8 \\ 0.8 \\ 0.8 \end{bmatrix}$$

$$j^* = 1$$

- تست ρ

$$S = \frac{X_4^T . T_1}{\|X_4\|} = \frac{2}{3} > \rho$$

$$T_1(4) = T_1(3) . X_4$$

$$= [1\ 0\ 1\ 0]^T$$

$$W_1 = [0.4 \quad 0 \quad 0.4 \quad 0]$$

ثابت ماندن P پس به حالت پایدار رسیده‌است

X_3 X_2 X_4 X_1

شبکه عصبی

مثال

$$X_1 = [1,\ 1,\ 0,\ 0,\ 0,\ 0,\ 1]^T$$
$$X_2 = [0,\ 0,\ 1,\ 1,\ 1,\ 1,\ 0]^T$$
$$X_3 = [1,\ 0,\ 1,\ 1,\ 1,\ 1,\ 0]^T$$
$$X_4 = [0,\ 0,\ 0,\ 1,\ 1,\ 1,\ 0]^T$$
$$X_5 = [1,\ 1,\ 0,\ 1,\ 1,\ 1,\ 0]^T$$

- بردارهای ورودی را در نظر بگیرید:
- با داشتن اندازه ورودی هفت

و یک خروجی

vigilance parameter $\rho = 0.7$

برای اولین بردار ورودی داریم:

$[1, 1, 0, 0, 0, 0, 1]^T$

$$y_1 = \frac{1}{8}\cdot 1 + \frac{1}{8}\cdot 1 + \frac{1}{8}\cdot 0 + \frac{1}{8}\cdot 0 + \frac{1}{8}\cdot 0 + \frac{1}{8}\cdot 0 + \frac{1}{8}\cdot 1 = \frac{3}{8}$$

Y1 برنده است (رقابتی صورت نگرفته)

$$S = \frac{X_1^T T_1(0)}{\|X_1\|} = \frac{3}{3} = 1 > \rho$$



$$T_{j^*}(1) = T_1(1) = T_1(0).X_1$$

$$W(1) = \begin{bmatrix} \frac{1}{3.5} & \frac{1}{3.5} & 0 & 0 & 0 & 0 & \frac{1}{3.5} \end{bmatrix}^T$$

$$T(1) = \begin{bmatrix} 1 & 1 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}^T$$

- اعمال ورودی دوم: $[0, 0, 1, 1, 1, 1, 0]^T$

$$y_1 = \frac{1}{3.5}\cdot 0 + \frac{1}{3.5}\cdot 0 + 0\cdot 1 + 0\cdot 1 + 0\cdot 1 + 0\cdot 1 + \frac{1}{3.5}\cdot 0 = 0$$

$$S = \frac{X_2^T . T_1(1)}{\|X_2\|} = \frac{0}{4} = 0 < \rho$$



واحدی جدید اضافه می‌شود

$$W(2)=\begin{bmatrix} 1/3.5 & 1/3.5 & 0 & 0 & 0 & 0 & 1/3.5 \\ 0 & 0 & 1/4.5 & 1/4.5 & 1/4.5 & 1/4.5 & 0 \end{bmatrix}^T$$

$$T(2)=\begin{bmatrix} 1 & 1 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 1 & 1 & 1 & 0 \end{bmatrix}^T$$

- ورودی سوم اعمال می‌شود $[1, 0, 1, 1, 1, 1, 0]^T$

Y2 برنده است پس تست برای پارامتر vigilance صورت می‌گیرد

$$y_1=\frac{1}{3.5};\qquad y_2=\frac{4}{4.5}$$

$$S=\frac{X_3^T T_2(2)}{\|X_3\|}=\frac{4}{5}=0.8>\rho$$



- به روزرسانی وزن‌ها

$$W(3)=\begin{bmatrix} 1/3.5 & 1/3.5 & 0 & 0 & 0 & 0 & 1/3.5 \\ 0 & 0 & 1/4.5 & 1/4.5 & 1/4.5 & 1/4.5 & 0 \end{bmatrix}^T$$

$$T(3)=\begin{bmatrix} 1 & 1 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 1 & 1 & 1 & 0 \end{bmatrix}^T$$

