

Prediction of Reading Achievement using Multiple Regression based on Machine Learning

Mahdi Kazemi, Ph.D. Student in Assessment & Measurement, Allameh Tabataba'i University, Tehran, Iran. (Corresponding Author). **E-mail:** mahdi_kazemi@atu.ac.ir

Masoud Geramipour, Associate Professor, Curriculum Studies & Assessment and Measurement Dept., Faculty of Psychology and Educational Sciences, Kharazmi University, Tehran, Iran. **E-mail:** mgramipour@khu.ac.ir

Abstract

This study employed a machine learning-based multiple regression model to predict reading literacy scores among 4th-grade students in Iran. The primary objective was to create a predictive model for reading literacy using advanced machine learning techniques. Data was sourced from the 2021 Progress in International Reading Literacy Study (PIRLS), encompassing a substantial sample of 5,943 students, selected through a stratified random sampling method to ensure representativeness. The analysis was conducted using the Scikit-learn package in Python, which facilitated the application of the multiple regression technique. The results indicated that specific attitude variables, notably self-confidence in reading and familiarity with digital devices, emerged as significant predictors of reading literacy scores. This finding underscores the importance of psychological and technological factors in literacy development. The research highlights the need for educational policymakers to prioritize initiatives aimed at enhancing students' self-confidence in reading and their comfort with digital tools. By addressing these areas, it is anticipated that literacy outcomes can be significantly improved, ultimately contributing to better educational achievements and lifelong learning skills among students.

Keywords: Elementary School, Machine Learning, Multiple Regression, PIRLS, Reading Achievement

پیش‌بینی پیشرفت خواندن با استفاده از رگرسیون چندگانه مبتنی بر یادگیری ماشین

مهدی کاظمی*

مسعود گرامی‌پور**

* دانشجوی دکتری رشتهٔ سنجش و اندازه‌گیری، دانشگاه علامه طباطبایی، تهران، ایران (نویسندهٔ مسئول).
mahdi_kazemi@atu.ac.ir

** دانشیار گروه مطالعات برنامه‌ریزی درسی و سنجش و اندازه‌گیری، دانشگاه خوارزمی، تهران، ایران.
mgramipour@khu.ac.ir

تاریخ پذیرش مقاله: ۱۴۰۴/۵/۲۸

تاریخ شروع بررسی: ۱۴۰۴/۱/۲۴

تاریخ دریافت مقاله: ۱۴۰۳/۱۲/۳۰

چکیده

این مطالعه از مدل رگرسیون چندگانه مبتنی بر یادگیری ماشین برای پیش‌بینی نمرات سواد خواندن دانش‌آموزان کلاس چهارم در کشور ایران استفاده کرد. هدف اصلی، ایجاد یک مدل پیش‌بینی برای سواد خواندن با استفاده از رویکرد یادگیری ماشین بود. داده‌های پژوهش حاضر مستخرج از پرلز سال ۲۰۲۱ و شامل نمونه قابل توجهی از ۵۹۴۳ دانش‌آموز است که از طریق روش نمونه‌گیری تصادفی طبقه‌ای برای اطمینان از معرف بودن نمونه انتخاب شده‌اند. تحلیل داده‌ها با کاربست تکنیک رگرسیون چندگانه در یادگیری ماشین انجام شد. نتایج نشان داد که متغیرهای نگرش دانش‌آموزان و مشخصاً اعتماد به نفس در خواندن و آشنایی با دستگاه‌های دیجیتال، می‌توانند نمرات سواد خواندن را به شکل معناداری پیش‌بینی کنند. این یافته بر اهمیت عوامل روان‌شناختی و فناوری در توسعهٔ سواد خواندن تأکید می‌کند. پژوهش حاضر لزوم ابتکار عمل سیاست‌گذاران آموزشی در جهت افزایش اعتماد به نفس دانش‌آموزان در خواندن و انس آن‌ها با ابزارهای دیجیتال را برجسته می‌کند. با تمرکز بر این زمینه‌ها، انتظار می‌رود که نمرات سواد خواندن به‌طور چشمگیری افزایش یابد و در نهایت به موفقیت‌های بهتر تحصیلی و توسعهٔ مهارت‌های یادگیری مادام‌العمر در دانش‌آموزان کمک شود.

کلمات کلیدی: پرلز، پیشرفت خواندن، دورهٔ ابتدایی، رگرسیون چندگانه، یادگیری ماشین

مهارت خواندن یکی از توانمندی‌های لازم برای پیشرفت فرد در تمامی حیطه‌های زندگی از جمله حوزه‌های تحصیلی و شغلی است. توانایی درک و فهم متون مختلف، روزنهٔ بسیاری از علوم گوناگون را به روی انسان باز کرده و وی را به تجربهٔ غنی بشر در زمان‌های گذشته متصل می‌سازد. در وصف اهمیت این مهارت، همین بس که در فقدان آن، پیشرفت بشر در زمینه‌های متفاوت از جمله علوم ریاضی، تجربی، انسانی و ... غیرممکن به نظر می‌رسد (سادات‌رضایی و همکاران، ۱۳۹۳). مهارت خواندن همچنین با کمک به رشد فرد در جنبه‌های شناختی، هیجانی و اجتماعی، نقش مهمی در تحول انسان، به‌خصوص کودکان و نوجوانان، بازی می‌کند (حسن‌آبادی و همکاران، ۱۳۹۶). در طول ۵۰ سال گذشته، مردم جهان در کشورهای مختلف نگران پیشرفت خواندن بوده‌اند (آلینگتون و مک‌گیل-فرانزن، ۲۰۲۱). یک دلیل موجه برای نگرانی در مورد پیشرفت خواندن دانش‌آموزان، شواهدی است که نشان می‌دهد دانش‌آموزان در حال حاضر، چه داخل و چه خارج از مدرسه، نسبت به والدین، پدربزرگ و مادر بزرگ خود، زمان کمتری را صرف خواندن می‌کنند (آیینگار و بال، ۲، ۲۰۰۷).

مطالعهٔ پرلز، که توسط انجمن بین‌المللی ارزشیابی پیشرفت تحصیلی (IEA)^۳ اجرا می‌شود، مهارت درک خواندن دانش‌آموزان کلاس چهارم را، که سن آن‌ها معمولاً بین ۹ تا ۱۰ است، می‌سنجد و از سال ۲۰۰۱ به بعد هر ۵ سال ۱ بار اجرا می‌شود. پرلز به سنجش دو هدف اصلی از خواندن می‌پردازد: (۱) تجربه ادبی^۴ و (۲) کسب اطلاعات و استفاده از آن. این نکته قابل ذکر است که تعداد یکسانی از سؤالات به هر ۲ هدف خواندن اختصاص داده شده است (لنکیت^۵ و همکاران، ۲۰۱۵). داده‌های جمع‌آوری شده توسط مطالعهٔ بین‌المللی پیشرفت سواد خواندن (پرلز)^۶ اطلاعات قابل مقایسه‌ای را در مورد کیفیت خواندن کودکان که توسط یک آزمون جامع سواد خواندن ارزیابی می‌شود، ارائه می‌کنند. این آزمون بر ۴ فرآیند درک مطلب همه‌جانبه که توسط دانش‌آموزان پایه چهارم استفاده می‌شود، متمرکز است: (۱) تمرکز بر اطلاعات دقیق و واضح و بازیابی آن‌ها؛ (۲) استنباط‌های مستقیم؛ (۳) تفسیر و ادغام ایده‌ها و اطلاعات و (۴) ارزیابی و نقد عناصر محتوایی و متنی (مولیس و مارتین^۷، ۲۰۱۹).

مطالعهٔ پرلز، عوامل مؤثر بر پیشرفت خواندن دانش‌آموزان را در ۵ سطح مختلف بررسی می‌کند: (۱) دانش‌آموز، (۲) والدین، (۳) معلم، (۴) مدرسه و (۵) برنامهٔ درسی. تحقیق ماروکو^۸ (۲۰۲۱) نشان می‌دهد که متغیرهای مربوط به دانش‌آموزان^۹ همسان‌ترین^{۱۰} متغیرهای تبیین‌کننده^{۱۱} عملکرد خواندن در بین کشورهای شرکت‌کننده هستند؛ به این معنا که در اغلب کشورهای اثر مشابهی بر عملکرد خواندن داشته‌اند. نامداری پژمان و کیامنش (۱۳۹۰) بدین نتیجه رسیدند که عوامل شناختی فردی (خودپندارهٔ خواندن و نگرش به خواندن) رابطهٔ مثبت و معنی‌داری با ابعاد درک مطلب دانش‌آموزان دارند. زارع و نقش (۱۳۹۵) نیز به نتیجه‌ای مشابه با پژوهش نامداری پژمان و کیامنش (۱۳۹۰) دست یافتند. آن‌ها بیان کردند که متغیرهای خودپنداره و نگرش از سطح دانش‌آموز رابطهٔ مستقیم و معنی‌داری با عملکرد خواندن دارند. ما^{۱۲} و همکاران (۲۰۲۲) در پژوهش خود رابطهٔ قابل توجه و معنی‌داری را بین متغیرهای

-
1. Allington & McGill-Fronzen
 2. Iyengar & Ball
 3. International association for the Evaluation of Educational Achievement (IEA)
 4. literacy experience
 5. Lenkeit
 6. Progress in International Reading Literacy Study (PIRLS)
 7. Mullis & Martin
 8. Marôco
 9. Student-level variables
 10. The most consistent
 11. Explanatory variables
 12. Ma

مربوط به سطح دانش آموز و عملکرد خواندن ادعا کردند. همچنین لیو^{۱۳} و همکاران (۲۰۲۳) اظهار داشتند که عوامل شناختی فردی تأثیر مستقیمی بر توانایی خواندن دانش‌آموزان دارند.

