



تحلیل هوشمند تصاویر زیست پزشکی

نیم سال اول ۰۳-۰۲

مدرس: محمدحسین رهبان

تمرین ششم

- مهلت ارسال پاسخ تا ساعت ۲۳:۵۹ روز مشخص شده است.
- در طول ترم امکان ارسال با تاخیر تمرین ها بدون کسر نمره تا سقف ۱۲ روز وجود دارد. محل بارگزاری جواب تمرین ها بعد از ۴ روز بسته خواهد شد و پس از گذشت این مدت، پاسخ های ارسال شده پذیرفته نخواهند شد.
- توجه داشته باشید که نوت بوک های شما باید قابلیت باز اجرای ۱۰۰ درصد داشته باشند و در صورت نیاز به نصب یک کتابخانه یا دسترسی به یک فایل، مراحل نصب و دانلود (از یک محل عمومی) در نوت بوک وجود داشته باشد.
- هم فکری در انجام تمرین مانعی ندارد، فقط توجه داشته باشید که پاسخ تمرین حتما باید توسط خود شخص نوشته شده باشد. همچنین در صورت هم فکری در هر تمرین، در ابتدای جواب تمرین نام افرادی که با آن ها هم فکری کرده اید را حتما ذکر کنید.
- برای پاسخ به سوالات نظری در صورتی که از برگه خود عکس تهیه می کنید، حتما توجه داشته باشید که تصویر کاملا واضح و خوانا باشد. در صورتی که خوانایی کافی را نداشته باشد، تصحیح نخواهد شد.
- محل بارگذاری سوالات نظری و عملی در هر تمرین مجزا خواهد بود. به منظور بارگذاری بایستی تمارین تئوری در یک فایل زیپ با نام `IABI_Theo_hw6_[First-Name]_[Last-Name]_[Student-Id].zip` و تمارین عملی نیز در یک فایل مجزای زیپ با نام `IABI_Prac_hw6_[First-Name]_[Last-Name]_[Student-Id].zip` بارگذاری شوند.
- در صورت وجود هرگونه ابهام یا مشکل، در کوثرای درس آن مشکل را بیان کنید و از پیغام دادن مستقیم به دستیاران آموزشی خودداری کنید.

بخش نظری (۵۰ نمره)

سوال اول (۱۵ نمره)

۱. در عرصه مدل های deep generative، مدل های Diffusion Models، Generative Adversarial Networks (GANs) و Variational Autoencoders (VAEs) از نظر معماری، روش های آموزش، کیفیت و تنوع نمونه ها و همچنین کاربرد در زمینه های مختلف بررسی کرده و مقایسه کنید؟
۲. چرا از Diffusion model ها در تحلیل تصاویر پزشکی استفاده می شود؟ مزایای استفاده از آنها چیست؟
پاسخ:

مقایسه مدل‌های generator عمیق - یعنی مدل‌های Diffusion Models، Generative Adversarial Networks (GANs) و Variational Autoencoders (VAEs) - می‌تواند از چندین زاویه مختلف بررسی شود: معماری، روش‌های آموزش، وفاداری و تنوع نمونه‌ها، و کاربرد در دامنه‌های مختلف.

• معماری:

- GANs از دو شبکه عصبی تشکیل شده‌اند: generator و discriminator. generator داده‌ها را ایجاد می‌کند و discriminator آنها را ارزیابی می‌کند. این یک محیط آموزش پویا ایجاد می‌کند که در آن generator به طور مداوم یاد می‌گیرد تا داده‌های بهتر (واقع‌گرایانه‌تر) تولید کند.

- VAEs دارای ساختار encoder-decoder هستند. encoder داده‌ها را به فضای کم‌بعدی فشرده می‌کند و decoder داده‌ها را از این فضا بازسازی می‌کند. encoder و decoder با هم آموزش داده می‌شوند تا خطای بازسازی را به حداقل برسانند و latent space را منظم نگه دارند.

- مدل‌های Diffusion دارای یک فرآیند متمایز شامل دیفیوژن رو به جلو و دیفیوژن رو به عقب هستند. در فاز رو به جلو یا forward، به تدریج نویز به داده‌ها اضافه می‌شود تا تنها نویز باقی بماند. در فاز رو به عقب یا backward، این فرآیند به منظور تولید داده از نویز معکوس می‌شود.

