پاسخ به نظرات داور گرامی

کد مقاله: MSCI-2009-1387(R2)

عنوان مقاله: کاربرد بسط تیلور در کاهش حجم مدل های یادگیری عمیق

با سلام و عرض ادب

مطابق دفعات قبل، پاسخ اینجانب با رنگ آبی مشخص شده است. تصویر برخی موارد اصلاح شده در لابلای پاسخ‌ها درج شده است.

نظرات دور سوم داور محترم:

برخی از نکات ذکر شده توسط داور انجام شده و برخی انجام نشده اند، برخی از ابتدای فهرست، برخی از میانه و برخی از انتها. منظور از بند 13 که نویسنده محترم سوال نمودند، جعبه شماره 13 در فرم ارزیابی است که نظرات به صورت متنی در آن درج شده اند.

شماره های مورد اشاره نیز مربوط به آخرین ارزیابی است که تا 31 شماره گذاری شده اند.

نویسنده محترم لطفا به موارد (17) و (1) مجددا توجه نماین. همچنین اصلاح انجام شده روی همان مورد اصلی مورد تاکید نیز کافی نیست و هنوز همان تاکید روی شماره (13) باقی است.

لذا (ویرایش شده ی) همان مکاتبات پیشین به شرح زیر ارسال می گردد.

مطلوب است اصلاحات مواردی که انجام شده اند برای داور اعلام گردند.

مورد (13) در زیر هنوز نیاز به رسیدگی دارد.

در نسخه‌ی قبل داوری، من فهرست نظرات داور گرامی را نداشتم. «جعبه شماره 13» که داور محترم اشاره فرمودند، اصلا برای من قابل مشاهده نبود. هم از سامانه‌ی نشریه و هم از طریق ایمیل اطلاع دادم که ظاهرا لیستی از نظرات بوده و من دریافت نکرده‌ام که چون جوابی دریافت نکردم و اصلاحات جزیی خواسته شده بود، فرض را بر نظرات قبلی گذاشته و پاسخ داده بودم.

1ـ با توجه به اصلاح عنوان و سوال نویسنده محترم از داور، نظر داور همچنان بر اضافه نمودن مثال مورد کاربرد آن در عنوان مقاله نیز هست به گونه ای که این عنوان مقاله به دو کاربرد ریاضیات اشاره خواهد داشت، هم در علوم کامپیوتر و هم در هنر و شاید تشخیص اصالت آثار، و از آنجا که یادگیری عمیق، شنیده شده تر از شبکه عصبی پیچشی و بیشتر مورد اقبال برای عموم است و با توجه به پاسخ (63ـ پیشین)، می تواند به صورت زیر اصلاح یابد:

(الف) کاربرد بسط تیلور در کاهش حجم یک مدل یادگیری عمیق برای...

(ب) و البته همان شبکه عصبی پیچشی نیز کاملا بلامانع و هنوز مناسب است و کماکان به کاربرد در علوم کامپیوتر اشاره دارد، و با افزودن بخش کاربرد به انتهای عنوان.

انتخاب بین این دو به نویسنده محترم واگذار می شود که شاید گرایش به عنوان دوم باشد.

انجام شد. البته چون عنوان طولانی جدید، سربرگ صفحات را خراب می‌کرد، مجبور شدم عنوان کوتاه‌تر قبلی را برای سربرگ‌ها بکار ببرم.

2ـ موردی نگارشی از نظر داور و نویسنده محترم دور مانده بود. بین ذکر مرجع و نویسه قبل از آن (چه فارسی، چه انگلیسی، چه شماره پانویس، ...) یک فاصله لازم است. لطفا در صفحه اول و دوم، در بخش اول اصلاح شوند. در صفحه اول، بعد از مرجع 7 یک فاصله درج گردد

انجام شد. گرچه برخی موارد را به عمد چسبانیده بودم. مثل مرجع ۱۵ در صفحه‌ی دوم که وقتی فاصله قرار می‌دهم، شماره مرجع در ابتدای خط بعدی قرار می‌گیرد که جالب به نظر نمی‌رسد. برخی از دیگر موارد هم به این دلیل بوده که پاراگراف مربوطه نصف نشود (یکی دو خطش به صفحه بعدی نرود). برای شکیل ماندن کلیات ظاهر مقاله گاهی برخی جزییات را قربانی می‌کنم. با اینحال، نکات گفته شده انجام شدند.