چارچوب پرسش‌نامه‌های زمینه‌ای پرلز ۲۰۲۱ اطلاعاتی را که باید از طریق پرسش‌نامه‌های پرلز ۲۰۲۱ گردآوری شوند، به‌دقت بیان می‌کند و توضیحات مختصری نیز ارائه می‌دهد. این چارچوب مطابق با پنج عامل که تأثیر گسترده‌ای بر پیشرفت خواندن دانش‌آموزان دارند، سازماندهی می‌شود. این پنج مورد مذکور عبارتند از: بافتار خانه، بافتار مدرسه، بافتار کلاس درس، ویژگی‌های دانش‌آموز و بافتار ملی (مولیس و مارتین، ۲۰۱۹). مطالعه حاضر، با توجه به مسئله آن و شکاف پژوهشی مستخرج از ادبیات پژوهشی، با تمرکز بر ویژگی‌های دانش‌آموز (یعنی، علاقه دانش‌آموزان به خواندن، اعتمادبه‌نفس دانش‌آموزان در خواندن و آشنایی با استفاده از دستگاه‌های دیجیتال) انجام می‌شود.

با وجود پیشرفت‌های چشمگیر در تحلیل داده‌های آموزشی، مانند داده‌های گردآوری‌شده در مطالعه پرلز، روش‌های سنتی آماری مانند تحلیل واریانس و رگرسیون‌های ساده اغلب در مواجهه با داده‌های چندبعدی و روابط پیچیده میان متغیرها با محدودیت‌هایی روبه‌رو هستند. این محدودیت‌ها به‌ویژه در تحلیل متغیرهای فردی دانش‌آموزان، مانند علاقه به خواندن، اعتمادبه‌نفس در خواندن و آشنایی با ابزارهای دیجیتال، که در مطالعه پرلز ۲۰۲۱ بررسی شده‌اند، آشکار می‌شوند (مولیس و مارتین، ۲۰۱۹). رویکردهای نوین مانند یادگیری ماشین^{۱۴}، با ارائه الگوریتم‌های انعطاف‌پذیر و قدرتمند، امکان تحلیل دقیق‌تر این داده‌ها و شناسایی الگوهای پنهان را فراهم می‌کنند، که این مهم می‌تواند به درک عمیق‌تری از عوامل مؤثر بر سواد خواندن منجر شود (لوان^{۱۵}، ۲۰۰۲).

یادگیری ماشین، به‌عنوان شاخه‌ای از علوم کامپیوتر، با بهره‌گیری از الگوریتم‌هایی که از داده‌ها بدون نیاز به برنامه‌ریزی صریح یاد می‌گیرند، توانایی تحلیل روابط پیچیده و چندمتغیره را دارد که در روش‌های سنتی کمتر دیده می‌شود (بای^{۱۶} و همکاران، ۲۰۱۸). در زمینه آموزش، این رویکرد می‌تواند با تحلیل داده‌های پرلز، تأثیر متغیرهای فردی مانند خودپنداره و نگرش به خواندن را با دقت بیشتری بررسی کند. برای مثال، الگوریتم‌های یادگیری ماشین می‌توانند الگوهای پنهان در داده‌های پرلز ۲۰۲۱ را شناسایی کرده و پیش‌بینی‌های دقیق‌تری از عملکرد سواد خواندن دانش‌آموزان ارائه دهند، که این امر به طراحی مداخلات آموزشی هدفمند کمک می‌کند.

با توجه به این قابلیت‌ها، مطالعه حاضر از الگوریتم رگرسیون خطی چندگانه، برای تحلیل داده‌های پرلز ۲۰۲۱ استفاده کرده است. این روش، با امکان مدل‌سازی هم‌زمان چندین متغیر پیش‌بین و کنترل متغیرهای مخدوش‌کننده^{۱۷}، دقت پیش‌بینی را افزایش می‌دهد و بینش‌های داده‌محوری را برای بهبود سواد خواندن دانش‌آموزان ایرانی فراهم می‌سازد (بندر^{۱۸}، ۲۰۰۹). در بخش‌های بعدی، جزئیات این الگوریتم و نحوه کاربرد آن در تحلیل داده‌های پرلز تشریح خواهد شد تا نشان دهد چگونه یادگیری ماشین می‌تواند تحلیل‌های آموزشی را ارتقا دهد و به شناخت دقیق‌تر عوامل مؤثر بر عملکرد خواندن کمک کند.

لازم به تأکید است که هرچند الگوریتم رگرسیون خطی چندگانه در اصل یکی از روش‌های آماری کلاسیک به شمار می‌رود، اما در این پژوهش در چارچوبی مورد استفاده قرار گرفته که هدف آن صرفاً تحلیل روابط آماری نیست، بلکه پیش‌بینی دقیق عملکرد خواندن با تکیه بر داده‌های واقعی و ارزیابی دقت مدل نیز مدنظر بوده است. به همین دلیل، بررسی پیش‌فرض‌های آماری در اینجا نه به‌عنوان الزام سنتی روش، بلکه به‌منظور اطمینان از ثبات و کیفیت مدل پیش‌بینی انجام شده است. از این منظر، استفاده از این الگوریتم در این مطالعه به‌مثابه ابزاری برای مدل‌سازی داده‌محور تلقی می‌شود که در امتداد رویکردهای نوین تحلیل، از جمله یادگیری ماشین، قابل تفسیر است.

13. Liu

14. Machine learning

15. Luan

16. Bai

17. Confounder

18. Bender

همان‌طور که اشاره شد، یادگیری ماشین یکی از شاخه‌های علوم کامپیوتر است که رایانه‌ها را قادر به «یادگیری» بدون برنامه‌نویسی دقیق می‌سازد (بای و همکاران، ۲۰۱۸). این هدف با مطالعه علمی الگوریتم‌ها و مدل‌های آماری واقع می‌شود. الگوریتم‌های یادگیری در بسیاری از برنامه‌هایی که ما به‌صورت روزانه از آن‌ها استفاده می‌کنیم، وجود دارند (ماهش^{۱۹}، ۲۰۲۰). یادگیری ماشین، در واقع، شاخه‌ای در حال تکامل از الگوریتم‌های محاسباتی است که برای تقلید از هوش انسان از طریق یادگیری از محیط اطراف طراحی شده است. این الگوریتم‌ها به‌عنوان پرکاربردترین عوامل در عصر جدید، که اصطلاحاً عصر کلان‌داده‌ها^{۲۰} نام دارد، در نظر گرفته می‌شوند. فنون مبتنی بر یادگیری ماشین در حوزه‌های متنوعی از جمله شناسایی الگو^{۲۱}، بینایی کامپیوتری^{۲۲}، مهندسی فضاپیما، امور مالی، تبلیغات و زیست‌شناسی محاسباتی^{۲۳} تا کاربردهای زیست‌پزشکی^{۲۴} و پزشکی با موفقیت به کار گرفته شده‌اند. توانایی الگوریتم‌های یادگیری ماشین در یادگیری از بافت حاضر و تعمیم آن به وظایف نادیده به نتایج بهتری منجر می‌شود (ال نقا^{۲۵} و همکاران، ۲۰۱۵). این نکته جذابیت زیادی برای متخصصین در علوم مختلف از جمله روان‌شناسی و علوم تربیتی دارد.

یادگیری نظارتی^{۲۶}، به‌عنوان یکی از مهم‌ترین شاخه‌های یادگیری ماشین، بخش قابل‌توجهی از فعالیت‌های پژوهشی در این حوزه را به خود اختصاص داده است. ویژگی اصلی این رویکرد، استفاده از داده‌های آموزشی برجسب‌گذاری شده است، به این معنا که برای هر نمونه آموزشی، پاسخ صحیح (یا برجسب خروجی) از پیش مشخص شده است (کانینگهام^{۲۷} و همکاران، ۲۰۰۸). در مطالعه حاضر، از این رویکرد از طریق الگوریتم رگرسیون خطی چندگانه برای تحلیل داده‌های پرلز ۲۰۲۱ استفاده شده است تا تأثیر متغیرهای فردی مانند علاقه به خواندن و اعتمادبه‌نفس در خواندن بر عملکرد سواد خواندن دانش‌آموزان بررسی شود. این روش امکان تحلیل دقیق روابط پیچیده میان متغیرها را فراهم می‌کند که در بخش‌های بعدی به‌طور مفصل تشریح خواهد شد.

یادگیری نظارتی شامل روش‌هایی است که برای پیش‌بینی یا طبقه‌بندی یک خروجی مورد نظر طراحی شده‌اند. مشابهت‌هایی در وظایف/اهداف گسترده رویکردهای آماری سنتی و یادگیری ماشین نظارتی وجود دارد. در عین حال، اغلب به این هم‌پوشانی توجه نمی‌شود، زیرا ادبیات پژوهش یادگیری ماشین از مجموعه اصطلاحات متفاوتی استفاده می‌کند. برای مثال، ادبیات پژوهش یادگیری ماشین به‌جای بحث و گفتگو در مورد متغیرهای پیش‌بین یا هم‌پراش^{۲۸} برای یک متغیر برون‌داد^{۲۹} یا متغیر وابسته، به ویژگی‌هایی^{۳۰} که می‌توانند برای پیش‌بینی یا طبقه‌بندی خروجی‌ها یا اهداف به کار روند، استناد می‌کند (جیانگ^{۳۱} و همکاران، ۲۰۲۰).

