



دانشگاه صنعتی امیرکبیر
دانشکده مهندسی کامپیوتر



«مبانی و کاربردهای هوش مصنوعی ترم پاییز ۱۴۰۴»

تمرین پنجم

در انجام تمرین‌ها به نکات زیر توجه فرمائید:

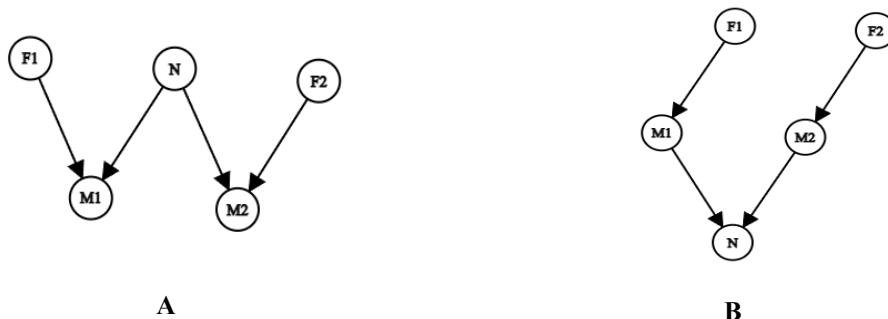
- ۱- مطابق قوانین دانشگاه هر نوع کپی‌برداری و اشتراک کار دانشجویان غیرمجاز است و پاسخ به تمرین‌ها باید به صورت انفرادی و بدون استفاده از ابزارهای هوش مصنوعی انجام شود، در صورت مشاهده چنین مواردی با طرفین شدیداً برخورد خواهد شد.
- ۲- پاسخ خود را در قالب یک فایل PDF به صورت تایپ‌شده و یا دستنویس (مرتب و خوانا) در سامانه کورسز آپلود نمائید.
- ۳- در صورت هر گونه سوال یا ابهام می‌توانید از طریق راه‌های ارتباطی گفته‌شده با تدریس‌یارهای طراح این تمرین در ارتباط باشید.
- ۴- توجه نمائید پاسخ تمرین‌ها تنها در صورت آپلود در سامانه کورسز پذیرفته خواهد شد و ارسال پاسخ از طریق ایمیل یا تلگرام بررسی نخواهد شد.
- ۵- فایل تمرین را با فرمت StudentID_AI_HW05.pdf تا ساعت ۲۳:۵۹ روز ۱۴۰۴/۱۰/۰۵ فقط در بخش مربوطه در سایت درس آپلود نمائید.

Bayes' Net

سوال اول

دو دوربین روی دو فانوس دریایی در نقاط مختلف یک بندر قرار دارند. این دو دوربین می‌خواهند تعداد کشتی‌های موجود در یک ناحیه از اقیانوس را اندازه‌گیری کنند. به احتمال e هر دوربین ممکن است تعداد کشتی‌ها را یک مورد بیشتر یا کمتر اندازه‌گیری کند. هر دوربین به احتمال f ممکن است خراب بشود که در این صورت تعداد کشتی‌ها توسط هر دوربین سه عدد یا بیشتر، کمتر شمرده می‌شود یا اگر تعداد کشتی‌ها کمتر سه باشند کلاً نمی‌توانند بشمارند. فرض کنید تعداد کشتی‌ها N است و دوربین اول را با M_1 و دوربین دوم با M_2 و خرابی دوربین اول و دوم را با متغیرهای F_1 و F_2 نشان می‌دهیم.

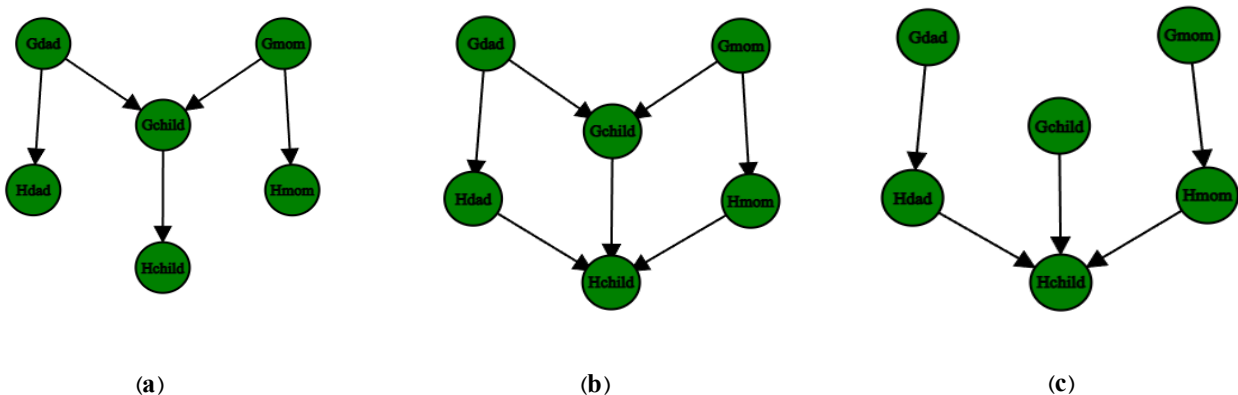
الف) کدام یک از شبکه‌های بیزین زیر برای مدلسازی مسئله مناسب است؟ دلیل بیاورید.



