

سنجش و ارزیابی روش‌ها و فنون تدریس اعضای هیأت علمی در آموزش دانشگاهی به منظور تضمین کیفیت آموزش

❖ محمد مرادی^۱*

صفحه: ۱۹۴-۱۷۰

چکیده

هدف این پژوهش، ارائه چارچوبی داده‌محور برای سنجش و ارزیابی روش‌ها و فنون تدریس اعضای هیأت علمی در آموزش دانشگاهی با تمرکز بر نقش این ارزیابی‌ها در تضمین کیفیت آموزشی است. برخلاف اغلب مطالعات پیشین که عمدتاً بر تحلیل‌های توصیفی یا نظرات خبرگان متکی بوده‌اند، در این پژوهش از داده‌های واقعی حاصل از ارزشیابی تدریس توسط دانشجویان و نمرات عملکرد آموزشی آنان استفاده شده است. جامعه آماری شامل داده‌های دروس ارائه‌شده در دانشگاه‌های قم و بزرگمهر فائز طی سال‌های ۱۳۹۵ تا ۱۴۰۴ است. ابتدا معیارهای ارزشیابی تدریس با استناد به دستورالعمل‌های وزارت علوم استخراج شد. سپس با بهره‌گیری از الگوریتم‌های داده‌کاوی، به‌ویژه خوشه‌بندی، دروس با الگوهای ارزشیابی مشابه گروه‌بندی شدند و معیارهای مرتبط با خوشه‌های دارای بالاترین عملکرد تحصیلی شناسایی گردید. در ادامه، وزن و اهمیت نسبی هر معیار به صورت داده‌محور محاسبه و نرمال‌سازی شد. نتایج تحلیل نشان داد که معیار «واکنش منطقی و معقول به پیشنهادها، انتقادات و دیدگاه‌های دانشجویان» با وزن ۰.۳۰۷۳۰۳ بیشترین تأثیر را بر بازدهی آموزشی دانشجویان دارد. پس از آن، معیار «تسلط استاد بر موضوع درس» با وزن ۰.۳۰۷۷۷۳ در رتبه دوم قرار گرفت. در مرحله پایانی، راهکارهایی برای بهبود معیارهای با اولویت بالا تدوین و با استفاده از تکنیک‌های تصمیم‌گیری چندمعیاره شامل AHP اولویت‌بندی گردید. یافته‌ها نشان می‌دهد که رویکرد تلفیقی مبتنی بر داده‌کاوی و تحلیل تصمیم‌محور، می‌تواند ابزاری کارآمد و دقیق در جهت بهبود نظام ارزشیابی آموزشی و ارتقای کیفیت تدریس در آموزش عالی فراهم آورد.

واژگان کلیدی: ارزشیابی تدریس، کیفیت آموزش، سیاست‌گذاری آموزش، داده‌کاوی، تصمیم‌گیری چندمعیاره

■ پذیرش مقاله: ۱۴۰۳/۰۹/۰۴

■ دریافت مقاله: ۱۴۰۳/۰۶/۱۹

^۱ استادیار گروه مهندسی کامپیوتر، دانشکده مهندسی، دانشگاه بزرگمهر فائز، قزوین، ایران.

مقدمه

در دهه‌های اخیر، آموزش عالی با تحولات عمیقی در سطح جهانی مواجه شده است. رشد سریع دانش، گسترش فناوری‌های نوین، و نیاز به تربیت نیروی انسانی ماهر موجب شده تا نقش دانشگاه‌ها در توسعه پایدار، بیش از پیش مورد توجه قرار گیرد. در این میان، کیفیت آموزش به عنوان یکی از ارکان اصلی موفقیت نظام آموزش عالی شناخته می‌شود، و یکی از مهم‌ترین عوامل موثر بر آن، روش‌ها و فنون تدریس اعضای هیأت علمی است (نصیر و همکاران^۱، ۲۰۲۱؛ خضری و همکاران^۲، ۲۰۲۵؛ گامپ^۳، ۲۰۰۷).

روش تدریس نه تنها بر میزان درک، یادگیری و انگیزش دانشجویان تأثیر می‌گذارد، بلکه می‌تواند عاملی تعیین‌کننده در دستیابی به اهداف آموزشی و تربیت فارغ‌التحصیلانی توانمند باشد. از این رو، سنجش و ارزیابی دقیق و علمی عملکرد تدریس اعضای هیأت علمی، ضرورتی انکارناپذیر در نظام تضمین کیفیت آموزش در دانشگاه‌ها محسوب می‌شود (ایکرامو و تورسونبایوا^۴، ۲۰۲۲؛ بازوهوری و همکاران^۵، ۲۰۲۱؛ بارخدا و همکاران، ۱۴۰۲؛ نگوین^۶، ۲۰۲۴؛ شهامت و همکاران، ۱۴۰۲).

با این حال، نظام آموزش عالی ایران با چالش‌های خاصی در حوزه ارزیابی کیفیت تدریس مواجه است؛ از جمله مقاومت برخی اساتید در برابر فرآیندهای ارزشیابی، ناکارآمدی نسبی سامانه‌های موجود، و محدودیت‌های ناشی از وابستگی صرف به نظرات دانشجویان (ارسلان و همکاران^۷، ۲۰۲۴؛ اسپورن و همکاران^۸، ۲۰۱۳). بر اساس گزارش رسمی وزارت علوم در سال ۱۴۰۲، بیش از ۶۷٪ اعضای هیأت علمی، از کارایی و عادلانه بودن نظام‌های فعلی ارزشیابی ابراز نارضایتی کرده‌اند؛ این آمار به روشنی بیانگر نیاز به بازنگری و طراحی چارچوب‌های نوین مبتنی بر داده و تحلیل‌های علمی است (سند تحول راهبردی آموزش عالی، ۱۴۰۲).

در پاسخ به این ضرورت، پژوهش حاضر با رویکردی داده‌محور، به دنبال توسعه چارچوبی نظام‌مند برای تحلیل و ارزیابی شیوه‌های تدریس اعضای هیأت علمی در دانشگاه‌های کشور است. در این مسیر، داده‌های واقعی حاصل از سامانه‌های ارزشیابی آموزشی و عملکرد تحصیلی دانشجویان مورد تحلیل قرار می‌گیرند تا معیارهای کلیدی مؤثر بر کیفیت تدریس شناسایی، وزن‌دهی و اولویت‌بندی شوند. همچنین تلاش می‌شود

1. Nseer et al
2. Khezri et al
3. Gump
4. Ikramov & Tursunbaeva
5. Bazohoori et al
6. Nguyen
7. Arslan et al
8. Spooren et al

راهکارهایی عملی و مبتنی بر شواهد برای بهبود کیفیت تدریس استخراج گردد. در راستای دستیابی به این هدف، سوالات پژوهش به صورت زیر تدوین شده‌اند:

۱. معیارهای کلیدی مؤثر بر کیفیت تدریس در داده‌های واقعی کدامند؟
۲. چگونه می‌توان وزن و اهمیت این معیارها را به صورت داده‌محور محاسبه کرد؟
۳. راهکارهای اولویت‌دار برای بهبود کیفیت تدریس، مبتنی بر این معیارها، چه هستند؟
۴. چارچوب داده‌محور پیشنهادی چگونه می‌تواند در طراحی سیاست‌ها و برنامه‌های ارتقاء کیفیت آموزش در نظام آموزش عالی مورد استفاده قرار گیرد؟

پاسخ به این سوالات با بهره‌گیری از ترکیب تکنیک‌های داده‌کاوی (به‌ویژه خوشه‌بندی) و روش‌های تصمیم‌گیری چندمعیاره^۱ (MADM) صورت خواهد گرفت. به این ترتیب، پژوهش ضمن ارائه تصویری واقعی و تحلیلی از وضعیت تدریس در دانشگاه‌ها، می‌تواند بستری علمی برای تدوین سیاست‌های ارتقاء کیفیت، طراحی برنامه‌های توسعه حرفه‌ای اعضای هیأت علمی، و اصلاح نظام ارزشیابی آموزشی در آموزش عالی کشور فراهم آورد.

در ادامه، بخش دوم مقاله به مروری انتقادی بر پیشینه پژوهش اختصاص دارد. در بخش سوم، روش‌شناسی تحقیق در قالب فازهای اجرایی تشریح شده است. یافته‌های پژوهش در بخش چهارم ارائه و تحلیل شده‌اند. سرانجام، در بخش پنجم، نتیجه‌گیری و پیشنهادهایی برای آینده ارائه می‌شود.

