

December 28, 2014

شبکه های عصبی - فصل سوم

Hadi Zayandehroodi



شبکه های عصبی



موضوعات

• فصل دوم (ادامه)

- شبکه های پسخور یا برگشتی
- مسائل حل شده فصل ۲ (خوانده شود و بحث در کلاس حل تمرین)

• فصل سوم: مسئله تشریحی شناسایی الگو

- شناسایی الگو
- معرفی سه شبکه های نمونه: پرسپترون، همینگ و هاپفیلد
- استفاده از آن در حل مسئله فوق



فصل ۳

مسئله تشریحی

(شناسایی الگو)

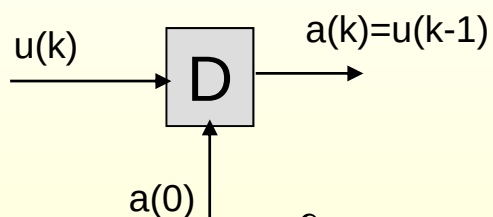


شبکه های پسخور یا برگشتی (ادامه فصل ۲)

- تفاوت با شبکه های پیشخور:

— وجود حداقل یک سیگنال برگشتی از یک نرون به همان نرون یا نرونهای همان لایه و یا لایه های قبل

- استفاده از بلوک تاخیر D (در شبکه های گسسته)



- سوال: بلوک متناظر با عنصر تاخیر در شبکه های پیوسته چیست؟
- نکته:

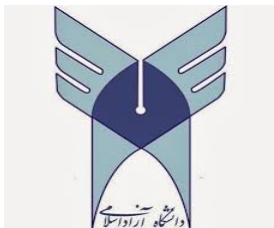
شبکه های پسخور از توانایی بالقوه بیشتری نسبت به شبکه های پیشخور برخوردارند و بهتر می توانند رفتار مربوط به ویژگیهای زمانی سیستمها را نشان دهند. (چرا؟)

- در فصل سوم کمی بیشتر با این شبکه ها آشنا می شویم. توضیحات مفصلتر در فصول ۵ و ۶ کتاب آمده است.



شناسایی الگو

- فاکتور مهم در طراحی سیستمهای اطلاعاتی
- موضوع مشترک تحقیقاتی برای اکثر رشته ها از زبان شناسی تا رشته های مهندسی
- ساختار کلی یک فرایند شناسایی الگو و تصمیم گیری:
 - چیزی یا موضوعی اتفاق افتاده است
 - اتفاق، توسط سیگنالی نمایندگی می شود.
 - سیگنال مشاهده شده با سیگنال ارسالی یکسان نیست.
 - اتخاذ تصمیم بر مبنای سیگنال مشاهده شده (نه ارسالی)
- مثال: (فرد بیمار)
 - اتفاق: بیماری
 - مشاهده کننده : پزشک
 - سیگنالها : تستهای آزمایشگاهی یا علائم بیماری
 - سیگنال آمیخته با نویز: اشکالات آزمایشی (چه در وسایل و چه در بدن فرد بیمار)
 - تشخیص بیماری بر مبنای سیگنال آمیخته با نویز



شناسایی الگو

- شناسایی الگو چیزی جز جدا سازی داده ها یا الگو های ورودی بین دستجات مختلف نیست.
- مراحل طراحی یک سیستم شناسایی الگو
 - **کد گذاری:** بیان ویژگیهای شی مورد نظر (که از طریق اندازه گیری به دست می آیند)، توسط بردارهای ورودی
 - بیان هندسی:** تخصیص یک نقطه در فضای چند بعدی اقلیدسی به هر شی
 - **استخراج شاخص:** استخراج مشخصه های مهم از روی بردار های ورودی و کاهش ابعاد الگوها
 - عناصری از بردارهای مشخصه که در تمامی طبقات مشترکند، حذف می شوند.
 - **طبقه بندی الگوها:** روندی برای تصمیم گیری بهینه اتوماتیک
- تخصیص الگوهای ورودی به یکی از طبقاتی که فضای اقلیدسی برای تصمیم گیری به تعداد متناهی از آنها تقسیم شده است



روشهای کلاسیک و شبکه های عصبی

- روشهای کلاسیک

- باید تمامی الگوهای یادگیری قبل از حل مسئله طبقه بندی الگو در اختیار باشند.
- یعنی، ابتدا باید یک **مدل ریاضی** از مشاهدات داشت تا پس از **ارزیابی مدلها** بر اساس داده های واقعی طراحی انجام شود.

