

A Bidirectional GRU-Based Framework for Early Prediction of Student Academic Failure in E-Learning Environments

Masoumeh Rezaei ^{1✉}

1. Assistant Professor of Computer Engineering, University of Sistan and Baluchestan, Zahedan, Iran (Corresponding Author). Email: mrezaei@ece.usb.ac.ir

Article Info	ABSTRACT
Article type: Research Article	Early prediction of students' academic performance in e-learning systems plays a crucial role in reducing dropout rates, improving the learning process, and designing targeted educational interventions. Despite recent advances in machine learning and deep learning methods, many existing studies are either unable to effectively model the temporal dependencies in students' behavioral data or require a large portion of semester-long data to make reliable predictions. In this paper, a framework based on a Bidirectional Gated Recurrent Unit network is proposed for the early prediction of at-risk students. The proposed method utilizes behavioral data from the Open University dataset and relies solely on interaction records collected during the first 42 days of the course. It extracts and analyzes temporal activity patterns of students. During the preprocessing stage, multiple statistical and temporal features are extracted, including cumulative activity counts, moving averages, activity change rate, number of active days, and time since last activity. The data are then structured into sequential formats, normalized, and enhanced before being fed into a Bi-GRU-based architecture, which effectively learns temporal dependencies and behavioral dynamics of students. To address data imbalance, a weighted loss function is employed, and early stopping is used to prevent overfitting. Experimental results demonstrate that the proposed method achieves an accuracy of 67.37%, an AUC of 73.54%, and an F1-score of 63.99%, indicating its effectiveness in identifying at-risk students at early stages of the course. The findings suggest that modeling student behavioral sequences using deep temporal learning approaches can provide an effective tool for early warning systems in educational environments.
Article history:	
Received: 6 April 2026	
Received in revised form: 9 May 2026	
Accepted: 6 June 2026	
Published online: 21 June 2026	
Keywords: Early prediction of academic performance, Learning analytics, Bidirectional GRU network, E-learning, Educational data mining	

Cite this article: Rezaei, M., (2026). A Bidirectional GRU-Based Framework for Early Prediction of Student Academic Failure in E-Learning Environments. *Journal of Creative and Effective Education*, 2 (1), 27-40.

DOI: 10.22111/cee.2026.55414.1027



© The Author(s).

Publisher: University of Sistan and Baluchestan.

چارچوب مبتنی بر شبکه GRU دوسویه برای پیش‌بینی زود هنگام افت تحصیلی دانشجویان در محیط‌های یادگیری الکترونیکی

✉ معصومه رضائی^۱

۱. استادیار گروه مهندسی کامپیوتر، دانشگاه سیستان و بلوچستان، زاهدان، ایران. (نویسنده مسئول) رایانامه: mrezaei@ece.usb.ac.ir

اطلاعات مقاله	چکیده
نوع مقاله: مقاله پژوهشی	پیش‌بینی زود هنگام عملکرد تحصیلی دانشجویان در سامانه‌های آموزش الکترونیکی نقش مهمی در کاهش افت تحصیلی، بهبود فرآیند یادگیری و طراحی مداخلات آموزشی هدفمند ایفا می‌کند. با وجود پیشرفت روش‌های یادگیری ماشین و یادگیری عمیق، بسیاری از پژوهش‌های پیشین یا قادر به مدل‌سازی دقیق وابستگی‌های زمانی رفتار دانشجویان نیستند و یا برای انجام پیش‌بینی به بخش بزرگی از داده‌های ترم تحصیلی نیاز دارند. در این پژوهش، یک چارچوب مبتنی بر شبکه بازگشتی دوسویه GRU برای پیش‌بینی زود هنگام دانشجویان در معرض خطر ارائه شده است. روش پیشنهادی با استفاده از داده‌های رفتاری مجموعه داده دانشگاه آزاد بریتانیا و تنها با استفاده از تعاملات ثبت شده طی ۴۲ روز نخست دوره آموزشی، الگوهای زمانی فعالیت دانشجویان را استخراج و تحلیل می‌کند. در مرحله پیش‌پردازش، ویژگی‌های آماری و زمانی متعددی شامل مجموع تجمعی فعالیت‌ها، میانگین‌های متحرک، نرخ تغییر فعالیت، تعداد روزهای فعال و فاصله زمانی از آخرین فعالیت استخراج شدند، سپس داده‌ها به صورت توالی‌های زمانی مدل‌سازی شده و پس از نرمال‌سازی و تقویت ویژگی‌ها، به یک معماری مبتنی بر شبکه GRU دوسویه وارد شدند تا وابستگی‌های زمانی و تغییرات رفتاری دانشجویان به صورت مؤثر یادگیری شوند. برای مقابله با عدم توازن داده‌ها از تابع هزینه وزن‌دهی شده و برای جلوگیری از بیش‌برازش از توقف زود هنگام استفاده شد. نتایج آزمایش‌ها نشان داد که روش پیشنهادی به صحت ۶۷/۳۷ درصد، مقدار AUC برابر ۷۳/۵۴ درصد و مقدار F1 برابر ۶۳/۹۹ درصد دست یافته است که نشان‌دهنده توانایی مناسب مدل در شناسایی دانشجویان در معرض خطر در مراحل ابتدایی دوره آموزشی است. یافته‌های پژوهش نشان می‌دهد که تحلیل توالی‌های رفتاری دانشجویان با استفاده از مدل‌های عمیق زمانی می‌تواند ابزار مؤثری برای سامانه‌های هشدار زود هنگام آموزشی فراهم کند.
تاریخ دریافت: ۱۴۰۵/۱/۱۷	
تاریخ ویرایش: ۱۴۰۵/۲/۱۹	
تاریخ پذیرش: ۱۴۰۵/۳/۱۶	
تاریخ انتشار: ۱۴۰۵/۳/۳۱	
واژه‌های کلیدی:	
پیش‌بینی زود هنگام عملکرد تحصیلی، تحلیل یادگیری، شبکه GRU دوسویه، آموزش الکترونیکی، داده کاوی آموزشی	

استناد: رضائی، معصومه. (۱۴۰۵). چارچوب مبتنی بر شبکه GRU دوسویه برای پیش‌بینی زود هنگام افت تحصیلی دانشجویان در محیط‌های یادگیری الکترونیکی. آموزش خلاق و اثربخش، ۲ (۱)، ۴۰-۲۷.

DOI: 10.22111/cee.2026.55414.1027



ناشر: دانشگاه سیستان و بلوچستان.