۲- پیشینه پژوهش

در این بخش به بررسی پژوهش‌های مرتبط پرداخته شده است. مهدی و همکاران (۱۴۰۰) به بررسی تطبیقی نظام ارزشیابی آموزشی اعضای هیأت علمی در دانشگاه‌های علوم پزشکی ایران پرداختند. بدین منظور از روش مطالعات تطبیقی با استفاده از الگوی جرج بردی در آن استفاده گردید. پس از بررسی و مطالعه داده‌های بدست آمده، نویسندگان نتیجه گرفتند استفاده از ارزشیابی کمی و کیفی در کنار استفاده از منابع مختلف و مقایسه با سیستم‌های ارزشیابی سایر دانشگاه‌ها منجر به بهبود کیفیت ارزشیابی اساتید خواهد شد. ایزدی‌ثابت و همکاران (۱۴۰۱) در پژوهشی به طراحی، اعتباریابی و پایاسازی ابزار ارزشیابی عملکرد آموزشی اعضای هیأت علمی دانشگاه‌های علوم پزشکی پرداختند. در این مطالعه از ترکیبی از روش‌های پژوهشی نظیر مطالعه مروری، مصاحبه با صاحب‌نظران و نیز طراحی و استانداردسازی ابزار ارزشیابی و نیز رویکردهای کیفی استفاده گردید. درنهایت، نویسندگان ابزار ارزشیابی دارای ۷۴ مولفه ارزشیابی در ۵ حیطه

1. Multi-Criteria Decision-Making (MCDM)

موسوم به اهداف آموزشی، محتوای آموزشی، فرایند یاددهی-یادگیری، ارزشیابی آموزشی، آموزش بالینی/آزمایشگاهی ارائه کردند.

یو و همکاران^۱ (۲۰۲۲) در یک مطالعه با عنوان «ارزیابی کیفیت تدریس بر اساس الگوریتم بهبودیافته شبکه عصبی BP، از مدل شبکه عصبی پیش‌خور برای تحلیل کیفیت آموزش استفاده کردند. این مدل در ارزیابی ترکیبی دوره‌های SPOC و MOOC در یکی از دانشگاه‌های چین به کار گرفته شد و نتایج نشان داد که مدل پیشنهادی دارای نرخ همگرایی بالا و خطای پایین بوده است. تحلیل‌های انجام‌شده مؤید آن است که خروجی این سیستم قادر است کیفیت تدریس دروس را با دقت مناسبی بازتاب دهد.

یو (۲۰۲۲) به موضوع ارزیابی و تحلیل کیفیت تدریس پرداخت. او روش‌های ارزیابی کیفیت تدریس، روش‌های تعیین شاخص‌های کیفیت و رویکردهای مختلف در فرآیند ارزیابی به همراه مزایا و معایب آن‌ها تجزیه و تحلیل کرد. علاوه بر این، توجه ویژه‌ای به اهمیت استفاده از فناوری‌های نوین و مدرن برای افزایش کیفیت تدریس شد.

منسها و همکاران^۲ (۲۰۲۳) به بررسی عواملی که سبب کاهش کیفیت آموزش می‌شود، پرداختند. بدین منظور از مطالعات میدانی و طراحی پرسشنامه استفاده شد. طبق یافته‌ها، «بیماری»، «مشکلات مالی ناشی از حقوق پایین» و «نظارت ضعیف»، از دلایل اصلی کاهش کیفیت آموزش شناسایی شد.

بابلو^۳ (۲۰۲۴) در مطالعه‌ای با عنوان «یادگیری ماشین در ارزیابی خودکار: ارتقاء عینیت و کارایی در ارزیابی‌های آموزشی»، کاربرد الگوریتم‌های یادگیری ماشین در بهبود دقت، سرعت و بی‌طرفی نظام‌های ارزیابی آموزشی را بررسی کرد. این مقاله با مرور ادبیات، تحلیل فناوری‌های موجود و بررسی مطالعات موردی، نشان می‌دهد که سیستم‌های خودکار مبتنی بر یادگیری ماشین می‌توانند جایگزینی مؤثر برای روش‌های سنتی ارزیابی باشند. با این حال، محدودیت‌هایی مانند سوگیری الگوریتم‌ها، دغدغه‌های اخلاقی و لزوم نظارت انسانی نیز مطرح شده است.

سعید^۴ (۲۰۲۴) به بررسی معیارهای ارزشیابی تدریس در دانشگاه پرداخت. بدین منظور از مطالعات میدانی و مصاحبه و مشاهده استفاده شد. بر اساس نتایج بدست آمده، معیارهای «تسلط بر موضوع»، «تجربه و به‌روز ماندن با روندهای جدید»، «تعامل فعال با منابع یادگیری» و «شیوه‌های بازخورد و تأمل» جزو معیارهای کلیدی

1. Yu et al

2. Mensaha et al

3. Bablu

4. Saeed

انتخاب شدند که اثربخشی تدریس را بهبود می‌بخشند. جدول ۱، پیشینه پژوهش را بر اساس نام پژوهشگران، سال انتشار، موضوع پژوهش، روش پژوهش و نتایج آن نشان می‌دهد.

جدول ۱: پیشینه پژوهش

پژوهشگر / سال	موضوع پژوهش	روش پژوهش	نتایج کلیدی
مهدی و همکاران (۱۴۰۰)	بررسی تطبیقی نظام ارزشیابی اعضای هیأت علمی در دانشگاه‌های علوم پزشکی	مطالعات تطبیقی (الگوی جرج بردی)	استفاده ترکیبی از ارزشیابی کمی و کیفی و منابع مختلف موجب بهبود کیفیت ارزیابی می‌شود.
ایزدی‌ثابت و همکاران (۱۴۰۱)	طراحی و اعتباریابی ابزار ارزشیابی عملکرد آموزشی اعضای هیأت علمی	روش ترکیبی (مروری، مصاحبه، طراحی ابزار)	طراحی ابزار با ۷۴ مولفه در ۵ حوزه شامل اهداف، محتوا، فرآیند یاددهی-یادگیری
یو و همکاران (۲۰۲۲)	ارزیابی کیفیت تدریس با استفاده از الگوریتم بهبود یافته شبکه عصبی BP و SPOC	مدل‌سازی شبکه عصبی و تحلیل داده‌های MOOC	مدل دارای همگرایی سریع و خطای پایین است و نتایج دقیق‌تری از کیفیت تدریس ارائه می‌دهد.
یو (۲۰۲۲)	ارزیابی و تحلیل کیفیت تدریس و شاخص‌های آن	مروری و تحلیلی	بررسی مزایا و معایب رویکردهای مختلف ارزیابی، تأکید بر نقش فناوری‌های نوین در تدریس
منسها و همکاران (۲۰۲۳)	عوامل کاهش کیفیت آموزش	مطالعه میدانی - پرسشنامه	بیماری، مشکلات مالی و ضعف نظارت از مهم‌ترین عوامل کاهش کیفیت آموزش
بابلو (۲۰۲۴)	کاربرد یادگیری ماشین در ارزیابی خودکار آموزشی	مرور ادبیات، تحلیل فناوری‌ها، مطالعات موردی	بهبود دقت و سرعت ارزیابی‌ها اما نیاز به نظارت انسانی و بررسی مسائل اخلاقی دارد.
سعید (۲۰۲۴)	شناسایی معیارهای اثربخش تدریس در آموزش عالی	میدانی - مصاحبه و مشاهده	تسلط بر موضوع، روزآمد بودن، تعامل فعال و شیوه‌های بازخورد به‌عنوان عوامل مؤثر در اثربخشی تدریس

بررسی پیشینه پژوهش نشان می‌دهد که اغلب مطالعات پیشین یا به تحلیل رابطه معنادار میان یک یا چند معیار ارزشیابی و پیامدهای آموزشی نظیر عملکرد تحصیلی دانشجویان پرداخته‌اند، یا با اتکا به نظرات

