Supreme National Defense University Defense Logistics and Technology

https://amfad.sndu.ac.ir E-ISSN: 2980-8073

Vol. 9, No. 3, Issue. 31, Autumn 2026 **Received:** 2026/02/10 **Accepted:** 2026/06/02

Research paper

The Role of Green Artificial Intelligence in Enhancing Energy Efficiency and Reducing Energy Consumption in Defense Applications

Amirhesam Afsharkohan
Department of Electrical and Computer Engineering, Ferdows Branch, Islamic Azad University, Ferdows, Iran.
Hamidreza Ghaffari
Department of Electrical and Computer Engineering, Ferdows Branch, Islamic Azad University, Ferdows, Iran (Corresponding Author).
Email: hamidghaffary53@iau.ac.ir

Abstract

This quantitative experimental study aims to analyze energy consumption, carbon dioxide emissions, and model performance under different optimization scenarios. Data were obtained from four standard datasets, namely Iris, Digits, Wine, and CIFAR-10, and the experiments were conducted in a laboratory computational environment. Accordingly, the spatial scope of the study was defined by the computational environment, while the temporal scope corresponded to the experimental period. The analytical procedure involved evaluating the models under different preprocessing scenarios and comparing accuracy, precision, recall, F1-score, energy consumption, and carbon emissions. The increasing energy consumption and carbon dioxide emissions associated with the development of machine learning models and deep neural networks underscore the need for Green Artificial Intelligence approaches. This study evaluates energy consumption, carbon dioxide emissions, and model performance across four scenarios: baseline, duplicate-data removal, principal component analysis (PCA), and 50% sampling. For the Iris and Wine datasets, accuracy, precision, recall, and F1-score remained equal or close to 1 across all scenarios, while energy consumption decreased from 0.000005 kWh to 0.000003 kWh. For the Digits dataset, accuracy ranged from 0.96 to 0.97, while sampling reduced energy consumption to 0.000007 kWh. In the second part of the study, a convolutional neural network (CNN) was applied to the CIFAR-10 dataset and achieved an accuracy of 0.56; however, its energy consumption and carbon emissions were substantially higher than those observed for the smaller datasets. The results indicate that optimization scenarios applied to small datasets can reduce energy consumption without compromising model accuracy, whereas larger datasets require more advanced optimization approaches.

Keywords: Artificial Intelligence; Energy Consumption; Carbon Dioxide Emissions; Machine Learning; Energy Efficiency.

سال نهم، شماره سوم (پیاپی ۳۱)، پاییز ۱۴۰۵، صص. ۲۰۷-۲۳۸
تاریخ دریافت: ۱۴۰۴/۱۱/۲۱ ـ تاریخ پذیرش: ۱۴۰۵/۰۳/۱۲

مقاله پژوهشی

نقش هوش مصنوعی سبز در ارتقای بهره‌وری و کاهش هزینه‌های انرژی دفاعی

امیرحسام افشارکهن
گروه مهندسی برق و کامپیوتر، واحد فردوس، دانشگاه آزاد اسلامی، فردوس، ایران
حمیدرضا غفاری
گروه مهندسی برق و کامپیوتر، واحد فردوس، دانشگاه آزاد اسلامی، فردوس، ایران (نویسنده مسئول)
Email: hamidghaffary53@iau.ac.ir

چکیده

این پژوهش از نوع کمی-تجربی است و با هدف تحلیل مصرف انرژی، انتشار دی‌اکسید کربن و عملکرد مدل‌ها در سناریوهای مختلف بهینه‌سازی انجام شد. داده‌ها از چهار مجموعه داده استاندارد Iris، Digits، Wine و CIFAR-10 گردآوری و آزمایش‌ها در محیط محاسباتی آزمایشگاهی اجرا گردید؛ بنابراین قلمرو مکانی بستر محاسباتی و قلمرو زمانی دوره اجرای آزمایش‌ها است. روش تحلیل شامل اجرای مدل‌ها تحت سناریوهای مختلف پیش‌پردازش و مقایسه شاخص‌های دقت، صحت، نرخ بازخوانی، امتیاز F1، مصرف انرژی و انتشار کربن است. افزایش مصرف انرژی و انتشار دی‌اکسیدکربن ناشی از توسعه مدل‌های یادگیری ماشین و شبکه‌های عصبی عمیق، ضرورت توجه به رویکردهای هوش مصنوعی سبز را برجسته کرده است. این پژوهش مصرف انرژی، انتشار دی‌اکسیدکربن و عملکرد مدل‌ها را در چهار سناریوی پایه، حذف داده‌های تکراری، PCA و نمونه‌برداری ۵۰ درصدی بررسی کرده است. در مجموعه‌های Iris و Wine، نرخ صحت، دقت، نرخ بازخوانی و امتیاز F1 در تمام سناریوها برابر یا نزدیک به ۱ باقی ماند و مصرف انرژی از kWh۰/۰۰۰۰۰۵ به kWh۰/۰۰۰۰۰۳ کاهش یافت. در مجموعه Digits نیز دقت بین ۰/۹۶ تا ۰/۹۷ متغیر بود و نمونه‌برداری مصرف انرژی را به kWh۰/۰۰۰۰۰۷ کاهش داد. در بخش دوم، شبکه عصبی پیچشی روی CIFAR-10 با دقت ۰/۵۶ اجرا شد؛ اما مصرف انرژی و انتشار کربن آن به‌طور قابل توجهی بالاتر از داده‌های کوچک بود. نتایج نشان می‌دهد سناریوهای بهینه‌سازی در داده‌های کوچک موجب کاهش انرژی بدون افت دقت می‌شوند؛ اما در داده‌های بزرگ نیاز به روش‌های پیشرفته‌تر وجود دارد.

کلیدواژه‌ها: هوش مصنوعی، مصرف انرژی، دی‌اکسید کربن، یادگیری ماشین، بهره‌وری

20.1001.1.28212606.1405.9.3.6.2 https://amfad.sndu.ac.ir/ E-ISSN: 2980-8073
صحت مطالب بر عهده نویسنده مقاله است و بیانگر دیدگاه دانشگاه عالی دفاع ملی نیست.

مقدمه

در دهه‌های اخیر، رشد شتابان فناوری‌های دیجیتال و به‌ویژه هوش مصنوعی، نقش تعیین‌کننده‌ای در تحولات علمی، صنعتی و دفاعی ایفا کرده است (گیدلاند، ۲۰۲۵). هوش مصنوعی اکنون دیگر یک ابزار صرفاً پژوهشی یا دانشگاهی نیست، بلکه به‌عنوان یک توانمندساز راهبردی در بسیاری از حوزه‌های حیاتی، از مدیریت زنجیره تأمین گرفته تا فرماندهی و کنترل، از امنیت سایبری تا مدل‌سازی سناریوهای دفاعی و حتی در بهداشت و سلامت، مورد استفاده قرار می‌گیرد (بورراچیا و همکاران، ۲۰۲۵). این گستره وسیع کاربردها سبب شده است که توسعه و استقرار سامانه‌های هوشمند به‌عنوان یک اولویت راهبردی برای کشورها و به‌ویژه نیروهای مسلح تعریف شود. بااین‌حال، این رشد پُرشتاب چالشی بزرگ مصرف بالای انرژی و هزینه‌های زیست‌محیطی و اقتصادی مرتبط با آموزش و اجرای مدل‌های هوش مصنوعی را نیز با خود همراه کرده است (گیدلاند، ۲۰۲۵). گزارش‌های متعددی در سطح جهانی نشان داده‌اند که آموزش مدل‌های بزرگ «یادگیری عمیق»[1]، گاه مصرف انرژی معادل صدها خانوار در یک سال را به‌همراه دارد و میزان انتشار «دی‌اکسید کربن»[2] آن‌ها با صنایع سنگین قابل مقایسه است. چنین وضعیتی نه‌تنها از منظر زیست‌محیطی نگران‌کننده است، بلکه در محیط‌های نظامی و دفاعی که محدودیت منابع و ضرورت بهره‌وری حداکثری حرف اول را می‌زند، یک تهدید راهبردی به شمار می‌رود (جگادسوری و راتهی‌پریا، ۲۰۲۵)؛ بنابراین بررسی و طراحی الگوهای مصرف بهینه انرژی در سامانه‌های هوش مصنوعی به‌ویژه در زمینه‌های مرتبط با دفاع، امنیت و صنایع راهبردی، ضرورتی اجتناب‌ناپذیر است.

با وجود پیشرفت‌های چشمگیر در کارایی الگوریتم‌ها، همچنان مشکل اصلی در بسیاری از مدل‌های هوش مصنوعی نیاز به محاسبات حجیم و پُرهزینه است (پرآنتو و همکاران، ۲۰۲۵). از یک‌سو افزایش «دقت»[3] و «صحت»[4] مدل‌ها معمولاً با پیچیدگی بیشتر و مصرف

1. Deep Learning
2. Carbon Dioxide (CO_2)
3. Precision
4. Accuracy

انرژی بالاتر همراه است؛ از سوی دیگر، بسیاری از کاربردهای دفاعی نیازمند سامانه‌هایی هستند که نه‌تنها دقیق و سریع باشند، بلکه در محیط‌هایی با منابع محدود مانند تجهیزات میدانی یا سامانه‌های حمل‌ونقل نظامی قابل اجرا باشند. این تضاد میان نیاز به‌دقت بالا و ضرورت کاهش مصرف انرژی یک خلأ پژوهشی جدی ایجاد کرده است (گیدلاند، ۲۰۲۵). در ادبیات جهانی، گرچه از مفاهیمی همچون «هوش مصنوعی سبز»[1] و محاسبات پایدار سخن به میان آمده است؛ اما کم‌تر پژوهشی به‌صورت کاربردی و عددی به مقایسه روش‌های مختلف «داده‌کاوی»[2] و «یادگیری ماشین»[3] از منظر هم‌زمان کارایی و مصرف انرژی پرداخته است. به‌ویژه در کشور ما که با محدودیت‌های منابع انرژی و ضرورت ارتقای بهره‌وری دفاعی مواجه است، فقدان چنین پژوهش‌هایی کاملاً محسوس است. ازاین‌رو، این مقاله می‌کوشد با طراحی و اجرای یک چهارچوب آزمایشی مشخص، به این پرسش پاسخ دهد که چه روش‌ها و سناریوهایی در فرایند آموزش مدل‌های هوش مصنوعی می‌توانند ضمن حفظ دقت بالا، مصرف انرژی و درنتیجه الگوی مصرف دفاعی را بهینه کنند.