3ـ صفحه اول، بخش اول، خط ـ5، «برده شده‌اند؛ که ...» -> «برده شده‌اند، که ...»

اصلاح شد.

4ـ ص2، خط 3، « ... ارائه شده است.» -> در ادامه، بخشی از پاسخ (45ـ پیشین) می تواند به صورت یک جمله اضافه شود که «هرس و کاهش حجم مدل می تواند یک مدل اولیه مثلا 500 مگابایت را به مثلا 50 مگابایت برای استفاده در یک برنامه اندرویدی تبدیل کند». لذا همین جا کم کم دلیل های کاربردی نیاز به کاهش مدل هم ملموس تر می شود.

انجام شد. جمله‌ی « هرس و کاهش اندازه‌ی مدل می‌تواند حجم یک مدل اولیه‌ی مثلاً 500 مگابایتی را به 50 مگابایت کاهش دهد؛که برای استفاده در یک برنامه‌ی اندرویدی مناسب‌تر است.» اضافه شد.

5- یادآوری (11ـ پیشین): ص2، پارارگراف 1، خط ـ6 و ـ5: «...در ادامه‌ی مقاله ذکر خواهد شد» -> «در بخش 2.1 آمده است.».

انجام شد. البته بخش ۳ هست.

6ـ ص3، بخش 2، پاراگراف اول، حاوی برخی تکرار و عدم تقدم مطالب شده است. داور، این پاراگراف را به صورت زیر پیشنهاد می کند (پانویسها در زیر درج نشده‌اند. مناسب است پانویس برای شبکه‌های عصبی مصنوعی و هوش مصنوعی نیز اضافه شود). نویسنده محترم می‌توانند این پاراگراف را به هر صورت بهتری اصلاح نمایند.

مدل‌های شبکه‌های عصبی مصنوعی، سال‌ها به امید ایجاد عملکردی مشابه مغز انسان در موضوعاتی نظیر تشخیص صحبت و تصویر در شاخه‌های هوش مصنوعی مورد مطالعه قرار گرفته‌اند. ساختار این مدل‌ها که الهام‌گرفته از شبکه‌های عصبی زیستی است، متشکل از تعدادی عنصر محاسباتی (نورون‌ها، گره‌ها، نودها) است که از طریق وزن‌هایی تطبیقی به هم متصل شده و به موازات یکدیگر کار می‌کنند. ساده‌ترین ساختار گره معمولا به صورت غیرخطی است که در آن، هر کدام از ورودی‌ها در وزن خاصی ضرب شده و حاصل‌جمع آنها با تابعی غیرخطی، خروجی را به دست می‌آورد (شکل 2). هر شبکه عصبی، علاوه بر معماری (توپولوژی) شبکه و مشخصات گره یا نود (یعنی نوع تابع آن، موسوم به تابع فعالیت)، با قاعده‌ی یادگیری یا آموزش نیز مشخص می شود. عمل یادگیری در شبکه‌های عصبی به معنی تنظیم درست وزن‌ها است به گونه‌ای که با ورود داده‌های مشخص، پاسخ مورد انتظار را تولید کند. قواعد یادگیری متشکل از مجموعه‌ای از وزن‌ها است و مشخص می‌کند که در هنگام آموزش شبکه، وزن‌ها چطور باید بهنگام شوند. آموزش شبکه، برای بالابردن دقت و درستی تشخیص است. روش‌های طراحی شبکه و قواعد آموزش آن، موضوع بسیاری از تحقیقات گذشته و اکنون است. با دسترسی به داده‌های آموزشی حجیم و لذا افزایش تعداد نودها و اتصالات، تنظیم وزن اتصالات بسیار زمان‌بر و نیازمند توان پردازشی زیاد همچون کارت‌های گرافیکی در کنار ابداع روش‌های بهتر بهینه‌سازی و انواع جدیدتر توابع فعالیت، و باعث اوج‌گیری مجدد شبکه‌های عصبی در قالب یادگیری عمیق شده است.