© نویسندگان. رضائی، معصومه

۱- مقدمه

گسترش سریع سامانه‌های آموزش الکترونیکی و محیط‌های یادگیری مجازی در سال‌های اخیر، موجب تولید حجم عظیمی از داده‌های آموزشی شده است. این داده‌ها اطلاعات ارزشمندی درباره نحوه تعامل دانشجویان با محیط یادگیری، الگوهای رفتاری آنان و روند پیشرفت تحصیلی ایشان در اختیار قرار می‌دهند. تحلیل این داده‌ها می‌تواند به شناسایی الگوهای یادگیری، ارزیابی وضعیت آموزشی دانشجویان و بهبود فرایندهای تصمیم‌گیری در نظام‌های آموزشی کمک کند (سیمنز^۱ و بیکر^۲، ۲۰۱۲؛ فرگوسن^۳، ۲۰۱۲). در این میان، پیش‌بینی عملکرد تحصیلی دانشجویان به یکی از مهم‌ترین موضوعات پژوهشی در حوزه تحلیل یادگیری و داده‌کاوی آموزشی تبدیل شده است؛ زیرا شناسایی زود هنگام دانشجویان در معرض افت تحصیلی، امکان طراحی مداخلات آموزشی مناسب و کاهش احتمال شکست یا ترک تحصیل را فراهم می‌سازد.

در سال‌های گذشته، پژوهشگران از روش‌های مختلف یادگیری ماشین برای پیش‌بینی عملکرد تحصیلی استفاده کرده‌اند. الگوریتم‌هایی نظیر درخت تصمیم، ماشین بردار پشتیبان، جنگل تصادفی و رگرسیون لجستیک از جمله روش‌های پرکاربرد در این حوزه بوده‌اند و در بسیاری از مطالعات نتایج قابل قبولی ارائه کرده‌اند (کوینلان^۴، ۲۰۱۴؛ کورتس^۵ و واپنیک^۶، ۱۹۹۵؛ بریمن^۷، ۲۰۰۱؛ هاسمر^۸ و دیگران، ۲۰۱۳). با وجود این، این روش‌ها عمدتاً بر داده‌های ایستا متکی هستند و توانایی محدودی در مدل‌سازی وابستگی‌های زمانی و تغییرات رفتاری دانشجویان در طول زمان دارند (سا^۹، ۲۰۱۶).

پیشرفت‌های اخیر در حوزه یادگیری عمیق، زمینه استفاده از شبکه‌های عصبی بازگشتی را برای تحلیل داده‌های آموزشی فراهم کرده است. به‌ویژه، معماری‌های LSTM^{۱۰} و GRU^{۱۱} به دلیل توانایی در مدل‌سازی داده‌های ترتیبی و استخراج وابستگی‌های زمانی، توجه بسیاری از پژوهشگران را به خود جلب کرده‌اند (هوخریتر^{۱۲} و اشمیدهوربر^{۱۳}، ۱۹۹۷؛ چو^{۱۴} و همکاران، ۲۰۱۴). این مدل‌ها قادرند الگوهای رفتاری دانشجویان را در طول زمان تحلیل کرده و اطلاعات ارزشمندی برای پیش‌بینی عملکرد تحصیلی فراهم آورند. با این حال، بخش قابل توجهی از پژوهش‌های پیشین از داده‌های کامل یا بخش عمده‌ای از یک نیمسال تحصیلی برای آموزش مدل استفاده کرده‌اند و توجه کمتری به مسئله پیش‌بینی زود هنگام عملکرد دانشجویان داشته‌اند.

^۱ . Siemens

^۲ . Baker

^۳ . Ferguson

^۴ . Quinlan

^۵ . Cortes

^۶ . Vapnik

^۷ . Breiman

^۸ . Hosmer

^۹ . Saa

^{۱۰} . Long Short-Term Memory

^{۱۱} . Gated Recurrent Unit

^{۱۲} . Hochreiter

^{۱۳} . Schmid Huber

^{۱۴} . Cho

در سال‌های اخیر، مجموعه داده OULAD^۱ به یکی از مهم‌ترین مجموعه داده‌های مرجع در حوزه تحلیل یادگیری تبدیل شده است (کوزلیک و همکاران، ۲۰۱۷). این مجموعه داده شامل اطلاعات جمعیت‌شناختی، سوابق ارزیابی و تعاملات دانشجویان با محیط یادگیری مجازی است و امکان بررسی رفتار یادگیرندگان در مقیاس گسترده را فراهم می‌کند. پژوهش‌های متعددی با استفاده از این مجموعه داده به تحلیل رفتار دانشجویان و پیش‌بینی عملکرد تحصیلی آنان پرداخته‌اند (هلستا^۲ و همکاران، ۲۰۱۸؛ کوزلیک و همکاران، ۲۰۱۷؛ توماسویچ^۳ و همکاران، ۲۰۲۰). با وجود نتایج امیدوارکننده، همچنان چالش‌هایی نظیر استخراج مؤثر الگوهای زمانی، مدل‌سازی تغییرات رفتاری دانشجویان و دستیابی به پیش‌بینی دقیق در مراحل ابتدایی دوره آموزشی وجود دارد.

بر این اساس، پژوهش حاضر یک چارچوب مبتنی بر شبکه عصبی بازگشتی دوسویه (Bi-GRU) GRU برای پیش‌بینی زودهنگام عملکرد تحصیلی دانشجویان ارائه می‌کند. در روش پیشنهادی، تنها داده‌های مربوط به ۴۲ روز نخست دوره آموزشی مورد استفاده قرار می‌گیرند تا امکان شناسایی سریع دانشجویان در معرض خطر فراهم شود. همچنین، مجموعه‌ای از ویژگی‌های زمانی و رفتاری از تعاملات دانشجویان استخراج شده و به صورت توالی‌های زمانی به مدل ارائه می‌شوند تا الگوهای رفتاری دانشجویان با دقت بیشتری مدل‌سازی شوند. لذا هدف اصلی پژوهش، توسعه مدلی است که بتواند با استفاده از داده‌های محدود ابتدای دوره، امکان شناسایی به‌موقع دانشجویان در معرض خطر و فراهم‌سازی بستر مداخلات آموزشی مؤثر را فراهم آورد.

مهم‌ترین دستاوردهای پژوهش حاضر عبارت‌اند از:

- ارائه یک چارچوب مبتنی بر یادگیری عمیق برای پیش‌بینی زودهنگام عملکرد تحصیلی دانشجویان در محیط‌های آموزش الکترونیکی؛
 - مدل‌سازی وابستگی‌های زمانی رفتار دانشجویان با استفاده از معماری دوسویه Bi-GRU
 - استخراج و به‌کارگیری مجموعه‌ای از ویژگی‌های زمانی و رفتاری شامل ویژگی‌های تجمعی، میانگین‌های متحرک، نرخ تغییرات و شاخص‌های فعالیت آموزشی؛
 - انجام پیش‌بینی زودهنگام تنها با استفاده از تعاملات ۴۲ روز نخست دانشجویان؛
 - دستیابی به عملکرد مناسب در شناسایی دانشجویان در معرض خطر در مجموعه داده OULAD.
- لازم به ذکر است ساختار مقاله به گونه‌ای تنظیم شده است که پس از مرور مبانی نظری و پژوهش‌های پیشین، روش پیشنهادی تشریح می‌شود؛ سپس نتایج ارزیابی تجربی ارائه و تحلیل شده و در پایان، نتیجه‌گیری، محدودیت‌ها و پیشنهادهایی برای پژوهش‌های آینده مطرح می‌شود.

