

به نام خدا



دانشگاه تهران  
پردیس دانشکده‌های فنی  
دانشکده برق و کامپیوتر



مدل‌های مولد عمیق

مدرس: دکتر مصطفی توسلی‌پور

تمرین شماره ۱

مهر ماه ۱۴۰۴

۴	..... سوال اول : PGM
۴	..... بخش اول
۴	..... زیربخش اول (۳ نمره)
۴	..... زیربخش دوم (۳ نمره)
۵	..... زیربخش سوم (۱۰ نمره)
۵	..... بخش دوم
۵	..... زیربخش اول (۳ نمره)
۵	..... زیربخش دوم (۱ نمره)
۶	..... زیربخش سوم (۲ نمره)
۶	..... زیربخش چهارم (۱ نمره)
۶	..... زیربخش پنجم (۳ نمره)
۶	..... زیربخش ششم (۶ نمره)
۷	..... بخش سوم
۷	..... زیربخش اول (۳ نمره)
۷	..... زیربخش دوم (۸ نمره)
۷	..... زیربخش سوم (۲ نمره)
۸	..... بخش چهارم (۱۰ نمره)
۹	..... سوال دوم :
۹	..... بخش اول
۹	..... زیر بخش اول (۲ نمره)
۹	..... زیر بخش دوم (۱ نمره)
۹	..... زیر بخش سوم (۲ نمره)
۹	..... زیر بخش چهارم (۱۰ نمره)
۱۰	..... زیر بخش پنجم (۲ نمره)
۱۰	..... زیر بخش ششم (۱۲ نمره)
۱۱	..... زیر بخش هفتم (۸ نمره)
۱۱	..... زیر بخش هشتم (۳ نمره)
۱۲	..... بخش دوم
۱۲	..... زیر بخش اول (۵ نمره)
۱۲	..... زیر بخش دوم (۵ نمره)
۱۲	..... زیر بخش سوم (۵ نمره)
۱۳	..... مراجع
۱۴	..... نکات تحویل



### بخش اول

با توجه به توضیحات زیر به سوالات در ادامه پاسخ دهید.

ابتلا و شدت نوعی بیماری در هر فرد متاثر از سیستم ایمنی فرد است، بطوری که اگر فرد سیستم ایمنی ضعیفی داشته باشد، شدت بیماری در فرد بیشتر می شود. همچنین این بیماری در فصل های سرد نیز شدت بیشتری در بیمار پیدا می کند. در صورتی که شدت بیماری در فرد زیاد باشد، پیشنهاد می شود از نوعی دارو گران استفاده شود و در غیر این صورت، دارویی ارزان قیمت پیشنهاد می شود. با این حال اگر فرد توانایی مالی خرید دارو گران قیمت را نداشته باشد، از دارو ارزان استفاده می کند. اگر شدت بیماری فرد زیاد باشد، احتمال مرگ او افزایش می یابد. همچنین دارو گران قیمت، نرخ بهبودی بیشتری دارد و احتمال مرگ را کاهش می دهد.

با توجه به توضیحات گفته شده در بالا آن را با یک شبکه بیزی مدل می کنیم. متغیرها در شبکه بیزی به صورت زیر است:

- متغیر M: قدرت سیستم ایمنی
- متغیر I: شدت بیماری
- متغیر T: دارو مورد استفاده برای درمان
- متغیر S: فصل
- متغیر F: استطاعت مالی فرد
- متغیر D: احتمال مرگ فرد

---

### زیربخش اول (۳ نمره)

با توجه به توضیحات گفته شده، شبکه بیزی آن را رسم کنید (می توانید گراف را روی کاغذ رسم کرده و عکس آن را در گزارش قرار دهید).

---

### زیربخش دوم (۳ نمره)

توزیع احتمال توام را با توجه به متغیرهای تعریف شده و شبکه بیزی بدست آورید.