- اعمال چهارمین ورودی $[0, 0, 0, 1, 1, 1, 0]^T$

$$y_1 = 0; \qquad y_2 = \frac{3}{4.5}$$

Y2 برنده است پس تست برای پارامتر vigilance صورت می‌گیرد

$$S = \frac{X_4^T . T_2(3)}{\|X_4\|} = \frac{3}{3} = 1 > \rho$$



شبکه عصبی

- به روزرسانی وزن‌ها

$$W(4)=\begin{bmatrix} 1/3.5 & 1/3.5 & 0 & 0 & 0 & 0 & 1/3.5 \\ 0 & 0 & 0 & 1/3.5 & 1/3.5 & 1/3.5 & 0 \end{bmatrix}^T$$

$$T(4)=\begin{bmatrix} 1 & 1 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 & 0 \end{bmatrix}^T$$

- اعمال ورودی پنجم $[1, 1, 0, 1, 1, 1, 0]^T$

$$y_1=\frac{2}{3.5};\qquad y_2=\frac{3}{3.5}$$

Y2 برنده است پس تست برای پارامتر vigilance صورت می‌گیرد

$$S=\frac{X_5^T T_2(4)}{\|X_5\|}=0.6<\rho$$



$$S=\frac{X_5^T T_1(4)}{\|X_5\|}=0.4<\rho$$







شبکه عصبی

- تست vigilance برای هر دو گروه قابل قبول نیست پس می‌باید گروه جدیدی ایجاد گردد:

$$W(5)=\begin{bmatrix} 1/3.5 & 1/3.5 & 0 & 0 & 0 & 0 & 1/3.5 \\ 0 & 0 & 0 & 1/3.5 & 1/3.5 & 1/3.5 & 0 \\ 1/5.5 & 1/5.5 & 0 & 1/5.5 & 1/5.5 & 1/5.5 & 0 \end{bmatrix}^T$$

$$T(5)=\begin{bmatrix} 1 & 1 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 & 0 \\ 1 & 1 & 0 & 1 & 1 & 1 & 0 \end{bmatrix}^T$$

دومین epoch

در epoch دوم، اولین بردار مجدداً اعمال می‌شود.

$[1, 1, 0, 0, 0, 0, 1]^T$

$$y_1 = \frac{3}{3.5}; \qquad y_2 = 0; \qquad y_3 = \frac{2}{5.5}$$

Y1 برنده است پس تست برای پارامتر vigilance صورت می‌گیرد

$$S = \frac{X_1^T T_1(5)}{\|X_1\|} = 1 > \rho$$



چون ورودی همانند بردار وزن است به‌روزرسانی لازم نیست.





دومین epoch

در epoch دوم، دومین بردار مجدداً اعمال می‌شود:

$[0, 0, 1, 1, 1, 1, 0]^T$

$$y_1 = 0; \qquad y_2 = \frac{3}{3.5}; \qquad y_3 = \frac{3}{5.5}$$

Y2 برنده است پس تست برای پارامتر vigilance صورت می‌گیرد

$$S = \frac{X_2 . T_2(6)}{\|X_2\|} = 1 > \rho$$



چون ورودی همانند بردار وزن است به‌روزرسانی لازم نیست.





دومین epoch

در epoch دوم، سومین بردار مجدداً اعمال می‌شود:

$[1, 0, 1, 1, 1, 1, 0]^T$

$$y_1 = \frac{1}{3.5}; \qquad y_2 = \frac{3}{3.5}; \qquad y_3 = \frac{4}{5.5}$$

Y2 برنده است پس تست برای پارامتر vigilance صورت می‌گیرد

$$S = \frac{X_3^T T_2(7)}{\|X_3\|} = 0.6 < \rho$$



از میان او ۳، ۳ بزرگ‌تر است پس تست را برای این گروه صورت می‌دهیم:

$$S = \frac{X_3^T T_3(7)}{\|X_3\|} = 0.8 > \rho$$







- با انتخاب ورودی جدید در گروه ۳ مجدداً وزن‌ها به‌روزرسانی می‌شود.

$$V(8)=\begin{bmatrix} 1/3.5 & 1/3.5 & 0 & 0 & 0 & 0 & 1/3.5 \\ 0 & 0 & 0 & 1/3.5 & 1/3.5 & 1/3.5 & 0 \\ 1/4.5 & 0 & 0 & 1/4.5 & 1/4.5 & 1/4.5 & 0 \end{bmatrix}^T$$

$$T(8)=\begin{bmatrix} 1 & 1 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 & 1 & 1 & 0 \end{bmatrix}^T$$

ادامه...

