به نام خدا

تمرین دوم درس هوش مصنوعی

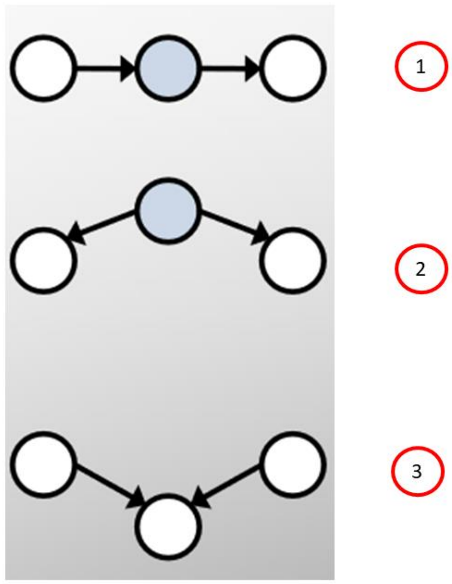
استاد: دکتر رهبان

نویسنده: سید علیرضا میررکنی

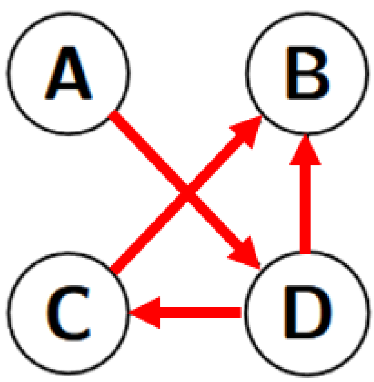
شماره دانشجویی: ۴۰۱۱۰۶۶۱۷

دانشکده مهندسی کامپیوتر دانشگاه صنعتی شریف – بهار ۱۴۰۳

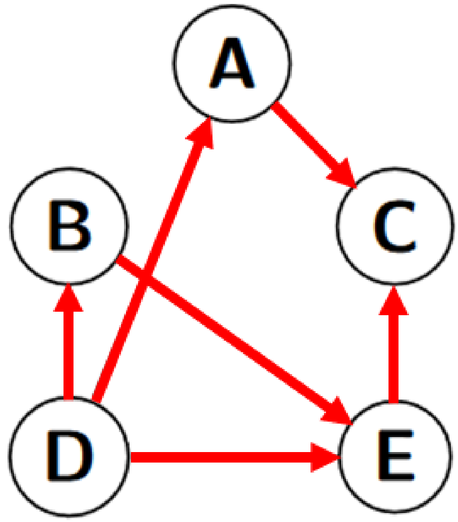
سوال ۱: در ابتدا دقت کنید که برای مستقل شدن متغیر های A و B به شرط متغیر C، باید هر مسیر بدون جهتی که در شبکه بیزی بین دو متغیر A و B وجود دارد، inactive باشد و برای این منظور باید حداقل یکی از سه الگوی زیر در آن مسیر مشاهده شود.



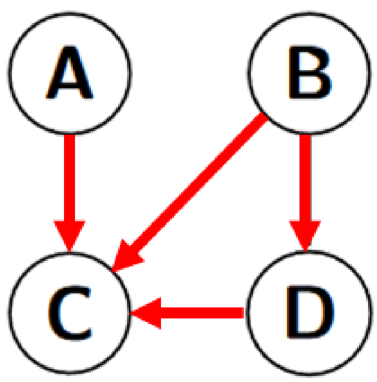
حال در هر مورد جهت دهی مطلوبی برای یال ها در شبکه بیزی معرفی می کنیم، به طوری که خواسته مسئله را ارضا نماید.



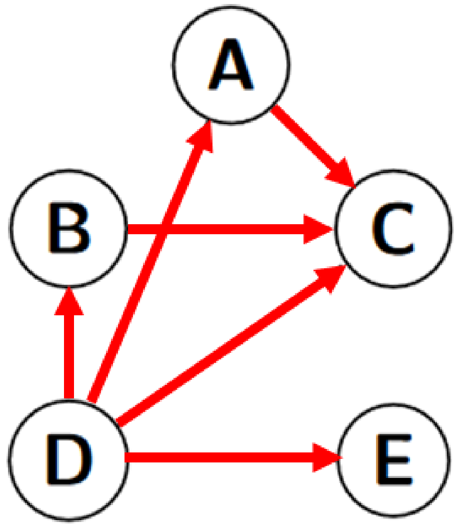
در شکل بالا، مسیر های بدون جهت از A به B مسیر های ADB و $ADCB$ هستند. در مسیر ADB، سه راس ADB الگوی شماره ۱ را تولید می کنند. در مسیر $ADCB$، سه راس ADC الگوی شماره ۱ را تولید می کنند. پس در این جهت دهی، همه مسیر های بدون جهت بین دو متغیر A و B به شرط داشتن متغیر D، inactive هستند و در نتیجه جهت دهی بالا مطلوب می باشد.



