

روشی نوین در ارزشیابی و رتبه‌بندی معلمان با استفاده از درخت تصمیم‌گیری

■ احد جعفرنژاد* ■ حسن‌رضا زین‌آبادی** ■ سارا عابدی کوشکی***

چکیده:

ارزشیابی معلمان یکی از سیاست‌های آموزشی مهم محسوب می‌شود و راهبردی قوی برای بهبود آموزش در جهت رسیدن به پیشرفت تحصیلی بهتر است. از جمله مدل‌های ارزشیابی برای رسیدن به این هدف، مدل «چارچوبی برای آموزش» دنیلسون است. این مدل دربرگیرنده چهار حیطه برنامه‌ریزی، محیط کلاسی، نحوه آموزش و مسئولیت‌پذیری است که هر یک به‌نحوی توصیف‌کننده ویژگی‌ها و رفتارهای معلمان می‌باشند. در روش کنونی این مدل ارزشیابی، مقادیر معیارها با اعداد ثابت سنجیده می‌شود. همچنین پیچیدگی فرآیند ارزشیابی، احتمال بروز خطای انسانی را از جانب ارزیاب افزایش می‌دهد. به‌عنوان راه‌حلی مناسب، در این مقاله، یک سیستم نرم‌افزاری ارزشیابی با استفاده از منطق «فازی» و درخت تصمیم‌گیری به‌عنوان پارادایم‌های هوش مصنوعی پیشنهاد شده است. در مدل پیشنهادی، امتیاز هر معیار از مدل دنیلسون (که پیش از این با اعداد ثابت سنجیده می‌شد) به‌صورت فازی سنجیده می‌شود و دسته‌بندی معلمان با استفاده از درخت تصمیم انجام خواهد شد. این مدل نرم‌افزاری، به دلیل استفاده از متغیرهای کیفی فازی، فرآیند ارزشیابی را برای ارزیاب بسیار آسان‌تر از روش قبل و با دقت بسیار بالایی انجام می‌دهد. در این پژوهش فرآیند ارزشیابی و رتبه‌بندی معلمان چند مدرسه از ناحیه ۴ کرج، در بازه زمانی ابتدای مهر تا انتهای اسفند ۹۳ توسط این سیستم ارزشیابی مورد بررسی قرار گرفته است. در مجموع، این مدل می‌تواند به‌عنوان یک سیستم پشتیبان تصمیم‌گیری و نرم‌افزاری دقیق، توسط مدیران در ارزشیابی و رتبه‌بندی معلمان مورد استفاده قرار گیرد.

کلید واژه‌ها: رتبه‌بندی معلمان، هوش مصنوعی، منطق فازی، درخت تصمیم‌گیری

□ تاریخ دریافت مقاله: ۹۴/۷/۹ □ تاریخ شروع بررسی: ۹۴/۱۰/۷ □ تاریخ پذیرش مقاله: ۹۵/۴/۸

* کارشناسی ارشد کامپیوتر، آموزش و پرورش استان خراسان شمالی..... Ahad.jafarnezhad@gmail.com
** دانشیار گروه مدیریت آموزشی، دانشکده مدیریت، دانشگاه خوارزمی تهران..... Zeinabadi_hr@khu.ac.ir
*** دانشجوی دکتری مدیریت آموزشی، دانشگاه سمنان، (نویسنده مسئول)..... Sara.abedi.kooshki@gmail.com

مقدمه

ارزشیابی معلمان یک سیاست مهم آموزشی و راهبردی قوی برای بهبود بخشیدن به آموزش دانش‌آموزان در جهت پیشرفت تحصیلی بهتر است. نظر به اهمیت نقش ارزشیابی در توسعه حرفه‌ای معلمان و استادان، ارزشیابان آموزشی و بخش‌های مسئول در این زمینه، در بسیاری از کشورها، تلاش زیادی برای افزایش پیشرفت تحصیلی، کیفیت و برابری انجام داده‌اند (گرینی و کلاگان، ۲۰۰۸). کلیر و اشباخر^۲ (۲۰۰۱) معتقدند اثربخش بودن تلاش‌های اصلاحی مدارس برای افزایش موفقیت تحصیلی، بستگی زیادی به کیفیت آموزش دارد. اما هنوز روش‌های مفید و کارآمدی برای توصیف فعالیت‌های کلاسی و به‌طور مستقیم برای ارزیابی اثر تلاش‌های اصلاحی بر محیط یادگیری دانش‌آموزان وجود ندارد. تویتنس و دووس^۳ (۲۰۱۱) نیز از آموزش و یادگیری به‌عنوان قلب تعلیم و تربیت نام برده‌اند. علاوه بر این، دنیلسون و مک‌گروت^۴ (۲۰۰۰) معتقدند که آموزش، بر مهارت‌های قابل قبول برای دانش‌آموزان از قبیل تفکر انتقادی، حل مسئله، یادگیری مادام‌العمر، یادگیری مشارکتی و فهم عمیق، عامل تأثیرگذاری تلقی می‌شود (ص. ۱۴). نظر به اهمیت آموزش و یادگیری، معلم با کیفیت بالا به‌عنوان مجری آموزش و عامل تأثیرگذار بر یادگیری دانش‌آموزان، به یکی از اهداف مهم سیاست‌های آموزشی در قرن ۲۱ تبدیل شده است (منجیت^۵، ۲۰۱۱).

با افزایش توجه به مسئولیت و توانایی معلمان، ارزیابی عملکرد معلمان، توسط پژوهشگران مختلف، به‌عنوان یک راهکار برای چند موضوع مدنظر قرار گرفت که عبارت‌اند از: افزایش بهبود معلمان، قضاوت درباره آن‌ها و حذف معلمان ناکارآمد (اوالوسا و اسیل^۶، ۲۰۰۷)، پرداخت حقوق و مزایا، فهم نحوه تصمیم‌گیری مدیران درباره عملکرد معلمان (کیمبل و میلانوسکی^۷، ۲۰۰۹)، بهبود اثربخشی معلمان و حمایت از توسعه حرفه‌ای آن‌ها (دنیلسون، ۲۰۰۱؛ استرانگ و توکر^۸، ۲۰۰۳؛ نقل از تویتنس و دووس، ۲۰۱۱)، البته نارضایتی گسترده‌ای در میان مدیران مدارس و معلمان بر سر نحوه ارزشیابی معلمان وجود دارد (دنیلسون، ۲۰۰۱). تویتنس و دووس (۲۰۱۱)، معتقدند اگر معلمان خودشان بازخورد دقیق و مفیدی از آموزش خود دریافت نکنند به بازخوردها اهمیتی نخواهند داد و برای یادگیری حرفه‌ای و نیز برای تغییر در روش آموزشی خویش مسئولیتی را تقبل نمی‌کنند.