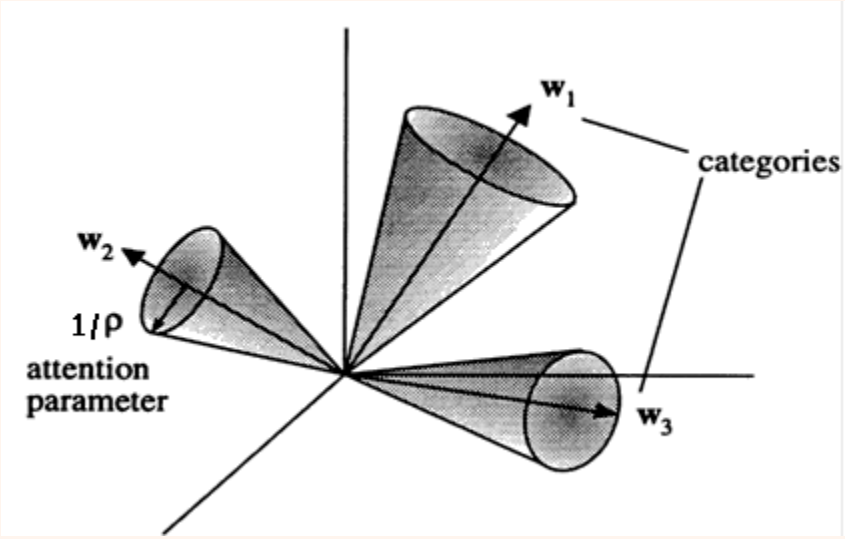
- به همین ترتیب با اعمال بردار چهارم گروه دوم برنده شده که تست را با موفقیت می‌گذراند، ولی در وزن تغییری ایجاد نمی‌کند.
- بر اثر اعمال بردار پنجم گروه سوم برنده شده که باز در وزن اثری ندارد و پس از آن در شبکه تغییری ایجاد نشده، به تعادل می‌رسد.