در شکل بالا، مسیر های بدون جهت از A به B مسیر های $ACEB$، $ACEDB$، ADB و $ADEB$ هستند. در مسیر $ACEB$، سه راس ACE الگوی شماره ۳ را تولید می کنند. در مسیر $ACEDB$، سه راس ACE الگوی شماره ۳ را تولید می کنند. در مسیر ADB، سه راس ADB الگوی شماره ۲ را تولید می کنند. در مسیر $ADEB$، سه راس ADE الگوی شماره ۲ را تولید می کنند. پس در این جهت دهی، همه مسیر های بدون جهت بین دو متغیر A و B به شرط داشتن متغیر D، inactive هستند و در نتیجه جهت دهی بالا مطلوب می باشد.



در شکل بالا، مسیر های بدون جهت از A به B مسیر های ACB و $ACDB$ هستند. در مسیر ACB، سه راس ACB الگوی شماره ۳ را تولید می کنند. در مسیر $ACDB$، سه راس BDC الگوی شماره ۱ را تولید می کنند. پس در این جهت دهی، همه مسیر های بدون جهت بین دو متغیر A و B به شرط داشتن متغیر D، inactive هستند و در نتیجه جهت دهی بالا مطلوب می باشد.



در شکل بالا، مسیر های بدون جهت از A به B مسیر های ADB، $ADCB$، ACB و $ACDB$ هستند. در مسیر ADB، سه راس ADB الگوی شماره ۲ را تولید می کنند. در مسیر $ADCB$، سه راس ADC الگوی شماره ۲ را تولید می کنند. در مسیر ACB، سه راس ACB الگوی شماره ۳ را تولید می کنند. در مسیر $ACDB$، سه راس CDB الگوی شماره ۲ را تولید می کنند. پس در این جهت دهی، همه مسیر های بدون جهت بین دو متغیر A و B به شرط داشتن متغیر D، inactive هستند و در نتیجه جهت دهی بالا مطلوب می باشد.

سوال ۲: در ابتدا دقت کنید که می توان نوشت: (عملیات marginalization)

$$
\begin{aligned}
P(A,C,D|+f) \propto P(A,C,D,+f) &= \sum_{B,E,G} P(A,C,D,+f,B,E,G) \\
&= \sum_{B,E,G} P(A)P(B|A)P(C|B)P(E|B)P(D|E)P(+f|E,D)P(G|+f) \\
&= \sum_{B,G} P(A)P(B|A)P(C|B)P(G|+f)\sum_{E} P(E|B)P(D|E)P(+f|E,D) \\
&= \sum_{B} P(A)P(B|A)P(C|B)\sum_{G} P(G|+f)\sum_{E} P(E|B)P(D|E)P(+f|E,D) \\
&= P(A)\sum_{B} P(B|A)P(C|B)\sum_{G} P(G|+f)\sum_{E} P(E|B)P(D|E)P(+f|E,D)
\end{aligned}
$$

حال دقت کنید که به طور معادل نوشت:

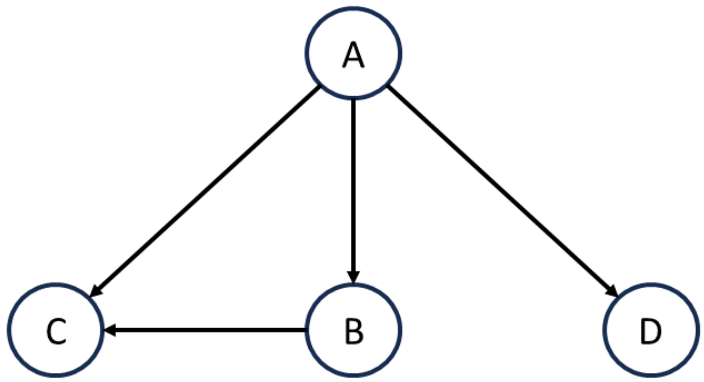
$$
\begin{aligned}
&P(A)\sum_{B} P(B|A)P(C|B)\sum_{G} P(G|+f)\sum_{E} P(E|B)P(D|E)P(+f|E,D) \\
&= P(A)\sum_{B} P(B|A)P(C|B)\sum_{G} P(G|+f)\sum_{E} P(E,D,+f|B) \\
&= P(A)\sum_{B} P(B|A)P(C|B)\sum_{G} P(G|+f)\, P(D,+f|B) \\
&= P(A)\sum_{B} P(B|A)P(C|B)\, 1 P(D,+f|B) \\
&= P(A)\sum_{B} P(B|A)P(C|B)\, P(D,+f|B) = P(A)\sum_{B} P(B,C,D,+f|A) \\
&= P(A)P(C,D,+f|A) = P(A,C,D,+f)
\end{aligned}
$$

