فصل نهم

مدل مخفی مارکوف

1-7: مقدمه

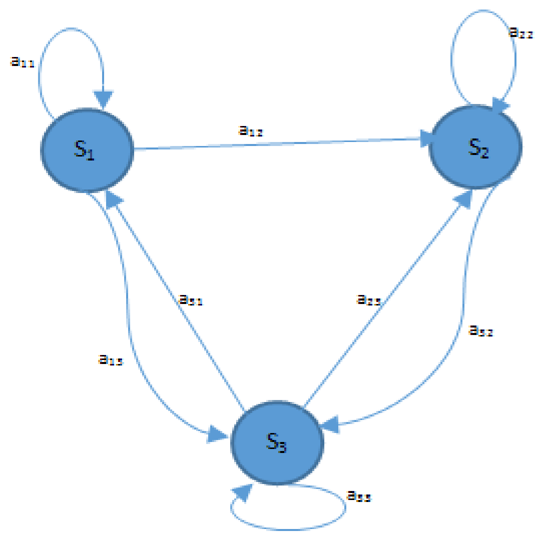
فرض کنید که شما در یک اتاق هستید و در اتاق دیگری، فـردی سـکه‌هـایی را بـه هـوا پرتـاب می‌کند. وی به شما نتایج هر پرتاب را می‌گوید، اما از چگونگی پرتاب سکه‌ها و سالم بودن آن‌هـا حرفی نمی زند. در این حالت شما با فرایند مخفی انداختن سکه‌هـا و بـا دنبالـه‌ای از مشـاهدات شیر یا خط مواجه هستید. مساله‌ای که اینجا مطـرح مـی‌شـود چگـونگی سـاختن مـدل مخفـی مارکوف به منظور بیان این فرایند تصادفی است.

قبل از تشکیل مدل مخفی مارکوف، لازم است مطالبی در مورد مدل مارکوف گسسته بدانیم.

2-7: فرآیند مارکوف گسسته

یک سیستم مانند شکل زیر را که در هر لحظه در یکی از حالات $S_1,S_2,S_3,\ldots,S_N$ است در نظر بگیرید. در زمان‌های گسسته و با فواصل منظم، حالت سیستم با توجه به مجموعه ای از احتمالات تغییر می‌کند. برای زمان‌های $t = 1,2,3,\ldots$، حالت در لحظه t را با q نشان می دهیم. برای یک توصیف مناسب از سیستم فعلی نیاز به دانستن حالت فعلی در کنار تمام حالات قبلی می باشد. برای یک حالت خاص از زنجیره مارکوف مرتبه اول، توصیف احتمالاتی تنها با حالت فعلی و حالت قبلی مشخص می شود.

رابطه (1-7) = = = ,... = , = =



شکل7-1 زنجیره مارکوف

اگر فرض کنیم رابطه 7- 1 مستقل از زمان باشد، داریم:

$$\forall \;,\; : \quad = \quad = \quad = \qquad \text{(رابطه 7-2)}$$

و نیز اگر داشته باشیم:

$$\geq 0 \qquad \sum \quad = 1 \qquad \text{(رابطه 7-3)}$$

فرایند تصادفی فوق را مدل مارکوف قابل مشاهده می‌گویند زیرا خروجی مـدل، مجموعـه‌ای از حالات است که قرار گرفتن در آن‌ها متناظر با یک مشـاهده مـی باشـد. مـا مـی تـوانیم دنبالـه مشاهدات مورد انتظار خود را تولید کنـیم و احتمـال وقـوع آن در زنجیـره مـارکوف را محاسـبه نماییم.

طول حافظه‌ای که مقادیر احتمالِ ممکن برای حالت بعـدی بـه کمـک آن محاسـبه مـی‌شـود را مرتبه مدل مارکوف می نامند. برای مثال، حالت بعدی در یک مـدل مـارکوف از درجـه 2 (مـدل

مارکوف مرتبه دوم) به دو حالت قبلی آن بستگی دارد. همچنین طرف سمت راست رابطه 7-1 بیانگر یک مدل مارکوف مرتبه اول است. برای مثال فرایند انداختن یک سکه ناسالم را می‌توان با مدل مارکوف مرتبه صفر بیان نمود زیرا حافظه این فرایند صفر می باشد.

7-3: مدل مخفی مارکوف (HMM)

تا اینجا ما مدل مارکوف را که در آن هر حالت، متناظر با یک رویداد قابل مشاهده بود معرفی نمودیم. در این بخش تعریف فوق را گسترش می‌دهیم، به این صورت که در آن، مشاهدات توابع احتمالاتی از حالت‌ها هستند. در این صورت مدل حاصل، یک مدل تصادفی با یک فرایند تصادفی مخفی است و تنها توسط مجموعه‌ای از فرایندهای تصادفی که دنباله مشاهدات را تولید می‌کنند قابل مشاهده است.