مطالعات پیشین نشان داده‌اند که استفاده از روش‌های ساده‌تر یا داده‌های کاهش‌یافته می‌تواند مصرف انرژی را تا حد قابل توجهی کاهش دهد، اما این کاهش معمولاً با افت دقت همراه است. برخی پژوهشگران پیشنهاد داده‌اند که با استفاده از روش‌هایی مانند فشرده‌سازی مدل‌ها، انتخاب زیرمجموعه داده‌ها و یا بهره‌گیری از تکنیک‌های کاهش بُعد مانند تحلیل مؤلفه‌های اصلی، می‌توان تعادلی میان دقت و بهره‌وری انرژی برقرار کرد. همچنین ابزارهایی مانند «کُدکربن»[4] به‌عنوان یک راهکار نوین برای سنجش مستقیم مصرف انرژی و برآورد انتشار کربن در فرایندهای محاسباتی معرفی شده‌اند (بورراچیا و همکاران، ۲۰۲۵). این ابزارها امکان می‌دهند پژوهشگران نه‌تنها خروجی الگوریتم‌ها از نظر دقت، صحت، بلکه اثرات زیست‌محیطی و انرژی‌بر آن‌ها را نیز بررسی کنند. با این وجود، بخش عمده‌ای از این پژوهش‌ها در محیط‌های عمومی و بدون در نظر گرفتن نیازهای راهبردی دفاعی انجام شده

1. Green Artificial Intelligence (Green AI)
2. Data Mining
3. Machine Learning
4. CodeCarbon

است. با توجه به شرایط کنونی کشور و الزامات دفاعی، پرداختن به موضوع ارتقای بهره‌وری انرژی در هوش مصنوعی از دو جنبه حائز اهمیت است. نخست، جنبه راهبردی آن است که نیروهای مسلح نیازمند الگوریتم‌ها و سامانه‌هایی هستند که ضمن ارائه دقت و صحت بالا، با کم‌ترین مصرف منابع عمل کنند تا قابلیت استقرار در محیط‌های عملیاتی را داشته باشند. دوم، جنبه زیست‌محیطی و اقتصادی است؛ کاهش مصرف انرژی به معنای کاهش هزینه‌ها و همچنین کاهش اثرات منفی زیست‌محیطی است که با سیاست‌های کلان کشور در حوزه انرژی و محیط‌زیست هم‌راستا است. پژوهش حاضر با بهره‌گیری از مجموعه داده‌های کلاسیک مانند «Iris»[۱]، «Digits»[۲]، «Wine»[۳] و یک مجموعه داده پیچیده‌تر مانند «CIFAR-10»[۴]، تلاش کرده است با مقایسه سناریوهای مختلف شامل «پایه»[۵]، «پاک‌سازی داده‌ها»[۶]، «تحلیل مؤلفه‌های اصلی»[۷] و «نمونه‌برداری»[۸] تصویری روشن از رابطه میان کارایی الگوریتم‌ها و مصرف انرژی ارائه دهد.

اهداف اصلی این پژوهش عبارت‌اند از: اندازه‌گیری و مقایسه هم‌زمان کارایی و مصرف انرژی در چندین سناریوی آموزشی، ارائه چهارچوبی عملیاتی برای سنجش الگوی مصرف دفاعی در حوزه هوش مصنوعی. نوآوری این کار در آن است که به‌جای تمرکز صرف بر معیارهای سنتی مانند نرخ صحت، نرخ دقت، نرخ «بازخوانی»[۹] و امتیاز «F1»[۱۰]، بُعد انرژی و انتشار کربن نیز به‌عنوان یک شاخص اصلی وارد تحلیل شده است. این نگاه چندبُعدی می‌تواند الگوی جدیدی در پژوهش‌های دفاعی و فناورانه کشور ایجاد کند.

چهارچوب پیشنهادی این پژوهش بر سه گام اصلی استوار است. نخست، آماده‌سازی

1. Iris dataset for machine learning, available at https://archive.ics.uci.edu/dataset/53/iris
2. Digits dataset for machine learning, available at https://www.kaggle.com/datasets/karnikakapoor/digits
3. Wine dataset for machine learning, available at https://archive.ics.uci.edu/dataset/109/wine
4. CIFAR-10 dataset for machine learning, available at https://keras.io/api/datasets/cifar10/
5. Baseline
6. Cleaning
7. Principal Component Analysis (PCA)
8. Sampling
9. Recall
10. F1 Score

داده‌ها و تعریف سناریوها که برای هر داده، چهار سناریو شامل پایه، پاک‌سازی داده‌ها، تحلیل مؤلفه‌های اصلی و نمونه‌برداری طراحی و اجرا شده است. دوم، آموزش مدل‌ها و سنجش کارایی که با استفاده از الگوریتم‌های متداول مانند «جنگل تصادفی»[1] و «شبکه عصبی پیچشی»[2]، معیارهای صحت، دقت، بازخوانی و F1 محاسبه شده‌اند (تقندیکی و همکاران، ۱۴۰۴) (اله دادی و همکاران، ۱۴۰۴). سوم، اندازه‌گیری انرژی و انتشار کربن که با استفاده از کتابخانه کُدکربن، مصرف انرژی در هر سناریو ثبت و انتشار «دی‌اکسید کربن»[3] براساس ضریب انتشار ایران برآورد شده است. این فرایند امکان مقایسه دقیق میان کارایی و بهره‌وری انرژی را فراهم کرده است.

درمجموع، این پژوهش در پی پاسخ به یک نیاز راهبردی و فوری است: چگونه می‌توان بهره‌وری و الگوی مصرف دفاعی در حوزه هوش مصنوعی را ارتقا داد، به‌گونه‌ای که دقت و قابلیت‌های عملیاتی حفظ شوند؛ اما مصرف انرژی و هزینه‌های جانبی به حداقل برسند. نتایج اولیه نشان می‌دهند که برخی روش‌ها مانند پاک‌سازی و نمونه‌برداری می‌توانند بدون افت محسوس دقت، به کاهش مصرف انرژی کمک کنند، درحالی‌که روش‌هایی مانند تحلیل مؤلفه‌های اصلی اگرچه بُعد داده‌ها را کاهش می‌دهند، ممکن است منجر به افت نسبی دقت شوند. تحلیل این تعادل‌ها می‌تواند به طراحی الگوریتم‌های هوش مصنوعی کاراتر و سازگارتر با نیازهای دفاعی کمک کند.

۱. مبانی نظری

این بخش تلاش دارد تا ضمن معرفی مبانی اصلی مرتبط با بهره‌وری انرژی در هوش مصنوعی و یادگیری ماشین، پیوند میان مفاهیم کلیدی همچون کارایی الگوریتم‌ها، انتخاب مجموعه داده‌ها، ابزارها و کتابخانه‌های نرم‌افزاری و در نهایت مرور مطالعات گذشته را روشن سازد.

1. Random Forest
2. Convolutional Neural Network (CNN)
3. Carbon Dioxide (CO_2)

۱-۱. کتابخانه کُدکربن

کتابخانه کُدکربن به‌عنوان ابزاری نوین برای اندازه‌گیری و پایش مصرف انرژی در اجرای مدل‌های هوش مصنوعی استفاده شده است. این ابزار می‌تواند میزان انرژی مصرف‌شده توسط سخت‌افزار (CPU و GPU) و حافظه را در طول فرایند آموزش ثبت کند و آن را براساس ضریب انتشار انرژی الکتریکی به میزان انتشار دی‌اکسیدکربن تبدیل نماید (گیدلاند، ۲۰۲۵). یکی از قابلیت‌های ارزشمند کُدکربن، امکان تنظیم ضریب انتشار متناسب با موقعیت جغرافیایی است (ورانا و موندال، ۲۰۲۵). به همین دلیل در این پژوهش مقدار ۰/۵۷ به‌عنوان ضریب انتشار ایران لحاظ شد تا نتایج به شرایط واقعی کشور نزدیک‌تر باشد.

۱-۲. Scikit-learn

کتابخانه Scikit-learn ازجمله پراستفاده‌ترین کتابخانه‌های متن‌باز در زمینهٔ یادگیری ماشین است. این کتابخانه شامل الگوریتم‌های گوناگون دسته‌بندی، رگرسیون، خوشه‌بندی و کاهش بُعد است (لِمنکووا، ۲۰۲۵). در پژوهش حاضر، الگوریتم جنگل تصادفی و روش تحلیل مؤلفه‌های اصلی از این کتابخانه استفاده شده‌اند. دلیل انتخاب Scikit-learn علاوه بر جامعیت الگوریتم‌ها، سهولت ادغام با ابزارهای مدیریت داده مانند Pandas و ابزارهای مصورسازی مانند Matplotlib بوده است.

۱-۳. PyTorch

چهارچوب PyTorch به‌عنوان یکی از قدرتمندترین ابزارهای پیاده‌سازی شبکه‌های عصبی عمیق انتخاب شد. این کتابخانه با پشتیبانی از محاسبات پویا و اجرای موازی روی GPU، امکان آموزش مدل‌های پیچیده را فراهم می‌سازد (چن و همکاران، ۲۰۲۵). در این پژوهش از PyTorch برای طراحی و آموزش شبکه عصبی پیچشی روی مجموعه داده CIFAR-10 استفاده شد.

۱-۴. Pandas و Matplotlib

کتابخانه Pandas برای مدیریت و سازماندِهی داده‌ها و Matplotlib برای مصورسازی نتایج استفاده شدند. Pandas با ساختار DataFrame خود امکان پردازش داده‌ها به شکل جدول‌محور را تسهیل کرد و Matplotlib نیز خروجی‌های تصویری همچون نمودارهای میله‌ای، پراکندگی و مقایسه‌ای را فراهم آورد (ساپر و وارتاک، ۲۰۲۰).