- شبکه های عصبی

- مستقیماً با داده های واقعی کار می کنند.
- لذا به طراحیهای **مدل آزاد** یا **تخمین زننده های جهانی مدل آزاد** موسومند.

- سوال:

اطلاعاتی که شبکه های عصبی بر مبنای آن عمل تصمیم گیری را انجام می دهد، چگونه و در کجا ذخیره می گردد؟



صورت مسئله

- **هدف:** جداسازی سه نوع میوه سیب، پرتقال و گلابی در یک انبار به طور خودکار
- سنسورهای موجود برای ثبت ویژگی میوه ها
 - شکل :
 - گرد $\leftarrow 1$ ، بیضوی $\leftarrow -1$
 - زبری و صافی سطح :
 - صاف $\leftarrow 1$ ، زبر $\leftarrow -1$
 - وزن :
 - کمتر از ۲۰۰ گرم $\leftarrow 1$ ، بیشتر از ۲۰۰ گرم $\leftarrow -1$
- **کد گذاری**

نشان دادن **سیب**، **پرتقال** و **گلابی** به ترتیب با بردارهای $\underline{p}^1$ ، $\underline{p}^2$ ، $\underline{p}^3$ به عنوان بردارهای مرجع (مولفه بردارها به ترتیب شکل، زبری و وزن می باشد)

$$\underline{p}^1 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \quad \underline{p}^2 = \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix} \quad \underline{p}^3 = \begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix}$$



صورت مسئله

- استخراج شاخص

عنصر سوم در همه طبقات مشترک است. لذا با حذف این مولفه به بردارهای دو بعدی می‌رسیم.

$$\underline{p}^1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \quad \underline{p}^2 = \begin{bmatrix} 1 \\ -1 \end{bmatrix} \quad \underline{p}^3 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$$

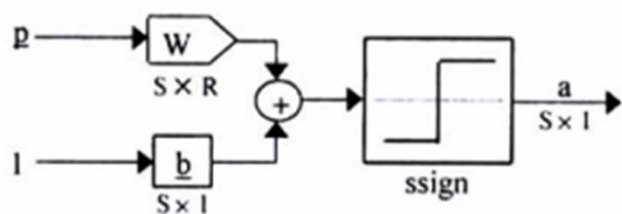
- یک مسئله ساده شناسایی الگوی باینری

- دریافت یک بردار ورودی ۲ بعدی و تصمیم‌گیری درباره نوع میوه
- معرفی سه شبکه عصبی پرسپترون (پیشخور)، همینگ (رقابتی) و هاپفیلد (حافظه انجمنی بازگشتی) و استفاده از آنها در حل مسئله



پرسپترون تک لایه

با تابع تحریک آستانه ای دو مقداره متقارن



پرسپترون تک لایه

- ساختار در فصل پیش دیده شد.
- توانایی در تفکیک الگوهای ورودی
- جداسازی الگوهای خطی

– مرز جدا ساز نواحی در حالت **تک نرونی**

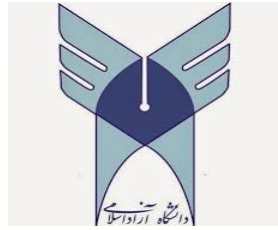
یک فوق صفحه و برای حالت دو بعدی یک خط راست است.

$$b + \sum_n w_n p_n = 0$$

– در حالت کلی برای هر نقطه دلخواه q روی ناحیه مرزی داریم:

$$\langle \underline{q}, \underline{w} \rangle = -b$$

– $\langle \underline{a}, \underline{b} \rangle$ نماد ضرب داخلی دو بردار بوده و همواره تصویر $\underline{b}$ روی $\underline{a}$ مضربی از ضرب داخلی دو بردار است.



پرسپترون تک لایه

- **تذکر ۱:** خط مرزی همواره **عمود بر بردار وزن** بوده و محل قرار گرفتن مرز توسط بردار بایاس تعیین می شود.

