

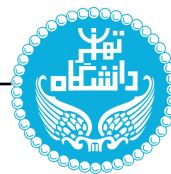
## تمرین شماره ۲

طراحان: علی پادیاو، مهدی جمال‌خواه، حسام  
رمضانیان

مهلت تحویل: شنبه ۱۸ فروردین ۱۴۰۳، ساعت ۲۳:۵۹

## هوش مصنوعی

مدرسین: دکتر فدایی و  
دکتر یعقوب‌زاده

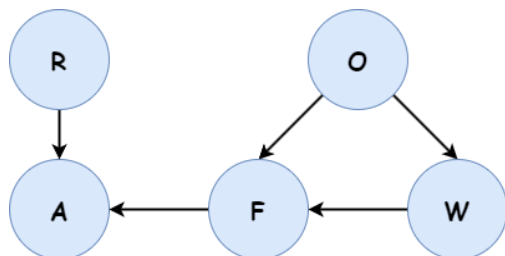


## بخش کتبی

### Bayes Nets

#### سوال اول

برای بررسی احتمالات مربوط به تصادفات جاده‌ای یک شبکه بیزی طراحی کرده ایم که در آن قدیمی بودن خودرو با  $O$ ، کارکرد بالای خودرو با  $W$ ، فرسودگی خودرو با  $F$ ، بارش باران با  $R$ ، تصادف کردن خودرو با  $A$  نمایش داده شده است. با توجه به این شبکه بیزی و جداول احتمالات داده شده به سوالات زیر پاسخ دهید:



O	P(O)
+	0.5
-	0.5

O	W	P(W O)
+	+	0.9
+	-	0.1
-	+	0.2
-	-	0.8

R	P(R)
+	0.2
-	0.8

F	R	A	P(A F,R)
+	+	+	0.9
+	+	-	0.1
+	-	+	0.7
+	-	-	0.3
-	+	+	0.4
-	+	-	0.6
-	-	+	0.3
-	-	-	0.7

O	W	F	P(F O,W)
+	+	+	0.9
+	+	-	0.1
+	-	+	0.6
+	-	-	0.4
-	+	+	0.7
-	+	-	0.3
-	-	+	0.2
-	-	-	0.8

الف) جدول احتمال مربوط به کارکرد بالای خودرو (W) را بنویسید.

ب) احتمال خواسته شده را محاسبه کنید :

$$P(+o,-w,+f,-r,+a)$$

ج) درست یا نادرست بودن عبارت های زیر را با ذکر دلیل مشخص کنید:

۱- در صورتی که از فرسودگی خودرو اطلاع داشته باشیم ، احتمال تصادف در جاده از احتمال قدیمی بودن خودرو مستقل است.

۲- احتمال فرسودگی خودرو و بارش باران به شرط اطلاع از تصادف خودرو از یکدیگر مستقل هستند.

۳- احتمال فرسودگی خودرو به شرط دانستن کارکرد و قدیمی بودن خودرو از احتمال تصادف کردن خودرو مستقل است.

۴- احتمال بارش باران از فرسودگی خودرو مستقل است.

د) با استفاده از Variable Elimination احتمال قدیمی بودن خودرو به شرط تصادف نکردن آن را به دست آورید. (در این قسمت نیازی به محاسبه مقدار عددی این احتمال نیست و تنها نوشتن مراحل Variable Elimination کافیت)

## HMM

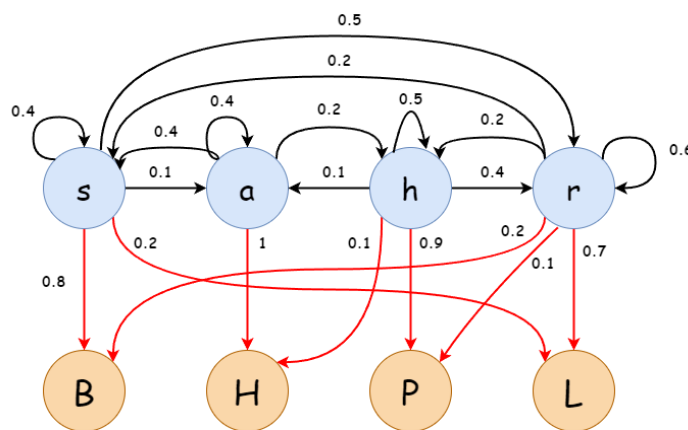
### سوال اول

فرض کنید یک HMM برای پلتفرم پخش آنلاین موسیقی طراحی کرده‌ایم که در آن مجموعه Hidden State ها را حالات روحی کاربر در نظر گرفته ایم ، حالات روحی افراد می‌تواند غمگین (s) ، عصبانی (a) ، شاد (h) ، آرام (r) باشد و مجموعه Observation ها را ژانر موسیقی که فرد در یک مرتبه مراجعه گوش می‌دهد در نظر گرفته‌ایم که میتواند Blues، Heavy Metal، Pop، Lo-Fi باشد. این دو مجموعه را به صورت زیر تعریف میکنیم:

Hidden States :  $X = \{s, a, h, r\}$

Observations :  $Y = \{B, H, P, L\}$

نمودار حالات این HMM به صورت زیر است و احتمال قرار داشتن فرد در هر کدام از حالات روحی در اولین مراجعه به پلتفرم ( $t=1$ ) برابر است. (اعداد نوشته شده بر روی یال ها بیانگر احتمالات است.)



