



تحلیل الگوهای یادگیری و پیش‌بینی عملکرد تحصیلی دانش‌آموزان در آموزش مجازی با بهره‌گیری از الگوریتم‌های داده‌کاوی

عاطفه السادات قریشی ^۱ یونس وکیل الرعایا ^{۲*} فریده حق شناس کاشانی ^۳ علی اکبر امین بیدختی ^۴	تاریخ چاپ نهایی: ۱ آبان ۱۴۰۵ تاریخ چاپ اولیه: ۸ خرداد ۱۴۰۵ تاریخ پذیرش: ۶ خرداد ۱۴۰۵ تاریخ بازنگری: ۲۷ اردیبهشت ۱۴۰۵ تاریخ ارسال: ۹ اسفند ۱۴۰۴	شیوه استناددهی: قریشی، عاطفه السادات، وکیل الرعایا، یونس، حق شناس کاشانی، فریده، و امین بیدختی، علی اکبر. (۱۴۰۵). تحلیل الگوهای یادگیری و پیش‌بینی عملکرد تحصیلی دانش‌آموزان در آموزش مجازی با بهره‌گیری از الگوریتم‌های داده‌کاوی. یادگیری هوشمند و تحول مدیریت، ۴(۴)، ۲۴-۱.
---	--	---

چکیده

هدف این پژوهش شناسایی الگوهای یادگیری دانش‌آموزان در محیط آموزش مجازی و پیش‌بینی عملکرد تحصیلی آنان با استفاده از الگوریتم‌های داده‌کاوی و یادگیری ماشین بود. این پژوهش با رویکرد داده‌محور و با بهره‌گیری از روش‌های داده‌کاوی آموزشی انجام شد. جامعه آماری شامل دانش‌آموزان دوره دوم متوسطه استان سمنان بود که از میان آنان داده‌های مربوط به ۴۰۰ دانش‌آموز فعال در محیط آموزش مجازی گردآوری شد. داده‌ها از سامانه‌های آموزش برخط، پایگاه‌های اطلاعاتی مدارس و فرم‌های الکترونیکی استخراج گردید و شامل متغیرهایی نظیر خودتنظیمی یادگیری، مهارت‌های دیجیتالی، حمایت خانوادگی، کیفیت بازخورد معلم، انگیزش تحصیلی و عملکرد قبلی بود. پس از مرحله پاکسازی، نرمال‌سازی و حذف داده‌های ناقص، الگوریتم K-Means برای خوشه‌بندی الگوهای یادگیری و الگوریتم‌های جنگل تصادفی، XGBoost، شبکه عصبی MLP، درخت تصمیم، KNN، SVM، رگرسیون لجستیک و Naive Bayes برای پیش‌بینی عملکرد تحصیلی مورد استفاده قرار گرفتند. ارزیابی مدل‌ها با شاخص‌های Recall، Accuracy، Precision و F1-Score انجام شد. نتایج خوشه‌بندی سه الگوی اصلی یادگیری شامل «یادگیرندگان مستقل و موفق»، «یادگیرندگان در معرض خطر» و «یادگیرندگان با پتانسیل نهفته» را شناسایی کرد. خوشه اول دارای خودتنظیمی، مهارت دیجیتالی و حمایت خانوادگی بالا بود، در حالی که خوشه دوم با ضعف در مهارت‌های خودتنظیمی و شرایط نامناسب یادگیری مشخص شد. الگوریتم جنگل تصادفی با دقت ۹۲ درصد، Recall برابر با ۹۲٫۸ درصد و F1-Score معادل ۹۱ درصد بهترین عملکرد را در پیش‌بینی عملکرد تحصیلی نشان داد. الگوریتم XGBoost در رتبه دوم و شبکه عصبی مصنوعی در رتبه سوم قرار گرفتند. تحلیل اهمیت ویژگی‌ها نشان داد که متغیرهای خودتنظیمی یادگیری، کیفیت بازخورد معلم و حمایت خانوادگی بیشترین تأثیر را بر عملکرد تحصیلی دانش‌آموزان دارند. یافته‌های پژوهش نشان داد که الگوهای یادگیری دانش‌آموزان در محیط‌های آموزش مجازی ناهمگن بوده و ترکیبی از عوامل شناختی، انگیزشی، تعاملی و محیطی در موفقیت تحصیلی آنان نقش دارد. همچنین، الگوریتم‌های داده‌کاوی به‌ویژه جنگل تصادفی، توانایی بالایی در تحلیل داده‌های آموزشی و پیش‌بینی عملکرد تحصیلی دارند و می‌توانند مبنایی برای طراحی آموزش‌های شخصی‌سازی‌شده و تصمیم‌گیری هوشمند در نظام‌های آموزشی مجازی فراهم سازند.

واژگان کلیدی: آموزش مجازی، یادگیری آنلاین، داده‌کاوی آموزشی، پیش‌بینی عملکرد تحصیلی، یادگیری ماشین، الگوریتم جنگل تصادفی

مشخصات نویسندگان:

۱. گروه مدیریت، واحد قزوین، دانشگاه آزاد اسلامی، قزوین، ایران
۲. گروه مدیریت، واحد مهدیشهر، دانشگاه آزاد اسلامی، مهدیشهر، ایران
۳. گروه مدیریت، واحد تهران مرکزی، دانشگاه آزاد اسلامی، تهران، ایران
۴. گروه مدیریت، دانشگاه سمنان، سمنان، ایران

پست الکترونیکی: y.vakil@semnaniau.ac.ir

© ۱۴۰۵ تمامی حقوق انتشار این مقاله متعلق به



نویسنده است.

انتشار این مقاله به‌صورت دسترسی آزاد مطابق با گواهی CC BY-NC 4.0 صورت گرفته است.

Analysis of Learning Patterns and Prediction of Students' Academic Performance in Virtual Education Using Data Mining Algorithms

Atefeh Sadat Ghoreyshi ¹ Younos Vakil Alroaia ^{2*} Farideh Haghshenas Kashani ³ Ali Akbar Amin beidokhti ⁴	Submit Date: 28 February 2026 Revise Date: 17 May 2026 Accept Date: 27 May 2026 Initial Publish: 29 May 2026 Final Publish: 23 October 2026	How to cite: Ghoreyshi, A. S., Vakil Alroaia, Y., Haghshenas Kashani, F., & Amin Beidokhti, A. A. (2026). Analysis of Learning Patterns and Prediction of Students' Academic Performance in Virtual Education Using Data Mining Algorithms. <i>Intelligent Learning and Management Transformation</i> , 4(4), 1-24.
---	---	--

Abstract

The present study aimed to identify students' learning patterns in virtual education environments and predict their academic performance using data mining and machine learning algorithms. This study employed a data-driven approach using educational data mining techniques. The statistical population consisted of secondary school students in Semnan Province, Iran, from whom data related to 400 students actively participating in virtual learning environments were collected. Data were extracted from online learning management systems, school databases, and electronic questionnaires and included variables such as self-regulated learning, digital skills, family support, teacher feedback quality, academic motivation, and prior academic performance. After preprocessing procedures including data cleaning, normalization, and removal of incomplete records, the K-Means algorithm was used for clustering learning patterns. In addition, Random Forest, XGBoost, MLP neural network, Decision Tree, SVM, KNN, Logistic Regression, and Naive Bayes algorithms were applied to predict academic performance. Model performance was evaluated using Accuracy, Recall, Precision, and F1-Score indices. The clustering analysis identified three major learning patterns: "independent and successful learners," "at-risk learners," and "learners with hidden potential." The first cluster demonstrated high levels of self-regulation, digital competence, and family support, whereas the second cluster was characterized by weak self-regulation skills and inadequate learning conditions. Among predictive models, the Random Forest algorithm achieved the best performance with 92% accuracy, 92.8% recall, and an F1-score of 91%. XGBoost ranked second, followed by the MLP neural network. Feature importance analysis revealed that self-regulated learning, teacher feedback quality, and family support were the most influential predictors of students' academic performance. The findings demonstrated that students' learning patterns in virtual education environments are highly heterogeneous and influenced by a combination of cognitive, motivational, interactive, and environmental factors. Furthermore, data mining algorithms, particularly Random Forest, showed high capability in analyzing educational data and predicting academic performance. These approaches can facilitate personalized learning design and intelligent decision-making in virtual education systems.

Keywords: Virtual education, Online learning, Educational data mining, Academic performance prediction, Machine learning, Random Forest algorithm

Authors' Information:

y.vakil@semnaniau.ac.ir

1. Department of Management, Qa.C., Islamic Azad University, Qazvin, Iran
2. Department of Management, MaSh.C., Islamic Azad University, Mahdishahr, Iran
3. Department of business Administration, Ct.C., Islamic Azad University, Central Tehran, Iran
4. Department of Management, Semnan University, Iran



© 2026 the authors. This is an open access article under the terms of the [CC BY-NC 4.0 License](https://creativecommons.org/licenses/by-nc/4.0/).

مقدمه

در دهه‌های اخیر، توسعه شتابان فناوری‌های دیجیتال و گسترش ابزارهای ارتباطی نوین، نظام‌های آموزشی را با تحولی بنیادین مواجه ساخته است. آموزش دیگر محدود به کلاس‌های سنتی و تعاملات حضوری نیست، بلکه فناوری‌های اطلاعاتی و ارتباطی بستری را فراهم کرده‌اند که یادگیری در هر زمان و مکان امکان‌پذیر باشد. تحول دیجیتال در آموزش نه تنها شیوه‌های انتقال دانش را تغییر داده، بلکه مفاهیم مرتبط با تعامل آموزشی، نقش معلم، مشارکت یادگیرنده و ارزیابی تحصیلی را نیز دگرگون کرده است (Jiao et al., 2025). در این میان، آموزش مجازی به عنوان یکی از مهم‌ترین جلوه‌های تحول آموزشی، به سرعت در نظام‌های آموزشی جهان گسترش یافته و به بخشی جدایی‌ناپذیر از فرآیند یاددهی-یادگیری تبدیل شده است (Jiang et al., 2025). آموزش مجازی با بهره‌گیری از پلتفرم‌های آنلاین، سامانه‌های مدیریت یادگیری و ابزارهای دیجیتال، امکان ارائه محتوا، تعامل آموزشی و ارزشیابی یادگیرندگان را در محیط‌های الکترونیکی فراهم می‌کند و می‌تواند فرصت‌های یادگیری انعطاف‌پذیر و شخصی‌سازی شده را توسعه دهد (Haleem et al., 2022). اهمیت این تحول به‌ویژه در دوران همه‌گیری کووید-۱۹ آشکار شد؛ زمانی که بسیاری از کشورها برای حفظ تداوم آموزش، ناچار به استفاده گسترده از بسترهای آموزش مجازی شدند (Alam et al., 2022). در این شرایط، آموزش مجازی نه تنها به عنوان یک راه‌حل اضطراری، بلکه به عنوان الگویی نوین برای آینده آموزش مورد توجه قرار گرفت (Imran et al., 2025).