- **تذکر ۲:** تصویر بردار سازنده هر نقطه روی خط مرزی بر بردار وزن، برابر با مضربی از منفی جمله بایاس ($-b$) است. (چرا؟)

- **نتیجه:** برای تمام نقاطی که **بالای**

(با توجه به جهت بردار وزن) خط مرزی قرار

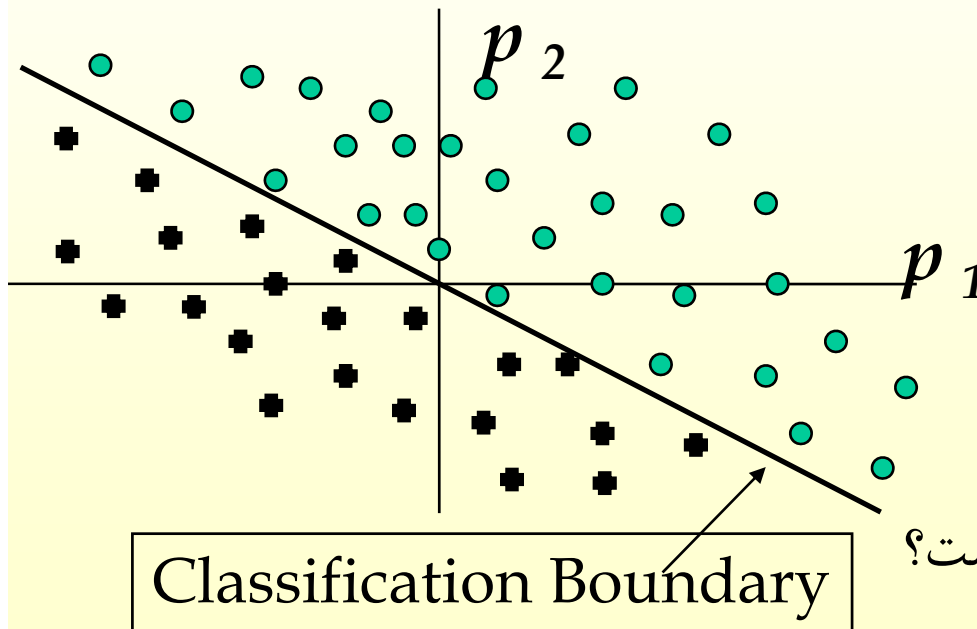
دارند، ورودی خالص نرون **مثبت** است

و برای تمام نقاط زیر آن منفی می باشد.

- **سوال**

در شکل روبرو ورودی خالص نرون چگونه است؟

آیا پاسخ یکتاست؟





پرسپترون تک لایه

- در حالتی که شبکه دارای S نرون میانی است، ماتریس وزن دارای S سطر بوده و هر سطر یک خط (فوق صفحه) مرزی را تعیین می کند.

- سوال

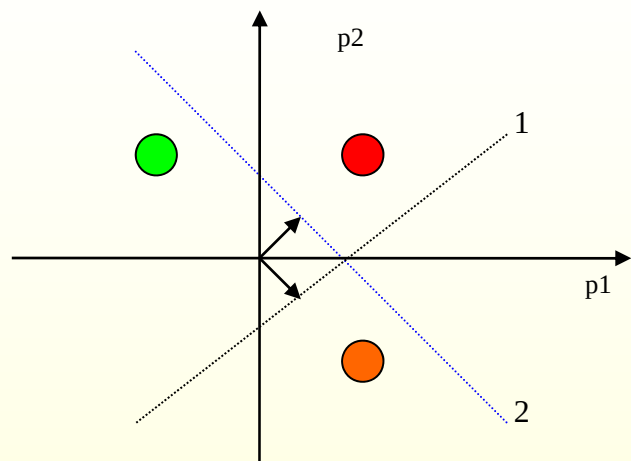
حداکثر تعداد نواحی که به طور خطی از هم تفکیک پذیرند برای پرسپترون با S نرون میانی و بردار ورودی دارای n عنصر، چه مقدار است؟

- مثال تشریحی شناسایی الگو

- چون دارای ۳ گروه هستیم، از ۲ نرون میانی استفاده می کنیم (چرا؟). ماتری وزن W نیز 2×2 خواهد بود.
- با توجه به شکل مشخص است که جوابهای بیشمار وجود دارد. یکی از آنها نمایش داده شده است



مثال تشریحی با پرسپترون



- $W(1,:) = [1, -1]$

- $W(2,:) = [1 \ 1]$

- $\underline{b} = [-1; -1]$

- به این ترتیب ناحیه قرمز (مربوط به سیب) دارای

خروجی $a = [-1; 1]$

ناحیه نارنجی (مربوط به پرتقال) خروجی $a = [1; -1]$

ناحیه سبز (مربوط به گلابی) خروجی $a = [-1; -1]$

می باشد.

تمرین ۱: درستی این بردارهای وزنی و بایاسها چک شود.