• روش‌های آموزش:

- GANs از آموزش مخاصمه‌ای استفاده می‌کنند. generator و discriminator به طور همزمان آموزش داده می‌شوند، به طوری که generator سعی می‌کند داده‌هایی تولید کند که discriminator نتواند از داده‌های واقعی تمیز دهد و discriminator در تمیز دادن واقعی از جعلی بهتر می‌شود.

- VAEs آموزش داده می‌شوند تا Evidence lower bound را بر روی احتمال داده‌ها به حداکثر برسانند. این شامل بازسازی ورودی از latent space است در حالی که اطمینان حاصل می‌شود که latent space یک توزیع مشخص را دنبال می‌کند، معمولاً گاوسی است.

- مدل‌های Diffusion با دیفیوژن تدریجی داده‌ها آموزش داده می‌شوند. مدل یاد می‌گیرد تا فرآیند دیفیوژن رو به عقب را پیش‌بینی کند، به عبارت دیگر به طور موثر یاد می‌گیرد تا داده‌ها را با معکوس کردن فرآیند نویز تولید کند.

• صحت و تنوع نمونه‌ها:

- GANs برای تولید نمونه‌های به شدت طبیعی شناخته شده‌اند اما ممکن است از تنوع کم رنج ببرند، که منجر به **mode collapse** می‌شود. آن‌ها می‌توانند تصاویر بسیار واقع‌گرایانه‌ای تولید کنند اما ممکن است تمام تنوع داده‌ها را نشان ندهند.

- VAEs معمولاً نمونه‌های متنوع‌تری تولید می‌کنند اما گاهی اوقات ممکن است تصاویری کمتر واقع‌گرایانه یا با وضوح کمتر تولید کنند، که این مشکل از تاکید آن بر پوشش کامل latent space سرچشمه می‌گیرد.

- مدل‌های Diffusion با معکوس کردن فرایند دیفیوژن و حذف نویزها قادر به تولید نمونه‌های با کیفیت و متنوع هستند. البته از آنجایی که خروجی این مدل‌ها باید مرحله به مرحله بهبود پیدا کنند و از نویز به تصویر تبدیل شوند تولید تصویر با استفاده از این روش بسیار زمانبر خواهد بود.

• کاربرد در دامنه‌های مختلف:

- GANs اغلب در جایی که کیفیت نمونه‌های تولید شده نقش حیاتی دارند، مانند در تولید تصاویر هنری یا تصویرسازی واقع‌گرایانه تصویر استفاده می‌شوند. در واقع آن‌ها در زمینه‌هایی که جزئیات و بافت‌های واقعی مهم‌ترین فاکتور هستند، برتری دارند.

- VAEs در زمینه‌هایی که به **latent space** منظم نیاز است مفید هستند، مانند یادگیری semi-supervised و موارد مشابه.

- مدل‌های Diffusion پتانسیل زیادی در زمینه‌هایی نشان داده‌اند که کیفیت و تنوع هر دو حیاتی هستند، مانند در کشف دارو یا وظایف پیچیده تولید تصویر. آن‌ها همچنین در دامنه‌هایی که می‌توان فرآیند تولید داده را کنترل و به دقت تنظیم کرد، مورد بررسی قرار می‌گیرند، مانند تبدیل متن به تصویر یا در وظایف تولید شرطی.

در نتیجه، انتخاب بین GANs، VAEs و مدل‌های Diffusion به شدت به نیازمندی و چیزی که از مدل انتظار داریم وابسته است. GANs ممکن است برای وظایفی که نیاز به خروجی‌های با کیفیت بالا دارند انتخاب شوند، VAEs برای latent space ساختاریافته و مدل‌های دیفیوژن برای کیفیت و تنوع بالای آن‌ها، هرچند با هزینه محاسباتی بیشتر.

البته این صرفاً این یک نمونه پاسخ مورد قبول است و پاسخ لازم نیست دقیقاً به این صورت باشد. در نهایت ارزیابی پاسخ شما بر اساس درستی آن و میزان پاسخگویی آن به صورت سوال صورت می‌گیرد.