رگرسیون خطی چندگانه یکی از رایج‌ترین روش‌های آماری در مدل‌سازی روابط بین چند متغیر مستقل و یک متغیر وابسته پیوسته است. همان‌طور که پیش‌تر اشاره شد، در شرایطی که متغیرهای متعددی در پیش‌بینی عملکرد دخیل هستند و باید متغیرهای مخدوش‌کننده کنترل شوند، این روش قابلیت بررسی تأثیر نسبی هر متغیر را در حالی که سایر متغیرها ثابت نگه داشته شده‌اند، فراهم می‌سازد (بندر، ۲۰۰۹؛ اشنایدر^{۳۲} و همکاران، ۲۰۱۰؛ شوبر و وتر^{۳۳}، ۲۰۲۰). مدل یک رگرسیون چندگانه که حاوی ترکیب خطی متغیرهای درون‌داد است به شکل زیر می‌باشد:

-
19. Mahesh
 20. Big data
 21. Pattern recognition
 22. Computer vision
 23. Computational biology
 24. Biomedical
 25. El Naqa
 26. Supervised learning
 27. Cunningham
 28. Covariates
 29. Output variable
 30. Features
 31. Jiang
 32. Schneider
 33. Schober & Vetter

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + e$$

در حوزه‌های مختلف علمی، طیف وسیعی از شاخص‌ها و معیارهای ارزیابی عملکرد مدل‌های رگرسیون معرفی و مورد استفاده قرار گرفته‌اند؛ معیارهایی مانند میانگین خطای مطلق^{۳۴}، خطای جذر میانگین مربعات^{۳۵}، ضریب تعیین^{۳۶} و خطای نسبی مطلق^{۳۷}. باین‌حال، تاکنون معیاری واحد و مورد توافق که به‌عنوان دقیق‌ترین و کارآمدترین شاخص در همه زمین‌ها شناخته شود، وجود ندارد. در اغلب پژوهش‌ها، فهرستی از شاخص‌های ارزیابی به‌همراه توصیفاتمانند «رایج‌ترین»، «محبوب‌ترین» یا «پرکاربردترین» ارائه می‌شود؛ باین‌حال، هیچ‌یک از این شاخص‌ها به‌عنوان «بهترین» معیار ارزیابی به‌صورت قطعی معرفی نشده‌اند، چراکه انتخاب شاخص به نوع داده، هدف پژوهش و زمینه کاربرد وابسته است. بررسی‌های تحلیلی زیادی وجود دارد که ده‌ها معیار را پوشش می‌دهند (بوچکاتف^{۳۸}، ۲۰۱۹). برای ارزیابی عملکرد مدل رگرسیون چندگانه مورد استفاده در این پژوهش، از چند شاخص رایج و معتبر در ادبیات یادگیری ماشین و تحلیل داده بهره گرفته شده است. این شاخص‌ها شامل میانگین خطای مطلق، میانگین مربعات خطا^{۳۹}، خطای جذر میانگین مربعات، نمره واریانس تبیین‌شده^{۴۰} و ضریب تعیین هستند. این معیارها به منظور سنجش دقت مدل و مقایسه عملکرد آن‌ها به کار رفته‌اند.

با توجه به هدف مطالعه حاضر، یعنی بررسی تأثیر عوامل فردی دانش‌آموزان نظیر علاقه به خواندن، اعتمادبه‌نفس در خواندن و آشنایی با ابزارهای دیجیتال بر عملکرد سواد خواندن، تلاش شده است از روش‌های نوین تحلیل داده و مشخصاً الگوریتم‌های یادگیری ماشین، بهره گرفته شود. این پژوهش نه با هدف مقایسه تجربی الگوریتم‌ها، بلکه با تأکید بر انعطاف‌پذیری مفهومی این رویکرد در مواجهه با داده‌های چندمتغیره و ساختار پیچیده روابط بین متغیرها طراحی شده است. از میان روش‌های موجود، مدل رگرسیون خطی چندگانه - که از الگوریتم‌های پایه در یادگیری نظارتی به شمار می‌رود - به کار گرفته شده است تا امکان تحلیل هم‌زمان چند متغیر فردی فراهم شود. به‌علاوه، تمرکز پژوهش بر پیش‌بینی عملکرد خواندن با داده‌های واقعی بوده است و بررسی پیش‌فرض‌های آماری صرفاً برای اطمینان از کیفیت مدل انجام شده، نه به‌عنوان الزام روش‌های کلاسیک. این انتخاب، با تکیه بر ادبیات پژوهشی مرتبط با قدرت الگوریتم‌های داده‌محور در شناسایی الگوهای پنهان و کنترل متغیرهای مداخله‌گر، صورت گرفته است و می‌تواند بینشی نو برای طراحی مداخلات آموزشی فراهم سازد.

روش پژوهش

پژوهش حاضر از نوع مطالعات توصیفی-همبستگی و ضمناً یک تحلیل ثانویه^{۴۱} است. جامعه مورد بررسی در این پژوهش، شامل کلیه آزمودنی‌های مطالعه بین‌المللی پرلز در کشور ایران در سال ۲۰۲۱ میلادی است که تمام‌شماری^{۴۲} و کلیه آزمودنی‌ها نیز جهت اجرای تحلیل سرشماری^{۴۳} شدند.

برای اندازه‌گیری متغیرهای پژوهشی در مطالعه حاضر، از آزمون پیشرفت سواد خواندن و پرسش‌نامه دانش‌آموز استفاده شد. سنجش پرلز برای ارائه تصویری کامل از پیشرفت سواد خواندن دانش‌آموزان مشارکت‌کننده در هر کشور طراحی شده است. این روند شامل پیشرفت در هدف خواندن، فرایند درک مطلب و همچنین پیشرفت عمومی خواندن است. در راستای هدف ارائه تصویری کامل از درک مطلب، سنجش پرلز ۲۰۲۱، در نسخ دیجیتال و کاغذی، حاوی ۱۸ متن خواندن و سؤالات همراه (معروف به گویه‌ها) است، که

-
34. Mean absolute error (MAE)
 35. Root mean squared error (RMSE)
 36. Coefficient of determination (R^2)
 37. Relative absolute error (RAE)
 38. Botchkarev
 39. Mean squared error (MSE)
 40. Explained Variance Score
 41. Secondary data analysis
 42. Total population sampling
 43. Census

نیمی از آن‌ها خواندن برای تجربه ادبی و نیمی دیگر خواندن برای کسب و استفاده از اطلاعات را می‌سنجند (مولیس و مارتین، ۲۰۱۹).

هر دانش‌آموز مشارکت‌کننده در پرلز ۲۰۲۱ یک پرسش‌نامه دانش‌آموز را تکمیل کرده است. پرسش‌نامه دارای سؤالاتی در مورد زندگی دانش‌آموزان در خانه و مدرسه شامل اطلاعات جمعیت‌شناختی پایه، محیط خانه آن‌ها، محیط مدرسه برای یادگیری و نگرش ایشان نسبت به خواندن است. پرسش‌نامه همچنین از دانش‌آموزان سؤالاتی درباره خودکارآمدی آن‌ها برای استفاده از دستگاه‌های دیجیتال پرسیده است (رینولدز و مارتین^{۴۴}، ۲۰۲۳). در این پرسش‌نامه که از دانش‌آموزان سؤالاتی درباره خود و همچنین طرز تفکرشان پرسیده شده است، ۳ سؤال به‌عنوان نمونه^{۴۵} و پس از آن ۱۹ سؤال اصلی ذکر شده است. پرسش‌نامه مذکور به‌شیوه امتیازپارهای^{۴۶} تدوین گشته است، به این معنا که سؤالات مختلف، بسته به نوع محتوا و هدف اندازه‌گیری، دارای مقیاس‌های پاسخ‌گویی متفاوت (مثلاً سه‌گزینه‌ای، چهارگزینه‌ای یا پنج‌گزینه‌ای) هستند.

در پژوهش حاضر، سه متغیر اصلی از سطح فردی دانش‌آموزان مورد بررسی قرار گرفته‌اند: (۱) علاقه دانش‌آموزان به خواندن، (۲) اعتمادبه‌نفس آن‌ها در خواندن و (۳) آشنایی با استفاده از دستگاه‌های دیجیتال. این متغیرها بر مبنای چارچوب نظری مطالعه پرلز و با توجه به یافته‌های پیشین درباره رابطه متغیرهای شناختی، انگیزشی و فناوری‌محور با عملکرد خواندن، به‌عنوان متغیرهای مستقل پژوهش انتخاب شده‌اند. در ادامه، ابزار سنجش این متغیرها و ویژگی‌های روان‌سنجی آن‌ها معرفی می‌شود.

برای تعیین اعتبار^{۴۷} مقیاس‌های پرسش‌نامه‌های زمینه‌ای از ضریب آلفای کرونباخ استفاده شده است. ضریب آلفای متغیرهای علاقه دانش‌آموزان به خواندن، اعتمادبه‌نفس دانش‌آموزان در خواندن و آشنایی با استفاده از دستگاه‌های دیجیتال، به ترتیب، برابر با ۰/۷۴، ۰/۷۰ و ۰/۷۸ محاسبه شدند. لازم به ذکر است که با توجه به اینکه پژوهش حاضر تحلیل داده‌های ثانویه است، ضرایب آلفای کرونباخ گزارش شده از منابع رسمی منتشر شده توسط انجمن بین‌المللی ارزشیابی پیشرفت تحصیلی در مورد مطالعه پرلز ۲۰۲۱ استخراج شده‌اند. رابطه مثبت با پیشرفت خواندن یکی از ابعاد مهم روایی مقیاس‌های پرسش‌نامه‌های زمینه‌ای پرلز است. میانگین همبستگی‌ها بسته به مقیاس‌ها و در کشورهای گوناگون با یکدیگر متفاوت است (بین^{۴۸} و رینولدز، ۲۰۲۳). ضریب همبستگی متغیرهای علاقه دانش‌آموزان به خواندن، اعتمادبه‌نفس دانش‌آموزان در خواندن و آشنایی با استفاده از دستگاه‌های دیجیتال در کشور ایران، به ترتیب، برابر با ۰/۱۶، ۰/۴۲ و ۰/۲۱ بود.