تمرین ۲: شبکه ورودی پرتقال بیضوی را چگونه ارزیابی می کند؟ چرا؟



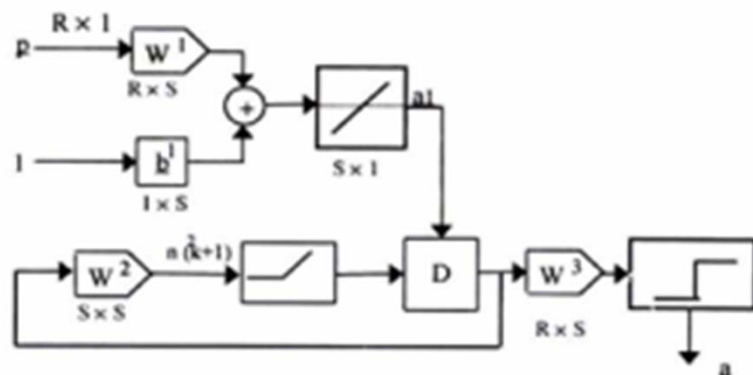
سوال

۱) آیا می توان همواره دسته ها را در فضای ورودی توسط پرسپترون تک لایه مجزا کرد؟

۲) آیا روند به کار رفته برای یافتن وزنهای شبکه پرسپترون تک لایه برای مسائل با ابعاد وسیعتر نیز قابل کاربرد است؟



شبکه همینگ



شبکه همینگ

- دارای هر دو ساختار پیشخور و پسخور
- دارای ۳ لایه
- برای شناسایی الگوهای باینری
- تشخیص اینکه کدام الگوی مرجع بیشترین نزدیکی را به الگوی ورودی دارد.

- اگر به شبکه ورودی p^i (یکی از الگوهای مرجع) اعمال شود، خروجی نیز p^i خواهد بود.
- در مواردی که به شبکه بردار ورودی اختیاری (مثلاً پرتقال بیضوی) اعمال می شود انتظار بر این است که الگوی مرجعی که بیشترین نزدیکی را با ورودی دارد، در خروجی ظاهر شود.
- دو پروسه محاسباتی: ذخیره سازی الگوها و بازیابی اطلاعات انجام می گیرد.



شبکه همینگ : لایه اول

- شبکه پیشخور با تابع تبدیل خطی
- همبستگی یا ضرب داخلی بین بردارهای مرجع با بردار ورودی را محاسبه می کند.
بردارهای مرجع الگوهایی هستند که قصد شناسایی آنها را داریم.

$$W^1 = \begin{bmatrix} (\underline{p}^1)^T \\ (\underline{p}^2)^T \\ (\underline{p}^3)^T \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & -1 \\ -1 & 1 \end{bmatrix}$$

- ذخیره سازی الگوهای مرجع در لایه اول
- تعداد نورونهای این لایه برابر تعداد الگوها
- R تعداد عناصر بردار ورودی
- به نوعی **فاصله همینگ** سنجیده می شود.

$$p_H^j = 1/2 \sum_{i=1}^R | \underline{p}_i^j - \underline{p}_i |$$



شبکه همینگ : لایه اول

- تمامی عناصر بردار خروجی لایه اول مقادیری بین 0 و $2R$ دارند. چرا؟
- نرون با بزرگترین مقدار خروجی متناظر با الگوی مرجعی خواهد بود که بیشترین شباهت (کمترین فاصله همینگ) را با الگوی ورودی دارد. چرا؟
- در مرحله بعد باید خروجی ماکزیمم از بقیه جدا گردد. این عمل توسط لایه برگشتی انجام می شود.



شبکه همینگ : لایه دوم

- لایه برگشتی (لایه WTA) (لایه رقابتی)

$$\underline{a}^2(0) = \underline{a}^1$$

$$\underline{a}^2(k+1) = \text{posl}(W^2 \underline{a}^2(k))$$

$$\text{posl}(x) = \begin{cases} x & x \geq 0 \\ 0 & x < 0 \end{cases}$$

$$w_{ij}^2 = \begin{cases} 1 & i = j \\ -\varepsilon & i \neq j \end{cases}, 0 < \varepsilon < 1$$

- انتخاب ماتریس W^2

$$\varepsilon = \frac{1}{s-1}$$

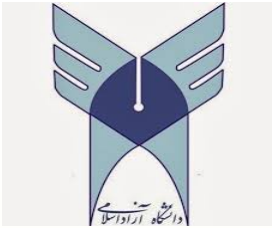
- برای رسیدن به سرعت همگرایی بیشتر معمولاً داریم:
- دلیل انتخاب ماتریس W^2 به شرح فوق، در بخش اثبات همگرایی شبکه همینگ در فصل ۳ آمده است.
- تمرین: تحقیق کنید که خروجی لایه دوم به ورودی سیب، برابر $[2;0;0]$ می باشد.



