فصلنامه پژوهش و برنامه‌ریزی در آموزش‌عالی، دوره 21، شماره 3، 1394، صص.120-97

طراحی سامانه شناسایی عوامل مؤثر در پذیرش داوطلبان دورهٔ دکترای تخصصی با استفاده از VPRST

فاطمه قربانی[1] و غلامعلی منتظر[2*]

چکیده

افزایش تقاضا برای ورود به دورهٔ دکترای تخصصی و سرعت کُند افزایش ظرفیت دانشگاهی متناسب با این رشد، موجب می‌شود که داوطلبان در فرایندی گزینشی برای حضور در این دوره قرار گیرند. عوامل بسیاری در انتخاب دانشجویان برای دورهٔ دکترای تخصصی تأثیرگذار است و از آنجا که بسیاری از این عوامل ذهنی و برخاسته از دیدگاههای استادان است، این موضوع با ابهام و عدم قطعیت جدی همراه است. شناسایی اثرگذارترین عوامل در انتخاب دانشجویان و نیز تعیین میزان اثر هریک از عوامل ذهنی مد نظر استادان در پذیرش دانشجویان دکترای تخصصی علاوه بر شفاف کردن فرایند، تأثیری چشمگیر در صرفه‌جویی در زمان و هزینه‌های دانشگاههای پذیرنده دارد. در این پژوهش سامانه‌ای با استفاده از نظریهٔ مجموعه‌های نادقیق با دقت متغیر (VPRST) برای شناسایی عوامل مؤثر در گزینش دانشجویان متقاضی ورود به دورهٔ دکترای تخصصی طراحی شد. در طراحی این سامانه عواملی که در فرایند گزینش دانشجویان ارزیابی می‌شوند، شناسایی و مجموعهٔ داده‌ای با استفاده از اطلاعات سالهای پیش آماده شد. سپس، این مجموعه با استفاده از نظریهٔ مجموعه‌های نادقیق با دقت متغیر تحلیل و عوامل مؤثر در پذیرش دانشجویان دکترای تخصصی شناسایی شد. نتایج این پژوهش برای استخراج قواعد پذیرش داوطلبان استفاده می‌شود و امکان تصمیم‌گیری در ارزیابی و پذیرش داوطلبان را فراهم می‌آورد. به‌منظور ارزیابی عملکرد این سامانه از داده‌های آزمون سالهای 1387 تا 1393 دانشگاه تربیت مدرس استفاده شد. نتایج نشان می‌دهد از بین 12 عامل که در مصاحبه بررسی شد، فقط سه عامل تعداد مقالات، استعداد و قدرت بیان و امتیاز آزمون کتبی در گزینش داوطلبان اثرگذارند و بر مبنای این سه عامل می‌توان دربارهٔ وضعیت 71% داوطلبان تصمیم‌گیری کرد. ضمن اینکه استفاده از عامل قدرت تجزیه و تحلیل علمی به‌عنوان چهارمین عامل اثرگذار می‌تواند به افزایش دقت تصمیم‌گیری تا سطح 91% بینجامد. این سامانه

1. دکتری مهندسی فنّاوری اطلاعات دانشگاه تربیت مدرّس، تهران، ایران: fatemeh.ghorbani@modares.ac.ir
2. دانشیار دانشکده فنی و مهندسی دانشگاه تربیت مدرّس، تهران، ایران.
* مسئول مکاتبات: montazer@modares.ac.ir
دریافت مقاله: 1392/9/7 پذیرش مقاله : 1393/11/28

می‌تواند به‌عنوان یک سامانهٔ تصمیم‌یار برای استادان و دانشگاهها در ارزیابی و پذیرش دانشجویان به‌کار آید.

کلیدواژگان: گزینش دانشجو، استخراج خصیصه، دورهٔ دکترای تخصصی، مجموعه‌های نادقیق با دقت متغیر.

مقدمه

به‌دلیل انبوه داوطلبان ورود به دانشگاههای دولتی و به‌منظور برقراری عدالت آموزشی و گزینش شایسته‌ترین افراد و نیز تخصیص امکانات آموزش‌عالی به افراد مستعد، گزینش دانشجو در تمام دوره‌های آموزش‌عالی ایران به‌صورت متمرکز یا نیمه متمرکز صورت می‌پذیرد. آمار نشان می‌دهد که در سال 1392 حدود 1/1 میلیون تن در آزمون ورودی دوره کارشناسی، حدود 900000 تن در آزمون دوره کارشناسی ارشد و حدود 218000 تن نیز در آزمون دورهٔ دکترای تخصصی شرکت کرده‌اند که از این تعداد حدود 522000 تن در رشته‌های متمرکز دورهٔ کارشناسی، حدود 115000 تن در دورهٔ کارشناسی ارشد و حدود 12000 تن در دورهٔ دکترای تخصصی پذیرفته شده‌اند (http://www.sanjesh.org).

گزینش داوطلبان یکی از فرایندهای پیچیدهٔ تصمیم‌گیری است که اغلب به ارزیابی جامعی از عملکرد آنان نیاز دارد. معیارهای چندگانهٔ گزینش از جمله توان علمی، سوابق تحصیلی، علایق پژوهشی، استعداد فردی و ... باید به‌صورت همزمان در نظر گرفته شوند و بدیهی است به‌دلیل ذهنی بودن این عوامل و برداشتهای مختلف استادان از این موضوع، تصمیم‌گیری آنان معمولاً توأم با نوعی عدم قطعیت است که همین موضوع فرایند تصمیم‌گیری نهایی را بسیار مشکل می‌سازد.

در اغلب مقالاتی که به موضوع گزینش و پذیرش دانشجویان پرداخته شده، هدف یافتن مجموعه‌ای بهینه از معیارها بوده است که بتوان با آنها به‌صورت عادلانه و علمی دانشجویان برتر را رتبه‌بندی و گزینش کرد. این معیارها باید با اهداف و رسالت دانشگاه سازگار باشند و پیش‌بینی درستی از موفقیت افراد پذیرفته شده ارائه دهند؛ به‌عبارتی دیگر، افرادی انتخاب شوند که احتمال موفقیت آنها در این مقطع تحصیلی بیشتر باشد (Wilson, 1999). بدین منظور، لازم است معیارهای مختلفی در فرایند پذیرش دانشجویان مورد ارزیابی قرار گیرد.

بیشتر پژوهشهای انجام شده در حوزهٔ پیش‌بینی موفقیت دانشجویان نشان می‌دهد عوامل علمی (معدل دوره‌های مختلف آموزشی و امتیاز آزمون‌های مختلف) و عوامل عمومی (سن، جنسیت، منطقه جغرافیایی و ...) می‌تواند بر موفقیت دانشجویان اثرگذار باشد. از این‌رو، ارزیابی این عوامل در فرایند پذیرش داوطلبان می‌تواند در موفقیتهای بعدی آنان نقش بسزایی داشته باشد. انتخاب معیارهای ارزیابی مناسب

به انتخاب بهترین داوطلبان و کاهش مخاطرهٔ شکست دانشجویان در فعالیتهای علمی منجر می‌شود (Ali, 2008).

معیارهای انتخاب داوطلبان از دانشگاهی به دانشگاه دیگر متفاوت است، با وجود این، بیشتر دانشگاهها سوابق آموزشی و پژوهشی داوطلبان را معیار مناسبی برای ورود آنان به دوره‌های آموزشی می‌دانند (Gordon, Williams, Hudson & Stewart, 2010). این معیارها باید در روند پذیرش داوطلبان مورد ارزیابی قرار گیرند.

