وزارت ارتباطات و فناوری اطلاعات شرکت ارتباطات زیرساخت	عنوان مقاله: آینده ارتباط هوش مصنوعی ، انسان و هزینه های زیست محیطی آن تهیه کننده/گان:فریده محرمی مدرک و رشته تحصیلی : فوق لیسانس توسعه اقتصادی و برنامه ریزی رشته شغلی : حسابدار اداره کل/دفتر : معاونت راهبردی و توسعه کسب و کارهای نوین

آینده ارتباط هوش مصنوعی ، انسان و هزینه های زیست محیطی آن

چکیده

هزینه های زیست محیطی و محدودیتهای انرژی مسائلی را برای آینده توسعه یادگیری ماشین (ML[1]) و هوش مصنوعی (AI)ایجاد کرده است. تا به حال ، بحث در مورد اثرات زیست محیطی ML/AI[2] فاقد چشم اندازی فراتر از اندازه گیری های کمّی هزینه های تحقیقاتی مرتبط با انرژی بوده است. بنابر اساسهای پایه گذاری شده توسط شوارتز و همکاران[3] (۲۰۱۹) در گام اول Green AI استدلال ما شامل دو پدیده در هم تنیده ی قابلیت تعمیم رایگان و آینده ای می شود که در آن ML/AI اکثریت موارد استنتاجهای قیاسی قابل سنجش را اجرا می کند .قابلیت تعمیم رایگان به تفاوت تقاضاهای شناختی کاری که انجام می شود و دقت بکاربرده شده در مدل ML/AI اطلاق می گردد . اگر دومی بیشتر از اولی باشد ، مدل به دلیل بهینه شدن برای دستیابی به بهترین دقت ممکن، ناکارامد می شود و عملکرد آن برای محیط زیست مضر می گردد. آینده تحت سلطه اجزای غیرانسانی، استفاده از ML/AI را بسیار گسترده توصیف می کند که بسیاری از استنباطهای قیاسی به تعمیم ML/AI مجهز می شوند. این مقاله استدلال می کند که بحث حاضر در رابطه با هزینه های زیست محیطی تحقیق و کاربردهای بی اثر ML/AI (موضوع قابلیت تعمیم رایگان) با آینده ای که مشخصه بارز آن هوش مصنوعی همه گیر می باشد، مستحق توسعه و گسترش است.

۱. مقدمه

به عنوان یک رشته علمی شناخته شده ، ML به دنبال توسعه "ابزار برای عملکرد مطلوب"[4] است . با توجه به وظیفه و دلایلی که فراگیری آن را می تواند راحت تر نماید ، "ابزار بهترین عملکرد"،جایگاه خود را با بدست آوردن قابلیت تعمیم کسب می کند

---

[1] Machine learning
[2] Artificial intellegent/Machine learning
[3] Schwartz et al
[4] tool-for-optimal-action