نظر به اهمیت موضوع، هدف این پژوهش توصیف دیدگاه‌های مختلف درباره ارزشیابی، و سپس معرفی یک مدل نرم‌افزاری مبتنی بر تکنیک‌های هوش مصنوعی، برای ارزشیابی معلمان و اساتید می‌باشد. یکی از دلایل نبود اتفاق نظر درباره مفهوم ارزشیابی، وجود مفاهیمی حاکی از عدم قطعیت در روش‌های اندازه‌گیری و سنجش عملکرد است. بر این اساس، در این پژوهش تلاش شده است تا یک مدل ارزشیابی کارآمد و مؤثر برای تخمین میزان عملکرد و دسته‌بندی معلمان، در جهت دسترسی به اهداف، ارائه شود. در این مدل با توجه به وجود عدم قطعیت در مفاهیم و روش‌های اندازه‌گیری

عملکرد، از عبارات کلامی و منطق فازی^۹ برای قطعیت بخشیدن به موارد غیرقطعی استفاده شده است. همچنین، جهت کاهش دخالت ذهن، برای دسته‌بندی از درخت تصمیم‌گیری^{۱۰} به‌عنوان یکی از تکنیک‌های قدرتمند هوش مصنوعی کمک گرفته شده است.

■ دلایل انتخاب مدل دنیلسون (چارچوبی برای آموزش)

با توجه به اهمیت شناسایی ویژگی‌ها و کارکردها و تأثیر آن بر روی کارکنان سازمان، مدل‌های مختلفی برای پژوهش‌های مرتبط با رفتار سازمانی طراحی و ارائه گردیده است.

جدید بودن مدل دنیلسون نسبت به مدل‌های پیشین و مبنای رفتارگرایی این مدل و همچنین کامل بودن آن از نظر شاخص‌های اندازه‌گیری و سنجش کامل ابعاد رفتار سازمانی، از دلایل انتخاب این مدل به‌شمار می‌آید. قابلیت استفاده از مدل دنیلسون در تمام سطوح سازمانی یکی دیگر از ویژگی‌های این مدل است. دنیلسون (۱۹۹۶، ص. ۱)، با مرور معیارها و ویژگی‌های معلمان کارآمد در میان پژوهش‌های تجربی و نظری انجام‌شده، مدل جدیدی به نام «چارچوبی برای آموزش»^{۱۱} در جهت بهبود و ارتقای یادگیری دانش‌آموزان ارائه داد. دنیلسون (۲۰۰۱) معتقد است تنها با درک این موضوع که «یادگیرندگان چگونه یاد می‌گیرند» مربیان قادر خواهند بود برنامه‌هایی را برای آموزش طراحی کنند که حداکثر یادگیری از آن حاصل شود؛ ضمناً به نظر او و هدف اولیه آموزش باید مشارکت یادگیرنده در ساخت دانش باشد و مسئولیت هر معلم را استفاده از منابع برای رسیدن به این هدف می‌داند (دنیلسون، ۱۹۹۶، ص. ۲۵). بنابراین باید گفت توسعه مدل چارچوبی برای آموزش بر اساس تئوری یادگیری ساختن‌گرایی صورت گرفته و به‌وسیله هیئت ملی استانداردهای آموزش حرفه‌ای برای بررسی دانش و مهارت‌های معلمان طراحی گردیده است (منجینت، ۲۰۱۱). این مدل دارای چهار حیطه آماده‌سازی، محیط کلاسی، آموزش و مسئولیت‌پذیری است. این حیطه‌ها جمعاً دارای ۲۲ بعد می‌باشند و هر بعد نیز خود دارای ۲ تا ۵ مؤلفه است. در مجموع این چهار حیطه شامل ۶۶ مؤلفه می‌باشند. مؤلفه‌ها توصیف‌کننده ویژگی‌ها و رفتارهای معلمان هستند. شیوه تحلیل نتایج به‌دست‌آمده در این مدل بدین صورت است که معلمان بر اساس رتبه‌های به‌دست‌آمده در چهار حیطه، در چهار گروه از نظر شایستگی عملکرد قرار می‌گیرند. سطوح عملکردی به ترتیب عبارت‌اند از نوآموز، آماتور، حرفه‌ای و پیشرفته است.

هنمن، میلانوسکی، کیمبل و اودن^{۱۲} (۲۰۰۶) معتقدند که این مدل قابلیت کاربرد در همه سطوح تحصیلی و موضوعات آموزشی را داراست و دامنه گسترده‌ای از عملکرد معلمان از شروع تا پایان فعالیت‌ها را در بر می‌گیرد. پژوهش‌های صورت گرفته بر اساس مؤلفه‌های این مدل، حاکی از این است که معلمان که نمره‌های بالایی در این مدل به دست آورده‌اند دانش‌آموزان آن‌ها نیز از وضعیت تحصیلی مناسبی برخوردارند (استرانگ، گارگانی و حتیفلیگلو^{۱۳}، ۲۰۱۱).

