

چگونه می توان از یادگیری انتقالی کوانتمی استفاده کرد و چالش های آن چیست؟

بخشی از پروژه پایانی محاسبات کوانتمی ترم 4031

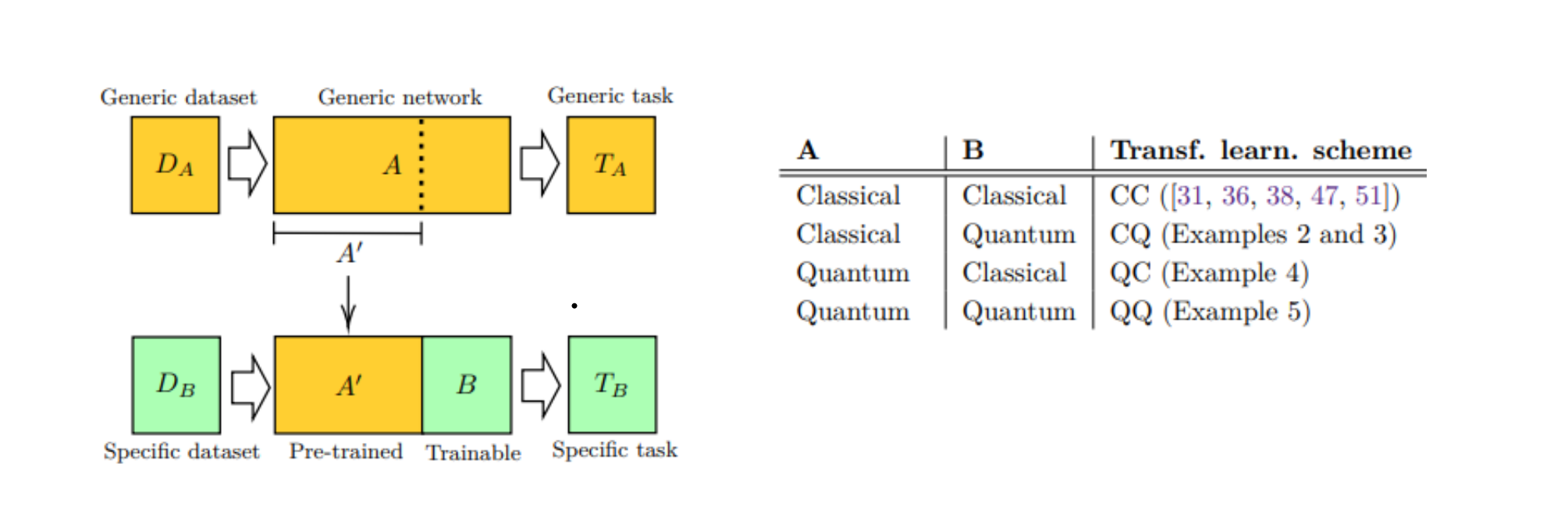
نام: کسری فولادی (گروه تک نفره)، شماره دانشجویی: 402442297

در این خلاصه نویسی از مقاله quantum transfer learning به این سوال پاسخ می دهیم که یادگیری انتقالی چیست ایده اصلی در یادگیری انتقالی کوانتمی چیست و یک مورد استفاده یادگیری انتقالی کوانتمی را بررسی می کنیم، در آخر به چالش های این روش می پردازیم.

مفهوم یادگیری انتقالی را که به طور گسترده در الگوریتم‌های یادگیری ماشینی مدرن به کار می‌رود، و در این مقاله نویسندگان آن را به زمینه در حال ظهور شبکه‌های عصبی هیبرید متشکل از عناصر کلاسیک و کوانتومی گسترش می دهند. پیاده سازی های متفاوتی از یادگیری انتقال ترکیبی را پیشنهاد در این مقاله نشان داده شده اند، که عمدتاً به این شکل است که در آن یک شبکه کلاسیک از پیش آموزش دیده شده توسط یک مدار کوانتومی تغییری نهایی اصلاح و تقویت می شود.