. سپس چنین ابزاری،از استنتاج هایی که می تواند فراتر از شواهد (داده های آموزشی) تعمیم یابد، پشتیبانی می کند . یعنی می تواند استنباطهایی را روی نمونه های جدید انجام دهد مشروط بر اینکه این موارد از همان منبع باشد یا به میزان زیادی مشابه توزیع احتمال شواهد (به عنوان داده های آموزشی) باشند . این رشته دارای یک هدف شناختی دوگانه است. اول ، بررسی نظری آن شامل فرضیات رسمی در مورد یادگیری که منجر به تعمیم می شود ، به وسیله یادگیری آماری احراز یا تایید می شود. (Kawaguchi, Kaelbling, & Bengio, 2019; Vapnik, 1995). دوم از دیدگاه تجربی این رشته به دنبال افزایش دقت استنباطهایی است که کلیات بدست آمده را ارائه می دهند .به نظر می رسد در حال حاضر ، تقویت متقابل این دو هدف فرعی نسبتا سخت باشد و عدم تقارن ایجاد کند. اگرچه با سرمایه گذاری زیاد در تحقیق نظری، این حوزه همچنان مسلط بر دومین هدف فرعی باقی می ماند. زمانی که نتایج تجربی قوی از درک تئوری کامل فراتر می رود ، تابلوهای راهبری (لیدربوردها) ، با وظیفه مشخص و قابلیت ردیابی دقت به اقدامی برای ارزیابی پیشرفت شناختی این حوزه تبدیل می شوند. تاکید بر یک هدف واحد-دقت- ایده ساده ای را القاء می کند مبنی بر اینکه تابلوهای راهبری به مانند نردبان هستند. هرچه مراحل بیشتر باشد ما به کاهش بار شناختی برای انجام وظایف مختلف از طریق بکاربردن "ابزارهای تقریبا کامل برای عملکرد بهینه" نزدیکتر خواهیم بود . اگر هدف معرفتی ML را به عنوان فهم کلیات تعبیر کنیم تابلوهای انگیزشی مشکلاتی را برای خود رشته و همچنین مشکلات قابل توجهی را برای محیط زیست به وجود می آورد. این موضوع به هزینه محاسبه شده منابع به وسیله الگوریتم ها مربوط می شود که صعود به موقعیت های بالاتر در تابلوهای راهبری را ممکن می سازد.در سناریوهایی که فهم نظری از نتایج تجربی عقب تر است، آخرین پیشرفت های علمی که معمولا از تجربیات آزمون و خطاء بوجود می آید اغلب منجر به نتایج اکتشافی (غیر مستدل) کاملا تصادفی می شود. تابلوهای راهبری در مواجهه با خلاء تئوریکی متخصصین با وسوسه حدس و گمانهای بعد از آن روبرو می شوند که ممکن است نقشی را که توضیحات نظری (تئوریکی) ایفاء می کنند ، را بپذیرند. (رجوع کنید به Lipton & Steinhardt, 2019) زمانیکه بطور تصادفی همراه با این موضوع سوء تفسیر در مورد منابع نتایج تجربی نیز رخ دهد ، به دنبال آن از تنظیم و میزان کردن ابرپارامترها ممکن است تغییرگزارش پیشرفتهای معماری عصبی ،رخ دهد(همان مرجع). به جای دستیابی به هدف شناختی کلیات ، تابلوهای راهبری (لیدر بورد ها) صرفا ممکن است فرضیه های بعد از آن، مطابق با نتایج اکتشافات ناشی از آزمون و خطاء کاملا تصادفی را ترغیب نماید. بطور باور نکردنی سرعت لیدر بوردها افزایش می یابد،پردازش زبان طبیعی NLP[۵] در بین بهترین نمونه های حاضر است (مراجعه کنید به Strubell, Ganesh, & McCallum, 2019 برای تخمین هزینه های زیست محیطی پردازش طبیعی زبان طبق آخرین یافته های علمی.) که اقدامات بد را به آژیر خطر تبدیل می کند که می تواند بطور قابل ملاحظه ای مانع هدف معرفتی دوگانه این رشته بشود.

---

۵ Natural language processing

۲- ارتباط هوش مصنوعی انسان و هزینه های زیست محیطی آن

کلیات به دنبال حداقل سازی ریسک تجربی است که بایستی استنتاج های دقیق مرتبط با کار مورد نظر را تضمین نماید. بنابراین کلیات یادگیری به دنبال حداقل سازی ریسک استقرایی مرتبط با کار می باشد. از بدنه رویکردهای نظری به قیاس کل از جزء(Norton's material theory (2003))،با فرض اینکه استنتاج های استقرایی در حقایق محلی موجود در دامنه های مشخص پایه گذاری شده باشند ،قابلیتی را برای مشخص نمودن شکل ارتباط آینده هوش مصنوعی و انسان نشان می دهد. قابل قبول است که استدلال کنید یک ML موفق دارای کلیات، یک طرح القایی را ایجاد می کند و ریسک آن را با محدود کردن آن به حقایق محلی (بومی) که در داده های آموزش یافت می شود،به حداقل می رساند. مطابق گفته نورتون ML موفق ممکن است از منظر معرفتی (شناختی) توضیح داده شود و به یاری تعداد زیادی از طرحهای استقرایی محلی در چند طرح جهانی دشوار تایید شود .صرف نظر از اینکه یادگیری ماشین یا انسان دارای کلیات را در نظر بگیریم یا نه ، هرگونه حداقل سازی ریسک تجربی با بومی سازی طرح استقرایی شامل هزینه های زیست محیطی است.