- ب) توزیع شرطی $P(M_1|N)$ بنویسید. فرض کنید N اعداد صحیح 1 تا 3 و M_1 اعداد صحیح 0 تا 4 هستند. هر قسمت از توزیع شرطی شما باید برحسب پارامترهای f و e نوشته شوند.
- ج) فرض کنید که $M_1 = 1$ و $M_2 = 3$ باشد. در این صورت مقادیر ممکن که N می‌تواند داشته باشد را بیابید.
- د) با توجه به فرض‌های قسمت قبل کدام مقدار N بیشترین احتمال را دارد؟ آیا می‌توان آن را حساب کرد؟ اگر جواب بله است چگونه حساب می‌کنید؟ اگر جواب خیر است چه اطلاعات بیشتری نیاز است تا آن را حساب کنیم؟

سوال دوم

فرض کنید H_x متغیری باشد که چپ‌دست یا راست‌دست بودن فرد x را با مقدار های $\{ \text{چپ‌دست} \}$ یا $\{ \text{راست‌دست} \}$ نشان می‌دهد. یک فرضیه عمومی این است که راست‌دست یا چپ‌دست بودن از طریق یک سازوکار وراثتی ساده به نسل بعد منتقل می‌شود. یعنی یک ژن به نام G_x وجود دارد که مقدار آن ۱ یا r است و وضعیت چپ‌دست یا راست‌دست بودن با احتمال s مشابه ژنی باشد که در بدن او قرار دارد. علاوه بر این ژن فرزند به احتمال مساوی از والدین به ارث می‌رسد، اما این احتمال هم وجود دارد که با یک جهش تصادفی به احتمال m ، وضعیت دست عوض شود.



الف) کدامیک از سه شبکه داده شده ادعا می‌کند که سه ژن پدر، مادر و فرزند از هم مستقل هستند؟

ب) کدامیک از سه شبکه داده شده روابط استقلال‌ای را نشان می‌دهد که با فرضیه ارائه شده در مورد نحوه به ارث رسیدن وضعیت دست، تطابق دارد؟

ج) از بین این سه نمودار کدامیک بهترین توصیف را از فرضیه مورد نظر دارد؟

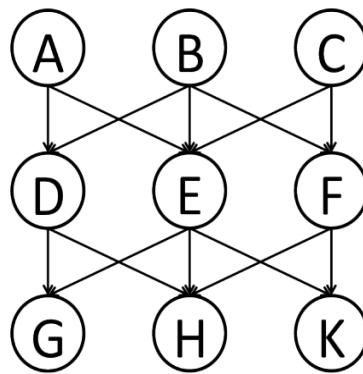
د) آیا می‌توان گره جدیدی با توجه به اطلاعات داده شده به شبکه a اضافه کرد؟

ه) CPT مربوط به گره G_{child} در شبکه a را بدست بیاورید.

و) فرض کنید احتمال اینکه پدر و مادر ژن چپ‌دست را داشته باشند، هر دو برابر q باشد. در این صورت در شبکه a با شرطی سازی بر روی گره های والد، احتمال اینکه فرزند ژن چپ‌دست داشته باشد را برحسب فقط m و q محاسبه نمایید.

سوال سوم

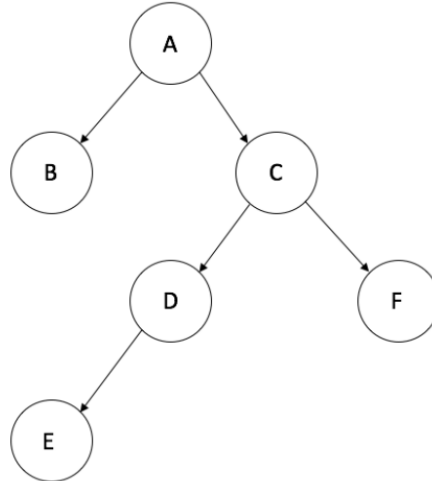
با توجه به شبکه بی‌زین زیر، در هر مورد مشخص کنید که آیا شبکه استقلال شرطی داده شده را تضمین می‌کند که درست است یا نادرست است یا کلاً نمی‌توان آن را مشخص کرد.