این رویکرد به ویژه در عصر کنونی فناوری کوانتوم در مقیاس متوسط جذاب است زیرا امکان پیش‌پرداز ی بهینه داده‌های با ابعاد بالا (مثلاً تصاویر) با هر شبکه کلاسیک پیش‌رفته و پیش‌رفته را فراهم می کند. در مقاله چندین مثال اثبات مفهوم از کاربرد آسان یادگیری انتقال کوانتمی برای تشخیص تصویر و طبقه بندی حالت کوانتمی ارائه می شود. که در این خلاصه نویسی حالتی که اطلاعات کلاسیک به وسیله لایه کوانتمی پردازش می شود را بررسی می کنیم.

یادگیری انتقالی یک نمونه معمولی از تکنیک هوش مصنوعی است که در ابتدا توسط هوش بیولوژیکی الهام گرفته شده است. این از مشاهده ساده ناشی می شود که دانش به دست آمده در یک زمینه خاص می تواند به یک منطقه متفاوت منتقل شود. به عنوان مثال، وقتی زبان دومی را یاد می گیریم، از صفر شروع نمی کنیم، بلکه از دانش زبانی قبلی خود استفاده می کنیم. گاهی اوقات یادگیری انتقالی تنها راه برای نزدیک شدن به وظایف شناختی پیچیده است، به عنوان مثال، قبل از یادگیری مکانیک کوانتومی، توصیه می شود ابتدا جبر خطی را مطالعه کنید. این ایده کلی با موفقیت برای طراحی شبکه های عصبی مصنوعی نیز به کار گرفته شده است و ثابت شده که در بسیاری از موقعیت ها، به جای آموزش یک شبکه کامل از ابتدا، کارآمدتر است که از یک شبکه عمیق از پیش آموزش دیده شروع شود و سپس تنها برخی از لایه های نهایی برای یک کار خاص بهینه سازی شوند. و بقیه لایه ها هیچ تغییری نمی کنند. تصویر زیر یک مثال است و هر کدام از بخش ها می توانند کلاسیک یا کوانتومی باشند. که در اینجا ما حالتی که B کوانتمی باشد و A کلاسیک باشند را بررسی می کنیم.(CQ)



قبل از اینکه مثالی از CQ را بررسی بکنیم بیایم ببینیم شبکه های هیبریدی چگونه کار می کنند.

بخش کلاسیک:

یک مدل بسیار موفق در یادگیری ماشین کلاسیک، مدل شبکه‌های عصبی deep fed-forward است. که شبکه های عصبی بر این مینا ساخته می شوند بلوک i یک شبکه عمیق یک لایه نامیده می شود و بردار های ورودی شامل n_i-1 عدد حقیقی را به بردار های خروجی شامل n_i عدد حقیقی نگاشت می کند. ساختار معمولی آن شامل یک عملیات آفین است که بر روی بردار ورودی اعمال می شود و بر تمام عضو های حاصل آن یک تابع فعال ساز غیر خطی اعمال می شود

$$L_{n_{i-1} \rightarrow n_i} \colon x \rightarrow y = \phi(Wx + b)$$

پس هر بلوک کلاسیک شامل d لایه را می توانیم ضرب تعدادی عملگر L به صورت زیر نوشت ه از چپ بر راست اعمال می شوند.

$$C = L_{n_{d-1} \rightarrow n_d} * \cdots * L_{n_0 \rightarrow n_1}$$

بخش کوانتمی:

برای هر بلوک کوانتمی ما سه بخش اصلی خواهیم داشت:

- لایه تناظر: یک بخش که اطلاعاتی که به دست ما رسیده را درون کیوبیت ها انکد کند.
- لایه (های) محاسبه: یک بخش که محاسبات کوانتمی لازم را انجام بدهد. (تقریبا چیزی معادل عملگر هایی که در بلوک های کلاسیک داریم)
- لایه استخراج: یک بخش که اطلاعات بدست آمده را برای ما از دل کیوبیت ها بیرون بکشد و حالت کوانتمی را به یک بردار کلاسیک نگاشت کنیم.