■ منطق فازی

منطق فازی را نخستین بار لطفعلی عسگرزاده (باکو، ۱۹۲۱ م) در سال ۱۹۶۵ م با انتشار مقاله‌ای با عنوان مجموعه‌های نادقیق مطرح کرد (تشسنة لب، صفاپور و افیونی، ۱۳۸۹). وی هدف از ابداع منطق نادقیق را توصیف پدیده‌های نادقیق و مبهم به وسیله مدل‌های ریاضی برای به‌کارگیری و استنتاج نتایج مناسب از آن‌ها بیان می‌کند. منطق فازی، بر اساس تئوری مجموعه‌های فازی، تلاش دارد تا روشی را در ارائه و استدلال در دنیای واقعی، را در مواجهه با عدم قطعیت، دنبال کند. عدم قطعیت می‌تواند به دلیل عمومیت، شانس، ابهام، ضعف و یا نقص در دانش باشد. مجموعه فازی را در ریاضیات می‌توان با نسبت دادن مقدار به هر امر ممکن، در مجموعه مرجع، تعریف کرد که این مقادیر درجه عضویت را در مجموعه فازی نشان می‌دهد. در تئوری کلاسیک مجموعه‌ها، یک عضو یا به یک مجموعه تعلق دارد و یا ندارد که این مجموعه‌ها را مجموعه‌های صریح^{۱۴} گویند. اما در مجموعه‌های فازی همانند نگاهت، هر عضو مجموعه مرجع فازی را به بازه صفر و یک می‌نگارد (ویجی آلاکشمی پای، ۱۳۹۱). به‌عنوان مثال، «مجموعه فازی افراد قdblند» را در نظر بگیریم؛ فردی با قد ۱۸۰ سانتی‌متر، کاملاً عضو این مجموعه بوده و درجه عضویت او یک است، اما فرد دیگری با قد ۱۶۰ سانتی‌متر عضو این مجموعه نیست و درجه عضویت او صفر است. ولی افرادی با قد ۱۷۰ سانتی‌متر با درجه عضویت ۰/۷ و قد ۱۵۰ سانتی‌متر با درجه عضویت ۰/۲ عضوی از مجموعه فازی افراد قdblند هستند.

■ درخت تصمیم‌گیری

درخت تصمیم یکی از مشهورترین و قدیمی‌ترین روش‌های ساخت مدل دسته‌بندی است. در الگوریتم‌های مبتنی بر درخت تصمیم، دانش خروجی به‌صورت یک درخت از حالات مختلف مقادیر ویژگی‌ها ارائه می‌شود (صنعی آباده، محمودی و طاهرپرور، ۱۳۹۱). درخت تصمیم قادر به تولید توصیفات قابل‌درک برای انسان، از روابط موجود در یک مجموعه داده است و می‌تواند برای وظایف دسته‌بندی و پیش‌بینی نیز به‌کار رود. این ساختار تصمیم‌گیری می‌تواند به شکل تکنیک‌های ریاضی و محاسباتی که به توصیف، دسته‌بندی و عام‌سازی یک مجموعه از داده‌ها کمک می‌کنند نیز معرفی شوند. از نقاط قوت درخت تصمیم‌گیری می‌توان به موارد ذیل اشاره کرد (چین و یین، ۲۰۰۹):

۱. فهم مدل ایجادشده توسط درخت تصمیم‌گیری آسان است؛ به عبارت دیگر با اینکه ممکن است الگوریتم‌هایی که درخت را ایجاد می‌کنند چندان ساده نباشند ولی فهم نتایج آن آسان می‌باشد؛
۲. درخت تصمیم‌گیری این توانایی را دارد که پیش‌بینی‌های خود را در قالب یک رشته قوانین ارائه دهد.
۳. نیاز به محاسبات خیلی پیچیده‌ای برای دسته‌بندی داده‌ها ندارد؛
۴. برای انواع مختلف داده‌ها قابل‌استفاده است؛

۵. درخت تصمیم‌گیری نشان می‌دهد که چه فیلدها یا متغیرهایی تأثیرات مهمی در پیش‌بینی کردن و دسته‌بندی دارند.

■ روش‌شناسی پژوهش

در این پژوهش ارزشیابی از اهداف کسب‌شدهٔ دو نظام «بخوانیم و بنویسیم» و «خوانداری-نوشتاری» انجام شد و نتایج مقایسه گردید. از آنجاکه در پژوهش حاضر به بررسی رابطهٔ چند متغیر پرداخته شد، روش تحقیق از نوع علی-مقایسه‌ای بوده و در آن از آزمون‌های پارامتری و غیرپارامتری مناسب، حسب مورد، استفاده شد. در این نوع پژوهش، با توجه به متغیر وابسته به بررسی علل احتمالی وقوع پرداخته می‌شود. به بیان دیگر این نوع تحقیق گذشته‌نگر است و سعی بر آن دارد از معلول به علت احتمالی پی برد (سرمد، بازرگان و حجازی، ۱۳۸۹). در انجام این پژوهش دو گروه نمونه با یکدیگر مقایسه شدند و عوامل مؤثر در نتایج هر گروه مشخص گردید.

■ مبانی نظری و پیشینهٔ تحقیق

طی چند سال اخیر تحقیقات زیادی در خصوص ارزیابی عملکرد آموزشی و ارزشیابی تحصیلی انجام شده است. در هر یک از این تحقیقات ضرورت و شایستگی روش ارزشیابی بررسی و سعی شده است تا میزان اعتبار و اثربخشی آن بهبود یابد. این تحقیقات بیانگر این واقعیت است که افزایش دقت و کارایی روش‌های ارزشیابی همواره مورد توجه اعضای هیئت‌علمی، دانشجویان و محققان بوده است و خواهد بود (آدم و اسیرا^{۱۶}، ۲۰۰۷). در این تحقیقات جنبه‌های مختلف ارزشیابی مانند شرایط امتحان، اعمال مقدماتی، ابزارهای مورد استفاده، ابعاد تربیتی و روان‌شناختی و آثار هر کدام از آنها بر فرآیند یادگیری مطالعه شده است (فراستخواه، بازرگان و طباطبایی، ۲۰۰۷). آدم و اسیرا (۲۰۰۷)، یک مدل فازی برای ارزشیابی و انتخاب کارمند نمونه، با توجه به میزان شایستگی و صلاحیت، ارائه کردند. آنها با استفاده از روش تحلیل سلسله مراتبی^{۱۷} به عوامل و معیارهای مؤثر در این ارزشیابی و انتخاب توجه کردند. محقر، امین ناصری و معماریانی (۲۰۰۵)، با استفاده از منطق فازی یک مدل ریاضی برای ارزشیابی عملکرد و کارایی مجلس شورای اسلامی در دوره‌های مختلف ارائه کردند. لازارویچ^{۱۸} (۲۰۰۱) برای انتخاب و ارزیابی کارمند نمونه از تلفیق روش‌های هوش مصنوعی و منطق فازی استفاده کرد. دریگاس، کورمنوس، رتاروس و کورمنوس^{۱۹} (۲۰۰۴)، با استفاده از شبکه‌های عصبی فازی یک سیستم خبره برای ارزشیابی عملکرد و انتخاب کارمند نمونه ارائه و در این مدل از یک فرآیند تحلیل سلسله مراتبی استفاده کردند. هوانگ و لی^{۲۰} (۱۹۹۴)، شهرضایی (۲۰۱۰) و بالأخره میرفخرالدینی، اولیا و جمالی (۲۰۰۹)، نیز چند مدل تصمیم‌گیری فازی با اهداف چندگانه ارائه کردند. هدف همهٔ این مدل‌ها ارائهٔ چند روش ارزشیابی عملکرد و مقایسهٔ نتایج آنها با یکدیگر بوده است.