با ویرایشاتی انجام شد.

7ـ ص3، بخش 2، پاراگراف 2، خط ـ2 و ـ1: «در ادامه، مدل پرسپترون و شبکه‌های عصبی پیچشی به صورت مختصر بیان می‌شوند.» و یادآوری (25ـ پیشین: «...آموزش می‌بیند که ... قادر به ... است و ...») و بخشهایی از پاسخ (60ـ پیشین) و (63ـ پیشین)، این پاراگراف به صورت زیر پیشنهاد می‌شود:

یکی از ابزارهای حوزه یادگیری عمیق، شبکه عصبی پیچشی است که با دقت زیاد قادر به تفکیک نمونه‌ها در قالب دسته‌های از پیش مشخص است. مثال مشهوری که امروزه در حوزه یادگیری عمیق استفاده می‌شود، دسته‌بندی تصاویر سگ و گربه است. با استفاده از صدها تصویر متفاوت سگ و گربه، یک شبکه عصبی پیچشی آموزش می‌بیند تا مشخص کند که یک تصویر جدید (سگ یا گربه) متعلق به کدام دسته است. روال یادگیری در بسیاری از انواع شبکه‌ها منجمله شبکه‌های عصبی پیچشی، بر پایه‌ی مدل پرسپترون است. در ادامه، مدل پرسپترون و شبکه‌های عصبی پیچشی به صورت مختصر بیان می‌شوند.

انجام شد.

8ـ ص4، پاراگراف 2، به نوعی تکرار پاراگراف قبل از آن است.

ویرایش شد.

9ـ ص4، پاراگراف 3، خط ـ2، «مشخص کننده‌ی» -> «مشخص‌کننده‌ی»

سه مورد اصلاح شد.

10ـ ص4، پاراگراف 4، شروع با: «به عنوان مثال، ».

پاراگراف ویرایش شد و با ویرایش جدید به گمانم نیازی به این عبارت نیست.

11ـ ص4، پاراگراف 4، خط 2: «مشخصه‌ی» -> «مشخص‌کننده‌ی»

اصلاح شد.

12ـ ص4، پاراگراف 4، خط ـ3، «...خط فوق عمل ...» -> «...خط فوق، عمل ...»

ویرایش شد.

13ـ ص4، پارگراف 4، خط 1 و 2، شاید مناسب‌تر به این صورت نوشته شود؟ -> «به عنوان مثال خط $w_1x_1+w_2x_2-a=b$با $w_1=w_2=1$، $a=0.5$ و $b=0$ و با فرض اینکه True و False به ترتیب با یک و صفر نشان داده شوند، مشخص‌کننده‌ی مرز تصمیم‌گیری ...». افزودن این پاراگراف مفید بوده است، ولی شفاف بیان نشده است و هنوز هم مشکل (27ـ پیشین) را حل نکرده است که مثلا خب چرا نمایانگر این عبارت منطقی است و نمایانگر عبارت منطقی دیگری نیست و ربط مقادیر a و b به این عبارت منطقی و شکل و جای خط جداکننده و نتیجه آن چگونه است که بتوان با تعمیم به آن تابع چندمتغیره هم رسید. این شفاف نبودن این پاراگراف، توسط دانش نویسنده محترم و کتب یادگیری ماشین، شبکه های عصبی، ... قابل رفع است و افزودن یا تغییربرخی جملات یا شکلها و کمی بازنویسی به واضح شدن این پاراگراف و پاراگراف بعدی کمک می کند. این مختصر مطالب زیر از اولین لینکهای جستجو در اینترنت نیز شاید کمک کننده باشد:...