سوال دوم (۱۰ نمره)

۱. چرا برای نمونه‌برداری از توزیع با استفاده از تابع امتیاز، به فرآیند دیفیوژن نیاز داریم و چه اتفاقی می‌افتد اگر بدون اضافه کردن نویز نرمال و کانولوشن توزیع اصلی با توزیع نرمال، سعی در استفاده مستقیم از تابع امتیاز کنیم؟

۲. حال در ادامه سوال قبل، توزیع احتمال $p(x, t)$ که به صورت زیر تعریف می‌شود را در نظر بگیرید:

$$p(x, t) = \sum_{j=1}^M w_j \frac{1}{\sqrt{2\pi (s_j^2 + \sigma^2 t)}} \exp \left\{ -\frac{[x - \mu_j]^2}{2 (s_j^2 + \sigma^2 t)} \right\}.$$

تابع امتیاز مربوطه $s(x, t)$ را برای این توزیع بدست آورید و آن را برای یک زمان t دلخواه با $t = 0$ مقایسه کنید.

$$s(x, t) = -\frac{1}{p(x, t)} \sum_{j=1}^M w_j \frac{(x - \mu_j)}{s_j^2 + \sigma^2 t} \frac{1}{\sqrt{2\pi (s_j^2 + \sigma^2 t)}} \exp \left\{ -\frac{[x - \mu_j]^2}{2 (s_j^2 + \sigma^2 t)} \right\}$$

سوال سوم (۱۵ نمره)

با توجه به مطالب گفته شده در کلاس و همچنین مطالعه بخش تحلیل خطا از کتاب Machine Learning Yearning by Andrew Ng به سوالات زیر پاسخ دهید:

۱. با ذکر یک مثال از تحلیل تصاویر پزشکی، توضیح دهید چگونه تحلیل خطا می‌تواند از صرف وقت بیهوده برای بهبود کیفیت مدل بر روی دسته‌ی خاصی از دیتاها جلوگیری کند؟

پاسخ:

تحلیل خطا در تصاویر پزشکی یک عنصر بسیار مهم در بهبود عملکرد مدل‌های یادگیری عمیق است. برای توضیح، فرض کنید یک مدل شناخت تصاویر رادیولوژیک (مثل اشعه ایکس) را آموزش داده‌ایم ولی مشاهده کردیم که در تصاویر بیماران با سن بالا، دقت مدل به شدت کاهش می‌یابد. حالا فرض کنید تحلیل خطا نشان دهد که مدل در تصاویر بیماران سنین بالا، اغلب با تصمیم‌های اشتباهی روبرو می‌شود. مثلاً ممکن است مدل در تصاویر افراد سالم با سن بالا، نمایش‌های غیرعادی را به عنوان علائم بیماری تشخیص دهد یا برعکس، علائم بیماری را در تصاویر افراد جوان به نادرست تشخیص دهد. حالا فرض کنید تصاویر با سن بالا در داده‌های آموزشی به میزان کافی نبوده و یا تنوع کافی نداشته باشند، در نتیجه مدل در مواجهه با این نوع تصاویر از عملکرد ضعیفی برخوردار باشد. اگر تحلیل خطا نشان دهد که مشکل اصلی در دسته خاصی از داده‌ها (مثل سن بالا) است، می‌توانیم به جای اینکه به طور کلی مدل را دوباره آموزش دهیم، تمرکز خاصی بر روی افزایش داده‌ها یا تعادل توزیع داده‌ها در این دسته از تصاویر بگذاریم. به این ترتیب، با تحلیل خطا، می‌توانیم وقت و انرژی خود را در بهبود کیفیت مدل در دسته‌های خاص مصرف کنیم و از صرف وقت بیهوده بر روی دسته‌های دیگری که بهبود نیاز ندارند، جلوگیری کنیم.