■ ضرورت و اهمیت انجام تحقیق

یکی از مهم‌ترین ضعف‌های مدل دنیلسون و سایر مدل‌های ارزشیابی، این است که اهمیت هر معیار در تصمیم‌گیری نهایی قابل‌سنجش نیست. به عبارت دیگر، شناسایی اینکه کدام عامل یا معیار، تصمیمات مدیران را بیشتر تحت تأثیر قرار می‌دهد تا حدی دشوار است. سیستم پیشنهادی این مشکل را برطرف می‌کند، زیرا اهمیت و وزن هر معیار به صورت متغیرهای زبان طبیعی (کم‌اهمیت، بااهمیت متوسط، بااهمیت و خیلی بااهمیت) قابل‌محاسبه می‌باشد. در واقع در موضوع ارزیابی عملکرد آموزشی اساتید و معلمان، مثل سایر موضوعات علوم انسانی و مدیریت، قطعیت ناچیز است. از طرف دیگر، متغیرهای مداخله‌گر غیرقابل‌کنترل یا با قابلیت کنترل کم در این ارزیابی به صورت مستقیم یا غیرمستقیم تأثیر دارند. بر این اساس است که استفاده از روش‌های دقیق ریاضی و آمار برای این موضوع مناسب نیست. در این پژوهش با توجه به ماهیت غیرقطعی بودن موضوع ارزشیابی، روش امتیازدهی به معیارها، با استفاده از منطق و سیستم‌های فازی پیشنهاد و در نهایت، با هدف کاهش قضاوت ذهن و افزایش دقت تصمیم‌گیری، جهت دسته‌بندی معلمان بر اساس معیارهای کیفی، از درخت تصمیم‌گیری استفاده شده است. در واقع ارزشیابی بر مبنای مدل دنیلسون، به کمک معیارهای کیفی (فازی شده) و الگوریتم درخت تصمیم‌گیری انجام می‌شود.

■ روش

با توجه به شواهد موجود، گام‌های ذیل جهت رسیدن به هدف نهایی در نظر گرفته می‌شود (شکل ۱).



شکل ۱. فرآیند ارزشیابی و رتبه‌بندی مبتنی بر مدل پیشنهادی

جمع‌آوری داده‌های مورد نیاز این پژوهش در دو مرحله و توسط یک فرم نظرسنجی و یک فرم ارزشیابی انجام شد. فرم‌های نظرسنجی جهت اولویت‌بندی معیارهای مطرح در تدریس، جهت ورود به سیستم فازی (توسط اساتید برتر و نمونه) و فرم‌های ارزشیابی معلمان، جهت آموزش مدل (توسط مدیران مدارس) تکمیل شد.

در واقع برای تعریف تابع فازی و نیز تعریف میزان اهمیت هر معیار، نیاز به نظر افراد خبره می‌باشد. لذا در مرحله اول فرم‌های نظرسنجی توسط ۱۰ نفر از معلمان و اساتید نمونه که در سال تحصیلی ۹۴-۹۳ در آموزش و پرورش ناحیه ۴ کرج مشغول به کار بودند تکمیل شد. انتخاب استادان از بین استادان نمونه بر مبنای سنجش سطح تدریس ایشان انجام شد. این افراد توسط اداره آموزش و پرورش معرفی شدند. در این بخش از آنان خواسته شد که با توجه به سوابق تدریس و نیازهایی که در راستای تدریس احساس نموده‌اند، بر اساس مدل دنیلسون (۱۹۹۶) به اولویت‌بندی معیارهای مطرح در تدریس بپردازند. بدین صورت که هر نفر از ایشان به هر کدام از معیارها امتیاز ۱ تا ۱۰ دادند. در نهایت حاصل جمع امتیازات هر معیار در جدول ۱ درج شد. لازم به ذکر است که چیدمان معیارها در فرم‌های نظرسنجی به صورت تصادفی انجام شد تا هیچ‌گونه پیش‌داوری از طرف ایشان انجام نگرفته باشد. تابع اول فازی (میزان اهمیت هر معیار) با توجه به اطلاعات جدول ۱ طراحی شد.

در مرحله دوم از جمع‌آوری اطلاعات مورد نیاز، نمونه فرم‌های ارزشیابی، با استفاده از مصاحبه و پرسش‌نامه از مدیران ۴ مدرسه از آموزش و پرورش ناحیه ۴ کرج، که با روش نمونه‌گیری خوشه‌ای انتخاب شده بودند، جمع‌آوری شد. جامعه آماری شامل ۱۰۶ نفر از معلمان و دبیران این مدارس بود که با استفاده از مدل چارچوبی دنیلسون، پس از امتیازدهی توسط مدیران، رتبه‌بندی ایشان انجام شد و سپس رتبه‌ها برای ساخت درخت تصمیم‌گیری مورد استفاده قرار گرفت.

همان‌طور که ذکر شد با توجه به اولویت‌بندی‌های تعیین‌شده توسط استادان در پرسش‌نامه‌ها و همچنین نتایج تحقیقات مراجع (فنلر، ۱۹۷۱؛ شوئر، ۱۳۷۷؛ جویس و ویل، ۱۳۷۹)، طراحی توابع فازی انجام گرفت. با توجه به نتایج حاصل از پرسش‌نامه‌ها از بین ۲۲ معیار آموزشی اولیه، ۱۴ مورد آن اصلی و مهم ارزیابی شده بود که این معیارها با توجه به همان نتایج اولویت‌بندی استادان مشخص گردید. جدول ۱ مرتب‌سازی این معیارها را، با توجه به حاصل جمع امتیازدهی استادان، نمایش می‌دهد. نحوه امتیازدهی بر اساس نظر استادان در مورد اهمیت و اولویت‌بندی معیارها در تدریس دروس مختلف می‌باشد و حداکثر امتیاز ۱۰ و حداقل امتیاز ۱، منظور گردیده است. حاصل جمع امتیازات داده‌شده توسط همه استادان محاسبه شده و معیار با بیشترین امتیاز در صدر جدول ۱ قرار گرفته است و دیگر معیارها به ترتیب در سطرهای بعدی جدول منظور گردید.