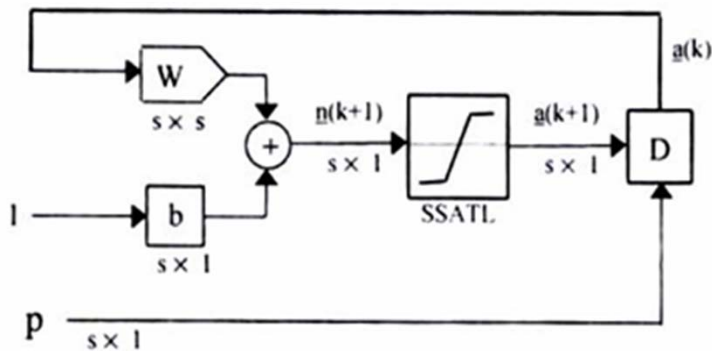
شبکه همینگ : لایه سوم

- شبکه پیشخور با تابع تحریک آستانه متقارن دو مقداره
- پس از همگرایی لایه دوم، بردار مرجع ذخیره شده در خروجی شبکه ظاهر می گردد.
- **بازیابی اطلاعات**
- انتخاب وزنهای لایه سوم:
 $w_{ij}^3 = p_i^j$
- – P_i^j عنصر i ام از الگوی مرجع j ام می باشد.
- مثلاً اگر ورودی سیب باشد، خروجی شبکه الگوی مربوط به سیب یعنی $[1;1]$ خواهد بود. **بررسی شود.**
- **توجه:**

حتماً باید لایه دوم همگرا شده باشد



شبکه هاپفیلد



شبکه هاپفیلد

- شبکه تک لایه بازگشتی
- عملکرد همه نرونها مشابه یکدیگر
- عدم وجود نرون ورودی یا خروجی
- مقادیر اولیه نرونها توسط ورودی تعیین شده و سپس شبکه با تکرار به یکی از الگوهای مرجع همگرا می شود.

$$\underline{a}(k+1) = ssatl(\underline{w}\underline{a}(k) + \underline{b})$$

$$\underline{a}(0) = \underline{p}$$

$$ssatl(x) = \begin{cases} -1 & x < -1 \\ x & -1 \leq x \leq 1 \\ 1 & x > 1 \end{cases}$$



شبکه هاپفیلد

- انتظار داریم که اگر در لحظه 0 به شبکه الگوی سیب [1;1] اعمال شود، شبکه به سمت [1;1] همگرا گردد.
- یک انتخاب خوب برای b , W عبارتست از:

$$W = \begin{bmatrix} 1.2 & 0 \\ 0 & 1.5 \end{bmatrix} \quad \underline{b} = \begin{bmatrix} 0.1 \\ 0.2 \end{bmatrix}$$

- مسلماً برای مسائل مشکلتر نیاز به **روند طراحی** وزن‌ها می باشد
- **تمرین:** پاسخ شبکه به ورودیهای سیب، پرتقال و گلابی بررسی شود.
- **سوال:** پاسخ شبکه به ورودی پرتقال بیضوی چیست؟
- **عیب شبکه هاپفیلد:** همگرایی به الگویی غیر از الگوهای مرجع



مهمترین ویژگیها

- شبکه پرسپترون جواب را کد می کند.
- شبکه همینگ به پاسخ مناسب منتهی می شود و همواره به یکی از الگوهای مرجع همگرا می شود، و الگوی دارای بیشترین تشابه به ورودی را بر می گزیند.
- شبکه هاپفیلد به پاسخ همگرا می شود که ممکن است جزو الگوهای ذخیره شده نباشد.



مهمترین نکات

- انتخاب مناسب پارامترها
 - انواع الگوریتمهای یادگیری برای تنظیم پارامترها از تفاوت‌های مهم انواع شبکه‌ها با یکدیگر است
- انتخاب ساختار مناسب شبکه‌های عصبی
 - چه زمانی باید از ساختارهای پیش‌خور یا پس‌خور استفاده کرد؟
 - چگونه می‌توان فهمید شبکه‌های پس‌خور نهایتاً همگرا می‌شوند؟
 - پایداری کدامیک از شبکه‌های پیش‌خور یا پس‌خور راحت‌تر تضمین می‌شود؟
 - آیا می‌توان وزن‌ها و بایاس‌های شبکه همینگ را، وقتی که الگوهای مرجع را نمیدانیم، آموزش داد؟