دوباره به مثال ابتدای فصل بیندیشید. اگر مشاهدات، حاصل از انداختن یک سکه باشد، می‌توان با یک مدل دو حالته مساله را بررسی نمود.

یک مدل مخفی مارکوف را می‌توان با تعیین پارامترهای زیر ایجاد نمود:

- **تعداد حالات ممکن:** تعداد حالتها در موفقیت مدل نقش به سزایی دارد و در مدل مخفی مارکوف هر حالت با یک رویداد متناظر است. برای اتصال حالت‌ها روش‌های متفاوتی وجود دارد که در عمومی‌ترین شکل، تمام حالت‌ها به یکدیگر متصل می‌شوند و از یکدیگر قابل دسترسی می‌باشند.
- **تعداد مشاهدات در هر حالت:** تعداد مشاهدات برابر است با تعداد خروجی‌هایی که سیستم مدل‌شده خواهد داشت.
- N **تعداد حالتهای مدل**
- M **تعداد سمبلهای مشاهده در الفبا**
- **ماتریس انتقال حالت:** یک مجموعه از احتمالات در بین حالت‌ها

$$\forall , : \quad = \quad = \quad =$$

که دارای شرایط زیر می باشد.

$$\geq 0$$

$$= 1$$

اگر مقدار a_{ij} برابر صفر باشد، به مفهوم عدم وجود ارتباطی بین دو حالت است.

- **توزیع احتمال مشاهدات:** یک توزیع احتمال برای هر یک از حالت‌ها داریم:

$$(\ ) = \quad = \quad = \quad 1 \leq \ \leq \ ,1 \leq \ \leq$$

که در آن بیانگر kامین سمبل مشاهده‌شده در الفبا است و بیانگر بردار پارامترهای ورودی فعلی می‌باشد. در مورد مقادیر احتمال حالت‌ها نیز شرایط موجود در نظریه احتمال باید رعایت گردند.

- **توزیع احتمال حالت آغازین:**

$$= \ (\ = \ ), \quad 1 \leq \ \leq$$

به این ترتیب ما می‌توانیم یک مدل مخفی مارکوف با توزیع احتمال گسسته را با استفاده از سه‌گانه زیر مشخص نماییم.

$$= (\ ,\ ,\ )$$

بررسی مثال زیر، مدل مخفی مارکوف را بیشتر مشخص می کند.

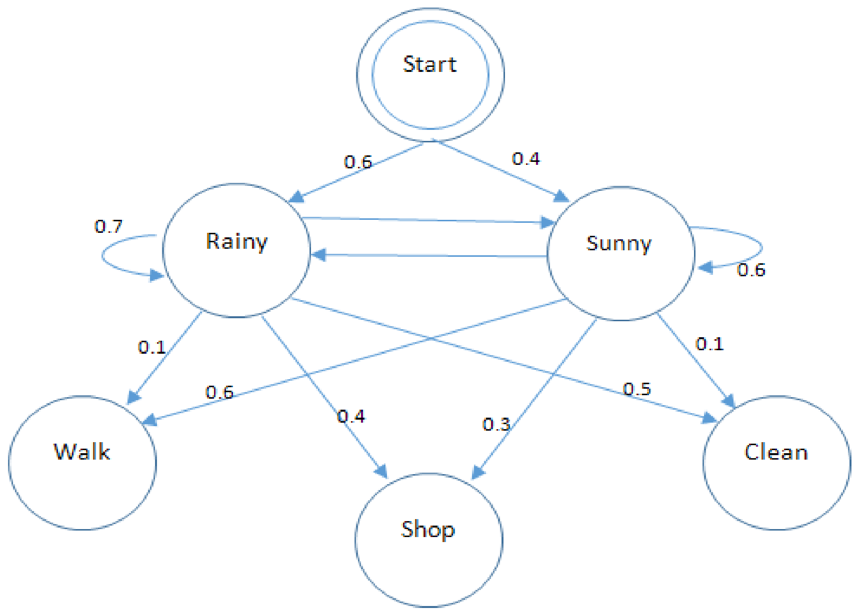
دو دوست به نام‌های آلیس و باب را در نظر بگیرید. آنها دور از هم زندگی کرده و هر روز درباره کارهای روزمره‌شان با هم تلفنی صحبت می‌کنند. فعالیت‌های باب شامل "قدم زدن در پارک"،"خرید کردن" و "تمیز کردن آپارتمان" می‌شود. انتخاب این‌که هر روز کدام کار را انجام دهد منحصرا بستگی به هوای همان روز دارد. آلیس اطلاع دقیقی از هوای محل زندگی باب