۲- مبانی نظری و پیشینه تحقیق

تحلیل یادگیری و داده‌کاوی آموزشی از مهم‌ترین حوزه‌های پژوهشی در آموزش الکترونیکی محسوب می‌شوند که با هدف استخراج دانش از داده‌های آموزشی، شناسایی الگوهای رفتاری یادگیرندگان و بهبود فرآیند یاددهی - یادگیری

^۱. Open University Learning Analytics Dataset

^۲. Hlosta

^۳. Tomasevic

توسعه یافته‌اند. گسترش سامانه‌های مدیریت یادگیری و محیط‌های آموزش مجازی موجب تولید حجم عظیمی از داده‌های آموزشی شده است که می‌توانند اطلاعات ارزشمندی درباره وضعیت تحصیلی، میزان مشارکت و رفتار یادگیری دانشجویان فراهم کنند. در این میان، پیش‌بینی عملکرد تحصیلی دانشجویان به یکی از مهم‌ترین کاربردهای تحلیل یادگیری تبدیل شده است؛ زیرا شناسایی زود هنگام دانشجویان در معرض افت تحصیلی امکان طراحی مداخلات آموزشی مناسب و کاهش نرخ شکست یا ترک تحصیل را فراهم می‌سازد.

پژوهش‌های اولیه در این حوزه عمدتاً بر استفاده از الگوریتم‌های کلاسیک یادگیری ماشین متمرکز بودند. رومرو^۱ و ونچورا^۲ (۲۰۱۰) در یک مطالعه مروری جامع نشان دادند که داده‌کاوی آموزشی می‌تواند نقش مؤثری در تحلیل رفتار یادگیرندگان و پیش‌بینی عملکرد تحصیلی ایفا کند. در ادامه، الگوریتم‌هایی نظیر درخت تصمیم، ماشین بردار پشتیبان، رگرسیون لجستیک و جنگل تصادفی به‌طور گسترده در پیش‌بینی عملکرد دانشجویان مورد استفاده قرار گرفتند. اگرچه این روش‌ها در استخراج روابط میان متغیرهای آموزشی موفق بودند، اما در مدل‌سازی وابستگی‌های زمانی و تغییرات رفتاری دانشجویان در طول زمان با محدودیت‌هایی مواجه بودند.

در حوزه مدل‌سازی داده‌های ترتیبی، بنجیو^۳ و همکاران (۱۹۹۴) نشان دادند که شبکه‌های عصبی بازگشتی متداول در یادگیری وابستگی‌های بلندمدت با مشکل مواجه هستند. به‌منظور رفع این محدودیت، (هوخریتر و اشمیدهور، ۱۹۹۷) معماری حافظه کوتاه‌مدت - بلندمدت را معرفی کردند که توانایی بیشتری در حفظ اطلاعات بلندمدت داشت. پس از آن، چو و همکاران (۲۰۱۴)، GRU را ارائه کردند که با ساختاری ساده‌تر و تعداد پارامترهای کمتر، عملکردی قابل مقایسه با LSTM در مدل‌سازی داده‌های ترتیبی ارائه می‌داد. این پیشرفت‌ها زمینه استفاده از شبکه‌های عصبی عمیق در تحلیل رفتار یادگیرندگان و پیش‌بینی عملکرد تحصیلی را فراهم کرد.

در حوزه تحلیل یادگیری، سیمنز و بیکر (۲۰۱۲) بر همگرایی حوزه‌های تحلیل یادگیری و داده‌کاوی آموزشی تأکید کردند و نقش این دو رویکرد را در ارتقای کیفیت آموزش مورد توجه قرار دادند. با گسترش کاربرد یادگیری عمیق در آموزش، پژوهشگران به استفاده از شبکه‌های عصبی بازگشتی برای مدل‌سازی رفتار دانشجویان روی آوردند. پیچ^۴ و همکاران (۲۰۱۵) مدل ردیابی عمیق دانش^۵ را معرفی کردند که با استفاده از شبکه‌های عصبی بازگشتی، فرآیند یادگیری دانشجویان را مدل‌سازی می‌کرد. نتایج این پژوهش نشان داد که مدل‌های عمیق قادرند الگوهای پنهان یادگیری را با دقت بیشتری نسبت به روش‌های سنتی استخراج کنند. در همین سال، هلستا^۶ و همکاران (۲۰۱۵) با بررسی رفتار دانشجویان در محیط‌های یادگیری مجازی نشان دادند که الگوهای تعامل دانشجویان با سامانه‌های آموزشی می‌تواند شاخص مناسبی برای پیش‌بینی عملکرد تحصیلی آنان باشد.

سا (۲۰۱۶) با استفاده از تکنیک‌های داده‌کاوی آموزشی به پیش‌بینی عملکرد دانشجویان پرداخت و گزارش کرد که ویژگی‌های استخراج‌شده از داده‌های آموزشی توانایی مناسبی در تفکیک دانشجویان موفق از دانشجویان در معرض خطر دارند. در ادامه، کوزلیک و همکاران، (۲۰۱۷) مجموعه داده OULAD را معرفی کردند. این مجموعه داده به

1. Romero

2. Ventura

3. Bengio

4. Piech

5. Deep Knowledge Tracing

6. Hlosta

دلیل برخورداری از اطلاعات جمعیت‌شناختی، نمرات ارزیابی و داده‌های رفتاری دانشجویان در محیط یادگیری مجازی، به یکی از مهم‌ترین مجموعه داده‌های مرجع در حوزه تحلیل یادگیری تبدیل شد و در بسیاری از پژوهش‌های بعدی مورد استفاده قرار گرفت.

توماسویچ و دیگران (۲۰۲۰) با مرور و مقایسه روش‌های مختلف داده‌کاوی و یادگیری ماشین برای پیش‌بینی عملکرد تحصیلی دانشجویان نشان دادند که الگوریتم‌های کلاسیک اگرچه نتایج قابل قبولی ارائه می‌کنند، اما در مدل‌سازی رفتارهای پویا و وابستگی‌های زمانی محدودیت دارند. این موضوع موجب افزایش توجه پژوهشگران به بهره‌گیری از مدل‌های عمیق مبتنی بر توالی شد.