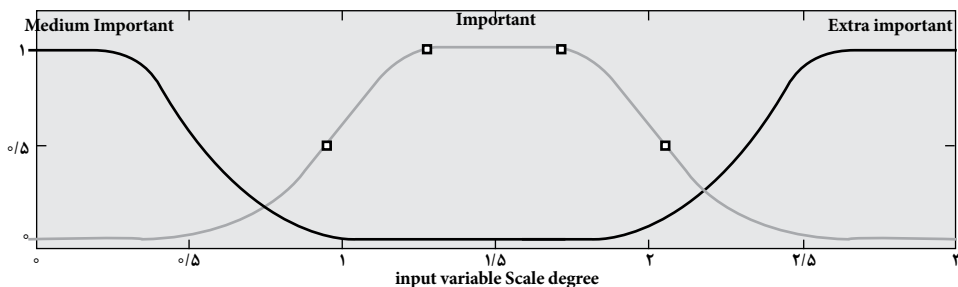
روشی نوین در ارزشیابی و رتبه‌بندی معلمان با استفاده از درخت تصمیم‌گیری

جدول ۱. اولویت‌بندی معیارهای مدل چارچوبی برای آموزش دنیلسون (۱۹۹۶)

ردیف	مدل دنیلسون	امتیاز (۱۰۰)	فازی سازی معیارها
۱	داشتن دانش درباره محتوا و تعلیم و تربیت	۹۲	فوق العاده بااهمیت
۲	انعطاف‌پذیری و اشتیاق برای انجام کار	۹۰	فوق العاده بااهمیت
۳	ایجاد فرهنگ یادگیری	۸۲	فوق العاده بااهمیت
۴	درگیر کردن دانش‌آموزان در یادگیری	۷۵	فوق العاده بااهمیت
۵	تبادل اطلاعات به‌طور واضح و دقیق	۷۳	فوق العاده بااهمیت
۶	استفاده از سؤال و فنون مباحثه	۶۳	بااهمیت
۷	چگونگی ارزیابی دانش‌آموزان	۵۱	بااهمیت
۸	داشتن دانش درباره منابع	۴۷	بااهمیت
۹	رشد و توسعه حرفه‌ای	۴۴	بااهمیت
۱۰	ایجاد محیطی از احترام و روابط دوستانه	۲۹	با اهمیت متوسط
۱۱	تفکر عمیق درباره تدریس	۲۶	با اهمیت متوسط
۱۲	داشتن دانش درباره دانش‌آموزان	۲۲	با اهمیت متوسط
۱۳	انتخاب اهداف آموزشی	۲۱	با اهمیت متوسط
۱۴	مدیریت فرآیندهای کلاسی	۲۰	با اهمیت متوسط

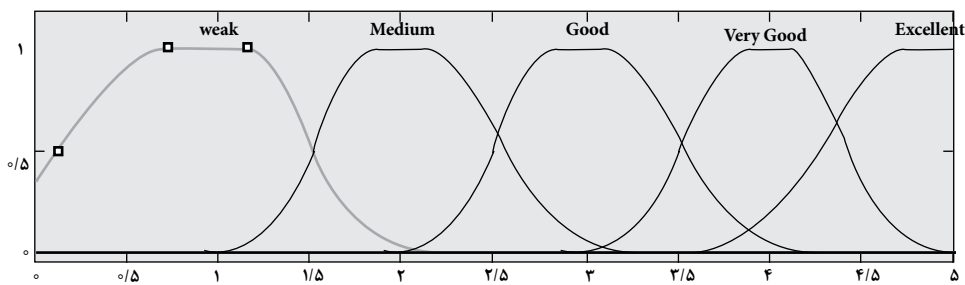
فازی سازی معیارها

دو تابع فازی برای مقداردهی «کیفی» به معیارهای ارزشیابی تعریف شد. تابع فازی اول، با استفاده از اطلاعات جدول ۱ تعیین می‌کند که معیارها از چه میزان اهمیتی در سنجش نهایی برخوردارند (شکل ۲). در واقع این تابع به‌مثابه یک ضریب برای مقدار هر معیار عمل می‌کند.



شکل ۲ تابع اول سیستم فازی: میزان اهمیت هر معیار

تابع دوم برای بیان مقدار کیفی (به‌جای عدد ثابت) به معیارهای ارزشیابی تعریف شد. در واقع برای بیان معیار سنجیده شده در نمونه مورد سنجش از کلمات عالی، خیلی خوب، خوب، متوسط و ضعیف استفاده شد که به‌صورت توابع فازی پیاده‌سازی گردید (شکل ۳).



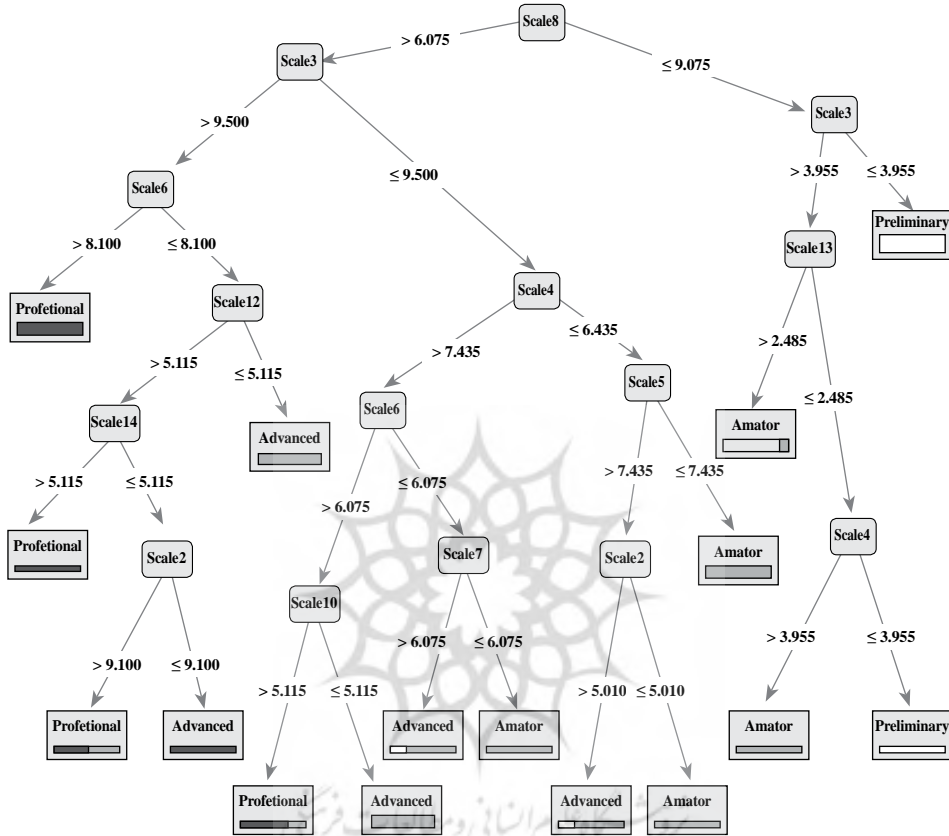
شکل ۳ تابع دوم سیستم فازی: سنجش کیفی هر معیار