نداشته ولی از تمایلات کلی وی آگاه‌است (بنا به نوع هوا چه کاری را دوست دارد انجام دهد). بر اساس گفته‌های باب در پایان هر روز آلیس سعی می‌کند هوای آن روز را حدس بزند. آلیس هـوا را یک زنجیره گسسته مارکوف می‌پندارد که دو حالت "بـارانی" و "آفتـابی" دارد. امـا بـه طـور مستقیم هوا را مشاهده نمی‌کند. بنابرین حالات هوا بـرای او مخفـی اسـت. در هـر روز احتمـال اینکه باب به "قدم زدن"،"خرید کردن" و "تمیز کردن"بپردازد بستگی بـه هـوا داشـته و دارای یک احتمال مشخص است. مشاهدات مساله شرح فعالیت‌هایی است که باب در انتهای هر روز به آلیس می‌گوید.

آلیس اطلاعات کلی درباره هوای محل زندگی باب و این‌که در هر نوع آب و هوا با چـه احتمـالی چه کار انجام می‌دهد را دارد. به عبارت دیگر پارامترهای مساله HMM معلوم هستند که در کد زیر مشاهده می‌شوند:

```
states = ('Rainy', 'Sunny')
observations = ('walk', 'shop', 'clean')
start_probability = {'Rainy': 0.6, 'Sunny': 0.4}
transition_probability = {
  'Rainy': {'Rainy': 0.7, 'Sunny': 0.3},
  'Sunny': {'Rainy': 0.4, 'Sunny': 0.6},       }
emission_probability = {
  'Rainy': {'walk': 0.1, 'shop': 0.4, 'clean': 0.5},
  'Sunny': {'walk': 0.6, 'shop': 0.3, 'clean': 0.1},       }
```

در این کد start_probability نمایانگر احتمال رخ‌دادن هر یک از حالات (بارانی یـا آفتـابی) در روز اولی که باب با آلیس صحبت می‌کند می‌باشد. transition_probability نمایانگر تغییـر هـوا بر اساس قواعد زنجیره مارکوف است. به طور مثال اگر امروز بارانی باشد بـه احتمـال 30% فـردا آفتابی است. emition_probability نمایانگر این است که باب علاقه دارد که در هـر هـوایی چـه کار کند به طور مثال در هوای بارانی با احتمـال 50% آپارتمـانش را تمیـز کـرده و یـا در هـوای آفتابی با احتمال 60% در پارک قدم می‌زند.



شکل7-2

4-7: فرضیات تئوری مدل مخفی مارکوف

برای اینکه مدل مخفی مارکوف از لحاظ ریاضی و محاسباتی قابل بیان باشد فرض‌های زیر در مورد آن در نظر گرفته می‌شود.

- **فرض مارکوف**

با داشتن یک مدل مخفی مارکوف، احتمال انتقال از حالت **i** به حالت **j** به صورت زیر تعریف می‌شود:

$$= (\quad | \quad =)$$

به بیان دیگر فرض می‌شود که حالت بعدی تنها به حالت فعلی بستگی دارد. در واقع مدل حاصل از فرض مارکوف یک مدل مرتبه صفر می‌باشد.

در حالت کلی، حالت بعدی می‌تواند به k حالت قبلی وابسته باشد که مدل HMM مرتبه kام می‌باشد.

- **فرض ایستایی**

در اینجا فرض می‌شود که احتمال انتقال در بین حالات از زمان واقعی رخداد انتقال مستقل است. در این صورت داریم:

$$= \quad = \quad = \quad = \quad =$$

- **فرض استقلال خروجی**

در این حالت فرض می‌شود که خروجی (مشاهدات) فعلی به صورت آماری از خروجی قبلی مستقل است. می توان این فرض را با داشتن دنباله‌ای از خروجی‌ها بیان نمود:

$$= \quad , \quad , \dots ,$$
$$(\quad | \quad , \quad , \dots \quad , \quad) = \qquad (\quad | \quad , \quad)$$

اگر چه بر خلاف دو فرض دیگر این فرض اعتبار کمتری دارد. در برخی حالات این فرضیه چندان معتبر نیست و موجب می‌شود که مدل HMM با ضعف‌های عمده‌ای مواجه گردد.

5-7: مسائل پیاده سازی مدل مخفی مارکوف

برای اینکه مدل HMM در دنیای واقعی قابل استفاده باشد باید سه مساله مهم حل شود. این سه مساله عبارتند از:

- **مساله ارزیابی**[1]

با داشتن دنباله مشاهدات ,…, , = و مدل (, ,) = چگونه (|) یعنی احتمال تولید دنباله مشاهدات توسط λ، را محاسبه نماییم؟ به طور مثال در شناسایی گفتار، مساله ارزیابی برای شناسایی کلمات جدا[2] استفاده می شود.