۱-۵. الگوریتم جنگل تصادفی

الگوریتم جنگل تصادفی با انتخاب تصادفی ویژگی‌ها و داده‌ها برای ساخت هر درخت، از بیش‌برازش جلوگیری می‌کند و در مقایسه با روش‌های تک‌درختی کارایی بهتری دارد (مالالا و همکاران، ۲۰۲۵). در این پژوهش، جنگل تصادفی به‌عنوان الگوریتم اصلی طبقه‌بندی داده‌های Iris، Digits و Wine انتخاب شد. این انتخاب به‌دلیل دقت بالا، سهولت پیاده‌سازی و توانایی در پردازش داده‌های کوچک تا متوسط صورت گرفت.

۱-۶. تحلیل مؤلفه‌های اصلی

روش تحلیل مؤلفه‌های اصلی یکی از پرکاربردترین روش‌ها برای کاهش بُعد داده‌ها است. در این روش، ترکیب خطی جدیدی از ویژگی‌های داده ساخته می‌شود که بیشترین واریانس را حفظ کند. کاهش بُعد نه‌تنها سرعت محاسبات را افزایش می‌دهد، بلکه می‌تواند منجر به کاهش مصرف انرژی نیز شود (وو و همکاران، ۲۰۲۵).

۱-۷. شبکه عصبی پیچشی

شبکه‌های عصبی پیچشی یکی از معماری‌های رایج در پردازش تصویر هستند که با بهره‌گیری از لایه‌های کانولوشن و pooling قادر به استخراج ویژگی‌های سلسله‌مراتبی از تصاویر هستند. معماری مورد استفاده در این پژوهش شامل دو لایه کانولوشن، یک لایه pooling و دو لایه Fully Connected بود (تقندیکی و همکاران، ۱۴۰۴).

این معماری برای مجموعه داده CIFAR-10 طراحی و پیاده‌سازی شد. انتخاب شبکه‌های عصبی پیچشی به‌دلیل توانایی بالا در پردازش تصاویر و نیز نیاز بالای محاسباتی آن صورت گرفت تا امکان ارزیابی دقیق‌تر مصرف انرژی فراهم شود.

۱-۸. مجموعه داده‌های مورد استفاده

❖ **Iris**

مجموعه داده Iris شامل ۱۵۰ نمونه گل زنبق است که با چهار ویژگی (طول و عرض کاسبرگ و گلبرگ) و سه کلاس مشخص می‌شود (آگراوال و همکاران، ۲۰۲۵).

❖ **Digits**

مجموعه داده Digits شامل تصاویر ۸×۸ پیکسلی از اعداد ۰ تا ۹ است که درمجموع ۱۷۹۷ نمونه دارد (احمد و همکاران، ۲۰۲۵).

❖ **Wine**

مجموعه داده Wine شامل ۱۷۸ نمونه از سه نوع مختلف نوشیدنی است که هر نمونه با ۱۳ ویژگی شیمیایی توصیف می‌شود. این مجموعه داده به‌دلیل تعداد ویژگی‌های بیشتر، چالش متفاوتی برای الگوریتم‌ها ایجاد می‌کند و امکان تحلیل دقیق‌تر اثر کاهش بُعد از طریق PCA را فراهم می‌آورد (نلسون و همکاران، ۲۰۲۴).

❖ **CIFAR-10**

مجموعه داده CIFAR-10 شامل ۶۰ هزار تصویر رنگی ۳۲×۳۲ در ۱۰ کلاس مختلف است. این داده به‌دلیل حجم و پیچیدگی بالا، انتخابی مناسب برای ارزیابی مدل‌های یادگیری عمیق به شمار می‌رود (گوتام و همکاران، ۲۰۲۴).

۲. پیشینهٔ پژوهش

در مطالعه (استروبِل و همکاران، ۲۰۲۰) نشان دادند که حتی آموزش مدل‌های نسبتاً ساده NLP حدود ۳۹ پوند دی‌اکسیدکربن، مدل‌های بزرگ‌تر مانند Transformer حدود ۱۹۲ پوند

دی‌اکسیدکربن و مدل‌های بسیار پیچیده مانند NAS تا ۶۲۶،۱۵۵ پوند دی‌اکسیدکربن تولید می‌کنند. این نتایج اهمیت انتخاب معماری و مقیاس مدل‌ها را در کاهش اثرات زیست‌محیطی برجسته می‌کند.

در مطالعه (پترسون و همکاران، ۲۰۲۱) با بررسی مدل‌هایی مانند T5، Meena و GPT-3 نشان دادند که انتخاب مرکز داده، نوع سخت‌افزار و معماری می‌تواند ردپای کربنی را تا ۱۰۰۰ برابر کاهش دهد. یافته‌های این پژوهش، ضرورت گزارش شفاف مصرف انرژی و دی‌اکسیدکربن در مطالعات یادگیری ماشین را برجسته کرده است.

در مطالعه (شوارتز و همکاران، ۲۰۲۰) با معرفی مفهوم هوش مصنوعی سبز به رشد نمایی محاسبات در یادگیری عمیق اشاره کردند. آن‌ها دریافتند که بین سال‌های ۲۰۱۲ تا ۲۰۱۸ محاسبات هرچند ماه دو برابر شده و این موضوع به افزایش ۳۰۰،۰۰۰ برابری توان محاسباتی منجر شده است. پیام اصلی این مطالعه، لزوم توجه به کارایی انرژی در کنار دقت مدل‌ها است.

در یک مرور سیستماتیک، در مطالعه (وردکیا و همکاران، ۲۰۲۳) با بررسی ۹۸ مطالعه اصلی در حوزه هوش مصنوعی سبز گزارش دادند که از سال ۲۰۲۰ تاکنون، این حوزه به‌طور چشمگیری رشد کرده است. نتایج نشان دادند که بسیاری از پژوهش‌ها روی تنظیم ابرپارامترها، نظارت بر ردپای کربنی و ارزیابی پایداری متمرکز بوده‌اند. درعینِ‌حال، مشارکت صنعتی در این حوزه همچنان محدود است.

از دیدگاه عملی مطالعه (اول حق و همکاران، ۲۰۲۵) نشان داد که پذیرش هوش مصنوعی سبز توسط کسب‌وکارهای کوچک و متوسط (SMEs) می‌تواند عملکرد پایدار سازمانی را تقویت کند. این تحقیق تأکید می‌کند که نوآوری زیست‌محیطی می‌تواند پلی میان فناوری‌های هوش مصنوعی و پایداری سازمانی ایجاد کند.

در حوزه طراحی مدل‌ها، در مطالعه (بولو و همکاران، ۲۰۲۴) راهکارهایی مانند انتخاب مدل‌های پویا (MMS) و خروج زودهنگام (EE) را برای کاهش مصرف انرژی معرفی کردند و نشان دادند که این روش‌ها می‌توانند ضمن کاهش مصرف انرژی، دقت پیش‌بینی‌ها را نیز تا

۵٪ بهبود دهند.

در مطالعه (شودل و همکاران، ۲۰۲۵) تمرکز خود را بر چرخه کامل عمر سیستم‌های توصیه‌گر گذاشتند و بیان کردند که مرحله استنتاج سهم بزرگی در مصرف انرژی دارد. این مطالعه بر لزوم ارزیابی تمام مراحل توسعه و اجرای مدل‌ها تأکید دارد.

از سوی دیگر، در مطالعه (کاستلانو–نیوس و گارسیا–فورت، ۲۰۲۳) در زمینهٔ AutoML سبز نشان دادند که حتی فرایندهای خودکار انتخاب مدل نیز می‌توانند با استفاده از جست‌وجوی بیزی و تصادفی بهینه‌سازی شوند تا مصرف انرژی کاهش یابد.

در حوزه پزشکی، مطالعه‌ای (اردومپوزانیس و پاپاکوستاس، ۲۰۲۴) نشان داد که فاین‌تیون مدل ViT گوگل در کاربردهای تصویربرداری پزشکی مصرف انرژی قابل‌توجهی دارد. هر آزمایش حدود ۰/۱۸ کیلووات‌ساعت انرژی و ۰/۰۶۶ کیلوگرم دی‌اکسیدکربن تولید کرده است. این نتایج اهمیت انتخاب مناسب مدل و سخت‌افزار را در کاربردهای حساس مانند پزشکی نشان می‌دهد.

در نهایت، مطالعه (فربر و لمپرت، ۲۰۲۲) نخستین آنتولوژی هوش مصنوعی سبز را معرفی کرد که معیارهایی مانند مصرف انرژی، زمان اجرا و ردپای کربنی مدل‌ها را شامل می‌شود. این چهارچوب می‌تواند به‌عنوان استانداردی برای مدیریت و سیاست‌گذاری در حوزه هوش مصنوعی سبز مورد استفاده قرار گیرد.

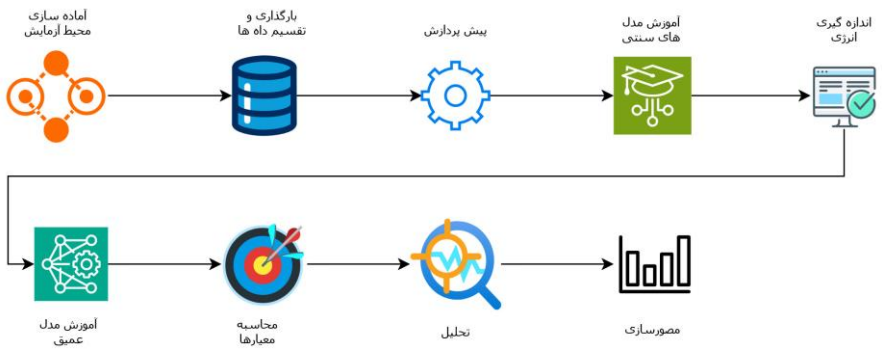
به‌طورکلی، مرور پیشینه نشان می‌دهد که پژوهش‌های متعددی در حوزه کاهش مصرف انرژی و ردپای کربنی در یادگیری ماشین انجام شده‌اند، اما همچنان چالش‌هایی مانند تعمیم نتایج به محیط‌های واقعی، مشارکت محدود صنعتی و نبود چهارچوب‌های استاندارد باقی‌مانده است. پژوهش حاضر با ترکیب ابزارهای نوین مانند کُدکربن، داده‌های متنوع و الگوریتم‌های کلاسیک و عمیق، گامی در جهت پر کردن این خلأها برمی‌دارد. جدول (۱)، مروری کلی بر مطالعات انجام شده را نشان می‌دهد.

