

به نام خدا



دانشگاه تهران
پردیس دانشکده‌های فنی
دانشکده برق و کامپیوتر



درس: مدل‌های مولد عمیق

مدرس: دکتر مصطفی توسلی‌پور

تمرین شماره ۱

مهرماه ۱۴۰۳

۳.....	بخش اول – PGM
۳.....	زیربخش اول
۴.....	زیربخش دوم
۴.....	زیربخش سوم
۵.....	زیربخش چهارم
۵.....	زیربخش پنجم
۶.....	بخش دوم – VAE
۶.....	زیربخش اول – پیاده سازی
۶.....	سوال اول
۶.....	سوال دوم
۷.....	سوال سوم
۹.....	زیربخش دوم – تئوری
۹.....	سوال اول
۱۰.....	سوال دوم
۱۰.....	سوال سوم
۱۱.....	نکات تحویل

بخش اول – PGM

زیربخش اول (۱۵ نمره)

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

فرض کنید که مسئول نگهداری و تامین سلامت سرورهای یک شرکت هستید. برای حفظ کارکرد مناسب سرورها لازم است که دما محیطی که سرورها در آن قرار دارند از یک حد آستانه‌ای کمتر باشد. دما اتاق سرور ها مدام توسط یک سیستم کنترل دما پایش می‌شود. اگر دما از یک حد آستانه بیشتر شد، یک پیغام متنی برای مسئول نگهداری سرورها ارسال می‌شود.

دمای اتاق سرورها در دو حالت زیر ممکن است از حد آستانه تعریف شده عبور کند:

(۱) در اتاق سرور ها برای مدت طولانی باز بماند.

(۲) سیستم تهویه اتاق سرور ها دچار نقص شود.

هنگامی که در اتاق سرورها برای مدت طولانی باز بماند، آژیر اخطار به صدا در خواهد آمد.

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 و همچنین تمامی متغیر ها دودویی هستند.

متغیر T: دمای اتاق از یک حد آستانه بیشتر شود.

متغیر O: در اتاق سرور برای مدت طولانی باز بماند.

متغیر M: پیامی متنی دریافت شود.

متغیر A: سیستم تهویه هوا دچار نقص شود.

متغیر S: آژیر اخطار به صدا در بیاید.

(۱) با توجه به توضیحات گفته شده، شبکه بیزی را آن را رسم کنید.

(۲) توزیع احتمال توام با توجه به متغیر های تعریف شده و شبکه بیزی بدست آورید.

(۳) با توجه به شبکه بیزی که رسم کردید، صحت عبارتهای زیر را با ذکر دلیل بیان کنید.

$$O \perp A \quad (۳.۱)$$

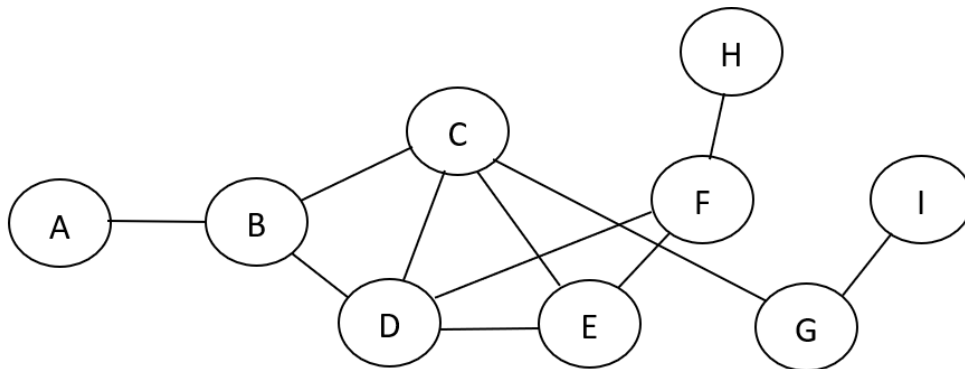
$$O \perp A \mid M \quad (۳.۲)$$

$$S \perp M \quad (۳.۳)$$

$$S \perp M \mid O \quad (۳.۴)$$

زیربخش دوم (۸ نمره)

با توجه به گراف مارکوف زیر درستی و نادرستی موارد بیان شده را با ذکر دلیل بیان کنید.



