



Pattern Recognition

HW4 : HMM and BN

Spring 99



۱. Markov Chain

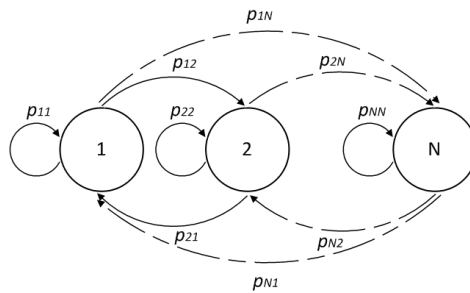
دنباله z_1, z_2, \dots, z_n را یک دنباله مارکوف محدود می‌نامیم اگر z_i ها متغیر تصادفی باشند به طوری که $z_i \in \{s_1, s_2, \dots, s_m\}$ و s_i یکی از حالت های زنجیره مارکوف است. با توجه به فرض مارکوف مرتبه اول می‌توان نوشت،

$$P\{z_i = s_k | z_{i-1} = s_{k-1}, z_{i-2} = s_{k-2}, \dots, z_1 = s_1\} = P\{z_i = s_k | z_{i-1} = s_{k-1}\}$$

ماتریس انتقال حالت یک زنجیر مارکوف را با $\Pi_{m \times m} = [p_{ij}]$ نشان می‌دهیم که در آن $p_{ij} = P\{z_n = s_j | z_{n-1} = s_i\}$ احتمال آن است که از s_i به s_j منتقل شویم. ماتریس انتقال باید خواص زیر را داشته باشد،

$$p_{ij} > 0, \quad \sum_{j=1}^m p_{ij} = 1$$

حالت یک زنجیر مارکوف را می‌توان با بردار حالت آن به صورت $\underline{W}^n = [w_1^n, w_2^n, \dots, w_m^n]$ توصیف کرد که در آن $w_k^n = P\{z_n = s_k\}$ می‌باشد. همچنین \underline{W}^1 حالت ابتدایی زنجیره است که بیان گر احتمال های ابتدایی حضور در هر استیت می‌باشد. مدل ساده‌ای از یک زنجیره مارکوف در شکل ۱ آمده است.



شکل ۱: مدل یک زنجیره مارکوف با N حالت

گزاره زیر را اثبات کنید،

$$\underline{W}^n = \underline{W}^1 \Pi^n$$

۲. برای محاسبه احتمال یک دنباله مشاهده شده از مدل HMM به صورت استاندارد از یکی از forward probability یا backward probability استفاده می‌شود.

$$\beta_j(t) = \begin{cases} 0 & t = T, j \neq \text{final status} \\ 1 & t = T, j = \text{final status} \\ \sum_{i=1}^c \beta_i(t+1) a_{ij} b_{jk} v(t+1) o.w & \end{cases} \quad \alpha_j(t) = \begin{cases} 0 & t = 1, j \neq \text{initial status} \\ 1 & t = 1, j = \text{initial status} \\ \sum_{i=1}^c \alpha_i(t-1) a_{ij} b_{jk} v(t) o.w & \end{cases}$$

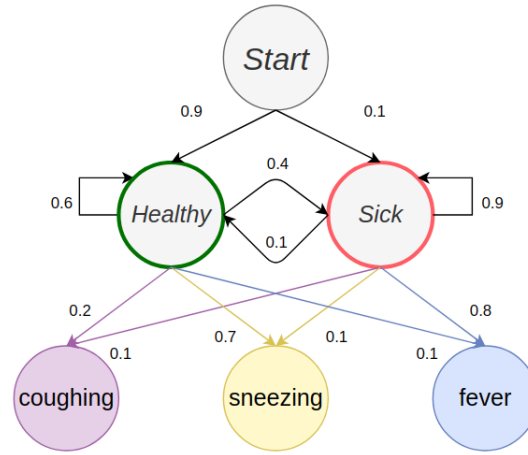
ثابت کنید که احتمال خواسته شده را که با $P\{V^T\}$ نشان می‌دهیم، می‌توان با ترکیب $\beta_i(T')$ و $\alpha_i(T')$ به صورت زیر بدست آورد.

$$P\{V^T\} = \sum_{i=1}^c \alpha_i(T') \beta_i(T') \quad \forall T' < T$$

برای مدل کردن علامت های یک بیماری، مدل شکل ۲ پیشنهاد شده است که در آن فرد به صورت احتمالاتی در یکی از شرایط سلامتی و یا بیماری قرار می گیرد و در هر حالت علامت هایی از خود نشان می دهد. احتمال دنباله FFSCFCSCSCFF را با استفاده از الگوریتم forward بدست آورید.