۲. دو زیرمجموعه Blackbox و Eyeball چه هستند و تقسیم‌بندی دیتای ولیدیشن به این دو چه فایده‌ای دارد؟

پاسخ:

اگر دیتاست ولیدیشن به اندازه کافی بزرگ باشد می‌توان آن را به دو بخش Blackbox و Eyeball تقسیم کرد. بخش Eyeball قسمتی از دیتاست است که بر روی آن تحلیل خطا را انجام می‌دهیم و سپس با توجه به این تحلیل خطا تغییرات لازم را بر روی مدل ایجاد می‌کنیم اما چون نمی‌خواهیم به این خطاها overfit شویم بهتر است دیتاست دیگری داشته باشیم که روی آن تحلیل خطا انجام ندهیم و فقط برای ارزیابی از آن استفاده کنیم. این زیرمجموعه از دیتاست را Blackbox می‌نامیم.

۳. مفهوم error analysis by parts را با ذکر یک مثال در حوزه تصاویر پزشکی توضیح دهید که چگونه این روش می‌تواند به اولویت‌دهی برای بهبود کیفیت بخش‌های مختلف سیستم کمک کند.

پاسخ:

تجزیه و تحلیل خطا بر اساس بخش‌ها (Error Analysis by Parts) یک روش است که در ارزیابی یک سیستم یا مدل، خطاها و نقص‌ها را بر اساس بخش‌های مختلف یا عناصر تشکیل‌دهنده آن سیستم تجزیه و تحلیل می‌کند. این روش به ما امکان می‌دهد تا به صورت دقیق‌تر متوجه شویم که در کدام بخش‌ها یا قسمت‌های سیستم خطاها رخ می‌دهند و این اطلاعات می‌تواند به ما کمک کند تا اولویت‌بندی و اصلاح مشکلات را بر اساس اهمیت و تأثیر آن‌ها بر کل سیستم انجام دهیم.

یک مثال در حوزه تصاویر پزشکی می‌تواند مرتبط با یک سیستم تشخیص تصاویر باشد. فرض کنید یک مدل با دقت نسبتاً مناسب ایجاد شده است، اما می‌خواهیم بفهمیم کدام بخش‌ها از تصویر ممکن است باعث خطا شوند و چگونه می‌توانیم بهبودهای لازم را اعمال کنیم.

پیش‌پردازش تصویر: ممکن است مشخص شود که خطاها از مراحل پیش‌پردازش تصاویر ناشی می‌شوند، مثلاً نویزهای موجود در تصاویر یا مشکلات در استخراج ویژگی‌ها.

معماری مدل: اگر تحلیل نشان دهد که خطاها بیشتر در معماری مدل ایجاد می‌شوند، ممکن است نیاز باشد که ساختار شبکه عصبی یا پارامترهای مدل بهبود یابد.

آموزش و داده‌ها: اگر خطاها ناشی از داده‌های آموزش یا روش‌های آموزش باشند، لازم است دقت کنیم که آیا داده‌های موجود متناسب با واقعیت‌های مختلف هستند یا نه.

پس‌پردازش و تفسیر نتایج: در صورتی که خطاها بیشتر در مرحله تفسیر و پس‌پردازش نتایج مدل رخ می‌دهند، نیاز است که الگوریتم‌ها یا روش‌های تفسیر بهبود یابند.

با انجام تجزیه و تحلیل به این شکل، می‌توان بهبودهای هدفمندتر و کارآمدتری را در سیستم اعمال کرد و از منابع بهینه‌تری برخوردار شد.

سوال چهارم (۱۰ نمره)

در این تمرین، شما با چالش‌ها، تبدیل و تقریب توابع هدف در مدل‌های مبتنی بر امتیاز در زمینه فرایندهای تصادفی سروکار دارید. مدل‌های مبتنی بر امتیاز اغلب در یادگیری ماشین و تحلیل داده‌های تصادفی کاربرد دارند. هدف شما این است که درک کنید چرا تابع هدف تغییر یافته نسبت به نسخه اولیه ترجیح داده می‌شود.