می توان استدلال کرد که پژوهش های ML دستیابی به کلیات طرحهای محلی استقرایی مورد تایید حقایق محلی (از شواهد و مدارک یعنی داده های آموزش) را به دنبال دارد که حداقل سازی ریسک تجربی را ممکن می سازد. بطور مشابه همانطور که توسط نورتون (۲۰۰۳) مطرح گردید برای بشر طرح استقرایی جهانی وجود ندارد و استنتاجهای استقرایی صرفا در محدوده های خاصی توسط حقایق محلی تایید می شوند. درک گسترده تر و عمیق تر از حقایق محلی،دستیابی شناختی بشر و همچنین ماشین ها را افزایش می دهد. وجه تمایز مهم اینست که انسانها به دلیل فشارهای تکاملی ، در مقایسه با نمونه ML ناکارآمد به عنوان یادگیرنده های کم هزینه باقی می مانند (به عنوان مثال آزمایش اخیر Open AI Five(Open AI, 2019) که در آن ۴۵۰۰۰ سال آموزش انسانی برای شکست بهترین بازیکنان انسانی Dota 2 computer game مورد نیاز است) .به دلایل زیادی این اغراق نیست که فرض کنید تلاش شناختی (معرفتی) بشر در حال پشت سر گذاشتن دوره تحولات بی سابقه ای می باشد. در حالیکه ممکن است این اتفاق نیز بیفتد که از مجموع تعداد قابل توجهی استنتاج های استقرایی، اکثر آنها به وسیله ابزارهایی برای اقدام بهینه بکارگرفته شوند. مهم تر از همه ، به احتمال زیاد تعداد کل استنتاج ها زیاد خواهد شد زیرا که بشر مشتاقانه بخشهای متزلزل تلاشهای شناختی متداول و علمی اش را بطور یکسان افزایش خواهد داد و حتی فراتر از آن احتمالا چیزهای جدیدی را اختراع خواهد کرد. در نتیجه ، هزینه های زیست محیطی بومی سازی انسانی و مصنوعی ریسک استقرایی که با تمام کلیات استقرا مطابقت دارد نیز افزایش خواهد یافت. اما در مقایسه با استنتاج های استقرایی بشر ، طرح های استقرایی که توسط کلیات ML در دسترس قرار می گیرند، درصورت آسیب به محیط زیست، با قیمت های قابل ملاحظه ای همراه خواهد بود .

۳- هزینه های زیست محیطی پژوهش مربوط به یادگیری ماشین

بطور اجمالی ، بکار بردن دقت برای عمل به آخرین پیشرفتهای علمی(SOTA) که تابلوهای راهبری(لیدر بوردها) را با استفاده از همه ابزارهای موجود ارتقاء می دهد، تعمیم کلیات ML را به عنوان یک تلاش مضر برای محیط زیست در نظر می گیرد. با استفاده از منابع محاسباتی زیاد می توان دقت بالا را خریداری نمود. این موضوع رشد سه برابری را فراهم می کند و بطور معمول منجر به دقت جدید SOTA می شود. (Schwartz, Dodge, Smith, & Etzioni, 2019) اما منابع محاسباتی با هزینه های زیست محیطی همراه می باشند ،و زمانیکه رابطه بین مقیاس آزمایش ها و دقت حاصله نامطلوب می شود از اهمیت زیادی برخوردار می گردند. (رشد نمایی آزمایش برای افزایش خطی تقریبی در دقت ، رجوع کنید به همان منبعی که قبلا ذکر گردید.) شوارتز و سایرین (۲۰۱۹) هزینه افزایش دقت تابلوهای راهبری را به صورت رشد سه برابر آزمایشها رسمیت بخشیده، فرموله کرده و در قالب یک طرح بیان می دارد.

هزینه (نتیجه)∝ E.D.H (۱)

هزینه کل آزمایش ML با افزایش E ، افزایش هزینه پردازش یک نمونه، (D)،حجم داده های آموزشی(H)و ، انواع آزمایشات انجام شده برای یافتن دقت جدید SOTA (همان منبع قبلی) بطور خطی افزایش پیدا می کند. ارتباط نامطلوب بین هزینه و دقت که توسط معادله (۱) نشان داده می شود به لحاظ تجربی به صورت روند کلان مصرف منابع محاسباتی در حال رشد موجود در ML و همچنین پیچیدگی رو به افزایش آزمایشات ML قابل تشخیص است. (Amodei & Hernandez, 2018; Sastry, Clark, Brockman, & Sutskever, 2019) این پیچیدگی با تعداد دفعاتی که مدل ML اولیه که برای یافتن یک برازش خوب معماری و ابر پارامترهای مدل برای داده ها یعنی برای انجام وظیفه مورد نظر ، مجددا آزموده می شود ، افزایش می یابد . بنابراین به این ترتیب به دقت جدیدSOTA می رسد. در سازمانهایی که تحقیق پیرامون معماری/ابرپارامترها توسط انسان هدایت می شود ما هزاران چرخه آموزشی را در هر آزمایش مشاهده می کنیم .(عبارت سوم معادله (۱) رجوع کنید به استروبل و همکاران ۲۰۱۹). در سازمانهایی که جستجوی تکاملی مصنوعی جایگزین هدایت و راهنمایی انسان پیرامون فضاهای معماری / ابرپارامترها شده است(یعنی پژوهش معماری عصبی)،تعداد مدل های آموزش داده شده به ده ها هزار می رسد. (رجوع کنید به Real, Aggarwal, Huang, & Le, 2019; So, Liang, & Le, 2019) بقیه این بخش با استفاده از تجزیه و تحلیل اولیه استروبل و همکاران (۲۰۱۹) به دنبال برطرف کردن یک مفهوم با نیت خوب ولی در نهایت نادرست است که فرض می کند یک دوره آموزشی در خصوص یک مدل قبلا توسعه داده شده ML در یک مجموعه داده های معیار می تواند به نحوی نماینده هزینه های زیست محیطی مرتبط با پژوهش ML باشد. برای توضیح بهتر این نکته ، بایستی نگاهی به یکی از برجسته ترین تابلوهای راهبری (Image Net)، (یک چالش طبقه بندی تصاویر و یک ML مدل برجسته) بیندازیم. حتی اگر در حال حاضر امکان آموزش ResNet-50 (He, Zhang, Ren, & Sun, 2016) به دقت رتبه بندی بالا در تابلوی راهبری (لیدربورد)Image Net در ۲:۴۳ دقیقه (بهینه سازی برای سرعت) یا برای ۱۲,۶۰ دلار (بهینه سازی برای