$$I \perp F \quad (۱)$$

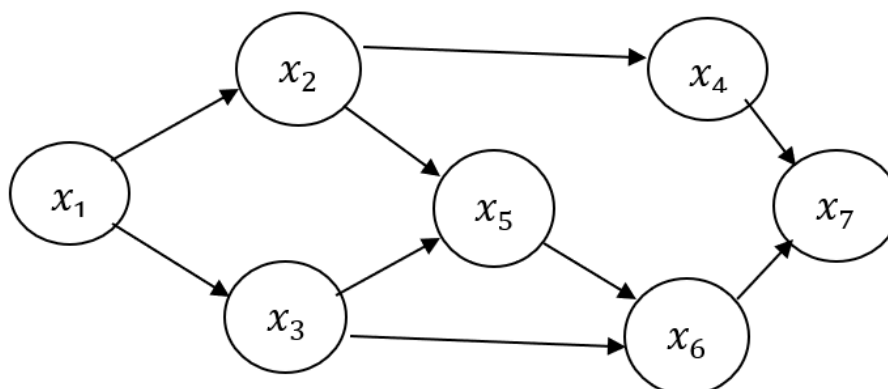
$$B \perp I \mid F \quad (۲)$$

$$A \perp G \mid C, E \quad (۳)$$

$$P(B|C, D, F) = P(B|C, D) \quad (۴)$$

زیربخش سوم (۱۲ نمره)

با توجه به گراف بی‌زی زیر به سوالات مربوط به آن پاسخ دهید.



(۱) توزیع توام متغیرها را با توجه به گراف بی‌زی، فرم فاکتور شده آن را بنویسید.

(۲) Markov blanket مربوط به متغیر x_6 را بدست بیاورید.

(۳) گراف مارکوف مربوط به گراف بی‌زی بالا را رسم کنید.

۴) با توجه به گراف مارکوف بدست آمده در قسمت ۳ با ذکر دلیل بیان کنید آیا گراف بدست آمده perfect I-map مربوط به گراف بی‌زی است؟ چرا؟

۵) با توجه به گراف مارکوف بدست آمده در قسمت ۳ بیان کنید آیا گراف بدست آمده chordal است؟ چرا؟

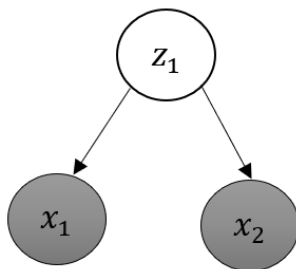
زیربخش چهارم (۵ نمره امتیازی)

درستی یا عدم درستی گزاره زیر را برای گراف G ، در هر دو حالت که گراف G بی‌زی یا گراف مارکوف باشد، اثبات کنید.

اگر یک لیست استقلال L را داشته باشیم و گراف G ، perfect I-map برای این لیست استقلال باشد. اگر گراف G' با حذف یک یال از گراف G ساخته شود. گراف G' ، I-map لیست استقلال L است.

زیربخش پنجم (۱۵ نمره)

شبکه بی‌زی رو به رو را با توزیع های زیر در نظر بگیرید



$$p(z_1) = N(0,1)$$

$$p(x_1|z_1) = N(x_1|z_1, \sigma_1^2)$$

$$p(x_2|z_1) = N(x_2|z_1, \sigma_2^2)$$

با استفاده از variational inference توزیع احتمال پسین $p(z_1|x_1, x_2)$ را تخمین بزنید. خانواده توزیع های تخمینی را توزیع لاپلاس در نظر بگیرید. توجه کنید جواب فرم بسته دارد.

بخش دوم – VAE

زیربخش اول – پیاده‌سازی

سوال اول: (۳ نمره)

در مدل‌های VAE، هدف یادگیری توزیع نهانی است که به توزیع واقعی داده‌ها نزدیک باشد. از آنجا که محاسبه مستقیم احتمال درست‌نمایی داده‌ها ($\log p(x)$) نیاز به محاسبات سنگینی دارد، از تابع «حد پایین شواهد» (ELBO) استفاده می‌کنیم. فرمول ELBO به شکل زیر است:

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

در اینجا:

$E_{q(z|x)}[\log p(x|z)]$ نشان‌دهنده بازسازی داده‌ها است.

D_{KL} فاصله کولبک-لیبلر بین توزیع تقریبی $q(z|x)$ و توزیع نهان $p(z)$ است.