زیربخش سوم (۱۰ نمره)

با توجه به شبکه بیزی که رسم کردید، درستی یا نادرستی عبارت‌های زیر را با ذکر دلیل بیان کنید.

a.  $F \perp D$

b.  $S \perp D \mid I$

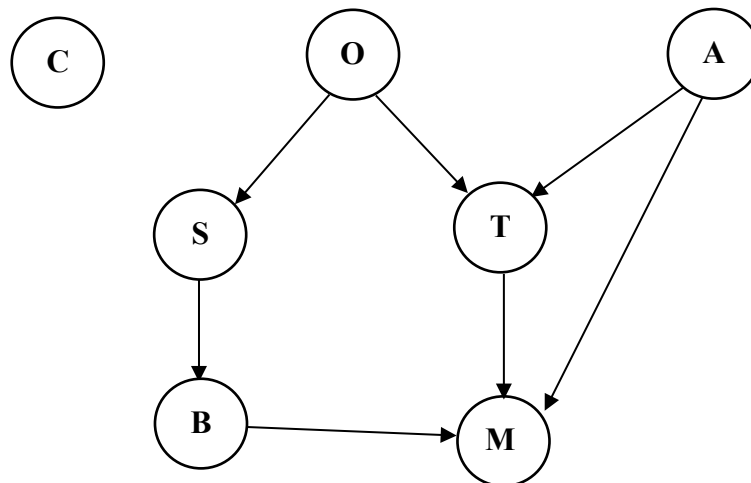
c.  $M \perp F$

d.  $M \perp F \mid T$

e.  $M \perp T \mid \{D, I\}$

بخش دوم

با توجه به گراف بیزی زیر به سوالات پاسخ دهید.



زیربخش اول (۳ نمره)

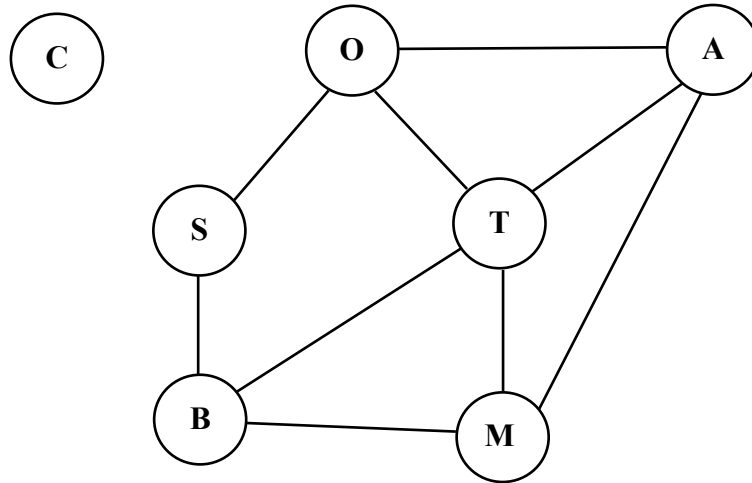
توزیع توام متغیرها را با توجه به گراف بیزی بنویسید.

زیربخش دوم (۱ نمره)

markov blanket مربوط به متغیر T را بنویسید.

زیربخش سوم (۲ نمره)

گراف مارکوف زیر را در نظر بگیرید. آیا این گراف Perfect I-Map مربوط به گراف بیزی است؟ چرا؟



زیربخش چهارم (۱ نمره)

با توجه به گراف مارکوف زیربخش سوم، بیان کنید آیا این گراف chordal است؟ چرا؟

زیربخش پنجم (۳ نمره)

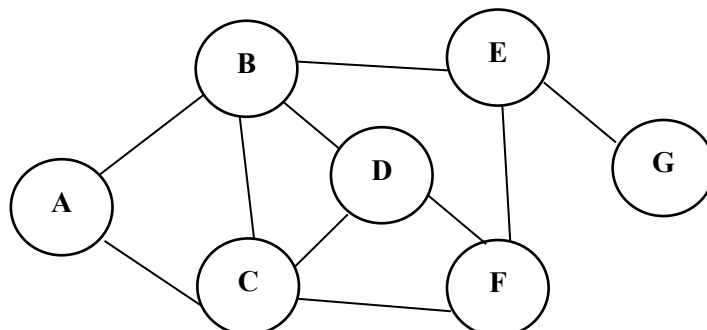
توزیع احتمال توام گراف مارکوف زیربخش سوم را بر اساس maximal cliques بنویسید.