انتخاب متون مناسب و ترجمه آن متون به زبان‌های مختلف در کشورهای گوناگون، یکی از چالش‌های احراز اعتبار در آزمون پیشرفت سواد خواندن پرلز ۲۰۲۱ محسوب می‌شد، چراکه اگر عناصر متنی مورد استفاده در سؤالات کشورهای مختلف با یکدیگر متفاوت بود یا ترجمه متن، اسباب تغییر قابل توجه معنای آن را فراهم می‌کرد، اعتبار آزمون زیر سؤال می‌رفت. به‌منظور اطمینان از هم‌ارزی مفهومی و فرهنگی متون، فرایندی چندمرحله‌ای اجرا شد که شامل ترجمه اولیه توسط کشورها، بازترجمه^{۴۹} به زبان انگلیسی، بررسی تخصصی توسط تیم‌های بین‌المللی و انجام اصلاحات زبانی و فنی بود. این اقدامات با هدف حفظ معنای اصلی، جلوگیری از سوگیری فرهنگی و اطمینان از مقایسه‌پذیری نتایج در کشورهای مختلف صورت گرفت (ابز^{۵۰} و همکاران، ۲۰۲۳). همچنین جهت احراز روایی متونی که برای سنجش سواد خواندن دانش‌آموزان پایه چهارم به کار می‌رود، از متخصصان برنامه درسی کشورهای مختلف کمک گرفته شد. آن‌ها نیز متونی که واجد بیشترین انطباق با معیارهای مناسب برای پایه چهارم، همچون تناسب موضوع با پایه تحصیلی،

44. Reynolds & Martin

45. Direction

46. Partial credit

47. Reliability

48. Yin

49. Backtranslation

50. Ebbs

نبود کنش افتراقی^{۵۱} در مورد جنسیت، نژاد، مذهب و... دانش‌آموزان، بودند را برای ارزیابی انتخاب کردند (فیشبین^{۵۲} و همکاران، ۲۰۲۴).

داده‌های این مطالعه پس از دریافت از وبسایت انجمن بین‌المللی ارزشیابی پیشرفت تحصیلی، ابتدا در نرم‌افزار IEA IDB Analyzer ادغام شدند. متغیر وابسته در این پژوهش، نمره سواد خواندن، براساس متغیر ASRREA01، اولین مقدار احتمالی سواد خواندن کلی (PLAUSIBLE VALUE: OVERALL READING PV1)، از پایگاه داده پرلز ۲۰۲۱ تعریف شد. این متغیر به‌عنوان یکی از پنج مقدار احتمالی ارائه شده در پرلز برای برآورد توانایی سواد خواندن دانش‌آموزان انتخاب شد، زیرا معیاری استاندارد و پرکاربرد در تحلیل‌های پرلز است و برای ساده‌سازی تحلیل‌های رگرسیون چندگانه در این مطالعه مناسب بود. انتخاب تنها یک مقدار احتمالی به‌جای میانگین‌گیری از پنج مقدار احتمالی، با هدف کاهش پیچیدگی محاسبات و تمرکز بر تحلیل روابط بین متغیرها انجام گرفت.

برای اطمینان از تعمیم‌پذیری نتایج به کل جمعیت دانش‌آموزی ایران، وزن کلی دانش‌آموزان مندرج در پایگاه داده پرلز ۲۰۲۱، با استفاده از متغیر TOTWGT، در تمامی تحلیل‌ها اعمال شد. این وزن‌ها، که برای تعدیل تفاوت‌های نمونه‌گیری و اطمینان از نمایندگی جمعیت طراحی شده‌اند، در فرآیند تحلیل‌های رگرسیون در نرم‌افزار پایتون به کار گرفته شدند تا دقت برآوردها افزایش یابد و نتایج به جامعه هدف تعمیم داده شوند. سپس محاسبه مقدار شاخص‌های توصیفی، مانند میانگین، انحراف معیار و... و پیش‌پردازش^{۵۳} داده‌ها شامل بررسی و مدیریت داده‌های گمشده و شناسایی مقادیر پرت^{۵۴} برای تحلیل رگرسیون و همچنین بررسی مفروضات رگرسیون در نرم‌افزار SPSS انجام شدند. پیاده‌سازی الگوریتم رگرسیون چندگانه مبتنی بر یادگیری ماشین نیز در نرم‌افزار پایتون انجام گرفت. تمامی متغیرهای پیش‌بین به‌شیوه هم‌زمان^{۵۵} وارد مدل شدند. دست‌کاری داده‌ها^{۵۶} با کاربست بسته‌های Pandas و Numpy، اجرای رگرسیون چندگانه یادگیری ماشین با استفاده از بسته Scikit-learn، محاسبه شاخص‌های مختلف همچون مقدار احتمال و مقدار t ضرایب رگرسیون با استفاده از بسته Statsmodels و دیداری‌سازی مدل نیز با استفاده از بسته Matplotlib انجام گرفت.

یافته‌های پژوهش

به‌منظور ارائه تصویری روشن از ویژگی‌های جمعیت‌شناختی نمونه، پیش از ورود به تحلیل متغیرهای اصلی، اطلاعات توصیفی مربوط به جنسیت و سن دانش‌آموزان شرکت‌کننده گزارش می‌شود. داده‌های مطالعه حاضر در مجموع از ۲۸۴۵ دختر (۴۷/۸۷٪) و ۳۰۹۸ پسر (۵۲/۱۳٪) دانش‌آموز تشکیل شده است. دامنه سنی مشارکت‌کنندگان در این مطالعه نیز، تقریباً از ۸ تا ۱۴ سال و میانگین سنی آن‌ها ۱۰/۲۰ سال بود. همچنین انحراف معیار این متغیر برابر با ۰/۵۰ بود. آماره‌های توصیفی مربوط به متغیرهای پیش‌بین و ملاک پژوهش، از جمله میانگین، انحراف معیار و سایر شاخص‌ها نیز در جدول ۱ فراهم آمده است.

جدول ۱. شاخص‌های توصیفی متغیرهای پیش‌بین و ملاک

متغیرها	میانگین	انحراف معیار	کمینه	بیشینه	کجی	کشیدگی
نمره سواد خواندن	۴۱۵/۵۹	۹۷/۳۵	۱۰۰/۸۶	۶۹۷/۲۳	-۰/۳۷	-۰/۰۸
اعتماد به نفس در خواندن	۱۰/۲۵	۲/۰۳	۲/۹۷	۱۳/۴۵	۰/۳۰	-۰/۹۸
علاقه‌مندی به خواندن	۱۱/۱۸	۱/۹۶	۲/۱۱	۱۴/۸۸	۰/۳۱	۰/۰۹
آشنایی با استفاده از دستگاه‌های دیجیتال	۹/۹۶	۲/۱۱	۲/۹۵	۱۴/۷۸	۰/۲۵	۱/۲۱

51. Differential functioning

52. Fishbein

53. Pre-processing

54. Outliers

55. Enter method

56. Data manipulation

همان طور که قابل مشاهده است، میانگین و انحراف معیار متغیر ملاک، نمره سواد خواندن، به ترتیب، برابر با ۴۱۵/۵۹ و ۹۷/۳۵ بود. همچنین میانگین و انحراف معیار متغیرهای پیش بین، یعنی اعتماد به نفس در خواندن، علاقه مندی به خواندن و آشنایی با استفاده از دستگاه های دیجیتال، به ترتیب، برابر با ۱۰/۲۵ و ۲/۰۳، ۱۱/۱۸ و ۱/۹۶، و ۹/۹۶ و ۲/۱۱ بود. پیش از تحلیل اصلی داده ها، آزمون همبستگی پیرسون نیز به منظور تعیین روابط میان متغیرهای پژوهشی انجام گرفته و نتایج آن در جدول ۲ فراهم آمده است.

جدول ۲. جدول همبستگی میان متغیرهای پژوهشی

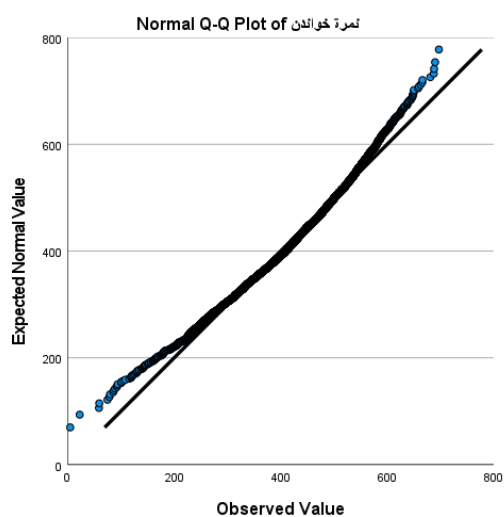
نمره سواد خواندن	آشنایی با استفاده از دستگاه های دیجیتال	علاقه دانش آموزان به خواندن	اعتماد به نفس دانش آموزان در خواندن
نمره سواد خواندن	همبستگی پیرسون	-	-
آشنایی با استفاده از دستگاه های دیجیتال	همبستگی پیرسون	۰/۲۰**	-
	معناداری (دوطرفه)	<۰/۰۰۱	-
علاقه دانش آموزان به خواندن	همبستگی پیرسون	۰/۲۱**	-
	معناداری (دوطرفه)	<۰/۰۰۱	<۰/۰۰۱
اعتماد به نفس دانش آموزان در خواندن	همبستگی پیرسون	۰/۲۰**	۰/۳۳**
	معناداری (دوطرفه)	<۰/۰۰۱	<۰/۰۰۱