هزینه) وجود داشته باشد، بدان معنی نیست که سخت افزارها و یا فنون آموزش بهتر باعث کاهش هزینه های صعود در تابلوهای راهبری(لیدر بوردها) می شود. برعکس ، رسیدن به رده های بالاتر و در نتیجه دقت نیازمند مدل های بزرگتر و داده های بیشتری است. در مورد مجموعه داده های معیار[۶] مثل Image Net که برای تابلوهای راهبری(لیدربورد) مورد نیاز است به معنای دستیابی داده ها برای "پیش آموزش" می باشد که امکان کسب آن همراه با سایر موارد پیشرفت و دقت بالاتر وجود دارد. در سال ۲۰۱۹ تورون، ودالدی و دوز[۷] ۹۴۰ میلیون تصویر عمومی برای آزمایش پیش آموزش با نظارت ضعیف در خصوص معماری ResNeXt-101 32 × 48d شامل ۸۲۹ میلیون پارامتر (۲۵,۶ میلیون پارامتر versus ResNet-50) و نتایج خوب تنظیم شده برای Image Net استفاده کردند . این آزمایش در حال حاضر (ژانویه ۲۰۲۰) در رتبه چهارم تابلو راهبری(لیدربورد) Image Net قراردارد. منابع موردنیاز برای انجام چنین آزمایشی با یک دوره آموزش واحد مدل ML استاندارد و قبلا توسعه داده شده (مثل ResNet-50) در یک مجموعه داده های معیار قابل مقایسه می باشد. پیامد جایگاههای برتر در لیدر بوردها بطور شفاف، آزمایشهای بزرگتر (داده ها و مدلهای ML بطور یکسان) ، افزایش تقاضای محاسباتی، و در نتیجه افزایش تاثیرات زیست محیطی می باشد. بنابراین، حتی اگر پیشرفتی را در راندمان سطح پایه مشاهده کنیم نشانه ناچیز بودن تاثیرات زیست محیطی ML نمی باشد. همانطور که در روند کلان بیان شده است آمودی و هرناندز در سال ۲۰۰۸ (Amodei and Hernandez) نشان دادند که از زمان شروع دوره یادگیری عمیق (DL) در ML میزان منابع محاسباتی صرف شده در بزرگترین آزمایشها که مربوط به جایگاههای برتر در تابلوهای راهبری است ، هر ماه ۳/۴ برابر می شود. از سال ۲۰۱۲ که Alex Net (Krizhevsky, Sutskever, & Hinton, 2012) برای اولین بار بر روشهای غیر[۸]DL در خصوص تابلوی راهبری ImageNet مسلط شد ، مقدار محاسبه به بیش از ۳۰۰۰۰۰ برابر افزایش یافت و در حال حاضر این آمار بصورت روزانه گزارش می شود که در آن آزمایش در یک پتافلاپ بر ثانیه[۹] (عملیات نقطه شناور در ثانیه آمودی و هرناندز ۲۰۰۸) در چندین GPU یا TPU واحد پردازش گرافیک ، واحد پردازش TENSOR ، پتافلاپ بر ثانیه صرف شده که با سرعت بیشتری و موازی با افزایش آزمایشات ، افزایش می یابند ، اجرا می شود. این مقیاس از پتافلاپ بر ثانیه یا روز زمان اجرا به صدها برابر پایین تر یا حتی بالاتر متغیر است. (این به معنای ده ها هزار سال در چارچوب مرجع زمانی می باشد. OpenAI, 2019) در صورتی که آزمایشهایی که بیشتر مورد درخواست است، به بیش از هزار پتافلاپ بر ثانیه یا روز برسد، (آمودی و هرناندز ۲۰۰۸) به معنای اینست که حد بالای امروز نماینده آزمایش خوب نیست).آمودی و هرناندز ۲۰۰۸ پیشنهاد نمودند که حداقل در کوتاه مدت احتمال ادامه روند کلان وجود دارد زیرا که ما در مورد بهبود در نسبت فلاپ برثانیه در هر وات و همچنین فرصتهای