تاکنون روشهای آماری مختلفی از قبیل تحلیل تمایز[3] و رگرسیون به‌صورت سنتی برای پیش‌بینی موفقیتهای علمی متقاضیان به‌کار رفته‌اند (Graham,1991). تحلیل چند معیاره[4] نیز روشی رایج برای گزینش یا رتبه‌بندی متقاضیان بر حسب چند معیار استفاده شده است (Roy, 1996). کارلسون و فولر (Carlsson & Fuller, 1997) از میانگین‌گیری وزن‌دارشدهٔ ترتیبی[5] در انتخاب و گزینش متقاضیان دوره دکتری رشته مهندسی کامپیوتر استفاده‌کرده‌اند. این مسئله نوعی تصمیم‌گیری چند معیاره - چند خبره[6] است. تعداد خصیصه‌های مورد نظر شش مورد بوده و از یازده استاد برای ارزیابی دانشجویان استفاده شده است. لوپز از روش الکتره 3[7] و الگوریتم ژنتیک برای ایجاد رابطهٔ اولویت[8] و رتبه‌بندی میان متقاضیان ورود به دانشگاه استفاده و فرایند گزینش دانشجو را به‌صورت یک مسئله تحلیل چند معیاره فرمول‌بندی کرده است. اساس روش الکتره مبتنی بر تعریف مقادیر آستانه[9] برای مدلسازی مفاهیم فازی و عدم قطعیت در مسئله است. تصمیم‌گیری چند خصیصه‌ای[10] روشی مؤثر برای رتبه‌بندی یا انتخاب یک یا چند گزینه از بین تعدادی محدود نسبت به چند معیار است (Lopez, 2005). در فرایند انتخاب دانشجو برای دریافت بورس تحصیلی[11] به‌عنوان یک مسئله تصمیم‌گیری چند خصیصه‌ای فرمول‌بندی شده است. واندام و همکاران (Vandamme, Meskens & Superby, 2007) از روشهای شبکهٔ عصبی و درخت تصمیم‌گیری برای پیش‌بینی عملکرد دانشجویان در اولین سال تحصیلی در دورهٔ کارشناسی استفاده کردند. گوردون و همکاران (Gordon et al., 2010) از روشهای آماری و بررسی رابطهٔ همبستگی بین عوامل سن، تجربهٔ کاری، جنسیت و میزان موفقیت تحصیلی آنان برای شناسایی عوامل مؤثر در موفقیت دانشجویان استفاده کرده‌اند.

---

۳. Discriminant Analysis
٤. Multi Criteria Analysis
٥. Ordered Weighted Averaging (OWA)
٦. Multi Expert-Multi Criteria Decision Making
۷. ELECTRE۳
۸. Outranking
۹. Threshold
۱۰. Multi Attribute Decision Making (MADM)
۱۱. Scholarship

تا کنون روشهای متنوعی برای ارزیابی نظام‌مند متقاضیان در تصمیم‌گیریها توسعه داده شده است. این روشها به مسائلی محدود می‌شود که چندین معیار ارزیابی وجود دارند و شامل برنامه‌ریزی آرمانی[12]، نظریهٔ کاربرد چند خصیصه‌ای[13]، مدلهای امتیازدهی[14]و فرایند تحلیل سلسله‌مراتبی[15]هستند (Matthew & Nydick, 1997). مدلهای امتیازدهی شاید قدیمی‌ترین و آشناترین دسته از مدلهای چند معیاره هستند. الزامات اولیهٔ این روش شامل ارائهٔ فهرستی از معیارهای ارزیابی، رسیدن به توافق کلی در باره وزنهای هر معیار و امتیازدهی هر متقاضی به هر یک از معیارهاست؛ سپس، امتیاز میانگین محاسبه و برای رتبه‌بندی و گزینش استفاده می‌شود. معمولاً این‌گونه مدلهای امتیازدهی یا از اعداد ترتیبی 1 تا 5 یا از مقیاس بازه‌ای برای ارزیابی متقاضیان در باره معیاری مشخص استفاده می‌کنند. هر چند جمع کلی این‌گونه امتیازهای ترتیبی و عددی برای تمام معیارها برای حصول به نتایج معناداری برای رتبه‌بندی متقاضیان امکان‌پذیر نیست. چنین محاسباتی فقط در صورتی امکان‌پذیر است که تمام مقیاسهای اندازه‌گیری مطلق باشند؛ لذا، به‌طور معمول بخشهای مسئول در پذیرش دانشگاههای دنیا [و از جمله ایران] از روش امتیازدهی ساده‌شده‌ای استفاده می‌کنند که در آن تمام مقیاسها عددی‌اند.

شایان ذکر است که استفاده از روشهای تصمیم‌گیری شناسایی مؤثرترین عوامل در حوزه‌های مختلف از جمله انتخاب دعوت‌شدگان برای مصاحبه کاری، پذیرش درخواست اقامت در کشورهای مختلف و سایر حوزه‌های مشابه نیز مطرح می‌شود که با توجه به متفاوت بودن زمینه این مقاله آنها بررسی نشده‌اند.

به‌رغم پژوهشهای بسیاری که در آنها رابطهٔ بین عوامل مختلف علمی و غیرعلمی و موفقیت دانشجویان و نیز رابطهٔ بین این عوامل و عوامل مورد ارزیابی در فرایند پذیرش بررسی شده‌اند، در بیشتر این پژوهشها مدلسازی مفاهیم مبهم مورد توجه بوده و ارزیابی میزان اهمیت عوامل چندان اهمیت نداشته است. ضمن اینکه در بیشتر این پژوهشها به بررسی رابطهٔ کلی بین عوامل و موفقیت دانشجویان پرداخته شده و به قواعد بین عوامل و پذیرش دانشجویان توجهی نشده است. این در حالی است که شناسایی مهم‌ترین عوامل کمک شایانی به تسریع روند بررسی داوطلبان می‌کند و در عین‌حال، به داوطلبان نیز امکان می‌دهد که با بهبود و افزایش نمرات خود در این دسته از عوامل، برنامه‌ریزی مناسب‌تری برای موفقیت در آزمون داشته باشند. در این مطالعه مهم‌ترین عوامل مؤثر در انتخاب و رتبه‌بندی دانشجویان با استفاده از نظریهٔ مجموعه‌های نادقیق با دقت متغیر[16] (Ziarko, 1993) که

---

۱۲. Goal Programming (GP)
۱۳. Multi Attribute Utility Theory (MAUT)
۱٤. Scoring Models
۱٥. Analytic Hierarchy Process (AHP)
۱٦. Variable Precision Rough Set Theory (VPRST)

توسعه‌ای از مدل اولیه مجموعه‌های نادقیق[17] (Pawlak, 1998) است، استخراج و ارزش هر کدام در فرایند تصمیم‌گیری مشخص شد. نظریهٔ مجموعه‌های نادقیق برای حل مسئله به اطلاعات اولیه (مثل درجهٔ عضویت در نظریهٔ فازی و احتمال اولیه در شبکهٔ بیز) نیاز ندارد و فقط با اطلاعات درونی و ذاتی هر یک از اقلام داده‌ای کار می‌کند. توانایی استخراج قواعد صریح و ساده به‌صورت«اگر ... آن‌گاه» و ارزیابی اهمیت و ارزش خصیصه‌ها از دیگر مشخصه‌های این نظریه است (Jensen & Shen, 2001). با شناسایی عوامل اثرگذار، قواعد تصمیم‌گیری بر مبنای این عوامل با تحلیل پایگاههای داده استخراج و از این قواعد برای پیش‌بینی نتایج سایر داده‌های مشابه استفاده می‌شود. در این مقاله از این روش برای شناسایی عوامل اثرگذار بر پذیرش داوطلبان در آزمون دکترای تخصصی استفاده و سپس، بر مبنای عوامل استخراج شده و رابطهٔ بین آنها و نتیجهٔ آزمون داوطلبان، قواعد پذیرش داوطلبان شناسایی و این قواعد در طراحی سامانهٔ پشتیبان تصمیم‌گیری استفاده شده است.