با وجود گسترش سریع آموزش مجازی، پژوهش‌ها نشان می‌دهند که انتقال از آموزش سنتی به محیط‌های دیجیتال همواره با چالش‌هایی همراه بوده است. یکی از مهم‌ترین این چالش‌ها، تفاوت در میزان آمادگی فناوری، مهارت‌های دیجیتال و توانایی خودتنظیمی دانش‌آموزان است که می‌تواند کیفیت یادگیری را تحت تأثیر قرار دهد (van de Werfhorst et al., 2022). شکاف دیجیتال میان مدارس و دانش‌آموزان، دسترسی نابرابر به اینترنت و ابزارهای فناوری، و تفاوت در مهارت‌های استفاده از سامانه‌های آموزشی، از جمله عواملی هستند که موجب نابرابری آموزشی در محیط‌های مجازی می‌شوند (Rahbar Karbasdehi & Rahbar Karbasdehi, 2021). در ایران نیز تجربه آموزش مجازی در دوران کرونا نشان داد که بسیاری از مدارس و خانواده‌ها آمادگی کافی برای استفاده مؤثر از آموزش آنلاین را نداشته‌اند و این موضوع در برخی موارد به کاهش کیفیت یادگیری و افت تحصیلی منجر شده است (Salimi et al., 2021). گزارش‌های وزارت آموزش و پرورش نیز نشان می‌دهد که در دوره آموزش مجازی، میانگین عملکرد تحصیلی دانش‌آموزان در برخی دروس کاهش یافته و سطح مشارکت آموزشی در بسیاری از مدارس کمتر از حد انتظار بوده است (Ministry of Education of, 2022). افزون بر این، پژوهش‌ها نشان داده‌اند که موفقیت در محیط‌های آموزش مجازی به عوامل متعددی از جمله انگیزش تحصیلی، مهارت‌های شناختی و فراشناختی، تعامل آموزشی و کیفیت بازخورد معلم وابسته است (Mousizadeh Ahaghi, 2025).

عملکرد تحصیلی یکی از شاخص‌های کلیدی در ارزیابی کیفیت یادگیری و اثربخشی نظام‌های آموزشی است و تحت تأثیر عوامل فردی، اجتماعی، انگیزشی و آموزشی قرار دارد (Costa et al., 2024). در محیط‌های آموزش مجازی، این عوامل پیچیدگی بیشتری پیدا می‌کنند؛ زیرا یادگیرندگان باید در غیاب نظارت مستقیم معلم، توانایی مدیریت زمان، برنامه‌ریزی، خودتنظیمی و تعامل مؤثر با محیط یادگیری را داشته باشند (Zhao et al., 2025). مطالعات نشان داده‌اند که راهبردهای یادگیری خودتنظیمی، مانند برنامه‌ریزی، پایش پیشرفت و مدیریت شناختی، رابطه معناداری با موفقیت تحصیلی در محیط‌های آنلاین دارند (Jin et al., 2023). از سوی دیگر، تعامل آموزشی و مشارکت فعال دانش‌آموزان نیز از عوامل تعیین‌کننده کیفیت یادگیری در آموزش مجازی محسوب می‌شود (Kolahi Zadeh et al., 2025). پژوهش‌های انجام‌شده در حوزه آموزش آنلاین نشان می‌دهند که حضور آموزشی، اجتماعی و شناختی در محیط‌های یادگیری مجازی می‌تواند سطح مشارکت دانش‌آموزان را افزایش دهد و موجب بهبود عملکرد تحصیلی آنان شود (Alshammari & Alrehaili, 2025). همچنین، محیط‌های یادگیری تعاملی و مبتنی بر ارزشیابی خلاق می‌توانند درگیری شناختی و انگیزش تحصیلی دانش‌آموزان را ارتقا دهند (Ghanaati & Jamebozorg, 2023).

در سال‌های اخیر، داده‌کاوی آموزشی و تحلیل یادگیری به عنوان دو رویکرد نوین برای تحلیل داده‌های آموزشی و بهبود تصمیم‌گیری در نظام‌های آموزشی مورد توجه قرار گرفته‌اند (Cerezo et al., 2024). داده‌کاوی آموزشی با استفاده از الگوریتم‌های یادگیری ماشین و روش‌های تحلیل داده، امکان استخراج الگوهای پنهان در رفتار یادگیرندگان و پیش‌بینی عملکرد تحصیلی آنان را فراهم می‌کند (Ataei et al., 2024). این رویکرد می‌تواند اطلاعات ارزشمندی درباره نحوه تعامل دانش‌آموزان با محیط یادگیری، سطح مشارکت، کیفیت انجام تکالیف و عوامل مؤثر بر موفقیت یا افت تحصیلی ارائه دهد (Maniyan et al., 2024). پژوهش‌ها نشان داده‌اند که الگوریتم‌های یادگیری ماشین مانند جنگل تصادفی، شبکه‌های عصبی، XGBoost و ماشین بردار پشتیبان توانایی بالایی در پیش‌بینی عملکرد تحصیلی دارند (Al Shanableh et al., 2026). همچنین، استفاده از مدل‌های داده‌کاوی می‌تواند زمینه طراحی آموزش شخصی‌سازی‌شده و ارائه بازخوردهای هدفمند را فراهم سازد (Azhineh et al., 2025). در این میان، تحلیل الگوهای یادگیری دانش‌آموزان اهمیت ویژه‌ای دارد؛ زیرا دانش‌آموزان در محیط‌های آموزش مجازی رفتارهای یادگیری متفاوتی از خود نشان می‌دهند و این تفاوت‌ها می‌تواند بر موفقیت تحصیلی آنان اثرگذار باشد (Song & Vermunt, 2021). مطالعات بین‌المللی و داخلی متعددی به بررسی نقش داده‌کاوی آموزشی و یادگیری ماشین در تحلیل عملکرد تحصیلی پرداخته‌اند. برای مثال، پژوهش یاغچی نشان داد که الگوریتم‌های یادگیری ماشین قادرند عملکرد تحصیلی دانشجویان را با دقت قابل‌قبولی پیش‌بینی کنند (Yagci, 2022). همچنین، وحید و همکاران با استفاده از مدل‌های یادگیری عمیق، نشان دادند که داده‌های حاصل از سامانه‌های یادگیری مجازی می‌توانند پیش‌بینی دقیقی از عملکرد تحصیلی ارائه دهند (Waheed et al., 2020). سالاری و همکاران نیز در پژوهشی بر روی داده‌های دانشگاهی گزارش کردند که الگوریتم درخت تصمیم و جنگل تصادفی عملکرد مناسبی در پیش‌بینی موفقیت تحصیلی دانشجویان دارند (Salari et al., 2022).

(2024). افزون بر این، تحلیل الگوهای ترتیبی تعامل یادگیرندگان در محیط‌های آموزشی نشان داده است که استفاده از ابزارهای راهنما و بازخوردهای تعاملی می‌تواند احتمال موفقیت در یادگیری را افزایش دهد (Verstege et al., 2024). از سوی دیگر، کیم و همکاران نشان دادند که نحوه استفاده دانش‌آموزان و معلمان از پلتفرم‌های آموزشی با میزان ماندگاری و مشارکت دانش‌آموزان در محیط یادگیری رابطه مستقیم دارد (Kim et al., 2020).

در ایران نیز پژوهش‌های مختلفی به بررسی پیامدهای آموزش مجازی و نقش فناوری در یادگیری پرداخته‌اند. خسروی و همکاران آموزش مجازی را بستری برای گسترش عدالت آموزشی و افزایش انعطاف‌پذیری یادگیری معرفی کرده‌اند (Khosravi et al., 2025). زارع‌نسب و جامه‌بزرگ نیز بر فرصت‌ها و چالش‌های استفاده از فناوری‌های نوین و هوش مصنوعی در آموزش تأکید کرده‌اند (Zare Nasab & Jamebozorg, 2025). زینلی و دانشمند نشان دادند که توسعه مهارت‌های حرفه‌ای و فناوری‌محور معلمان می‌تواند کارایی آموزشی را بهبود بخشد (Zeinali & Daneshmand, 2025). همچنین، مطالعاتی مانند پژوهش رستم‌زاده و همکاران نشان داده‌اند که استفاده از محیط‌های یادگیری دیجیتال می‌تواند انگیزش پیشرفت و عملکرد تحصیلی دانش‌آموزان را ارتقا دهد (Rostamzadeh et al., 2023). تربتی‌نژاد و همکاران نیز بر اهمیت تعامل آموزشی، توانمندسازی معلمان و بهبود زیرساخت‌های فنی در اثربخشی آموزش مجازی تأکید کرده‌اند (Torbati Nejad et al., 2022). جعفری‌اقدم و همکاران در مطالعه‌ای بر دانشجویان دانشگاه تهران گزارش کردند که آموزش مجازی می‌تواند در برخی شرایط موجب افزایش انعطاف‌پذیری یادگیری و بهبود معدل تحصیلی شود (Jafari Aghdam et al., 2023). افزون بر این، مطالعات نشان داده‌اند که انگیزش شناختی-رفتاری و تعامل فعال دانش‌آموزان با محیط یادگیری، نقش مهمی در مشارکت و موفقیت تحصیلی آنان دارد (Singh et al., 2022). با وجود گسترش پژوهش‌ها در حوزه آموزش مجازی و داده‌کاوی آموزشی، هنوز شکاف‌های پژوهشی قابل توجهی وجود دارد. بسیاری از مطالعات داخلی بیشتر بر دیدگاه‌ها و تجربیات معلمان و دانش‌آموزان تمرکز داشته‌اند و کمتر از داده‌های واقعی رفتاری و الگوریتم‌های داده‌کاوی برای تحلیل الگوهای یادگیری استفاده کرده‌اند (Salimi et al., 2021). همچنین، بخش قابل توجهی از پژوهش‌های داده‌کاوی آموزشی در سطح آموزش عالی انجام شده و ویژگی‌های خاص دانش‌آموزان نوجوان کمتر مورد توجه قرار گرفته است (Towner et al., 2023). این در حالی است که نوجوانان از نظر شناختی، عاطفی و اجتماعی دارای ویژگی‌هایی متفاوت از دانشجویان هستند و الگوهای یادگیری آنان در محیط‌های دیجیتال می‌تواند متفاوت باشد. افزون بر این، بسیاری از سامانه‌های آموزش مجازی همچنان بر اساس رویکردهای یکسان طراحی شده‌اند و تفاوت‌های فردی دانش‌آموزان را در نظر نمی‌گیرند (Dietrich et al., 2021). در نتیجه، نیاز به پژوهش‌هایی وجود دارد که با استفاده از روش‌های داده‌محور، بتوانند الگوهای یادگیری دانش‌آموزان را شناسایی کرده و عوامل مؤثر بر عملکرد تحصیلی آنان را تحلیل کنند. چنین

پژوهش‌هایی می‌توانند زمینه طراحی محیط‌های یادگیری هوشمند، آموزش شخصی‌سازی‌شده و تصمیم‌گیری مبتنی بر داده در نظام آموزشی را فراهم سازند (Rezai et al., 2025).

بر این اساس، هدف پژوهش حاضر تحلیل الگوهای یادگیری دانش‌آموزان در محیط آموزش مجازی و پیش‌بینی عملکرد تحصیلی آنان با استفاده از الگوریتم‌های داده‌کاوی و یادگیری ماشین است.