الف) $A \perp C$ ب) $A \perp C \mid E$ ج) $A \perp C \mid G$ د) $A \perp K$ ه) $A \perp G \mid D, E, F$ و) (امتیازی) $A \perp B \mid D, E, F$ ز) $A \perp C \mid D, F, K$ ح) $A \perp G \mid D$ ط) $A \perp C \mid H$ ی) $G \perp F \mid H, E$

سوال چهارم

شبکه بی‌زی زیر را در نظر بگیرید. همه متغیرها (دو‌حالتی) هستند:

$$A, B, C, D, E, F \in \{+, -\}$$



(i) [درست/نادرست] بررسی کنید آیا F و D مستقل‌اند یا نه:

$$F \perp D$$

(منظور از «درست/نادرست» این است که با توجه به ساختار BN (و با منطق d-separation/استقلال شرطی گرافی) تصمیم بگیرید این استقلال برقرار است یا خیر.)

(ii) [درست/نادرست] بررسی کنید آیا F و E با شرط D مستقل می‌شوند یا نه:

$$F \perp E \mid D$$

(یعنی آیا با دانستن مقدار D ، دانستن E هیچ اطلاعات اضافه‌ای درباره F نمی‌دهد؟)

(iii) می‌خواهیم کوئری زیر را حساب کنیم:

$$P(F \mid +e)$$

(یعنی evidence این است که $E = +e$). تصمیم می‌گیریم با elimination variable حل کنیم و ترتیب حذف را این‌طور انتخاب می‌کنیم: ابتدا D و سپس C . پس از حذف D و C ، یک factor به‌دست می‌آورید که آن را این‌گونه نام‌گذاری می‌کنیم:

$$f(+e, F \mid A, B)$$

عبارت $f(+e, F \mid A, B)$ را فقط بر حسب احتمال‌های شرطی گذشته در BN (یعنی CPT‌های مربوطه مثل $P(B \mid A)$ ، $P(C \mid A)$ ، $P(D \mid B, C)$ ، $P(E \mid D)$ ، $P(F \mid C)$ ، ...) بنویسید.

(iv) حالا $f(+e, F | A, B)$ را دارید و می‌خواهید ادامه variable elimination را انجام دهید: ابتدا B و سپس A را حذف می‌کنید تا factor نهایی زیر ساخته شود:

$$f(+e, F)$$

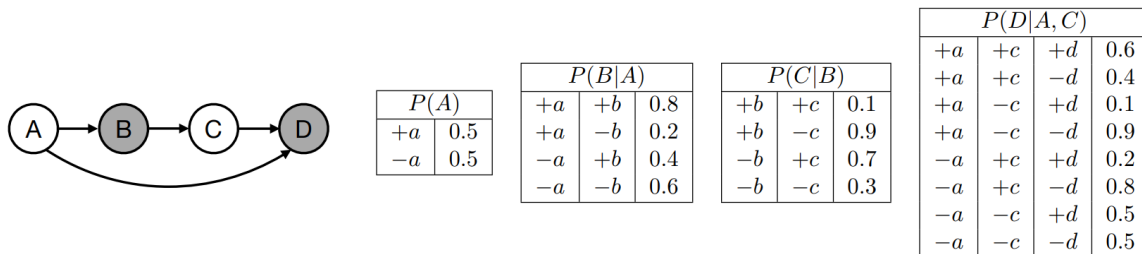
$f(+e, F)$ را بر حسب $f(+e, F | A, B)$ و CPT های مرتبطی که در این دو مرحله لازم می‌شوند بنویسید.

(v) گام آخر نرمال‌سازی است تا به $P(F | +e)$ برسیم. عبارت نهایی $P(F | +e)$ را با نرمال‌سازی $f(+e, F)$ بنویسید.

سوال پنجم

شبکه بی‌زی زیر را در نظر بگیرید که در آن شواهد (evidence) داریم:

$$B = +b, \quad D = +d$$



الف) Gibbs sampling

فرض کنید Gibbs sampling انجام می‌دهیم. همچنین فرض کنید مقداردهی اولیه همه متغیرها این است:

$$A = +a, B = +b, C = +c, D = +d$$

سپس متغیر C را un-assign می‌کنیم، یعنی مقدار C را موقتاً «نامعین» می‌گذاریم تا دوباره resample شود:

$$A = +a, B = +b, C = ?, D = +d$$

در این مرحله، احتمال‌های زیر را حساب کنید:

$$P(C = +c \text{ at next step}), \quad P(C = -c \text{ at next step})$$

(در Gibbs sampling، هنگام resample کردن یک متغیر (اینجا C) باید از توزیع شرطی آن با توجه به

بقیه مقدارهای فعلی نمونه (و evidence) نمونه‌گیری کنیم:

$$P(C | A = +a, B = +b, D = +d)$$

و معمولاً فقط CPT های مربوط به ** blanket Markov همان متغیر وارد محاسبه می شوند.