بدین منظور، ساختار مقاله بدین صورت تنظیم شده است که در بخش دوم به تبیین مسئلهٔ گزینش دانشجوی دورهٔ دکترای تخصصی در دانشگاههای ایران و برخی آیین‌نامه‌های مربوط پرداخته شده است. در بخش سوم خلاصه‌ای از مفاهیم نظریهٔ مجموعه‌های نادقیق و مجموعه‌های نادقیق با دقت متغیر بیان و چگونگی کاربرد این نظریه در مسئلهٔ گزینش دانشجو در بخش چهارم آورده شده است. بخش پنجم به ارزیابی روش پیشنهادی اختصاص دارد که با استفاده از مجموعهٔ اطلاعات آزمونهای دکترای تخصصی سالهای 1387 تا 1392 دانشگاه تربیت مدرّس انجام شده و در نهایت، نتیجه‌گیری ارائه شده است.

تعریف مسئله

در شکل‌گیری نحوه پذیرش دانشجویان دوره دکترای تخصصی در دانشگاهها دو دیدگاه اصلی وجود دارد: یکی دیدگاه استادان است که خواستار اختیار عمل بیشتر دانشگاه برای پذیرش دانشجویان هستند و دیگری دیدگاه دانشجویانی است که خواستار شفاف‌سازی روند پذیرش از طریق تأثیر بیشتر آزمون کتبی هستند؛ طبیعی است که نمی‌توان هیچ یک از این دو دیدگاه را مطلق دانست و باید یک روش ترکیبی در این خصوص اتخاذ شود. بدین منظور، آزمون دکترای تخصصی دانشگاههای ایران در دو مرحلهٔ آزمون کتبی و مصاحبه انجام می‌شود. آزمون کتبی از سال1390 به‌صورت سراسری است و سازمان سنجش آموزش کشور آن را برگزار می‌کند. پیش از سال 1390 هر دانشگاه آزمون مستقلی برگزار می‌کرد و دانشجو می‌پذیرفت. در آن زمان دانشگاهها با سلیقه‌ها و معیارهای متفاوت دانشجویان دوره دکتری را می‌پذیرفتند. یکی از ایرادات این روش آن بود که دانشجو حداکثر می‌توانست در آزمون چند دانشگاه شرکت کند که این امر شانس قبولی وی را کاهش می‌داد و وی فرصت قبولی زیادی را از

۱۷. Rough Set Theory

دست می‌داد. اما با روش جدید عدالت نسبی بیشتری برای داوطلبان وجود دارد. در این آزمون سه محور دروس تخصصی، آزمون زبان و آزمون استعداد تحصیلی در نظر گرفته می‌شود. درسها و مواد آزمون از مواردی است که متخصصان شورای عالی آن را برنامه‌ریزی می‌کنند. سیاست فعلی وزارت علوم، تحقیقات و فناوری آن است که سنجش به‌صورت متمرکز صورت گیرد و پذیرش دانشجوی دکتری به دانشگاهها واگذار شود؛ بدین معنا که یک غربال‌گری اولیه به شکل آزمون کتبی، افراد را برای مصاحبه جدا و آنها را به دانشگاهها هدایت می‌کند. این آزمون برای پذیرش دانشجوی دکتری کفایت نمی‌کند و یک سنجش ثانویه نیز برای اندازه‌گیری مشخصه‌های دیگر مانند میزان مهارت، انگیزه، نگرش و احساس مسئولیت دانشجو ضروری است که این سنجش فقط از طریق بررسی سوابق تحصیلی در دانشگاه امکان‌پذیر است. برای اینکه دانشجوی دوره دکتری بتواند در دوران تحصیل با فراغ بال و اختیار کامل با استاد همکاری کند و این همکاری به تولید علم منجر شود، لازم است بررسی سوابق تحصیلی (مصاحبه، سوابق آموزشی و پژوهشی) در دانشگاه و به دست استادان در پذیرش دانشجو تأثیرگذار باشد. بررسی سوابق تحصیلی داوطلب از سوی دانشگاه موجب می‌شود که استاد بتواند دانشجوی مورد نظر خود را با توجه به پارامترهای مشخص انتخاب کند. از طرف دیگر، برای آنکه داوطلبان احساس بی‌عدالتی نکنند و این تصور به وجود نیاید که استاد و گروه آموزشی یک دانشجو را از قبل برای ورود به دوره دکتری انتخاب کرده‌اند، یک آزمون اولیه نیز برگزار می‌شود. بررسی سوابق تحصیلی باید با در نظر گرفتن سوابق تحصیلی و پژوهشی فرد، مرتبط بودن با رشته و علاقه‌مند بودن داوطلب صورت گیرد.

افراد مجاز به انتخاب رشته با توجه به رشته انتخابی و نمره کسب شده در آزمون به دانشگاهها دعوت می‌شوند. در مرحله اول حدود چهار تا پنج برابر ظرفیت پذیرش، داوطلب مجاز به انتخاب رشته می‌شود و سپس، در مرحله بعدی حدود 2/5 برابر ظرفیت نیز برای بررسی سوابق تحصیلی به دانشگاه معرفی می‌شوند که این امر موجب می‌شود از حجم داوطلبان کاسته و بررسی سوابق تحصیلی از سوی دانشگاه برای دانشجویان و استادان آسان‌تر شود. در مرحله اول که داوطلبان مجاز به انتخاب رشته می‌شوند، می‌توانند با توجه به نمره خود انتخاب رشته کنند و سپس، سازمان سنجش با در نظر گرفتن سه معیار نمره آزمون و انتخابهای فرد، معدل کارشناسی و کارشناسی‌ارشد افراد مجاز را به دانشگاهها معرفی می‌کند. به همین دلیل اگر نمره یک داوطلب با رشته‌ای که انتخاب کرده است مطابقت نداشته باشد، برای بررسی سوابق تحصیلی به آن دانشگاه معرفی نخواهد شد. در آزمون دوره دکتری رشته مرتبط تعریف شده است. بنابراین، داوطلبان نمی‌توانند از هر رشته‌ای برای شرکت در آزمون رشته‌های دیگر اقدام کنند، مگر اینکه آن رشته‌ها بین رشته‌ای باشد. با در نظر گرفتن این توضیحات معدل کارشناسی و کارشناسی‌ارشد دانشجویان تراز می‌شود و در نمره کنکور با ضریب 20 درصدی اثرگذار خواهد شد.

در سال 1390 بررسی سوابق تحصیلی پس از معرفی دانشجو به دانشگاه محوریت صد درصدی در پذیرش دانشجو داشت، در سال 1391 وزن سوابق تحصیلی به دو سوم کاهش یافت و در سال 1392

وزن سوابق تحصیلی در پذیرش دانشجوی دکتری بـه 50 درصـد رسـید و در نهایـت، در سـال 1393 بـا بررسی روند آزمون و تحلیل موقعیت گذشته، بار دیگر وزن سـوابق تحصـیلی دانشـگاهها بـه 70 درصـد افزایش پیدا کرد.

علاوه بر پذیرش از طریق آزمون سراسری، در هر دانشگاه بخشی از ظرفیت هر یک از رشته‌ها نیز بـه دانش‌آموختگان استعدادهای درخشان اختصاص دارد. دانشگاهها هم طبق آیین‌نامه اجبار ندارنـد کـه هـر سال دانشجویان استعداد درخشان را بدون آزمون در دوره دکتری بپذیرند، بلکـه هـر دانشـگاه بـر اسـاس امکانات و ظرفیت هر سال تحصیلی خود تصمیم می‌گیرد که بـا در نظـر گـرفتن سـوابق تحصـیلی یـک دانشجو و وضعیت علمی وی، این دسته از دانشجویان را بپذیرد(www.isna.ir).