خبرگان و به کارگیری تکنیک‌های تصمیم‌گیری چندمعیاره، به تعیین وزن و اهمیت معیارهای ارزشیابی اقدام کرده‌اند. با این حال، در بیشتر این مطالعات، وزن‌دهی به معیارها بر پایه داده‌های واقعی حاصل از ارزشیابی تدریس اعضای هیأت علمی توسط دانشجویان و همچنین عملکرد آموزشی دانشجویان (نمرات دروس)، به صورت داده‌محور مورد بررسی قرار نگرفته است. همچنین در برخی پژوهش‌های صورت گرفته مانند مطالعه یو و همکاران (۲۰۲۲)، که از شبکه‌های عصبی برای پیش‌بینی کیفیت تدریس استفاده شده است، انتخاب بهینه تعداد نرون‌ها در لایه پنهان و نیاز به تنظیمات پیچیده، دقت مدل را وابسته به پارامترهای حساس و داده‌های بزرگ کرده است. همچنین، اگرچه این روش در محیط‌های برخط مانند SPOC و MOOC نتایج نسبتاً خوبی ارائه داده، اما در محیط دانشگاهی سنتی با داده‌های محدود و پراکنده، عملکرد آن چالش‌برانگیز است. پژوهش حاضر با بهره‌گیری از روش‌های داده‌کاوی شفاف‌تر مانند درخت تصمیم، نه تنها امکان تفسیر ساده و مستقیم معیارهای مؤثر بر کیفیت تدریس را فراهم می‌آورد، بلکه برای سیاست‌گذاری آموزشی نیز قابلیت کاربرد بالاتری دارد. همچنین، رویکرد ترکیبی داده‌کاوی و AHP در این پژوهش، امکان وزن‌دهی عینی‌تر معیارها را فراهم می‌کند. به عنوان مثالی از پژوهش دیگر مانند مطالعه بابلو (۲۰۲۴)، اگرچه مزایایی چون افزایش سرعت، دقت و عینیت ارزیابی‌ها را نشان می‌دهد، اما به دلیل استفاده از مدل‌های پیچیده و اغلب جعبه‌سیاه، با چالش‌هایی نظیر نبود شفافیت در فرایند تصمیم‌گیری، احتمال سوگیری داده‌محور و نیاز مداوم به نظارت انسانی مواجه است. در مقابل، روش ارائه‌شده در این مقاله با بهره‌گیری از تحلیل داده‌های واقعی تدریس و استفاده از مدل‌های قابل تفسیر، نه تنها امکان درک بهتر سازوکارهای اثرگذار بر کیفیت آموزش را فراهم می‌آورد، بلکه به مدیران دانشگاهی امکان ارائه بازخورد دقیق و عملی به اعضای هیأت علمی را می‌دهد. بدین ترتیب، پژوهش حاضر با رویکرد داده‌محور و تفسیرپذیر، شکاف موجود میان تحلیل‌های فناورانه و سیاست‌گذاری آموزشی را به صورت مؤثرتری پوشش می‌دهد.

افزون بر این، در پژوهش‌های پیشین به ندرت به ارائه راهکارهای عملیاتی برای ارتقای کیفیت آموزش بر اساس اولویت و اهمیت معیارهای شناسایی‌شده پرداخته شده است. پژوهش حاضر در تلاش است با استخراج و تحلیل معیارهای ارزشیابی تدریس اعضای هیأت علمی، جمع‌آوری داده‌های میدانی شامل نتایج ارزشیابی دانشجویان و نمرات آن‌ها، تعیین وزن و اهمیت هر معیار با رویکرد داده‌محور، و در نهایت ارائه راهکارهای بهبود کیفیت آموزشی مبتنی بر مهم‌ترین معیارها، چارچوبی کاربردی در راستای ارتقاء نظام ارزیابی و کیفیت آموزش دانشگاهی ارائه دهد.

روش‌شناسی تحقیق

پژوهش حاضر از نظر هدف، کاربردی و از حیث ماهیت و روش، توصیفی-تحلیلی با رویکرد داده‌محور محسوب می‌شود. در این مطالعه، فرآیند تحقیق در قالب یک زنجیره نظام‌مند طراحی شده که از شناسایی مفاهیم نظری آغاز و به تحلیل داده‌ها و ارائه راهکارهای اجرایی ختم می‌شود. در مرحله آغازین، معیارهای ارزشیابی تدریس اعضای هیأت علمی دانشگاه‌ها با استناد به دستورالعمل‌های رسمی وزارت علوم، تحقیقات و فناوری و فرم‌های رایج ارزشیابی موجود در سامانه گلستان استخراج شدند. این معیارها به عنوان پایه‌ای برای گردآوری و تحلیل داده‌ها مورد استفاده قرار گرفتند.

در ادامه، داده‌های ۵۲۷ درس مربوط به ارزشیابی تدریس اعضای هیأت علمی توسط دانشجویان از دو دانشگاه قم و بزرگمهر قائنات بر اساس معیارهای استخراج شده در گام قبل، به صورت تصادفی از بین دروس ارائه شده در بازه زمانی ۱۴۰۴-۱۳۹۵ در مقطع کارشناسی انتخاب شدند. معیار خروج، دروس با ارزشیابی ناقص بود. همچنین به منظور کنترل سوگیری نظرات دانشجویان، ارزشیابی‌هایی که برای تمامی معیارها نمرات کم‌تر از ۱۳ یا برابر ۲۰ ثبت کرده بودند، از بررسی خارج شدند. هدف از این مرحله، بررسی ارتباط میان شاخص‌های ارزشیابی و بازدهی آموزشی در بستر واقعی آموزش دانشگاهی بود.

برای تحلیل داده‌های گردآوری شده، از تکنیک‌های داده‌کاوی به‌ویژه الگوریتم‌های خوشه‌بندی^۱ استفاده شد تا دروس بر اساس الگوهای ارزشیابی مشابه در خوشه‌هایی همگن طبقه‌بندی شوند. به منظور خوشه‌بندی داده‌ها از الگوریتم K-means استفاده شد. همچنین شاخص‌های سیلوئت^۲، کالینسکی-هاراباز^۳، دیویس بولدین^۴ و آرنج^۵ برای تعیین تعداد بهینه خوشه‌ها و اعتبارسنجی خوشه‌های ایجاد شده مورد استفاده قرار گرفت. شاخص سیلوئت برای خوشه‌های ایجاد شده ۰٫۲۹ بود که نشان از خوشه‌بندی با مطلوبیت متوسط داشت. لازم به ذکر است این مقدار، بیشترین مقدار شاخص سیلوئت در بین تعداد خوشه‌های مختلف بوده است. خوشه‌هایی که میانگین نمرات دانشجویان بالاتری داشتند، به عنوان خوشه‌های با بازدهی آموزشی بالاتر در نظر گرفته شدند و سپس از آن‌ها برای محاسبه وزن و اهمیت نسبی هر یک از معیارهای ارزشیابی استفاده شد. این تحلیل به صورت داده‌محور و بدون دخالت قضاوت ذهنی انجام گرفت تا روایی و اعتبار نتایج افزایش یابد.

1. Clustering
2. Silhouette
3. Calinski-Harabasz
4. Davies-Bouldin
5. Elbow

در گام پایانی، برای معیارهایی که از نظر آماری دارای بیشترین وزن و اهمیت بودند، مجموعه‌ای از راهکارهای اجرایی جهت بهبود کیفیت آموزشی تدوین شد. این راهکارها با استفاده از مطالعات میدانی، تحلیل محتوای نظری و نظرسنجی از خبرگان حوزه آموزش طراحی شدند و سپس با بهره‌گیری از تکنیک‌های تصمیم‌گیری چندمعیاره، شامل تکنیک تحلیل سلسله مراتبی^۱ (AHP)، اولویت‌بندی و رتبه‌بندی گردیدند. به‌منظور اعتبارسنجی نتایج بدست آمده از تکنیک AHP، از نرخ ناسازگاری^۲ استفاده شد. نرخ ناسازگاری برابر ۰.۸۷ بود که نشان می‌دهد سازگاری مقایسه‌ها قابل قبول است. شکل ۱، فرآیند تحقیق و فازهای آن را نشان می‌دهد.

فاز ۱: شناسایی معیارهای ارزشیابی تدریس

استخراج معیارهای رسمی از اسناد وزارت علوم و فرم‌های سامانه گلستان دانشگاه‌ها

فاز ۲: جمع‌آوری داده‌های میدانی

گردآوری داده‌های ارزشیابی اساتید و نمرات دانشجویان
دانشگاه‌های مورد بررسی: قم و بزرگمهر قائنات
بازه زمانی: ۱۳۹۵ تا ۱۴۰۴

فاز ۳: تحلیل داده‌ها با داده‌کاوی

خوشه‌بندی دروس بر اساس الگوهای ارزشیابی مشابه
شناسایی خوشه‌ها با میانگین نمره بالاتر دانشجویان

فاز ۴: محاسبه وزن و اهمیت معیارها

تحلیل داده‌محور بر اساس خوشه‌های دارای عملکرد آموزشی بالاتر
تعیین وزن نسبی معیارها

فاز ۵: ارائه و اولویت‌بندی راهکارها

ارائه راهکار برای معیارهای مهم
استفاده از تکنیک‌های تصمیم‌گیری چندمعیاره شامل تکنیک تحلیل سلسله مراتبی برای اولویت‌بندی راهکارها