- **مساله کدگشایی**[3]

باداشتن دنباله مشاهدات ,…, , = و مدل (, ,) = چگونه دنباله حالات بهینه ,…, , = برای تولید ,…, , = را بدست آوریم؟ شناسایی گفتار پیوسته نمونه ای از مساله کدگشایی است.

- **مساله آموزش** [4]

چگونه پارامترهای مدل (, ,) = را بدست آوریم؟ مساله آموزش نیز برای استفاده از مدل HMM در کاربردهای مختلف شناسایی گفتار استفاده می‌شود.

1-5-7: مساله ارزیابی

در این حالت با داشتن مدل (, ,) = و دنباله مشاهدات ,…, , = باید مقدار (|) را پیدا نماییم. می‌توانیم این مقدار را با روش‌های آماری مبتنی بر پارامترها محاسبه نماییم. البته این کار به محاسباتی با پیچیدگی ( ) احتیاج دارد. این تعداد محاسبات حتی برای مقادیر متوسط t نیز بسیار بزرگ است. برای حل مسئله روش دیگری ارائه شده که از متغیر کمکی ( ) با نام متغیر پیشرو استفاده می‌کند.

متغیر پیشرو به صورت یک احتمال از دنباله مشاهدات ,…, , = تعریف می‌شود که در حالت i خاتمه می یابد. به بیان ریاضی:

---

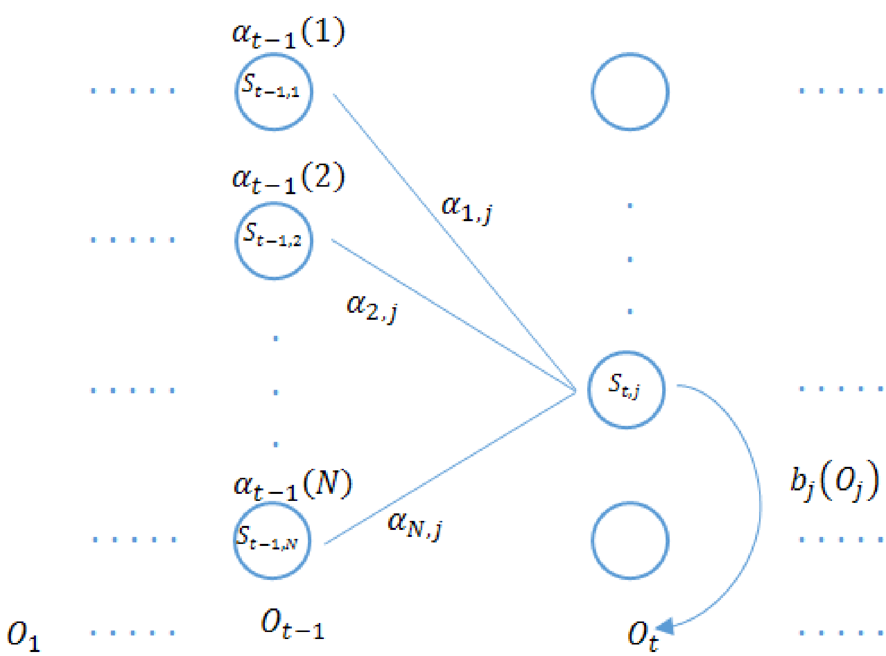
[1] Evaluation Problem
[2] isolated word recognition
[3] Decoding problem
[4] Learning problem

$$( ) = \{ \quad , \quad ,\dots, \quad , \quad = \mid \}$$

آنگاه داریم:

$$( ) = \quad (\quad) \qquad ( ) \quad ,1 \leq \quad \leq \quad ,1 \leq \quad \leq \quad -1$$

$$( ) = \qquad (\quad),1 \leq \quad \leq$$



شکل3-7

با داشتن این رابطه بازگشتی می توانیم مقدار زیر را محاسبه نماییم.

$$( ),1 \leq \quad \leq$$

و آنگاه احتمال (|) به صورت زیر محاسبه خواهد شد:

$$(\mid) = \qquad ( )$$

پیچیدگی محاسباتی روش فوق که به الگوریتم پیشرو معروف است برابر با ( ) می‌باشد، که در مقایسه با حالت محاسبه مستقیم که قبلا گفته شد و دارای پیچیدگی نمایی بود، بسیار سریعتر است.