اشاره به این نکته ضروری است که معیارهای گزینش مـورد اسـتفاده در فراینـد پـذیرش در نظامهـای آموزش‌عالی با توجه به نوع رشته‌ها و نوع دانشـگاه متفـاوت اسـت. از ایـن‌رو، شناسـایی عـواملی کـه در پذیرش دانشجویان در تمام دانشگاههای کشور مد نظر باشد، دشوار به نظر می‌رسد. در این مطالعه سـعی شده است تا سامانه‌ای به‌منظور شناسایی این عوامل طراحی شود که با استفاده از داده‌هـای دانشـگاههای مختلف، عوامل مؤثر در پذیرش دانشجویان در هر دانشگاه با توجه به داده‌هـای موجـود همـان دانشـگاه شناسایی شوند. نتایج این سامانه می‌تواند بسته به نوع دانشگاه و نوع رشته متفاوت باشد؛ به‌عبارت دیگـر، در رشته‌های مختلف و در دانشگاههای مختلف عوامل متفاوتی بر پذیرش دانشجویان اثرگـذار اسـت کـه این سامانه با استفاده از داده‌های مربوط امکان شناسایی این عوامل را دارد. نکته مهم در این تحقیـق آن بود که سامانه طراحی شده کاملاً عمومی باشد و پارامترهای آن به شکل متغیـر تعریـف شـود تـا امکـان بهره‌گیری از آن در هر یک از دانشگاهها و با ضرایب مختلف پذیرش وجود داشته باشد.

نظریهٔ مجموعه‌های نادقیق

نظریهٔ مجموعه‌های نادقیق روشی ریاضی برای تحلیـل داده و داده‌کـاوی اسـت. ایـن نظریـه را پـاولاک (Pawlak, 1988) در سال 1982 میلادی برای استدلال از داده‌های ناقص و غیردقیق ارائه کرده اسـت و مبتنی بر مفاهیم تقریبهای بالا و پایین مجموعه‌هاست. نظریـهٔ مجموعـه‌هـای نـادقیق در حـوزه‌هـای یادگیری ماشینی (Fayyad & Piatetsky-Shapiro, 1996)، اکتسـاب دانـش (Leung, Wu & Zhang, 2006)، تحلیـل تصـمیم (Salvatore Greco, Matarazzo & Slowinski, 2001; Tuan-Fang, Duen-Ren & Gwo-Hshiung, 2006)، کشف دانش از پایگاه‌هـای داده (Sinh Hoa & Hung Son,1988)، سامانه‌های خبـره (Shusaku, 1988)، طراحـی سـامانه‌هـای تصـمیم یـار[18](Swiniarski, Hunt, Chalvet & Pearson, 1995)، داده کـاوی(Kusiak, 2001) و

۱۸. Decision Support Systems

استخراج الگو (Sinh Hoa & Hung Son, 1988) کاربردهای بسیاری دارد. مزیت اصلی این نظریه در مقایسه با سایر روشهای هوشمند مانند نظریه گواه[19] یا نظریه فازی[20] این است که هیچ‌گونه نیازی به اطلاعات اضافی و اولیه درباره داده ندارد (Pawlak, 1988).

با روش مجموعه نادقیق می‌توان دانش (مهم‌ترین و ارزشمندترین بخش اطلاعات) را از حداقل داده‌ها استخراج کرد، ارزش و اهمیت داده‌ها را با معیاری عددی ارزیابی کرد و قوعد ساده و قابل فهمی تولید ساخت که تفسیرهای سرراستی از نتایج به‌دست آمده ارائه دهد. در این روش با ساخت زیرمجموعه‌ای حداقل از خصیصه‌های مستقل، که فروکاست[21] نامیده می‌شود، با همان کیفیت دسته‌بندی جدول اطلاعاتی اصلی حجم جدول اطلاعاتی کاهش می‌یابد (Pawlak, 2002). در ادامه خلاصه‌ای از مفاهیم اصلی نظریهٔ مجموعه‌های نادقیق ارائه شده است[22].

سامانهٔ اطلاعاتی: سامانه (جدول) اطلاعاتی مهم‌ترین مفهوم در نظریهٔ مجموعهٔ نادقیق است. در سامانهٔ اطلاعاتی داده‌ها به کمک جدول بیان می‌شوند که در آن ستونها نشان‌دهندهٔ خصیصه و ردیفها نشان‌دهندهٔ اشیا هستند و درایه‌های جدول ارزش خصیصه‌ها را برای هر یک از اشیا بیان می‌کنند (Ponce & Karahoca, 2009). اشیا می‌توانند اشیای واقعی، حالتها، مفاهیم انتزاعی، لحظه‌های زمان و... باشند و خصیصه‌ها نشان‌دهندهٔ خصوصیتهایی از اشیا مثل وزن، رنگ و ... هستند؛ به عبارتی، هر ردیف نشان‌دهندهٔ اطلاعاتی دربارهٔ شیء آن ردیف است. چنین جدولی به‌عنوان سامانهٔ اطلاعاتی یا جدول داده شناخته می‌شود (Walczak & Massart, 1991; Yao, 1996). در جدولهای اطلاعاتی بین دو نوع خصیصه به نامهای شرط[23] و تصمیم تمایز قایل می‌شویم و چنین جدولی به نام «جدولِ تصمیم» شناخته می‌شود. ردیف‌های جدول تصمیم به‌صورت جملات «اگر...آن گاه» بیان می‌شوند که شرایط لازم برای مشخص کردن خروجی سامانه (مقدارخصیصهٔ تصمیم) را نشان می‌دهند. هر ردیفِ جدولِ تصمیم یک قاعدهٔ تصمیم را مشخص می‌کند. هنگامی که شرایط ذکرشده بر خصیصه‌های شرطی اعمال می‌شوند، تصمیمی گرفته یا عملی انجام می‌شود. سامانهٔ اطلاعاتی را با زوج مرتب $\boldsymbol{S = (U, A)}$ نشان می‌دهیم که U مجموعه مرجع اشیا و A مجموعه‌ای از خصیصه‌هاست. به هرخصیصه از A $\boldsymbol{(a \in A)}$، مجموعه V_a از ارزش‌ها، که دامنهٔ a نامیده می‌شود، اختصاص می‌یابد. در جدول 1 نمونه‌ای از جدول تصمیم نشان داده شده است که در آن هر ردیف سه اطلاعات دمای بدن، سردرد و حالت تهوع را برای بیماران نشان می‌دهد. ستون آخر جدول تصمیم دربارهٔ سرماخوردگی بیمار را نشان می‌دهد.

---

۱۹. Evidence Theory
۲۰. Fuzzy Theory
۲۱. Reduct

22. برای توضیحات بیشتر Hu, ۱۹۹۵ را ببینید.

۲۳. Condition

در این جدول U={4،3،2،1} مجموعهٔ بیماران است و{دمای بدن، سردرد، حالت تهوع}=A است. آخرین ستون جدول نیز نتیجه وضعیت بیمار؛ یعنی داشتن یا نداشتن سرماخوردگی را مشخص می‌کند.