ویرایش‌های زیادی انجام شد. دو شکل ۳ و ۴ اضافه شدند. چون شکل ۲ باید قبل از اینها می‌آمد، از صفحه‌ی بعد به این صفحه منتقل شد. نمایش خط جداساز برای ترکیب فصلی، زیباتر از ترکیب فصلی بود، لذا ترکیب فصلی با ترکیب عطفی جایگزین شد. هدف مقاله، توضیح کامل شبکه‌های عصبی نبوده است؛ به نظر می‌رسد الان توضیحات در حد مقاله کافی هستند و خواننده اگر نیازمند اطلاعات بیشتر است می‌تواند به کتابها و منابع دیگر مراجعه کند.

14ـ ص7، پاراگراف اول، خط1، جمله اول حذف شود. جمله دوم -> «از معروف‌ترین توپولوژی‌های ابداع شده از شبکه‌های پیچشی، مدل VGG ... ».

ویرایش شد.

15ـ ص7، پاراگراف 2، خط ـ4، «... در بخش پیش توضیح داده شد»: شماره بخش ذکر گردد.

انجام شد.

16ـ ص7، بخش 3، خط 1، «همانگونه که ملاحظه شد،...» -> «همانگونه که در جدول 1 ملاحظه شد، ...».

انجام شد.

17ـ ص 7، خط ـ2، منظور داور در (41ـ پیشین) این خط و در مورد واژه‌ی متبادر نبوده است و مناسب است این جمله به صورت قبلی بازنویسی شود. منظور، نماد برای مفهوم «تعداد عناصر غیرصفر مدل» در ص9(فعلی)، خط 3 است. توصیف نماد به صورت «تعداد وزن‌های صفر در مدل» ؟ گه لذا هرچه مقدار آن کمتر باشد، فشردگی بیشتر است؟ یا نماد نیاز به اصلاح دارد؟

جمله‌ی قبلی جایگزین شد.

برای بخش دوم، پاراگراف مربوطه به صورت زیر ویرایش شد، امیدوارم که واضح‌تر شده باشد. (ابتدای ص ۱۰):

۲.۳. استفاده از بسط تیلور در انتخاب وزن‌ها برای کاهش حجم مدل. فرض کنید $\mathbf{W}$ بردار وزن‌های شبکه و $E(\mathbf{W})$ تابع هزینه‌ی مسئله‌ی طبقه‌بندی باشد. کاهش حجم مدل، به معنی کاستن تعداد درایه‌های $\mathbf{W}$ است. برای بیان ریاضی مسئله، به جای حذف درایه‌ها، مؤلفه‌های مورد هرس از $\mathbf{W}$ را صفر کرده و بردار جدید را $\mathbf{W}'$ می‌نامیم. صفر شدن وزن یک اتصال مترادف با حذف (هرس) ارتباط بین دو نود مربوطه است. اگر به فرض $\mathbf{W}$ شامل هیچ درایه‌ی برابر با صفری نباشد، نرم صفر $\mathbf{W}'$ ($\|\mathbf{W}'\|_0$)، تعداد مؤلفه‌های غیرصفر مدل، بعد از هرس را نشان می‌دهد. به این ترتیب میزان تغییر خطای شبکه، ناشی از هرس کردن تعدادی از اتصالات آن را می‌توان به صورت $|E(\mathbf{W}) - E(\mathbf{W}')|$ نشان داد. فرض کنید میزان کاهش حجم نهایی مشخص است، یعنی به عنوان مثال برای B مفروض، مایل هستیم $\|\mathbf{W}'\|_0 \leq B$. کمتر بودن B معادل کمتر بودن مؤلفه‌های غیرصفر بردار وزن و به عبارت دیگر کمتر بودن اتصالات فعال مدل است. اگر اتصالات با وزن برابر با صفر — به عنوان اتصالات غیرفعال — از بردار وزن حذف شوند، مدلی کوچک‌تر خواهیم داشت. حجم مدل جدید ناشی از هرس اتصالات برابر با تعداد درایه‌های غیرصفر باقیمانده (نرم صفر بردار) است. با مفروضات فوق، مسئله‌ی کاهش تعداد اتصالات شبکه را می‌توان به صورت زیر نوشت:

$$\min_{\mathbf{W}'} |E(\mathbf{W}) - E(\mathbf{W}')| \quad \text{s.t.} \ \|\mathbf{W}'\|_0 \leq B \tag{۱۱}$$