هر بلوک کوانتومی را می توان با یک مدار کوانتمی متغیر پیاده سازی کرد. در ابتدا لازم است که با یک تابع بردار از عدد های حقیقی را به یک حالت کوانتمی نگاشت کنیم.

$$\mathcal{E} \colon x \rightarrow |x\rangle = E(x)|0\rangle$$

برای مثال می توان به ازای هر عضو بردار یک کیوبیت خاص را به مقدار معینی دوران داد.

برای پیاده سازی هر لایه محاسباتی کوانتمی هم می توان یک عملگر یونیتاری تعریف کرد

$$\mathcal{L}_i \colon |x\rangle \rightarrow |y\rangle = U(w)|x\rangle$$

که در آن دبلو یک آرایه از اعداد حقیقی است (بعد از هر لایه کوانتومی مقدار اولیه ی کیوبیت های به دلیل خواص ماتریس های یونیتاری قابل باز یابیست) بنا بر این می توانیم لایه های محاسباتی را هم به این شکل نمایش دهیم:

$$Q = \mathcal{L}_q * \cdots * \mathcal{L}_1$$

در لایه استخراج هم برای اینکه از حالت نهایی هم خبر دار شویم همه کیوبیت ها را اندازه گیری می کنیم این را هم می توانیم به شکل یک عملگر تعریف کنیم که حالت کلی سیستم ما را به تعدادی حالت نهایی تناظر می دهد:

$$\mathcal{M} \colon |x\rangle \rightarrow y = \langle x|Y|x\rangle, Y = [y_1, \ldots, y_{n_q}]$$

پس شکل کلی یک بلوک کوانتمی به مطابق زیر است:

$$Q = \mathcal{M} * \mathcal{Q} * \mathcal{E}$$

مثال از استفاده یادگیری انتقالی کوانتمی(کلاسیفیکیشن مورچه و زنبور):

ابتدا یک شبکه کلاسیک مولد را آموزش می دهیم سپس برخی از لایه های آن را عوض می کنیم و همه لایه ها بجز چند تارو فریز می کنیم در این مثال نویسندگان از یک کلاسیفیکیشن بر روی تصاویر انجام می دهند که برای اولینبار با شبکه هیبریدی انجام می شود و روش شبیه ساز و دو کامپیوتر کوانتمی متفاوت ران شده و نتایج را بررسی کرده است.

<i>D</i> _A = ImageNet: a public image dataset with 1000 classes [13].	
<i>A</i> = ResNet18: a pre-trained residual neural network introduced by Microsoft in 2016 [20].	
<i>T</i> _A = Classification (1000 labels).	
<i>A</i> ' = ResNet18 without the final linear layer, obtaining a pre-trained extractor of 512 features.	<i>B</i> = <i>Q̃</i> = <i>L</i> _{4→2} ◦ <i>Q</i> ◦ <i>L</i> _{512→4} : i.e., a 4-qubit dressed quantum circuit (9) with 512 input features and 2 real outputs.
<i>D</i> _{<i>B</i>} = Images of two classes: <i>ants</i> and <i>bees</i> (<i>Hy-menoptera</i> subset of ImageNet), separated into a training set of 245 images and a testing set of 153 images.	<i>T</i> _{<i>B</i>} = Classification (2 labels).

بلوک های کلاسیک همان معماری رایج شبکه های عصبی عمیق را دارند با تابع فعال ساز تانژانت هایپر بویلیک و در بهینه سازی تمام بلوک ها از بهینه ساز آدام استفاده شده، و باقی جزئیات در باره نحوه پیاده سازی بلوک کوانتمی می باشد.