با توجه به تساوی های بالا، در ابتدا جداول متغیر های E، D و F را ادغام می کنیم تا متغیر E، sum out شود. برای این منظور؛ $P(E|B)$، $P(D|E)$ و $P(+f|E,D)$ را در هم ضرب می کنیم تا $P(E,D,+f|B)$ حاصل شود. سپس با جمع زدن روی تمامی مقادیر متغیر E، $P(D,+f|B)$ را به دست می آوریم.

سپس از جدول متغیر G استفاده می کنیم تا خود این متغیر را sum out کنیم، به این شکل که $P(G|+f)$ را با استفاده از این جدول به دست می آوریم و با جمع زدن روی تمامی مقادیر متغیر G، به مقدار ۱ می رسیم (در حقیقت نیازی به استفاده از جدول نمی باشد، چرا که می دانیم حاصل نهایی برابر ۱ خواهد بود).

در ادامه با ادغام جدول به دست آمده برای $P(D,+f|B)$ در مرحله اول و همچنین جداول متغیر های B و C، متغیر B را sum out می کنیم. برای این منظور، $P(B|A)$، $P(C|B)$ و $P(D,+f|B)$ را در هم ضرب می کنیم تا $P(B,C,D,+f|A)$ حاصل شود. سپس با جمع زدن روی تمامی مقادیر متغیر B، $P(C,D,+f|A)$ را به دست می آوریم و در نهایت با ضرب کردن آن در $P(A)$ مقدار $P(A,C,D,+f)$ محاسبه می شود.

سوال ۳: در ابتدا با توجه به جداول داده شده، می توان دریافت که شبکه بیزی به شکل زیر می باشد:



الف) در ابتدا وزن هر کدام از نمونه ها را محاسبه می کنیم. نمونه ها را از بالا به پایین، به ترتیب با ۱ تا ۵ شماره گذاری می کنیم.

$$weight(sample_1) = P(+b|+a) \times P(-c|+a,+b) = 0.8 \times 0.6 = 0.48$$
$$weight(sample_2) = P(+b|-a) \times P(-c|-a,+b) = 0.3 \times 0.75 = 0.225$$
$$weight(sample_3) = P(+b|-a) \times P(-c|-a,+b) = 0.3 \times 0.75 = 0.225$$
$$weight(sample_4) = P(+b|+a) \times P(-c|+a,+b) = 0.8 \times 0.6 = 0.48$$
$$weight(sample_5) = P(+b|+a) \times P(-c|+a,+b) = 0.8 \times 0.6 = 0.48$$

حال دقت کنید که در نمونه های شماره ۱، ۴ و ۵ مقدار a برابر $True$ شده است (به عبارت دیگر در این نمونه ها حالت $+a$ اتفاق افتاده است). پس می توان نوشت:

$$P(+a|+b,-c) = \frac{weight(sample_1) + weight(sample_4) + weight(sample_5)}{\sum_{i=1}^{5} weight(sample_i)}$$
$$= \frac{3 \times 0.48}{3 \times 0.48 + 2 \times 0.225} = \frac{16}{21} \approx 0.7619$$

ب) ۵ نمونه داریم که در ۳ تا از آن ها (نمونه های شماره ۱، ۲ و ۵) مقدار d برابر $True$ شده است (به عبارت دیگر در این نمونه ها حالت $+d$ اتفاق افتاده است). پس می توان نوشت:

$$P(+d) = \frac{3}{5} = 0.6$$

پ) طبق روش نمونه گیری $Gibbs$ می توان نوشت:

$$P(+b\ in\ next\ sampling) = P(+b|+a,+c,+d) = P(+b|+a,+c) = \frac{P(+b,+a,+c)}{P(+a,+c)}$$

$$= \frac{P(+a)P(+b|+a)P(+c|+a,+b)}{\sum_B P(+a)P(B|+a)P(+c|B,+a)} = \frac{P(+a)P(+b|+a)P(+c|+a,+b)}{P(+a)\sum_B P(B|+a)P(+c|B,+a)}$$

$$= \frac{P(+b|+a)P(+c|+a,+b)}{\sum_B P(B|+a)P(+c|B,+a)} = \frac{0.8 \times 0.4}{0.8 \times 0.4 + 0.2 \times 0.8} = \frac{2}{3} \approx 0.67$$