---

[۶] benchmark

[۷] Touvron, Vedaldi, Douze, & Jégou,

[۸] Deep learning

[۹] petaflop/s

موجود برای استفاده بهتر از همسانی آزمایشات ML بحث نکرده ایم. و هنوز حرف برای گفتن وجود دارد . بنابراین صرفا با درنظرگرفتن محدودیتهای اقتصادی ، کسب راندمان سطح پایه بزرگتر یعنی دوره آموزشی سریع تر و ارزانتر از مدل ML قبلا توسعه داده شده در مجموعه داده های معیار ، محتمل است که منجر به آزمایشات بیشتر، پیگیری دقت بالاتر و صعود در تابلوهای راهبری شود. زمانیکه خود را به شروع روند نزدیکتر ببینیم، برآورد انتشار CO_2 ناشی از آزمایشات به موقع است. تولید گازهای گلخانه ای ناشی از مصرف انرژی برای برآوردن تقاضای محاسباتی آزمایشات می باشد و بنابراین سهم ML به تماس بشر با طبیعت سیستم زمین را تغییر می دهد.

استروبل و همکاران در سال ۲۰۱۹ (Strubell et al) مصرف برق آزمایش ML را بصورت مجموع برق مصرف شده توسط سخت افزار ضربدر نیاز برق اضافی برای برای حفظ زیرساخت ها برآورد نمودند(یعنی سرمایش و غیره). مقدار CO_2 منتشرشده بوسیله آزمایش با ضرب مجموع مصرف برق در متوسط انتشار CO_2 در هر کیلو وات ساعت که توسط آژانس حفاظت از محیط زیست آمریکا ارائه می گردد، محاسبه می شود.(Strubell et al. (2019) EPA) . پس از آن استروبل و همکاران در سال ۲۰۱۹ انواع منتخبی از آزمایشهای ML به متوسط CO_2 منتشر شده طی مسیر رفت و برگشت از نیویورک تا سانفرانسیسکو (1 passenger, 1984 lbs [900 kg]) ، یک سال از عمر متوسط انسان /یک سال عمر متوسط یک آمریکایی (11,023 lbs [5 tons]/36,156 lbs [16.4 tons]) یا متوسط عمر ماشین شامل سوخت مصرفی را باهم مقایسه نمودند. با درنظرگرفتن ML در پردازش زبان طبیعی (NLP) زیادی انتشار CO_2 ناشی از این مسئله هست که آیا تحقیق در خصوص فضاهای معماری و ابرپارامترها توسط محققان بشری انجام می شود یا توسط تحقیق تکاملی مصنوعی .با درنظرگرفتن مثالی از نمونه قبلی آزمایش ارائه شده توسط استروبل و همکاران مبتنی بر توسعه مدل NLP جدید ، میزان انتشارات CO_2 ، ۷۸۴۶۸ lbs برآورد می شود .(۳۵/۶ تن با استفاده از متوسط انتشارات CO_2 ایالات متحده در هر کیلو وات ساعت توسط EPA منتشر شده است.) تولید گازهای گلخانه ای که از ۴۷۸۹ مدل در طی آزمایش جمع آوری شده (عبارت سوم معادله اول) منجر به بهترین مدل هماهنگ ، گردیده است ،که SOTA را در برخی از تابلوهای راهبری(لیدر بورد) بوجود می آورد. برای آزمایش NLP از آخرین گروه آزمایش مبتنی بر معماری عصبی پژوهش یعنی تحقیق تکاملی برای یافتن بهترین مدل ، استروبل و همکاران در سال ۲۰۱۹ مقدار انتشار CO_2 را ۶۲۶۱۵۵ lbs (۲۸۴ تن با استفاده از همان روش برآورد بالا) برآورد نموده اند. شاید به عنوان داستان اخطارآمیز موثق آزمایش اخیر (منگ و همکاران Meng et al) با ادعای کاربرد ۵۱۲ GPUs برای سه ماه، بتواند به عنوان پیش نمایش برای مدل ترجمه ماشینی در حدود ۴۰ میلیارد جمله را انجام دهد. این پیش نمایش یک بهبود متوسط SOTA را در خصوص مجموعه داده های معیار به دنبال داشت، این آزمایش فاقد (همکاری و کمک از ناحیه) مبانی است و مجموعه داده های بزرگ مورد استفاده برای این پیش نمایش هنوز منتشر نشده است . از آنجا که نمی توان برای تایید نتایج آزمایش را تکرار نمود و این سوال آشکار وجود دارد که آیا باید آزمایش تکرار شود یا نه . تنها نتیجه ملموس و