بلوک کوانتمی رو اگر بررسی بکنیم لایه انکد کردنش به این صورت پیاده سازی می شود که:

$$\mathcal{E}(x) = \left(\otimes_{i=1}^4 R_y\left(x_i\frac{\pi}{2}\right)H\right)|0000\rangle$$

و برای لایه های محاسباتی هم 6 لایه داریم که هر لایه به شکل زیر پیاده می شود:

$$Q = \mathcal{L}_6 * \cdots * \mathcal{L}_1$$



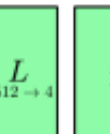
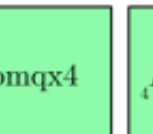
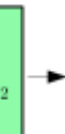


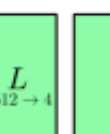
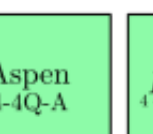

$$\mathcal{L}(w) \colon |x\rangle \rightarrow |y\rangle = K\left(\bigotimes_{i=1}^4 R_y(w_i)\right)|x\rangle, \quad K = (CNOT \otimes \mathbb{I}_{3,4})(\mathbb{I}_{2,1} \otimes CNOT)(\mathbb{I}_1 \otimes CNOT \otimes \mathbb{I}_4)$$

و برای لایه استخراج هم تعریف می کنیم:

$$Y = \begin{bmatrix} Z \otimes \mathbb{I} \otimes \mathbb{I} \otimes \mathbb{I} \\ \mathbb{I} \otimes Z \otimes \mathbb{I} \otimes \mathbb{I} \\ \mathbb{I} \otimes \mathbb{I} \otimes Z \otimes \mathbb{I} \\ \mathbb{I} \otimes \mathbb{I} \otimes \mathbb{I} \otimes Z \end{bmatrix}, (Z: \text{tha pauli gate } Z)$$

برای بهینه کردن پارامتر ها هم از سافت مکس لاگ استفاده شده که با کامپیوتر کلاسیک پارامتر های کلاسیک لایه های را به روز می کنیم و نکته جدید و خاصی ندارد.

در زیر می توانید دقت بدست آمده مدل در سه محیطی که در آن آزمایش شده و یک شماتیک از ساختار و عملکرد این شبکه ببینید. همچین این شبکه بر روی کامپیوتر های خانگی هم از نظر سرعت عملکرد قابل قبولی دارد.

					→ “bee”
					→ “ant”

QPU	Accuracy
Simulator	0.967
ibmqx4	0.95
Aspen-4-4Q-A	0.80

چالش ها:

از چالش های این روش میتوان به این اشاره کرد که فناوری امروزه جهان به سختی می تواند صرفا تسک های کلاسیفیکیشن را در مدار کوانتمی به خوبی اجرا کند و اجرا کردن در شبیه ساز ها هم معمولاً زمان و هزینه زیادی می برد طوری که نسخه های کلاسیک منطقی تر هستند.

برای مثال من در پیاده سازی یک شبکه مولد تخصصی که مولد آن کوانتمی باشد چون دیتاست ساده ای پیدا نکردم مجبور شدم با تعداد کیوبیت بالا محاسبه را انجام بدهم که هزینه زمانی فوق العاده بالایی داشت و کم کردن کیوبیت ها هم کیفیت کار را پایین می آورد بنا بر این مجبور شدم کدم که شامل 22 کیوبیت بود رو روی ورک اسپیس دوستم روی سوپر کامپیوتر دانشگاه آریزونا ران کنم و با توجه به اینکه 22 کیوبیت تقریبا در حدود تعداد پارامتر های نسخه کلاسیک به من داد (2 به توان 22 تقریبا برابر تعداد پارامتر های نسخه کلاسیکی هست که نوشتم) و درسته که کیفیت اپوک کوانتمی 20 از کیفیت اپوک 80 کلاسیک بیشتر بود اما باز هم کلاسیک به صر فه بود.

و چالش دیگر هندل کردن نویز های کامپیوتر های کوانتمی و ارزان کردن فناوری ان است .