

## مدل های یادگیری ماشین کم مصرف (هوش مصنوعی سبز)

خانم آر. کارتیکا<sup>۱</sup>, جی. اس. سریجا<sup>۲</sup>, ک. جاگاتیسواران<sup>۳</sup><sup>۱</sup>استادیار، گروه علوم کامپیوتر، کالج هنر و علوم سری کریشنا، کوییمباتور، هند<sup>۲</sup>دانشجوی کارشناسی علوم کامپیوتر، کالج هنر و علوم سری کریشنا، کوییمباتور، هند

## چکیده

پیشرفت سریع هوش مصنوعی (AI) موجب تحولات انقلابی در حوزه هایی مانند سلامت، مالی، آموزش و سامانه های خودران شده است. با این حال، این مزایا هزینه های زیست محیطی قابل توجهی دارند، چرا که آموزش و استقرار مدل های یادگیری ماشین در مقیاس بزرگ نیازمند توان محاسباتی عظیم بوده و منجر به مصرف بالای انرژی و انتشار کربن می شود. این چالش منجر به ظهور مفهوم «هوش مصنوعی سبز» شده است که بر طراحی مدل هایی تأکید دارد که علاوه بر دقت بالا، مصرف انرژی و اثرات زیست محیطی را نیز کاهش دهد. این مقاله به بررسی راهکارهای مختلف برای دستیابی به بهره وری انرژی در یادگیری ماشین می پردازد. تکنیک هایی مانند هرس مدل، کوانتیزاسیون و تقطیر دانش برای کاهش پیچیدگی مدل بدون افت محسوس دقت بررسی شده اند. همچنین روش هایی مانند جستجوی معماری عصبی (NAS) و آموزش با دقت ترکیبی به عنوان ابزارهای نویدبخش برای ایجاد تعادل میان عملکرد و پایداری معرفی می شوند. بهینه سازی سخت افزار با استفاده از شتاب دهنده های کم مصرف، دستگاه های هوش مصنوعی لبه ای و مراکز داده با انرژی تجدید پذیر نیز مورد توجه قرار گرفته است. مطالعات تطبیقی و موردنی در حوزه های طبقه بندی تصویر و پردازش زبان طبیعی نشان می دهد که مدل های بهینه شده از نظر انرژی می توانند تا ۵٪ درصد کاهش مصرف انرژی را بدون افت محسوس دقت تجربه کنند. یافته ها تأکید دارند که اتخاذ رویکردهای هوش مصنوعی سبز هم از نظر فنی امکان پذیر و هم از نظر زیست محیطی ضروری است. در نهایت، این مقاله هوش مصنوعی سبز را به عنوان مسیری برای نوآوری پایدار و مسئولانه معرفی می کند که پیشرفت فناوری را با اهداف جهانی اقلیمی همسو می سازد. واژگان کلیدی: هوش مصنوعی سبز، بهره وری انرژی، یادگیری ماشین، فشرده سازی مدل، محاسبات پایدار، رد پای کربنی، تقطیر دانش، کوانتیزاسیون، هرس، جستجوی معماری عصبی (NAS)، هوش مصنوعی لبه ای، محاسبات سازگار با محیط زیست.

## 1. مقدمه

هوش مصنوعی به یکی از محركهای اصلی تحول دیجیتال تبدیل شده است و امكان خودکارسازی، شخصیسازی و تصمیمگیری را در مقیاس‌های بی‌سابقه فراهم کرده است. مدل‌هایی مانند GPT، BERT و ترانسفورمرهای تصویری بزرگ عملکردی فراتر از انسان در درک زبان طبیعی و بینایی ماشین ارائه می‌دهند. با این حال، این پیشرفت‌ها چالش‌های پایداری قابل توجهی ایجاد می‌کنند، بهویژه با گسترش جهانی استفاده از هوش مصنوعی. دولتها، صنایع و جوامع علمی به طور فزاینده‌ای بر لزوم اتخاذ شیوه‌های مستو‌لانه در توسعه هوش مصنوعی تأکید دارند که آسیب‌های زیست‌محیطی را به حداقل برساند. «هوش مصنوعی سبز» با تمرکز بر نوآوری‌های مبتنی بر بهره‌وری، به این نیاز پاسخ می‌دهد. برخلاف «هوش مصنوعی قرمز» که صرفاً به دنبال دقت بیشتر با هر هزینه‌ای است، هوش مصنوعی سبز توسعه مدل‌هایی را تشویق می‌کند که با تقاضای محاسباتی کمتر، عملکرد بهینه ارائه دهند. این تغییر رویکرد برای همسویی هوش مصنوعی با اهداف اقلیمی جهانی در عین حفظ ارزش عملی آن، حیاتی است.

