

○

محاسبه $P(A|w)$

چهار روش DTW، مدل مخفی مارکوف، شبکه های عصبی و سیستم های ترکیبی

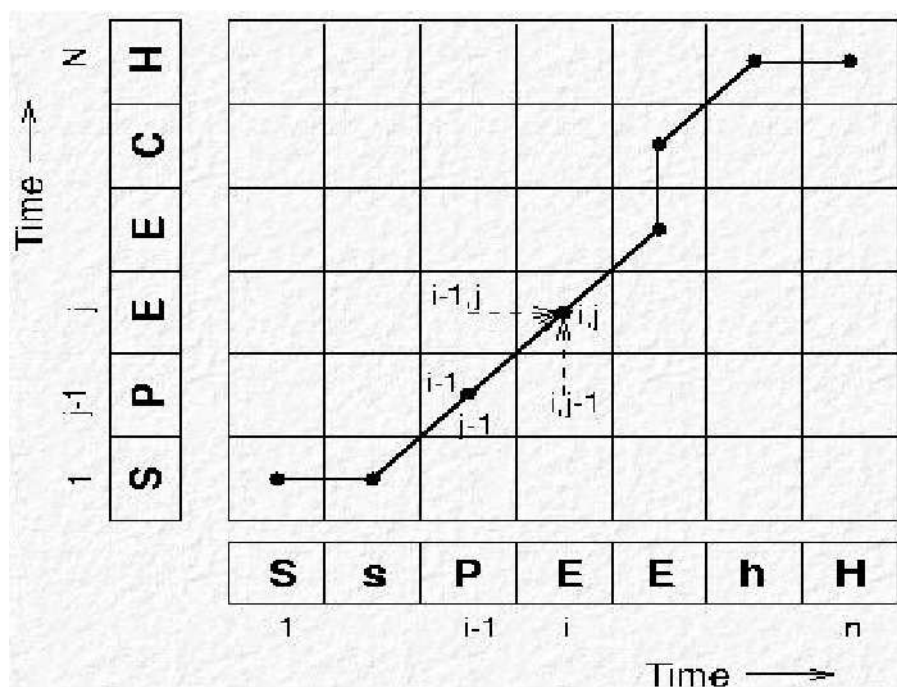
1. Dynamic Time Warping

در این روش سعی می شود که فاصله کلی بین دو نمونه گفتار محاسبه شود.

به عبارتی یک فایل صوتی چه فاصله ای با یک فایل صوتی دیگر دارد.

برای فایل های صوتی نیاز به یک تطبیق زمانی است.

مثال: تطبیق زمانی بین دو کلمه تلفظ شده speech به روش DTW در تصویر 6 نشان داده شده است.

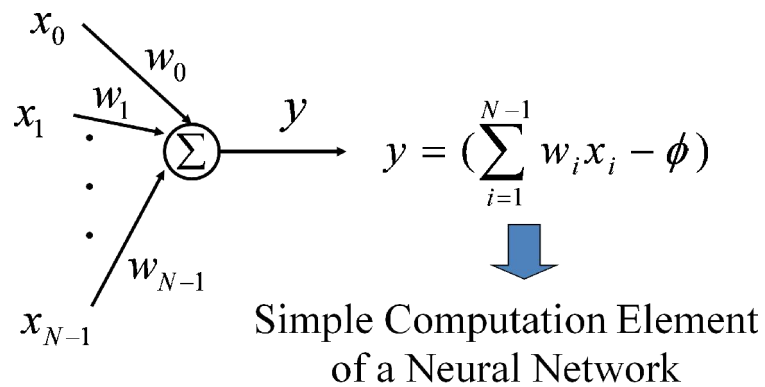


تصویر 6 - انجام تطبیق زمانی بین دو تلفظ speech

2. Artificial Neural Network

یکی از روش های محاسبه $P(A|w)$ استفاده از شبکه عصبی است.

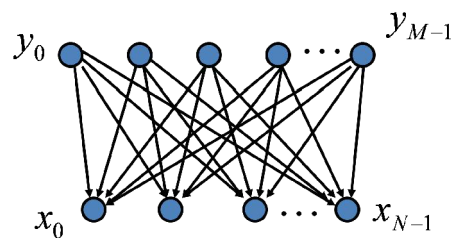
پایه ای ترین واحد شبکه عصبی نرون گفته می شود. یک نرون در شکل 7 نشان داده شده است.