جدول 1- نمونه‌ای از جدول تصمیم

بیمار	دمای بدن	سر درد	حالت تهوع	سرماخوردگی
1	بالا	دارد	ندارد	دارد
2	بسیار بالا	دارد	دارد	دارد
3	بالا	ندارد	ندارد	ندارد

رابطهٔ عدم تمایز: نقطهٔ آغازین نظریهٔ مجموعه‌های نادقیق رابطهٔ عدم تمایز[24] است که از اطلاعات مربوط به اشیای مورد بحث ایجاد می‌شود؛ رابطهٔ عدم تمایز روی مجموعهٔ $\boldsymbol{B \subseteq A}$ که با $\boldsymbol{I_B}$ نشان داده می‌شود، به‌صورت زیر تعریف می‌شود:

$$x\ I_B\ y \ \Leftrightarrow a(x) = a(y)\ \forall\ a \in B \qquad (1)$$

که در آن $a(x)$ مقدار خصیصهٔ a برای عنصر x را نشان می‌دهد. عناصری که رابطهٔ I_B با یکدیگر دارند، نسبت به خصیصه‌های مجموعهٔ B مشابه هستند (Walczak & Massart, 1991). خانوادهٔ همهٔ کلاسهای هم‌ارزی که از رابطهٔ I_B به‌دست می‌آیند، افراز مجموعه U توسط مجموعهٔ B را تولید می‌کنند که آن را با U/I_B (یا به‌صورت ساده‌تر با U/B) نشان می‌دهیم. کلاس هم‌ارزی عنصر $x \in B$ نسبت به رابطهٔ I_B را با $B(x)$، $[x]_B$ یا $[x]_{I_B}$ نشان می‌دهند. این مجموعه اعضایی از مجموعه مرجع را نشان می‌دهد که مقدار خصیصه‌های B آنها با مقدار این خصیصه‌ها برای متغیر x مساوی است.

روش مجموعه‌های نادقیق برای تحلیل داده‌ها منوط به دو مفهوم بنیادی به نام «تقریب بالا»[25] و «تقریب پایین»[26] یک مجموعه است. تقریب پایین X نسبت به B را با $B_*(X)$ یا $\underline{BX}$ نشان می‌دهیم. این تقریب شامل اشیایی است که به‌طور قطع متعلق به X هستند و به‌صورت زیر تعریف می‌شود:

$$B_*(X) = \{x \in U : B(x) \subseteq X\} \qquad (2)$$

تقریب بالای X نسبت به B را با $B^*(X)$ یا $\overline{BX}$ نشان می‌دهیم. تقریب بالا شامل اشیایی است که ممکن است متعلق به X باشند و به‌صورت زیر تعریف می‌شود:

---

۲٤. Indiscernibility
۲٥. Upper Approximation
۲٦. Lower Approximation

$$B^{*}(X)=\{x\in U : B(x)\cap X \neq f\} \quad (3)$$

ناحیهٔ مثبت مجموعهٔ X نسبت به مجموعه خصیصه‌های B نشان‌دهندهٔ اشیایی است که می‌توان با استفاده از دانش B با اطمینان گفت که به مجموعهٔ X تعلق دارند. این ناحیه با تقریب پایین X یکسان است، در واقع داریم:

$$POS_B(X)=B_{*}(X) \quad (4)$$

وابستگی و کاهش خصیصه‌ها: یکی از اهداف مهم تحلیل داده کشف وابستگی‌هاست. مجموعه D از خصیصه‌ها را وابسته به مجموعه C از خصیصه‌ها گوییم، هرگاه همهٔ مقادیر خصیصه‌ها از D به‌صورت یکتایی با مقادیر خصیصه‌های C مشخص شوند. این مفهوم را با $C \Rightarrow D$ نشان می‌دهند. در حالت کلی می‌توان وابستگی را به‌صورت زیر تعریف کرد:

فرض کنید C و D زیرمجموعه‌هایی از A (مجموعه خصیصه‌ها) باشند، می‌گوییم D با درجهٔ k $(0\le k\le 1)$ به C وابسته است و آن را با $C \Rightarrow_k D$ نشان می‌دهیم، هرگاه:

$$k=\gamma(C,D)=\frac{|Pos_C(D)|}{|U|} \quad (5)$$

که در آن $Pos_C(D) = \cup_{X\in U/D} C_{*}(X)$ ناحیهٔ مثبت افراز U/D نسبت به C است. این ناحیه، مجموعهٔ همه عناصری از U است که می‌توانند به‌صورت یکتا به بلوکهایی از افراز U/D با به‌کارگیری C طبقه‌بندی شوند. ضریب k نشان‌دهندهٔ تعداد اعضا نسبت به کل اعضای مجموعهٔ مرجع است که می‌توانند به طرز مناسبی با کمک خصیصه‌های شرطی به بلوکهایی از افراز U/D طبقه‌بندی شوند.

هدف از کاستن خصیصه‌ها در یک سامانهٔ اطلاعاتی به‌دست آوردن جدولی با خصیصه‌های کمتر است، به‌طوری که دانش موجود دربارهٔ اشیا (و به‌عبارتی توانایی کلاس‌بندی اعضا) تغییری نکند.

مدل مجموعهٔ نادقیق با دقت متغیر: با توجه به تعریف تقریبهای بالا و پایین، مدل اصلی مجموعه‌های نادقیق نسبت به داده‌های نویزدار حساس است (Hong, Wang & Wang, 2007). مدل مجموعهٔ نادقیق دقت متغیر توسیعی از نظریهٔ مجموعه نادقیق برای تحلیل و شناسایی الگوهای داده‌ای است که مقاومت بیشتری در برابر نویز دارد. این مدل را زیارکو (Ziarko, 1993) ارائه کرده است. در این مدل موضوعات با خطای کمتری از یک سطح از پیش تعیین شده طبقه‌بندی می‌شوند.

فرض کنید $X, Y \subseteq U$ خطای نسبی با رابطهٔ:

$$C(X,Y)=1-\frac{|X\cap Y|}{|X|} \quad (6)$$

تعریف می‌شود. شمولیت نادقیق با مجاز دانستن یک سطح نسبی از خطا در طبقه‌بندی از رابطهٔ 7 به‌دست می‌آید:

$$X \subseteq_{\beta} Y \ \ iff \quad C(X,Y)\le\beta \qquad 0\le\beta\le 0.5 \quad (7)$$

این مفهوم شمولیت نادقیق[27] خوانده می‌شود. با استفاده از $\subseteq_\beta$ به جای $\subseteq$، β- تقریب پایین و β - تقریب بالای مجموعه X نسبت به رابطهٔ هم‌ارزی B به ترتیب با رابطه‌های 8 و 9 تعریف می‌شوند (Ziarko,1993):

$$\underline{B}_\beta(X) = \{x|[x]_B \subseteq_\beta X\} = \left\{x\middle|\frac{|X \cap B(x)|}{|B(x)|} \geq 1-\beta\right\} \quad (8)$$

$$\overline{B}_\beta(X) = \left\{x\middle|\frac{|X \cap B(x)|}{|B(x)|} < 1-\beta\right\} \quad (9)$$

در این حالت ناحیهٔ مثبت برحسب β طبق رابطه 10 محاسبه می‌شود:

$$POS_{C,\beta}(D) = \bigcup_{X \in U/D} \underline{C}_\beta(X) \quad (10)$$

که در آن C و D مجموعه‌ای از خصیصه‌ها هستند. در این حالت ناحیهٔ مثبت مجموعه‌ای از اشیاست که تصمیم درباره مقدار خصیصه‌های D آنها با اطلاعات مجموعه خصیصه‌های C به‌گونه‌ای اتخاذ می‌شود که خطای تصمیم‌گیری بیشتر از β نباشد (Ziarko,1993). در مدل مجموعه‌های نادقیق با دقت متغیر تابع وابستگی عبارت است از:

$$\gamma_{C,\beta}(D) = \frac{|POS_{C,\beta}(D)|}{|U|} \quad (11)$$

محاسبهٔ فروکاستهای جدول تصمیم: با استفاده از نظریهٔ مجموعه‌های نادقیق می‌توان فروکاستهای یک جدول تصمیم را شناسایی و سپس، حداقل قواعد تصمیم‌گیری را با استفاده از خصیصه‌های موجود در فروکاست کمینه بیان کرد. فروکاست یکی از مهم‌ترین مفاهیم کاربردی مجموعه‌های نادقیق در داده‌کاوی است که مجموعه‌ای حداقل از خصیصه‌هاست که توان دسته‌بندی پایگاه دادهٔ اصلی را حفظ می‌کند (Pal & Skowron, 1999).