زیربخش ششم (۶ نمره)

درستی دو عبارت زیر را نشان دهید:

- اگر بخواهیم متغیر  $C$  را حذف کنیم، یعنی توزیع احتمال توأم روی سایر متغیرها را بدست آوریم، آنگاه:
- در حالت گراف بیزی، فقط کفایست فاکتور  $p(C)$  را حذف کنیم.
  - در حالت گراف مارکوف، در صورتی که  $\int_{-\infty}^{+\infty} \phi(C) dC = 1$  فقط کفایست فاکتور  $\phi(C)$  را حذف کنیم.

با توجه به گراف مارکوف زیر به سوالات پاسخ دهید.



زیربخش اول (۳ نمره)

توزیع احتمال توام گراف مارکوف را بر اساس maximal cliques بنویسید.

زیربخش دوم (۸ نمره)

درستی یا نادرستی موارد زیر را با ذکر دلیل بیان کنید:

a.  $G \perp A$

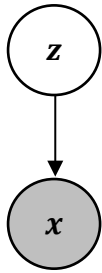
b.  $F \perp A \mid \{D, C\}$

c.  $G \perp C \mid E$

d.  $p(A \mid B, C) = p(A \mid B, C, E)$

زیربخش سوم (۲ نمره)

در توزیع توام بدست آمده در زیربخش اول، اگر تابع پتانسیل  $\phi(E, G)$  را ۵ برابر کنیم، توزیع احتمال حاصل چه تغییری می‌کند؟



شبکه بیزی روبرو را به همراه توزیع‌های شرطی زیر در نظر بگیرید.

$$p(z) = e^{-z} \quad z > 0$$

$$p(x|z) = z e^{-zx} \quad x > 0$$

میخواهیم با استفاده از Variational Inference، توزیع احتمال پسین  $p(z|x)$  را تقریب بزنیم. برای این کار، توزیع پسین را با توزیع زیر تقریب می‌زنیم:

$$q(z) = \theta^2 z e^{-\theta z} \quad z > 0$$

به علاوه برای توزیع  $q(z)$  داریم:

$$\mathbb{E}_q[z] = \frac{2}{\theta}$$

مقدار بهینه پارامتر  $\theta$  را بیابید.



## سوال دوم :

### بخش اول

مدل های VAE از جمله مدل های مولد هستند که با افزودن یک مؤلفه ی احتمالاتی به ساختار اتوانکدرها، قابلیت تولید داده های جدید را فراهم می کنند. در این بخش به پیاده سازی این مدل می پردازیم

#### زیر بخش اول (۲ نمره)

تابع هزینه استفاده شده برای آموزش این مدل مولد مطابق رابطه (۱) است. چرا به طور مستقیم احتمال درست نمایی داده ها را افزایش نمی دهیم و همچنین ترم های این عبارت چگونه به بهینه سازی مدل می انجامد؟

$$E_{q(z|x)}[\log p(x|z)] - D_{KL}(q(z|x)||p(z)) \quad (1)$$

#### زیر بخش دوم (۱ نمره)

دادگان مورد استفاده در این سوال `dsprites` است. توضیح مختصری درباره این دادگان بدهید و چند نمونه از آن را نمایش دهید.