شکل ۱: فرآیند تحقیق

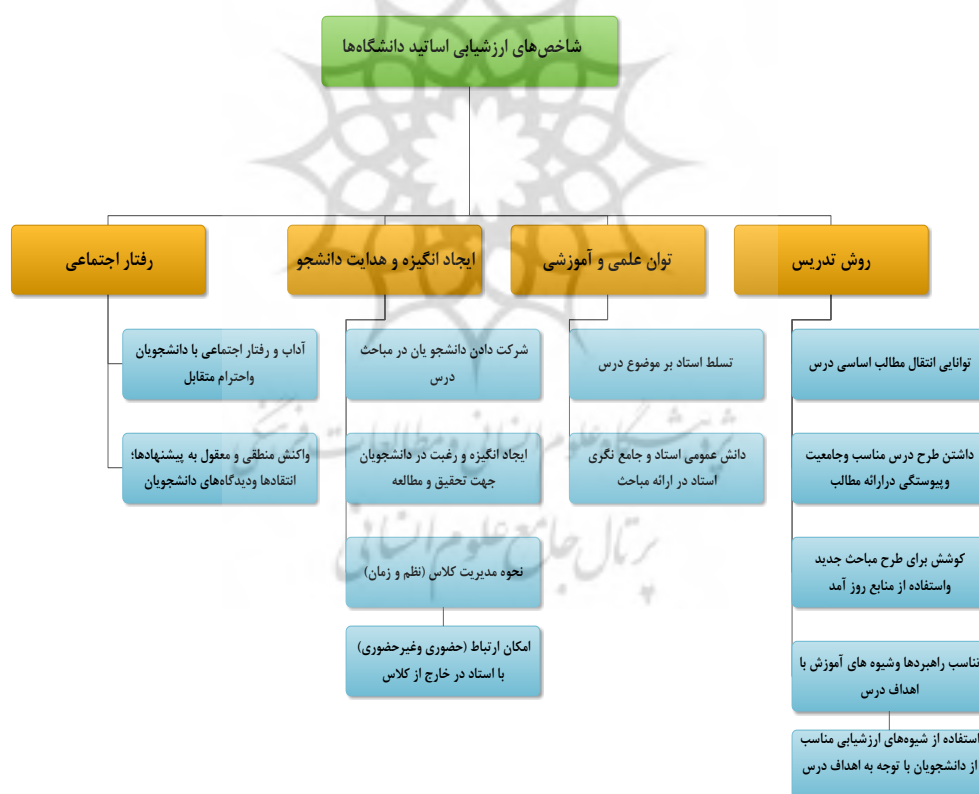
1. Analytic Hierarchy Process (AHP)
2. Inconsistency Ratio

یافته‌های تحقیق

در این بخش به ارائه یافته‌های پژوهش بر اساس فازهای بیان شده در روش پژوهش پرداخته شده است.

۱-۴- فاز اول: شناسایی معیارهای ارزشیابی تدریس

در این مرحله، با استناد به ضوابط و دستورالعمل‌های اعلامی از سوی وزارت علوم، تحقیقات و فناوری و بر اساس فرم‌های استاندارد ارزشیابی تدریس موجود در سامانه گلستان دانشگاه‌ها، مجموعه‌ای از معیارهای رسمی ارزشیابی عملکرد آموزشی اعضای هیأت علمی استخراج شد. این معیارها مبنای تحلیل‌های بعدی پژوهش قرار گرفتند. شکل ۲، معیارهای ارزشیابی را بر اساس ۴ بعد «روش تدریس»، «توان علمی-آموزشی»، «ایجاد انگیزه و هدایت دانشجو» و «رفتار اجتماعی» نشان می‌دهد.



شکل ۲: معیارهای ارزشیابی عملکرد آموزشی اعضای هیأت علمی دانشگاه‌ها

۲-۴- فاز دوم: گردآوری داده‌های میدانی

در این فاز، داده‌های مربوط به ارزشیابی تدریس اعضای هیأت علمی توسط دانشجویان در دو دانشگاه قم و بزرگمهر قائنات گردآوری شد. داده‌ها شامل نمرات ارزشیابی بر اساس معیارهای بدست‌آمده در فاز قبل و تفکیک‌شده بر اساس دروس، به همراه عملکرد تحصیلی دانشجویان (نمرات دروس متناظر) در بازه زمانی سال‌های ۱۳۹۵ تا ۱۴۰۴ بود. این داده‌ها به منظور تحلیل رابطه بین کیفیت تدریس و بازدهی آموزشی مورد استفاده قرار گرفت. شکل ۳، نمایی از داده‌های جمع‌آوری شده را نشان می‌دهد.

نام دانشجو	نام درس	نمرات ارزشیابی	نمرات دروس متناظر	نمرات عملکرد
1	1	1	1	1
2	2	2	2	2
3	3	3	3	3
4	4	4	4	4
5	5	5	5	5
6	6	6	6	6
7	7	7	7	7
8	8	8	8	8
9	9	9	9	9
10	10	10	10	10
11	11	11	11	11
12	12	12	12	12
13	13	13	13	13
14	14	14	14	14
15	15	15	15	15
16	16	16	16	16
17	17	17	17	17
18	18	18	18	18
19	19	19	19	19
20	20	20	20	20

شکل ۳: نمایی از داده‌های جمع‌آوری شده

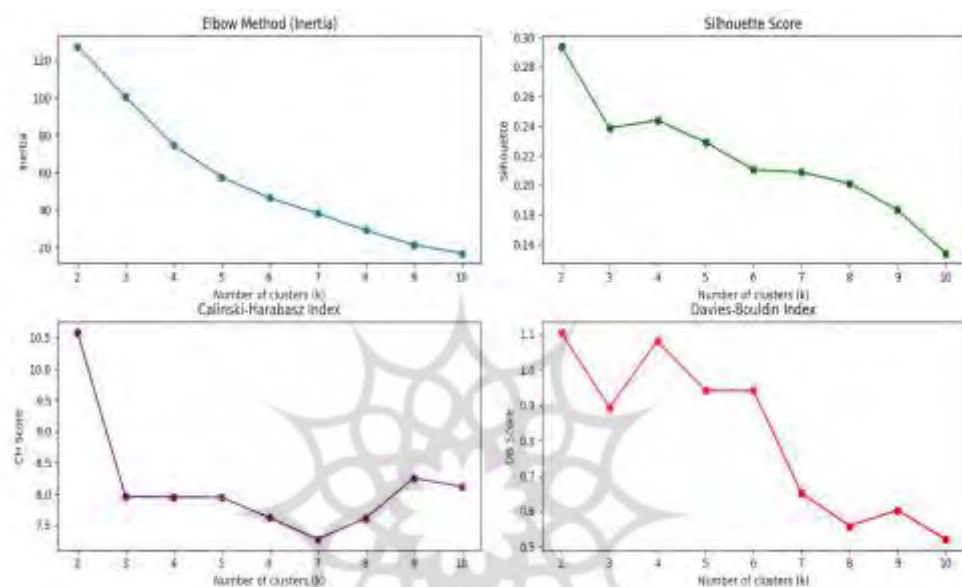
۳-۴- فاز سوم: تحلیل داده‌ها با استفاده از تکنیک‌های داده کاوی

در این مرحله، از تکنیک‌های داده کاوی به‌ویژه الگوریتم‌های خوشه‌بندی برای گروه‌بندی دروس بر اساس الگوهای مشابه در ارزشیابی استفاده شد. هدف از این تحلیل، شناسایی گروه‌هایی از دروس با ویژگی‌های ارزشیابی مشابه و بررسی همبستگی آن‌ها با عملکرد دانشجویان بود. دروس با ارزشیابی مشابه در یک خوشه قرار گرفتند و میانگین نمرات دانشجویان در هر خوشه به‌عنوان شاخص بازدهی آموزشی مورد استفاده قرار گرفت.

به‌منظور خوشه‌بندی دروس با وضعیت مشابه در نتایج نظرسنجی معیارهای مختلف، نیاز است ابتدا تعداد خوشه بهینه مشخص شود. بدین منظور از کد پایتون^۱ و شاخص‌های سیلوئت، کالینسکی-هاراباز، دیویس بولدین و آرنج استفاده شد (برای آشنایی با فرمول‌های ریاضی شاخص‌ها، به پیوست، مراجعه شود). شکل

1. Python

۴ و جدول ۲، نتایج حاصل از این شاخص‌ها را نشان می‌دهد که طبق نتایج بدست آمده، رای اکثریت یعنی ۲ خوشه به عنوان تعداد خوشه بهینه تعیین شد. لازم به ذکر است فرمول‌های ریاضی محاسبه هر یک از شاخص‌های ذکر شده در پیوست ارائه شده است.