** همبستگی در سطح ۰/۰۱ (به صورت دوطرفه) معنادار است.

شش مفروضه مهم که بایستی پیش از انجام تحلیل رگرسیون خطی چندگانه بررسی شوند عبارتند از: بهنجاری توزیع داده ها، هم خطی چندگانه^{۵۷}، بررسی داده های پرت، مدیریت داده های گم شده، بهنجاری توزیع باقی مانده ها و هم واریانس^{۵۸}. به دلیل بالا بودن حجم نمونه در مطالعه حاضر، به جای آزمون های استنباطی برای بررسی بهنجاری توزیع داده ها از شاخص های توصیفی کجی و کشیدگی استفاده شد، زیرا در صورت کاربرد آزمون های آمار استنباطی، همچون کولموگروف-اسمیرنوف^{۵۹}، افزایش خطای نوع اول باعث می شود که بهنجاری توزیع داده ها به اشتباه رد شود. مطابق با اطلاعات مندرج در جدول ۲، مقدار آماره های کجی و کشیدگی تمامی متغیرهای پژوهشی بین ۲- و ۲+ است، لذا تخطی معنادار از مفروضه بهنجاری توزیع داده ها صورت نگرفته است (جرج و

57. Multicollinearity
58. Homoscedasticity
59. Kolmogorov-smirnov

مالری ۶۰، ۲۰۲۱). همچنین نمودارهای چندک-چندک^{۶۱} و هیستوگرام^{۶۲} متغیر ملاک پژوهش، نمره سواد خواندن، به ترتیب، در اشکال ۱ و ۲ جهت بررسی دیداری فراهم آمده است.



شکل ۱. نمودار چندک-چندک متغیر ملاک

همان طور که از نمودار فوق پیدا است، توزیع نمرات سواد خواندن در بین دانش آموزان تقریباً بهنجار است. انحراف اندکی که از خط مرجع^{۶۳} در این نمودار مشاهده می شود، با در نظر گرفتن حجم بالای نمونه قابل چشم پوشی است. چراکه در اندازه های بالای نمونه، حتی اختلافات بسیار اندک با توزیع نرمال، می تواند سبب انحراف قابل توجه از خط مرجع در نمودار چندک-چندک شود (فاکس^{۶۴}، ۲۰۱۵).

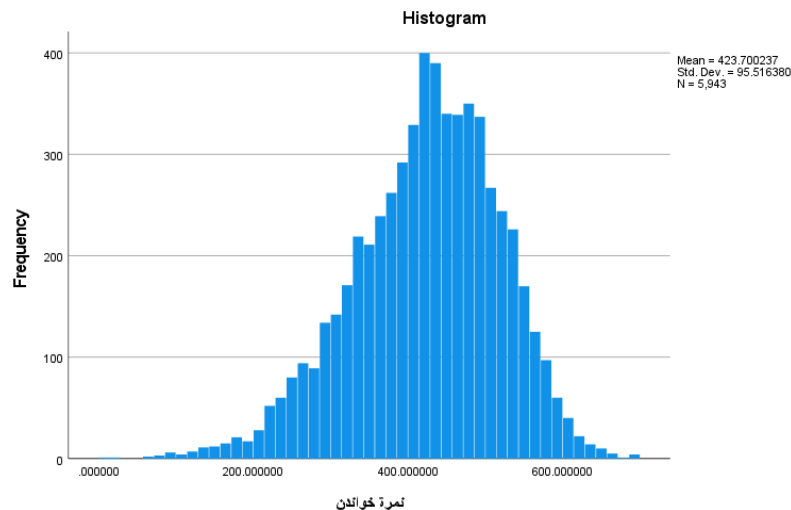
60. George & Mallery

61. Q-Q plot

62. Histogram

63. Reference line

64. Fox



شکل ۲. نمودار هیستوگرام توزیع متغیرهای ملاک

با توجه به نمودار هیستوگرام نمره خواندن نیز مجدداً می‌توان دریافت که این متغیر از توزیعی نسبتاً بهنجار برخوردار است. برای بررسی مفروضه هم‌خطی چندگانه، مقدار آمارگان تحمل^{۶۵} و عامل تورم واریانس^{۶۶} تمامی متغیرهای پژوهشی محاسبه شد. همان‌طور که در جدول ۳ مشخص است، در هیچ‌کدام از آن‌ها تخطی از مفروضه مذکور مشاهده نگشت. با توجه به آزمون آماری به‌کاررفته در مطالعه حاضر، برای شناسایی و حذف داده‌های پرت از مجموعه داده‌ها از روش فاصله ماھالانوبیس استفاده شد. بدین ترتیب، پس از محاسبه فاصله ماھالانوبیس نقاط داده‌ها، مشاهداتی که فاصله‌ای بیش از مقدار بحرانی داشتند، از مجموعه داده‌ها حذف شدند. یکی دیگر از مفروضات این نوع تحلیل، بررسی و مدیریت داده‌های گم‌شده است. این داده‌ها باید در مرحله پیش‌پردازش داده‌ها حذف و یا ترجیحاً جانهی^{۶۷} شوند. در این پژوهش نیز به‌منظور استفاده حداکثری از داده‌های گردآوری‌شده، از حذف مقادیر گم‌شده پرهیز و ارزش‌های آن‌ها، با توجه به مشخصه‌های داده‌ها، با استفاده از شیوه درون‌یابی خطی^{۶۸} جانهی شد. بهنجاری توزیع باقی‌مانده‌ها، یکی از مفروضات آماری تحلیل رگرسیون است. این مفروضه بیانگر آن است که برای اعتبار نتایج حاصل از تحلیل، بایستی توزیع باقی‌مانده‌های مدل تقریباً بهنجار باشد. البته لازم به ذکر است که تحلیل رگرسیون خطی تاحدودی نسبت به این مفروضه مقاوم^{۶۹} است. نمودارهای هیستوگرام و چندک-چندک باقی‌مانده‌های مدل رگرسیون، به‌ترتیب، در اشکال ۳ و ۴ قابل مشاهده است.

جدول ۳. جدول شاخص‌های هم‌خطی چندگانه متغیرهای پیش‌بین در مدل رگرسیون چندگانه

متغیرهای پیش‌بین	ضریب تحمل	عامل تورم واریانس
اعتمادبه‌نفس در خواندن	۰/۸۷	۱/۱۴
علاقه‌مندی به خواندن	۰/۸۷	۱/۱۵
آشنایی با استفاده از دستگاه‌های دیجیتال	۰/۹۴	۱/۰۷