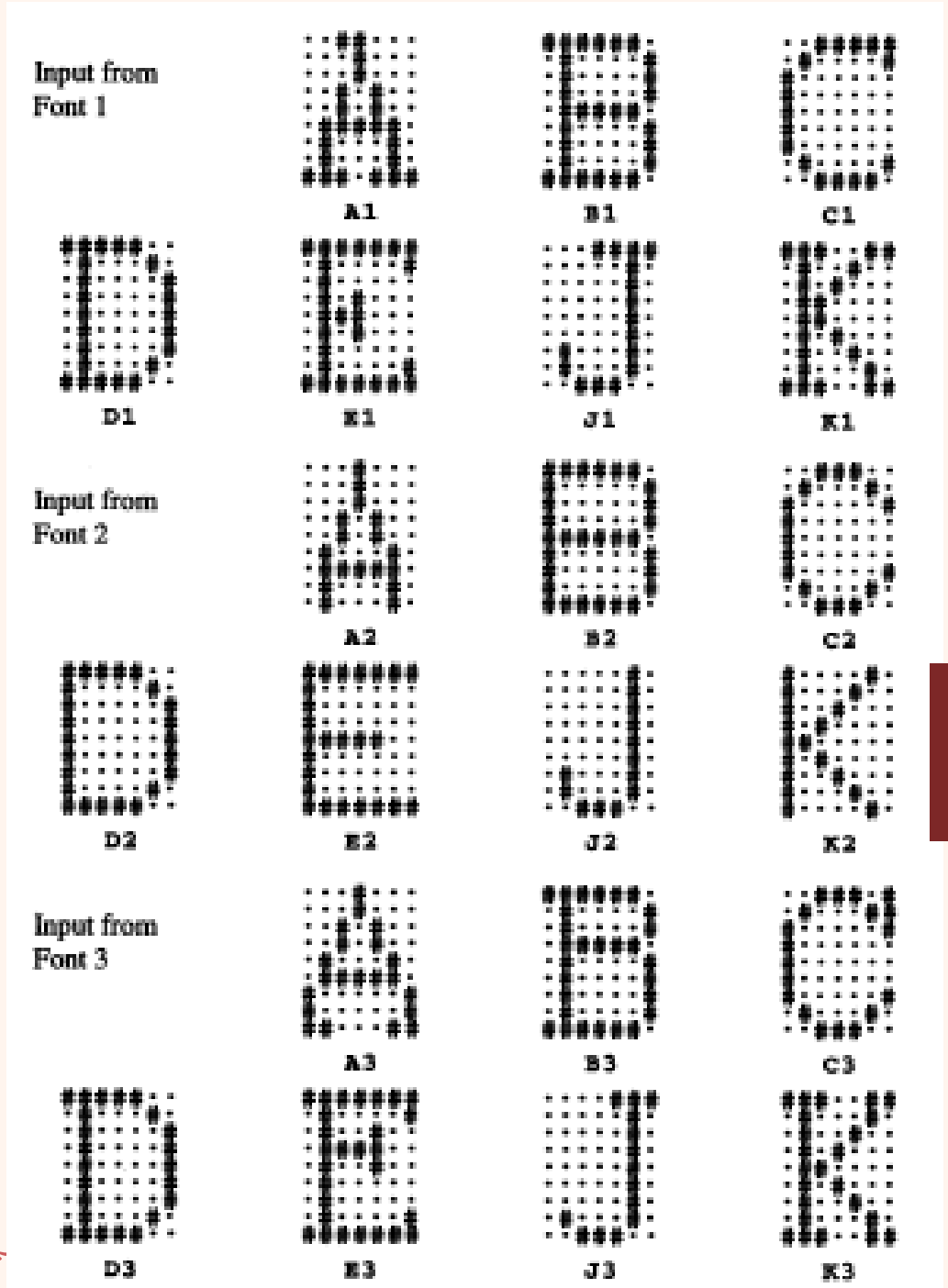


نکات

- اگر vigilance parameter کاهش یابد، گروه‌بندی داده‌ها فرمت دیگری خواهد داشت. در این صورت است که گروه‌بندی جدید کم‌تر اتفاق می‌افتد.
- برعکس با افزایش پارامتر مذکور تعداد گروه‌بندی ها افزایش خواهد یافت.