پس از مقداردهی به هر معیار و با توجه به درجه اهمیت آن، سنجش با استفاده از قوانین فازی انجام می‌شود.

استفاده از درخت تصمیم‌گیری برای دسته‌بندی معلمان

منطق فازی در ترکیب با سایر الگوریتم‌های هوش مصنوعی دارای کاربردهای زیادی است. در این پژوهش، منطق فازی با الگوریتم درخت تصمیم ترکیب می‌شود و ورودی‌های آن را جهت دسته‌بندی تشکیل می‌دهد. درختان تصمیم‌گیری به‌منظور پیشگویی یا کلاس‌بندی داده‌ها بر اساس مجموعه قوانین تصمیم ایجاد می‌شوند. دسته‌بندی داده‌ها با درختان تصمیم‌گیری یک فرآیند دومارحله‌ای است، که در مرحله اول که به آن مرحله آموزش گفته می‌شود، مدل ابتدا برای استفاده در دنیای واقعی باید آماده‌سازی و تست شود (در اصطلاح آموزش داده شود). مجموعه آموزشی این مدل شامل ۱۰۶ نمونه ارزیابی شده از معلمان هستند. منظور از آموزش و تست مدل، پیش‌بینی رتبه کارکنان برای هر یک از رکوردهاست. پس از آموزش و تست موفقیت‌آمیز مدل، می‌توان از آن جهت ارزشیابی و رتبه‌بندی کارکنان استفاده کرد.

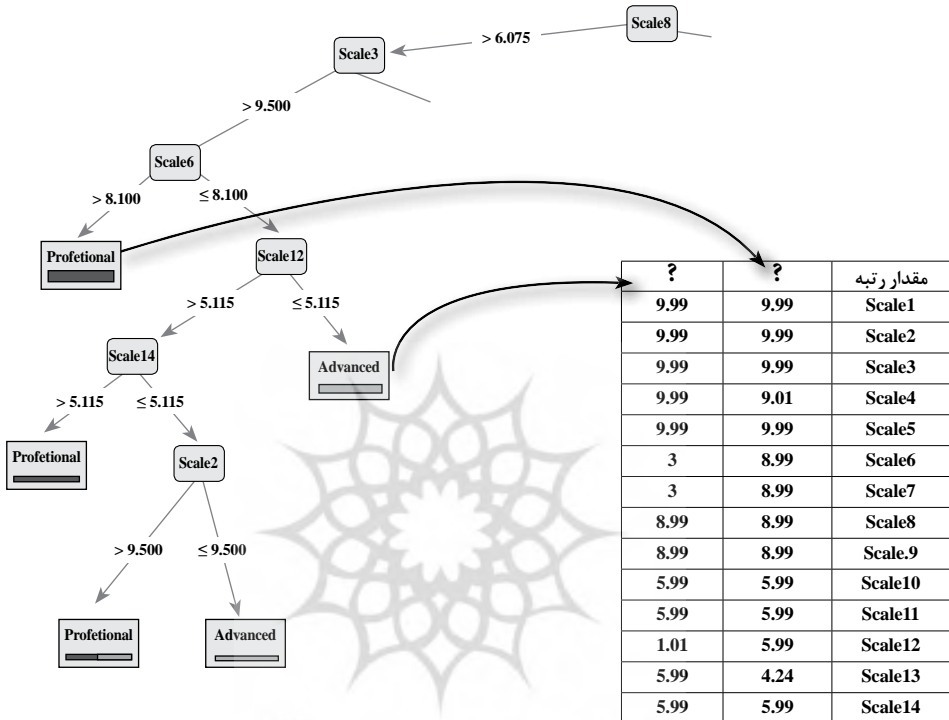
به این ترتیب در اینجا یک مجموعه شامل ۱۰۶ رکورد داریم و هدف ما ساخت یک درخت تصمیم بر اساس آن‌هاست. مطابق با آنچه ذکر شد درخت تصمیم‌گیری مربوط به داده‌های جمع‌آوری شده، مطابق شکل ۴ ترسیم شد.



شکل ۴ درخت تصمیم‌گیری رتبه‌بندی معلمان. پیاده‌سازی شده در نرم‌افزار Rapid Miner نسخه ۵

در درخت تصمیم رسم شده مربوط به داده‌ها، چهارده معیار کد شده از جدول ۱، با Scale1 تا Scale14 و رتبه‌های نهایی معلمان در مدل دنیلسون (شامل نوآموز، آماتور، پیشرفته و حرفه‌ای)، با کلمات، Professional، Advanced، Amatore، Preliminary. نمایش داده شده است. در الگوریتم‌های مبتنی بر درخت تصمیم، دانش خروجی به صورت یک درخت از حالات مختلف مقادیر رکوردها (نمونه‌های ارزیابی شده) ارائه می‌شود. نمایش دانش به شکل درخت سبب شده است که دسته‌بندی‌های مبتنی بر درخت تصمیم کاملاً قابل تفسیر باشند. به‌عنوان مثال با استفاده از درخت تصمیم می‌توان با دقت بالایی این نتیجه را گرفت که اگر مقدار معیار هشتم (داشتن دانش درباره منابع) کمتر از ۶۰۷۵ و مقدار معیار سوم (ایجاد فرهنگ یادگیری) کمتر از ۳/۹۵۵ باشد، فرد مربوطه در رتبه Preliminary قرار می‌گیرد. جهت روشن‌تر شدن نحوه پیش‌بینی و رتبه‌بندی رکوردها، در شکل ۵ شاخهٔ برش‌خوردهٔ سمت

چپ درخت تصمیم به همراه دو نمونه رکورد آزمایشی از رکوردهای ارزیابی شده، نمایش داده شده است. هدف، پیش‌بینی مقدار رتبه مربوط به این رکوردها، بر اساس درخت تصمیم موجود است.



