



- مهلت ارسال پاسخ تا ساعت ۲۳:۵۹ روز مشخص شده است.
- همکاری و هم‌فکری شما در انجام تمرین مانعی ندارد اما پاسخ ارسالی هر کس حتما باید توسط خود او نوشته شده باشد.
- در صورت هم‌فکری و یا استفاده از هر منابع خارج درسی، نام هم‌فکران و آدرس منابع مورد استفاده برای حل سوال مورد نظر را ذکر کنید.
- لطفا تصویری واضح از پاسخ سوالات نظری بارگذاری کنید. در غیر این صورت پاسخ شما تصحیح نخواهد شد.

سوالات نظری (۵ + ۶۰ نمره)

مسئله‌ی ۱. (۱۰ نمره)

(آ) یکی از راه‌های معمول برای تخمین گرادیان یک امید ریاضی، استفاده از رابطه‌ی زیر است:

$$\nabla_{\theta} \mathbb{E}_{z \sim q_{\theta}(z)} [f(z)] \approx \frac{1}{N} \sum_{i=1}^N f(z^i) \cdot \nabla_{\theta} \ln q_{\theta}(z^i)$$

که در آن هر z^i نمونه‌ی مستقلی از توزیع $q_{\theta}(z)$ می باشد. درستی این رابطه را نشان دهید و بیان کنید که چطور می توانیم از آن در VAE استفاده کنیم. با مراجعه به این مقاله مشکلی که در استفاده از این روش وجود دارد را بیان کنید.

(ب) به صورت شهودی بیان کنید که روش Reparameterization چگونه می‌تواند این مشکل را حل کند؟

(ج) در بسیاری از موارد تابع خطای رمزگشای VAE را خطای MSE در نظر میگیریم. این در حالی است که هدف ما بیشینه کردن تابع $\mathbb{E}_{z \sim q_{\theta}(z|x)} \ln p_{\theta}(x|z)$ می باشد. در چه صورتی و با چه فرض‌هایی این دو کار معادل یکدیگر هستند؟

مسئله‌ی ۲. (۱۵ نمره)

(آ) در یک VAE اگر داده ورودی از نوع باینری باشد (تصویری با پیکسل‌های ۰ و ۱)، می‌توان به جای توزیع گاوسی چندمتغیره روی خروجی کدگشا، از توزیع برنولی چندمتغیره استفاده کرد. توزیع خروجی کدگشا را به شکل برنولی چندمتغیره در نظر بگیرید و تابع فعال سازی آخرین لایه کدگشا را sigmoid در نظر بگیرید. اثبات کنید که در تابع هزینه این شبکه یک جمله‌ای به شکل Binary Cross Entropy ظاهر می‌شود.

(ب) تکنیک Reparameterization روی بسیاری از توزیع‌های پیوسته قابل اعمال است. تحقیق کنید که چگونه می‌توان از این تکنیک برای یک توزیع categorical استفاده کرد؟

(ج) یکی از نسخه‌های تغییر یافته VAE مقاله مربوط به beta-VAE می باشد. در این روش یک ضریب beta ای پشت یکی از توابع هزینه قرار میگیرد. درباره این روش تحقیق کنید و بیان کنید:

- (۱) با انجام چه روندی از محاسبات، این ضریب در تابع هدف این روش ظاهر می‌شود؟
- (۲) هدف از افزودن ضریب β چیست و اضافه کردن آن چه تاثیری روی ویژگی‌های فضای نهان یادگرفته‌شده توسط مدل دارد؟

مسئله‌ی ۳. (۱۰ نمره)

(آ) همانطور که می‌دانیم، دو مدل GAN و VAE از مهم‌ترین و شناخته‌شده‌ترین مدل‌های generative در یادگیری عمیق می‌باشند. یکی از مهم‌ترین کاربردهای آن‌ها، تولید تصاویر و data augmentation می‌باشد. در این صورت، با فرض استفاده از مجموعه داده یکسان و فرآیند آموزش نسبتاً کامل، آیا به طور کلی کیفیت تصاویر تولید شده توسط یکی از این مدل‌ها بر دیگری برتری دارد؟ لطفاً پاسخ خود را با دلیل و در صورت نیاز اثبات ریاضی تشریح نمایید. می‌توانید از این [مقاله](#) استفاده نمایید.

(ب) تابع ReLU یکی از پرکاربردترین توابع فعالسازی مورد استفاده در شبکه‌های یادگیری عمیق می‌باشد، اما در برخی کاربری‌های خاص همانند بخش Generative در روش GAN می‌تواند منجر به ایجاد مشکل در فرآیند آموزش شود، به عبارت دیگر شبکه مولد ما آموزش نخواهد دید. در این مورد استثنا، توصیه می‌شود به جای ReLU، از Leaky ReLU به عنوان تابع فعالسازی استفاده گردد. لطفاً علت بروز مشکل به هنگام استفاده از ReLU را به طور کامل توضیح داده و تشریح کنید که به چه علت استفاده از Leaky ReLU می‌تواند مشکل را حل کند.

مسئله‌ی ۴. (۱۵ نمره)

(آ) یکی از مشکلات شایع در شبکه‌های GAN مشکل Mode Collapse می‌باشد که باعث می‌شود شبکه GAN به یک حالت عدم آموزش رسیده و به طور متوالی خروجی‌های یکسان تولید کند. این مشکل را به طور کامل تشریح کرده و راه حل‌های احتمالی به منجر به غلبه بر این مشکل می‌شوند را بیان نمایید.