اگر D زیرمجموعه‌ای از خصیصه‌های تصمیم و C زیرمجموعه‌ای از خصیصه‌های شرایط باشد و درجهٔ وابستگی D به C برابر یک باشد، می‌توان چنین نتیجه گرفت که تصمیم در باره هر شیء را می‌توان با مجموعه ویژگیهای C مشخص کرد یا به عبارت دیگر، خصیصه‌های C برای تصمیم دربارهٔ نتیجهٔ هر شیء کافی هستند. با این مفهوم فروکاست، زیرمجموعه‌ای از خصیصه‌های شرایط مانند C خواهد بود که برای مجموعه خصیصه‌های تصمیم D در دو شرط زیر صدق می‌کند (Jensen, 2005):

1. $\gamma_C(D) = \gamma_R(D)$ که R مجموعه تمام خصیصه‌های شرایط است.
2. حذف هر خصیصه از C شرط (1) را تحت تأثیر قرار می‌دهد.

---

۲۷. Rough Inclusion

در مفهوم مجموعه‌های نادقیق، فروکاست با کمترین تعداد عضو اهمیت دارد که این فروکاست کمینه نامیده می‌شود. یافتن فروکاستهای کمینه و همچنین یافتن تمام فروکاستها مسئله‌ای NP است (Pal & Skowron, 1999). یک راه به‌دست آوردن فروکاست کمینه محاسبهٔ وابستگیهای تمام زیرمجموعه‌های ممکن R است. هر زیرمجموعه‌ای که برای آن داشته باشیم $\gamma(D) = 1$ یک فروکاست خواهد بود. کوچک‌ترین زیرمجموعه با این خاصیت فروکاست کمینه خواهد بود. الگوریتم فروکاست سریع[28]، محاسبهٔ فروکاست کمینه را بدون تولید تمام زیرمجموعه‌های ممکن امکان‌پذیر می‌سازد. این الگوریتم با مجموعه تهی $(C = \emptyset)$ شروع می‌شود و سپس، در رویکردی گام به گام خصیصه‌هایی را که افزودن آنها بیشترین افزایش در $\gamma_C(D)$ را سبب می‌شود، به مجموعه C اضافه می‌شود و این فرایند تا زمانی‌که بزرگ‌ترین مقدار ممکن برای $\gamma_C(D)$، معمولاً 1، به‌دست آید، ادامه می‌یابد (Jensen, 2005).

طراحی سامانهٔ شناسایی عوامل اثرگذار در پذیرش دانشجویان

همان‌گونه که اشاره شد، نظریهٔ مجموعه نادقیق و الگوریتم فروکاست سریع امکان شناسایی فروکاست کمینه از یک جدول یا سامانه اطلاعاتی را دارد. این مجموعه در واقع، خصیصه‌هایی را نشان می‌دهد که بیشترین اثر را در نتیجه دارند و به عبارت دیگر، نتیجه با استفاده از این خصیصه‌ها قابل پیش‌بینی است. با تعمیم مفاهیم این نظریه به فرایند پذیرش داوطلبان دوره دکتری می‌توان این‌گونه بیان کرد که هر دانشجو با مجموعه‌ای از خصیصه‌ها و ویژگیها که با نام سوابق تحصیلی شناخته می‌شود [مانند معدل کارشناسی، معدل کارشناسی‌ارشد، تعداد مقالات و ...] در آزمون سراسری دکترای تخصصی شرکت می‌کند، سپس، بر مبنای نمره آزمون کتبی در مصاحبه شرکت و نمرهٔ خاصی را نیز در مصاحبه کسب می‌کند. مجموع این نمرات وضعیت پذیرش یا عدم پذیرش دانشجو را مشخص می‌کند. نمرهٔ مصاحبه بر محورهای مختلف از جمله استعداد، قدرت بیان و تسلط بر روش علمی محاسبه می‌شود و در نمرهٔ سوابق علمی نیز مشخصه‌های مختلفی مانند معدل دوره‌های مختلف، تعداد مقالات، تعداد کتابهای تألیفی یا ترجمه شده و پروژه‌های انجام شده ارزیابی می‌شود. بدین ترتیب، می‌توان جدول اطلاعاتی از نمرات دانشجویان را تشکیل داد که برای هر دانشجو شامل نمرهٔ آزمون کتبی، نمرهٔ مصاحبه (زیرشاخصهای ارزیابی شده در مصاحبه) و نمرهٔ سوابق تحصیلی (زیرشاخصهای ارزیابی سوابق تحصیلی) باشد. با تحلیل این جدول اطلاعات می‌توان شاخصهای اصلی اثرگذار در پذیرش داوطلبان را شناسایی و بر اساس مقادیر مختلف این شاخصها قواعد پذیرش داوطلبان را استخراج کرد.

---

۲۸. Quick Reduct Algorithm

به‌منظور شناسایی عوامل مؤثر و استخراج مجموعه قواعد پذیرش داوطلبان، سامانه‌ای تحت وب طراحی و پیاده‌سازی شده است که در شکل 1 رابط کاربری آن نشان داده شده است..

شکل 1- رابط کاربری سامانه شناسایی عوامل مؤثر در پذیرش دانشجویان

ورودیها و خروجیهای سامانه عبارت‌اند از:

الف- ورودیهای سامانه

1. فایل اکسل ورودی: مجموعه داده‌ای آموزش، که با هدف شناسایی عوامل مؤثر در پذیرش دانشجویان استفاده می‌شود، در قالب فایل اکسل در سامانه وارد می‌شود. سطر اول این فایل نام متغیرها یا عواملی است که در باره هر دانشجو ارزیابی می‌شود و سایر سطرها اطلاعات مربوط به هر داوطلب را نشان می‌دهد.

2. دقت مورد انتظار در قواعد استخراج شده از داده‌ها: این ورودی معادل پارامتر β در الگوریتم فروکاست سریع است و نسبتی از داده‌ها را نشان می‌دهد که در شرایط قاعدهٔ استخراج شده صدق می‌کنند و نتیجه آنها مشابه نتیجه قاعده استخراج شده است؛ به عبارت دیگر، دقت مورد انتظار در قواعد استخراج شده از داده‌ها، نسبت داده‌هایی را نشان می‌دهد که قاعده استخراج شده در باره آنها صدق می‌کند. این پارامتر مقداری بین صفر و یک اختیار می‌کند.

3. حداقل پوشش داده‌ها: این پارامتر درصد داده‌هایی را نشان می‌دهد که تصمیم‌گیری (پیش‌بینی نتیجه) در باره آنها بر مبنای عوامل اثرگذار شناسایی شده (فروکاست کمینه) امکان‌پذیر است. برای مثال، انتخاب عدد 0/9 برای پارامتر حداقل پوشش داده‌ها بدان معناست که تصمیم‌گیری در باره نتیجه 90% داوطلبان آتی بر مبنای قواعد استخراج شده از داده‌ها امکان‌پذیر خواهد بود. این پارامتر نیز مقداری بین صفر و یک اختیار می‌کند.

ب- خروجیهای سامانه

خروجی سامانه شامل تکرارهای مختلف الگوریتم فروکاست سریع و شناسایی عوامل مؤثر در پذیرش داوطلبان است و از جدول قواعد استخراج شده است.