با توجه به توضیحات داده شده به سوالات زیر پاسخ دهید :

الف) ماتریس Transition ، Emission، Initial Probability را به ترتیب با نام های  $A, B, \pi$  تعریف کرده و مقادیر آن ها را مشخص کنید.

ب) با استفاده از الگوریتم Forward احتمال اینکه فرد در مراجعات خود به ترتیب از چپ به راست این رشته از ژانر ها را گوش دهد محاسبه کنید.  $O = \{B, B, L, H\}$

ج) به کمک روابط مربوط به Smoothing احتمال اینکه در مراجعه دوم ( $t=2$ ) کاربر در حالت روحی غمگین (s) قرار داشته باشد به شرط اینکه رشته ژانر O را مشاهده کرده باشیم را محاسبه کنید.

د) با استفاده از الگوریتم Viterbi رشته حالت های روحی که بالاترین احتمال را دارد . به شرط اینکه رشته ژانر O را مشاهده کرده باشیم را محاسبه کنید.

## بخش عملی

### مقدمه

Hidden Markov Models (HMM) یکی از ابزارهای مهم در زمینه پردازش سیگنال‌های زمانی و تشخیص الگو هستند. در این مدل‌ها هر حالت به یک وضعیت مشخص مرتبط می‌شود و احتمال حرکت از یک وضعیت به وضعیت دیگر با توجه به وضعیت فعلی تعیین می‌شود. از این رو، HMM به عنوان یک ابزار قدرتمند برای مدل‌سازی سیستم‌های پویا و تصمیم‌گیری در شرایط عدم قطعیت به‌شمار می‌آید. در زمینه تشخیص گفتار، HMM معمولاً برای مدل‌سازی دقیق‌تر وضعیت‌های مختلف گفتار، مانند حروف، کلمات یا فریم‌های زمانی کوتاه از گفتار، استفاده می‌شود.

هدف اصلی نیز در این پروژه توسعه یک سیستم تشخیص گفتار اعداد بر اساس HMM می‌باشد. برای اینکار مجموعه داده صوتی‌ای در اختیارتان قرار گرفته و شما باید مراحل پیش‌پردازش، استخراج ویژگی، پیاده‌سازی مدل و تحلیل آن را در طی پروژه انجام دهید.

### آشنایی با مجموعه داده

مجموعه داده‌ای که در این پروژه در اختیارتان قرار گرفته، شامل ضبط‌های گفتاری از 6 گوینده مختلف است که هر کدام 50 بار ارقام 0 تا 9 را به زبان انگلیسی تلفظ می‌کنند. نام هر فایل صوتی به صورت `{digitLabel}_{speakerName}_{index}.wav` است که به ترتیب عدد گفته شده، نام گوینده و شماره نمونه را نشان می‌دهد.

### پیش‌پردازش و استخراج ویژگی

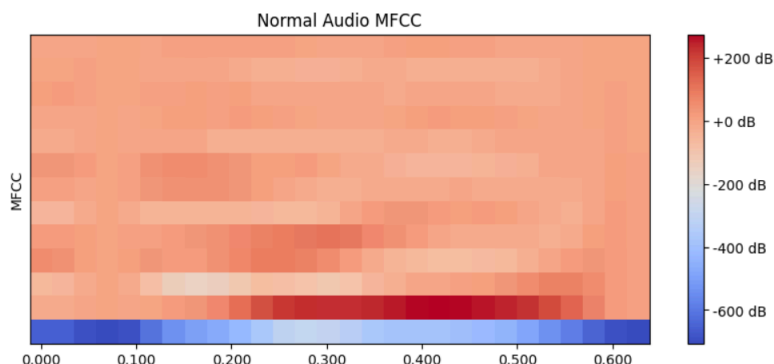
پیش‌پردازش داده‌ها اولین گام در هر پروژه داده محور است. هدف از پیش‌پردازش داده‌ها، بهبود کیفیت داده‌ها و آماده‌سازی آنها برای آموزش و آزمایش سیستم تشخیص گفتار است. مراحل معمول پیش‌پردازش داده‌ها می‌تواند شامل افزایش داده، حذف نویز، تقویت سیگنال، نرمال‌سازی و قطعه‌بندی باشد.