18ـ در ص8، مناسب است برخی بخش‌های پاسخ (10ـ پیشین) که در متن نیستند در جا(ها)ی مناسب در جملات پاراگراف اول تا سوم (یا ص2، مربوط به (10ـ پیشین)) درج شود.

پاراگراف زیر اضافه شد؛ و چون حجم بخش ۳ زیاد شده است، دو زیربخش به آن افزوده شد.

ممکن است این پرسش مطرح شود که درصورت حذف تعدادی از اتصالات، وضعیت شبکه به چه صورت درخواهد آمد؟ همانگونه که پیشتر ذکر شد، عموماً در عمل، گروهی از اتصالات حذف می‌شوند. هر لایه در یک شبکه‌ی عصبی پیچشی متضمن چندین فیلتر (گروهی از اتصالات) است. به عنوان نمونه، اولین لایه‌ی مدل VGG16 مطابق جدول ۱ دارای ۶۴ فیلتر است. حذف یک یا چند فیلتر از این لایه، خروجی و تعداد پارامترهای این لایه و ورودی به لایه‌ی بعدی را تحت تاثیر قرار خواهد داد. کاهش تعداد فیلترها، مترادف با کاهش حجم مدل است که هدف این نوشتار هم همین است.

19ـ ص8، خط ـ2، جمله اول: شماره بخش ذکر گردد.

انجام شد.

20ـ ص9، خط 4 و 5: یادآوری (42ـ پیشین)

پاراگراف کلاً ویرایش شده است.

21ـ ص9، مناسب است ابتدا و انتهای پاسخ (43ـ پیشین) به صورت مناسب در متن یا پانویس درج شود.

به گمان بنده بررسی حالات دیگر، خواننده را از هدف اصلی دور خواهد کرد و احتمالاً باعث گیج شدن وی می‌شود.

22ـ ص11، خط 2، فاصله: «مینیاتورفرشچیان» -> «مینیاتور فرشچیان»

فقط در یک جا «مینیاتورفرشچیان» به صورت متصل به هم نوشته شده، که باید به همین صورت باشد. در برنامه‌ی خزنده‌ای که برای جستجوی تصاویر در اینترنت از آن استفاده کرده‌ام، به دلایلی عبارات «مینیاتورفرشچیان» و «impressionism-claude-monet» دقیقاً به همین صورت مورد جستجو قرار گرفته‌اند.

23ـ ص11، خ ـ7، «قبلا اشاره شد» -> «در بخش ... اشاره شد»

انجام شد.

24ـ ص11، انتهای صفحه، مناسب است پاسخ (60ـ پیشین) اضافه گردد، مثلا:

این تشخیص نادرست توسط ماشین، نشانگر خطاهایی است که انسان هم ممکن است در تشخیص داشته باشد. یک تفاوت اصلی شبکه‌های عصبی پیچشی در یادگیری عمیق با سایر طبقه‌بندهای معمول مثل شبکه‌های عصبی چند لایه آن است که در شبکه‌های پیچشی، استخراج ویژگی توسط خود شبکه انجام می‌شود و تفسیر آن نیازمند بصری‌سازی لایه‌های میانی شبکه است. و در صورت نیاز، ارجاع به مرجع 17 انجام شود.

انجام شد.

25ـ ص12، پاراگراف اول یا در عنوان جول 2، اکنون مناسب است آن «...دستور summary در پای تورچ برای ایجاد خروجی ...» مربوط به جدول 1 می تواند در اینجا ذکر گردد.

انجام شد.