ماندگار احتمالا انتشار میزان قابل توجهی CO_2 است.سرانجام باید تاکید نمود که نتایج آزمایشهایML فاقد انتقال پذیری هستند. این بدان معنی است که آموزش/تنظیم دقیق برای استقرار مدلهای ML در حوزه های جدید یا برای انجام وظایف جدید لازم است . بنابراین انتظار می رود که میزان انتشار CO_2 افزایش یابد. (رجوع کنید به استروبل و همکاران ۲۰۱۹). میزان CO_2 منتشر شده ناشی از آزمایشهای ML در مقایسه با انتشار CO_2 مرتبط با مناطق متداول زندگی انسان، غیر مجاز می باشد. علاوه براین، توسعه انفجاری اخیر ML به عنوان یک رشته مورد نظر است که با افزایش مقالات پژوهشی در این زمینه می توان آن را توضیح داد . در پایان سال ۲۰۱۸، تعداد ارسال ها به بخش های arxiv.org مرتبط با ML (مخزن محبوب نسخه های الکترونیکی که توسط دانشگاه کرنل اداره می شود) به ۳۰۰۰ مقاله در ماه رسید (Dean, 2019). روند انفجاری مشابهی در تعداد کنفرانس های دانشگاهی ML وجود دارد. کنفرانس دانشگاهی پیشرو در زمینه ML ، NeurIPS در طی پنج سال گذشته به ۶۷۴۳ مورد ارسال رسیده است. (Beygelzimer, Fox, d'Alché-Buc, & Larochelle, 2019) هرچند همه آزمایشات به اندازه مثالهای فوق سخت نیستند. در مواجهه با واقعیت آزمایشهای بزرگ، که به یک روش استاندارد برای بهبود دقت تبدیل شده است ، این رشته تصدیق می کند که هدف معرفتی آن جستجو و تلاش برای فهم تجربی و نظری تعمیم هزینه های زیست محیطی را دربردارد. انجمن ML دو راه حل ممکن را ارائه داد. ابتدا ، منابع محاسباتی مورد نیاز برای بالا بردن دقت در لیدر بورد (تابلو راهبری) بایستی صرفا توسط انرژی حاصل از منابع تجدید پذیر تامین شود ، بنابراین ، پایداری ۱۰۰درصد را تضمین می کند.(رجوع کنید به .Hölzle, 2019) . با اینحال ، ادعای مصرف ۱۰۰درصد انرژی تجدید پذیر معمولا به مصرف سالیانه اطلاق می شود ولی مصرف واقعی در زمانهای معین مثلا در هنگام شب به جای انرژی خورشیدی هنوز از سوزاندن سوختهای فسیلی تامین می شود.(de Chalendar & Benson, 2019) .مصرف ۱۰۰ ٪ انرژی تجدیدپذیر نیازمند ذخیره مازاد انرژی در زمان حداکثر عرضه انرژی تجدید پذیر می باشد. دوم، مکانهای انتشار بایستی بطور صریح این تحقیق را انجام دهند که دقت رقابتی و البته غیر SOTA ،مقادیر معادله (۱) را کاهش می دهد.(رجوع کنید به Schwartz et al., 2019) ، همچنین یک تحقیق مرتبط با این موضوع گزارشی را در خصوص نتایج معتبر بدست آمده در طی آموزش به منظور تخمین بودجه محاسباتی مورد نیاز برای دقت و صحت اعتبار سنجی معین را پیشنهاد می کند. (رجوع کنید به Dodge, Gururangan, Card, Schwartz, & Smith, 2019)