روش‌شناسی

روش‌شناسی این پژوهش بر پایه رویکرد داده‌محور و با استفاده از الگوریتم‌های داده‌کاوی و یادگیری ماشین طراحی شده است. هدف اصلی، تحلیل الگوهای یادگیری دانش‌آموزان در بستر آموزش مجازی و شناسایی متغیرهایی است که بیشترین اثر را بر عملکرد تحصیلی آنان دارند. از آنجا که داده‌کاوی نیازمند حجم مناسب و متنوعی از داده‌هاست، جامعه آماری شامل تمامی دانش‌آموزان دوره دوم متوسطه استان سمنان در سال تحصیلی اخیر در نظر گرفته شد و داده‌های مربوط به ۴۰۰ دانش‌آموز که به‌طور کامل در محیط آموزش مجازی فعالیت داشتند، گردآوری شد. داده‌ها از طریق ترکیبی از پایگاه‌های اطلاعاتی مدارس، سامانه‌های آموزش برخط و فرم‌های الکترونیکی تکمیل‌شده توسط دانش‌آموزان و معلمان دریافت گردید و شامل مجموعه‌ای از متغیرهای فردی، شناختی، انگیزشی، تعاملی، فنی و عملکردی بود که مشخصات کامل آن‌ها در جدول ۱ ارائه شده است.

جدول ۱. دسته‌بندی متغیرهای مورد استفاده در پژوهش

دسته اصلی متغیرها	شرح متغیرها و نمونه شاخص‌ها
خودتنظیمی یادگیری	برنامه‌ریزی مطالعه، پایش پیشرفت، توانایی مدیریت زمان
مهارت‌های دیجیتال	توانایی کار با ابزارهای آموزشی دیجیتال، استفاده از پلتفرم مجازی، تسلط بر ابزارهای یادگیری
حمایت خانوادگی	پیگیری والدین، وجود حمایت عاطفی-تحصیلی در خانه
بازخورد معلم	میزان، کیفیت و به‌موقع بودن بازخوردهای ارائه‌شده به دانش‌آموز
انگیزش تحصیلی	علاقه، پشتکار، تلاش برای موفقیت در کلاس مجازی
عملکرد قبلی	نمره سال قبل، شاخص‌های عملکرد گذشته
فضای مطالعه	کیفیت محیط یادگیری در خانه (سکوت، تمرکز، امکانات مطالعه)
تعامل‌پذیری محتوا	میزان تعاملی بودن محتوا و فعالیت‌های آموزشی ارائه‌شده

پس از گردآوری داده‌ها، مرحله پیش‌پردازش با دقت کامل انجام شد تا داده‌ها برای اعمال الگوریتم‌های داده‌کاوی آماده شوند. ابتدا داده‌های ناقص، تکراری یا ناسازگار شناسایی و اصلاح شدند؛ داده‌های گمشده با استفاده از روش‌های استاندارد نظیر میانگین‌گذاری یا تخمین مبتنی بر

همسایگی نزدیک^۱ جایگزین گردید و مقادیر پرت بر اساس روش‌های آماری شامل بازه بین‌چارکی کنترل شد. سپس برای جلوگیری از تأثیر نامتناسب مقیاس متغیرها بر عملکرد مدل‌ها، داده‌ها با روش نرمال‌سازی کمینه-بیشینه در بازه صفر تا یک نرمال‌سازی شدند. در ادامه برای کاهش پیچیدگی مدل و حذف ویژگی‌های کم‌اهمیت، از تحلیل همبستگی و شاخص‌های اهمیت متغیرها استفاده شد. در مرحله تحلیل اصلی، مجموعه‌ای از الگوریتم‌های قدرتمند داده‌کاوی به کار گرفته شد که در پژوهش‌های آموزشی از کارایی و دقت بالایی برخوردارند. برای تحلیل الگوهای پنهان یادگیری، از الگوریتم خوشه‌بندی کی-مینز (K-Means) استفاده شد و تعداد خوشه‌ها با بهره‌گیری از معیارهای آرنج (Elbow) و سیلوئت (Silhouette) به صورت بهینه تعیین گردید. به منظور پیش‌بینی عملکرد تحصیلی و مدل‌سازی روابط پیچیده میان متغیرها، علاوه بر سه الگوریتم اصلی شامل درخت تصمیم کارتی (CART)^۲، جنگل تصادفی^۳ و شبکه عصبی مصنوعی چندلایه (MLP)^۴، مجموعه‌ای از الگوریتم‌های مقایسه‌ای شامل XGBoost، ماشین بردار پشتیبان (SVM)، الگوریتم KNN، رگرسیون لجستیک و نایو بیز نیز اجرا و عملکرد آن‌ها ارزیابی شد. انتخاب این الگوریتم‌ها به دلیل ترکیب قابلیت تفسیرپذیری، دقت بالا، پشتیبانی از روابط غیرخطی و امکان مقایسه جامع مدل‌ها صورت گرفت. داده‌ها با نسبت ۷۰ درصد برای آموزش و ۳۰ درصد برای آزمون تفکیک شدند و پارامترهای مدل‌ها پس از تنظیم‌های لازم بهینه‌سازی گردید. برای نمونه، شبکه عصبی با یک لایه پنهان حاوی ۱۰۰ نرون و تابع فعال‌سازی رلو آموزش داده شد و جنگل تصادفی با مجموعه‌ای از درخت‌های تصمیم با عمق بهینه اجرا شد. برای ارزیابی عملکرد مدل‌ها از شاخص‌های دقت، بازخوانی، امتیاز اف ۱ و مساحت زیر منحنی ROC استفاده شد. در مدل شبکه عصبی، شاخص‌های MSE و R^2 نیز محاسبه گردید. همچنین زمان آموزش و نسبت دقت به سرعت برای مقایسه عملیاتی مدل‌ها در نظر گرفته شد. به منظور افزایش اعتبار نتایج، از اعتبارسنجی متقابل ده‌تایی استفاده شد. اهمیت متغیرها از طریق شاخص اهمیت ویژگی در جنگل تصادفی و تحلیل حساسیت در شبکه عصبی بررسی گردید و قواعد استخراج شده از درخت تصمیم برای تفسیر نتایج مورد استفاده قرار گرفت. تمام تحلیل‌ها با استفاده از نرم‌افزار پایتون (Python) و کتابخانه‌های استاندارد اسکیکیت‌لرن (Scikit Learn)، پانداس (Pandas)، نام‌پای (NumPy) و مت‌پلات‌لیب (Matplotlib) انجام شد و در برخی مراحل برای مقایسه نتایج از محیط نرم‌افزارهایی نظیر رپیدماینر (RapidMiner) یا وکا (Weka) نیز بهره گرفته شد. اصول اخلاق پژوهش شامل محرمانگی داده‌ها، حذف اطلاعات هویتی و اخذ رضایت مدارس و والدین رعایت شد. این فرایند امکان تحلیل دقیق داده‌ها، کشف الگوهای مؤثر و ارائه مدل قابل اتکایی برای پیش‌بینی عملکرد تحصیلی در آموزش مجازی را فراهم کرد.

¹ KNN Imputation

² Classification and Regression Tree

³ Random Forest

⁴ Multilayer Perceptron

یافته‌ها

داده‌های این پژوهش شامل اطلاعات ۴۰۰ دانش‌آموز دوره دوم متوسطه استان سمنان است که در بستر آموزش مجازی فعالیت داشته‌اند. در مرحله پیش‌پردازش، داده‌ها ابتدا مورد پاک‌سازی اولیه قرار گرفتند؛ به گونه‌ای که رکوردهای ناقص و موارد دارای مقادیر گم‌شده غیرقابل بازیابی از مجموعه داده حذف شدند. سپس توزیع متغیرها مورد بررسی قرار گرفت و وجود داده‌های پرت با استفاده از شاخص‌های آماری مبتنی بر انحراف معیار ارزیابی شد. نتایج بررسی‌ها نشان داد که داده‌ها فاقد مقادیر پرت معنادار هستند و بنابراین نیازی به حذف یا اصلاح مقادیر پرت وجود نداشت. در ادامه، به منظور هم‌مقیاس‌سازی متغیرها و جلوگیری از تأثیر نامتوازن ویژگی‌ها در الگوریتم‌های مبتنی بر فاصله، تمامی متغیرهای عددی با استفاده از روش استانداردسازی Z-score نرمال‌سازی شدند. در این روش، مقادیر هر متغیر بر اساس میانگین و انحراف معیار آن متغیر مقیاس‌بندی می‌شوند که مطابق رابطه زیر محاسبه می‌شود:

$$Z = \frac{x - \mu}{\sigma}$$

که در آن X مقدار اولیه متغیر، μ میانگین و σ انحراف معیار آن متغیر است. متغیرهای مورد استفاده طیفی از ویژگی‌های فردی-شناختی، انگیزشی، تعاملی، الگوی حضور در سامانه آموزشی، کیفیت انجام تکالیف، وضعیت دسترسی فناوریانه و عملکرد تحصیلی را پوشش می‌دهند.

به منظور شناسایی الگوهای یادگیری دانش‌آموزان در محیط آموزش مجازی، از الگوریتم خوشه‌بندی K-Means استفاده شد. با توجه به ماهیت عددی متغیرها و انجام مرحله استانداردسازی داده‌ها، این الگوریتم گزینه‌ای مناسب برای کشف ساختارهای پنهان و گروه‌بندی الگوهای مشابه در داده‌ها محسوب می‌شود. برای تعیین تعداد بهینه خوشه‌ها از دو شاخص مکمل شامل روش Elbow و شاخص Silhouette استفاده شد. نتایج تحلیل نمودار Elbow نشان داد که مقدار مجموع مربعات درون‌خوشه‌ای (Within-Cluster Sum of Squares: WCSS) تا رسیدن به سه خوشه کاهش قابل توجهی دارد، اما پس از آن نرخ کاهش به‌طور محسوسی کند می‌شود که بیانگر نقطه شکست در نمودار است. علاوه بر این، مقدار شاخص Silhouette در حالت سه خوشه‌ای بیشترین مقدار نسبی را در میان حالت‌های بررسی شده نشان داد و بیانگر تفکیک‌پذیری مناسب خوشه‌ها و انسجام درونی قابل قبول آن‌ها بود. بر این اساس، مقدار $k=3$ به عنوان تعداد بهینه خوشه‌ها انتخاب شد.