شکل ۵. آزمایش مدل درخت تصمیم‌گیری بر روی دو رکورد آزمایشی مطالعات فزنی

همان‌طور که در شکل ۵ ملاحظه می‌شود فرآیند تعیین مقدار رتبه نمونه‌های آزمایشی از ریشه درخت تصمیم که شاخص ۸ است آغاز می‌شود. مقدار شاخص ۸ هر دو رکورد ۹,۹۹ می‌باشد، لذا با توجه به عدد روی یال سمت چپ (۶۰۷۵) شاخه سمت چپ درخت انتخاب خواهد شد. در ادامه، وضعیت معیار شاخص ۳ مورد بررسی قرار می‌گیرد که چون مقدار این معیار نیز در هر دو رکورد آزمایشی برابر ۹/۹۹ است، با توجه به عدد روی یال (۹/۵)، باز هم شاخه سمت چپ انتخاب خواهد شد. در ادامه، مقدار شاخص ۶ بررسی می‌شود که چون مقدار شاخص ۶ در رکورد آزمایشی اول ۸/۹۹ و بیشتر از ۸/۱۰ است، با انتخاب شاخه سمت چپ به گره Professional که برگ درخت محسوب می‌شود، خواهیم رسید. لذا رتبه رکورد آزمایشی مربوط به ستون اول با Professional مقادری می‌شود. در مورد رکورد آزمایشی ستون دوم نیز به دلیل اینکه مقدار شاخص ۶ کمتر از ۸/۱۰ است شاخه سمت راست انتخاب می‌شود و در نهایت با بررسی مقدار شاخص ۱۲ مقدار رتبه این رکورد

روشی نوین در ارزشیابی و رتبه‌بندی معلمان با استفاده از درخت تصمیم‌گیری

Advanced (رتبه رشته‌ای) در نظر گرفته می‌شود.

جدول ۲ ماتریس درهم‌ریختگی^{۲۱} مربوط به درخت تصمیم‌گیری را نمایش می‌دهد.

جدول ۲. ماتریس درهم‌ریختگی درخت تصمیم

دقت دسته‌بندی	حرفه‌ای	پیشرفته	آماتور	نوآموز	تعداد نمونه‌های صحیح	نام دسته
۱۰۰٪	۰	۰	۰	۲۴		نوآموز
۹۲٫۵۹٪	۰	۱	۲۵	۱		آماتور
۸۷٫۸۸٪	۱	۲۹	۳	۰		پیشرفته
۹۵٫۴۵٪	۲۱	۱	۰	۰		حرفه‌ای
۹۳٫۴۰٪ : پارامتر صحت ^{۲۲}						

در حوزه هوش مصنوعی، ماتریس درهم‌ریختگی به ماتریسی گفته می‌شود که در آن عملکرد مدل ساخته‌شده را نشان می‌دهد. هر ستون از ماتریس، تعداد نمونه‌های مقادیر دسته‌بندی‌شده را نشان می‌دهد. در صورتی که هر سطر نمونه‌ای واقعی (درست) را در بر دارد.

برای مثال سطر دوم ماتریس، مربوط به دسته آماتور، نشان می‌دهد که از کل ۱۰۶ نمونه موجود (معلمان ارزشیابی شده توسط مدیران)، تعداد ۲۷ نفر دارای رتبه آماتور هستند که مدل ساخته‌شده این پژوهش، تعداد ۲۵ نفر را درست و دو نمونه را اشتباه تشخیص داده است؛ به این معنی که درخت تصمیم‌گیری، یکی از نمونه‌ها را به اشتباه در دسته نوآموز و یکی از رکوردها را به اشتباه در دسته حرفه‌ای دسته‌بندی کرده است. همان‌طور که در جدول ۲ ملاحظه می‌شود دقت مدل ۹۳٪ است که این نشان‌دهنده دقت بالایی دسته‌بندی توسط مدل ساخته شده است. در خارج از مقوله هوش مصنوعی این ماتریس معمولاً ماتریس خطا نیز نامیده می‌شود.

■ نتیجه‌گیری ■

دستیابی نظام تعلیم و تربیت هر کشور به معلمان باکیفیت، بدون شناسایی نقاط ضعف و قوت و توسعه حرفه‌ای آن‌ها، همراه با آموزش‌های مداوم و متناسب با تغییرات ایجادشده در فناوری اطلاعات، امکان‌پذیر نمی‌باشد. از جمله راهکارهای مهم برای رسیدن به اهداف ذکرشده در اکثر کشورهای دنیا ارزشیابی مداوم معلمان است. مدل‌های متعددی از ارزشیابی برای تحلیل فعالیت‌های آموزشی معلمان مطرح شده است که در این پژوهش سعی شد با استفاده از یک مدل نرم‌افزاری به دسته‌بندی معلمان بر اساس معیارهای چارچوبی دنیلسون پرداخته شود. در واقع ترکیب پارادایم‌های هوش مصنوعی باعث سهولت و افزایش دقت و انعطاف‌پذیری در انجام پروژه‌ها می‌شود. در روش ارائه‌شده ابتدا معیارهای سنجش مدل دنیلسون که پیش از این با اعداد ثابت سنجیده می‌شد، توسط کارشناسان اولویت‌بندی و به متغیرهای کلامی فازی و کیفی تبدیل شد. پس از ساخت مدل نرم‌افزاری سیستم استنتاج فازی و درخت تصمیم، مدل ساخته شده ارزیابی شده و دقت و صحت آن اندازه‌گیری شد. نتایج نشان می‌دهد که مدل نرم‌افزاری ساخته‌شده نسبت به روش سنتی، به دلیل انجام مراحل در محیط نرم‌افزاری، استفاده از مقادیر کیفی برای معیارها و همچنین استفاده از تکنیک‌های هوش مصنوعی، سرعت، دقت (۹۳٪) و انعطاف‌پذیری بسیار بالاتری جهت انجام عملیات رتبه‌بندی و ارزیابی عملکرد کارکنان را دارد. این مدل بر اساس نوشته‌ها و مصاحبه با متخصصان حوزه آموزش، با چهارده معیار مهم‌تر، از معیارهای ارزیابی عملکرد مدل دنیلسون، ارزشیابی عملکرد معلمان و اساتید را انجام می‌دهد. در مجموع، این مدل می‌تواند به‌عنوان یک سیستم پشتیبان تصمیم‌گیری، توسط مدیران آموزشی، برای ارزیابی دقیق فعالیت‌های آموزشی، صلاحیت‌ها، رتبه‌بندی و پرداخت حقوق و مزایای اساتید و معلمان در نظام آموزشی مورد استفاده قرار گیرد.