جدول ۱: مروری بر مطالعات

ارتباط با پژوهش حاضر	یافته‌های کلیدی	موضوع–حوزه اصلی	نویسندگان (سال)
اهمیت انتخاب مدل‌های کوچک‌تر و بهینه	مدل‌های بزرگ NLP می‌توانند تا ۶۲۶،۱۵۵ پوند دی‌اکسیدکربن تولید کنند	NLP و مصرف انرژی	(استروبِل و همکاران، ۲۰۲۰)
نشان‌دهنده اهمیت زیرساخت در هوش مصنوعی سبز	انتخاب مرکز داده و سخت‌افزار می‌تواند ردپای کربنی را تا ۱۰۰۰ برابر کاهش دهد	مدل‌های بزرگ زبانی	(پترسون و همکاران، ۲۰۲۱)
مبنای نظری پژوهش حاضر	رشد ۳۰۰،۰۰۰ برابری محاسبات، لزوم توجه به کارایی انرژی	هوش مصنوعی سبز	(شوارتز و همکاران، ۲۰۲۰)
نشان‌دهنده خلأ در کاربرد عملی	صرفه‌جویی انرژی تا ۱۱۵٪ گزارش شده؛ مشارکت صنعتی کم	مرور سیستماتیک	(وردکیا و همکاران، ۲۰۲۳)
پیوند بین AI و پایداری اقتصادی	هوش مصنوعی سبز باعث بهبود عملکرد پایدار سازمانی می‌شود	SMEs و نوآوری سبز	(اول حق و همکاران، ۲۰۲۵)
اهمیت طراحی معماری کم‌مصرف	معرفی MMS و EE برای کاهش انرژی؛ افزایش ۵٪ دقت	شبکه‌های عصبی پویا	(بولو و همکاران، ۲۰۲۴)
ضرورت بررسی چرخه کامل مدل‌ها	استنتاج سهم بزرگی از مصرف انرژی دارد	سیستم‌های توصیه‌گر	(شودل و همکاران، ۲۰۲۵)
کاربرد AutoML در هوش مصنوعی سبز	جست‌وجوی بیزی و تصادفی برای کاهش مصرف انرژی	AutoML سبز	(کاستلانو–نیوس و گارسیا–فورت، ۲۰۲۳)
تأکید بر سخت‌افزار مناسب در پزشکی	هر فاین‌تیون ViT ≈ ۰/۱۸ kWh انرژی و ۰/۰۶۶ kg دی‌اکسیدکربن	تصویربرداری پزشکی	(اردومپوزانیس و پاپاکوستاس، ۲۰۲۴)
ابزار سیاست‌گذاری و مقایسه مدل‌ها	ارائه چهارچوب استاندارد برای انرژی و دی‌اکسیدکربن	آنتولوژی هوش مصنوعی سبز	(فربر و لمپرت، ۲۰۲۲)

۳. روش‌شناسی پژوهش

روش‌شناسی این پژوهش با هدف ارائه یک رویکرد جامع برای ارزیابی هم‌زمان کارایی الگوریتم‌های یادگیری ماشین و میزان مصرف انرژی و انتشار دی‌اکسیدکربن طراحی شده است. در این بخش تلاش می‌شود تا ضمن تشریح مراحل گام‌به‌گام، مشخص شود که چگونه داده‌ها، الگوریتم‌ها و ابزارهای اندازه‌گیری در کنار یکدیگر به‌کار گرفته شده‌اند تا تصویری دقیق از ارتباط میان دقت مدل‌ها و بهره‌وری انرژی به دست آید. مراحل پیاده‌سازی رویکرد پیشنهادی، از ۹ مرحله تشکیل شده است که در ادامه هر یک از این مراحل به‌صورت جداگانه نام برده و توضیح داده شده‌اند (شکل ۱).



شکل ۱: مراحل پیاده‌سازی رویکرد پیشنهادی

۳-۱. آماده‌سازی محیط آزمایش

اولین گام در این پژوهش آماده‌سازی محیط آزمایش بود تا اطمینان حاصل شود تمامی آزمایش‌ها در شرایط کنترل‌شده و قابل تکرار انجام می‌شوند. در این مرحله، کتابخانه‌های مورد نیاز شامل scikit-learn، PyTorch، کُدکربن، NumPy، Pandas و Matplotlib نصب و پیکربندی شدند. یکی از مزایای مهم این مرحله، ایجاد بستری پایدار برای اجرای الگوریتم‌ها و ثبت نتایج است، به‌گونه‌ای که خطاهای ناشی از ناسازگاری نسخه‌ها یا عدم تطابق محیطی به حداقل برسد. تعیین سخت‌افزار مورد استفاده (CPU یا GPU) و تنظیمات مربوط به آن اهمیت بالایی داشت، زیرا این عوامل مستقیماً بر مصرف انرژی و زمان اجرا اثر می‌گذارند.

۳-۲. «بارگذاری و تقسیم داده‌ها»[1]

پس از آماده‌سازی محیط، داده‌های پژوهشی بارگذاری شدند. در این تحقیق از سه مجموعه داده کلاسیک Iris، Digits و Wine به‌عنوان نمونه‌های کوچک و از مجموعه داده CIFAR-10 به‌عنوان یک داده پیچیده تصویری استفاده شد. استفاده از این تنوع داده‌ای باعث شد که نتایج تحقیق قابلیت تعمیم بالاتری داشته باشند، چراکه عملکرد الگوریتم‌ها در شرایط و ابعاد مختلف بررسی شد.

برای تضمین ارزیابی صحیح، داده‌ها به دو بخش آموزش و آزمون تقسیم شدند. نسبت ۸۰ به ۲۰ انتخاب شد تا بخش عمده‌ای از داده‌ها برای آموزش در اختیار مدل‌ها قرار گیرد و درعین‌حال مجموعه‌ای مستقل برای آزمون باقی بماند. این کار مانع از سرریز شد و کمک کرد عملکرد واقعی مدل‌ها بر روی داده‌های نادیده گرفته‌شده سنجیده شود.

۳-۳. پیش‌پردازش

در این مرحله، داده‌ها تحت چهار سناریوی متفاوت پردازش شدند. سناریو پایه به‌عنوان نقطه مرجع عمل کرد، یعنی داده‌ها بدون تغییر به مدل‌ها داده شدند. سناریو پاک‌سازی شامل حذف نمونه‌های تکراری یا ناسازگار بود که باعث شد داده‌ها کیفیت بالاتری داشته باشند و از بار محاسباتی غیرضروری کاسته شود. این موضوع در کاهش مصرف انرژی و بهبود دقت مدل تأثیرگذار است.

سناریوی سوم، تحلیل مؤلفه‌های اصلی بود که ابعاد داده‌ها را کاهش می‌داد. این کار با حفظ بیشترین واریانس موجود در داده‌ها، تعداد ویژگی‌های ورودی به مدل را کم می‌کرد و درنتیجه نیاز محاسباتی کاهش می‌یافت. در نهایت، سناریوی نمونه‌برداری اجرا شد که در آن نیمی از داده‌های آموزشی به‌صورت تصادفی انتخاب شدند. این روش به‌ویژه برای بررسی تأثیر کاهش حجم داده‌ها بر عملکرد و مصرف انرژی ارزشمند بود.

1. Train/Test Split

۳-۴. آموزش مدل‌های سنتی

برای داده‌های کوچک‌تر (Iris، Wine و Digits) الگوریتم جنگل تصادفی انتخاب شد. این الگوریتم یکی از قدرتمندترین روش‌های دسته‌بندی کلاسیک است که با استفاده از مجموعه‌ای از درخت‌های تصمیم عمل می‌کند. مزیت اصلی آن در مدیریت داده‌های با نویز بالا و کاهش خطر بیش‌برازش است. مدل برای هر یک از سناریوهای پیش‌پردازشی آموزش داده شد تا تأثیر مستقیم هر روش بر دقت، توان پیش‌بینی و مصرف انرژی بررسی گردد. یکی از نقاط قوت استفاده الگوریتم جنگل تصادفی از این بود که نتایج حاصل قابلیت مقایسه بالایی داشتند، زیرا الگوریتم نسبتاً پایدار و مقاوم است. این ویژگی اجازه داد تغییرات مشاهده‌شده در معیارهای عملکردی و انرژی را بیشتر به سناریوهای پیش‌پردازش نسبت دهیم تا به ضعف‌های ذاتی الگوریتم.

۳-۵. اندازه‌گیری انرژی و دی‌اکسیدکربن

برای سنجش دقیق مصرف انرژی و اثرات زیست‌محیطی، از کتابخانه کُدکربن استفاده شد. این ابزار در طول فرایند آموزش مدل‌ها، میزان انرژی مصرف‌شده را ثبت کرد و سپس آن را برحسب انتشار دی‌اکسیدکربن معادل محاسبه نمود. در این پژوهش، ضریب انتشار ویژه ایران (۰/۵۷ کیلوگرم دی‌اکسیدکربن برای هر kWh) به‌کار گرفته شد تا نتایج به شرایط واقعی نزدیک‌تر باشند.

این مرحله از اهمیت ویژه‌ای برخوردار بود، زیرا به پژوهش بُعدی فراتر از کارایی الگوریتم‌ها بخشید و امکان تحلیل پایداری زیست‌محیطی را فراهم ساخت. اندازه‌گیری انرژی و دی‌اکسیدکربن نشان داد که حتی تغییرات کوچک در داده‌ها مانند حذف نمونه‌های تکراری یا کاهش ابعاد با سناریو تحلیل مؤلفه‌های اصلی می‌توانند کاهش قابل‌توجهی در ردپای کربنی مدل‌ها ایجاد کنند. این یافته‌ها تأکید می‌کنند که انتخاب روش‌های بهینه نه‌تنها از منظر علمی بلکه از دیدگاه محیط‌زیستی نیز ضروری است.