با بیشینه کردن ELBO، بازسازی بهبود یافته و توزیع نهان به توزیع واقعی داده‌ها نزدیک‌تر می‌شود.

با انجام محاسبات ریاضیاتی مورد نیاز، اثبات کنید که ELBO به صورت یک حد پایین از احتمال درست‌نمایی داده‌ها ($\log p(x)$) است و چگونه کمینه‌سازی D_{KL} بین توزیع تقریبی و توزیع واقعی منجر به بهینه‌سازی مدل می‌شود.

سوال دوم: (۳ نمره)

همانطور که در سوال قبل گفته شد، برای آموزش یک مدل VAE، می‌توان عبارت ELBO را بیشینه کرد. به این منظور نیاز است تا از توزیع $q(z|x)$ نمونه‌برداری شود. این نمونه‌برداری فرآیند آموزش VAE را به یک فرآیند احتمالاتی^۱ مبدل می‌کند. انجام محاسبات بازگشتی^۲ و اعمال گرادیان در این محاسبات،

^۱ Probabilistic
^۲ Back propagation

امکان پذیر نخواهد بود. برای برطرف کردن این مشکل از چه تکنیکی استفاده می شود؟ با نوشتن روابط مرتبط، این تکنیک را توضیح دهید.

سوال سوم:

الف) (۱ نمره)

در این سوال، می خواهیم یک شبکه VAE را پیاده سازی کرده و بر روی دادگان تصویری صورت آموزش دهیم.

ابتدا تصاویر را از پوشه به اشتراک گذاشته شده دریافت کرده و پس از پیاده سازی کلاس `data loader`، چند تصویر تصادفی از مجموعه دادگان را نمایش دهید.

ب) (۱۰ نمره)

یک کلاس برای پیاده سازی مدل VAE تعریف کنید و از ساختار ذکر شده در جدول ۱ برای تعریف `encoder` و `decoder` مدل استفاده کنید. همچنین، ترفند `reparameterization trick` را در کلاس مدل پیاده سازی کرده و برای استفاده از آن، ساختار (لایه های) مورد نیاز را به `encoder` و `decoder` اضافه کنید.

جدول ۱- تعریف لایه های مورد استفاده در شبکه β -VAE

Encoder Layers	Decoder Layers
Conv2d(3,32)	Linear(4096, h_dim)
Conv2d(32,64)	Reshape()
Conv2d(64,64)	Conv2d(64,64)
Conv2d(64,64)	Conv2d(64,64)
Flatten()	Conv2d(64,32)
Linear(h_dim, 4096)	Conv2d(32,3)

در ایجاد ساختار شبکه، پس از هر لایه از تابع غیرخطی ساز Relu استفاده کرده و برای لایه‌های کانولوشنی، از batch normalization و همچنین dropout با مقدار دلخواه استفاده کنید.

(ج) (۱۰ نمره)

تابع هزینه مورد استفاده در VAE را پیاده‌سازی کنید و شبکه را بر روی دادگان صورت آموزش دهید. به منظور آموزش شبکه از مقادیر جدول ۲ استفاده کنید. همچنین، مقادیر محاسبه شده به ازای هر یک از بخش‌های KL و reconstruction در تابع هزینه را برای تصاویر آموزش و اعتبار سنجی در حین آموزش را در یک نمودار رسم کنید.

جدول ۲ - پارامترهای مورد استفاده در آموزش مدل

Parameter	Value
Image dimensions	(128, 128)
Batch size	128
Transformation	Normalize, to tensor
Train/test ratio	0.8/0.2
Optimizer	Adam
Initial Learning Rate	0.0005
Epochs	1000

(د) (۴ نمره)

پس از آموزش شبکه، ۳۲ تصویر را ابتدا با استفاده از تصاویر واقعی (اعمال decoder بر خروجی بردار نهان تصویر واقعی که از encoder عبور کرده است) و سپس با استفاده از نویز تصادفی (اعمال decoder بر بردار نهانی که با استفاده از نویز تصادفی ساخته شده است) بازسازی کرده و در گزارش خود نمایش دهید.