مثال



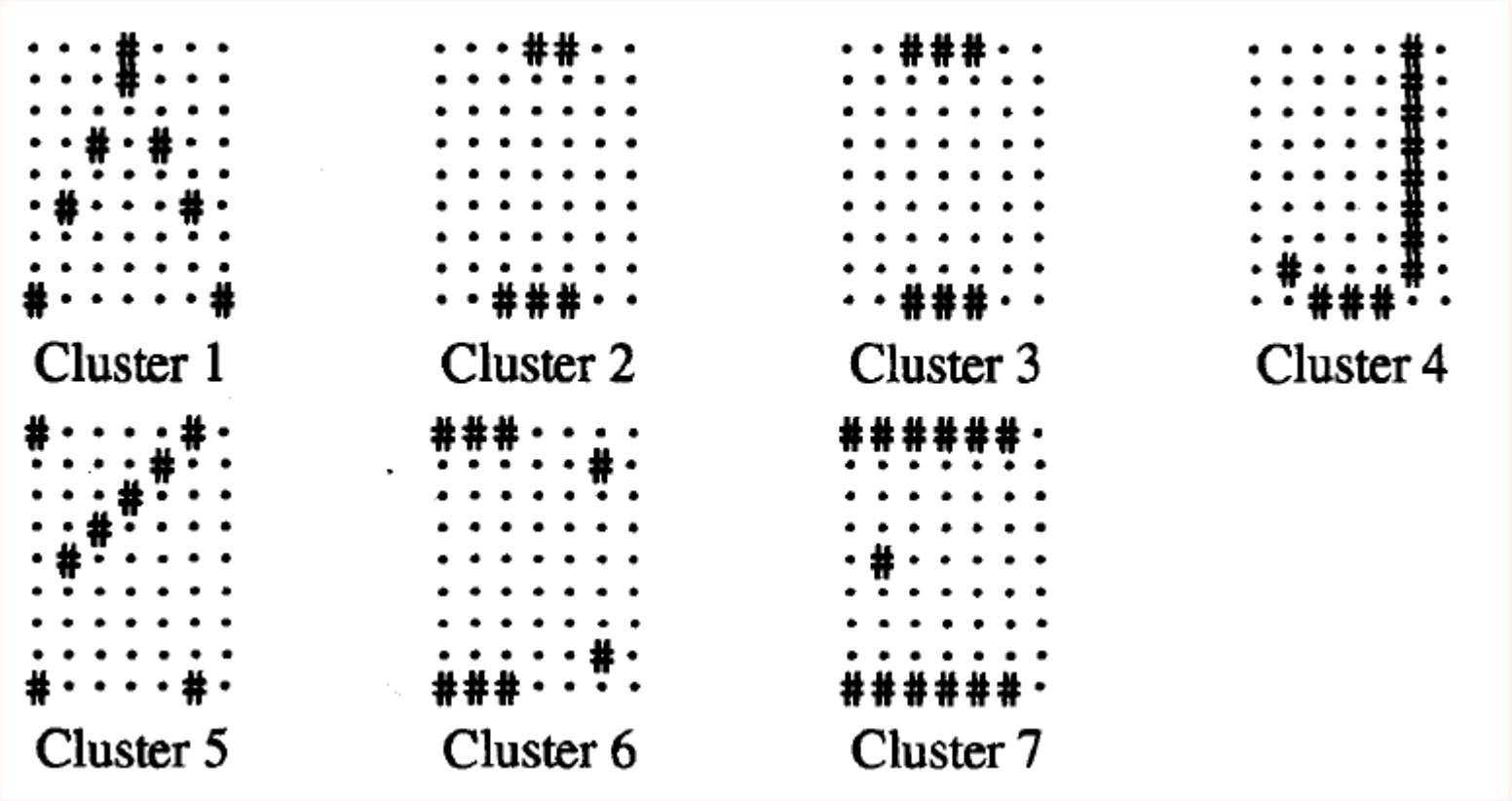
Training input patterns for character recognition examples

- با انتخاب ۱۰ گروه و با قرار دادن پارامتر vigilance=0.3 پس از سه مرحله سیستم به پایداری می‌رسد.
- ترتیب ورود داده‌ها همانند زیر در نظر گرفته شده است:

A1, A2, A3, B1, B2, B3, C1, C2, C3, . . . , J1, J2, J3, K1, K2, K3,

CLUSTER	EPOCH 1	EPOCH 2	EPOCH 3
1	A1, A2, A3	A1, A2, A3	A1, A2, A3
2	B1, B2, B3 C1, C2, C3 J1		
3	D1, D2, D3 E1, E2, E3	B1, B2, B3 C1, C2, C3	C1, C2, C3
4	J2, J3	J1, J2, J3	J1, J2, J3
5	K1, K2	K1, K2	K1, K2
6	K3	D1, D2, D3, K3	D1, D2, D3, K3
7		E1, E2, E3	B1, B2, B3 E1, E2, E3

- وزن‌های نهایی



- اگر ترتیب عوض شود با همان پارامترها خواهیم داشت:

A1, B1, C1, D1, E1, J1, K1, A2, B2, C2, D2, E2, J2, K2, A3, B3, C3, D3, E3, J3, K3,

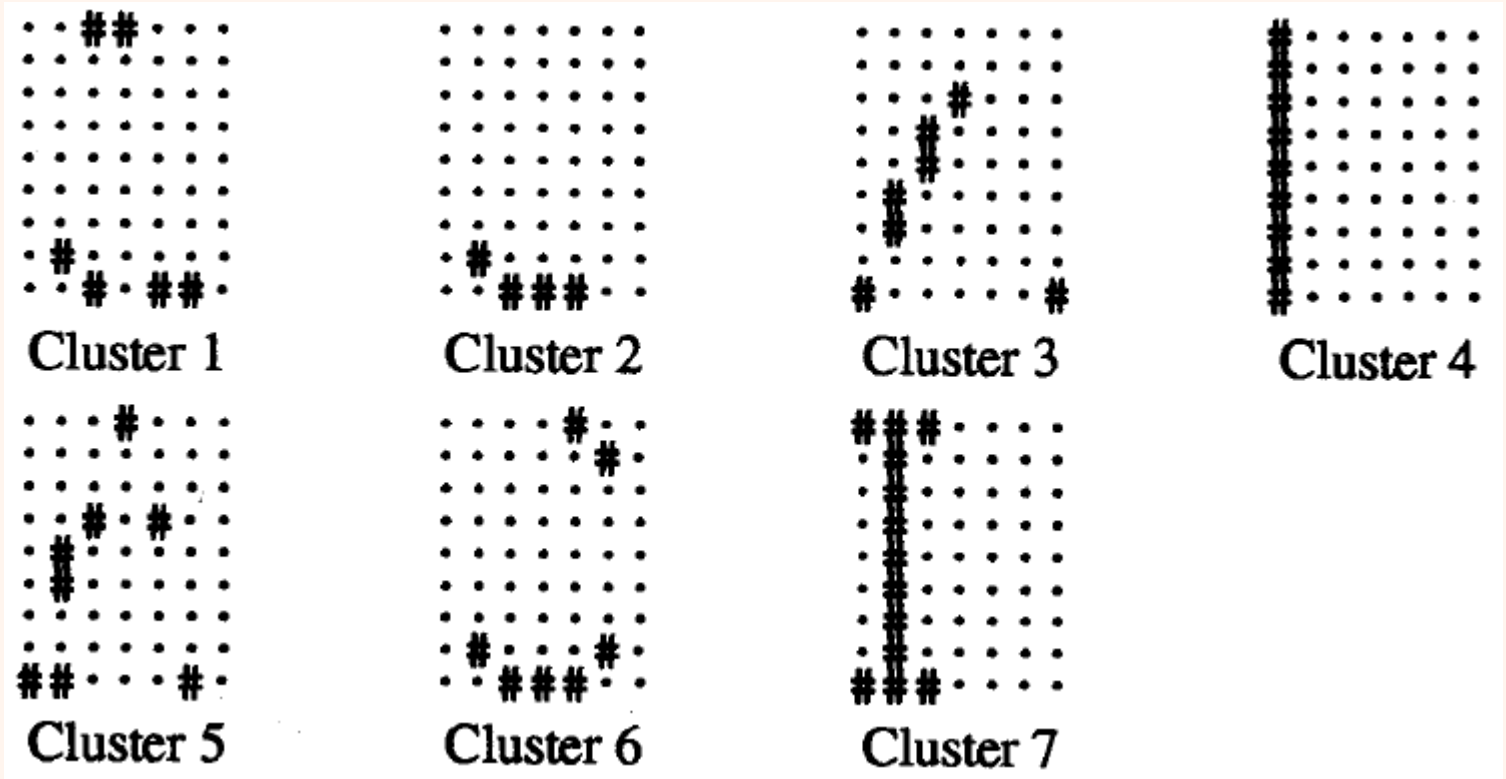
CLUSTER	EPOCH 1	EPOCH 2
1	A1, B1, C1	C1
2	D1, E1, J1 C2, J2	J2
3	K1, A2	A1, A2
4	B2, D2, E2, K2	B2, D2, E2, K2
5	A3, B3, E3	A3
6	C3, D3, J3	J1, C2, C3, J3
7	K3	B1, D1, E1, K1 B3, D3, E3, K3