1. به نظر شما قطعه‌بندی (segmentation) داده‌ها برای این دیتاست مفید است؟ چرا؟

استخراج و انتخاب ویژگی از مهم‌ترین مراحل هر پروژه هوش مصنوعی و یادگیری ماشین می‌باشند. استفاده از ویژگی مناسب می‌تواند تاثیر بالایی در خروجی مدل نهایی داشته باشد. همچنین ویژگی‌های بسیار متفاوتی را می‌توان از یک محتوای صوتی استخراج کرد که هر کدام اطلاعات گوناگونی را به ما می‌دهند. از جمله این ویژگی‌ها می‌توان mfcc، Zero Crossing Rate، mel-spectrogram و chroma features را نام برد.

2. در مورد هر کدام از این ویژگی‌ها تحقیق کنید و روابط بین آنها را توضیح دهید.

در این پروژه از ضرایب Mel Frequency Cepstral Coefficient (MFCC) برای دسته‌بندی موسیقی‌ها استفاده می‌کنیم. ضرایب MFCC به طور گسترده در زمینه‌های مرتبط با دسته‌بندی موسیقی و تشخیص گفتار استفاده شده‌اند.



در این قسمت مجموعه ضرایب MFCC را برای هر کدام از نمونه‌های داده شده استخراج نماییم. برای این کار می‌توانید از یکی از کتابخانه‌های موجود برای کار با صوت استفاده نماییم. پس از استخراج ویژگی‌ها، نمودار Heat Map مربوط به ضرایب برای یک نمونه از هر کدام از دسته‌ها رسم نماییم.

3. robustness و حساسیت ویژگی‌های MFCCs را نسبت به تغییرات در سیگنال‌های صوتی بررسی کنید.

4. آیا موارد خاصی وجود دارند که ضرایب MFCC کارایی کمتری داشته باشند؟

5. چرا در محاسبه MFCC فریم‌های استفاده شده با یکدیگر هم‌پوشانی دارند؟

6. چرا در اکثر پروژه‌های مرتبط با صوت تنها از ۱۲ یا ۱۳ ضریب ابتدایی MFCC استفاده می‌شود؟

## آشنایی با HMM

همانطور که گفته شد، HMM یک مدل آماری است که به عنوان یک روش مناسب برای تحلیل پدیده‌های متغیر بر اساس زمان شناخته می‌شود. هر مدل HMM از بخش‌های زیر تشکیل شده است:

*States*

*Observations*

*Transition Probabilities*

*Emission Probabilities*

1. توضیح دهید منظور از State ها و Observation چیست؟ در این تمرین State ها کدامند و Observation چگونه بدست می‌آید؟

2. مدل‌های HMM را میتوان بر اساس میزان وابستگی میان State های پنهان دسته‌بندی کرد، مدلی که در این تمرین به پیاده‌سازی آن می‌پردازید یک مدل First-Order HMM است. دلیل نامگذاری آن و همچنین ویژگی‌های آن را بررسی کنید و تفاوت آن با مدل‌های دیگر در این دسته‌بندی را بیان کنید.
3. درباره HMM تحقیق کنید و توضیح دهید که این مدل برای بررسی و تحلیل چه پدیده‌هایی مناسب است؟ چرایی این موضوع را توضیح دهید.
4. مدل HMM نیز مانند هر مدل دیگری دارای مزایا و معایبی است که آن را ویژه می‌کند. مزایا و معایب این مدل را بررسی کرده و هر کدام را مختصراً توضیح دهید.
5. انواع مختلفی از مدل‌های HMM وجود دارد، درباره آنها تحقیق کنید و چند مورد را بطور مختصر بررسی کنید.

## پیاده‌سازی مسئله

### بخش اول: Implementing with Libraries

در بخش اول، شما باید با استفاده از کتابخانه‌های آماده (hmmlearn) به طراحی و پیاده‌سازی یک مدل HMM بپردازید. سپس مدل را بر روی داده‌هایی که در اختیار دارید آموزش داده و در نهایت نتایج را با معیارهایی که در بخش ارزیابی و تحلیل معرفی می‌شوند، بررسی کنید.

### بخش دوم: Implementing from Scratch

در این بخش از تمرین به پیاده‌سازی یک مدل HMM از پایه می‌پردازید. یک فایل Notebook شامل کدهای ناقص مورد نیاز برای پیاده‌سازی این مدل برای شما آپلود شده است. با تکمیل بخش‌های مختلف این فایل، در نهایت به یک مدل HMM دست خواهید یافت. مدل پایانی را بر روی داده‌ها train کرده و نتایج آن را با معیارهای متفاوتی که در ادامه معرفی می‌شوند، بررسی و تحلیل کنید.