جدول ۲. نتایج خوشه‌بندی دانش‌آموزان با الگوریتم K-Means

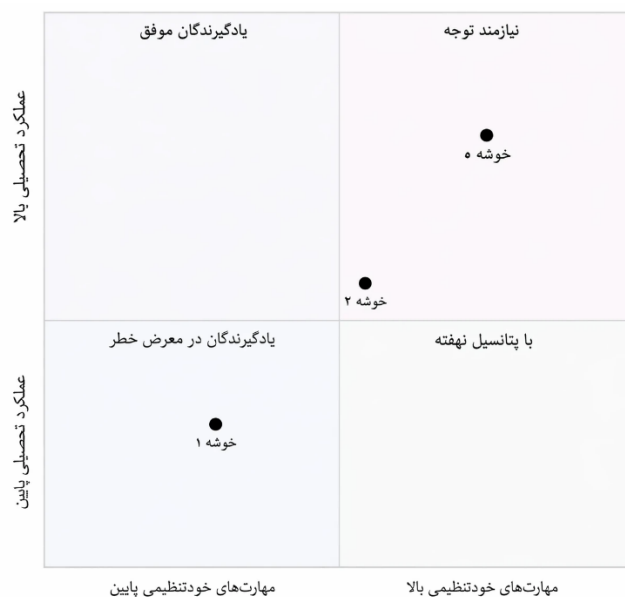
خوشه	عنوان خوشه	اندازه خوشه	پرو فایل ویژگی‌های شاخص
۰	یادگیرندگان مستقل و موفق	۳۲٪ (۱۲۸ نفر)	خودتنظیمی بالا، عملکرد قبلی بالا، مهارت‌های دیجیتالی قوی، حمایت خانوادگی بالا
۱	یادگیرندگان در معرض خطر	۲۵٪ (۱۰۰ نفر)	خودتنظیمی پایین، عملکرد قبلی متوسط به پایین، مهارت‌های دیجیتالی ضعیف، فضای مطالعه نامناسب
۲	یادگیرندگان با پتانسیل نهفته	۴۳٪ (۱۷۲ نفر)	انگیزش بالا، عملکرد قبلی متوسط، مهارت‌های دیجیتالی متوسط، اما بازخورد معلم پایین و محتوای غیرتعاملی

پس از اجرای الگوریتم K-Means با سه خوشه، میانگین متغیرهای اصلی در هر خوشه محاسبه و پروفایل رفتاری و یادگیری هر گروه تحلیل شد. نتایج خوشه‌بندی در جدول مربوطه ارائه شده است. بر اساس این نتایج، سه الگوی متمایز از یادگیری دانش‌آموزان در محیط آموزش مجازی شناسایی گردید.

خوشه اول (یادگیرندگان مستقل و موفق): این گروه حدود ۳۲ درصد از کل نمونه‌ها را شامل می‌شود. دانش‌آموزان این خوشه از سطح بالای خودتنظیمی در یادگیری، مهارت‌های دیجیتالی مناسب، مشارکت فعال در سامانه آموزشی و حمایت خانوادگی مطلوب برخوردار هستند. این ویژگی‌ها نشان می‌دهد که این دسته از دانش‌آموزان توانایی بالایی در سازگاری با محیط‌های یادگیری مجازی داشته و به‌طور مستقل فرآیند یادگیری خود را مدیریت می‌کنند.

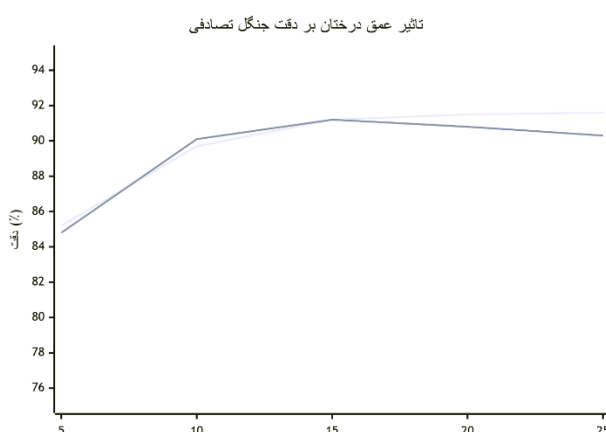
خوشه دوم (یادگیرندگان در معرض خطر): حدود ۲۵ درصد از دانش‌آموزان در این خوشه قرار دارند. اعضای این گروه از نظر مهارت‌های خودتنظیمی، توانایی‌های دیجیتالی و شرایط محیطی مطالعه در سطح پایین‌تری قرار دارند و میزان مشارکت آموزشی آن‌ها نیز محدودتر است. ترکیب این عوامل نشان‌دهنده گروهی از دانش‌آموزان است که برای بهره‌گیری مؤثر از محیط آموزش مجازی با چالش‌های بیشتری مواجه هستند. خوشه سوم (یادگیرندگان با پتانسیل نهفته): این خوشه بزرگ‌ترین گروه را تشکیل می‌دهد و حدود ۴۳ درصد از نمونه‌ها را در بر می‌گیرد. دانش‌آموزان این گروه از سطح انگیزش نسبتاً بالایی برخوردارند و از نظر توانایی‌های شناختی و مهارت‌های دیجیتالی در سطح متوسط قرار دارند، اما به دلیل دریافت بازخورد محدود از سوی معلم و تعامل کمتر با محتوای آموزشی، از ظرفیت کامل یادگیری خود در محیط مجازی بهره نمی‌برند. به‌طور کلی، نتایج خوشه‌بندی نشان می‌دهد که الگوهای یادگیری دانش‌آموزان در محیط آموزش مجازی ناهمگن است و دانش‌آموزان بر اساس ترکیبی از ویژگی‌های فردی، انگیزشی، تعاملی و مهارتی در گروه‌های متمایزی قرار می‌گیرند. شناسایی این الگوها می‌تواند زمینه‌ای برای طراحی راهبردهای آموزشی متناسب با نیازهای هر گروه از یادگیرندگان و بهبود اثربخشی آموزش مجازی فراهم کند.

شکل ۱ پراکندگی خوشه‌های به‌دست آمده از تحلیل K-Means را نشان می‌دهد. همان‌گونه که مشاهده می‌شود، خوشه اول در ناحیه‌ای قرار دارد که با سطوح بالاتر خودتنظیمی و مشارکت آموزشی مشخص می‌شود و بیانگر گروه یادگیرندگان مستقل و موفق است. در مقابل، خوشه دوم در محدوده مقادیر پایین‌تر این شاخص‌ها قرار دارد و نشان‌دهنده دانش‌آموزانی است که در بهره‌گیری از محیط آموزش مجازی با چالش مواجه‌اند. خوشه سوم نیز در موقعیتی میانی قرار گرفته و بیانگر گروهی از دانش‌آموزان با انگیزش بالا اما تعامل آموزشی محدود است که از ظرفیت یادگیری خود به‌طور کامل استفاده نمی‌کنند.



شکل ۱. پراکندگی خوشه‌های دانش‌آموزان حاصل از تحلیل خوشه‌بندی K-Means

به‌منظور پیش‌بینی عملکرد تحصیلی دانش‌آموزان در محیط آموزش مجازی از مدل جنگل تصادفی استفاده شد. این الگوریتم به دلیل دقت بالا در مسائل طبقه‌بندی، توانایی مدل‌سازی روابط غیرخطی میان متغیرها و همچنین قابلیت استخراج اهمیت ویژگی‌ها، به‌عنوان یکی از مدل‌های مناسب برای تحلیل داده‌های آموزشی انتخاب گردید. علاوه بر این، جنگل تصادفی با ترکیب مجموعه‌ای از درخت‌های تصمیم، توانایی کاهش بیش‌برازش و افزایش پایداری پیش‌بینی را دارد. برای ارزیابی عملکرد مدل، از شاخص‌های Accuracy، Precision، Recall و F1-Score استفاده شد. نتایج نشان داد که مدل جنگل تصادفی توانایی قابل قبولی در پیش‌بینی عملکرد تحصیلی دانش‌آموزان دارد و نسبت به بسیاری از مدل‌های دیگر مورد بررسی، دقت بالاتری ارائه می‌دهد (شکل ۲).



شکل ۲. تغییرات دقت مدل جنگل تصادفی در سطوح مختلف عمق درخت.

مقادیر شاخص‌های ارزیابی برای این مدل به صورت زیر محاسبه شد:

Overall Accuracy: $(۱۴۲ + ۱۳۴) / ۳۰۰ = ۹۲\%$

- Recall: $۱۴۲ / (۱۴۲ + ۱۱) = ۹۲.۸\%$

- Precision: $۱۴۲ / (۱۴۲ + ۱۳) = ۹۱.۶\%$

- F1-Score: $۲ * (۹۲.۸ * ۹۱.۶) / (۹۲.۸ + ۹۱.۶) = ۹۱.۰\%$

به منظور بررسی دقیق تر عملکرد مدل، ماتریس درهم ریختگی نیز تحلیل شد. نتایج نشان داد که بخش قابل توجهی از نمونه ها به درستی طبقه بندی شده اند و میزان خطای طبقه بندی نسبتاً پایین است. همچنین تحلیل خطاهای نوع اول و نوع دوم نشان داد که مدل در تشخیص صحیح دانش آموزان با عملکرد تحصیلی متفاوت عملکرد مناسبی داشته است.

"ماتریس درهم ریختگی - مدل جنگل تصادفی"

	پیش بینی مدل	پیش بینی مدل
مقدار واقعی	False Positive: 5	True Positive: 68
مقدار واقعی	False Negative: 7	True Negative: 48

شکل ۳: ماتریس درهم ریختگی مدل جنگل تصادفی

در فرآیند تنظیم پارامترهای مدل، عمق درخت های تصمیم در بازه ۵ تا ۲۰ مورد بررسی قرار گرفت. نتایج نشان داد که عمق ۱۵ بهترین عملکرد را از نظر دقت پیش بینی ارائه می دهد. افزایش عمق بیشتر از این مقدار نه تنها موجب بهبود قابل توجهی نشد، بلکه احتمال بیش برآزش را نیز افزایش داد.

به منظور مدل سازی روابط غیر خطی میان متغیرهای آموزشی، یک شبکه عصبی پرسپترون چندلایه (MLP) با یک لایه پنهان شامل ۱۰۰ نرون و تابع فعال ساز ReLU طراحی و آموزش داده شد. تقسیم داده ها با نسبت ۷۰٪ آموزش و ۳۰٪ آزمون انجام شد. نتایج ارزیابی مدل بر روی داده های آزمون در جدول ۳ ارائه شده است.

جدول ۳. شاخص‌های ارزیابی عملکرد مدل شبکه عصبی MLP بر روی داده‌های آزمون

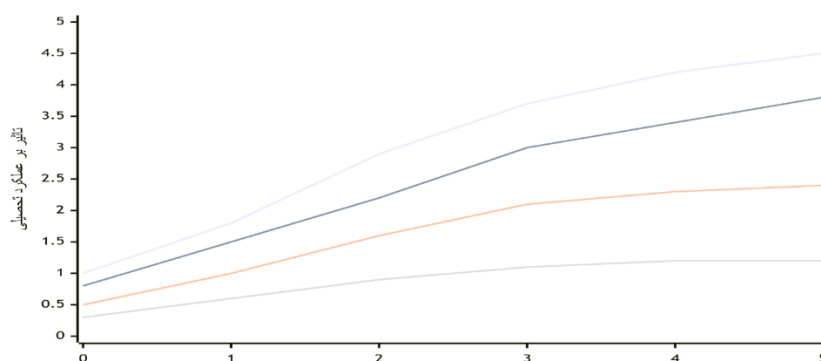
شاخص ارزیابی	نماد	مقدار
دقت طبقه‌بندی	Accuracy	۸۹.۸٪
ضریب تعیین	R ²	۰.۸۷۶
میانگین مربعات خطا	MSE	۰.۰۰۸۵

مقادیر فوق بیانگر توانایی مناسب مدل در پیش‌بینی عملکرد تحصیلی و تبیین واریانس متغیر وابسته است. در ادامه به‌منظور بررسی میزان تأثیر متغیرهای ورودی بر خروجی مدل، تحلیل حساسیت ورودی-خروجی انجام شد. نتایج نشان داد که متغیرهای خودتنظیمی یادگیری و کیفیت بازخورد معلم بیشترین سهم را در تغییرات پیش‌بینی عملکرد تحصیلی دارند. این یافته با نتایج حاصل از مدل جنگل تصادفی همسو است. همچنین شبکه عصبی الگوی غیرخطی قابل توجهی میان حمایت عاطفی خانواده و عملکرد تحصیلی شناسایی کرد؛ به گونه‌ای که برای دانش‌آموزان با سطح خودتنظیمی متوسط، افزایش حمایت خانوادگی با افزایش قابل توجه در عملکرد تحصیلی همراه است. این نتیجه نشان‌دهنده وجود اثرات تعاملی میان متغیرهای زمینه‌ای و مهارت‌های خودتنظیمی است. مقادیر حاصل از تحلیل حساسیت شبکه عصبی برای متغیرهای اصلی پژوهش در جدول ۴ ارائه شده است که نشان می‌دهد متغیر خودتنظیمی یادگیری بیشترین تأثیر را بر عملکرد تحصیلی دارد.