روشی مشابه روش فوق را می توان با تعیین متغیر پسرو، ( ) ، به عنوان احتمال جزئی دنباله مشاهدات , ..., , در حالت i تعریف نمود. متغیر پیشرو را می توان به شکل زیر نمایش داد.

$$( ) = \{ \quad , \quad , \ldots , \quad \mid \quad = \; , \; \}$$

مانند روش پیشرو یک رابطه بازگشتی به شکل زیر برای محاسبه ( ) وجود دارد.

$$( ) = \qquad ( ) \qquad (\qquad) , 1 \leq \quad \leq \quad , 1 \leq \quad \leq \quad - 1$$

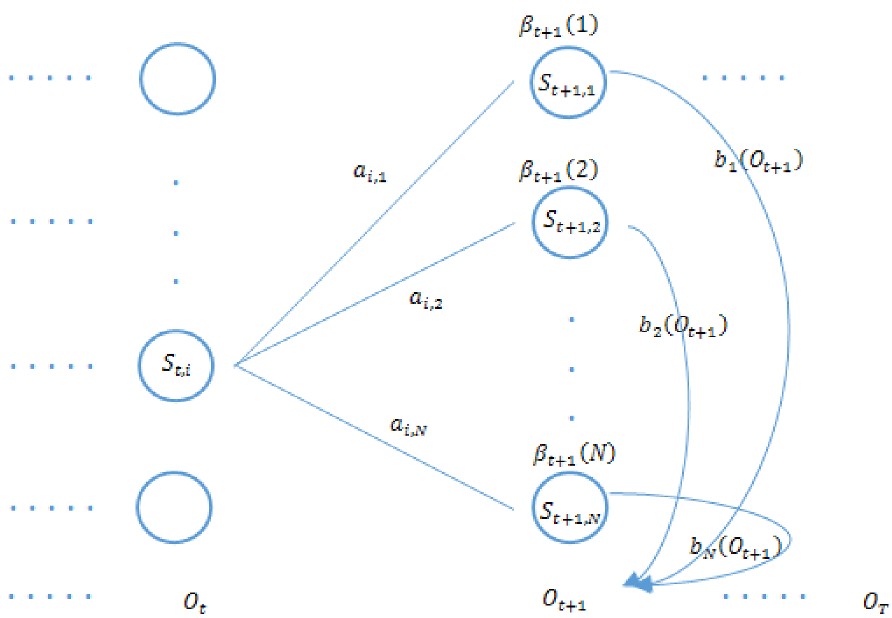
$$( ) = 1 , 1 \leq \quad \leq$$

در حالت کلی می توان ثابت کرد:

$$( ) \quad ( ) = \{ \; , \quad = \mid \} , 1 \leq \quad \leq \quad , 1 \leq \quad \leq$$

آنگاه می‌توان با کمک هر دو روش پیشرو و پسرو مقدار احتمال (|) را محاسبه نمود.

$$(\mid) = \qquad \{ \; , \quad = \mid \} = \qquad ( ) \quad ( )$$



شکل 7-4

7-5-2: مساله کد گشایی

در این حالت مـی‌خـواهیم بـا داشـتن دنبالـه مشـاهدات , … , , = و مـدل = ، (, ,) دنباله حالات بهینـه , … , , = بـرای تولیـد , … , , = را بدسـت آوریم.

یک راه حل این است که محتمل‌ترین حالت در لحظه t را بدست آوریم و تمام حالات را به این شکل برای دنباله ورودی تعیین کنیم. اما برخی مواقع این روش به ما یک دنباله معتبر و با معنا از حالات را نمی دهد به همین دلیل باید راهی پیدا نمود که یک چنین مشکلی نداشته باشد.

در یکی از این روش‌ها که با نام الگوریتم Viterbi شناخته می‌شود، دنباله حالات کامل با بیشترین مقدار نسبت شباهت پیدا می‌شود. در این روش برای ساده‌کردن محاسبات متغیر کمکی زیر را تعریف می‌نماییم.

$$(\ ) = \quad {}_{,\ ,\ldots,} \quad \{\ ,\ \ ,\ldots,\ \ ,\ \ =\ ,\ \ ,\ \ ,\ldots,\ \ |\ \}$$

که در شرایطی که حالت فعلی برابر با i باشد، بیشترین مقدار احتمال برای دنباله حالات و دنباله مشاهدات در زمان t را می‌دهد. به همین ترتیب می‌توان روابط بازگشتی زیر را نیز بدست آورد.

$$(\ ) = \ (\quad) \qquad (\ ) \quad ,1 \le \ \le \ ,1 \le \ \le \ -1$$

که در آن

$$(\ ) = \quad (\ ),1 \le \ \le$$

به همین دلیل روال پیداکردن دنباله حالات با بیشترین احتمال از محاسبه مقدار $\le\ ,(\ )$ $\le$ و با کمک رابطه فوق شروع می‌شود. در این روش در هر زمان یک اشاره‌گر به حالت برنده قبلی خواهیم داشت. در نهایت حالت *j را با داشتن شرط زیر بدست می آوریم.

$$* = \arg \qquad (\ )$$
$$\le \ \le$$

و با شروع از حالت *j ، دنباله حالات به شکل بازگشت به عقب و با دنبال‌کردن اشاره‌گر به حالات قبلی بدست می‌آید. با استفاده از این روش می‌توان مجموعه حالات موردنظر را بدست آورد. این الگوریتم را می‌توان به صورت یک جستجو در گراف که نودهای آن برابر با حالت‌های مدل HMM در هر لحظه از زمان می‌باشند نیز تفسیر نمود.