(ب) معماری کلی و تابع هزینه مدل W-GAN را تشریح نمایید و تفاوت‌های آن با مدل پایه GAN را توضیح دهید. آیا تابع هزینه این مدل کمکی به برطرف شدن مشکل Mode Collapse خواهد کرد؟ توضیح دهید.

مسئله‌ی ۵. (۵ + ۱۰ نمره)

(آ) یکی از چالش‌های مهم که مدل‌های GAN تاثیر به سزایی در برطرف شدن آن‌ها دارند موضوع image2image translation می‌باشد. یکی از مدل‌هایی که به طور خاص برای این امر توسعه داده شده است مدل Cycle GAN می‌باشد. معماری دوگانه این مدل را تشریح کرده و مزایای آن را در ارتباط با چالش مذکور نسبت به مدل پایه GAN بیان نمایید.

(ب) فرض کنید تحت شرایطی بسیار خاص از شما درخواست شده با تعداد داده محدود (در حدی که معماری پایه GAN به سختی به کمک آن train می‌شود) یک مدل Cycle GAN را train نمایید. در صورتی که الزام به استفاده از Cycle GAN باشد، راه حل پیشنهادی شما چیست؟ (دقت شود سوال ابتکاری بوده و پاسخ کاملاً یکتا ندارد، اما بایستی هر راه حل پیشنهادی با تحلیل کامل و حتی استفاده از روابط ریاضی در صورت نیاز تشریح گردد).

سوالات عملی (۵ + ۴۰ نمره)

مسئله ۶. (۲۰ نمره)

هدف این سوال طراحی یک شبکه ساده GAN می باشد که بتواند تصاویر دادگان MNIST را تولید نماید. فایل GAN.ipynb را براساس موارد خواسته شده تکمیل نمایید. دقت کنید که بخش هایی از نمره بستگی به نتایج بدست آمده دارند. باتوجه به نکات تمرین و مواردی که در کلاس عنوان شده، ساختار شبکه و تابع خطا و پارامترهای دیگر را طوری طراحی و انتخاب نمایید که در نهایت تصاویر باکیفیتی توسط Generator تولید شود و فرایند آموزش پایدار باشد. در نهایت فایل تکمیل شده به همراه نتایج را بعلاوه فایل پارامترهای شبکه Generator ای که آموزش داده اید به همراه دیگر بخش های تمرین ارسال نمایید.

مسئله ۷. (۵ + ۲۰ نمره)

Conditional VAE یکی از ورژن های modified شده VAE بوده که بر خلاف VAE کلاسیک، متغیرهای مورد نیاز را به صورت conditioned نسبت به برخی متغیرهای تصادفی تخمین می زند. در این تمرین هدف مقایسه خروجی این دو مدل بر روی مجموعه داده MNIST می باشد. لطفا کد هر دو روش پیاده سازی شده و خروجی آن ها از بعد میزان وضوح تصاویر تولید شده مقایسه گردد.

برای طراحی شبکه های VAE و CVAE، به طور کلی محدودیت چندانی وجود ندارد اما پیشنهاد می شود برای قسمت Encoder، سه لایه کانولوشن دو بعدی به ترتیب با ۱۶، ۳۲، و ۳۲ لایه به همراه MaxPool دو بعدی ۲ در ۲ پس از هرکدام طراحی شود. برای قسمت Decoder نیز دو لایه خطی به ترتیب با ۳۲ و ۶۴ لایه طراحی گردد. برای تابع هزینه لطفا از Binary Cross Entropy به همراه KL Divergence استفاده شود. برای Optimizer نیز از Adam استفاده گردد. مابقی پارامترها همانند mean می تواند به صورت customize شده انتخاب گردد و بسته به خروجی بهتر تغییر کند. برای راهنمایی بیشتر می توانید از کد موجود در [این لینک](#) استفاده کنید. لطفا کد را کپی نکرده و صرفا برای کمک و الهام گیری کد زنی خود از آن استفاده شود. توجه شود حتی ساختار پیشنهادی شبکه بسته به صلاح دید شخصی شما قابل تغییر بوده فقط توجه شود که نوشتن گزارش بخش عملی الزامی بوده و دارای نمره می باشد لذا حتما تمامی مراحل اعم از ساختار شبکه ها و پارامترها باید به طور کامل در گزارش توضیح داده شوند. برای پیاده سازی نیز تنها مجاز به استفاده از کتابخانه pytorch می باشید.

توضیحات کلی:

به وضوح مشورت و همفکری در حل سوالات هیچگونه ای مشکلی ایجاد نخواهد کرد اما جواب سوالات به هیچ عنوان نباید یکسان باشد. هر فرد باید جداگانه و Unique پاسخ های خود را بنویسد و در صورت شباهت بسیار زیاد نمره بین افرادی که پاسخ های بسیار مشابه دارند تقسیم خواهد گشت. در صورتی که کد شما در بخش عملی به خطا برخورد و خطا قابل برطرف کردن نبود، نمره شما بر اساس کیفیت کد به میزان قابل قبولی لحاظ خواهد گشت. لذا لطفا پاسخ سوال های عملی را خالی نگذارید. برای سوال عملی لطفا کد و گزارش هردو ارسال شوند. کمبود یکی از آن ها منجر به از دست رفتن نمره خواهد شد.

موفق باشید :