در مطالعات داخلی نیز توجه به پیش‌بینی عملکرد تحصیلی دانشجویان افزایش یافته است. سالاری، رادفر و فقیهی (۲۰۲۴) با استفاده از الگوریتم‌های یادگیری ماشین به تحلیل داده‌های آموزشی و پیش‌بینی عملکرد دانشجویان پرداختند. نتایج پژوهش آنان نشان داد که ویژگی‌های آموزشی و رفتاری دانشجویان می‌توانند مبنای مناسبی برای شناسایی دانشجویان در معرض افت تحصیلی و طراحی مداخلات آموزشی هدفمند باشند.

در پژوهش‌های جدیدتر، استفاده از معماری‌های پیشرفته یادگیری عمیق برای تحلیل داده‌های آموزشی مورد توجه قرار گرفته است. النسیان^۱ و همکاران (۲۰۲۵) مدل KAN Former را برای پیش‌بینی عملکرد دانشجویان در محیط‌های یادگیری مجازی ارائه کردند. این مدل با ترکیب سازوکار توجه و شبکه‌های Kolmogorov-Arnold توانست عملکرد بهتری نسبت به چندین مدل پایه در وظایف طبقه‌بندی دودویی و چندکلاسه ارائه دهد و دقت، یادآوری و امتیاز F1 بالاتری را به دست آورد. نتایج این پژوهش نشان‌دهنده ظرفیت بالای مدل‌های عمیق مبتنی بر توجه در شناسایی دانشجویان در معرض خطر و پشتیبانی از تصمیم‌گیری آموزشی است.

مرور مطالعات پیشین نشان می‌دهد که اگرچه پژوهش‌های متعددی در زمینه پیش‌بینی عملکرد تحصیلی دانشجویان انجام شده است، اما بخش قابل توجهی از آن‌ها بر الگوریتم‌های کلاسیک یادگیری ماشین یا استفاده از داده‌های بخش عمده‌ای از نیمسال تحصیلی متمرکز بوده‌اند. همچنین تعداد محدودی از مطالعات به پیش‌بینی زودهنگام عملکرد دانشجویان بر اساس داده‌های رفتاری هفته‌های ابتدایی دوره آموزشی و مدل‌سازی وابستگی‌های زمانی با استفاده از معماری‌های عمیق پرداخته‌اند. از این‌رو، پژوهش حاضر با بهره‌گیری از معماری Bi-GRU و استفاده از داده‌های ۴۲ روز نخست دوره آموزشی در مجموعه داده OULAD، درصدد ارائه مدلی برای پیش‌بینی زودهنگام عملکرد تحصیلی دانشجویان و شناسایی دانشجویان در معرض خطر است.

۳- روش‌شناسی و داده‌های تحقیق

در پژوهش حاضر، یک چارچوب مبتنی بر یادگیری عمیق برای پیش‌بینی زودهنگام عملکرد تحصیلی دانشجویان در محیط‌های آموزش آنلاین ارائه شده است. هدف اصلی روش پیشنهادی، شناسایی دانشجویان در معرض افت تحصیلی تنها با استفاده از تعاملات رفتاری هفته‌های ابتدایی دوره آموزشی است. برای این منظور، یک معماری مبتنی بر شبکه بازگشتی دوسویه GRU طراحی شد تا بتواند الگوهای زمانی رفتار دانشجویان را به صورت مؤثر مدل‌سازی کند.

^۱ . Alnasyan

چارچوب پیشنهادی شامل مراحل جمع‌آوری و ادغام داده‌ها، پیش‌پردازش، استخراج ویژگی‌های زمانی و رفتاری، تشکیل توالی‌های زمانی، تقویت ویژگی‌ها و آموزش مدل یادگیری عمیق می‌باشد.

۳-۱. مجموعه داده و آماده‌سازی اولیه داده‌ها

در این پژوهش از مجموعه داده دانشگاه آزاد بریتانیا^۱ استفاده شد (کوزیلک و همکاران، ۲۰۱۷). این مجموعه داده شامل اطلاعات جمعیت‌شناختی دانشجویان، وضعیت نهایی تحصیلی و تعاملات آن‌ها با محیط آموزش الکترونیکی است. دو جدول اصلی مورد استفاده شامل اطلاعات دانشجویان و تعاملات کاربران با محیط یادگیری بودند که بر اساس شناسه دانشجو با یکدیگر ادغام شدند. به‌منظور انجام پیش‌بینی زود هنگام، تنها داده‌های مربوط به ۴۲ روز ابتدایی دوره آموزشی انتخاب گردیدند. در ادامه، وضعیت نهایی دانشجویان به دو کلاس اصلی تبدیل شد. دانشجویان دارای وضعیت «قبولی» و «قبولی ممتاز» در گروه دانشجویان موفق و دانشجویان دارای وضعیت «رد شدن» یا «انصراف» در گروه دانشجویان در معرض خطر قرار گرفتند. پس از حذف داده‌های ناقص و نمونه‌های نامعتبر، ویژگی‌های دسته‌ای شامل جنسیت، منطقه جغرافیایی، سطح تحصیلات، بازه سنی و وضعیت ناتوانی به مقادیر عددی تبدیل شدند تا امکان استفاده از آن‌ها در مدل یادگیری عمیق فراهم گردد.

۳-۲. استخراج ویژگی‌های رفتاری و زمانی

به‌منظور افزایش توانایی مدل در تحلیل رفتار دانشجویان، مجموعه‌ای از ویژگی‌های زمانی و رفتاری از داده‌های خام استخراج شد. این ویژگی‌ها اطلاعات مهمی درباره روند تعامل دانشجویان با سامانه آموزشی فراهم می‌کنند و امکان مدل‌سازی رفتار یادگیری را بهبود می‌دهند. در این مرحله، ویژگی‌هایی شامل تعداد کلیک‌های روزانه، لگاریتم تعداد کلیک‌ها، مجموع تجمعی فعالیت‌ها، میانگین متحرک فعالیت‌ها در پنجره‌های زمانی مختلف، نرخ تغییر فعالیت‌ها، تعداد روزهای فعال دانشجو، شماره هفته آموزشی، روز هفته و فاصله زمانی از آخرین فعالیت استخراج گردیدند. مجموع تجمعی فعالیت‌های دانشجو مطابق رابطه زیر محاسبه شد:

$$C_t = \sum_{i=1}^t x_i \quad (1)$$

که در آن x_i تعداد فعالیت‌های دانشجو در روز i و C_t مجموع تجمعی فعالیت‌ها تا روز t است. برای تحلیل روند تغییرات رفتاری، نرخ تغییر فعالیت‌ها نیز محاسبه شد:

$$V_t = x_t - x_{t-1} \quad (2)$$

که در آن V_t بیانگر سرعت تغییر رفتار دانشجو در روز t می‌باشد. همچنین میانگین متحرک فعالیت‌ها برای کاهش نوسانات کوتاه‌مدت مطابق رابطه زیر استخراج شد:

$$M_t = \frac{1}{w} \sum_{i=t-w+1}^t x_i \quad (3)$$

^۱ The Open University of UK

که در آن W اندازه پنجره زمانی است. به منظور تحلیل میزان پایداری فعالیت دانشجویان، ویژگی فاصله از آخرین فعالیت نیز تعریف شد. این ویژگی نشان می‌دهد دانشجو چه مدت غیرفعال بوده است و می‌تواند نقش مهمی در پیش‌بینی افت تحصیلی داشته باشد.