3-5-7: مساله یادگیری

به طور کلی مساله یادگیری به این موضوع مـی پـردازد کـه چگونـه مـی تـوان پارامترهـای مدل HMM را تخمین زد تا مجموعه داده‌های آموزشی به بهترین نحو بـه کمـک مـدل HMM برای یک کاربرد مشخص بازنمایی شوند. به همین دلیل می توان نتیجه گرفت که میـزان بهینـه

بودن مدل HMM برای کاربردهای مختلف، متفاوت است. به بیـان دیگـر مـی تـوان از چنـدین معیار بهینه‌سازی متفاوت استفاده نمود که هر یک برای کاربرد خاصی مناسب‌تر است. دو معیـار بهینه‌سازی مختلف برای آموزش مدل HMM وجود دارد که معیار بیشترین شـباهت (ML)[5] و معیار ماکزیمم اطلاعات متقابل MMI[6] می‌باشند.

7-5-3-1: معیار بیشترین شباهت(ML)

در معیار ML ما سعی داریم که احتمال یک دنباله ورودی که به کلاس w تعلـق دارد را با داشتن مدل HMM همان کلاس بدست آوریم. این میزان احتمال برابر با نسبت شباهت کلـی دنباله مشاهدات است و به صورت زیر محاسبه می‌شود.

$$= \{ \quad | \quad \}$$

با توجه به رابطه فوق در حالت کلی معیار ML به صورت زیر تعریف می شود.

$$= \{ \ | \ \}$$

اگر چه هیچ راه حل تحلیلی مناسبی برای مدل (, ,) = وجود نـدارد کـه مقـدار را ماکزیمم نماید، لیکن می توانیم با استفاده از یک روال بازگشتی پارامترهای مـدل را بـه شـکلی انتخاب کنیم که مقدار ماکزیمم بدست آید.

7-5-3-2: الگوریتم حداکثر سازی امید ریاضی

الگوریتم حداکثر سازی امید ریاضی یا EM بـه عنـوان یـک نمونـه از الگـوریتم بـام – ولـش در آمـوزش مـدلهای HMM مـورد اسـتفاده قـرار مـی‌گیـرد. الگـوریتم EM دارای دو فـاز تحـت عنوان Expectation و Maximization است. مراحل آموزش مدل در الگوریتم EM به‌صورت زیـر است.

الف) مرحله مقداردهی اولیه: پارامترهای اولیه مدل را تعیین می نماییم.

---

[5] Maximum Likelihood
[6] Maximum Mutual Information

ب) مرحله امید ریاضی[7] : برای مدل موارد زیر را محاسبه می کنیم.

- مقادیر با استفاده از الگوریتم پیشرو
- مقادیر و با استفاده از الگوریتم پسرو

ج) مرحله ماکزیمم‌سازی: مدل را با استفاده از الگوریتم باز تخمین محاسبه می نماییم.

د) مرحله بروز رسانی

ه) بازگشت به مرحله امید ریاضی

روال فوق تا زمانی که میزان نسبت شباهت نسبت به مرحله قبل بهبود مناسبی داشته باشد ادامه می یابد.

7-5-4: معیار ماکزیمم اطلاعات متقابل

در معیار ML ما مدل HMM را تنها برای یک کلاس در هر لحظه بروز رسانی می‌نماییم و به مدل‌های دیگر توجه نمی کنیم. به همین دلیل این روش مفهوم تمایز را که در شناسایی الگو بسیار مورد توجه است در نظر نمی‌گیرد. به همین دلیل روش یادگیری ML، به خصوص هنگامی‌که نمونه‌های فاز آموزش با نمونه‌های ورودی در فاز شناسایی متفاوت هستند، تمایزات ضعیفی بین مدل‌ها ایجاد می‌کند. اینگونه ناهمخوانی‌ها به دو دلیل ایجاد می‌شوند. اول آنکه داده‌های آموزشی و آزمایشی ویژگی‌های آماری متفاوتی دارند و دوم این‌که در مرحله آموزش نمی‌توان پارامترهای مدل را به شکل قابل اطمینانی تخمین زد.