۴. هزینه های غیر موجه زیست محیطی قابلیتهای تعمیم

پیش بینی اثرات زیست محیطی بالقوه ناشی از اهداف معرفتی این رشته و تمایل به مقابله با آنها واقعا ستودنی است. با اینحال تصویر اثرات زیست محیطی مرتبط با ML کامل نمی باشد . این موضوع صرفا نوک کوه یخی می باشد، بقیه موارد مرتبط با الگوی تعمیم یادگیری ،بطور فراگیر **"ابزار برای عملکرد مطلوب"** را به منظور بالا بردن دقت بهینه می کند. برای بدست آوردن یک تصویر دقیق از تاثیر زیست محیطی ، جدا از هزینه تحقیق که بوسیله معادله (۱) نشان داده شده است ، لازم است

که هزینه ای را نیز که در حین استفاده از ابزار قابلیت تعمیم برای انجام وظیفه تعیین شده جمع می شود ، را نیز به حساب بیاوریم. در مرحله اول، برای برآورد نمودن انرژی مصرف شده بوسیله یک برنامه از ابزار قابلیت تعمیم ، محاسبه تعداد عملکردهای محاسباتی در هر استنتاج مورد نیاز است. دوم، این مقدار بایستی به تعداد استنتاج هایی که انتظار می رود توسط همه موارد آینده **ابزار بهترین عملکرد** به عنوان چهارمین عضو معادله (۱) انجام گردد، ضرب شود.

$$Cost(Result) \propto E \cdot D \cdot H \cdot cI \qquad (۱)$$

در(۱) Eq اصلاح شده ، C نماینده تعداد عملیات محاسباتی در هر استنتاج وI بیانگر تعداد کل استنتاج های آینده می باشد. صرفا پس از چنین اصلاحی همگرایی (۱) Eq در خصوص تاثیر زیست محیطی درست این ابزار شروع می شود. در این زمینه ، لازم است تاکید شود که با افزایش دقت، زمان مورد نیاز برای انجام یک استنتاج نیز افزایش می یابد (رجوع کنید به Bianco, Cadene, Celona, & Napoletano, 2018) . بنابراین ، مقدار منابع محاسباتی و انرژی مورد نیاز برای استقرار را افزایش می دهد. اگرچه اخیرا برخی نتایج دلگرم کننده ای حاصل شده است که نشان دهنده پیشرفت در نسبت دقت به زمان تاخیر است(رجوع کنید به Gupta & Tan, 2019) ، عبارت دوم و سوم (۱) Eq که برای پیشرفت مورد نیاز است همچنان پرهزینه می باشند.

با اطمینان می توان فرض نمود که اشتیاق ما به مجاز نمودن ML برای مراقبت از حتی ناچیزترین کارهای روزمره ،سرانجام منجر به استفاده فراگیر از کلیات ارائه شده توسط ابزارهای مختص عملکرد بهینه خواهد شد.اما چنین چشم اندازی حتی برآورد تقریبی چهارمین مقدار را مختل می نماید که متاسفانه نسخه اصلاح شده (۱) Eq را بصورت غیرعملی ارائه می دهد. با رسیدن به بن بست ، این مسئله بطور شفاف نیازمند رویکرد متفاوتی است که ممکن است از تغییر چشم انداز ذیل نمایان شود. جدا از به حداقل رساندن مقدار(۱) Eq ، یک رویکرد اصولی، تعریف مجدد مفهوم"ابزار برای عملکرد بهینه" می باشد. به جای بهینه کردن یک هدف واحد یعنی دقت "ابزار برای بهترین عملکرد" رعایت محدوده خاص وظیفه که تعداد حداکثر محاسبات را در هر استنتاج تعیین می نماید، ضروری است. این محدودیت مصرف انرژی ابزار را از بین خواهد برد و رد پای زیست محیطی آنرا تضعیف خواهد کرد. همچنین دقت قابلیت تعمیم آن را کاهش خواهد داد. در ارتباط با این موضوع ، هدف دوم به دنبال انعکاس این واقعیت است که حتی سطح کاهش داده شده قابلیت تعمیم ممکن است یک ابزار بهینه ارائه نماید. مثلا ، وظیفه دم دست ممکن است شامل فردی باشد که اصلاحاتی را (معمولا برای گسترش معرفتی یا شناختی) را انجام می دهد یا شامل یک خطای ناچیز که ابزار قدرتمندتری را نمی توان توجیه کرد.