#### زیر بخش سوم (۲ نمره)

در اتصال بخش انکودر به دیکودر در VAE ها، برخلاف اتوانکودرهای معمولی، ما یک نمونه از توزیع نهان را به جای مقدار پارامتری آن استفاده می کنیم. این کار باعث می شود که زنجیره ی گرادیان ها در انکودر قطع شود و به طور مستقیم اعمال گرادیان امکان پذیر نباشد. برای حل این مشکل چه راه حلی وجود دارد؟ توضیح دهید.

#### زیر بخش چهارم (۱۰ نمره)

با توجه به نکات گفته شده مدل VAE مورد نظر را آموزش دهید. برای آموزش میتوانید از شبکه پیشنهادی و هایپر پارامترهای پیشنهادی در جدول ۱ و ۲ استفاده کنید. در نهایت مقدار هر دو قسمت بازسازی و KL و همچنین مقدار کل تابع هزینه را در طول آموزش در یک نمودار نشان دهید.

پس از اتمام آموزش ۸ نمونه تصادفی از دادگان را از مدل عبور دهید و نسخه بازسازی شده و اولیه را با هم مقایسه کرده و عملکرد مدل را بررسی کنید.

جدول ۱: لایه‌های پیشنهادی برای شبکه

Encoder	Decoder
Conv2d(1, 32)	Linear (h_dim, 8192)
Conv2d(32, 64)	Conv2dT(128, 64)
Conv2d(64, 128)	Conv2dT(64, 32)
Flatten ()	Conv2dT(32, 1)
Linear (8192, h_dim)	Sigmoid()

جدول ۲: هایپرپارامترهای پیشنهادی

Parameter	Value
Epochs	50
Optimizer	Adam
Initial learning rate	0.001
Image dimension	(64,64)

#### زیر بخش پنجم (۲ نمره)

در ادامه توسعه مدل‌های VAE نسخه بعدی مورد بررسی مدل  $\beta - VAE$  است که رابطه ریاضی آن مطابق رابطه (۲) است. این مدل چه بهبودی نسبت به نسخه اصلی دارد؟

$$E_{q(z|x)}[\log p(x|z)] - \beta D_{KL}(q(z|x)||p(z)) \quad (2)$$

#### زیر بخش ششم (۱۲ نمره)

در نهایت در این قسمت به آموزش یک مدل  $\beta - VAE$  خواهیم پرداخت. برای این قسمت از دو مقدار بتای مختلف بالا و پایین (هر دو بالاتر از یک) برای آموزش استفاده کنید. مشابه قسمت قبل به ازای هر مقدار بتا یک نمودار برای توابع هزینه تشکیل دهید. همچنین ۸ نمونه تصادفی را به کمک این دو مدل بازسازی کرده و کنار هم نمایش دهید و نتایج را بررسی کنید.

---

#### زیر بخش هفتم (۸ نمره)

از مزایای مدل  $\beta - VAE$  نسبت به مدل VAE معمولی میتوان به مستقل تر بودن ابعاد فضای نهان اشاره کرد. برای اندازه گیری این متغیر میتوان از معیار [1] MIG استفاده کرد. در مورد این معیار توضیح دهید و سپس به ازای سه مدل موجود آن را به دست بیاورید و با هم مقایسه کنید. (این معیار را هم به ازای فاکتورهای مختلف و هم به صورت کلی به دست آورید).

---

#### زیر بخش هشتم (۳ نمره)

برای یکی از مدل های موجود ابعاد پنهان را به دست آورده و نمودار پراکندگی داده ها در فضای دو مؤلفه اصلی (تحلیل مؤلفه های اصلی در PCA) را رسم کنید. این نمودارها را تحلیل کرده و بگویید آیا با نتایج قسمت قبل همخوانی دارد یا خیر و در صورت عدم تطابق در مورد علت آن توضیح دهید.