شبکه عصبی

- وزن‌نهایی

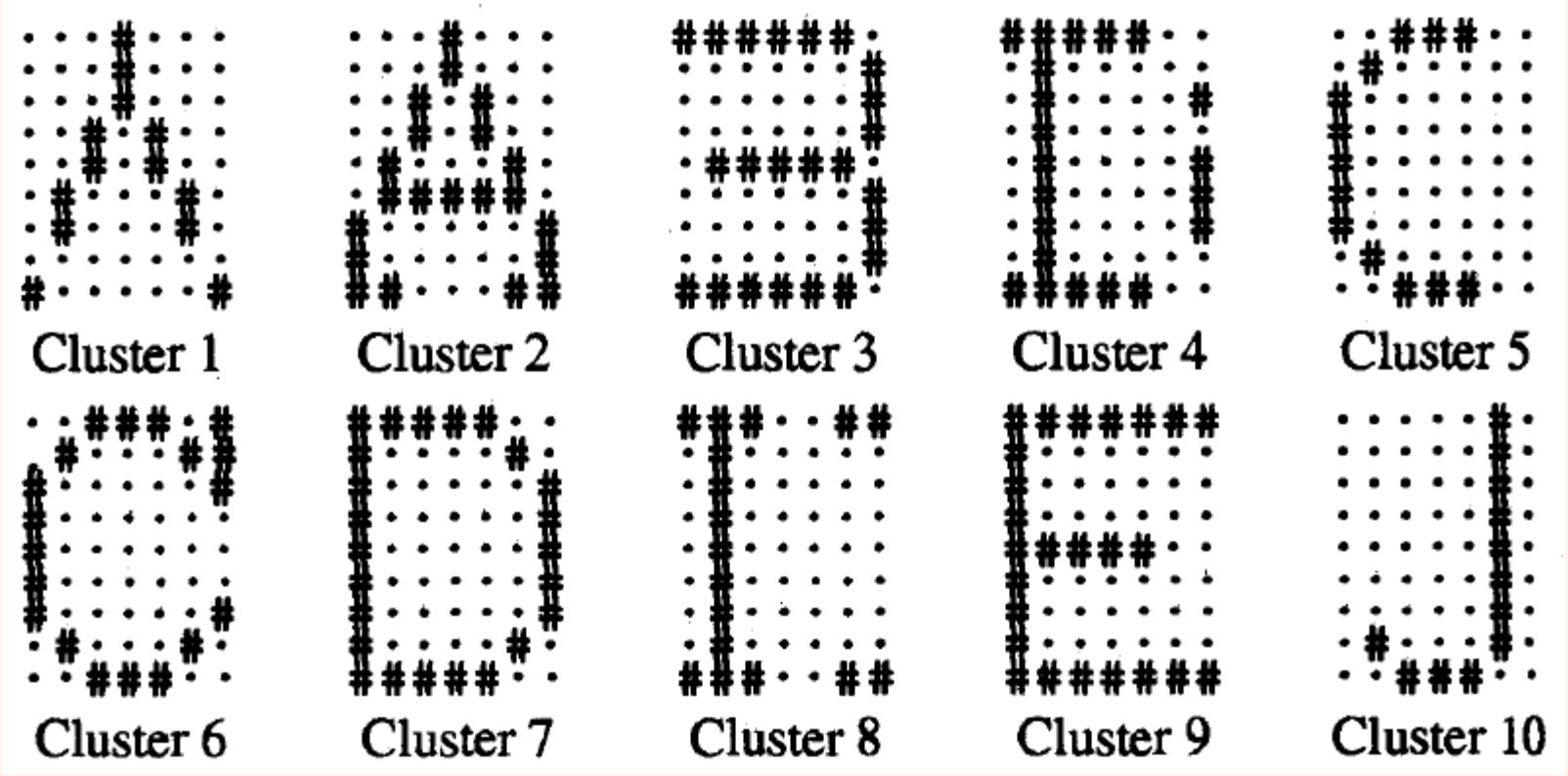

Cluster 1
Cluster 2
Cluster 3
Cluster 4
Cluster 5
Cluster 6
Cluster 7

- با انتخاب ۱۰ گروه و با قرار دادن پارامتر vigilance=0.7 پس از دو مرحله سیستم به پایداری می‌رسد.