به بیان رسمی ، انجام همه وظایف شناختی بطور مساوی خواستار تعمیم است ، در حالیکه بهینه سازی "ابزار برای عملکرد بهینه" صرفا برای دقت SoTA است که مازاد کل قابلیت تعمیم را ایجاد می کند. یادگیری تعمیم تک هدفی مهم از هدف معرفتی متغیر ناشی می شود که از افزایش دقت تابلوهای راهبری(لیدر بورد) استفاده می کند. از این رو این تعامل دوگانه قابلیت های عمومی سازی مجانی را ایجاد می کند که ترس از بین رفتن و در نتیجه مضر برای محیط زیست وجود داشته باشد. می توان استدلال کرد که یادگیری تعمیم چند هدفی می تواند حداقل این تاثیرات مضر را کاهش دهد. با روش ارزیابی پیچیدگی شناختی هر کار مورد نظر ، بایستی "ابزارهای مورد نظر برای عملکرد بهینه" را بر مبنای سبک و سنگین کردن مصرف انرژی و قابلیت تعمیم ،مورد استفاده قرار دهد، همچنین کاهش هزینه تحقیقات را که بوسیله Eq(1) نشان داده شده، تشویق کند. بطور متفاوت ، با توافق بر روی دقت معقول، که تقاضای شناختی کار مشخصی را تقریب می دهد ، محدود نمودن بودجه محاسباتی یادگیری تعمیم ((Eq(1) و متعاقبا بودجه زمان استقرار استنتاجات امکان پذیر می شود.

با پیش بینی احتمال استفاده همه جانبه از تعمیم های ML ، تعریف " هدف چندگانه" تضمین می کند که "ابزار برای عملکرد بهینه" دقت تصریح شده را بدون ارزیابی قابلیت های تعمیم مجانی به قیمت آسیب به محیط زیست به دست بدهد. دو تغییر در هدف معرفتی این رشته می تواند به چنین نتیجه ای کمک نماید. اول ، شاید بهتر باشد که مسابقه برای رسیدن به بالای تابلوهای راهبری(لیدربورد) کنار گذاشته شود مگر اینکه اتکاء به پیشرفتهای بینهایت کوچک را با استفاده از بالابردن مقادیر Eq(1)متوقف نماید. دوم، این رشته ممکن است در آستانه کشف مبانی نظری باشد که آینده یادگیری تعمیم را به تلاش اصولی تغییر دهد. (Frankle & Carbin, 2019; Jiang, Neyshabur, Mobahi, Krishnan, & Bengio, 2019; Nagarajan & Kolter,). در آن مورد ، به احتمال زیاد می توان اهداف شناختی و زیست محیطی را محقق نموده و به میزان قابل توجهی هزینه تحقیق نشان داده شده بوسیله Eq(1) را کاهش داد.

۵. نتیجه

با تحقق کامل چنین سناریویی می توان حوزه فنی را هم دید (Haff, 2014) ، یعنی بخشی از سیستم زمین که تمدن مدرن و ساکنان آن را تقویت می کند و بعد معرفتی یا شناختی آنرا رشد میدهد. با اشتیاق و شور و شوق فعلی برای گسترش شناختی ML/AI ،پیش بینی می شود آینده تلاشهای هر روزه و علمی بشر احتمالا بر اساس میزان زیادی از استنتاج های قیاسی غیر انسانی (غیر بشری) باشد. اگر آینده ارتباط هوش مصنوعی بشر به اثرات مضر (ناکارامد)زیست محیطی ML/AI وابسته باشد ، بعد شناختی حوزه فنی احتمالا خطرات آنتروپوسن (اشرف مخلوقات) را بصورت بدتر نمودن سیستم زمین تشدید خواهد نمود. با پی بردن به شکل احتمالی وقوع ارتباط هوش مصنوعی بشر ، امیدواریم که بشر بزرگترین از سرگیری تلاش شناختی اش را نشان دهد به این دلیل که روش علمی ممکن است هزینه های زیست محیطی را به دنبال داشته باشد ، اگر ابتکار عمل نظیر Green AI روی گوش های ناشنوا قرار بگیرد.

این مقاله یک روش فنی –فلسفی تفکر در مورد هزینه های زیست محیطی ML/AI را پیشنهاد می نماید ، که به دنبال پیشنهاد یک روش جایگزین برای اغلب عقاید هنجاری و قانونی برانگیزاننده می باشد . می توان استدلال کرد که اگر تحقیق و برنامه های کاربردی ML/AI ناکارآمد باقی بمانند ، بایستی برای هزینه های پیش بینی نشده زیست محیطی آماده باشیم. ناکارآمدی در یادگیری با هدف واحد ، به هر قیمتی به دنبال بهترین دقت ممکن است. یک راه حل ممکن می تواند بر اساس یادگیری چند هدفی ، که اولین هدف ، وظیفه منطقی تعریف شده ، کران بالای دقت را برای دومین هدف تامین می کند.