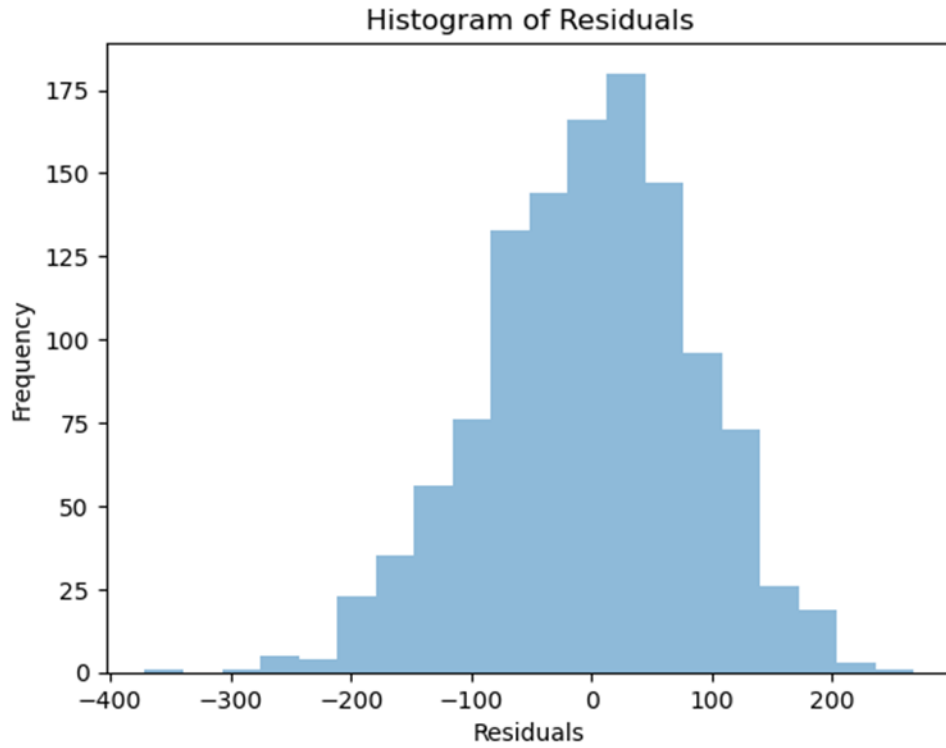
65. Tolerance

66. Variance inflation factor (VIF)

67. Imputation

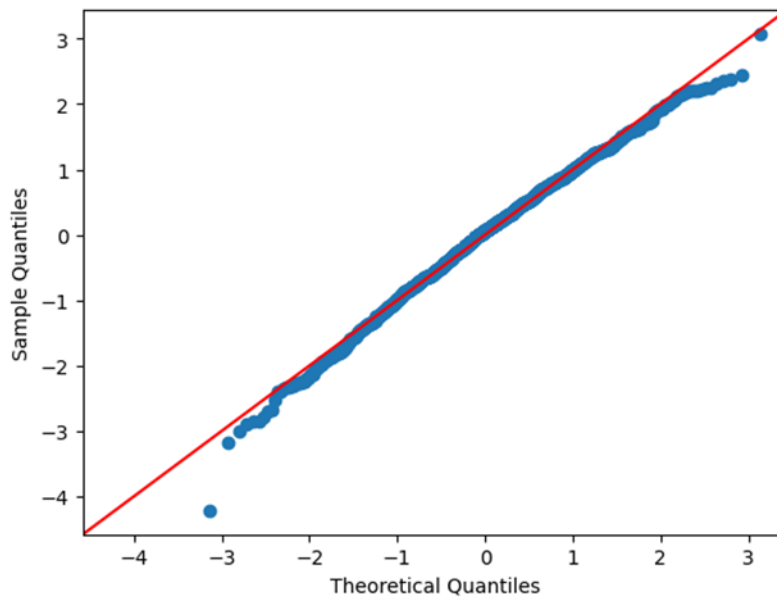
68. Linear interpolation

69. Robust



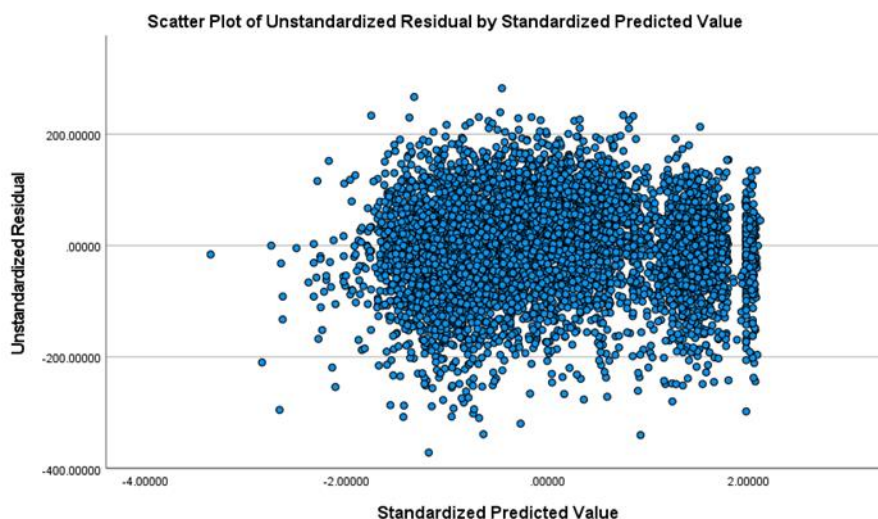
شکل ۳. نمودار هیستوگرام باقی‌مانده‌های مدل رگرسیون

همان‌طور که قابل مشاهده است، نمودار هیستوگرام باقی‌مانده‌ها در مدل رگرسیون مبتنی بر یادگیری ماشین، تقریباً بهنجار است.



شکل ۴. نمودار چندک-چندک باقی‌مانده‌های مدل رگرسیون

مطابق نمودار فوق نیز پیروی توزیع باقی مانده‌ها در مدل رگرسیون مبتنی بر یادگیری ماشین تأیید می‌گردد. انحراف اندک نقاط از خط مرجع نیز به علت حجم بالای نمونه بوده و مشکل ساز نیست. هم‌واریانس به معنای ثبات واریانس در میان مقادیر باقی مانده است. در واقع، مجموعه داده‌هایی را هم‌واریانس گویند که میزان واریانس خطا در آن‌ها تاحدودی در سطوح مختلف متغیرهای مستقل یکسان باشد. مفهوم مقابل آن، ناهم‌واریانس^{۷۰}، در شرایطی رخ می‌دهد که میان سطوح مختلف متغیر مستقل و مقادیر خطا، روند خاصی مشاهده شود (عیسی و نادال^{۷۱}، ۲۰۱۱). شکل ۵ وضعیت این مفروضه را در پژوهش حاضر نمایش می‌دهد.



شکل ۵. نمودار پراکنش مقادیر باقی مانده در مقابل مقادیر پیش‌بینی شده

با نظر به شکل ۴، می‌توان مشاهده کرد که درمجموع روند یا الگویی کلی میان مقادیر باقی مانده و پیش‌بینی شده قابل تصور نیست. بنابراین می‌توان این‌گونه نتیجه گرفت که از مفروضه هم‌واریانس در مجموعه داده‌های مورد استفاده در این مطالعه تخطی قابل توجهی صورت نگرفته است.

تحلیل‌های آماری رگرسیون در این پژوهش با استفاده از نرم‌افزار پایتون، نسخه ۳/۱۱ انجام شده‌اند. جدول ۴ حاوی مقادیر ضریب تعیین، مقدار احتمال و سایر شاخص‌های مرتبط با مدل توسعه‌یافته توسط یادگیری ماشین است.

جدول ۴. جدول خلاصه مدل رگرسیون خطی چندگانه مبتنی بر یادگیری ماشین

مقدار احتمال	آماره F	درجه آزادی	خطای استاندارد برآورد	ضریب تعیین تعدیل شده	ضریب تعیین	همبستگی چندگانه	مدل
<۰/۰۰۱	۴۳۰/۴	۳	۱۳۳۹/۵۷	۰/۱۸	۰/۱۹	۰/۴۳	نمره سواد خواندن

70. Heteroscedasticity

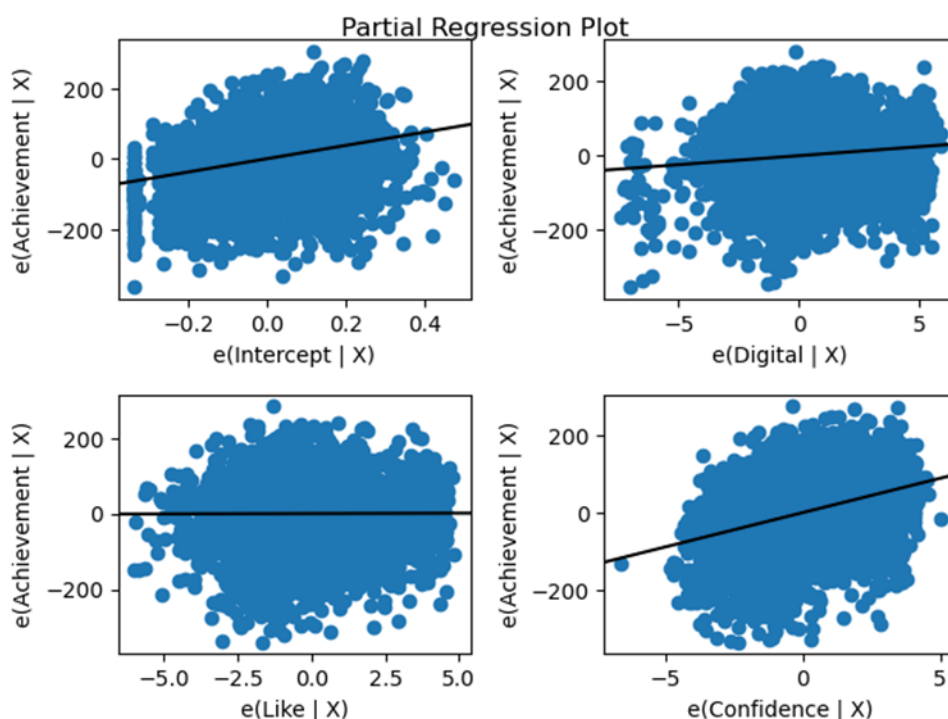
71. Issa & Nadal

با استناد به نتایج مندرج در جدول ۴، متغیرهای نگرش دانش‌آموزان (یعنی علاقه دانش‌آموزان به خواندن، اعتمادبه‌نفس آن‌ها در خواندن و آشنایی با استفاده از دستگاه‌های دیجیتال)، می‌توانند با ۹۹ درصد اطمینان نمره سواد خواندن دانش‌آموزان را به‌طور معنی‌داری پیش‌بینی کنند. همچنین با در نظر گرفتن مقدار ضریب تعیین مدل در جدول بالا، متغیرهای فوق‌الذکر می‌توانند به‌طور کلی تقریباً ۱۷٪ از پراکندگی نمره سواد خواندن را پیش‌بینی کنند. جدول ضرایب رگرسیون مدل در جدول ۵ فراهم آمده است.

جدول ۵. جدول ضرایب رگرسیون

متغیرها	ضریب برآورد	خطای استاندارد	مقدار t	مقدار احتمال
علاقه دانش‌آموزان به خواندن	۰/۰۵	۰/۶۵	۰/۰۸	۰/۹۴
اعتمادبه‌نفس دانش‌آموزان در خواندن	۱۸/۸۵	۰/۶۲	۳۰/۲۱	<۰/۰۰۱
آشنایی با استفاده از دستگاه‌های دیجیتال	۵/۷۶	۰/۵۷	۱۰/۰۵	<۰/۰۰۱

همان‌گونه که از آماره‌های جدول ۵ مشخص است، متغیرهای اعتمادبه‌نفس دانش‌آموزان در خواندن و آشنایی با استفاده از دستگاه‌های دیجیتال، برخلاف علاقه دانش‌آموزان به خواندن، می‌توانند با اطمینان بیش از ۹۹٪ نمره سواد خواندن را پیش‌بینی کنند. برای بررسی تأثیر انفرادی متغیرهای پیش‌بین بر متغیر ملاک، نمودارهای رگرسیون جزئی^{۷۲} مربوطه رسم و در شکل ۶ فراهم آمده است.