این بار احتمال های اولیه ورود به حالت بیماری و سلامتی را جابه جا کنید و احتمال مشاهده دنباله FFSCFCSCSCFF را مجددا محاسبه کنید. بنظر شما کدام یکی از مدل های HMM برای توصیف وضعیت فرد مناسب تر است؟

****در این سوال الگوریتم خواسته شده را باید خودتان پیاده سازی کنید و مجاز به استفاده از کتابخانه های آماده نیستید.**



شکل ۲: مدل پیشنهادی برای علامت های بیماری

۴. فرض کنید که تعداد زیادی دنباله از یک HMM استخراج شده است، در این مدل یکی از احتمال های انتقال در بین hidden states برابر * بوده است. این احتمال را به صورت $a_{i',j'}$ نشان می دهیم.

از روی این داده های استخراج شده، مدل HMM دیگری را با استفاده از الگوریتم forward-backward آموزش می دهیم. اثبات کنید که اگر در مدل دوم نیز مقدار اولیه $a_{i',j'}$ برابر * باشد، در طی آموزش این احتمال تغییری نخواهد کرد. به زبان دیگر شما نشان خواهید داد که اگر ساختار یک HMM معین باشد، و در مقدار دهی اولیه این احتمال برابر * انتخاب شود، در نهایت نیز * باقی خواهد ماند و آموزش HMM ثانویه حداقل برای احتمال های *، به درستی انجام شده است.

۵. Twitter Sentiment Analysis

مدل HMM بدلیل توانایی در مدل سازی داده های دنباله دار، در پردازش زبان های طبیعی مورد استفاده قرار می گیرد. در این تمرین می خواهیم با رویکرد بیزی، یک طبقه بند دوگانه طراحی کنیم تا جملات را به دو دسته مثبت و منفی تقسیم کند. در حقیقت در این تمرین از مدل HMM به عنوان یک مدل پارامتریک برای تخمین چگالی های شرطی در رابطه بیز استفاده می شود. مجموعه داده انتخابی برای این سوال twitter sentiment می باشد که نسخه ای از آن به همراه فایل تمرین در صفحه درس قرار داده شده است. پس از آموزش طبقه بند عملکرد آن را با معیار های متفاوتی همانند Precision، Recall و confusion matrix بسنجید و آن ها را گزارش کنید.

**** نکاتی در مورد پیاده سازی سوال:**

برای حل سوال می توانید از کتابخانه های آماده همانند `hmmlearn` و یا `sklearn` استفاده کنید. یکی از چالش های مهم در پردازش زبان های طبیعی کار با کلمات و جملات می باشد، چرا که معمولا مدل های قابل آموزش، مدل هایی عددی هستند. برای رفع این مشکل از ابزار های پردازش متن همانند `TextBlob`، `NLTK`، `spaCy` و جز آن استفاده می شود. همچنین از تکنیک `word embedding` برای عددی کردن کلمات استفاده می شود. در صورتی که با این مفهوم آشنایی ندارین پیشنهاد می شود اینجا را مطالعه کنید. در این مسئله نگاشت کردن کلمات به بردارهایی با بعد بالاتر از ۱ الزامی نمیباشد و بهتر است که مدل را ساده پیاده سازی کنید. برای کمک در پیاده سازی شما، نمونه ساده ای از `word embedding` همراه تمرین قرار داده شده است که می توانید با الگو گیری از آن ادامه پیاده سازی را انجام دهید. در صورتی که برای این قسمت پیاده سازی پیشرفته تری انجام دهید به شما نمره امتیازی تعلق می گیرد. ممکن است پردازش بر روی داده ها بسته به سخت افزار مورد استفاده بسیار طولانی شود در این صورت تنها از بخشی از مجموعه داده برای آموزش و آزمایش استفاده کنید.