جدول ۴. مقادیر تحلیل حساسیت متغیرها بر عملکرد تحصیلی

سطح متغیر	خودتنظیمی یادگیری	بازخورد معلم	حمایت عاطفی خانواده	سایر عوامل
۰	۱.۰	۰.۸	۰.۵	۰.۳
۱	۱.۸	۱.۵	۰.۹	۰.۵
۲	۲.۹	۲.۲	۱.۶	۰.۹
۳	۳.۷	۳.۰	۲.۱	۱.۱
۴	۴.۲	۳.۴	۲.۳	۱.۲
۵	۴.۵	۳.۸	۲.۴	۱.۲

به‌منظور نمایش بصری روند تغییرات میزان تأثیر متغیرها در سطوح مختلف، داده‌های جدول ۴ به‌صورت نمودار خطی در شکل ۴ ارائه شده است.



شکل ۴. نتایج تحلیل حساسیت شبکه عصبی مصنوعی در پیش‌بینی عملکرد تحصیلی

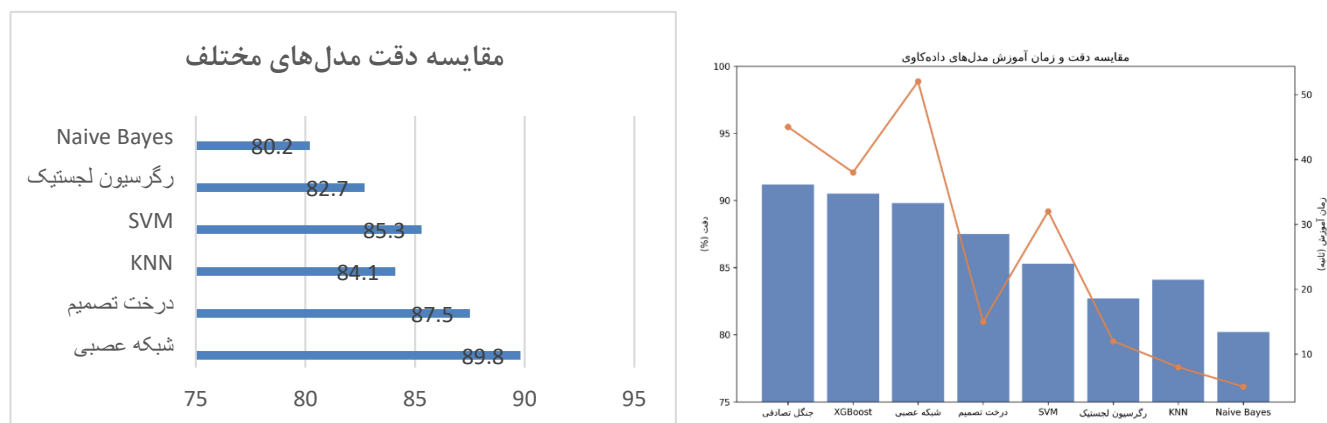
همان گونه که در شکل ۴ مشاهده می شود، متغیرهای خودتنظیمی و کیفیت بازخورد معلم بیشترین میزان اهمیت نسبی را در خروجی مدل دارند، در حالی که سایر متغیرها سهم کمتری در تغییرات پیش‌بینی شده عملکرد تحصیلی ایفا می کنند.

به منظور ارزیابی کارایی الگوریتم‌های مختلف یادگیری ماشین در تحلیل داده‌های آموزش مجازی، عملکرد چندین الگوریتم شامل جنگل تصادفی، XGBoost، شبکه عصبی مصنوعی، درخت تصمیم، KNN، ماشین بردار پشتیبان (SVM)، رگرسیون لجستیک و Naive Bayes مورد بررسی و مقایسه قرار گرفت. معیارهای ارزیابی شامل دقت (Accuracy)، بازخوانی (Recall)، دقت پیش‌بینی (Precision)، امتیاز F1 و زمان آموزش مدل‌ها بود.

جدول ۵. مقایسه عملکرد الگوریتم‌های داده‌کاوی در پیش‌بینی عملکرد تحصیلی

دقت (%)	Recall (%)	Precision (%)	F1-Score (%)	زمان آموزش (ثانیه)	مدل
۹۲	۹۲.۸	۹۱.۶	۹۱.۰	۴۵	جنگل تصادفی
۹۰.۵	۹۱.۳	۹۰.۸	۹۰.۲	۳۸	XGBoost
۸۹.۸	۹۰.۲	۸۹.۴	۸۹.۶	۵۲	شبکه عصبی
۸۷.۵	۸۸.۱	۸۶.۹	۸۷.۲	۱۵	درخت تصمیم
۸۴.۱	۸۵.۰	۸۳.۸	۸۳.۹	۸	KNN
۸۵.۳	۸۸.۲	۸۵.۴	۸۵.۱	۳۲	SVM
۸۲.۷	۸۳.۵	۸۱.۹	۸۲.۴	۱۲	رگرسیون لجستیک
۸۰.۲	۸۱.۰	۷۹.۸	۷۹.۹	۵	Naive Bayes

نتایج نشان داد که الگوریتم جنگل تصادفی با دستیابی به دقت ۹۲ درصد، Recall برابر با ۹۲/۸ درصد و F1-Score معادل ۹۱/۰ درصد، بهترین عملکرد را در میان مدل‌های مورد بررسی داشته است. این الگوریتم همچنین توانسته است تعادل مناسبی میان شاخص‌های Precision و Recall برقرار کند که نشان‌دهنده توانایی بالای آن در شناسایی صحیح الگوهای موجود در داده‌های آموزشی است. الگوریتم XGBoost نیز عملکرد قابل قبولی از خود نشان داد و با دقت ۹۰/۵ درصد در رتبه دوم قرار گرفت. شبکه عصبی مصنوعی نیز با دقت ۸۹/۸ درصد نتایج نسبتاً مناسبی ارائه داد، اگرچه زمان آموزش آن (۵۲ ثانیه) بیشتر از سایر مدل‌ها بود. در مقابل، الگوریتم‌هایی مانند درخت تصمیم، KNN و SVM عملکرد متوسطی از نظر شاخص‌های ارزیابی نشان دادند. با وجود اینکه الگوریتم KNN و Naive Bayes زمان آموزش بسیار کوتاهی داشتند (به ترتیب ۸ و ۵ ثانیه)، اما دقت آن‌ها نسبت به سایر مدل‌ها پایین‌تر گزارش شد. همچنین الگوریتم رگرسیون لجستیک و Naive Bayes کمترین میزان دقت را در میان مدل‌های مورد بررسی داشتند.



شکل ۵. مقایسه عملکرد الگوریتم‌های داده‌کاوی در پیش‌بینی عملکرد تحصیلی

به طور کلی، نتایج مقایسه نشان می‌دهد که الگوریتم‌های مبتنی بر مجموعه درخت‌ها، به‌ویژه جنگل تصادفی و XGBoost، توانایی بیشتری در مدل‌سازی روابط پیچیده میان متغیرهای آموزشی دارند. بر این اساس، الگوریتم جنگل تصادفی به عنوان مدل مناسب برای تحلیل داده‌های آموزش مجازی و شناسایی الگوهای یادگیری دانش‌آموزان انتخاب شد.

بحث و نتیجه‌گیری

پژوهش حاضر با هدف تحلیل الگوهای یادگیری دانش‌آموزان در محیط آموزش مجازی و پیش‌بینی عملکرد تحصیلی آنان با استفاده از الگوریتم‌های داده‌کاوی و یادگیری ماشین انجام شد. نتایج به‌دست آمده نشان داد که الگوهای یادگیری دانش‌آموزان در محیط‌های آموزش مجازی ناهمگن بوده و می‌توان دانش‌آموزان را بر اساس ویژگی‌های شناختی، انگیزشی، تعاملی و مهارتی در خوشه‌های متفاوتی دسته‌بندی کرد. تحلیل خوشه‌ای با استفاده از الگوریتم K-Means سه گروه اصلی شامل «یادگیرندگان مستقل و موفق»، «یادگیرندگان در معرض خطر» و «یادگیرندگان با پتانسیل نهفته» را شناسایی کرد. این یافته نشان می‌دهد که دانش‌آموزان در محیط‌های یادگیری مجازی رفتارها و الگوهای متفاوتی از مشارکت، تعامل و مدیریت یادگیری از خود نشان می‌دهند و این تفاوت‌ها می‌تواند بر موفقیت تحصیلی آنان تأثیر مستقیم داشته باشد. نتایج حاضر با یافته‌های سانگ و ورمونت همسو است که بیان کردند یادگیرندگان در محیط‌های آموزشی از الگوهای متفاوت یادگیری شامل الگوهای عمیق، سطحی و راهبردی استفاده می‌کنند و این تفاوت‌ها بر کیفیت یادگیری اثرگذار است (Song & Vermunt, 2021). همچنین، یافته‌های پژوهش حاضر با نتایج یان و همکاران هم‌راستا است که نشان دادند در محیط‌های یادگیری آنلاین، سطح تعامل و مشارکت یادگیرندگان می‌تواند الگوهای متفاوتی از رفتار یادگیری ایجاد کند (Yan et al., 2022). شناسایی خوشه یادگیرندگان مستقل و موفق نشان داد که برخورداری از مهارت‌های خودتنظیمی، توانایی مدیریت زمان، مهارت‌های دیجیتالی و حمایت خانوادگی می‌تواند زمینه موفقیت بیشتر دانش‌آموزان در محیط‌های آموزش

مجازی را فراهم سازد. این نتیجه با فراتحلیل ژائو و همکاران همسو است که راهبردهای یادگیری خودتنظیمی را از مهم‌ترین پیش‌بینی‌کننده‌های عملکرد تحصیلی در محیط‌های یادگیری آنلاین معرفی کردند (Zhao et al., 2025). همچنین، پژوهش جین و همکاران نشان داد که ابزارهای مبتنی بر هوش مصنوعی و بازخوردهای شخصی‌سازی‌شده می‌توانند توانایی خودتنظیمی یادگیرندگان را تقویت کرده و موجب ارتقای عملکرد تحصیلی شوند (Jin et al., 2023).