شکل ۶. نمودارهای رگرسیون جزئی

نمودارهای رسم شده نیز مؤید نتایج حاصل از تحلیل هستند. با بررسی دیداری خطوط برازش یافته می‌توان متوجه شد که متغیر اعتمادبه‌نفس دانش‌آموزان در خواندن دارای بیشترین میزان شیب در میان متغیرهای پیش‌بین است. متغیرهای آشنایی با استفاده از دستگاه‌های دیجیتال و علاقه دانش‌آموزان به خواندن، به ترتیب، در جایگاه‌های بعدی قرار دارند.

برای ارزیابی دقت پیش‌بینی مدل از شاخص‌های میزان میانگین خطای مطلق، خطای میانگین مربعات، خطای جذر میانگین مربعات، نمره واریانس تبیین شده و ضریب تعیین استفاده شد که به ترتیب، برابر با $۶۹/۸۴$ ، $۷۶۸۱/۳۶$ ، $۸۷/۶۴$ ، $۰/۱۷$ و $۰/۱۷$ محاسبه شدند. مقادیر نسبتاً پایین MAE و RMSE نشانگر میزان خطای قابل قبول در پیش‌بینی عملکرد خواندن است و ضریب تعیین مدل مشخص می‌کند که مدل توانسته است حدود ۱۷% از واریانس متغیر ملاک را تبیین کند.

از آنجایی که تقسیم‌بندی تصادفی^{۷۳} داده‌ها می‌تواند به تفاوت در نتایج مدل منجر شود، به منظور افزایش تکرارپذیری^{۷۴} تحلیل و کنترل اثر تغییرات ناشی از تقسیم تصادفی داده‌ها، مقدار پارامتر حالت تصادفی^{۷۵} در اجرای مدل رگرسیون مبتنی بر یادگیری ماشین روی عدد ۴ تنظیم شد. با این کار، هر بار که مدل اجرا شود، داده‌ها به همان صورت قبلی به مجموعه‌های آموزش و ارزیابی تقسیم خواهند شد و نتایج مدل تغییر نخواهد کرد. نسبت تقسیم داده‌ها نیز بر اساس پیشینه پژوهشی ۸۰ درصد برای آموزش و ۲۰ درصد برای ارزیابی در نظر گرفته شد. به عبارت دیگر، ۸۰٪ از نقاط داده‌ها به مجموعه آموزشی و ۲۰٪ به مجموعه ارزیابی اختصاص داده شد تا عملکرد مدل بر روی داده‌های دیده‌نشده سنجیده شود.

نتیجه‌گیری

در مدل توسعه یافته برای نمره سواد خواندن، متغیرهای نگرش دانش‌آموزان شامل علاقه به خواندن، اعتمادبه‌نفس در خواندن و آشنایی با استفاده از ابزارهای دیجیتال، در مجموع توانستند نمره سواد خواندن آن‌ها را به‌طور معنی‌داری پیش‌بینی کنند. با این حال، تحلیل‌های جداگانه نشان داد که علاقه به خواندن، برخلاف دو متغیر دیگر، تأثیر معنی‌داری بر نمره سواد خواندن نداشت. این یافته با برخی پژوهش‌ها مانند گسکه و اوزولا^{۷۶} (۲۰۰۸) که علاقه به خواندن را عامل مؤثر بر موفقیت خواندن معرفی کرده‌اند، متفاوت است. این تفاوت می‌تواند به دلیل تفاوت در تعریف مفهومی متغیر، ابزارهای اندازه‌گیری یا زمینه فرهنگی پژوهش‌ها باشد. همچنین ممکن است علاقه به خواندن به تنهایی برای بهبود مهارت‌های درک مطلب کافی نباشد و به عواملی چون فرصت، منابع آموزشی یا سبک‌های یادگیری وابسته باشد.

در مقابل، آشنایی با ابزارهای دیجیتال به‌عنوان متغیری معنی‌دار در پیش‌بینی نمره سواد خواندن ظاهر شد. این یافته با نتایج پژوهش‌هایی مانند چن^{۷۷} و همکاران (۲۰۱۵)، چو^{۷۸} و همکاران (۲۰۲۱) و لوپز^{۷۹} و همکاران (۲۰۲۲) هم‌راستا است. استفاده مؤثر از فناوری‌های دیجیتال می‌تواند به تجربه خواندن متنوع‌تری منجر شود و مهارت تحلیل و پردازش متون را افزایش دهد. همچنین، اعتمادبه‌نفس در خواندن نیز به‌عنوان قوی‌ترین عامل پیش‌بین، نقشی کلیدی در عملکرد سواد خواندن ایفا کرده است. این متغیر می‌تواند بازتابی از تجربه موفق قبلی، انگیزش درونی و خودکارآمدی دانش‌آموز باشد که همگی با عملکرد واقعی در خواندن مرتبط‌اند. یافته‌ای مشابه نیز در پژوهش قائدامینی و همکاران (۱۳۹۳) گزارش شده است؛ آن‌ها با استفاده از داده‌های پرلز ۲۰۰۶ نشان دادند که خودپنداره و نگرش دانش‌آموزان به خواندن تأثیر معنی‌داری بر پیشرفت سواد خواندن دارد، که این نتایج، هم‌راستا با نقش اعتمادبه‌نفس در مطالعه حاضر قابل تفسیر است.

73. Random splitting

74. Repeatability

75. Random state

76. Geske & Ozola

77. Chen

78. Cho

79. Lopes

یافته‌های این پژوهش همچنین قابلیت الگوریتم رگرسیون چندگانه، به‌عنوان یکی از روش‌های یادگیری ماشین، را در پیش‌بینی عملکرد تحصیلی دانش‌آموزان نشان می‌دهد. این الگوریتم، به‌ویژه در شرایطی که داده‌ها گسترده و متغیرها متعدد هستند، با امکان تحلیل هم‌زمان عوامل مختلف و شناسایی الگوهای پیچیده میان آن‌ها، مزیتی نسبت به روش‌های سنتی آماری دارد. یادگیری ماشین، برخلاف تحلیل‌های کلاسیک که وابسته به پیش‌فرض‌هایی نظیر نرمال بودن توزیع یا استقلال متغیرها هستند، می‌تواند بدون این محدودیت‌ها و با انعطاف بیشتر، عملکرد دقیق‌تری در مدل‌سازی و پیش‌بینی داشته باشد. در پژوهش حاضر، به‌کارگیری این روش ظرفیت بالقوه‌ای برای کاهش میزان خطای پیش‌بینی و افزایش توان تبیین مدل فراهم کرده است. چنین دقتی در تحلیل می‌تواند سیاست‌گذاران آموزشی را در شناسایی دقیق‌تر عوامل خطر، تدوین مداخلات هدفمند و ارتقای اثربخشی برنامه‌های آموزشی یاری رساند.

ادغام یادگیری ماشین در ارزیابی‌های آموزشی می‌تواند تحولی در نظام‌های ارزشیابی ملی و بین‌المللی ایجاد کند. برخلاف شیوه‌های سنتی که اغلب محدود به آزمون‌های استاندارد و شاخص‌های محدود هستند، یادگیری ماشین امکان تحلیل جامع‌تری از عملکرد، سوابق، و نیازهای یادگیری دانش‌آموزان فراهم می‌سازد. این رویکرد می‌تواند در بازطراحی برنامه درسی، تخصیص منابع، و انتخاب روش‌های آموزشی مؤثر نقشی اساسی ایفا کند.

در نهایت، یافته‌های این پژوهش نشان داد که از میان متغیرهای بررسی‌شده، «اعتمادبه‌نفس در خواندن» و «آشنایی با ابزارهای دیجیتال» سهم معنی‌داری در پیش‌بینی عملکرد خواندن دارند، درحالی‌که «علاقه به خواندن» در مدل نهایی معنی‌دار نبود. با توجه به این نتیجه، پیشنهاد می‌شود در مطالعات آینده، نقش واسطه‌ای یا تعاملی علاقه به خواندن در کنار سایر متغیرهای انگیزشی، از جمله اعتمادبه‌نفس و خودکارآمدی، به‌صورت دقیق‌تری بررسی شود. همچنین از آنجاکه این مطالعه از مدل رگرسیون خطی بهره گرفته است، به پژوهشگران توصیه می‌شود از مدل‌های غیرخطی یا مبتنی بر یادگیری ماشین، مانند درخت تصمیم یا شبکه‌های عصبی، برای تحلیل ساختارهای پیچیده‌تر و بررسی روابط احتمالی پنهان استفاده کنند. در نهایت، استفاده از مجموعه‌داده‌های طولی یا آمیخته، می‌تواند در آزمون پایداری الگوهای پیش‌بینی مؤثر باشد.

علاوه بر پیشنهادها، پژوهشی، نتایج این مطالعه می‌تواند در سطح اجرا نیز مورد توجه قرار گیرد. با توجه به اهمیت مؤلفه‌هایی همچون اعتمادبه‌نفس در خواندن و خودکارآمدی دیجیتال، طراحی مداخلات آموزشی مدرسه‌محور برای تقویت این عوامل توصیه می‌شود. همچنین، تجهیز کلاس‌های درس به ابزارهای دیجیتال و آموزش هدفمند معلمان برای بهره‌گیری از این فناوری‌ها، می‌تواند در بهبود سواد خواندن مؤثر باشد.

در مجموع، این یافته‌ها می‌توانند مبنایی برای طراحی سیاست‌های آموزشی مبتنی بر داده و توسعه برنامه‌هایی باشند که هم به ارتقای سواد خواندن دانش‌آموزان کمک می‌کنند و هم نقش عوامل انگیزشی و فناورانه را در یادگیری مدرسه‌ای پررنگ‌تر می‌سازند.