در مقابل، معیار یادگیری MMI در هر لحظه تمام مدل‌های مربوط به کلاس ها را مورد آموزش قرار می‌دهد. در این حالت پارامترهای مدل اصلی برای بازنمایی مناسب داده‌ها بروز رسانی می شوند، در حالی‌که پارامترهای سایر مدل‌ها به شکلی تغییر می کنند که به میزان کمتری نمونه های آموزشی را بازنمایی نمایند. این روال باعث می شود که تمایز بین مدل‌ها افزایش یابد و به

---

[7] Expectation

همین دلیل است که روش یادگیری MMI به گروه روش‌های یادگیری تمایزی تعلق دارد. حـال فرض کنید که یک مجموعه از مدل‌های HMM را در اختیار داریم:

$$= \{ \quad ,1 \leq \quad \leq \quad \}$$

مساله این است که عدم قطعیت شرطی کلاس را با داشتن دنباله مشاهدات O حداقل نمـاییم. این مساله معادل است با کم کردن اطلاعات شرطی کلاس بر حسب A و به صورت زیـر بیـان می‌شود:

$$(\mid \ , \ ) = -\log p\{ \mid \ , \ \}$$

در چهارچوب تئوری اطلاعات، مساله فوق مانند حداقل کردن آنتروپی شرطی است.

$$(\mid \ ) = \quad [\ (\mid \ )]$$

که در آن بیانگر مجموعه تمام کلاس‌ها است وO دنباله مشـاهدات را نشـان مـی دهـد. آنگـاه اطلاعات متقابل بین کلاس‌ها و دنباله مشاهدات

$$(\ , \ ) = \quad (\ ) - \quad (\mid \ )$$

باید حداکثر شود، ( ) ثابت است. به همین دلیل اسـت کـه ایـن روش را روش حداکثرسـازی اطلاعات متقابل گویند. این روش با نام حداکثر سازی احتمال پسین یا (MAP) نیز شناخته می شود زیرا در رابطه (47) مقدار احتمال $\{ \ , \ \mid \ \}$ باید حداکثر شود. رابطه(47) را می تـوان بـه کمک تئوری بیز به شکل زیر نیز بیان نمود.

$$= - \quad p\{ \mid \ , \ \} - \quad \frac{p\{ \mid \ , \ \}}{p\{ \ \mid \ \}} - \quad \frac{p\{ \mid \ , \ \}}{\sum \ p\{ \mid \ , \ \}}$$

که در آن یکی از کلاس‌های آموزشی را نشان می دهد.

$$= \ \{ \ , \ \mid \ \}$$

$$= \ \{ \ , \ \mid \ \}$$

$$= -$$

در ایـن حالـت مـی تـوان مقـدار را بـا اسـتفاده از روش مبتنـی بـر گرادیـان یـا روش بازتخمین ML حداقل نمود. برای مثال روش مبتنی بر گرادیان را می توان به صورت زیر تعریـف نمود. این روش با تلاش برای مینیمم کردن رابطه زیر شروع می شود.

$$J=$$

آنگاه می توان مساله فوق را به صـورت مسـاله محاسـبه — کـه در آن یکـی از پارامترهـای مدل‌های A است، ساده نمود.

$$- = - -$$

تکنیک مشابهی مانند آنچه در یادگیری ML استفاده شد، می‌تواند در اینجا نیـز مـورد اسـتفاده قرار گیرد. به کمک متغیرهای پیشرو و پسرو به صورت زیر داریم:

$$= \quad (\ ) \quad (\ )$$

$$= \quad (\ ) \quad (\ )$$

6-7: استفاده از مدل HMM در شناسایی گفتار

بحث شناسایی اتوماتیک گفتار را می توان از دو جنبه مورد بررسی قرار داد.

- از جنبه تولید گفتار
- از جنبه فهم و دریافت گفتار

مدل مخفی مارکوف (HMM) تلاشی برای مدل‌سازی آماری دستگاه تولید گفتار است و به همین دلیل به اولین دسته از روش‌های شناسایی گفتار تعلق دارد. در طول چندین سال گذشته این روش به عنوان موفق‌ترین روش در شناسایی گفتار مورد استفاده قرار گرفته است. دلیل اصلی این مساله این است که مدل HMM قادر است به شکل بسیار خوبی خصوصیات سیگنال گفتار را در یک قالب ریاضی قابل فهم تعریف نماید.

در یک سیستم ASR مبتنی بر HMM قبل از آموزش HMM یک مرحله استخراج ویژگی‌ها انجام می‌گیرد. به این ترتیب ورودی HMM یک دنباله گسسته از پارامترهای برداری است. بردارهای ویژگی می تواند به یکی از دو طریق بردارهای چندی‌سازی شده یا مقادیر پیوسته به مدل HMM آموزش داده شوند. می توان مدل HMM را به گونه ای طراحی نمود که هر یک از این انواع ورودیها را دریافت نماید. مساله مهم این است که مدل HMM چگونه با طبیعت تصادفی مقادیر بردار ویژگی سازگاری پیدا خواهد کرد.