۳-۶. آموزش مدل عمیق

برای مجموعه داده تصویری CIFAR-10که شامل ۶۰ هزار تصویر رنگی در ۱۰ کلاس مختلف است، از یک شبکه عصبی پیچشی ساده استفاده شد. این مدل با استفاده از لایه‌های کانولوشن، Pooling و Fully Connected طراحی گردید تا ویژگی‌های بصری پیچیده را استخراج و دسته‌بندی کند.

PyTorch به‌عنوان بستر پیاده‌سازی انتخاب شد، زیرا انعطاف‌پذیری و سرعت بالایی در آموزش شبکه‌های عصبی دارد. مدل با الگوریتم بهینه‌ساز Adam و تابع هزینه Cross-Entropy Loss آموزش داده شد. انتخاب این معماری و ابزارها امکان مقایسه‌ای جالب میان مدل‌های کلاسیک و مدل‌های عمیق فراهم کرد. نتایج نشان دادند که درحالی‌که شبکه عصبی پیچشی توانایی بالاتری در پردازش داده‌های تصویری دارد، هزینه انرژی و انتشار دی‌اکسیدکربن آن نیز بسیار بیشتر از مدل‌های سبک‌تر است.

۳-۷. محاسبه معیارهای عملکرد

برای ارزیابی کیفیت مدل‌ها از چهار معیار اصلی استفاده شد. نرخ صحت نشان‌دهنده نسبت پیش‌بینی‌های صحیح به کل پیش‌بینی‌ها است و دید کلی از عملکرد مدل ارائه می‌دهد. نرخ دقت میزان دقت مدل در شناسایی نمونه‌های مثبت را نشان می‌دهد و اهمیت ویژه‌ای در کاربردهایی دارد که خطای مثبت کاذب پُرهزینه است. در مقابل، نرخ بازخوانی بر توانایی مدل در شناسایی تمام نمونه‌های مثبت واقعی تأکید دارد. برای متوازن‌سازی این دو معیار، از امتیاز F1 استفاده شد که میانگین هارمونیک نرخ دقت و نرخ بازخوانی است.

این ترکیب از معیارها تصویری چندبُعدی از عملکرد مدل‌ها ارائه داد و مانع از آن شد که تنها بر یک شاخص خاص تمرکز شود.

۳-۸. تحلیل آماری و شاخص‌های ترکیبی کارایی-انرژی

برای دستیابی به بینشی جامع‌تر، علاوه بر مقایسه مستقیم معیارهای عملکرد و انرژی،

شاخص‌های ترکیبی طراحی شدند. این شاخص‌ها نرخ صحت، دقت، بازخوانی و امتیاز F1 به انتشار دی‌اکسیدکربن را محاسبه می‌کردند. هدف از این تحلیل‌ها ایجاد معیاری بود که هم‌زمان کارایی علمی و بهره‌وری انرژی را در نظر بگیرد.

این بخش نشان داد که برخی سناریوها اگرچه دقتی مشابه سایرین دارند، اما از نظر مصرف انرژی بسیار کارآمدترند. برای مثال، در برخی داده‌ها، سناریوهای پاک‌سازی و تحلیل مؤلفه‌های اصلی توانستند بدون کاهش محسوس کیفیت، ردپای انرژی را به شکل معناداری کم کنند. چنین تحلیل‌هایی اهمیت زیادی برای تصمیم‌گیری عملی دارند، زیرا در بسیاری از کاربردها انتخاب مدلی که تعادل بهتری میان دقت و پایداری دارد، از انتخاب مدلی با بیشترین دقت صرف ارزشمندتر است. نتایج نهایی در بخش شماره ۳ (تجزیه‌وتحلیل یافته‌ها) توضیح داده شده‌اند.

۳-۹. مصورسازی نتایج

برای انتقال بهتر یافته‌ها، نتایج عددی با استفاده از نمودارها مصورسازی شدند. نمودار پراکندگی رابطه میان نرخ صحت و انتشار دی‌اکسیدکربن را نشان داد و به‌وضوح تأثیر سناریوهای مختلف را آشکار کرد. نمودارهای میله‌ای نیز برای مقایسه دقیق معیارهای عملکردی و انرژی در میان مجموعه داده‌ها و سناریوها ترسیم شدند.

این مصورسازی‌ها دو مزیت اصلی داشتند: نخست، داده‌های پیچیده در قالبی ساده و قابل فهم برای خواننده ارائه شدند. دوم، الگوهایی که در جدول‌های عددی به‌سختی دیده می‌شدند، به شکل بصری برجسته شدند. برای نمونه، به‌وضوح مشخص شد که مدل شبکه عصبی پیچشی روی CIFAR-10 نسبت به مدل‌های کلاسیک مصرف انرژی بسیار بالاتری دارد. به این ترتیب، مصورسازی نقش کلیدی در شفاف‌سازی نتایج و جذاب‌تر کردن ارائه مقاله ایفا کرد. نتایج مصورسازی در بخش شماره ۳ (تجزیه‌وتحلیل یافته‌ها) نمایش داده شده‌اند.

۴. تجزیه‌وتحلیل یافته‌ها

در این پژوهش برای بررسی کارایی و بهره‌وری الگوریتم‌ها، نتایج به‌دست‌آمده از چهار مجموعه داده مختلف شامل Iris، Digits، Wine و CIFAR-10 تحت سناریوهای گوناگون پیش‌پردازش مورد تحلیل قرار گرفت که در جدول (۲)، مقادیر انواع پارامترهای ارزیابی نشان داده شده است.

جدول ۲: مقایسه دقت رویکرد پیشنهادی و کارهای انجام شده پیشین

مجموعه داده	سناریوها	نرخ صحت	نرخ دقت	نرخ بازخوانی	امتیاز F1	انرژی (kWh)	دی‌اکسیدکربن (kg)
Iris	Baseline	۱/۰۰	۱/۰۰	۱/۰۰	۱/۰۰	۰/۰۰۰۰۰۵	۰/۰۰۰۰۰۳
Iris	Cleaning	۱/۰۰	۱/۰۰	۱/۰۰	۱/۰۰	۰/۰۰۰۰۰۳	۰/۰۰۰۰۰۲
Iris	PCA	۱/۰۰	۱/۰۰	۱/۰۰	۱/۰۰	۰/۰۰۰۰۰۳	۰/۰۰۰۰۰۲
Iris	Sampling 50%	۱/۰۰	۱/۰۰	۱/۰۰	۱/۰۰	۰/۰۰۰۰۰۴	۰/۰۰۰۰۰۲
Digits	Baseline	۰/۹۷	۰/۹۷	۰/۹۷	۰/۹۷	۰/۰۰۰۰۰۹	۰/۰۰۰۰۰۵
Digits	Cleaning	۰/۹۷	۰/۹۷	۰/۹۷	۰/۹۷	۰/۰۰۰۰۰۹	۰/۰۰۰۰۰۵
Digits	PCA	۰/۹۶	۰/۹۶	۰/۹۶	۰/۹۶	۰/۰۰۰۰۱۳	۰/۰۰۰۰۰۸
Digits	Sampling 50%	۰/۹۷	۰/۹۷	۰/۹۷	۰/۹۷	۰/۰۰۰۰۰۷	۰/۰۰۰۰۰۴
Wine	Baseline	۱/۰۰	۱/۰۰	۱/۰۰	۱/۰۰	۰/۰۰۰۰۰۵	۰/۰۰۰۰۰۳
Wine	Cleaning	۱/۰۰	۱/۰۰	۱/۰۰	۱/۰۰	۰/۰۰۰۰۰۳	۰/۰۰۰۰۰۲
Wine	PCA	۱/۰۰	۱/۰۰	۱/۰۰	۱/۰۰	۰/۰۰۰۰۰۴	۰/۰۰۰۰۰۲
Wine	Sampling 50%	۰/۹۷	۰/۹۸	۰/۹۸	۰/۹۸	۰/۰۰۰۰۰۳	۰/۰۰۰۰۰۲
CIFAR-10	Baseline	۰/۵۶	۰/۵۶	۰/۵۶	۰/۵۵	۰/۰۰۰۲۲۷	۰/۰۰۰۱۵۵

یافته‌ها نشان می‌دهند که اگرچه معیارهای عملکردی همچون نرخ صحت، دقت، بازخوانی و امتیاز F1 در برخی موارد تقریباً ثابت باقی مانده‌اند؛ اما میزان مصرف انرژی و انتشار دی‌اکسیدکربن تفاوت‌های قابل توجهی داشته است. این امر نشان‌دهنده اهمیت سنجش هم‌زمان کارایی علمی و پایداری زیست‌محیطی در ارزیابی مدل‌ها است.

نتایج حاصل از مجموعه داده Iris نشان دادند که در تمامی سناریوها (پایه، پاک‌سازی داده‌ها، تحلیل مؤلفه‌های اصلی و نمونه‌برداری) نرخ صحت و سایر معیارهای عملکردی

به‌طور کامل برابر با ۱۰۰٪ بودند. این یافته بدین معناست که داده‌های Iris به‌دلیل ساختار ساده و قابل تفکیک، در برابر تغییرات پیش‌پردازشی حساسیت چندانی ندارند. بااین‌حال، در مقیاس انرژی تفاوت‌های معناداری مشاهده شد. برای مثال، مصرف انرژی در حالت پاک‌سازی و تحلیل مؤلفه‌های اصلی به ۰/۰۰۰۰۰۳kWh کاهش یافت که نسبت به حالت پایه (۰/۰۰۰۰۰۵ kWh) صرفه‌جویی قابل توجهی محسوب می‌شود. همین روند در انتشار دی‌اکسیدکربن نیز مشاهده شد. این موضوع نشان می‌دهد که حتی در شرایطی که نرخ صحت مدل تغییر نمی‌کند، انتخاب سناریوی مناسب می‌تواند اثرات زیست‌محیطی الگوریتم را کاهش دهد.