زیر بخش اول (۵ نمره)

یکی از انواع راهکارهای پیشنهاد شده برای افزایش کیفیت فضای نهان در VAE ها استفاده از VQ-VAE [2] است. تفاوت این مدل با مدل پایه را توضیح دهید و بگویید گسسته سازی در این مدل به چه معناست و چه مزیتی به همراه دارد.

زیر بخش دوم (۵ نمره)

توضیح دهید که در مقاله *VampPrior* [3] نویسندگان چگونه توزیع پسین را تخمین می‌زنند و چگونه این کار منجر به استفاده بهتر از متغیرهای نهان در مقایسه با توزیع نرمال استاندارد میشود.

زیر بخش سوم (۵ نمره)

مدل  $SC - VAE$  [4] را توضیح دهید و بگویید الگوریتم ISTA چه نقشی در این مدل دارد. تنک بودن بازنمایی های نهان چه مزیتی نسبت به مدل پایه دارد؟

1. Chen, R. T. Q., Li, X., Grosse, R. B., & Duvenaud, D. (2018). Isolating sources of disentanglement in variational autoencoders. In Advances in Neural Information Processing Systems (NeurIPS 2018).
2. van den Oord, A., Vinyals, O., & Kavukcuoglu, K. (2017). Neural discrete representation learning. In Advances in Neural Information Processing Systems (NeurIPS 2017).
3. Tomczak, J. M., & Welling, M. (2018). VAE with a VampPrior. In Proceedings of the Twenty-First International Conference on Artificial Intelligence and Statistics (AISTATS 2018) (Vol. 84, pp. 1214-1223). PMLR.
4. Xiao, P., Qiu, P., Ha, S., Bani, A., Zhou, S., & Sotiras, A. (2024). Sparse Coding-based Variational Autoencoder with Learned ISTA (SC-VAE). Pattern Recognition, 161, Article 111187.

## نکات تحویل

- مهلت ارسال این تمرین تا پایان روز "سه شنبه ۱۳ آبان ماه" خواهد بود.
- این زمان قابل تمدید نیست و در صورت نیاز می‌توانید از grace time استفاده کنید.
- در نظر داشته باشید که حداکثر مهلت آپلود تمرین در سامانه تا ۷ روز پس مهلت تحویل است و پس از آن سامانه بسته خواهد شد.
- پیاده سازی با زبان برنامه نویسی پایتون باید باشد و کدهای شما باید قابل اجرا بوده و به همراه گزارش آپلود شوند.
- انجام این تمرین به صورت یک نفره می‌باشد.
- در صورت مشاهده هر گونه تشابه در گزارش کار یا کدهای پیاده‌سازی، این امر به منزله تقلب برای طرفین در نظر گرفته خواهد شد.
- استفاده از کدهای آماده بدون ذکر منبع و بدون تغییر به منزله تقلب خواهد بود و نمره تمرین شما صفر در نظر گرفته می‌شود.
- در صورت رعایت نکردن فرمت گزارش کار نمره گزارش به شما تعلق نخواهد گرفت.
- تحویل تمرین به صورت دستنویس قابل پذیرش نیست.
- تمامی تصاویر و جداول مورد استفاده در گزارش کار باید دارای توضیح (caption) و شماره باشند.
- بخش زیادی از نمره شما مربوط به گزارش کار و روند حل مسئله است.
- لطفا گزارش، فایل کدها و سایر ضمائم مورد نیاز را با فرمت زیر در سامانه بارگذاری نمایید.
- HW1\_[Lastname]\_[StudentNumber].zip
- در صورت وجود سوال و یا ابهام می‌توانید از طریق رایانامه زیر با موضوع TAI\_HW1 با دستیاران آموزشی در ارتباط باشید:

○ سوال اول

[ma.moghimi202@gmail.com](mailto:ma.moghimi202@gmail.com)

○ سوال دوم

[S.m.moosavi000@ut.ac.ir](mailto:S.m.moosavi000@ut.ac.ir)

با آرزوی سلامتی و موفقیت روزافزون.