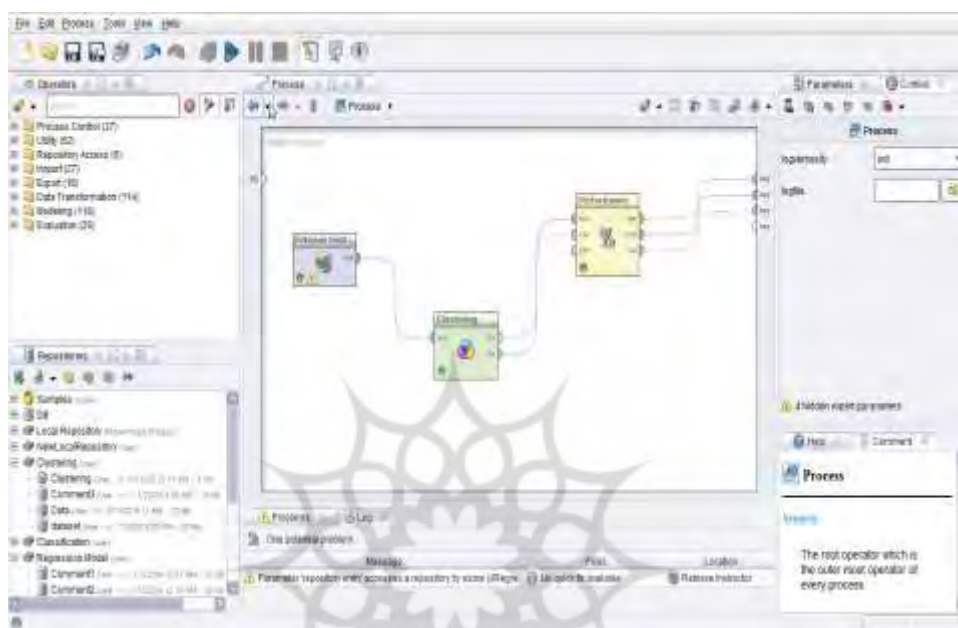


شکل ۴: تعداد خوشه بهینه بر اساس شاخص‌های سیلوئت، کالینسکی-هاراباز، دیویس بولدین و آرئج

جدول ۲: تعداد خوشه بهینه بر اساس شاخص‌های سیلوئت، کالینسکی، دیویس بولدین و البو

نام شاخص	تعداد خوشه بهینه
سیلوئت	۲
کالینسکی	۲
دیویس بولدین	۱۰
البو	۲
رای اکثریت	۲

پس از تعیین تعداد خوشه بهینه از نرم‌افزار رپیدماینر^۱ به منظور خوشه‌بندی دروس با وضعیت ارزشیابی مشابه استفاده شد. شکل ۵، عملگرهای^۲ استفاده شده به منظور خوشه‌بندی در نرم‌افزار رپیدماینر را نشان می‌دهد.



شکل ۵: عملگرهای استفاده شده در نرم‌افزار رپیدماینر به منظور خوشه‌بندی دروس با وضعیت ارزشیابی مشابه

۴-۴- فاز چهارم: محاسبه وزن و اهمیت معیارها

در این فاز، با استفاده از داده‌های به دست آمده از خوشه‌ای که بالاترین میانگین نمرات دانشجویان را دارا بود، وزن و اهمیت نسبی هر یک از معیارهای ارزشیابی تدریس تعیین شد. این تحلیل به صورت داده‌محور و بدون مداخله ذهنی انجام شد تا اعتبار و عینیت یافته‌ها حفظ شود. جدول ۳، وزن و اهمیت معیارهای ارزشیابی اعضای هیأت علمی دانشگاه‌ها را نشان می‌دهد.

1. Rapid Miner
2. Operators

جدول ۳: وزن و اهمیت معیارهای ارزشیابی اعضای هیأت علمی دانشگاه‌ها

اولویت	نام معیار ارزشیابی	وزن
۱	واکنش منطقی و معقول به پیشنهادها؛ انتقادات و دیدگاه‌های دانشجویان	۰,۰۷۸
۲	تسلط استاد بر موضوع درس	۰,۰۷۷۷
۳	داشتن طرح درس مناسب و جامعیت و پیوستگی در ارائه مطالب	۰,۰۷۷۵
۴	نحوه مدیریت کلاس (نظم و زمان)	۰,۰۷۷۳
۵	استفاده از شیوه‌های ارزشیابی مناسب از دانشجویان با توجه به اهداف درس	۰,۰۷۷۱
۶	امکان ارتباط (حضور و غیرحضور) با استاد در خارج از کلاس	۰,۰۷۷۱
۷	آداب و رفتار اجتماعی با دانشجویان و احترام متقابل	۰,۰۷۷
۸	دانش عمومی استاد و جامع‌نگری استاد در ارائه مباحث	۰,۰۷۶۹
۹	توانایی انتقال مطالب اساسی درس	۰,۰۷۶۶
۱۰	تناسب راهبردها و شیوه‌های آموزش با اهداف درس	۰,۰۷۶۵
۱۱	کوشش برای طرح مباحث جدید و استفاده از منابع روز آمد	۰,۰۷۶۴
۱۲	شرکت دادن دانشجویان در مباحث درس	۰,۰۷۶۱
۱۳	ایجاد انگیزه و رغبت در دانشجویان جهت تحقیق و مطالعه	۰,۰۷۵۷

همان‌طور که از جدول ۳ مشخص است، معیار «واکنش منطقی و معقول به پیشنهادها؛ انتقادات و دیدگاه‌های دانشجویان» با وزن ۰,۰۷۸ بیشترین تاثیر را بر بازدهی دانشجویان داشته است. معیارهای «تسلط استاد بر موضوع درس» با وزن ۰,۰۷۷۷ و «داشتن طرح درس مناسب و جامعیت و پیوستگی در ارائه مطالب» با وزن ۰,۰۷۷۵ به ترتیب رتبه‌های دوم و سوم را تشکیل می‌دهند. همچنین، نتایج محاسبه شاخص‌های آماری برای وزن معیارها به شرح زیر است:

≠ میانگین وزن‌ها: ۰,۰۷۶۹

≠ انحراف معیار (SD): ۰,۰۰۰۶۵

≠ فاصله اطمینان ۹۵٪ (CI): $\pm ۰,۰۰۰۳۵$

این مقادیر نشان می‌دهند که پراکندگی وزن‌ها بسیار کم است و نتایج از پایایی مناسبی برخوردارند.

۵-۴- فاز پنجم: ارائه و اولویت‌بندی راهکارها

برای معیار ارزشیابی «داشتن طرح درس مناسب و جامعیت و پیوستگی در ارائه مطالب» که جزو معیارهایی بود که وزن و اهمیت بالایی را در بهبود کیفیت آموزشی داشت، مجموعه‌ای از راهکارهای ارتقای پیشنهاد شد. این راهکارها با استفاده از مطالعات کتابخانه‌ای و میدانی، بررسی تجارب مشابه، و نظرسنجی از خبرگان تدوین گردید. سپس با بهره‌گیری از تکنیک‌های تصمیم‌گیری چندمعیاره شامل AHP، میزان اثربخشی و اولویت راهکارها نسبت به یکدیگر ارزیابی و رتبه‌بندی شد.

پس از مطالعات کتابخانه‌ای و پرسش از افراد خبره، معیارهای تاثیرگذار بر این شاخص به صورت زیر شناسایی شدند (فارل و اشکرفت^۱، ۲۰۲۴؛ فرهنگ و همکاران^۲، ۲۰۲۳؛ احدی ساروخلیل و همکاران، ۱۴۰۲؛ رضانی چاهک و همکاران، ۱۴۰۲):

۱. تحلیل و تنظیم اهداف آموزشی: هدف‌های آموزشی، بیان‌کننده وضعیت مطلوب یادگیرنده

در یک رویداد آموزشی هستند. هدف‌های مفید و قابل وصول باید فراگیرمحور (مخاطب یادگیرنده)، صریح و قابل فهم، قابل مشاهده و اندازه‌گیری و توصیفی از نتایج یادگیری باشند.

۲. تحلیل و تعیین محتوای آموزشی: محتوای درس بر اساس توانایی‌های دانشجویان و مدت

زمان موجود، تدوین می‌گردد.

۳. تحلیل و تنظیم روش آموزشی: روش تدریس با توجه به اهداف رفتاری و موضوع درس،

مفاهیم، اصول و قوانین هر علم، ارتباط با مسائل روز، توانایی فراگیران، علاقه و نگرش آن‌ها، انطباق با زمان آموزش و غیره انتخاب می‌شود.

1. Farrell & Ashcraft

2. Farhang et al

۴. **تحلیل و تنظیم وسایل آموزشی:** وسایل آموزشی به کلیه تجهیزات و امکاناتی اطلاق می‌شود که می‌توانند در محیط آموزشی شرایطی را به وجود آورند که در آن شرایط، یادگیری سریع‌تر، آسان‌تر، بهتر، بادوام‌تر و موثرتر صورت بگیرد.

پس از استخراج معیارهای تاثیرگذار بر «داشتن طرح درس مناسب و جامعیت و پیوستگی در ارائه مطالب»، به محاسبه وزن و اهمیت هر معیار با استفاده از تکنیک‌های تصمیم‌گیری چندمعیاره و مدل تحلیل سلسله‌مراتبی (AHP) با پرسش از افراد خبره پرداخته شد. شکل ۶ نمایی از مقایسه زوجی بین معیارها و شکل ۷ نمایی از ماتریس تصمیم^۱ را نشان می‌دهد.