(و) (۵ نمره)

ابعاد در فضای نهان دارای معنا بوده و هر یک بیانگر یکی از ویژگی‌های توزیع بدست آمده است. با توجه به اینکه دادگان مورد استفاده در آموزش مدل تنها تصاویر صورت انسان بوده و فقط شامل دو کلاس smile و non-smile بوده است، انتظار می‌رود که هر یک از ابعاد بردار نهان مرتبط با یکی از اجزا صورت باشد. همچنین انتظار می‌رود که یکی از این ابعاد بیانگر فرم لب، بر صورت باشد. یکی از ساده‌ترین راهکارها برای استفاده از این بعد به ترتیب زیر است.

ابتدا بردار میانگین تمام تصاویر کلاس smile و non-smile را به صورت جدا در فضای نهان بدست آورید (با استفاده از تمام تصاویر در این کلاس‌ها). با کسر بردار نهان کلاس smile از بردار کلاس non-smile انتظار می‌رود که بیشترین تفاوت در بعد مرتبط با ویژگی‌های لب رخ دهد. حال یک تصویر واقعی از دادگان را انتخاب کرده و با کسر یا جمع ضریب‌های متفاوت از بردار بدست آمده با بردار تصویر انتخاب شده در فضای نهان، ۶ تصاویر با درصد لبخندهای متفاوت تولید کرده و نمایش دهید.

زیربخش دوم – تئوری

سوال اول: (۶ نمره)

یکی از انواع مختلف VAE ها، مدل $\beta - VAE$ است که تأکید ویژه‌ای بر کشف عوامل نهان جداشده^۱ دارد. عوامل نهان جداشده به این نکته اشاره دارد که هر متغیر در بازنمایی نهان تنها به یک عامل تولید شده خاص حساس باشد و نسبت به سایر عوامل نسبتاً ثابت بماند.

در ادامه همان هدف در VAE ها، ما می‌خواهیم احتمال تولید داده‌های واقعی را بیشینه کنیم، در حالی که فاصله بین توزیع نهان واقعی و تخمین زده شده را کوچک نگه داریم (مثلاً کمتر از یک ثابت کوچک δ). با استفاده از مفهوم لاگرانژین و استفاده از شرایط KKT، ثابت کنید که تابع هزینه $\beta - VAE$ به صورت زیر خواهد بود:

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

^۱ Disentangled latent factors

سوال دوم: (۷ نمره)

در تابع هزینه $\beta - VAE$ ، پارامتر β کنترل کننده معاوضه بین کیفیت بازسازی و جداسازی عوامل نهفته است. افزایش مقدار β منجر به جداسازی بهتر ابعاد فضای نهان شده اما خطای بازسازی مدل را افزایش می دهد. برای برطرف کردن این مشکل، راهکارهای متفاوتی ارائه شده است که یکی از راهکارها در مقاله Disentangling by Factorization^۱ معرفی شده است. هدف نویسندگان از ارائه این مدل، ایجاد یک معاوضه بهتر بین خطای بازسازی و جداسازی ابعاد فضای نهان است. با انجام محاسبات مورد نیاز، نشان دهید که عبارت KL موجود در تابع هزینه $\beta - VAE$ در واقع از دو بخش mutual information بین توزیع واقعی و توزیع فضای نهان و KL بین $p(z)$ و $q(z)$ است. توضیح دهید که چگونه از این عبارت برای محاسبه تابع هدف factor-VAE استفاده شده است.

سوال سوم: (۶ نمره)

تابع هدف factor-VAE به صورت زیر بیان شده است.

$$\frac{1}{N} \sum_{i=1}^N [E_{q(z|x^{(i)})} [\log p(x^{(i)}|z)] - KL(q(z|x^{(i)})||p(z)) - \gamma KL(q(z)||\bar{q}(z))]]$$

محاسبه مقادیر $q(z)$ و $\bar{q}(z)$ نیاز به محاسبات سنگین و پیچیده در تمام ابعاد فضا داشته و عملاً غیرقابل حل است. به همین منظور، نویسندگان مقاله چند روش برای تخمین این دو توزیع پیشنهاد کرده اند. روش های پیشنهاد شده را نام برده و روش انتخابی را به طور مختصر توضیح دهید.

^۱ <https://arxiv.org/abs/1802.05983>

نکات تحویل

- مهلت ارسال این تمرین تا پایان روز "شنبه ۱۲ 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 ضمیمات 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

smousavichashmi@ut.ac.ir

○ سوال دوم

ftaherinezhad@ut.ac.ir یا تلگرام FatiTaheriN@

با آرزوی سلامتی و موفقیت روزافزون.