توجه کنید که بدلیل وجود hyper parameter هایی همانند تعداد استیت های لایه مخفی، نحوه نگاشت کلمات به بردار های عددی و جز آن، به بازه متنوعی از دقت عملکرد مدل می رسین. در این تمرین گزارش صحیح از عملکرد مدل معیار اصلی است و دقت های بالا اهمیت چندانی ندارد. البته در صورتی که به دقت بالایی رسیدین می توانید عملکرد دقیق مدل خود را با شرکت در این مسابقه بسنجید.

هدف بدست آوردن چگالی مشترک ۶ متغیر تصادفی X_1, X_2, \dots, X_6 می‌باشد. چگالی‌های شرطی زیر به عنوان راهنمایی به ما داده شده است. با توجه به روابط زیر شبکه بیزی متناظر را رسم کنید و چگالی احتمال $f_{X_1, X_2, \dots, X_6}(x_1, x_2, \dots, x_6)$ به صورت ساده شده بازنویسی کنید.

$$f(x_2|x_1) = f(x_2)$$

$$f(x_3|x_2, x_1) = f(x_3)$$

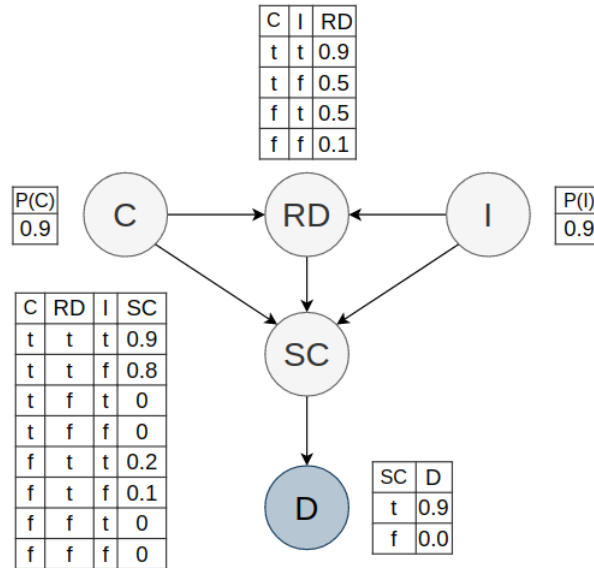
$$f(x_4|x_3, x_2, x_1) = f(x_4|x_3, x_2, x_1)$$

$$f(x_5|x_4, x_3, x_2, x_1) = f(x_5|x_4, x_3)$$

$$f(x_6|X_5, x_4, x_3, x_2, x_1) = f(x_6|x_5, x_1)$$

Bayesian Network Inference ۷.

با توجه به مطلبی که WHO در اینجا منتشر کرده است، ویروس COVID19(C) و influenza(I) در ایجاد بیماری‌های تنفسی respiratory disease(RD) به یک دیگر شباهت دارند. این نوع بیماری‌ها رنج وسیعی از بدون علامت asymptomatic تا شدید severe condition(SC) را شامل می‌شوند که در حالت شدید گاهی به مرگ می‌انجامد. از آنجایی که مدل‌های بیزی برای نمایش دانش اولیه در یک سیستم تصادفی مناسب هستند، شبکه بیزی شکل ۳ برای این مسئله طراحی شده است و جداول احتمال شرطی CPTs متغیرهای آن نیز داده شده اند. با توجه به این شبکه به قسمت‌های زیر پاسخ دهید.



شکل ۳: شبکه بیزی برای نمایش دانش اولیه در مورد دو ویروس

(آ) درستی گزاره‌های زیر را با توجه به ساختار شبکه بیزی بررسی کنید.

$$f(C, RD, I) = f(C)f(RD)f(I) \text{ i.}$$

$$f(D|SC) = f(D|SC, RD) \text{ ii.}$$

$$f(I|SC, C, RD) = f(I|SC, C, RD, D) \text{ iii.}$$

(ب) احتمال $P\{C = t, RD = t, I = f, SC = t, D = t\}$ را محاسبه کنید.

(ج) در حالت کلی در شبکه بیزی، یک متغیر با مشروط شدن به تعدادی از والد هایش از دیگر والد هایش مستقل نخواهد شد. اما در بعضی از مسئله‌ها، متغیر با مشروط شدن به مقادیر خاصی از تعدادی والد، از دیگر والد‌های خود مستقل می‌شود. این حالت را context-specific independence می‌نامیم. این نوع استقلال را در شبکه بیزی شکل ۳ بیابید.

