



در سال‌های اخیر، شبکه‌های عصبی سبب بهبود عملکرد سامانه‌های هوش مصنوعی شده‌اند. سامانه‌هایی که در کاربردهایی نظیر خودران‌ها، تشخیص گفتار، بینایی کامپیوتری و ترجمه خودکار به کار می‌روند. اما شبکه‌های عصبی حافظه و توان زیادی مصرف می‌کنند و به همین دلیل معمولاً روی سرورهایی در کلاود اجرا می‌شوند، به طوری که داده‌ها را از کامپیوترهای رومیزی یا دستگاه‌های همراه نظیر گوشی‌های هوشمند دریافت و نتیجه پردازش را برای آن‌ها ارسال می‌کنند.

سال گذشته، ویویان زی دانشیار مهندسی برق و علوم کامپیوتر دانشگاه ام‌آی‌تی و همکارانش از تراشه کامپیوتری کم‌مصرفی رونمایی کردند که برای شبکه‌های عصبی بهینه‌سازی شده بود و امکان پیاده‌سازی محلی سامانه‌های هوش مصنوعی قدرتمندی را فراهم می‌کرد که روی دستگاه‌های موبایل اجرا می‌شدند. حال او و همکارانش سعی کرده‌اند به این موضوع از جهت دیگری بپردازند و شبکه‌های عصبی با مصرف توان کمتر را طراحی کنند. آن‌ها در ابتدا با روش‌های تحلیلی سعی کردند مشخص کنند یک شبکه عصبی وقتی روی یک سخت‌افزار نوعی اجرا می‌شود، چقدر توان مصرف خواهد کرد. سپس آن‌ها از این یافته‌ها برای ساده‌سازی شبکه‌های عصبی استفاده کردند، به گونه‌ای که روی دستگاه‌های همراه دستی بهینه‌تر اجرا شوند. آن‌ها در این تحقیق کاهش 73 درصدی را در مصرف توان یک پیاده‌سازی استاندارد شبکه عصبی و کاهش 43 درصدی مصرف توان نسبت به روش بهینه‌سازی شبکه‌های عصبی که پیش از این استفاده کرده بودند مشاهده کردند. ابزاری که آن‌ها طراحی کردند، میزان مصرف توان یک شبکه عصبی را بر اساس معماری شبکه و مقادیر وزن‌های آن محاسبه می‌کند. بعد از اینکه مشخص شد انرژی چطور مصرف می‌شود، از نتایج این بررسی برای طراحی شبکه عصبی با مصرف انرژی استفاده کردند.

مطلب پیشنهادی



رایگان دانلود کنید: کتاب الکترونیکی «یادگیری ماشینی؛ سفری به اعماق هوشمندی»

لازم به ذکر است که محققان در گذشته برای کاهش توان مصرفی شبکه‌های عصبی از روشی به نام هرس کردن (Pruning) استفاده می‌کردند. در این روش، بسیاری از اتصالات بین گره‌ها که وزن کمی دارند و تأثیر بسیار اندکی در نتیجه نهایی شبکه عصبی خواهند داشت، حذف یا به اصطلاح هرس می‌شوند. ابزار جدید محققان ام‌آی‌تی کمک می‌کند تا فرآیند هرس کردن شاخه‌های شبکه به طور آگاهانه‌تر و با هدف کاهش مصرف توان انجام شود. هرچند بریدن شمار زیادی از اتصالات با وزن کم تأثیر اندکی روی خروجی نهایی شبکه عصبی خواهد داشت، اما بریدن و

هرس کردن همه آنها احتمالاً در خروجی تأثیر قابل توجهی خواهد گذاشت و این مطلوب نیست. بنابراین، در اجرای روش‌های هرس، تعیین اینکه کدام اتصالات حذف شوند و کدام‌ها باقی بمانند بسیار اهمیت دارد. بدین ترتیب، محققان ام‌آی‌تی شروع به هرس کردن لایه‌هایی از شبکه کردند که بیشترین مصرف توان را داشتند و به همین دلیل آنها روش خود را Energy-Aware Pruning نامیدند. در سال‌های اخیر، استفاده از یادگیری عمیق و شبکه‌های عصبی محبوبیت زیادی پیدا کرده و در حال حاضر تلاش‌ها در این حوزه بر توسعه معماری‌های بهینه‌ای متمرکز شده است تا امکان استفاده از شبکه‌های عصبی بر روی سکوهایی که قدرت محدود محاسباتی دارند فراهم شود. هرچند بیشتر بر روی کوچک‌سازی معماری شبکه عصبی یا کاهش توان محاسباتی مورد نیاز کار شده است، اما برای پیاده‌سازی چنین شبکه‌هایی روی ابزارهای محدودی نظیر تلفن‌های همراه، به دلیل محدودیت باتری و لزوم کاهش حرارت تولید شده از دستگاه، توجه به مصرف انرژی بیشترین اهمیت را پیدا می‌کند.

تاریخ انتشار:

02 شهریور 1396

نشانی منبع:

<https://www.shabakeh-mag.com/information-feature/9284/%D9%86%D9%82%D8%B4-%D8%B4%D8%A8%DA%A9%D9%87%E2%80%8C%D9%87%D8%A7%DB%8C-%D8%B9%D8%B5%D8%A8%DB%8C-%DA%A9%D9%85%E2%80%8C%D9%85%D8%B5%D8%B1%D9%81-%D8%AF%D8%B1-%D8%A8%D9%87%D8%A8%D9%88%D8%AF-%D8%B9%D9%85%D9%84%DA%A9%D8%B1%D8%AF-%D9%87%D9%88%D8%B4-%D9%85%D8%B5%D9%86%D9%88%D8%B9%DB%8C>