۳-۳. تشکیل توالی‌های زمانی و تقویت ویژگی‌ها

از آنجا که رفتار دانشجویان دارای ماهیت زمانی است، داده‌ها به صورت توالی‌های ترتیبی سازمان‌دهی شدند. برای هر دانشجو، یک توالی زمانی شامل فعالیت‌های روزانه طی ۴۲ روز ابتدایی تشکیل شد. در صورتی که طول توالی کمتر از ۴۲ روز بود، از مقدار صفر برای تکمیل توالی استفاده گردید تا تمامی نمونه‌ها دارای طول یکسان باشند. پس از استخراج ویژگی‌ها، هر دانشجو به صورت یک ماتریس زمانی با ابعاد 42×20 نمایش داده شد. در مجموع، ۱۲ ویژگی رفتاری و زمانی به همراه ۸ ویژگی دسته‌ای استخراج شد که در نهایت بردار ویژگی ۲۰ بعدی را تشکیل دادند. سپس برای تقویت نمایش ویژگی‌ها، اختلاف زمانی ویژگی‌ها نیز محاسبه و به بردار ویژگی افزوده شد. اختلاف زمانی مطابق رابطه زیر تعریف گردید:

$$D_t = X_t - X_{t-1} \quad (4)$$

که در آن X_t بردار ویژگی‌ها در زمان t است. در نهایت، ویژگی‌های اصلی و ویژگی‌های اختلاف زمانی با یکدیگر ترکیب شدند:

$$F_t = [X_t, D_t] \quad (5)$$

بردار تقویت‌شده F_t به عنوان ورودی نهایی مدل دوسویه GRU مورد استفاده قرار گرفت. این مرحله موجب شد مدل بتواند علاوه بر مقدار ویژگی‌ها، روند تغییرات رفتاری دانشجویان را نیز یاد بگیرد. پس از مرحله تقویت ویژگی، ابعاد نهایی هر توالی افزایش یافت و اطلاعات رفتاری غنی‌تری در اختیار مدل قرار گرفت.

۳-۴. معماری مدل پیشنهادی

معماری پیشنهادی این پژوهش بر پایه شبکه بازگشتی دوسویه GRU طراحی شده است. هدف اصلی این معماری، استخراج وابستگی‌های زمانی کوتاه‌مدت و بلندمدت از رفتار دانشجویان در هفته‌های ابتدایی دوره آموزشی است. ورودی مدل شامل توالی‌های زمانی مربوط به هر دانشجو می‌باشد. پس از مرحله تقویت ویژگی‌ها، داده‌ها به شبکه GRU دوسویه وارد می‌شوند. استفاده از ساختار دوسویه موجب می‌شود وابستگی‌های زمانی موجود در کل توالی رفتاری دانشجو به صورت مؤثرتری استخراج شوند.

در معماری پیشنهادی، از یک لایه GRU دوسویه با ۱۲۸ واحد پنهان در هر جهت استفاده شد. همچنین برای کاهش بیش‌برازش، نرخ حذف تصادفی برابر با ۰.۳ استفاده گردید. خروجی شبکه بازگشتی در هر گام زمانی مطابق رابطه زیر محاسبه می‌شود:

$$h_t = GRU(F_t, h_{t-1}) \quad (6)$$

در ساختار GRU، دروازه به‌روزرسانی و دروازه بازنشانی برای کنترل جریان اطلاعات استفاده می‌شوند. دروازه به‌روزرسانی مطابق رابطه زیر تعریف می‌شود:

$$z_t = \sigma(W_z x_t + U_z h_{t-1}) \quad (7)$$

و دروازه بازنشانی نیز به‌صورت زیر محاسبه می‌گردد:

$$r_t = \sigma(W_r x_t + U_r h_{t-1}) \quad (8)$$

حالت پنهان جدید شبکه مطابق رابطه زیر تولید می‌شود:

$$\tilde{h}_t = \tanh(W_h x_t + U_h (r_t \odot h_{t-1})) \quad (9)$$

و خروجی نهایی GRU برابر است با:

$$h_t = (1 - z_t) \odot h_{t-1} + z_t \odot \tilde{h}_t \quad (10)$$

پس از استخراج نمایش‌های زمانی توسط شبکه GRU دوسویه، میانگین خروجی تمامی گام‌های زمانی محاسبه شد تا یک نمایش کلی از رفتار دانشجو تولید گردد:

$$H = \frac{1}{T} \sum_{t=1}^T h_t \quad (11)$$

که در آن T تعداد گام‌های زمانی است. بردار نهایی سپس وارد بخش طبقه‌بندی شد. این بخش شامل یک لایه کاملاً متصل با ۱۲۸ نورون و تابع فعال‌ساز ReLU است. برای جلوگیری از بیش‌برازش، یک لایه حذف تصادفی با نرخ ۰.۴ نیز استفاده شد. در نهایت، خروجی مدل توسط یک لایه کاملاً متصل دوکلاسه تولید گردید تا احتمال تعلق دانشجو به کلاس «موفق» یا «در معرض خطر» محاسبه شود.

۳-۵. فرآیند آموزش مدل

برای آموزش مدل، داده‌ها به سه بخش آموزش، اعتبارسنجی و آزمون تقسیم شدند. به‌منظور جلوگیری از نشت اطلاعات، عملیات نرمال‌سازی تنها بر روی داده‌های بخش آموزش انجام شد و سپس همان پارامترها برای داده‌های اعتبارسنجی و آزمون مورد استفاده قرار گرفتند. به‌دلیل عدم‌توازن میان کلاس‌های داده، از تابع هزینه وزن‌دار کراس-آنتروپی استفاده شد تا مدل نسبت به کلاس اقلیت حساسیت بیشتری داشته باشد. برای بهینه‌سازی پارامترهای مدل از الگوریتم Adam با نرخ یادگیری 3×10^{-4} استفاده شد. همچنین برای جلوگیری از ناپایداری گرادیان‌ها، محدودسازی گرادیان با مقدار ۱ اعمال گردید. به‌منظور کاهش بیش‌برازش، راهبرد توقف زود هنگام نیز در فرآیند آموزش استفاده شد تا در صورت عدم بهبود عملکرد مدل در داده‌های اعتبارسنجی، آموزش متوقف گردد.