26ـ ص12، جمع بندی، خط ـ2، «... اولیه را داشت» -> «..اولیه را دارد»

«داشت» نبود. «داراست» بود. که بنا با نظر شما تبدیل به «دارد» شد.

27ـ ص12، پارارگراف اول، خط ـ4، «83درصد» -> «83 درصد» انجام شد.

28ـ ص12، پاراگراف دوم، خط 1، «جدول1» -> «جدول 1»

انجام شد. با اینکه عمداً فاصله را حذف کرده بودم که عدد ۱ به ابتدای خط بعد منتقل نشود.

29ـ پیشنهاد می شود در انتهای جمع‌بندی، سطری برای ذکر کاستی این روش و توسعه این مقاله یا مقوله ای دیگر در ادامه ی این کار به خواننده معرفی گردد.

انجام شد.

30ـ موارد متعدد از نگارش به صورت افعال اول شخص جمع و زمان آینده در متن (ص3، بخش 2.1؛ و بخش 3 در ص 8 و 9 و10) اصلاح گردند.

مواردی ویرایش شدند.

صفحه ۳ جدید:

۱.۲. مدل پرسپترون. روال اصلی یادگیری در شبکه‌های پرسپترونی مبتنی بر کمینه کردن خطای شبکه است. اگر هدف شبکه، طبقه‌بندی باشد، تابع هدف می‌تواند کاهش تعداد نمونه‌هایی باشد که به صورت نادرست دسته‌بندی شده‌اند[۲۰]. فرض کنید که مسئله‌ی موردنظر، طبقه‌بندی داده‌ها به دو کلاس است که برچسب دو کلاس (طبقه) به ترتیب صفر و یک درنظر گرفته شده و خروجی شبکه صفر یا یک است. اگر t_i[۲۱] برچسب درست طبقه‌ی نمونه‌ی iام و

ص ۹ جدید:

روش فوق در عمل کارایی نخواهد داشت؛ عموماً به جای حذف تکی، انتخاب و حذف نورون‌ها به صورت گروهی انجام می‌شود؛ اما انتخاب و حذف وزن‌ها با این روش، یک مسئله‌ی بهینه‌سازی ترکیبیاتی است. اگر در هر مرحله k اتصال از بین n اتصال انتخاب شوند، $\binom{n}{k}$ حالت برای انتخاب وجود دارد. به عنوان نمونه اگر مدل شامل یک میلیون اتصال بوده و هدف، انتخاب و حذف ده اتصال باشد، $\binom{10^6}{10} \approx 2.75 \times 10^{53}$ حالت ممکن برای انتخاب این ده اتصال وجود دارد که بررسی همه آنها عملی نیست.

پس از حذف تعدادی از اتصالات، شبکه نیازمند آموزش مجدد است؛ که این عمل نیز زمان‌بر است. اگر شبکه شامل n اتصال باشد و در هر مرحله تعداد کمی از اتصالات هرس شوند، بلافاصله نباید اقدام به انتخاب و حذف گروه بعدی اتصالات نمود. چون حذف اتصالات قبلی، ساختار شبکه را مقداری تغییر داده است و شبکه برای یک مجموعه داده‌ی مشخص، همان خروجی قبل از حذف را تولید نمی‌کند. پس از هرس تعدادی از اتصالات، شبکه باید مجدداً مقداری آموزش داده شود تا اثر ناشی از حذف اتصالات جبران گردد.

ص ۱۰:

پیش از این، بردار $\mathbf{W}$ که تعدادی از درایه‌های آن صفر باشند، با $\mathbf{W}'$ نمایش داده شد؛ برای حالت دو متغیره‌ی فوق‌الذکر و با فرض $\mathbf{W} = [w_1, w_2]$ و $\mathbf{W}' = [0, w_2]$، رابطه‌ی (۱۵) را می‌توان به صورت زیر نوشت:

مواردی هم در بندهای قبل ذکر شد.

31ـ در انتهای بیوگرافی، نقطه اضافه حذف گردد.

انجام شد.