یافته‌های پژوهش همچنین نشان داد که گروهی از دانش‌آموزان در خوشه «یادگیرندگان در معرض خطر» قرار می‌گیرند که از نظر مهارت‌های خودتنظیمی، مهارت‌های دیجیتال، شرایط محیطی مطالعه و سطح تعامل آموزشی در وضعیت ضعیف‌تری قرار دارند. این نتیجه بیانگر آن است که آموزش مجازی برای همه دانش‌آموزان به یک اندازه اثربخش نیست و تفاوت در شرایط فردی و محیطی می‌تواند موجب نابرابری آموزشی شود. این یافته با نتایج پژوهش ون‌دی ورفهورست و همکاران همخوانی دارد که شکاف دیجیتال و نابرابری در آمادگی فناوریانه مدارس و دانش‌آموزان را از مهم‌ترین چالش‌های آموزش آنلاین معرفی کردند (van de Werfhorst et al., 2022). همچنین، رحبر کرباسدهی و رحبر کرباسدهی نیز در بررسی مشکلات آموزش مجازی در ایران به عواملی مانند ضعف زیرساخت‌ها، دسترسی نابرابر به اینترنت و پایین بودن مهارت‌های فناوری اشاره کرده‌اند (Rahbar Karbasdehi & Rahbar Karbasdehi, 2021). نتایج این پژوهش نشان می‌دهد که ضعف در مهارت‌های دیجیتالی و نبود فضای مناسب مطالعه می‌تواند مشارکت دانش‌آموزان را کاهش داده و در نهایت موجب افت عملکرد تحصیلی شود. این یافته با نتایج پژوهش کیم و همکاران همسو است که نشان دادند تعامل مستمر دانش‌آموزان با سامانه‌های یادگیری و مشارکت فعال معلمان در محیط‌های آنلاین، نقش مهمی در کاهش ریزش تحصیلی و افزایش درگیری آموزشی دارد (Kim et al., 2020). همچنین، پژوهش سلیمی و همکاران درباره شبکه آموزشی شاد نیز نشان داد که بسیاری از دانش‌آموزان و معلمان ایرانی در استفاده مؤثر از آموزش مجازی با چالش‌های فنی و آموزشی مواجه بوده‌اند (Salimi et al., 2021).

از دیگر یافته‌های مهم پژوهش حاضر، شناسایی خوشه «یادگیرندگان با پتانسیل نهفته» بود. این گروه اگرچه از انگیزش نسبتاً بالایی برخوردار بودند، اما به دلیل دریافت بازخورد محدود از سوی معلم و تعامل پایین با محتوای آموزشی، نتوانسته بودند از ظرفیت کامل یادگیری خود استفاده کنند. این نتیجه نشان می‌دهد که صرف برخورداری از انگیزش برای موفقیت در محیط‌های آموزش مجازی کافی نیست و کیفیت تعامل آموزشی و طراحی محتوا نیز نقش تعیین‌کننده‌ای در عملکرد تحصیلی دارد. این یافته با نتایج پژوهش الشمری و الرحیلی همسو است که نشان دادند حضور آموزشی، اجتماعی و شناختی در دوره‌های آنلاین می‌تواند میزان مشارکت دانش‌آموزان را افزایش دهد (Alshammari & Alrehaili, 2025). همچنین، غناتی و جامه‌بزرگ گزارش کردند که طراحی محیط‌های یادگیری مبتنی بر ارزشیابی خلاق و تعامل فعال، موجب افزایش درگیری شناختی و بهبود یادگیری دانش‌آموزان می‌شود (Ghanaati & Jamebozorg, 2023). از سوی دیگر، پژوهش رزایی و همکاران

نشان داد که استفاده از رویکردهای یادگیری فعال می‌تواند انگیزش تحصیلی را افزایش داده و نگرش مثبت‌تری نسبت به یادگیری ایجاد کند (Rezai et al., 2025). بنابراین، نتایج پژوهش حاضر تأکید می‌کند که برای ارتقای کیفیت آموزش مجازی، باید به طراحی محیط‌های تعاملی، ارائه بازخورد مؤثر و افزایش مشارکت فعال دانش‌آموزان توجه ویژه‌ای شود.

در بخش پیش‌بینی عملکرد تحصیلی، نتایج نشان داد که الگوریتم جنگل تصادفی بالاترین دقت را در مقایسه با سایر مدل‌های یادگیری ماشین داشته و توانسته است عملکرد تحصیلی دانش‌آموزان را با دقت بالایی پیش‌بینی کند. این یافته با نتایج پژوهش آل‌شناپله و همکاران همسو است که گزارش کردند الگوریتم‌های مبتنی بر مجموعه درخت‌ها، به‌ویژه جنگل تصادفی، از توانایی بالایی در تحلیل داده‌های آموزشی برخوردارند (Al Shanableh et al., 2026). همچنین، سالاری و همکاران در مطالعه‌ای بر داده‌های دانشگاهی نشان دادند که الگوریتم‌های یادگیری ماشین، به‌ویژه درخت تصمیم و جنگل تصادفی، می‌توانند پیش‌بینی دقیقی از عملکرد تحصیلی ارائه دهند (Salari et al., 2024). یافته‌های این پژوهش با مطالعه یاغچی نیز همسو است که گزارش کرد مدل‌های یادگیری ماشین قادرند با تحلیل داده‌های رفتاری و آموزشی، عملکرد تحصیلی دانش‌آموزان را با دقت قابل‌قبولی پیش‌بینی کنند (Yagci, 2022). افزون بر این، پژوهش وحید و همکاران نشان داد که استفاده از مدل‌های یادگیری عمیق بر داده‌های کلان محیط‌های یادگیری مجازی می‌تواند به پیش‌بینی دقیق عملکرد تحصیلی منجر شود (Waheed et al., 2020). برتری الگوریتم جنگل تصادفی در پژوهش حاضر را می‌توان به توانایی آن در مدل‌سازی روابط غیرخطی، مدیریت حجم بالای داده‌ها و کاهش بیش‌برازش نسبت داد. همچنین، این الگوریتم توانسته است تعادل مناسبی میان شاخص‌های Recall و Precision برقرار کند که نشان‌دهنده قدرت بالای آن در شناسایی صحیح الگوهای موجود در داده‌های آموزشی است.

نتایج تحلیل اهمیت ویژگی‌ها نیز نشان داد که متغیرهای خودتنظیمی یادگیری، کیفیت بازخورد معلم و حمایت خانوادگی بیشترین نقش را در پیش‌بینی عملکرد تحصیلی دانش‌آموزان دارند. این یافته با نتایج پژوهش موسی‌زاده آهقی همسو است که نشان داد آموزش مجازی می‌تواند مهارت‌های فراشناختی و خودتنظیمی یادگیرندگان را توسعه دهد و این مهارت‌ها با موفقیت تحصیلی رابطه مستقیم دارند (Mousizadeh Ahaghi, 2025). همچنین، ژائو و همکاران در فراتحلیل خود تأکید کردند که خودتنظیمی از مهم‌ترین عوامل موفقیت در محیط‌های یادگیری آنلاین و ترکیبی است (Zhao et al., 2025). نقش کیفیت بازخورد معلم نیز با یافته‌های الشمری و الرحیلی هماهنگ است که تعامل آموزشی و حضور معلم را عامل مهمی در افزایش درگیری یادگیرندگان معرفی کردند (Alshammari & Alrehaili, 2025). افزون بر این، یافته‌های حاضر با پژوهش تربتی‌نژاد و همکاران نیز همسو است که نشان دادند اثربخشی آموزش مجازی وابسته به توانمندسازی معلمان، تعامل آموزشی و کیفیت ارائه محتوا است (Torbati Nejad et al., 2022). همچنین، نقش حمایت خانوادگی در موفقیت تحصیلی دانش‌آموزان را می‌توان با توجه به شرایط آموزش مجازی و وابستگی بیشتر یادگیرندگان به محیط خانه تبیین کرد. در محیط‌های آموزش آنلاین، خانواده‌ها نقش مهمی در

ایجاد فضای مناسب مطالعه، حمایت عاطفی و پیگیری فعالیت‌های آموزشی دارند و همین مسئله می‌تواند بر عملکرد تحصیلی دانش‌آموزان اثرگذار باشد.

نتایج پژوهش حاضر همچنین نشان داد که آموزش مجازی، در صورت طراحی مناسب و بهره‌گیری از فناوری‌های هوشمند، می‌تواند فرصت‌هایی برای شخصی‌سازی یادگیری و بهبود تصمیم‌گیری آموزشی فراهم سازد. این نتیجه با یافته‌های پژوهش سرزو و همکاران همخوانی دارد که داده‌کاوی آموزشی را ابزاری مؤثر برای استخراج الگوهای یادگیری و توسعه علم داده‌های آموزشی معرفی کردند (Cerezo et al., 2024). همچنین، مانیان و همکاران نشان دادند که سیستم‌های پشتیبان تصمیم مبتنی بر داده‌کاوی می‌توانند به مدیران آموزشی در شناسایی عوامل مؤثر بر کارایی آموزشی کمک کنند (Maniyan et al., 2024). پژوهش حاضر نیز نشان داد که تحلیل داده‌های آموزشی می‌تواند اطلاعات ارزشمندی درباره ویژگی‌های رفتاری و آموزشی دانش‌آموزان فراهم سازد و زمینه طراحی مداخلات آموزشی هدفمند را ایجاد کند. افزون بر این، نتایج پژوهش با مطالعات خسروی و همکاران و کاظمی‌مهر و صالحی همسو است که بر ظرفیت محیط‌های یادگیری دیجیتال برای ایجاد یادگیری انعطاف‌پذیر، تعاملی و مشارکتی تأکید کرده‌اند (Kazemi Mehr & Salehi, 2023; Khosravi et al., 2025). در مجموع، یافته‌های این پژوهش نشان می‌دهد که استفاده از رویکردهای داده‌محور و الگوریتم‌های یادگیری ماشین می‌تواند به درک بهتر رفتار یادگیرندگان، شناسایی دانش‌آموزان در معرض خطر و ارتقای کیفیت آموزش مجازی کمک کند.

یکی از محدودیت‌های پژوهش حاضر آن است که داده‌ها تنها از دانش‌آموزان دوره دوم متوسطه استان سمنان گردآوری شده و بنابراین تعمیم نتایج به سایر مقاطع تحصیلی یا مناطق جغرافیایی باید با احتیاط انجام شود. همچنین، بخشی از داده‌ها مبتنی بر خودگزارشی دانش‌آموزان و معلمان بود که ممکن است تحت تأثیر سوگیری پاسخ‌دهی قرار گرفته باشد. محدود بودن متغیرهای روان‌شناختی و اجتماعی مورد بررسی و استفاده از داده‌های مقطعی نیز از دیگر محدودیت‌های پژوهش محسوب می‌شود.