نکته: ممکن است نتایج شما در بخش اول و دوم فرق کند و مدل آماده (که در بخش اول از آن استفاده کردید) نتایج متفاوت و دقت بالاتری نسبت به مدل طراحی شده توسط شما داشته باشد. این اختلاف ممکن است چه دلایلی داشته باشد؟ درباره عوامل تاثیرگذار بر روی این اختلاف دقت تحقیق کنید. راهنمایی: این تفاوت می‌تواند در ساختار مدل‌ها یا پیش‌پردازش داده‌ها باشد.

## راهنمایی پیاده‌سازی:

### \_sate\_likelihood:

این متد احتمال observation-ها را برای هر state محاسبه می‌کند. می‌توانید برای تکمیل این بخش از Multivariate normal distribution استفاده کنید.

### **\_em\_step:**

ابتدای این متد مقادیر جدید alpha و beta محاسبه می‌شوند تا در الگوریتم Expectation-Maximization به کار گرفته شوند.

Expectation-Maximization شامل آپدیت کردن پارامترهای مدل از جمله:

B (observation probability),

transition\_matrix (transition probability),

initial\_prob (states probability at the first time)

می‌باشد. همچنین میانگین و کوواریانس هم دوباره محاسبه می‌شوند که این دو مورد به طور کامل پیاده سازی شده‌اند و نیاز به تکمیل ندارند.

### **train:**

مدل ما در فرایند train در چرخه re-estimate و Expectation-Maximization قرار می‌گیرد تا likelihood دیدن observation-های آن مدل بیشتر شود.

### **score:**

با استفاده از این متد احتمال دیده شدن یک observation را در مدل به دست می‌آوریم.

## ارزیابی و تحلیل

معیارهای زیادی برای سنجش و ارزیابی عملکرد مدل‌ها وجود دارد؛ مثل: Accuracy، Precision، Recall، F1 Score.

1. درباره هر کدام از معیارهای بالا تحقیق کنید و نحوه محاسبه هر یک را توضیح دهید. و مطالعه این دو لینک (لینک ۱ و لینک ۲) برای درک این موضوع به شما کمک خواهد کرد.

2. آیا محاسبه معیارهای ذکر شده برای این پروژه که multi-class است، چالشی دارد؟ اگر بله چه راه حلی برایش دارید؟

3. توضیح دهید که هر کدام از معیارها چگونه مدل را ارزیابی میکنند.

4. تفاوت میان Precision و Recall را بیان کنید و توضیح دهید چرا هر کدام به تنهایی برای ارزیابی مدل کافی نیست؟ برای هر یک مثالی بیاورید که در آن، این معیار مقدار بالایی دارد اما مدل عملکرد خوبی ندارد.

5. معیار F1 از چه نوع میانگین‌گیری استفاده میکند؟ تفاوت این نوع میانگین‌گیری با میانگین‌گیری عادی چیست و در اینجا چرا اهمیت دارد؟

6. مدل خود را بر اساس رقم گفته شده در فایل صوتی (target var = digitLabel) آماده کنید. Confusion Matrix رسم کنید و دو معیار Accuracy و Precision را محاسبه کنید. در آخر مقادیر به دست آمده را تحلیل کنید.

7. مدل خود را بر اساس گوینده (target var = speakerName) آماده کنید. Confusion Matrix رسم کنید و دو معیار Accuracy و Precision را محاسبه کنید. در آخر مقادیر به دست آمده را تحلیل کنید.

8. تفاوت نتایج بخش‌های 6 و 7 را بررسی کنید و علل آن را مشخص کنید.

تذکر: برای نشان دادن Confusion Matrix و محاسبه معیارها از کتابخانه‌های آماده استفاده نکنید.

## نکات پایانی

- توضیحات مربوط به هر بخش از پروژه را بطور خلاصه و در عین حال مفید در گزارش خود ذکر کنید. از ابزارهای تحلیل داده مانند نمودارها استفاده کنید.
- پس از مطالعه کامل و دقیق صورت پروژه، در صورت وجود هرگونه ابهام یا سوال با طراحان پروژه در ارتباط باشید.
- نتایج، گزارش و کدهای خود را در قالب یک فایل فشرده با فرمت AI\_CA2\_[stdNumber].zip در سامانه ای لرن بارگذاری کنید.
- محتویات پوشه باید شامل فایل jupyter-notebook، خروجی html و فایل‌های مورد نیاز برای اجرای آن باشد. از نمایش درست خروجی‌های مورد نیاز در فایل html مطمئن شوید.
- دقت کنید که نیازی به آپلود مجموعه داده‌ها در سامانه ای لرن نیست.