عملیات arc reversal در یک شبکه بیزی این امکان را ایجاد می‌کند که جهت یک یال که به صورت $X \rightarrow Y$ بوده است را به $Y \rightarrow X$ تغییر دهیم. برای این که چگالی مشترک شبکه بیزی تغییری نکند لازم است که تمامی والد های X ، والد های Y بشوند و تمامی والد های Y ، والد های X بشوند. در نتیجه این عملیات ممکن است تعدادی یال به شبکه بیزی افزوده شود.

فرض کنیم والد های X را به صورت $U \cup V$ و والد های Y را به صورت $V \cup W$ نشان دهیم به طوری که $U \cap W = \emptyset$. همچنین مجموعه نود هایی از شبکه که X یا Y را به عنوان اجداد خود دارند با D و آن هایی که X یا Y را به عنوان اجداد خود ندارند با \bar{D} نشان می‌دهیم.

نشان دهید که چگالی مشترک متغیر های شبکه بیزی با انجام عملیات arc reversal تغییری نمیکند.

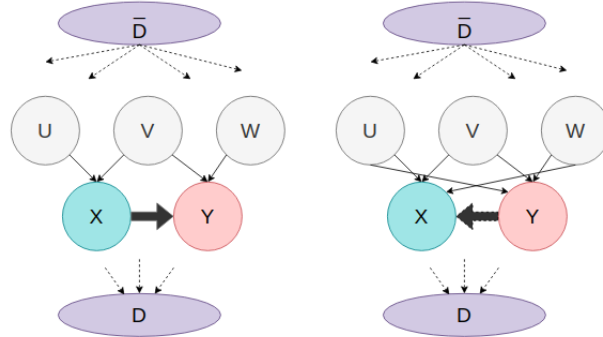
$$f'(D, \bar{D}, X, Y, W, U, V) = f(D, \bar{D}, X, Y, W, U, V)$$

**راهنمایی: ابتدا گزاره های زیر را اثبات کنید

$$f'(Y|U, V, W) = \sum_x f(Y|V, W, x)f(x|U, V)$$

$$f'(X|U, V, W, Y) = f(Y|X, V, W)f(X|U, V)/f(Y|U, V, W)$$

$$f'(X, Y|U, V, W) = f(X, Y|U, V, W)$$

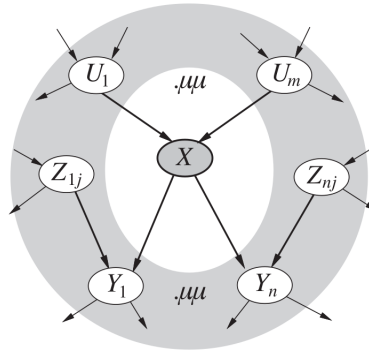


شکل ۴: نحوه تغییر شبکه بیزی توسط arc reversal

۹. Markov Blanket (امتیازی)

به مجموعه نود های والدان، بچه ها و والدان بچه ها برای یک نود در گراف شبکه بیزی پوشش مارکوف گفته می‌شود. اثبات کنید که هر نود مشروط بر پوشش مارکوف خود از تمامی نود های دیگر مستقل است.

$$f(X|MB(X), D) = f(X|MB(X)) \quad \forall D \rightarrow D \cap MarkovBlanket(X) = \emptyset$$



شکل ۵: مثالی از پوشش مارکوف برای یک شبکه بیزی

نکات پایانی:

۱. شما باید پاسخ های خود را با الگو PATREC-HW4-SID.zip در محل تعیین شده آپلود کنید
 ۲. گزارش شما معیار اصلی ارزیابی خواهد بود، در نتیجه دقت کنید کیفیت عکس ها مناسب باشند.
 ۳. کدهای خود را به تفکیک سوال ارسال کنید و استفاده از دیگر زبان های برنامه نویسی ممانعتی ندارد.
 ۴. هدف از انجام تمرین یادگیری مباحث درس می باشد، بنابر این تمرین را خودتان انجام دهید. در صورت کشف مشابهت بلا توجیه، با توجه به قوانین درس عمل خواهد شد.
 ۵. شما میتوانید سوالات خود را از طریق ایمیل sj.pakdaman@ut.ac.ir بپرسید
- موفق باشید