پیشنهاد می‌شود پژوهش‌های آینده با استفاده از نمونه‌های گسترده‌تر و چنداستانی، به بررسی الگوهای یادگیری دانش‌آموزان در مقاطع تحصیلی مختلف بپردازند. همچنین، استفاده از داده‌های طولی می‌تواند امکان تحلیل تغییرات رفتار یادگیری در طول زمان را فراهم سازد. بررسی نقش متغیرهایی مانند سلامت روان، سبک‌های فرزندپروری، ویژگی‌های شخصیتی و سواد دیجیتال در پیش‌بینی عملکرد تحصیلی نیز می‌تواند افق‌های جدیدی برای پژوهش‌های آینده ایجاد کند. افزون بر این، پیشنهاد می‌شود در مطالعات آینده از مدل‌های پیشرفته‌تر یادگیری عمیق و هوش مصنوعی برای تحلیل داده‌های آموزشی استفاده شود.

بر اساس نتایج پژوهش حاضر، پیشنهاد می‌شود نظام آموزشی در طراحی سامانه‌های آموزش مجازی به تفاوت‌های فردی دانش‌آموزان توجه بیشتری داشته باشد و محیط‌های یادگیری شخصی‌سازی شده را توسعه دهد. همچنین، تقویت مهارت‌های خودتنظیمی، آموزش سواد دیجیتال به دانش‌آموزان

و معلمان، و ارائه بازخوردهای مستمر و هدفمند می‌تواند اثربخشی آموزش مجازی را افزایش دهد. توجه به دانش‌آموزان در معرض خطر، طراحی برنامه‌های حمایتی برای آنان، و استفاده از ابزارهای داده‌کاوی آموزشی برای شناسایی زود هنگام افت تحصیلی نیز می‌تواند در بهبود کیفیت یادگیری و کاهش نابرابری آموزشی مؤثر باشد.

مشارکت نویسندگان

در نگارش این مقاله تمامی نویسندگان نقش یکسانی ایفا کردند.

تعارض منافع

در انجام مطالعه حاضر، هیچ‌گونه تضاد منافی وجود ندارد.

موازین اخلاقی

در تمامی مراحل پژوهش حاضر اصول اخلاقی مرتبط با نشر و انجام پژوهش رعایت گردیده است.

تشکر و قدردانی

از تمامی کسانی که در انجام این پژوهش ما را همراهی کردند تشکر و قدردانی به عمل می‌آید.

Extended Abstract

Introduction

The rapid advancement of digital technologies has fundamentally transformed educational systems worldwide and reshaped the processes of teaching and learning. In recent years, virtual education has emerged as one of the most influential educational paradigms, enabling learners to access educational resources beyond traditional temporal and geographical limitations (Jiao et al., 2025). The integration of information and communication technologies into educational environments has facilitated the development of online learning platforms, learning management systems, and interactive digital environments that support flexible and learner-centered education (Jiang et al., 2025). The significance of virtual education became particularly evident during the COVID-19 pandemic, when educational institutions across the globe were forced to shift from face-to-face instruction to online learning environments (Alam et al., 2022). This sudden transformation accelerated the digitalization of education and highlighted both the opportunities and challenges associated with virtual learning (Imran et al., 2025). Research has shown that online education can enhance flexibility, accessibility, and educational innovation; however, it may also intensify inequalities related to digital readiness, technological infrastructure, and students' self-regulation skills (van de Werfhorst et al., 2022). In many countries, including Iran, the rapid implementation of virtual learning exposed deficiencies in

educational infrastructure and digital competencies among teachers and students (Rahbar Karbasdehi & Rahbar Karbasdehi, 2021). Reports from educational institutions indicated that many learners experienced reduced engagement, lower academic motivation, and decreased academic achievement during the virtual learning period (Ministry of Education of, 2022).

Academic achievement is one of the most important indicators of educational effectiveness and reflects the extent to which learners attain educational goals (Costa et al., 2024). In online learning environments, academic achievement is influenced by multiple factors, including digital literacy, motivation, interaction quality, self-regulated learning strategies, and family support (Mousizadeh Ahaghi, 2025). Studies have emphasized that students who possess stronger self-regulation abilities, such as planning, monitoring, and time management, tend to achieve better academic outcomes in online and blended learning environments (Zhao et al., 2025). Furthermore, interaction and participation in virtual educational networks significantly contribute to students' engagement and academic success (Kolahi Zadeh et al., 2025). Alshammari and Alrehaili demonstrated that teaching, cognitive, and social presence positively influence student engagement in online courses (Alshammari & Alrehaili, 2025). Similarly, innovative and interactive learning environments can enhance learners' motivation and cognitive involvement (Ghanaati & Jamebozorg, 2023). The increasing complexity of virtual learning environments has also led researchers to focus on educational data mining and learning analytics as effective approaches for analyzing educational data and predicting students' academic performance (Cerezo et al., 2024).

Educational data mining utilizes machine learning algorithms and data analysis techniques to identify hidden patterns in educational data and support intelligent educational decision-making (Ataei et al., 2024). Through the analysis of behavioral, cognitive, and interactive learning data, educational data mining can predict students' academic performance, identify learners at risk, and support personalized learning design (Maniyan et al., 2024). Machine learning algorithms such as Random Forest, Artificial Neural Networks, XGBoost, and Support Vector Machines have demonstrated substantial predictive capability in educational contexts (Al Shanableh et al., 2026). Previous studies revealed that deep learning and machine learning models can accurately predict academic performance based on learners' interaction patterns within virtual learning environments (Waheed et al., 2020; Yagci, 2022). Moreover, adaptive learning systems based on educational data mining have been found to increase participation and reduce academic failure in online learning environments (Azhineh et al., 2025). Despite the growing body of literature, many previous studies have focused mainly on higher education students and have paid less attention to secondary school learners and their unique cognitive, emotional, and social characteristics (Towner et al., 2023). In addition, many studies in Iran have primarily examined teachers' and students' experiences of virtual learning rather than analyzing actual educational data using machine learning methods (Salimi et al., 2021). Therefore, there remains a need for research that investigates learning patterns among secondary school students in virtual learning environments through data-driven approaches and machine learning algorithms. Accordingly, the present study aimed to

analyze students' learning patterns in virtual education environments and predict their academic performance using educational data mining techniques and machine learning algorithms.

Methods and Materials

The present study employed a data-driven quantitative design based on educational data mining and machine learning approaches. The statistical population consisted of secondary school students in Semnan Province, Iran, who had actively participated in virtual learning environments during the academic year. Data related to 400 students were collected from online learning management systems, school databases, and electronic forms completed by students and teachers. The dataset included a wide range of variables related to behavioral, cognitive, motivational, and educational dimensions of learning. These variables included self-regulated learning, digital skills, family support, teacher feedback quality, academic motivation, previous academic performance, study environment quality, and interaction with educational content.

Data preprocessing procedures were carefully conducted before implementing machine learning algorithms. Incomplete and duplicated records were removed, and missing values were replaced using standard imputation techniques. Outliers were examined using statistical methods and were controlled to ensure the reliability of the dataset. Subsequently, all numerical variables were normalized using z-score standardization to avoid disproportionate influence of variables with different scales. To reduce model complexity and improve predictive accuracy, feature selection techniques based on correlation analysis and variable importance indices were applied.

To identify hidden learning patterns among students, the K-Means clustering algorithm was employed. The optimal number of clusters was determined using the Elbow and Silhouette methods. For predicting students' academic performance, several machine learning algorithms were implemented and compared, including Random Forest, XGBoost, Artificial Neural Network (MLP), Decision Tree, K-Nearest Neighbor (KNN), Support Vector Machine (SVM), Logistic Regression, and Naive Bayes. The dataset was divided into training and testing subsets with a ratio of 70% to 30%. Hyperparameter tuning procedures were conducted to optimize model performance. Model evaluation was performed using Accuracy, Recall, Precision, F1-Score, and ROC curve analysis. Ten-fold cross-validation was also used to improve the reliability and generalizability of the findings. All analyses were conducted using Python programming language and machine learning libraries such as Scikit-learn, Pandas, NumPy, and Matplotlib.

Findings

The results of the clustering analysis revealed three distinct learning patterns among students in virtual education environments. The first cluster, labeled "Independent and Successful Learners," comprised approximately 32% of the sample and was characterized by high levels of self-regulated learning, strong digital skills, active participation in online learning systems, and high family support. Students in this cluster demonstrated superior academic performance and effective adaptation to virtual learning environments. The second cluster, identified as "At-Risk Learners," included nearly 25% of the participants. This group exhibited

low self-regulation abilities, weak digital competencies, limited participation in educational activities, and inadequate study conditions at home. These students experienced greater difficulties in adapting to online education and demonstrated lower academic performance. The third cluster, referred to as “Learners with Hidden Potential,” represented the largest group, accounting for approximately 43% of the sample. Students in this cluster possessed relatively high motivation and moderate cognitive and digital skills; however, insufficient teacher feedback and limited interaction with educational content prevented them from fully utilizing their learning capacities.

The evaluation of machine learning algorithms demonstrated that the Random Forest model achieved the highest predictive performance among all examined models. The model obtained an accuracy rate of 92%, a recall value of 92.8%, a precision score of 91.6%, and an F1-score of 91%. The confusion matrix analysis indicated that the Random Forest model successfully classified the majority of students and exhibited relatively low classification errors. XGBoost ranked second in predictive performance with an accuracy of 90.5%, followed by the Artificial Neural Network model with an accuracy of 89.8%. Although the neural network demonstrated acceptable predictive capability, its training time was longer compared to other models. Decision Tree, SVM, and KNN models showed moderate predictive performance, whereas Logistic Regression and Naive Bayes produced the lowest accuracy values.

Feature importance analysis revealed that self-regulated learning, teacher feedback quality, and family support were the most influential predictors of students’ academic performance. Sensitivity analysis in the Artificial Neural Network model further confirmed that self-regulated learning exerted the strongest effect on academic achievement. In addition, the results indicated a nonlinear interaction between family emotional support and academic performance. Specifically, increased family support significantly improved academic outcomes among students with moderate levels of self-regulation. These findings demonstrated that academic achievement in virtual learning environments is influenced by a complex interaction of cognitive, motivational, social, and technological factors.

Discussion and Conclusion

The findings of the present study demonstrated that students’ learning patterns in virtual education environments are highly heterogeneous and influenced by a combination of individual, motivational, cognitive, and environmental characteristics. The identification of three distinct learning clusters suggests that students experience online learning environments differently and require differentiated educational strategies. Students classified as independent and successful learners possessed stronger self-regulation skills, digital competencies, and family support, which enabled them to manage their learning processes effectively and achieve better academic outcomes. In contrast, at-risk learners faced substantial challenges related to digital literacy, inadequate learning environments, and weak self-regulation abilities. These results indicate that virtual education may intensify educational inequalities when students do not have equal access to technological resources and supportive learning conditions.

The findings also highlighted the critical role of interaction and feedback in online learning environments. Learners with hidden potential demonstrated relatively high academic motivation; however, limited interaction with educational content and insufficient teacher feedback reduced their learning effectiveness. This result emphasizes that motivation alone is insufficient for successful online learning and that instructional design, interaction quality, and teacher presence remain essential components of educational effectiveness. Furthermore, the findings suggest that educational systems should move beyond one-size-fits-all approaches and adopt personalized learning strategies that consider learners' individual differences and behavioral patterns.