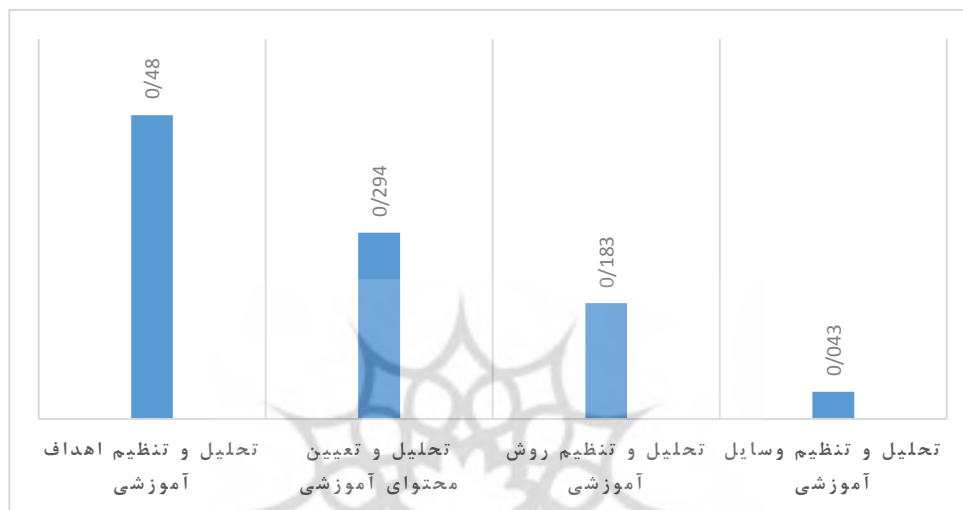
شکل ۶. نمایی از مقایسه زوجی بین معیارهای تاثیرگذار در تسلط استاد بر موضوع درس

	1	2	3	4
1	1	2.00	3.00	8.00
2	0.50	1	2.00	7.00
3	0.33	0.50	1	6.00
4	0.12	0.14	0.17	1

شکل ۷: نمایی از ماتریس تصمیم ایجادشده حاصل از مقایسه زوجی بین معیارهای تاثیرگذار در تعیین طرح درس

1. Decision Matrix

پس از مقایسه زوجی معیارها، وزن و اهمیت هر معیار تاثیرگذار در تعیین طرح درس محاسبه شد. شکل ۸، نمودار وزن معیارها را نشان می‌دهد. لازم به ذکر است نرخ ناسازگاری کم‌تر از ۰,۱ بود که پایایی پرسشنامه‌ها را تایید می‌کند.



شکل ۸. نمودار وزن و اهمیت معیارهای تاثیرگذار در تعیین طرح درس

همان‌طور که در نمودار شکل ۸ قابل مشاهده است، معیار «تحلیل و تنظیم اهداف آموزشی» با وزن ۰,۴۸ بیشترین اهمیت و تاثیر را در طراحی طرح درس دارد. این موضوع بدین دلیل است که در صورت تعیین اهداف اشتباه تمامی فعالیت‌های دیگر از جمله تعیین محتوا، روش و وسایل آموزشی نیز به اشتباه انجام خواهند شد.

بحث و نتیجه‌گیری

تضمین کیفیت آموزشی در آموزش عالی، در گرو ارزیابی دقیق، عینی و چندبعدی از عملکرد اعضای هیأت علمی است؛ امری که در سال‌های اخیر با چالش‌هایی نظیر اتکای صرف به نظرات دانشجویان، استفاده محدود از روش‌های تحلیلی، و فقدان پیوند بین نتایج ارزشیابی و بازدهی واقعی آموزشی مواجه بوده است. پژوهش حاضر با هدف پر کردن این شکاف، رویکردی داده‌محور و ترکیبی را در ارزیابی روش‌ها و فنون تدریس اعضای هیأت علمی ارائه داده است. با تلفیق داده‌های حاصل از ارزشیابی تدریس و عملکرد تحصیلی دانشجویان و با بهره‌گیری از الگوریتم‌های داده‌کاوی و تحلیل‌های تصمیم‌محور، چارچوبی نوآورانه برای تعیین وزن و اهمیت معیارهای ارزشیابی ارائه شد.

برخلاف بسیاری از مطالعات پیشین که صرفاً به تحلیل همبستگی میان یک یا چند معیار با بازدهی آموزشی پرداخته‌اند یا وزن‌دهی به معیارها را مبتنی بر قضاوت‌های خبرگان و تکنیک‌های تصمیم‌گیری چندمعیاره انجام داده‌اند، پژوهش حاضر از داده‌های واقعی ارزشیابی و نمرات دانشجویان به عنوان مبنای تجربی بهره گرفت. در این رویکرد، با استفاده از خوشه‌بندی دروس بر اساس شباهت الگوهای ارزشیابی، و تمرکز بر خوشه‌هایی با بالاترین عملکرد تحصیلی دانشجویان، معیارهایی که بیشترین نقش در تحقق کیفیت آموزشی ایفا می‌کنند شناسایی و به صورت نرمال‌شده وزن‌دهی شدند. این امر نه تنها به کاهش خطای انسانی و افزایش اعتبار تحلیل‌ها کمک کرد، بلکه امکان استخراج بینش‌های معنادار از داده‌های خام دانشگاهی را نیز فراهم آورد.

افزون بر این، پژوهش حاضر گامی فراتر نهاده و برای معیارهایی که از اهمیت بیشتری برخوردار بودند، مجموعه‌ای از راهکارهای اجرایی جهت بهبود کیفیت تدریس ارائه داد. این راهکارها با استفاده از روش‌های تصمیم‌گیری چندمعیاره شامل AHP اولویت‌بندی شدند تا در برنامه‌ریزی‌های آموزشی و سیاست‌گذاری‌های دانشگاهی قابل استفاده و اتکا باشند.

در مجموع، نتایج این پژوهش نشان می‌دهد که استفاده ترکیبی از داده‌کاوی و تحلیل تصمیم‌محور می‌تواند چارچوبی دقیق‌تر، معتبرتر و کاربردی‌تر برای ارزیابی عملکرد آموزشی اعضای هیأت علمی فراهم آورد؛ چارچوبی که می‌تواند مبنایی برای طراحی نظام‌های ارزشیابی هوشمند، توسعه حرفه‌ای اساتید، و ارتقای کیفی آموزش در دانشگاه‌ها باشد. بر این اساس، پیشنهاد می‌شود دانشگاه‌ها و نهادهای مسئول در سیاست‌گذاری آموزشی، به بهره‌گیری از این نوع رویکردهای داده‌محور توجه بیشتری نشان دهند تا فرآیند ارزیابی از سطحی تشریفاتی و ایستا به ابزاری اثربخش برای بهبود مستمر آموزش ارتقا یابد.

شایان توجه است که در برخی موارد، خوشه‌هایی مشاهده شد که علی‌رغم کسب امتیازات متوسط در معیارهای ارزشیابی تدریس، دارای عملکرد تحصیلی بالایی بودند. این پدیده می‌تواند ناشی از تأثیر متغیرهای مخدوش‌کننده‌ای باشد که در مدل لحاظ نشده‌اند؛ از جمله انگیزه ذاتی دانشجویان، پیش‌زمینه علمی قوی، یا عوامل بیرونی مانند منابع آموزشی مکمل و فضای رقابتی کلاس. البته برای کاهش سوگیری ناشی از داده‌های افراطی، در طراحی پژوهش، پاسخ‌هایی که در آن‌ها دانشجویان تمامی معیارهای ارزشیابی را به صورت غیرمتعارف پایین (کمتر از ۱۳) یا غیرمنطقی بالا (تماماً ۲۰) امتیاز داده بودند، از تحلیل حذف شد. با این حال، همچنان وجود الگوهای غیرهمخوان میان ارزشیابی و عملکرد تحصیلی می‌تواند مؤید لزوم استفاده از مدل‌های جامع‌تر در پژوهش‌های آینده باشد—مدل‌هایی که متغیرهای پنهان^۱ و مخدوش‌کننده

1. Latent Variables

را با روش‌هایی نظیر رگرسیون چندسطحی یا مدل‌سازی معادلات ساختاری لحاظ کنند تا تمایز دقیق‌تری میان اثر واقعی کیفیت تدریس و سایر عوامل زمینه‌ای فراهم گردد.

یکی دیگر از محدودیت‌های مهم پژوهش حاضر، مربوط به دامنه جغرافیایی و حجم نمونه است. با آن که داده‌های این مطالعه از دو دانشگاه مختلف گردآوری شده و تنوعی نسبی در رشته‌ها و گروه‌های آموزشی لحاظ شده است، اما این تعداد برای تعمیم مستقیم نتایج به کل نظام آموزش عالی کشور کافی نیست. تفاوت‌های فرهنگی، ساختاری و مدیریتی در دانشگاه‌های کشور می‌تواند بر الگوهای ارزشیابی، عملکرد تحصیلی دانشجویان و حتی نحوه تعامل استاد و دانشجو اثرگذار باشد. بنابراین، یافته‌های این پژوهش باید با احتیاط در سطح ملی تفسیر شوند و در عین حال، می‌توان آن‌ها را به عنوان الگوی اولیه‌ای برای توسعه نظام‌های ارزشیابی داده‌محور در سایر دانشگاه‌ها مدنظر قرار داد.