۴- تجزیه و تحلیل نتایج

برای ارزیابی جامع عملکرد مدل پیشنهادی، مجموعه‌ای از معیارهای ارزیابی استاندارد در مسائل طبقه‌بندی دودویی و به‌ویژه سیستم‌های هشدار زودهنگام آموزشی مورد استفاده قرار گرفت. این معیارها شامل صحت، دقت^۱، فراخوانی، AUC، F1-score، PR-AUC، MCC، Cohen's Kappa، صحت متوازن^۲ و معیار Brier هستند که هر کدام جنبه متفاوتی از عملکرد مدل را نشان می‌دهند.

جدول (۱): ارزیابی عملکرد مدل پیشنهادی بر اساس مجموعه‌ای از معیارهای ارزیابی

مقدار	معیار ارزیابی
67.37%	Accuracy
68.05%	Precision
60.39%	Recall
63.99%	F1-score
73.54%	AUC
74.17%	PR-AUC
0.346	MCC
0.344	Cohen's Kappa
67.10%	Balanced Accuracy
0.209	Brier Score

نتایج به‌دست‌آمده نشان می‌دهد که مدل پیشنهادی توانسته است یک عملکرد پایدار و قابل اتکا در مسئله پیش‌بینی زودهنگام افت تحصیلی ارائه دهد. مقدار صحت برابر با ۶۷/۳۷٪ نشان‌دهنده توانایی کلی مدل در طبقه‌بندی صحیح نمونه‌ها است. با توجه به اینکه مسئله دارای ماهیت نامتوازن است و کلاس دانشجویان در معرض خطر معمولاً پیچیدگی رفتاری بیشتری دارد، این مقدار از صحت را می‌توان در سطح قابل قبول ارزیابی کرد.

از منظر دقت برابر با ۶۸/۰۵٪، مدل نشان می‌دهد که در میان نمونه‌هایی که به‌عنوان «دانشجوی در معرض خطر» پیش‌بینی شده‌اند، درصد نسبتاً خوبی واقعاً متعلق به این کلاس هستند. این موضوع در کاربردهای آموزشی اهمیت دارد، زیرا باعث کاهش مداخلات اشتباه آموزشی برای دانشجویان سالم می‌شود.

در مقابل، مقدار فراخوانی برابر با ۶۰/۳۹٪ نشان می‌دهد که مدل توانسته بخش قابل توجهی از دانشجویان در معرض خطر را شناسایی کند، هرچند هنوز تعدادی از موارد مثبت واقعی از دید مدل پنهان مانده‌اند. این موضوع در مسائل پیش‌بینی زودهنگام که معمولاً با عدم‌توازن داده‌ها همراه هستند، یک رفتار رایج محسوب می‌شود. از آنجا که در سیستم‌های هشدار زودهنگام، افزایش فراخوانی معمولاً با کاهش دقت همراه است، این نتیجه نشان‌دهنده وجود یک موازنه طبیعی در عملکرد مدل است. بنابراین، با تنظیم آستانه تصمیم‌گیری می‌توان حساسیت مدل را در جهت افزایش پوشش دانشجویان در معرض خطر بهبود داد.

^۱. Precision

^۲. Balanced Accuracy

مقدار F1-score برابر با ۶۳.۹۹٪ نشان‌دهنده تعادل نسبی میان دقت و فراخوانی است و نشان می‌دهد که مدل در مجموع توانسته است بین دقت پیش‌بینی و پوشش کلاس اقلیت تعادل برقرار کند. با این حال، عدم تقارن میان دقت و فراخوانی نشان می‌دهد که تنظیم آستانه تصمیم‌گیری می‌تواند نقش مهمی در بهبود عملکرد نهایی داشته باشد. از سوی دیگر، مقدار AUC برابر با ۷۳.۵۴٪ نشان‌دهنده توانایی مناسب مدل در تفکیک بین دو کلاس در تمام آستانه‌های تصمیم‌گیری است. این معیار یکی از مهم‌ترین شاخص‌ها در مسائل طبقه‌بندی نامتوازن محسوب می‌شود و نشان می‌دهد که مدل قادر است الگوهای رفتاری دانشجویان موفق و در معرض خطر را به‌طور معناداری از هم جدا کند.

همچنین PR-AUC برابر با ۷۴.۱۷٪ اهمیت ویژه‌ای دارد، زیرا در داده‌های نامتوازن، این معیار نسبت به AUC اطلاعات دقیق‌تری درباره عملکرد مدل در کلاس اقلیت ارائه می‌دهد. مقدار به‌دست‌آمده نشان می‌دهد که مدل در شناسایی دانشجویان در معرض خطر از پایداری مناسبی برخوردار است.

در بخش معیارهای توافقی، مقدار MCC برابر با ۰.۳۴۶ و Cohen's Kappa برابر با ۰.۳۴۴ نشان‌دهنده همبستگی متوسط بین پیش‌بینی‌های مدل و برچسب‌های واقعی است. این دو معیار نسبت به صحت اطلاعات دقیق‌تری ارائه می‌دهند، زیرا اثر عدم‌توازن داده‌ها را نیز در نظر می‌گیرند. صحت متوازن برابر با ۶۷.۱۰٪ نشان می‌دهد که مدل عملکرد نسبتاً متعادلی در هر دو کلاس دارد و دچار سوگیری شدید به سمت کلاس اکثریت نشده است. در نهایت، مقدار معیار Brier برابر با ۰.۲۰۹ نشان‌دهنده کیفیت قابل قبول در کالیبراسیون احتمالات خروجی مدل است، به این معنا که احتمال‌های پیش‌بینی شده توسط مدل تا حد خوبی با واقعیت تجربی هم‌راستا هستند.

۱-۴. مقایسه با روش‌های پایه

برای ارزیابی مدل پیشنهادی و مقایسه با روش‌های پایه، از معیارهای استاندارد استفاده شد. نتایج حاصل از آزمایش‌ها در جدول ۲ ارائه شده است.