Hybrid: Rejection Sampling + Likelihood-Weighted Sampling (ب)

یک طرح نمونه گیری در نظر بگیرید که hybrid از rejection sampling و likelihood-weighted sampling است:

۱. ابتدا برای متغیرهای A و B rejection sampling انجام می دهیم (با در نظر گرفتن evidence مربوط، یعنی $B = +b$).

۲. سپس مقادیر نمونه برداری شده A و B را نگه می داریم و نمونه را برای متغیرهای C و D extend می کنیم و برای C, D از likelihood-weighted sampling استفاده می کنیم (یعنی evidence مربوط به $D = +d$ در وزن وارد می شود).

(i) چند نمونه کاندید برای بخش rejection (روی A, B) داده شده است. مشخص کنید کدام ها رد (reject) می شوند:

- $(-a, -b)$
- $(+a, +b)$
- $(+a, -b)$
- $(-a, +b)$

(ii) مستقل از قسمت (i)، یک مجموعه جدید از نمونه های کامل (برای A, B, C, D) داده شده است. وزن (weight) هر نمونه را تحت همین طرح hybrid محاسبه کنید:

- $-a, +b, -c, +d$
- $+a, +b, -c, +d$
- $+a, +b, -c, +d$
- $-a, +b, +c, +d$
- $+a, +b, +c, +d$

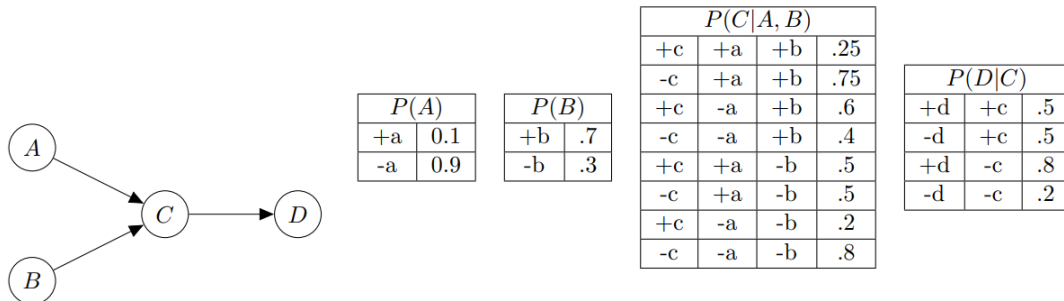
(iii) با استفاده از همین نمونه های وزندار، یک تخمین از عبارت زیر به دست آورید:

$$P(+a | +b, +d)$$

(در تخمین وزندار، سهم هر نمونه متناسب با weight آن است (پس شمردن ساده کافی نیست).)

سؤال ششم

فرض كنءء شبكه بىزى زىر و ءوزىع هائى مءناظر مءغىرهائى آن به شما ءاءه شده اسء:



الف) فرض كنءء شواءء (evidence) ءرىاءء كءءه اىم كه:

$$A = +a$$

اگر بءواءىم با Sampling Rejection نمونه بءءارى كنىم، به ءور امىءرىاضى (on expectation) چه ءرصدى از نمونه هاء رء (reject) ءواءنء شء؟

ب) اكنون فرض كنءء هر ءو مشاءءه زىر را ءارىم:

$$A = +a, \quad D = +d$$

ءر Likelihood Weighting، وزن (weight) نمونه هائى زىر را مءاسبه كنءء:

۱. $(+a, -b, +c, +d)$
۲. $(+a, -b, -c, +d)$
۳. $(+a, +b, -c, +d)$

ج) (امءىازى) با اسءءاءه از نمونه هائى قسءء قبل، مءءار زىر را ءءمىن بزنءء:

$$P(-b \mid +a, +d)$$

ء) فرض كنءء باىء به ءورء ءقرىبى (approximately) به ءو كوئرى اسءءءاجى زىر براءى همىن گراف باسء ءهىء:

$$P(C \mid +a) \quad \text{و} \quad P(C \mid +d)$$

شما مؤظف هستىء:

- يكى از اىن ءو كوئرى را با Likelihood Weighting
- و ءىگرى را با Gibbs Sampling باسء ءهىء.

ءر هر ءالء فقط مى ءوانىء ءءءاء كمى نمونه ءمع آورى كنءء؛ بءابراىن براءى بىشءرىن ءقء باىء الگورىءم را هوشمءءانه با ءوءه به مءاسب بوءءنش براءى كوئرى انءءاب كنءء.

- ۱) مشءء كنءء هر الگورىءم را براءى كءام كوئرى اسءءءاءه مى كنءء.
- ۲) انءءاب ءوء را ءوءىه (justify) كنءء.