میدانیم در روش‌های Score-based تابع هدف اولیه معمولاً به صورت زیر است:

$$J_{\text{naive}}(\theta) := \frac{1}{2} \int dx dt p(\mathbf{x}, t) [\mathbf{s}_{\theta}(\mathbf{x}, t) - \nabla_{\mathbf{x}} \log p(\mathbf{x}, t)]^2$$

اما این تابع مشکل عمده‌ای دارد: تخمین گرادیان لگاریتم $p(\mathbf{x}, t)$ دشوار است زیرا $p(\mathbf{x}, t)$ ممکن است به شدت به $p(\mathbf{x}, 0)$ توزیع هدف ناشناخته، وابسته باشد. برای حل این مشکل، به سراغ تابع هدف تغییر یافته، $J_{\text{mod}}(\theta)$ ، می‌رویم که همان global minimum را دارد اما کاربردی‌تر است.

$$J_{\text{mod}}(\theta) := \frac{1}{2} \int dx d\mathbf{x}^{(0)} dt p(\mathbf{x}, t | \mathbf{x}^{(0)}, 0) p(\mathbf{x}^{(0)}) [\mathbf{s}_{\theta}(\mathbf{x}, t) - \nabla_{\mathbf{x}} \log p(\mathbf{x}, t | \mathbf{x}^{(0)}, 0)]^2$$

اثبات کنید که $J_{\text{naive}}(\theta)$ و $J_{\text{mod}}(\theta)$ دارای global minimum های یکسان هستند. **راهنمایی:** در روند اثبات لازم است از Log-derivative trick استفاده کنید. چنانچه با آن آشنایی ندارید پیش از حل سوال این [لینک](#) را مطالعه کنید.

پاسخ:

اثبات تساوی با نشان دادن اینکه گرادیان‌های هر دو تابع $J_{\text{naive}}(\theta)$ و $J_{\text{mod}}(\theta)$ یکسان هستند صورت می‌گیرد:

$$\begin{aligned}
 \nabla_{\theta} J_{\text{naive}}(\theta) &= \int d\mathbf{x} dt p(\mathbf{x}, t) [\mathbf{s}_{\theta}(\mathbf{x}, t) - \nabla_{\mathbf{x}} \log p(\mathbf{x}, t)] \cdot \nabla_{\theta} \mathbf{s}_{\theta}(\mathbf{x}, t) \\
 &= \int d\mathbf{x} dt p(\mathbf{x}, t) \left[\mathbf{s}_{\theta}(\mathbf{x}, t) - \frac{\nabla_{\mathbf{x}} p(\mathbf{x}, t)}{p(\mathbf{x}, t)} \right] \cdot \nabla_{\theta} \mathbf{s}_{\theta}(\mathbf{x}, t) \\
 &= \int d\mathbf{x} dt [p(\mathbf{x}, t) \mathbf{s}_{\theta}(\mathbf{x}, t) - \nabla_{\mathbf{x}} p(\mathbf{x}, t)] \cdot \nabla_{\theta} \mathbf{s}_{\theta}(\mathbf{x}, t) \\
 &= \int d\mathbf{x} d\mathbf{x}^{(0)} dt p(\mathbf{x}^{(0)}) [p(\mathbf{x}, t | \mathbf{x}^{(0)}, 0) \mathbf{s}_{\theta}(\mathbf{x}, t) \\
 &\quad - \nabla_{\mathbf{x}} p(\mathbf{x}, t | \mathbf{x}^{(0)}, 0)] \cdot \nabla_{\theta} \mathbf{s}_{\theta}(\mathbf{x}, t) \\
 &= \int d\mathbf{x} d\mathbf{x}^{(0)} dt p(\mathbf{x}, t | \mathbf{x}^{(0)}, 0) p(\mathbf{x}^{(0)}) [\mathbf{s}_{\theta}(\mathbf{x}, t) \\
 &\quad - \nabla_{\mathbf{x}} \log p(\mathbf{x}, t | \mathbf{x}^{(0)}, 0)] \cdot \nabla_{\theta} \mathbf{s}_{\theta}(\mathbf{x}, t)
 \end{aligned}$$

به طور مشابه داریم:

$$\begin{aligned}
 \nabla_{\theta} J_{\text{mod}}(\theta) &= \int d\mathbf{x} d\mathbf{x}^{(0)} dt p(\mathbf{x}, t | \mathbf{x}^{(0)}, 0) p(\mathbf{x}^{(0)}) [\mathbf{s}_{\theta}(\mathbf{x}, t) \\
 &\quad - \nabla_{\mathbf{x}} \log p(\mathbf{x}, t | \mathbf{x}^{(0)}, 0)] \cdot \nabla_{\theta} \mathbf{s}_{\theta}(\mathbf{x}, t)
 \end{aligned}$$