شایان ذکر است که در بسیاری از دانشگاه‌های طراز اول جهان نظیر MIT، Stanford و Oxford، ارزشیابی عملکرد آموزشی اساتید به صورت ترکیبی و چندمنظوره انجام می‌شود؛ به طوری که علاوه بر نظرات دانشجویان، داده‌های عملکرد تحصیلی، بازخورد هم‌تایان^۱، و نتایج یادگیری نیز در ارزیابی دخیل هستند. برای مثال، در دانشگاه MIT، داده‌کاوی روی نتایج امتحانات نهایی و پیشرفت تحصیلی در طول ترم، مکمل ارزیابی‌های کیفی دانشجویان محسوب می‌شود و نتایج این ارزیابی‌ها مستقیماً در فرآیند ارتقاء اعضای هیأت علمی و طراحی توسعه حرفه‌ای آموزشی لحاظ می‌گردد (موسسه فناوری ماساچوست^۲، ۲۰۲۵). به همین ترتیب، دانشگاه آکسفورد نیز در چارچوب "Oxford Learning Institute" از مدل‌های تحلیل چندلایه‌ای بهره می‌گیرد تا تصویر جامع‌تری از کیفیت تدریس ارائه دهد (دانشگاه آکسفورد^۳، ۲۰۲۵). در این راستا، پژوهش حاضر نیز در مسیر چنین الگوهایی، با استفاده از روش‌های داده‌محور و تحلیل‌های خوشه‌ای، تلاش کرده است مدلی کارآمد، قابل اتکا و تکرارپذیر برای ارزیابی تدریس در دانشگاه‌های ایران ارائه دهد. همگرایی رویکرد پژوهش حاضر با نظام‌های ارزشیابی دانشگاه‌های پیشرو جهانی، مؤید آن است که استفاده از داده‌های عینی و روش‌های تحلیلی می‌تواند مبنای بازنگری جدی در نظام ارزشیابی تدریس در آموزش عالی کشور قرار گیرد.

با این حال، یافته‌های این پژوهش در مقایسه با برخی مطالعات پیشین، مانند پژوهش مارش^۴ (۲۰۰۷) که تسلط علمی استاد بر محتوای درس را کلیدی‌ترین عامل در اثربخشی تدریس معرفی می‌کرد، تفاوت

1. Peer-Review

2. Massachusetts Institute of Technology

3. University of Oxford

4. Marsh

معناداری نشان می‌دهد. در پژوهش حاضر، تعاملات انسانی و واکنش استاد به نظرات، پیشنهادات و انتقادات دانشجویان به عنوان مؤثرترین مؤلفه در کیفیت تدریس شناسایی شده است. این تفاوت می‌تواند ناشی از ابعاد فرهنگی، انتظارات بومی و الگوهای تعامل در نظام آموزش عالی ایران باشد، جایی که تعامل متقابل استاد-دانشجو، درک‌پذیری رفتاری، و پاسخ‌گویی به نیازهای دانشجویان از جایگاه برجسته‌تری برخوردار است. بنابراین، یافته‌های این پژوهش نه تنها نقش عوامل شناختی، بلکه اهمیت عوامل عاطفی، ارتباطی و بین‌فردی را در اثربخشی تدریس در محیط‌های فرهنگی خاص برجسته می‌سازند. این تفاوت تأکید می‌کند که مدل‌های جهانی ارزیابی تدریس، نیازمند بومی‌سازی و تطبیق با بافت فرهنگی و اجتماعی آموزش عالی هر کشور هستند.

در راستای توسعه و تعمیق یافته‌های این پژوهش، مسیرهای متعددی برای تحقیقات آینده قابل پیشنهاد است. نخست، بررسی تأثیر رشته تحصیلی یا گروه آموزشی بر وزن و اهمیت معیارهای ارزشیابی می‌تواند به طراحی مدل‌های تخصصی‌تر و بومی‌شده برای هر حوزه علمی کمک کند. دوم، پیشنهاد می‌شود از الگوریتم‌های یادگیری عمیق^۱ به‌ویژه در پردازش زبان طبیعی برای تحلیل نظرات متنی و کیفی دانشجویان بهره‌گیری شود تا لایه‌های پنهان‌تری از ادراک و نگرش دانشجویان نسبت به عملکرد اساتید استخراج گردد. سوم، مطالعات تطبیقی با دانشگاه‌های برتر بین‌المللی (نظیر MIT، Stanford یا Oxford) می‌تواند به شناسایی نقاط قوت و ضعف رویکردهای بومی در ایران کمک کند و بستر اصلاح و به‌روزرسانی سیاست‌های ارزشیابی آموزشی را فراهم آورد. این مسیرها، زمینه‌ساز ارتقای بیشتر نظام‌های ارزشیابی مبتنی بر داده در آموزش عالی خواهند بود.

پژوهشگاه علوم انسانی و مطالعات فرهنگی
پرتال جامع علوم انسانی

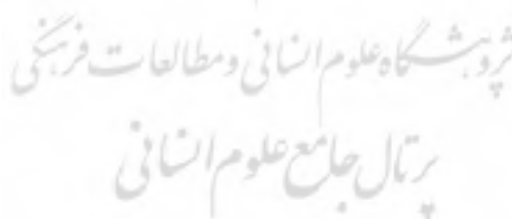
منابع

- ایزدی ثابت، فریده، سلمانی، معصومه، صفایی، زهرا، و غریبی، فرید. (۱۴۰۱). طراحی، اعتباریابی و پایاسازی ابزار ارزشیابی عملکرد آموزشی اعضای هیأت علمی دانشگاه‌های علوم پزشکی. کومش، ۲۴(۵) (پیاپی ۹۱)، ۶۶۱-۶۷۳.
- بارخدا، سیدجمال؛ ویسی، شعیب و شریفی، ژینا. (۱۴۰۲). بررسی چالش‌های تدریس و ارزشیابی در بستر فضای مجازی. مدیریت دانشگاهی، ۲(۴)، ۱۴۲-۱۵۶.
- حقیقی، مهدی، اکبری، احمد، چراغین، مسلم، و کریمی، محمد. (۱۴۰۰). بررسی تطبیقی نظام ارزشیابی آموزشی اعضای هیأت علمی در دانشگاه‌های علوم پزشکی ایران. مجله ایرانی آموزش در علوم پزشکی، ۲۱(۸۴)، ۲۳۷-۲۴۶.
- رمضانی چاهک، ف.، گرمابی، ب.، محمدی، ف.، و باجگیر، م. (۱۴۰۲). فواید طرح درس در فرایند تدریس، سیزدهمین کنفرانس بین‌المللی پژوهش‌های مدیریت و علوم انسانی، تهران.
- شهامت، نادر؛ زارعی، رضا؛ صالحی، مسلم و حسینی حسام آبادی، سیدمحمدادی. (۱۴۰۲). شناسایی ابعاد و مؤلفه‌های مدل تدریس اثربخش در مقطع ابتدایی. دوماهنامه علمی - پژوهشی رهیافتی نو در مدیریت آموزشی، ۱۴(۲)، ۱۴۱-۱۵۴.
- وزارت علوم، تحقیقات و فناوری. (۱۴۰۲). سند تحول راهبردی آموزش عالی جمهوری اسلامی ایران. تهران: معاونت آموزشی وزارت علوم، تحقیقات و فناوری.

- Arslan, B., Lehman, B., Tenison, C., Sparks, J. R., López, A. A., Gu, L., & Zapata-Rivera, D. (2024). Opportunities and challenges of using generative AI to personalize educational assessment. *Frontiers in Artificial Intelligence*, 7, 1460651.
- Bablu, T. A. (2024). Machine learning in automated assessment: Enhancing objectivity and efficiency in educational evaluations. *Journal of Advanced Computing Systems*, 4(7).
- Barkhoda, J., Vaisi, S. and Sharifi, Z. (2023). Investigating the challenges of teaching and evaluation in the context of cyberspace. *Journal of Academic Management*, 2(4), 165-142. [in Persian]
- Bazohoori, S., Yari haj atalo, J. and Maleki avarsin, S. (2021). Explaining teachers' self-management and self-efficacy based on transformational leadership with the mediating role of learning organization in Mashhad schools. *School Administration*, 9(1), 156-128.