شبکه عصبی نوع های ،

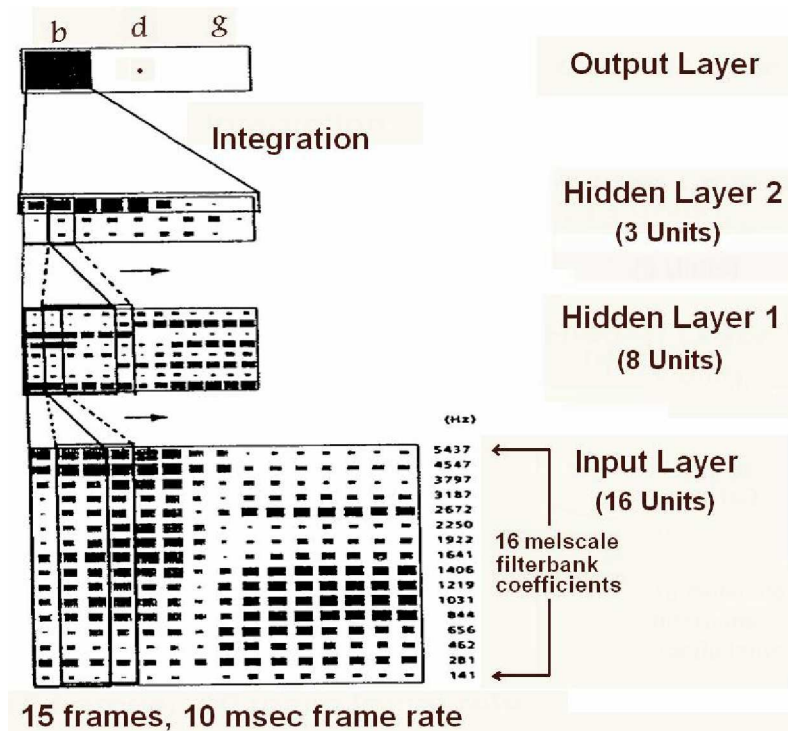
○ پرسپترون تک

Single Layer Perceptron



تصویر 8 - پرسپترون تک لایه

○ شبکه عصبی با تاخیر زمانی (TDNN): در تصویر 9 یک شبکه عصبی تاخیر دار بازنمایی واج های ب، د و گ ارائه شده است.



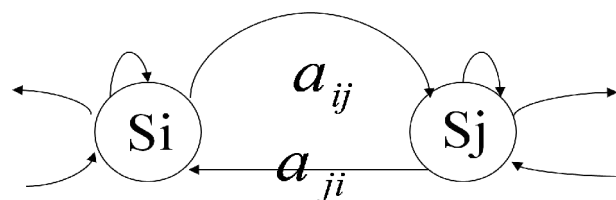
تصویر 9 - یک شبکه عصبی TDNN برای بازشناسی ب، د و گ

3. Hidden Markov Models

مدل های مخفی مارکوف بیشترین استفاده را برای محاسبه $P(A|w)$ استفاده می شود.

این مباحث در فصول آینده توضیح داده شده اند.

یک مدل مخفی ساده را در تصویر 10 مشاهده می کنید.



تصویر 10 - نمونه ای از مدل مخفی مارکوف

مدل مخفی مارکوف دارای یک سری پارامترها است.

- مشاهدات (observation)
- حالات (states)
- پرش بین حالات
- احتمال تولید یک مشاهده در یک حالت



3- حالات مختلف بازشناسی

- بازشناسی کلمات گسسته (Isolated Word Recognition) در مقابل بازشناسی گفتار پیوسته (Continuous Speech Recognition)
- وابسته به گوینده و مستقل از گوینده
- اندازه فرهنگ لغت
 - کوچک (کمتر از 100 لغت)
 - متوسط (بین 100 تا 1000 لغت)
 - بزرگ (بین 1000 تا 10000 لغت)
 - خیلی بزرگ (بزرگتر از 10000 لغت)
- فاکتورهای تولید خطا
 - پروزودی (بازشناسی باید مستقل از پروزودی باشد)
 - نویز (باید از نویز جلوگیری کند)
 - باید قابلیت بازشناسی گفتار محاوره ای را داشته باشد.

4- خودآزمایی

شماره سوال	نوع سوال	صورت سؤال	متن گزینه ها	پاسخ درست	مهلت پاسخگویی	اجازه عبور به قسمتهای بعدی در صورت اشتباه بودن پاسخ
1	تستی <input type="checkbox"/> جاخالی <input type="checkbox"/>	P(w) نشان دهنده است.		مدل زبانی	1 دقیقه	بله <input checked="" type="checkbox"/> خیر <input type="checkbox"/>

5 - خلاصه و نتیجه گیری:

در این فصل با بحث اولیه بازشناسی گفتار آشنا شدیم. ازجمله با مسئله بازشناسی گفتار به صورت احتمالی آشنا شدیم.



سپس یادگرفتیم که هر قسمت از احتمال ها ($P(A|w)$ و $P(w)$) را به روش هایی می توان محاسبه کرد.

- $P(w)$ را می توان بوسیله مدل زمانی محاسبه کرد.
- $P(A|w)$ را می توان بوسیله DTW، HMM، شبکه عصبی و روش های ترکیبی این ها محاسبه نمود.

6 – منابع درس:

- 1- Rabiner, “Fundamentals of Speech Recognition”
- 2- Huang, Acero, “Spoken Language Processing”
- 3- Deller, “Discrete-time processing of speech signals”