7-7: استفاده از HMM در شناسایی کلمات جداگانه

در حالت کلی شناسایی واحدهای گفتاری جدا از هم به کاربردی اطلاق می شود که در آن یک کلمه، یک زیر کلمه یا دنباله‌ای از کلمات به صورت جداگانه و به تنهایی شناسایی شود. باید توجه داشت که این تعریف با مساله شناسایی گفتار گسسته که در آن گفتار به صورت گسسته بیان می شود متفاوت است. در این بین شناسایی کلمات جداگانه کاربرد بیشتری به نسبت دو مورد دیگر دارد و دو مورد دیگر بیشتر در عرصه مطالعات تئوری مورد بررسی قرار می گیرند.

برای این کاربرد راه حل‌های مختلفی وجود دارد زیرا معیارهای بهینه‌سازی متفاوتی برای این منظور معرفی شده اند و الگوریتم‌های پیاده‌سازی شده مختلفی نیز برای هر معیار موجود است. این مساله را از دو جنبه آموزش و شناسایی مورد بررسی قرار می‌دهیم.

1-7-7: آموزش

فرض می کنیم که فاز پیش پردازش سیستم دنباله مشاهدات زیر را تولید نماید:

$$O = \{ \quad , \quad , \ldots , \quad \}$$

پارامترهای اولیه تمام مدل‌های HMM را با یـک مجموعـه از مقـادیر مشـخص مقـداردهی مـی نماییم.

$$, \; 1 \leq \quad \leq$$

در آغاز این مساله را برای حالت clamped در نظر بگیرید. از آنجایی که مـا بـرای هـر کـلاس از واحدها یک HMM داریم، می توانیم مـدل از کـلاس l را کـه دنبالـه مشـاهدات فعلـی بـه آن مربوط می شود، انتخاب نماییم.

$$= \quad = \sum \quad (\,) \quad (\,) = \sum \quad (\,)$$

$$= \quad = \quad [\quad (\,) \quad (\,)]$$

$$= \quad (\,)$$

که در آن بیانگر میزان شباهت دنباله مشاهدات فعلی به کلاس l در مدل است.

$$— = [\frac{}{\sum} \; -—] \sum \quad (\,) \quad (\quad) \quad (\,)$$

$$\frac{}{(\quad)} = [\frac{1}{\sum} - —] \frac{(\,) \quad (\,)}{(\quad)}$$

با توجه به موارد فوق، روال آمـوزش بـا اسـتفاده از روش مبتنـی بـر گرادیـان و معیـار MMI را می‌توان به شکل زیر خلاصه نمود.

1) هــر یــک از مــدل‌هــای HMM، $(\ , \ , \ ) =$ ، را بــا یکــی از دو روش تصــادفی یــا خوشه‌بندی K-means مقداردهی اولیه می کنیم.

2) با داشتن دنباله مشاهدات

I. مقادیر احتمال پیشرو و پسرو هر HMM را محاسبه می کنیم.
II. مقدار نسبت شباهت را محاسبه می کنیم.
III. مقدار گرادیان را بر حسب پارامترهای هر مدل محاسبه می نماییم.
IV. پارامترهای مدل را بروز رسانی می کنیم.

3) اگر همه نمونه‌های آموزشی استفاده نشده‌اند به گام 2 بر می گردیم.

4) مراحل 2 و 3 را تا رسیدن به شرط همگرایی ادامه می دهیم.

این مراحل را می توان به سادگی برای مدل‌های HMM دارای چگالی پیوسته نیز تغییر داد. این کار با انتشار گرادیان با استفاده از قانون زنجیره‌ای انجام می‌شود.

2-7-7: شناسایی

در مقایسه با آموزش، روال شناسایی بسیار ساده‌تر است.

1) الگوریتم دنباله مشاهدات مورد نظر را دریافت می کند.

I. مقادیر احتمالات پیشرو و پسرو را برای هر یک از مدل‌ها محاسبه می‌کنیم.
II. مقدار احتمال $1 \leq \quad \leq$ را برای تمام مدل‌ها محاسبه می‌کنیم.
III. سپس کلاس دنباله مشاهدات ورودی با استفاده از رابطه زیر تعیین می‌شود.

در این حالت نرخ شناسایی به‌صورت نسبت بین واحدهای شناسایی صـحیح بـه کـل واحـدهای آموزشی حساب می شود.

مراجع:

- A step-by-step tutorial on HMMs (University of Leeds)
- K. F. Lee And H. W. Hon, “Speaker-Independent Phone Recognition Using Hidden Markov Models,” IEEE Transactions On Acoustics, Speech, And Signal Processing, Vol. 31, No. 11, 1989.
- Seymore, Andrew McCallum, and Roni Rosenfeld. Learning Hidden Markov Model Structure for Information Extraction. AAAI 99 Workshop on Machine Learning for Information Extraction, 1999.
- J. Li, A. Najmi, R. M. Gray, Image classification by a two dimensional hidden Markov model, IEEE Transactions on Signal Processing, 48(2):517-33, February 2000.