A1, A2, A3, B1, B2, B3, C1, C2, C3, D1, D2, D3, E1, E2, E3, J1, J2, J3, K1, K2, K3

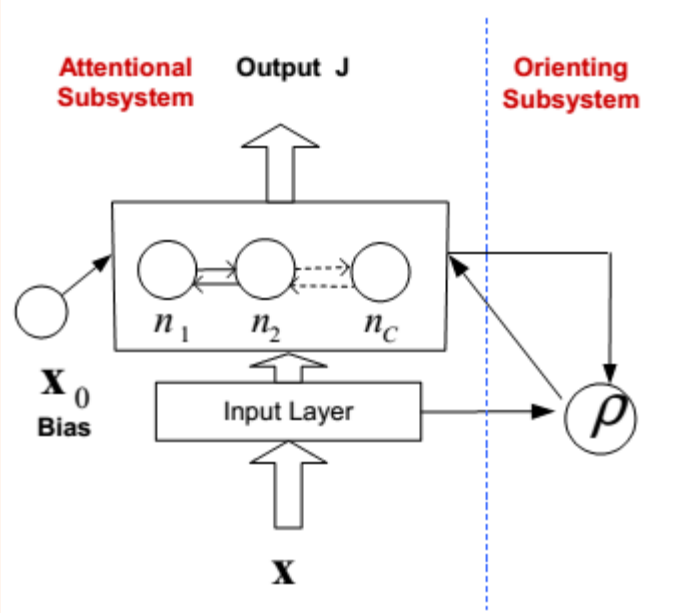
CLUSTER	EPOCH 1	EPOCH 2
1	A1, A2	A2
2	A3	A3
3	B1, B2	B1, B2
4	B3, D1, D3	B3, D1, D3
5	C1, C2, K2	C1, C2
6	C3	C3
7	D2	D2
8	E1, E3, K1, K3	E1, E3
9	E2	E2
10	J1, J2, J3	J1, J2, J3
CNC		A1, K2

Using a higher vigilance parameter of 0.70, but still allowing a maximum of only 10 cluster units, some patterns cannot be placed on clusters. (These are shown as CNC, for "could not cluster,"


Cluster 1
Cluster 2
Cluster 3
Cluster 4
Cluster 5
Cluster 6
Cluster 7
Cluster 8
Cluster 9
Cluster 10

ساختار ساده شده

- ساختار مطرح شده برای این شبکه به صورت وزن‌های روبه‌بالا و روبه‌پایین قابل ساده‌سازی با یک شبکه با وزن‌های روبه جلو است.



Architecture of a simplified ART

Baraldi, A. and E. Alpaydın, Simplified ART: A new class of ART algorithms. 1998.

ویرایش‌های مختلف ART

- ویرایش‌های مختلفی از این شبکه ارائه شده است که مهمترین آن‌ها ART2 برای داده‌های پیوسته، ARTMAP نمونه‌ی بانظارت، Fuzzy ARTMAP می‌باشد.
- همچنین نمونه‌های متنوعی توسط سایر پژوهشگران مطرح شده است:
  - Adaptive Hamming Net(AHN)
  - Gaussian ART(GA)
  - Simplified Fuzzy ARTMAP(SART)
  - μ ARTMAP
  - .....





الگوریتم ART2

- در این الگوریتم از مقادیر پیوسته استفاده می‌شود.
- وقتی پیوسته است برای اصلاح نمی‌توان از and استفاده کرد.
- معمولاً هنگامی‌که ورودی‌های نویزی داریم از این الگوریتم استفاده می‌شود.
- فرضیه‌ی اولیه به گونه‌ای است که p اگر اندازه‌ی ورودی باشد، یک پارامتر تعریف می‌شود که در صورتی‌که هر مولفه‌ی بردار ورودی از پارامتر θ کم‌تر باشد را به صفر می‌بریم.
- پس از صفر کردن بردارها را نرمال می‌کنیم.





ادامه...

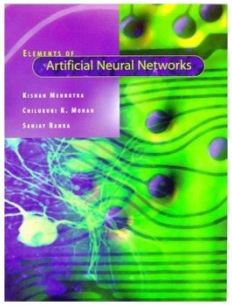
- الگوریتم ART2 شبیه به ART1 است با این تفاوت که در لایه‌ی ورودی به سبب داشتن ورودی‌های پیوسته می‌باید فرآیندهایی چون نرمالیزه کردن و کاهش نویز صورت پذیرد.
- در لایه‌ی دوم یا همان لایه‌ی رقابت همانند ART1 عمل می‌کند.





منابع مورد استفاده

- عمده‌ی مطالب این بخش برگرفته از بخش ۵.۴ کتاب زیر است:



- Elements of Artificial Neural Networks
  Kishan Mehrotra, Chilukuri K. Mohan and Sanjay Rank

- همچنین از مثال‌های کتاب Fausett نیز استفاده شده است.