1. عوامل مؤثر در پذیرش داوطلبان: در پایان اجرای الگوریتم، عوامل مؤثر در پذیرش داوطلبان بر مبنای دقت مورد نظر استخراج و در صفحه اول خروجی نشان داده می‌شوند. گامهای مختلف الگوریتم شامل عوامل انتخاب شده در هر گام و سطح پوشش داده‌ها بر مبنای مشخصه‌های شناسایی شده در این گام نیز به تفکیک مراحل اجرای الگوریتم در این خروجی قابل مشاهده است. در شکل 2 نمونه‌ای از این خروجی نشان داده شده است.

نتیجه گام 1

K:V1 = 0.0
K:V10 = 0.16393442622950818
K:V2 = 0.0
K:V11 = 0.1885245901639344
K:V3 = 0.0
K:V12 = 0.1721311475409836
K:V4 = 0.0
K:V5 = 0.0
K:V6 = 0.0
K:V7 = 0.0
K:V8 = 0.0
K:V9 = 0.16666666666666666

متغیرهای انتخاب شده به ترتیب اهمیت	استعداد و قدرت بیان
درصدی از جامعه مورد بررسی که می توان تنها با متغیر انتخاب شده، نتیجه را پیش بینی نمود	18.85

شکل 2- خروجی سامانه

2. جدول قواعد استخراج شده: این خروجی در قالب فایل اکسل قابل ذخیره‌سازی و مشاهده است. هر سطر جدول استخراج شده شامل یک قاعده است که در آن مقادیر متناظر با مشخصه‌های ورودی و نتیجه پیش‌بینی شده بر مبنای آن مقادیر بیان شده‌اند.

نتایج

به‌منظور ارزیابی عملکرد سامانهٔ طراحی شده در شناسایی عوامل مؤثر در پذیرش داوطلبان دورهٔ دکترای تخصصی، از اطلاعات آزمونهای سالهای 1387 تا 1390 دانشگاه تربیت مدرّس استفاده شده است. این دانشگاه تنها دانشگاه جامع تحصیلات تکمیلی در کشور ایران است. جامعیت این دانشگاه در تقریباً همهٔ رشته‌های تحصیلی، قرار گرفتن این دانشگاه در پایتخت، جایگاه مناسب پژوهشی دانشگاه در سطح بین‌المللی و نیز بهره‌مندی از امکانات تحقیقاتی و آزمایشگاهی مناسب سبب شده است تا تقاضای زیادی برای ورود به دوره‌های دکتری و کارشناسی‌ارشد این دانشگاه وجود داشته باشد. امتیاز نهایی داوطلبان ورود به دورهٔ دکترای تخصصی این دانشگاه [مانند سایر دانشگاههای کشور] بر مبنای سه امتیاز آزمون کتبی (آزمون سازمان سنجش)، امتیاز مصاحبه و امتیاز سوابق تحصیلی محاسبه می‌شود. وزن امتیازهای این سه بخش تا قبل از سال 1390 به‌صورت آزمون کتبی (T) 60 امتیاز، مصاحبه (I) 30 امتیاز و سوابق تحصیلی (H) 10 امتیاز محاسبه می‌شد. در سال 1390 بنا به آیین‌نامه وزارت علوم، تحقیقات و فناوری، امتیاز آزمون کتبی در نظر گرفته نشده و تمام امتیاز به نمرهٔ مصاحبه اختصاص داده شده است. این نسبت برای سالهای 1391 و 1392 به نسبت 50-50 و برای سال 1393 نسبت 30-70 بوده است. از آنجا که سال 1390 تنها سالی است که نمرهٔ آزمون کتبی در آن در نظر گرفته نشده، اطلاعات این سال از مجموعهٔ داده‌ای حذف شده است. در این روند از بین داوطلبانی که برمبنای امتیاز آزمون کتبی سراسری به مصاحبه راه یافته‌اند، مصاحبهٔ تخصصی به‌عمل می‌آید. در این مرحله از داوطلبان خواسته می‌شود تا در جلسهٔ مصاحبه تمام اطلاعات مرتبط با فعالیتهای آموزشی و پژوهشی خود را همراه داشته باشند. هیچ شیوهٔ رسمی مبنی بر اینکه چه اطلاعاتی ضروری و چه اطلاعاتی زاید است، وجود ندارد و به همین دلیل، دانشجویان اطلاعات بی‌شمار و تأییدیه‌های بسیاری را به همراه دارند و در جلسه مطرح می‌کنند. ضمن اینکه تعداد زیاد پذیرفته‌شدگان آزمون کتبی و فرصت کم مصاحبهٔ آنها سبب می‌شود که بعضاً در یک روز با بیش از 50 داوطلب مصاحبه شود. نامشخص بودن معیارهای اصلی پذیرش، ذهنی بودن ارزیابیها و خطاهای انسانی (ناشی از خستگی مصاحبه‌ها) همگی سبب می‌شود که فرایند مصاحبه یکی از مراحل دشوار و در عین حال، مبهم فعالیتهای دانشگاهی باشد. از این رو، در دانشگاه تربیت مدرس تلاش شده است که تا حدّ ممکن عوامل مؤثر بر پذیرش دانشجو احصا و امتیازبندی شود. در جدول 2 معیارهای تصمیم‌گیری دربارهٔ پذیرش داوطلبان دورهٔ دکترای تخصصی در این دانشگاه نشان داده شده است. امتیازات این شاخصها بر مبنای سالهای پیش از 1390 بیان شده است. در پایگاه دادهٔ نهایی با توجه به هنجارسازی داده‌های هر سال به‌طور مجزا، اثر مقدار این امتیازات از بین رفته است.

پس از مصاحبه با داوطلبان و با محاسبهٔ امتیاز آنان، متقاضیان بر اساس امتیاز کسب شده رتبه‌بندی و سپس، بر اساس میزان ظرفیت گروههای آموزشی پذیرفته می‌شوند. در جدول 3 بخشی از پایگاه داده آورده شده که حاوی اطلاعات در باره رشته‌های مختلف دانشکده فنی و مهندسی دانشگاه تربیت مدرّس در آزمونهای سالهای 1387 تا 1391 (بجز سال 1390) است. این پایگاه داده شامل اطلاعات 604

داوطلب است. در این جدول AC نشان دهنده پذیرفته شدن داوطلب و RE نشان دهنـده عـدم پـذیرش وی است.

جدول 2- معیارهای اصلی در فرایند گزینش دانشجو

معیار اصلی	زیرمعیارها	شاخص	امتیاز	امتیاز کل
آزمون کتبی	-	T	60	60
مصاحبه	تجزیه و تحلیل علمی	I_1	10	30
	تسلط بر روش علمی	I_2	10	
	استعداد و قدرت بیان	I_3	10	
سوابق تحصیلی	معدل دیپلم	h_1	0/25	10
	معدل کارشناسی	h_2	0/5	
	معدل کارشناسی‌ارشد	h_3	0/75	
	فعالیت آموزشی	h_4	0/5	
	تألیف کتاب	h_5	3	
	کار تحقیقی	h_6	1/5	
	ترجمه کتاب	h_7	2	
	تألیف مقاله	h_8	1/5	

همچنان‌که در جدول 3 ملاحظه می‌شـود، در برخـی از سـالها و برخـی رشـته‌هـا فقـط چنـد شـاخص نمره‌دهی شده و مابقی بی‌ارزش انگاشته شده‌اند و امتیاز آنها در سایر شاخصها سرشکن شـده اسـت. ایـن امر موجب شده است که در برخی موارد نمره یک شاخص بیش از سقف مذکور در جدول 1 منظـور شـده باشد. برای همگن‌سازی سامانه، این‌گونـه امتیازهـا هنجـار شـده اسـت. بایـد در نظـر داشـت کـه عمـل هنجارسازی گرچه سبب همگن‌شدن سامانه می‌شود، ولی بر میزان ابهام و نقص دانش موجود در سـامانه می‌افزاید.