- Farhang, A. P. Q., Hashemi, A. P. S. S. A., & Ghorianfar, A. P. S. M. (2023). Lesson plan and its importance in teaching process. *International Journal of Current Science Research and Review*, 6(08), 5901-5913.
- Farrell, T. S., & Ashcraft, N. (2024). *Lesson planning*. TESOL Press.
- Gump, S. E. (2007). Student Evaluations of Teaching Effectiveness and the Leniency Hypothesis: A Literature Review. *Educational Research Quarterly*, 30(3), 56-69.
- Haghighi, M., Akbari, A., Cherabin, M., & Karimi, M. (2021). A comparative study of the educational evaluation system of faculty members in Iranian universities of medical sciences. *Iranian Journal of Medical Education*; 21 :237-246. [in Persian]
- Ikramov, S. K., & Tursunbaeva, K. I. (2022). TEACHING METHODS AND TECHNOLOGY AFFECTING THE QUALITY OF EDUCATION. *CURRENT RESEARCH JOURNAL OF PEDAGOGICS*, 3(02), 5-8.
- Izadisabet, F., Salmani, M., Safaei, Z., & Gharibi, F. (2022). Development and psychometric features of a questionnaire to evaluate the educational performance of faculty members in medical universities. *KOOMESH*, 24(5 (91)), 661-673. [in Persian]
- Jiawei, H., & Micheline, K. (2006). *Data mining: concepts and techniques*. Morgan kaufmann.
- Khezri, M., Abdollahi, F. and Hasani, R. (2025). Validating an Evaluation Model for Quality Assurance of High Schools Education. *School Administration*, 12(4), 101-131.
- Marsh, H. W. (2007). Students' evaluations of university teaching: dimensionality, reliability, validity, potential biases and usefulness. In the *Scholarship of Teaching and Learning in Higher Education: An Evidence-based Perspective*, edited by R. P. Perry and J. C. Smart, pp.319-383. Dordrecht: Springer.
- Massachusetts Institute of Technology. (n.d.). *Subject evaluation: Office of the Registrar*. Retrieved July 24, 2025, from <https://registrar.mit.edu/classes-grades-evaluations/subject-evaluation>
- Mensaha, P. O., Kumahb, A., Korantengc, F. O., & Antwia, F. A. EVALUATING QUALITY TEACHING AND LEARNING, PERCEIVED TEACHERS' ABSENTEEISM IN SENIOR HIGH SCHOOLS. *Education, Sustainability & Society (ESS)*, 6(1), 1-6.
- Nguyen, T. D. (2024). Examining the Impact of University Teaching Quality on Learning Engagement: A Cross-sectional Study in Vietnam. *Vietnam Journal of Education*, 102-111.
- Nseer, A. H., Al-Hiyari, H. A., & Abu Ashour, K. M. (2021). Teaching methods used by faculty members at Yarmouk University with students. *Turkish Online Journal of Qualitative Inquiry*, 12(6).

- Ramezani Chahak, F., Garmabi, B., Mohammadi, F., & Bajgir, M. (2023). The benefits of the lesson plan in the teaching process, *International Conference on Management and Humanities Research*, Tehran. [in Persian]
- Saeed, A. (2024). A Case Study: Evaluating the Quality of Teaching at a Private University in Karachi. *Journal of Education and Humanities Research (JEHR)*, University of Balochistan, Quetta, 17(1), 116-124.
- Shahamat, N., zarei, R., salehi, M., & hoseini hesam abadi, S. M. H. (2023). Identify the Dimensions and Components of An Effective Teaching Model in Elementary School. *Journal of New Approaches in Educational Administration*, 14(2), 154-141. [in Persian]
- Spooren, P., Brockx, B., & Mortelmans, D. (2013). On the validity of student evaluation of teaching: The state of the art. *Review of Educational Research*, 83(4), 598-642.
- University of Oxford. (n.d.). *Evaluating teaching: Centre for Teaching and Learning*. Retrieved July 24, 2025, from <https://wwwctl.ox.ac.uk/evaluating-teaching>
- Yu, Y. (2022). Evaluating teaching quality in colleges and universities of public art education using the AHP fuzzy comprehensive method. *Scientific Programming*, 2022(1), 3529311.
- Yu, C., Zhao, B., & Ding, K. (2022, April). Research on Course Teaching Quality Evaluation Based on Improved BP Neural Network Algorithm. In *Proceedings of the 7th International Conference on Information and Education Innovations* (pp. 54-59).



Assessment and Evaluation of University Faculty Teaching Methods and Techniques for Ensuring Educational Quality

Mohammad Moradi^{1*}

Abstract

The purpose of this study is to present a data-driven framework for assessing and evaluating the teaching methods and techniques of university faculty members, with a focus on the role of such evaluations in ensuring educational quality. Unlike most previous studies that primarily relied on descriptive analyses or expert opinions, this research utilizes real data obtained from student evaluations of teaching and corresponding academic performance scores. The statistical population includes course data from Qom University and Bozorgmehr University of Qaenat over the period of 2016 to 2025. First, the evaluation criteria were extracted based on the official guidelines of the Ministry of Science, Research, and Technology. Then, using data mining algorithms—particularly clustering—courses were grouped based on similar evaluation patterns, and the criteria associated with clusters exhibiting the highest academic performance were identified. The relative weight and importance of each criterion were subsequently calculated and normalized in a data-oriented manner. The analysis revealed that the criterion “Reasonable and logical response to student suggestions, criticisms, and viewpoints” had the highest impact on student performance, with a weight of 0.07803. This was followed by “Mastery of the subject” (weight: 0.07774) and “Having a proper lesson plan with coherence and comprehensive content delivery” (weight: 0.07746). In the final phase, targeted strategies were proposed for the high-priority criteria and ranked using multi-criteria decision-making (MCDM) methods, including the Analytic Hierarchy Process (AHP). The findings indicate that the integrated approach of data mining and decision analysis offers a precise and effective tool for enhancing the academic evaluation system and improving teaching quality in higher education.

Keywords: Teaching Evaluation; Educational Quality; Educational Policy; Data Mining; Multi-Criteria Decision-Making (MCDM)

¹- Assistant Professor, Department of Computer Engineering, Faculty of Engineering, Bozorgmehr University of Qaenat, Qaenat, Iran.

پیوست

به منظور افزایش شفافیت و امکان تکرارپذیری نتایج پژوهش، در این پیوست فرمول‌های ریاضی شاخص‌های ارزیابی استفاده شده شامل سیلوئت، کالینسکی-هاراباز، دیویس-بولدین و روش آرنج ارائه شده است (جیاوی و میچلین، ۲۰۰۶).

۱- شاخص سیلوئت

برای هر نمونه i :

$a(i) \neq$: میانگین فاصله بین نمونه i و سایر نقاط در همان خوشه

$b(i) \neq$: کمترین میانگین فاصله بین نمونه i و تمام نقاط در سایر خوشه‌ها

فرمول:

$$S(i) = \frac{b(i) - a(i)}{\max\{a(i), b(i)\}}$$

امتیاز کلی سیلوئت، میانگین $S(i)$ برای تمام نمونه‌ها است.

۲- شاخص کالینسکی-هاراباز

فرمول:

$$CH = \frac{n-i}{k-1} \cdot \frac{Tr(B_k)}{Tr(W_k)}$$

که در آن:

$n \neq$: تعداد کل نمونه‌ها

$k \neq$: تعداد خوشه‌ها

$Tr(B_k) \neq$: مجموع مربعات بین خوشه‌ای

$Tr(W_k) \neq$: مجموع مربعات درون خوشه‌ای

مقدار بیشتر این شاخص، خوشه‌بندی بهتری را نشان می‌دهد.

۳- شاخص دیویس-بولدین

فرمول:

$$DB = \frac{1}{k} \sum_{i=1}^k \max_{j \neq i} \left(\frac{\sigma_i + \sigma_j}{d(c_i, c_j)} \right)$$

که در آن:

k : تعداد خوشه‌ها $\neq$

σ_i : میانگین فاصله بین نقاط خوشه i و مرکز آن $\neq$

$d(c_i, c_j)$: فاصله بین مراکز خوشه‌های i و j $\neq$

مقدار کمتر DB نشان‌دهنده خوشه‌بندی بهتر است.

۴- روش آرنج

در این روش، از تابع اینرسی^۱ که به صورت مجموع مربع فاصله‌ها بین نمونه‌ها و مرکز خوشه آن‌ها تعریف می‌شود، استفاده می‌گردد:

$$\text{Inertia} = \sum_{i=1}^k \sum_{x \in C_i} \|\mu_i - x\|^2$$

که در آن:

k : تعداد خوشه‌ها $\neq$

C_i : مجموعه نمونه‌های موجود در خوشه i -ام $\neq$

$x \in C_i$: هر نمونه (نقطه داده) x که در خوشه i قرار دارد. $\neq$

μ_i : مرکز خوشه i -ام، که به طور معمول میانگین همه نمونه‌های داخل خوشه i است. $\neq$

$\|\mu_i - x\|^2$: مربع فاصله اقلیدسی بین نمونه x و مرکز خوشه μ_i . این مقدار نشان می‌دهد که هر نقطه داده چقدر

از مرکز خوشه‌اش دور است.

نقطه‌ای که کاهش مقدار Inertia از حالت خطی خارج می‌شود (زاویه‌دار می‌شود)، به عنوان بهترین k در نظر گرفته می‌شود.