در مجموعه داده Digits که نسبت به Iris پیچیده‌تر است، عملکرد مدل در سطح بالایی باقی ماند و در سناریوهای پایه، پاک‌سازی و نمونه‌برداری نرخ صحت برابر با ۰/۹۷ ثبت شد. تنها در حالت تحلیل مؤلفه‌های اصلی اندکی افت، نرخ صحت به ۰/۹۶ مشاهده گردید. این کاهش نشان‌دهنده آن است که کاهش ابعاد داده با تحلیل مؤلفه‌های اصلی در این مجموعه موجب از دست رفتن بخشی از اطلاعات مهم شده است. بااین‌حال، از نظر انرژی، اجرای سناریوی تحلیل مؤلفه‌های اصلی نیازمند ۰/۰۰۰۰۱۳ kWh بود که از مصرف سناریوی پایه بیشتر است. این تناقض بیانگر آن است که انتخاب روش پیش‌پردازش همیشه به بهبود هم‌زمان نرخ صحت و انرژی منجر نمی‌شود و باید میان این دو جنبه تعادل برقرار کرد.

مجموعه داده Wine نتایجی مشابه Iris به‌همراه داشت. در سناریوهای پایه، پاک‌سازی و تحلیل مؤلفه‌های اصلی نرخ صحت مدل برابر با ۱۰۰٪ بود؛ اما در حالت نمونه‌برداری اندکی افت مشاهده شد و نرخ صحت به ۰/۹۷ کاهش یافت. بااین‌حال، شاخص‌های نرخ دقت، نرخ بازخوانی و امتیاز F1 همچنان در سطح ۰/۹۸ باقی ماندند که نشان‌دهنده پایداری الگوریتم در مواجهه با کاهش حجم داده‌ها است. از منظر انرژی، پاک‌سازی و نمونه‌برداری بهینه‌ترین نتایج را ثبت کردند، به‌گونه‌ای که مصرف انرژی به ۰/۰۰۰۰۰۳ kWh رسید. این یافته اهمیت استفاده از روش‌های داده‌محور ساده را در کاهش هزینه محاسباتی بدون آسیب جدی به نرخ صحت نشان می‌دهد. در نهایت، تحلیل مجموعه داده CIFAR-10که به‌مراتب پیچیده‌تر از سه

مجموعه دیگر است، تصویری متفاوت ارائه داد. در این مجموعه که نیازمند استفاده از مدل‌های عمیق‌تر مانند شبکه عصبی پیچشی است، نرخ صحت مدل تنها ۰/۵۶ ثبت شد و سایر معیارها نیز در همین حدود باقی ماندند. این میزان در مقایسه با نرخ صحت‌های تقریباً کامل در داده‌های کلاسیک بسیار پایین‌تر است و نشان می‌دهد که چالش‌های مرتبط با داده‌های تصویری بسیار بیشترند. مهم‌تر از آن، مصرف انرژی در این بخش به ۰/۰۰۰۲۷۲ kWh و انتشار دی‌اکسیدکربن به ۰/۰۰۰۱۵۵ kgرسید که چندین برابر بیشتر از داده‌های ساده‌تر بود. این امر بیانگر بار محاسباتی سنگین شبکه‌های عصبی عمیق و ضرورت توجه جدی به بهینه‌سازی مصرف انرژی در چنین کاربردهایی است.

به‌طورکلی، نتایج این پژوهش نشان دادند که:

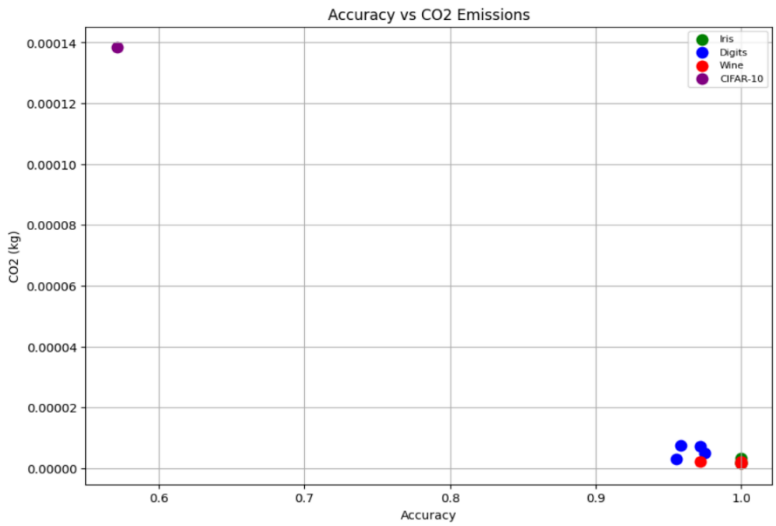
۱. در داده‌های ساده مانند Iris و Wine، پیش‌پردازش‌های پاک‌سازی و تحلیل مؤلفه‌های اصلی می‌توانند مصرف انرژی را کاهش دهند بدون این‌که بر نرخ صحت اثر بگذارند؛

۲. در داده‌های متوسط مانند Digits، برخی روش‌ها مثل تحلیل مؤلفه‌های اصلی ممکن است باعث افت جزئی در نرخ صحت شوند و حتی مصرف انرژی بیشتری ایجاد کنند؛

۳. در داده‌های پیچیده و تصویری مانند CIFAR-10، استفاده از مدل‌های عمیق با وجود توانایی پردازش تصویر، هزینه انرژی و انتشار دی‌اکسیدکربن بسیار بالایی دارد.

این تحلیل‌ها به‌روشنی نشان می‌دهند که انتخاب رویکرد مناسب نه‌تنها باید براساس معیارهای سنتی مانند نرخ صحت انجام گیرد، بلکه معیارهای زیست‌محیطی نظیر انرژی و انتشار دی‌اکسیدکربن نیز باید در فرایند تصمیم‌گیری لحاظ شوند. چنین نگرشی امکان دستیابی به الگوریتم‌هایی پایدارتر و کارآمدتر را فراهم می‌آورد.

نمودار پراکندگی ارائه‌شده در شکل (۲)، رابطه میان نرخ صحت مدل‌ها و میزان انتشار دی‌اکسیدکربن را برای مجموعه داده‌های مختلف نشان می‌دهد.



شکل ۲: رابطه میان نرخ صحت مدل‌ها و میزان انتشار دی‌اکسیدکربن را برای مجموعه داده‌های مختلف

مشاهده نقاط در این نمودار بیانگر آن است که داده‌های ساده مانند Irisو Wine در محدوده‌ای نزدیک به صحت کامل (۱/۰) قرار دارند، درحالی‌که انتشار دی‌اکسیدکربن آن‌ها در پایین‌ترین سطح ممکن ثبت شده است.

این وضعیت نشان می‌دهد که برای داده‌هایی با ساختار ساده، می‌توان به بالاترین نرخ صحت دست یافت بدون آنکه هزینه قابل توجهی از نظر انرژی یا اثرات زیست‌محیطی ایجاد شود. در مقابل، نقطه مربوط به مجموعه داده CIFAR-10 در سمت چپ نمودار قرار گرفته و نرخ صحتی در حدود ۰/۵۶ را نشان می‌دهد. درعینِ‌حال، میزان انتشار دی‌اکسیدکربن آن به‌طور قابل توجهی بالاتر از سایر داده‌ها است. این اختلاف چشمگیر بیانگر آن است که مدل‌های عمیق نظیر شبکه عصبی پیچشی با وجود بار محاسباتی زیاد، الزاماً به بهبود نرخ صحت منجر نمی‌شوند؛ بلکه در مواردی حتی با کاهش صحت و افزایش انتشار دی‌اکسیدکربن همراه هستند. ازاین‌رو، استفاده از چنین مدل‌هایی در کاربردهای پیچیده باید با استراتژی‌های بهینه‌سازی انرژی همراه شود تا هزینه‌های زیست‌محیطی آن توجیه‌پذیر باشد.

همچنین، مجموعه داده Digits اگرچه صحتی کم‌تر از Iris و Wine دارد (حدود ۰/۹۶ تا ۰/۹۷)؛ اما همچنان انتشار دی‌اکسیدکربن آن در سطحی بسیار پایین باقی‌مانده است. این موقعیت میانی نشان می‌دهد که داده‌های با پیچیدگی متوسط به‌نوعی نقطه تعادل میان صحت بالا و انتشار کربن اندک را فراهم می‌کنند.

به‌طورکلی، تحلیل این نمودار نشان می‌دهد که:

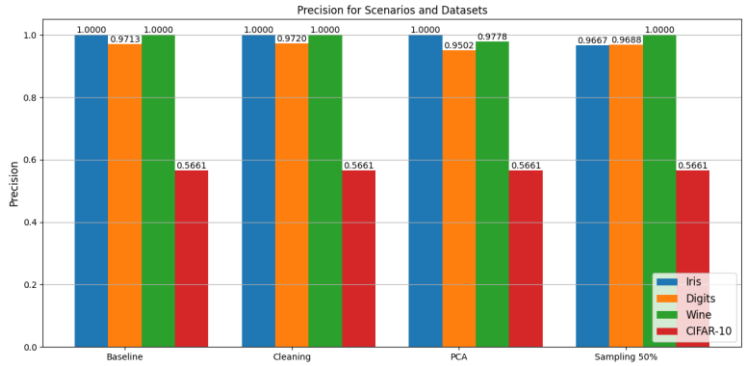
۱. مدل‌های ساده روی داده‌های ساده (Iris و Wine) بهترین ترکیب صحت و کارایی انرژی را دارند؛
۲. مدل‌های متوسط روی داده‌های نیمه‌پیچیده (Digits) می‌توانند تعادل نسبی ایجاد کنند؛
۳. مدل‌های عمیق روی داده‌های پیچیده (CIFAR-10) با هزینه بالای انرژی همراه‌اند و کارایی نسبی آن‌ها از منظر صحت به‌تنهایی توجیه‌کننده این مصرف بالا نیست.

این نتایج اهمیت رویکرد هوش مصنوعی سبز را دوچندان می‌سازد، زیرا نشان می‌دهد تصمیم‌گیری در انتخاب مدل باید علاوه بر معیارهای سنتی عملکرد، شامل شاخص‌های زیست‌محیطی نیز باشد. در غیر این صورت، استفاده از مدل‌های بسیار پیچیده ممکن است به افزایش قابل‌توجه دی‌اکسیدکربن بدون دستاورد علمی متناسب منجر شود.