## 2. روش‌شناسی

برای تحلیل و پیشنهاد راهکارهای هوش مصنوعی کم‌صرف، این پژوهش از یک روش سه‌محوره استفاده کرده است:

### ۲/۱ بهینه‌سازی الگوریتمی

- فشرده‌سازی مدل: هرس وزن‌های زائد، کوانتیزاسیون به دقت پایین و تجزیه تتسورها
- تقطیر دانش: آموزش مدل‌های کوچک‌تر با استفاده از مدل‌های بزرگ‌تر
- استراتژی‌های آموزش بهینه: توقف زودهنگام، نرخ یادگیری تطبیقی، آموزش با دقت ترکیبی

### ۲/۲ بهینه‌سازی سخت‌افزاری

- استفاده از شتاب‌دهنده‌های کم‌صرف مانند FPGA، NVIDIA Jetson، Google TPU Edge
- معماری‌های تخصصی مانند Flash Attention برای کاهش بار حافظه و محاسبات

### ۲/۳ پایش و سنجش انرژی

- ابزارهایی مانند Experiment Impact Tracker و Code Carbon برای اندازه‌گیری ردپای کربنی
- تحلیل تطبیقی روی مجموعه‌داده‌های GLUE، ImageNet، CIFAR-10

## 3. مدل‌سازی و تحلیل

مرحله مدل‌سازی و تجزیه و تحلیل این تحقیق بر ارزیابی چگونگی تأثیر تکنیک‌های مختلف هوش مصنوعی سبز بر مواد نه بین دقت، بهره وری انرژی و ردپای کربن متتمرکز است. یک رویکرد چند لایه برای بررسی بهینه‌سازی معماري مدل، پیچیدگی مجموعه داده و استراتژی‌های استقرار سخت افزار اتخاذ شد.

### a. مدل‌های پایه در مقابل مدل‌های بهینه شده

مدل‌های پایه مانند ResNet-50 (طبقه‌بندی تصویر)، BERT (پردازش زبان طبیعی) و معماري‌های مبتنی بر ترانسفورماتور با استفاده از روش‌های مرسوم با دقت استاندارد ممیز شناور 32 بیتی آموزش داده شدند. سیس نسخه‌های بهینه شده این مدل‌ها با استفاده از تکنیک‌های هرس، کوانتیزاسیون و تقطیر داشتند.

#### 1. مدل‌های پایه: برای حداقل دقت و بدون ملاحظات انرژی آموزش دیده اند.

#### 2. مدل‌های بهینه: کاهش پارامترها و طراحی‌های آگاه از انرژی، دستیابی به تعادل بین دقت و کارایی.

### b. مطالعه موردی: طبقه‌بندی تصویر

#### 1. ResNet-50 هرس شده: دقت - 74٪، انرژی آموزشی - ~ 95 کیلووات ساعت (کاهش انرژی ~ 52٪).

#### 2. ResNet-50 کوانتیزه شده (8 بیتی): دقت - 73.5٪، سرعت استنتاج در دستگاه‌های لبه 3 × بهبود یافته است.

این نشان می‌دهد که هرس و کوانتیزاسیون به طور قابل توجهی انرژی مورد نیاز را با حداقل افت عملکرد کاهش می‌دهد.

### c. مطالعه موردی: پردازش زبان طبیعی

#### 1. DistilBERT (پارامترهای 66M): دقت - 82٪، انرژی - ~ 160 کیلووات ساعت (کاهش انتشار کربن ~ 65٪).

#### 2. TinyBERT (پارامترهای 14M): دقت - 79٪، انرژی - ~ 70 کیلووات ساعت (مناسب برای استقرار موبایل).

این نشان می‌دهد که تقطیر داشن یک رویکرد موثر برای ایجاد مدل‌های NLP کوچک‌تر و در عین حال قدرتمند است.

### d. مطالعه موردی: استقرار لبه

دستگاه های تلفن همراه و سیستم های اینترنت اشیا اغلب با محدودیت های شدید انرژی مواجه هستند. **MobileNetV3** در جتسون ناتو (صرف برق 4 وات): 90 درصد دقت ResNet-50 را با  $10 \times$  انرژی کمتر به دست آورد.

**Net-Lite**: برای پردازنده های تلفن همراه بهینه شده است و به نتایج پیشرفته با تأخیر کمتر دست می یابد. این نتایج تأیید می کند که روش های هوش مصنوعی سیز برای کاربردهای هوش مصنوعی لبه مانند هوایپیماهای بدون سرنشین خودران، دستگاه های مراقبت های بهداشتی هوشمند و سیستم های نظارتی بلاذرنگ بسیار مهم هستند.