منابع

- جویس، بروس و ویل، مارشال. (۱۳۷۹). *الگوهای تدریس* (ترجمه محمدرضا برنجی). تهران: انتشارات صحیفه.
- شوثر، لول. (۱۳۷۷). *اندازه‌گیری و ارزشیابی در آموزش و پرورش* (ترجمه حمزه گنجی، چاپ ششم). تهران: انتشارات بعثت.
- صنیعی‌آباد، محمد، محمودی، سینا و طاهرپرور، محدثه. (۱۳۹۱). *داده‌کاوی کاربردی*. تهران: انتشارات نیاز دانش.
- فنلر، جرج. (۱۹۷۱). *آشنایی با فلسفه آموزش و پرورش* (ترجمه فریدون بازرگان). انتشارات: دانشگاه تهران.
- ویجی‌الاکشمی پای، راجاسکاران. (۱۳۹۱). *شبکه‌های عصبی، منطق فازی و الگوریتم ژنتیک: ترکیب و کاربرد* (ترجمه محمود کشاورز). تهران: انتشارات نورپردازان.
- Adem, G., & Esra, K. (2007). A fuzzy model for competency-based employee evaluation and selection. *Computer and Industrial Engineering*, 52, 143-161.
- Avalos, B., & Assael, J. (2007). Moving from Resistance to Agreement: The Case of the Chilean Teacher Performance Evaluation. *Educational Research*, 45, 254-266.
- Chih, F., & yen, J., (2009). Earnings management prediction: a pilot study of combining neural networks and decision trees. *Expert Systems with Applications*, 36, 7183-7191.
- Clare, L., & Aschbacher, P. R. (2001). Exploring the Technical Quality of Using Assignments and Student Work as Indicators of Classroom Practice. *Educational Assessment*, 7, 1, 39-59.
- Danielson, C. (1996). *Enhancing Professional Practice: A framework for Teaching*. Alexandria, VA: Association for Supervision and Curriculum Development.
- Danielson, C. (2001). New Trends in Teacher. *Evaluation In: Educational Leadership*, 58(5), 12-15
- Danielson, C., & McGreal, T. L. (2000). *Teacher Evaluation: To Enhance Professional Practice*. Alexandria, VA: Association for Supervision and Curriculum Development & Princeton, NJ: Educational Testing Service
- Drigas, A., Kouremenos, S., Vrettaros, S., & Kouremenos, I.D. (2004). An expert system for job matching of the unemployed. *Expert System with Application*, 26, 217- 224.
- Farasakhah, M., Bazargan, A., & Tabatabaee, M. (2007). Comparative analysis of higher education quality Assurance systems: Dimensions of similarity and diversity in the global experiences. *Quarterly Journal of Research and Planning in Higher Education*, 13(2), 1-19 (in Persian).
- Greaney, V., & Kellaghan, T. (2008). *Assessing National Achievement Levels in Education*. Washington, DC: The World Bank.
- Heneman, H., Milanowski, A., Kimball, S. M., & Odden, A. R. (2006). *Standards-Based Teacher Evaluation as a Foundation for Knowledge and Skill-based Pay*. Philadelphia, PA: University of Pennsylvania, Graduate School of Education, Consortium for Policy Research in Education.
- Hwang, Ch. L., & Lai, Y. J. (1994). *Fuzzy MODM*. New York: Berlin Heidelberg.
- Kimball, S. M., & Milanowski, A. (2009). Within a Standards-Based Evaluation System. *Educational Administration Quarterly*, 45, 1-8.
- Lazarevic, S. P. (2001). Personal selection fuzzy models. *International Transaction in Operational Research*, 8, 89-105.
- Mirfakhradin, S.H., Owlia, M.S., & Jamali, R. (2009). Reverse engineering quality management in center learning higher. *Quarterly Journal of Research and Planning in Higher Education*, 15, 131-157 (in Persian).
- Mohaghar, A., Aminnasery, M.R., & Memariani, A. (2005). A new models for evaluation efficiency group five with fuzzy sets. *Journal Engineering Tarbiat Modares*, 11, 11-27.
- Mangiante, E. M. S. (2011). Teachers Matter: Measures of Teacher Effectiveness in low-Income Minority schools. *Educ Asses Eval Acc*, 23, 41-63.
- Shahrezaei, S.R. (2010). Review evaluation in learning higher iron with presentation a performance. *Quarterly Journal of Research and Planning in Higher Education*, 16, 41-60 (in Persian).
- Strong, M., Gargani, J., & Hacifazlioglu, Ö. (2011). Do We Know a Successful Teacher When we see one? Experiments in the Identification of Effective Teachers. *Journal of Teacher Education*, 62, 367-382.
- Tuytens, M., & Devos, G. (2011). Stimulating Professional Learning Through Teacher Evaluation: An Impossible Task for the School Leader? *Teaching and Teacher Education*, 27(5), 891-899.

پی‌نوشت‌ها

- | | | |
|-------------------------|--|--|
| 1. Greaney & Kellaghan | 9. Fuzzy Logic | 16. Adem & Esra |
| 2. Clare & Aschbacher | 10. Decision Tree | 17. AHP |
| 3. Tuytens & Devos | 11. Framework for Teaching | 18. Lazarevic |
| 4. Denilson & Macgrut | 12. Heneman, Milanowski, Kimball & Odden | 19. Drigas, Kouremenos, Vrettaros & Kouremenos |
| 5. Mangiante | 13. Strong, Gargani & Hacifazlioglu | 20. Hwang & Lai |
| 6. Avalosa & Assael | 14. Crisp sets | 21. Confusion Matrix |
| 7. Kimball & Milanowski | 15. Chih & yen | 22. Accuracy |
| 8. Stronge & Tucker | | |