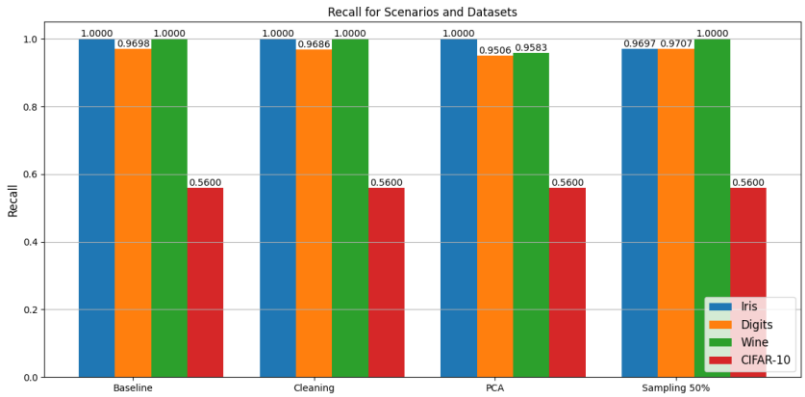
برای درک بهتر تأثیر سناریوهای مختلف پیش‌پردازش بر عملکرد مدل‌ها، معیارهای نرخ صحت، دقت، بازخوانی و امتیاز F1 در کنار شاخص‌های مصرف انرژی (kWh) و انتشار دی‌اکسیدکربن (kg) به تفکیک برای هر مجموعه داده رسم و تحلیل شدند. نمودارهای ستونی ارائه‌شده در شکل‌های ۳ تا ۶ نشان می‌دهند که تغییر سناریو چگونه بر تعادل بین کارایی محاسباتی و کیفیت پیش‌بینی تأثیر می‌گذارد.



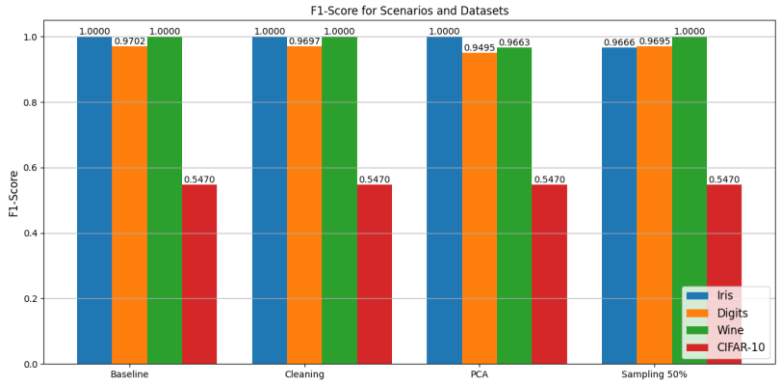
شکل ۳: سناریوهای مختلف پیش‌پردازش بر نرخ صحت



شکل ۴: سناریوهای مختلف پیش‌پردازش بر نرخ دقت



شکل ۵: سناریوهای مختلف پیش‌پردازش بر نرخ بازخوانی



شکل ۶: سناریوهای مختلف پیش‌پردازش بر امتیاز F1

به‌طورکلی، تحلیل نمودارها نشان می‌دهد که در داده‌های ساده، پاک‌سازی داده‌ها و تحلیل مؤلفه‌های اصلی بهترین روش‌ها برای کاهش مصرف انرژی و حفظ دقت هستند. در داده‌های متوسط، نمونه‌گیری و پاک‌سازی داده‌ها تعادل بهتری بین کیفیت و مصرف انرژی ایجاد می‌کنند؛ اما در داده‌های پیچیده، نیاز به بهینه‌سازی‌های عمیق‌تر در سطح معماری مدل و سخت‌افزار وجود دارد، زیرا روش‌های ساده کارایی کافی ندارند. این یافته‌ها تأکید می‌کنند که انتخاب روش پیش‌پردازش مناسب باید براساس ویژگی‌های داده و پیچیدگی مدل انجام شود و توجه صرف به پارامترهای ارزیابی بدون در نظر گرفتن مصرف انرژی می‌تواند به تصمیم‌های ناپایدار منجر شود.

۴-۱. جمع‌بندی

یافته‌های این پژوهش نشان داد که ترکیب سناریوهای پیش‌پردازشی سبک‌وزن (مانند پاک‌سازی و تحلیل مؤلفه‌های اصلی) با مدل‌های سنتی مانند جنگل تصادفی و استفاده از ابزار کُدکربن برای اندازه‌گیری مصرف انرژی می‌تواند به نتایجی دست یابد که هم از نظر نرخ صحت و هم از نظر کارایی انرژی قابل‌توجه هستند. درحالی‌که پژوهش‌های پیشین عمدتاً بر افزایش نرخ صحت و دقت مدل‌ها بدون توجه به مصرف انرژی متمرکز بوده‌اند، رویکرد حاضر تلاش می‌کند توازنی میان کیفیت پیش‌بینی و پایداری زیست‌محیطی برقرار کند.

به‌عنوان نمونه، استروبِل و همکاران (۲۰۱۹) و پترسون و همکاران (۲۰۲۱) نشان دادند که آموزش مدل‌های بزرگ NLP و شبکه‌های عمیق می‌تواند تا صدها هزار پوند دی‌اکسیدکربن تولید کند، درحالی‌که رویکرد پیشنهادی ما توانست با استفاده از مجموعه داده‌های متنوع، میزان انتشار را در محدوده میلی‌گرم تا گرم نگه دارد و هم‌زمان دقت بسیار بالا (تا ۱۰۰٪ در مجموعه‌های ساده‌تر) ارائه دهد. همچنین، برخلاف شوارتز و همکاران (۲۰۱۹) که صرفاً بر مفهوم هوش مصنوعی سبز و اهمیت آن تأکید داشتند، این پژوهش یک پیاده‌سازی عملی و قابل تکرار از آن را ارائه می‌کند.

مقایسه با وردکیا و همکاران (۲۰۲۳) نیز نشان می‌دهد که اگرچه بسیاری از رویکردهای موجود در محیط‌های آزمایشگاهی باقی مانده‌اند، روش حاضر یک مسیر عملیاتی برای ترکیب کارایی انرژی با الگوریتم‌های سنتی و عمیق فراهم می‌آورد. درعینِ‌حال، نتایج این پژوهش هم‌سو با مطالعات بولو و همکاران (۲۰۲۴) است که نشان دادند انتخاب معماری بهینه می‌تواند منجر به صرفه‌جویی چشمگیر در انرژی شود.

جدول (۳)، خلاصه‌ای از مقایسه رویکرد پیشنهادی با پژوهش‌های شاخص گذشته را ارائه می‌دهد.

جدول ۳: مقایسه رویکرد پیشنهادی با سایر مطالعات

پژوهش / رویکرد	دقت مدل‌ها	تمرکز بر مصرف انرژی	میزان انتشار دی‌اکسیدکربن گزارش‌شده	نوع داده/مدل	مزیت یا محدودیت اصلی
(استروبِل و همکاران، ۲۰۲۰)	بالا در NLP	خیر	تا ۶۲۶،۰۰۰ پوند	NLP, Transformer	هزینه بسیار زیاد انرژی و کربن
(پترسون و همکاران، ۲۰۲۱)	بسیار بالا	بخشی	تا ۱۰۰۰ برابر کاهش با بهینه‌سازی زیرساخت	T5, GPT-3	نیازمند مراکز داده پیشرفته
(شوارتز و همکاران، ۲۰۲۰)	متغیر	بله (مفهومی)	گزارش کلی (بدون عدد دقیق)	یادگیری عمیق	تأکید بر فلسفه هوش مصنوعی سبز

مزیت یا محدودیت اصلی	نوع داده/مدل	میزان انتشار دی‌اکسیدکربن گزارش‌شده	تمرکز بر مصرف انرژی	دقت مدل‌ها	پژوهش / رویکرد
عمدتاً آزمایشگاهی	متنوع	صرفه‌جویی تا ۱۱۵٪	بله	متفاوت	(وردکیا و همکاران، ۲۰۲۳)
تمرکز بر inference	شبکه‌های پویا	بهینه‌سازی پویا در inference	بله	بالا	(بولو و همکاران، ۲۰۲۴)
تعادل میان دقت و کارایی انرژی؛ قابل تکرار و ساده	جنگل تصادفی، شبکه عصبی پیچشی	در حد میلی‌گرم تا گرم (۰/۰۰۰۰۰۲– ۰/۰۰۰۱۵۵ کیلوگرم)	بله (عملیاتی)	بالا (۰/۹۶– ۱/۰ در داده‌های ساده، ۰/۵۶ در CIFAR-10)	رویکرد پیشنهادی (این پژوهش)

این مقایسه نشان می‌دهد که برخلاف بسیاری از کارهای قبلی که تنها به افزایش نرخ صحت یا معرفی چهارچوب‌های نظری پرداخته‌اند، رویکرد پیشنهادی این پژوهش یک مسیر عملی، ساده و شفاف برای کاهش مصرف انرژی و ردپای کربنی ارائه می‌کند، بدون آنکه کیفیت پیش‌بینی قربانی شود. این مزیت به‌ویژه برای کاربردهای دفاعی و امنیتی که بهره‌وری و الگوی مصرف انرژی حیاتی هستند، اهمیت دوچندان دارد.