جدول (۲): مقایسه عملکرد مدل پیشنهادی با روش‌های پایه

مدل	Accuracy	F1-score	AUC
رگرسیون لجستیک	58.45%	66.37%	60.55%
جنگل تصادفی	62.79%	68.83%	65.15%
ماشین بردار پشتیبان	61.86%	66.58%	63.48%
XGBoost	62.61%	67.48%	64.95%
مدل پیشنهادی	67.37%	63.99%	73.54%

نتایج جدول ۲ نشان می‌دهد که مدل پیشنهادی توانسته است در معیارهای کلیدی عملکرد، به‌ویژه صحت و AUC، نسبت به تمامی روش‌های کلاسیک مورد مقایسه عملکرد بهتری ارائه دهد. مقدار صحت برابر با ۶۷/۳۷ بیانگر توانایی مناسب مدل در طبقه‌بندی صحیح نمونه‌های دانشجویان است و در مقایسه با بهترین روش کلاسیک (جنگل تصادفی) با دقت ۶۲/۸۰ بهبود قابل توجهی را نشان می‌دهد.

مهم‌ترین مزیت مدل پیشنهادی در مقدار AUC مشاهده می‌شود که برابر با ۷۳/۵۴ است. این مقدار نسبت به سایر روش‌ها افزایش محسوسی دارد و نشان می‌دهد که مدل پیشنهادی در تفکیک کلاس‌های «در معرض خطر» و «غیر در معرض خطر» در تمامی آستانه‌های تصمیم‌گیری عملکرد پایدارتری دارد. از سوی دیگر، مقدار دقت برابر با ۶۸/۰۶

نشان می‌دهد که مدل در پیش‌بینی نمونه‌های مثبت (دانشجویان در معرض خطر) نرخ خطای نسبتاً کنترل‌شده‌ای دارد.

مقدار فراخوانی برابر با ۶۰/۳۹٪ نشان می‌دهد که بخشی از دانشجویان در معرض خطر توسط مدل شناسایی نشده‌اند. با این حال، این نتیجه باید در چارچوب موازنه بین فراخوانی و دقت تفسیر شود، زیرا افزایش حساسیت مدل معمولاً منجر به افزایش خطای مثبت کاذب می‌شود. از آنجا که مدل پیشنهادی دارای AUC نسبتاً بالا (۰/۷۳/۵۴) است، می‌توان نتیجه گرفت که کاهش فراخوانی ناشی از انتخاب آستانه پیش‌فرض تصمیم‌گیری بوده و نه ضعف ذاتی مدل در یادگیری الگوها.

اگرچه برخی روش‌های کلاسیک مانند جنگل تصادفی در معیار F1 مقدار کمی بالاتر از مدل پیشنهادی نشان می‌دهند، این اختلاف باید با توجه به ماهیت نامتوازن داده‌ها و وابستگی معیار F1 به آستانه تصمیم‌گیری تفسیر شود. در مسائل هشدار زودهنگام آموزشی، معیارهایی مانند AUC و PR-AUC اهمیت بیشتری دارند، زیرا توانایی مدل در رتبه‌بندی نمونه‌ها را مستقل از آستانه ارزیابی می‌کنند. از این منظر، مدل پیشنهادی با دستیابی به AUC بالاتر نسبت به تمامی روش‌های پایه، نشان‌دهنده توانایی بهتر در تفکیک دانشجویان در معرض خطر از سایرین در تمام آستانه‌های تصمیم‌گیری است.

در مجموع، نتایج تجربی نشان می‌دهد که مدل پیشنهادی در مقایسه با روش‌های کلاسیک، به‌ویژه در معیارهای AUC و صحت، عملکرد بهتری ارائه می‌دهد و برای مسئله پیش‌بینی زودهنگام افت تحصیلی در محیط‌های یادگیری الکترونیکی گزینه مناسبی محسوب می‌شود. با این حال، کاهش نسبی در فراخوانی نشان می‌دهد که همچنان فضای بهبود برای افزایش حساسیت مدل نسبت به دانشجویان در معرض خطر وجود دارد که می‌تواند در مطالعات آینده مورد توجه قرار گیرد.

۵- بحث و نتیجه‌گیری

در این پژوهش، یک چارچوب مبتنی بر شبکه عصبی بازگشتی از نوع GRU دوسویه برای پیش‌بینی زودهنگام عملکرد تحصیلی دانشجویان در محیط‌های یادگیری الکترونیکی ارائه شد. هدف اصلی پژوهش، بهره‌گیری از داده‌های رفتاری و تعاملی دانشجویان در هفته‌های ابتدایی دوره آموزشی و استخراج الگوهای زمانی مؤثر در شناسایی دانشجویان در معرض افت تحصیلی بود. برای این منظور، مجموعه‌ای از ویژگی‌های رفتاری و زمانی از داده‌های تعامل دانشجویان با محیط یادگیری مجازی استخراج و تنها داده‌های مربوط به ۴۲ روز نخست دوره آموزشی به‌عنوان ورودی مدل مورد استفاده قرار گرفت.

نتایج آزمایش‌ها نشان داد که مدل پیشنهادی توانسته است در پیش‌بینی زودهنگام دانشجویان در معرض خطر عملکرد قابل قبولی ارائه دهد. این مدل به صحت ۶۷/۳۷ درصد، صحت متوازن ۶۷/۱۰ درصد، مقدار AUC برابر با ۷۳/۵۴ درصد و PR-AUC برابر با ۷۴/۱۷ درصد دست یافت. همچنین مقادیر MCC برابر با ۰/۳۴۶ و ضریب کاپای کوهن برابر با ۰/۳۴۴ بیانگر وجود توافق مناسب میان پیش‌بینی‌های مدل و برچسب‌های واقعی هستند. این نتایج نشان می‌دهد که مدل پیشنهادی قادر است با استفاده از داده‌های محدود مربوط به مراحل ابتدایی دوره آموزشی، الگوهای رفتاری مرتبط با موفقیت یا افت تحصیلی دانشجویان را شناسایی کند.

مقایسه نتایج پژوهش حاضر با مطالعات پیشین نشان می‌دهد که یافته‌های این پژوهش با نتایج پژوهش‌های هلستا و همکاران (۲۰۱۸)، سا (۲۰۱۶)، سالاری و همکاران (۲۰۲۴) و النسیان و همکاران (۲۰۲۵) همخوانی دارد؛ زیرا تمامی این مطالعات بر اهمیت داده‌های رفتاری و تعاملی دانشجویان در پیش‌بینی عملکرد تحصیلی تأکید کرده‌اند. همچنین نتایج این پژوهش مؤید یافته‌های پیچ و همکاران (۲۰۱۵) است که نشان دادند مدل‌های مبتنی بر یادگیری عمیق توانایی مناسبی در استخراج الگوهای پنهان موجود در داده‌های آموزشی دارند. به‌طور کلی، یافته‌های پژوهش حاضر نشان می‌دهد که استفاده از داده‌های رفتاری ثبت‌شده در سامانه‌های یادگیری الکترونیکی می‌تواند مبنای مناسبی برای توسعه سامانه‌های هوشمند پیش‌بینی عملکرد تحصیلی باشد.