منابع

حسن‌آبادی، حمیدرضا، طلایی، ابراهیم، سیدمیرزایی جهقی، آزاده، و برارپور، گلرخ. (۱۳۹۶). شواهد تجربی از تعامل متن و خواننده در دانش‌آموزان جهشی پایه چهارم: آیا سواد خواندن معیاری برای تسریع تحصیلی محسوب می‌شود؟. *روانشناسی تحولی: روانشناسان ایرانی*، ۱۴(۵۴)، ۱۴۶-۱۳۵. https://journals.iau.ir/article_539236.html

زارع، حسین، و نقش، سیمین. (۱۳۹۵). بررسی سهم متغیرهای سطح دانش آموز و معلم بر عملکرد خواندن دانش‌آموزان پایه چهارم براساس داده‌های پرلز ۲۰۰۶. *دوماهنامه علمی - پژوهشی رهیافتی نو در مدیریت آموزشی*، ۷(۲۶)، ۵۲-۳۹. <https://dorl.net/dor/20.1001.1.20086369.1395.7.26.3.4>

سادات‌رضایی، مهناز، کیامنش، علیرضا، و زهرا، نقش. (۱۳۹۳). بررسی سهم متغیرهای سطح دانش آموز و مدرسه بر عملکرد خواندن دانش‌آموزان براساس داده‌های پرلز ۲۰۰۶. *فصلنامه تعلیم و تربیت*، ۳۰(۴)، ۱۴۰-۱۲۳. <http://qjoe.ir/article-1-189-fa.html>

قائدامینی، رقیه، کیامنش، علیرضا، و قربانی، رقیه. (۱۳۹۳). رابطه بین وضعیت اقتصادی-اجتماعی خانواده، فعالیت‌های خواندن در خانه، خودپنداره و نگرش دانش‌آموزان به خواندن با عملکرد سواد خواندن دانش‌آموزان (بر اساس مطالعات پرلز ۲۰۰۶). *پژوهش در برنامه‌ریزی درسی (دانش و پژوهش در علوم تربیتی-برنامه‌ریزی درسی)*، ۱۱(۴۱)، ۸۸-۱۰۳. <https://sanad.iau.ir/Journal/jsre/Article/898018>

نامداری پژمان، مهدی، و کیامنش، علیرضا. (۱۳۹۰). ارتباط عوامل شناختی فردی و خانوادگی با ابعاد درک مطلب خواندن دانش‌آموزان چهارم ابتدایی شرکت‌کننده در مطالعه پرلز ۲۰۰۶. *مطالعات برنامه‌ریزی درسی*، ۵(۲۰)، ۵۷-۳۷. <https://sid.ir/paper/101172/fa>

Allington, R. L., & McGill-Franzen, A. M. (2021). Reading volume and reading achievement: A review of recent research. *Reading Research Quarterly*, 56(S1), 231–238. <https://doi.org/10.1002/rrq.404>

Bai, Y., Liu, J., Wang, S., & Yang, F. (2018). Machine learning applied to star-galaxy-QSO classification and stellar effective temperature regression. *The Astronomical Journal*, 157(1), Article 9. <https://doi.org/10.3847/1538-3881/aaf009>

Bender, R. (2009). Introduction to the use of regression models in epidemiology. In M. Verma (Ed.), *Cancer epidemiology: Methods in molecular biology* (Vol. 471, pp. 179-195). Humana Press. https://doi.org/10.1007/978-1-59745-416-2_9

Botchkarev, A. (2019). Performance metrics (error measures) in machine learning regression, forecasting and prognostics: Properties and typology. *Interdisciplinary Journal of Information, Knowledge, and Management*, 14, 45–79. <https://doi.org/10.48550/arXiv.1809.03006>

- Chen, G., Kumar, V., Huang, R., & Kong, S. C. (Eds.). (2015). *Emerging issues in smart learning*. Springer Berlin Heidelberg. <https://doi.org/10.1007/978-3-662-44188-6>
- Cho, B. Y., Hwang, H., & Jang, B. G. (2021). Predicting fourth grade digital reading comprehension: A secondary data analysis of (e) PIRLS 2016. *International Journal of Educational Research*, 105, Article 101696. <https://doi.org/10.1016/j.ijer.2020.101696>
- Cunningham, P., Cord, M., & Delany, S. J. (2008). Supervised learning. In M. Cord & P. Cunningham (Eds.), *Machine learning techniques for multimedia: Case studies on organization and retrieval* (pp. 21-49). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-540-75171-7_2
- Ebbs, D., Flicop, S., Hidalgo, M. M., & Netten, A. (2023). Systems and instrument verification in PIRLS 2021. In M. von Davier, I. V. S. Mullis, B. Fishbein, & P. Foy (Eds.), *Methods and procedures: PIRLS 2021 technical report* (pp. 5.1-5.24). Boston College, TIMSS & PIRLS International Study Center. <https://doi.org/10.6017/lse.tpisc.tr2103.kb2485>
- El Naqa, I., Li, R., & Murphy, M. J. (Eds.). (2015). *Machine learning in radiation oncology: Theory and applications*. Springer International Publishing. <https://doi.org/10.1007/978-3-319-18305-3>
- Fishbein, B., Yin, L., & Foy, P. (2024). *PIRLS 2021 user guide for the international database* (2nd ed.). Boston College, TIMSS & PIRLS International Study Center. <https://pirls2021.org/data>
- Fox, J. (2015). *Applied regression analysis and generalized linear models*. Sage Publications.
- George, D., & Mallery, M. (2021). *SPSS for Windows step by step: A simple guide and reference* (17th ed.). Routledge. <https://doi.org/10.4324/9781003205333>
- Geske, A., & Ozola, A. (2008). Factors influencing reading literacy at the primary school level. *Problems of Education in the 21st Century*, 6, 71-77. <https://www.scientiasocialis.lt/pec/node/112>
- Issa, M. A., & Nadal, K. L. (2011). Homoscedasticity. In S. Goldstein & J. A. Naglieri (Eds.), *Encyclopedia of child behavior and development* (p. 752). Springer. https://doi.org/10.1007/978-0-387-79061-9_1382
- Iyengar, S., & Ball, D. (2007). *To read or not to read: A question of national consequence*. Washington, DC: National Endowment for the Arts.

<https://www.arts.gov/impact/research/publications/read-or-not-read-question-national-consequence>

- Jiang, T., Gradus, J. L., & Rosellini, A. J. (2020). Supervised machine learning: A brief primer. *Behavior Therapy, 51*(5), 675-687. <https://doi.org/10.1016/j.beth.2020.05.002>
- Lenkeit, J., Chan, J., Hopfenbeck, T. N., & Baird, J. A. (2015). A review of the representation of PIRLS related research in scientific journals. *Educational Research Review, 16*, 102-115. <https://doi.org/10.1016/j.edurev.2015.10.002>
- Liu, D., Wang, L., Xu, Z., Li, M., Joshi, R. M., Li, N., & Zhang, X. (2023). Understanding Chinese children's word reading by considering the factors from cognitive, psychological and ecological factors. *Contemporary Educational Psychology, 73*, Article 102163. <https://doi.org/10.1016/j.cedpsych.2023.102163>
- Lopes, J., Oliveira, C., & Costa, P. (2022). Determinantes escolares y de los estudiantes en el rendimiento lector: Un análisis multinivel con estudiantes portugueses. *Revista de Psicodidáctica, 27*(1), 29-37. <https://doi.org/10.1016/j.psicod.2021.05.001>
- Luan, J. (2002). Data mining and knowledge management in higher education: Potential applications (ED474143). ERIC. <https://eric.ed.gov/?id=ED474143>
- Ma, L., Xiao, L., & Hau, K. T. (2022). Teacher feedback, disciplinary climate, student self-concept, and reading achievement: A multilevel moderated mediation model. *Learning and Instruction, 79*, Article 101602. <https://doi.org/10.1016/j.learninstruc.2022.101602>
- Mahesh, B. (2020). Machine learning algorithms: A review. *International Journal of Science and Research (IJSR), 9*(1), 381-386. <https://doi.org/10.21275/ART20203995>
- Marôco, J. (2021). What makes a good reader? Worldwide insights from PIRLS 2016. *Reading and Writing, 34*(1), 231-272. <https://doi.org/10.1007/s11145-020-10068-8>
- Mullis, I. V. S., & Martin, M. O. (Eds.). (2019). *PIRLS 2021 assessment frameworks*. Boston College, TIMSS & PIRLS International Study Center. <https://timssandpirls.bc.edu/pirls2021/frameworks/>
- Reynolds, K., & Martin, M. O. (2023). Updating the PIRLS 2021 instruments for describing the contexts for student learning. In M. von Davier, I. V. S. Mullis, B. Fishbein, & P. Foy (Eds.),

Methods and procedures: PIRLS 2021 technical report (pp. 2.1-2.8). Boston College, TIMSS & PIRLS International Study Center. <https://doi.org/10.6017/lse.tpisc.tr2102.kb8382>

Schneider, A., Hommel, G., & Blettner, M. (2010). Linear regression analysis: Part 14 of a series on evaluation of scientific publications. *Deutsches Ärzteblatt International*, 107(44), 776-782. <https://doi.org/10.3238/arztebl.2010.0776>

Schober, P., & Vetter, T. R. (2020). Confounding in observational research. *Anesthesia & Analgesia*, 130(3), 636-643. <https://doi.org/10.1213/ANE.0000000000004627>

Yin, L., & Reynolds, K. A. (2023). Creating and interpreting the PIRLS 2021 context questionnaire scales. In M. von Davier, I. V. S. Mullis, B. Fishbein, & P. Foy (Eds.), *Methods and procedures: PIRLS 2021 technical report* (pp. 15.1-15.161). Boston College, TIMSS & PIRLS International Study Center. <https://doi.org/10.6017/lse.tpisc.tr2103.kb6994>