۵. نتیجه‌گیری و پیشنهاد

پژوهش حاضر با هدف بررسی و ارزیابی رابطه میان عملکرد مدل‌های یادگیری ماشین و یادگیری عمیق با مصرف انرژی و انتشار کربن انجام شد و براساس مجموعه داده‌های متنوع (Iris، Digits، Wine و CIFAR-10) و سناریوهای گوناگون پیش‌پردازشی (پایه، پاک‌سازی داده‌ها، تحلیل مؤلفه‌های اصلی و نمونه‌برداری) مورد آزمایش قرار گرفت. نتایج نشان داد که در داده‌های کوچک‌تر مانند Iris و Wine، مدل‌های سنتی همچون الگوریتم جنگل تصادفی قادر بودند با مصرف انرژی بسیار اندک و انتشار کربن در حد میلی‌گرم، به نرخ صحت کامل

(۱۰۰٪) دست یابند. این موضوع بیانگر آن است که استفاده از مدل‌های پیچیده و پُرهزینه همواره ضروری نیست و در بسیاری از کاربردها می‌توان با روش‌های سبک‌وزن به نتایجی کاملاً رقابتی دست یافت. در مقابل، در داده‌های پیچیده‌تر مانند CIFAR-10، مدل عمیق شبکه عصبی پیچشی اگرچه صحت حدود ۵۶٪ به دست آورد، اما مصرف انرژی و انتشار دی‌اکسیدکربن آن به‌طور قابل‌توجهی بیشتر از مدل‌های ساده‌تر بود. این یافته بار دیگر تأکید می‌کند که توازن میان کیفیت پیش‌بینی و پایداری انرژی باید به‌عنوان یک اصل بنیادین در طراحی سامانه‌های هوشمند مدنظر قرار گیرد.

از نظر سناریوهای پیش‌پردازشی، پاک‌سازی و تحلیل مؤلفه‌های اصلی نه‌تنها کیفیت مدل را کاهش ندادند بلکه در برخی موارد با کاهش مصرف انرژی همراه بودند. به‌ویژه، سناریوی تحلیل مؤلفه‌های اصلی در برخی داده‌ها توانست بدون افت عملکرد، دی‌اکسیدکربن را کاهش دهد. این مسئله نشان می‌دهد که استفاده هوشمندانه از پیش‌پردازش داده‌ها می‌تواند به‌عنوان یک ابزار کارآمد در مدیریت مصرف انرژی مدل‌های هوش مصنوعی عمل کند.

براساس این نتایج، می‌توان چند پیشنهاد کلیدی برای آینده مطرح کرد:

- ✓ کاربرد در حوزه‌های حساس: در صنایع دفاعی و امنیتی که پایداری انرژی و بهره‌وری اهمیت حیاتی دارند، استفاده از مدل‌های سبک‌وزن و الگوریتم‌های کلاسیک می‌تواند به‌عنوان یک راهکار کم‌هزینه و مطمئن مورد توجه قرار گیرد.
- ✓ ترکیب مدل‌ها: استفاده از رویکردهای ترکیبی مانند (Ensemble یاHybrid Models) می‌تواند صحت مدل‌های ساده‌تر را در داده‌های پیچیده افزایش داده و هم‌زمان مصرف انرژی را مدیریت کند.
- ✓ بهینه‌سازی زیرساخت‌ها: طراحی چهارچوب‌های بهینه برای اجرای مدل‌ها ازجمله استفاده ازGPU های کم‌مصرف، فشرده‌سازی مدل و pruning می‌تواند نقش مؤثری در کاهش دی‌اکسیدکربن داشته باشد.
- ✓ گسترش به داده‌های واقعی‌تر: بررسی مجموعه داده‌های حجیم‌تر و پیچیده‌تر در آینده می‌تواند میزان تعمیم‌پذیری رویکرد پیشنهادی را روشن‌تر کند.

۵-۱. محدودیت‌های پژوهش

یافته‌های این پژوهش در چهارچوب داده‌ها، سناریوهای بهینه‌سازی و بستر محاسباتی مورد استفاده قابل تفسیر است. بدیهی است تغییر در نوع مجموعه داده، معماری مدل‌ها یا زیرساخت اجرایی می‌تواند بر میزان مصرف انرژی و عملکرد نهایی اثرگذار باشد؛ ازاین‌رو، تعمیم نتایج به سایر شرایط، مستلزم آزمون در محیط‌ها و پیکربندی‌های متفاوت خواهد بود.

در نهایت، این پژوهش نشان می‌دهد که توسعه مدل‌های هوش مصنوعی تنها نباید به افزایش صحت و دقت معطوف باشد، بلکه باید هم‌زمان ملاحظات زیست‌محیطی و بهره‌وری انرژی نیز در طراحی و انتخاب الگوریتم‌ها لحاظ شوند. این تغییر نگرش می‌تواند گامی مؤثر در جهت تحقق الگوی مصرف دفاعی و پایداری هوش مصنوعی سبز باشد.

فهرست منابع

اله دادی، حمید؛ لطیف منش، حجت اله؛ معصومی اصل، اسد (۱۴۰۴). ***کاربردهای هوش مصنوعی در علوم زراعی و بهنژادی گیاهی: راهبردها، فرصت‌ها و چالش‌ها***. فصلنامه آماد و فناوری دفاعی، ۸(۲).

تقندیکی، کاظم (۱۴۰۴). ***تشخیص وضعیت بیماران مبتلا به هپاتیت C با استفاده از الگوریتم یادگیری عمیق CNN***. فصلنامه آماد و فناوری دفاعی، ۸(۲).

References

Agrawal, V. K., Bogireddy, S. R., Murari, H., Patil, L. N., Panwar, V. S., & Patil, Y. M. (2025, May). *Handling Missing Data in the Iris Dataset: Imputation Techniques and Their Effects on Model Accuracy*. In 2025 Global Conference in Emerging Technology (GINOTECH) (pp. 1-4). IEEE.

Ahmed, T., Khan, M., Khan, K., Syed, I., & Ullah, S. S. (2025). *A comprehensive voice dataset for Hindko digit recognition*. Data in Brief, 58, 111220.

Borraccia, S., Masciari, E., & Napolitano, E. V. (2025, August). *A Hybrid Approach to Estimating AI Carbon Emissions*. In International Conference on Database and Expert Systems Applications (pp. 329-337). Cham: Springer Nature Switzerland.

Bullo, M., Jardak, S., Carnelli, P., & Gündüz, D. (2024). *Energy-Aware Dynamic Neural Inference*. arXiv preprint arXiv:2411.02471.

Castellanos-Nieves, D., & García-Forte, L. (2023). *Improving automated machine-learning systems through green ai*. Applied Sciences, 13(20), 11583.

Chen, Y., Wan, Z., Zhuang, Y., Liu, N., Lo, D., & Yang, X. (2025). *Understanding the OSS Communities of Deep Learning Frameworks: A Comparative Case Study of PyTorch and TensorFlow*. ACM Transactions on Software Engineering and Methodology, 34(3), 1-30.

Färber, M., & Lamprecht, D. (2022). *The Green AI Ontology: An Ontology for Modeling the Energy Consumption of AI Models*. ISWC (Posters/Demos/Industry),

Gautam, A., Lohumi, Y., & Gangodkar, D. (2024, July). *Achieving near-perfect accuracy in CIFAR-10 classification*. In 2024 Second international conference on advances in information technology (ICAIT) (Vol. 1, pp. 1-6). IEEE.

Gidlund, J. (2025). *Green AI: Reducing the Carbon Footprint of Automated Testing*.

Jegadeeswari, K., & Rathipriya, R. *Green AI Practices in Multi-objective Hyperparameter Optimization for Sustainable Machine Learning*.

Lemenkova, P. (2025). *Automation of image processing through ML algorithms of GRASS GIS using embedded Scikit-Learn library of Python*. Examples and Counterexamples, 7, 100180.

Mallala, B., Ahmed, A. I. U., Pamidi, S. V., Faruque, M. O., & Reddy, R. (2025). *Forecasting global sustainable energy from renewable sources using random forest algorithm*. Results in Engineering, 25, 103789.

Nelson, J., Coleman, R., Gravesen, P., Silacci, M., Velasquez, A., Marinelli, K., & Boulton, R. (2024). *Analysis of a Commercial Red Wine Fermentation Dataset with a Wine Kinetic Model*. Fermentation, 11(1), 4.

Ordoumpozanis, K., & Papakostas, G. A. (2024). *Green AI: Assessing the Carbon Footprint of Fine-Tuning Pre-Trained Deep Learning Models in Medical Imaging*. 2024 International Conference on Innovation and Intelligence for Informatics, Computing, and Technologies (3ICT),

Patterson, D., Gonzalez, J., Le, Q., Liang, C., Munguia, L.-M., Rothchild, D., So, D., Texier, M., & Dean, J. (2021). *Carbon emissions and large neural network*

training. arXiv preprint arXiv:2104.10350.

Preanto, S. A., Razi, A., Shorna, S. A., Moni, R., Dua, J., & Dwivedi, V. K. (2025, April). *A Sustainable Approach to Predict Coronary Artery Disease with Machine Learning and Explainable AI*. In 2025 12th International Conference on Computing for Sustainable Global Development (INDIACom) (pp. 1-6). IEEE.

Sapre, A., & Vartak, S. (2020). *Scientific computing and data analysis using NumPy and Pandas*. International Research Journal of Engineering and Technology, 7, 1334-1346.

Schodl, J., Lesota, O., Tommasel, A., & Schedl, M. (2025). *Investigating Carbon Footprint of Recommender Systems Beyond Training Time*. Proceedings of the Nineteenth ACM Conference on Recommender Systems,

Schwartz, R., Dodge, J., Smith, N. A., & Etzioni, O. (2020). *Green ai*. Communications of the ACM, 63(12), 54–63.

Strubell, E., Ganesh, A., & McCallum, A. (2020). *Energy and policy considerations for modern deep learning research*. Proceedings of the AAAI conference on artificial intelligence

ul Haq, F., Suki, N. M., Setini, M., Masood, A., & Khan, T. A. (2025). *Adopting green AI for SME sustainability: mediating role of green investment and moderation by green servant leadership*. Sustainable Futures, 101002.

Verdecchia, R., Sallou, J., & Cruz, L. (2023). *A systematic review of Green AI*. Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, 13(4), e1507.

Vrana, V. G., & Mondal, S. R. (2025). *The Carbon Code: Decoding AI's Role in Climate Mitigation*. In Generative AI for a Net-Zero Economy: Managing Climate Change and Business Innovation in the Digital Era (pp. 1-18). Singapore: Springer Nature Singapore

Wu, Z., Jing, L., Wu, B., & Jin, L. (2025). *A PCA-AdaBoost model for E-commerce customer churn prediction*. Annals of operations research, 350(2), 537-554.