از منظر تحلیلی، عملکرد مدل پیشنهادی را می‌توان به توانایی معماری Bi-GRU در مدل‌سازی وابستگی‌های زمانی میان فعالیت‌های دانشجویان نسبت داد. برخلاف بسیاری از الگوریتم‌های کلاسیک یادگیری ماشین که داده‌ها را به‌صورت ایستا تحلیل می‌کنند، شبکه‌های بازگشتی قادرند تغییرات رفتاری دانشجویان را در طول زمان در نظر بگیرند و روابط زمانی میان تعاملات آموزشی را استخراج کنند. استفاده از ساختار دوسویه نیز موجب می‌شود اطلاعات موجود در کل توالی رفتاری دانشجو به شکل مؤثرتری مورد استفاده قرار گیرد. نتایج به‌دست‌آمده نشان می‌دهد که حتی با استفاده از داده‌های ۴۲ روز نخست دوره آموزشی نیز می‌توان دانشجویان در معرض خطر را پیش از پایان نیمسال شناسایی کرد و امکان انجام مداخلات آموزشی زود هنگام را فراهم ساخت.

یافته‌های این پژوهش دارای کاربردهای عملی مهمی برای دانشگاه‌ها، مراکز آموزش الکترونیکی و طراحان سامانه‌های مدیریت یادگیری است. مدل پیشنهادی می‌تواند به‌عنوان هسته اصلی سامانه‌های هشدار زود هنگام آموزشی مورد استفاده قرار گیرد تا دانشجویان دارای ریسک افت تحصیلی در مراحل ابتدایی دوره شناسایی شوند. همچنین نتایج حاصل می‌تواند به مدرسان و مدیران آموزشی در طراحی برنامه‌های حمایتی، مشاوره‌ای و آموزشی هدفمند برای دانشجویان در معرض خطر کمک کند و در نهایت موجب بهبود کیفیت یادگیری و کاهش نرخ افت تحصیلی شود. برای پژوهش‌های آینده، استفاده از معماری‌های پیشرفته‌تر یادگیری عمیق مبتنی بر سازوکار توجه^۱، ترکیب داده‌های رفتاری با اطلاعات ارزیابی‌ها و ویژگی‌های جمعیت‌شناختی، بررسی روش‌های تفسیرپذیری مدل‌های عمیق و ارزیابی مدل بر روی مجموعه داده‌های آموزشی دیگر پیشنهاد می‌شود. همچنین مطالعه تأثیر بازه‌های زمانی مختلف بر عملکرد پیش‌بینی زود هنگام می‌تواند زمینه مناسبی برای توسعه پژوهش‌های آتی فراهم کند.

از جمله محدودیت‌های این پژوهش می‌توان به استفاده از تنها یک مجموعه داده استاندارد، محدود بودن اطلاعات به تعاملات ثبت‌شده در محیط یادگیری مجازی و عدم دسترسی به برخی عوامل فردی، انگیزشی و روان‌شناختی مؤثر بر عملکرد تحصیلی دانشجویان اشاره کرد. علاوه بر این، تمرکز پژوهش بر داده‌های ۴۲ روز نخست دوره آموزشی سبب می‌شود بخشی از اطلاعات رفتاری دانشجویان در ادامه دوره در مدل لحاظ نشود. با وجود این محدودیت‌ها، نتایج پژوهش نشان داد که رویکرد پیشنهادی می‌تواند چارچوب مناسبی برای توسعه سامانه‌های هوشمند پیش‌بینی زود هنگام و حمایت آموزشی در محیط‌های یادگیری الکترونیکی فراهم آورد.

^۱ . Attention

References:

- Alnasyan, B., Basher, M., Alassafi, M., & Alnasyan, K. (2025). Kanformer: an attention-enhanced deep learning model for predicting student performance in virtual learning environments. *Social Network Analysis and Mining*, 15(1), 25.
- Bengio, Y., Simard, P., & Frasconi, P. (1994). Learning long-term dependencies with gradient descent is difficult. *IEEE transactions on neural networks*, 5(2), 157-166.
- Breiman, L. (2001). Random forests. *Machine learning*, 45(1), 5-32.
- Cho, K., Van Merriënboer, B., Gulçehre, Ç., Bahdanau, D., Bougares, F., Schwenk, H., & Bengio, Y. (2014, October). Learning phrase representations using RNN encoder-decoder for statistical machine translation. In *Proceedings of the 2014 conference on empirical methods in natural language processing (EMNLP)* (pp. 1724-1734).
- Cortes, C., & Vapnik, V. (1995). Support-vector networks. *Machine learning*, 20(3), 273-297.
- Ferguson, R. (2012). Learning analytics: drivers, developments and challenges. *International journal of technology enhanced learning*, 4(5-6), 304-317.
- Hlosta, M., Herrmannova, D., Vachova, L., Kuzilek, J., Zdrahal, Z., & Wolff, A. (2018). Modelling student online behaviour in a virtual learning environment. *arXiv preprint arXiv:1811.06369*.
- Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. *Neural Computation*, 9(8), 1735-1780.
- Hosmer Jr, D. W., Lemeshow, S., & Sturdivant, R. X. (2013). *Applied logistic regression*. John Wiley & Sons.
- Kuzilek, J., Hlosta, M., & Zdrahal, Z. (2017). Open university learning analytics dataset. *Scientific data*, 4(1), 170171.
- Piech, C., Bassen, J., Huang, J., Ganguli, S., Sahami, M., Guibas, L. J., & Sohl-Dickstein, J. (2015). Deep knowledge tracing. *Advances in neural information processing systems*. Association for Computing Machinery, 201-204.
- Quinlan, J. R. (2014). *C4. 5: programs for machine learning*. Elsevier.
- Romero, C., & Ventura, S. (2010). Educational data mining: a review of the state of the art. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (applications and reviews)*, 40(6), 601-618.
- Saa, A. A. (2016). Educational data mining & students' performance prediction. *International journal of advanced computer science and applications*, 7(5).
- Salari, M., Radfar, R., & Faghihi, M. (2024). Predicting students' performance using machine learning algorithms and educational data mining (a case study of Shahed University). *Business Intelligence Management Studies*, 13(47), 315-366.
- Siemens, G., & Baker, R. S. D. (2012, April). Learning analytics and educational data mining: towards communication and collaboration. In *Proceedings of the 2nd international conference on learning analytics and knowledge* (pp. 252-254).
- Tomasevic, N., Gvozdenovic, N., & Vranes, S. (2020). An overview and comparison of supervised data mining techniques for student exam performance prediction. *Computers & education*, 143, 103676.