Another important finding was the superior predictive performance of the Random Forest algorithm compared to other machine learning models. The high predictive accuracy of Random Forest indicates that ensemble-based algorithms are highly effective in modeling complex nonlinear relationships among educational variables. The ability of machine learning models to identify students at risk and predict academic outcomes demonstrates the practical value of educational data mining for intelligent educational decision-making. In particular, the identification of self-regulated learning, teacher feedback quality, and family support as the strongest predictors of academic achievement suggests that educational interventions should focus on strengthening these dimensions. Enhancing students' self-regulation skills, improving digital literacy, increasing interaction within virtual classrooms, and supporting families in the learning process may substantially improve the effectiveness of online education.

Overall, the present study demonstrated that educational data mining and machine learning techniques provide powerful tools for understanding learning behaviors, identifying hidden learning patterns, and predicting academic performance in virtual education environments. The findings contribute to the growing body of research on educational data science and support the development of data-driven, personalized, and intelligent educational systems capable of improving learning quality and reducing educational inequalities in digital learning environments.

References

- Al Shanableh, N., Alzyoud, M. S., Khalil, A., Benlamine, M., Kraidia, I., Damrah, S., & Tabakhi, A. M. (2026). Forecasting students' academic performance in educational data using machine learning techniques. *International Journal of Information and Communication Technology Education*, 22(1). <https://doi.org/10.4018/IJICTE.399756>
- Alam, M., Al Mamun, M., Pramanik, M. N. H., Jahan, I., Khan, M. R., Dishy, T. T., Akter, S. H., Jothi, Y. M., Shanta, T. A., & Hossain, M. J. (2022). Paradigm shifting of education system during COVID-19 pandemic: A qualitative study on education components. *Heliyon*, 8(12), e11927. <https://doi.org/10.1016/j.heliyon.2022.e11927>
- Alshammari, S. H., & Alrehaili, T. A. (2025). The effect of teaching, social, and cognitive presence on student engagement in online courses: A structural equation modelling approach. *Acta Psychologica*, 258, 105183. <https://doi.org/10.1016/j.actpsy.2025.105183>
- Ataei, A., Rahimi, G., & Shirzad, A. (2024). Applications of data mining in education. *Quarterly Journal of Strategic Research in Teaching and Education*, 1(13), 59-77.
- Azhineh, A., Esmaeili, Z., Seyedahmadi, M. S., & Souri, M. (2025). Designing an adaptive cognitive learning algorithm using educational data mining to predict and improve students' academic performance in virtual environments. Twenty-Second National Conference on Law, Social Sciences and Humanities, Psychology and Counseling,
- Cerezo, R., Lara, J. A., Azevedo, R., & Romero, C. (2024). Reviewing the differences between learning analytics and educational data mining: Towards educational data science. *Computers in human Behavior*, 154, 108155. <https://doi.org/10.1016/j.chb.2024.108155>

- Costa, A., Moreira, D., Casanova, J., Azevedo, A., Goncalves, A., Oliveira, I., Azevedo, R., & Dias, P. C. (2024). Determinants of academic achievement from the middle to secondary school education: A systematic review. *Social Psychology of Education*, 27, 3533-3572.
- Dietrich, J., Greiner, F., Weber Liel, D., Berweiger, B., Kaempfe, N., & Kracke, B. (2021). Does an individualized learning design improve university student online learning? A randomized field experiment. *Computers in human Behavior*, 122, 106819. <https://doi.org/10.1016/j.chb.2021.106819>
- Ghanaati, R., & Jamebozorg, Z. (2023). Designing the learning environment of virtual education based on creative assessment methods and its impact on the learning of the second year elementary school students. *Journal of Psychological Science*, 22(132), 111-132. <https://doi.org/10.52547/JPS.22.132.2503>
- Haleem, A., Javaid, M., Qadri, M. A., & Suman, R. (2022). Understanding the role of digital technologies in education: A review. *Sustainable Operations and Computers*, 3, 275-285. <https://doi.org/10.1016/j.susoc.2022.03.004>
- Imran, M., Almusharraf, N., & Abbasova, M. Y. (2025). Digital learning transformation: A study of teachers' post COVID-19 experiences. *Social Sciences & Humanities Open*, 11, 101228. <https://doi.org/10.1016/j.ssaho.2024.101228>
- Jafari Aghdam, F., Jadidzadeh, A., Naji Esfahani, S. M., & Pourtaghi Rastgoo Moghadam, S. (2023). The effect of virtual education on students' academic performance during the coronavirus pandemic: A case study of economics students at the University of Tehran. *Quarterly Journal of Economic Research*, 58(4), 565-592.
- Jiang, S., Li, H., & Gan, D. (2025). Technology acceptance model for online education: Identifying interdisciplinary topics and their evolution based on BERTopic model. *Social Sciences & Humanities Open*, 12, 101831. <https://doi.org/10.1016/j.ssaho.2025.101831>
- Jiao, H., Wang, T., Libaers, D., Yang, J., & Hu, L. (2025). The relationship between digital technologies and innovation: A review, critique, and research agenda. *Journal of Innovation & Knowledge*, 10(1), 100638. <https://doi.org/10.1016/j.jik.2024.100638>
- Jin, S. H., Im, K., Yoo, M., Roll, I., & Seo, K. (2023). Supporting students' self-regulated learning in online learning using artificial intelligence applications. *International Journal of Educational Technology in Higher Education*, 20, 37. <https://doi.org/10.1186/s41239-023-00392-3>
- Kazemi Mehr, B., & Salehi, M. (2023). The classroom in the digital world: Exploring educational innovations based on web-based learning models. First International Conference on New Horizons in Education in the Third Millennium,
- Khosravi, Z., Nazari, Z., & Aminnejad, F. (2025). Innovative virtual learning models in Iran's education system. First International Conference on New Horizons in Learning Psychology and Education from the Teacher's Perspective,
- Kim, D., Lee, Y., Leite, W. L., & Huggins Manley, A. C. (2020). Exploring student and teacher usage patterns associated with student attrition in an open educational resource supported online learning platform. *Computers & Education*, 156, 103961. <https://doi.org/10.1016/j.compedu.2020.103961>
- Kolahi Zadeh, N., Heshmati Amin, S., & Kheiri, Z. (2025). Investigating the effects of student interaction and participation in virtual educational networks on improving their academic achievements. Fourth National Conference on New Developments in Developmental and Educational Psychology, Bandar Abbas.
- Maniyan, S., Ghousi, R., & Haeri, A. (2024). Data mining based decision support system for educational decision makers: Extracting rules to enhance academic efficiency. *Computers and Education: Artificial Intelligence*, 6, 100242. <https://doi.org/10.1016/j.caeai.2024.100242>
- Ministry of Education of, I. (2022). *Report on students' academic performance during the virtual education period*.
- Mousizadeh Ahaghi, M. (2025). Analyzing the effect of e-learning and virtual education environments on the development of learners' cognitive, metacognitive, and self-regulation skills. Sixth International Conference on Educational Sciences, Counseling, Psychology, and Social Sciences,
- Rahbar Karbasdehi, F., & Rahbar Karbasdehi, E. (2021). Virtual education of students during the 2019 coronavirus pandemic: Problems and solutions. *Journal of Yazd Center for Medical Sciences Education Studies and Development*, 16(3), 224-225.
- Rezai, A., Ahmadi, R., Ashkani, P., & Hosseini, G. H. (2025). Implementing active learning approach to promote motivation, reduce anxiety, and shape positive attitudes: A case study of EFL learners. *Acta Psychologica*, 253, 104704. <https://doi.org/10.1016/j.actpsy.2025.104704>
- Rostamzadeh, F., Shojaabadi, S., & Asadi, S. (2023). The effect of virtual education on achievement motivation and academic achievement of elementary school students in Urmia County in the 2021-2022 academic year. International Conference on Management and Industry,
- Salari, M., Radfar, R., & Faghihi, M. (2024). Predicting students' performance using machine learning and educational data mining algorithms: A case study of Shahed University. *Intelligent Business Management Studies*, 13(47), 315-366.
- Salimi, S., Bahari, A., & Moudi, B. (2021). Analyzing elementary teachers' lived experiences of the strengths and weaknesses of the Shad educational network during the COVID-19 pandemic. *New Educational Approaches*, 16(2), 81-98.
- Singh, M., James, P. S., Paul, H., & Bolar, K. (2022). Impact of cognitive behavioral motivation on student engagement. *Heliyon*, 8(7), e09843. <https://doi.org/10.1016/j.heliyon.2022.e09843>
- Song, Y., & Vermunt, J. D. (2021). A comparative study of learning patterns of secondary school, high school and college students. *Studies in Educational Evaluation*, 68, 100958. <https://doi.org/10.1016/j.stueduc.2020.100958>
- Torbati Nejad, H., Shahvarani, M., & Hojgozari, R. (2022). Investigating the effect of virtual education on students' learning with emphasis on strategies to increase effectiveness. *New Advances in Psychology, Educational Sciences and Education*, 5(46), 129-138.
- Towner, E., Chierchia, G., & Blakemore, S. J. (2023). Sensitivity and specificity in affective and social learning in adolescence. *Trends in Cognitive Sciences*, 27(7), 642-655. <https://doi.org/10.1016/j.tics.2023.04.002>
- van de Werfhorst, H. G., Kessenich, E., & Geven, S. (2022). The digital divide in online education: Inequality in digital readiness of students and schools. *Computers and Education Open*, 3, 100100. <https://doi.org/10.1016/j.caeo.2022.100100>

- Verstege, S., Zhang, Y., Wierenga, P., Paquette, L., & Diederens, J. (2024). Using sequential pattern mining to understand how students use guidance while doing scientific calculations. *Technology, Knowledge and Learning*, 29, 897-920. <https://doi.org/10.1007/s10758-023-09668-8>
- Waheed, H., Hassan, S. U., Aljohani, N. R., Hardman, J., Alelyani, S., & Nawaz, R. (2020). Predicting academic performance of students from VLE big data using deep learning models. *Computers in human Behavior*, 104, 106189. <https://doi.org/10.1016/j.chb.2019.106189>
- Yagci, M. (2022). Educational data mining: Prediction of students' academic performance using machine learning algorithms. *Smart Learning Environments*, 9, 11. <https://doi.org/10.1186/s40561-022-00192-z>
- Yan, F., Wu, N., & Iliyasa, A. M. (2022). Framework for identifying and visualising emotional atmosphere in online learning environments in the COVID-19 era. *Applied Intelligence*, 52, 9406-9422. <https://doi.org/10.1007/s10489-021-02916-z>
- Zare Nasab, M., & Jamebozorg, Z. (2025). Challenges and opportunities of using artificial intelligence in elementary education: From the perspective of novice teachers. *Technology and Knowledge Studies in Education*, 5(1), 35-50.
- Zeinali, B., & Daneshmand, A. (2025). Factors affecting technology-based educational and research efficiency of elementary teachers with emphasis on professional skill development. *Research and Innovation in Elementary Education*, 7(1), 22-42.
- Zhao, Y., Li, Y., Ma, S., Xu, Z., & Zhang, B. (2025). A meta-analysis of the correlation between self-regulated learning strategies and academic performance in online and blended learning environments. *Computers & Education*, 230, 105279. <https://doi.org/10.1016/j.compedu.2025.105279>