#### 4. نتایج و بحث

##### 4/1 صرفهジョビイ انرژی

- صرفهجویی ۴۰٪-۶۰٪ با فشردهسازی
- کاهش ۷۰٪ اندازه مدل با نقطیز دانش
- افزایش ۳-۵ برابر سرعت استنتاج با کوانتنیزاسیون

##### 4/2 تأثیرات زیستمحیطی

- کاهش ۵۰٪ انتشار کربن
- مراکز داده با انرژی تجدیدپذیر توسط

##### 4/3 کاربردهای صنعتی

- سلامت: تشخیص پزشکی با مدل های کمصرف
- مالی: تشخیص تقلب بلاذرنگ
- IoT و لبه: دستیار های هوشمند، پهپادهای خودران

##### 4/4 بحث

- برخی روش های فشردهسازی ممکن است دقت را بیش از حد کاهش دهند
- نیاز به معیار های استاندارد برای نسبت انرژی به دقت

**Figure 1: Comparative Analysis Table**

| Model Type         | Accuracy (%) | Energy (kWh) | Carbon Reduction (%) | Deployment Suitability |
|--------------------|--------------|--------------|----------------------|------------------------|
| ResNet-50 Baseline | 76           | 200          | —                    | Data Centers           |
| Pruned ResNet-50   | 74           | 95           | 52                   | Edge Devices / Cloud   |
| BERT Baseline      | 84           | 500          | —                    | Cloud NLP Systems      |
| DistilBERT         | 82           | 160          | 65                   | Mobile Apps / Edge NLP |
| MobileNetV3        | 71           | 20           | 90                   | IoT & Smart Devices    |

## 5. نتیجه‌گیری

- این مطالعه نشان می‌دهد که هوش مصنوعی سبز برای رشد پایدار دیجیتال ضروری است. با بهره‌گیری از تکنیک‌هایی مانند هرس، کوانتیزاسیون و تنظیر داشت، سامانه‌های هوش مصنوعی می‌توانند ریاضی محاسباتی و زیست‌محیطی خود را کاهش دهن، در حالی که عملکرد رقابتی خود را حفظ می‌کنند. راهبردهای آگاه از سخت‌افزار و استفاده از انرژی‌های تجدیدپذیر نیز پایداری را بیشتر تقویت می‌کنند.
- این پژوهش تأکید می‌کند که هوش مصنوعی سبز صرفاً یک تنظیم فنی نیست، بلکه یک تغییر پارادایم بهسیوی نوآوری مسئولانه است. مسیرهای آینده شامل توسعه جستجوی معماری عصبی آگاه از محیط زیست (eco-aware NAS)، بهبود چارچوب‌های سنجش انرژی و ترویج سیاست‌هایی است که توسعه هوش مصنوعی با آگاهی از کربن را تشویق می‌کنند. با ادامه رشد استفاده از هوش مصنوعی، گنجاندن اصول پایداری در آن کلید تضمین همسوی فناوری با تعهدات اقلیمی جهانی خواهد بود.

## 6. منابع

- .1 O. Etzioni, R. Schwartz, J. Dodge, N. A. Smith, «هوش مصنوعی سبز»، Communications of the ACM، جلد 63، شماره 12، صفحات 54–53، ۲۰۲۰.
- .2 A. McCallum, A. Ganesh, E. Strubell، «ملاحظات انرژی و سیاست در یادگیری عمیق برای NLP»، مجموعه مقالات نشست سالانه انجمن زبان‌شناسی محاسباتی، صفحات ۴۵۶۳–۴۳۶۳، ۲۰۱۹.
- .3 D. Patterson و همکاران، «انتشار کربن و آموزش شبکه‌های عصبی بزرگ»، arXiv:2104.10350، ۲۰۲۱.
- .4 H. Li, J. Liu, Z. Lin, Y. Xu، «جستجوی معماری عصبی آگاه از انرژی برای یادگیری عمیق کارآمد»، IEEE/CVF کنفرانس، صفحات ۴۹۲۰–۴۹۲۰، ۲۰۲۱.
- .5 H. Grahn, G. Riley, C. Rodrigues, N. Garcia-Martin ماشین»، Journal of Parallel and Distributed Computing، جلد ۱۳۴، صفحات ۷۵–۸۸، ۲۰۱۹.
- .6 J. Zou, M. Abid, A. Ghorbani، «تفسیر شبکه‌های عصبی شکننده است»، کنفرانس AAAI در هوش مصنوعی، ۲۰۱۹.
- .7 V. J. Reddi و همکاران، «MLPerf: مجموعه معیار برای عملکرد یادگیری ماشین»، arXiv:1910.01500، ۲۰۱۹.
- .8 A. Green, C. L. Hutchison, J. K. Ward، «گزارش پایداری گوگل، «مراکز داده کربن‌خنثی و آموزش هوش مصنوعی»، ۲۰۲۲.
- .9 E. Choi, R. Hu, M. Henderson، «یادگیری ماشین پایدار: کاهش ریاضی کربنی هوش مصنوعی»، IEEE Access، جلد 10، صفحات ۶۵۱۲–۱۵۶۱، ۲۰۲۲.
- .10 E. Choi, R. Hu, M. Henderson، «بهسیوی هوش مصنوعی آگاه از کربن: سنجش و کاهش اثر انرژی یادگیری ماشین»، کنفرانس AAAI در هوش مصنوعی، ۲۰۲۱.