پس از تکمیل پایگاه داده، هدف یـافتن عـواملی از شاخصـهای معرفـی شـده در جـدول 2 اسـت کـه بیشترین اثر را در پذیرش داوطلب دارند؛ به‌عبارت دیگر، هدف شناسایی عواملی است کـه بـرای پـذیرش داوطلب کافی هستند و نیازی به بررسی سایر عوامل نخواهد بود. برای این منظور و شناسایی این عوامـل از سامانهٔ طراحی شده در جدول 4 استفاده شده است.

برای شناسایی اثرگذارترین ویژگیها در انتخاب دانشجویان دکترای تخصصی ابتدا الگـوریتم فروکاسـت سریع روی مجموعه داده‌های مذکور در جدول 2 بیان می‌شود. برای اجرای الگوریتم فروکاست سـریع در اولین گام لازم است که مجموعه داده‌ای گسسته سازی شود. برای ایـن منظـور میـانگین (μ) و انحـراف معیار (σ) امتیاز هر یک از شاخصها در سالهای مختلف محاسبه و سپس، امتیاز داوطلبـان در هـر سـال و هر رشته بر این مبنا گسسته‌سازی شده است. بدین ترتیب که چنانچـه امتیـاز داوطلـب کمتـر از $\mu - \sigma$

باشد، مقدار 1- ، بین $\mu - \sigma$ و $\mu + \sigma$ مقدار 0 و برای امتیازهای بـالاتر از $\mu + \sigma$ نمـره 1+ در نظـر گرفته می‌شود. در جدول 4 بخشی از داده‌های گسسته‌سازی شده نشان داده شده است.

جدول 3- بخشی از پایگاه دادهٔ داوطلبان دورهٔ دکترای تخصصی

شماره داوطلب	h_1	h_2	h_3	h_4	h_5	h_6	h_7	h_8	I_1	I_2	I_3	T	نتیجه
1	0	3/50	2/75	0	0	0	0	0	7/775	3/5	4/5	36/96	AC
2	0	4	2/3	0/525	0	0	0	0	13	4/1	4	29/32	AC
3	0	1	1/5	0/909	0	0	0	0/5	9	4	4/045	24/17	RE
4	0	2	2/4	0/5	0	0	0	0	8	4	2/1	18/99	RE
5	0	1/25	1/25	0	0	0	0	0/625	8	4	0/75	18/25	RE
6	0	2/25	2/5	0	0	0	0	0/563	10	3	2/5	16/85	RE
7	0	3	1/375	0	0	0	0	0	7	4/1	5	16/6	RE
8	0	0/5	1	0	0	0	0	0	4/67	4/83	5/83	27/12	RE
9	0	0/5	1	0	0	0	0	1	3/92	3/08	4	28/32	RE
10	0	0/5	1	0	0	1	1	0	9/3	9/16	9/16	37/02	AC
11	0	0/5	1	0	0	0	0/5	2	8/83	8/83	8/58	33/84	AC
12	0	0/5	2	0	0	0	0	3	6/17	6/333	6/5	25/98	RE
13	0	0/5	1	0	0	0	0	2	3/83	3/83	4/16	27/36	RE
14	0	0/5	1	0/5	0	0	0	3	6/58	6/33	6/33	29/04	RE
15	0	0/5	2	0	0	0	0	1	9	9	9	25/92	AC
16	0	0/5	1	0	0	0	0	3	4/41	3/91	3	24	RE
17	0/6	1/6	2/1	0/6	0	0	0	0	10	10	10	35/25	AC
18	0/8	1/2	1/8	0/3	0	0	0	0	3	2	3	35/06	RE
19	0/5	1	1/8	0/6	0	0	0	0	2	2	3	35/06	RE
20	0/8	1/2	2/4	0/6	1/5	0	0	0	10	10	10	32/5	AC
21	0/7	1/2	1/8	0/3	0	0	0	0	10	10	10	32/06	AC
22	0/5	0/75	0/75	1	0	1/5	0	1/5	9	8	9	35/4	AC
23	0/5	0/75	1/5	0/25	0	1/5	0	1/5	9	8	8	30/5	AC
24	0/5	1	0/75	0/25	0	1/5	0	1/5	9	9	8	44/4	AC

RE	29/6	5	4	3	0/5	0	0/25	1/5	1	0/25	0/25	0/5	25
AC	36/9	8	9	9	1/5	0	1/5	0	0/25	0/75	0/75	0/5	26
RE	30/3	6	4	4	0/75	0	1/5	0	0/75	0/75	0/75	0/5	27
AC	27/89	10	10	10	5	0	0	0	0	1/42	0/94	0/7	28
RE	31/13	6	6	6	3	0	0	0	0	0/93	0/74	0/84	29
RE	35/78	4	4	4	3	0	0	0	0	0/65	0	0/4	30

پس از گسسته‌سازی پایگاه داده، این مجموعه در قالب فایل اکسل به‌عنوان ورودی سامانه استفاده و مقدار β برابر با 0/2 و حداقل پوشش داده‌ها 0/9 انتخاب شد. شایان ذکر است که در این پایگاه شاخص نتیجه داوطلب به‌عنوان خصیصهٔ تصمیم در نظر گرفته شد؛ یعنی {نتیجه}=D و 12 شاخص دیگر خصیصه‌های شرط هستند. در جدول 5 نتایج تکرارهای مختلف این الگوریتم نشان داده شده است. همان‌طور که در جدول 5 مشخص است، در اولین گام، شاخص I_3 که استعداد و قدرت بیان است، به‌عنوان معنادارترین شاخص مشخص شده است؛ معناداری این شاخص 0/1989 بود و این بدان معناست که با استفاده از فقط این شاخص می‌توان درباره وضعیت 19/89% داوطلبان تصمیم گرفت و این تصمیم در حداقل 0/8 (1-β) موارد و به‌عبارت دیگر، 80% موارد صحیح خواهد بود. ادامهٔ روند الگوریتم فروکاست $C = \{h_8, I_3, T\}$؛ یعنی امتیاز آزمون کتبی، امتیاز مقاله و استعداد و قدرت بیان را به‌عنوان فروکاست کمینه معرفی می‌کند.

جدول 4- قسمتی از پایگاه داده اصلی گسسته‌سازی شده

شماره داوطلب	h_1	h_2	h_3	h_4	h_5	h_6	h_7	h_8	I_1	I_2	I_3	T	نتیجه
1	0	0	0	+1	+1	+1	+1	0	0	0	0	+1	RE
2	0	-1	0	+1	+1	+1	+1	-1	0	0	0	0	RE
3	+1	0	+1	+1	+1	+1	+1	0	+1	+1	+1	0	AC
4	+1	0	+1	+1	+1	+1	+1	0	+1	+1	+1	-1	AC
5	0	-1	0	+1	+1	+1	+1	0	-1	-1	0	0	RE
6	0	0	0	+1	+1	+1	+1	0	0	0	0	0	RE
7	+1	0	0	+1	+1	+1	+1	0	0	0	0	-1	RE
8	+1	0	0	0	0	0	0	+1	0	0	0	0	RE
9	+1	0	0	0	0	0	0	+1	0	0	0	0	RE
10	+1	0	+1	0	0	0	0	0	+1	+1	+1	0	AC

RE	0	0	0	0	0	0	0	0	0	+1	0	+1	11
AC	-1	+1	+1	+1	0	0	0	0	0	0	0	+1	12
RE	0	0	0	0	0	0	0	0	0	0	0	+1	13
RE	-1	0	0	0	0	+1	+1	+1	+1	0	-1	+1	14
RE	0	0	0	0	0	0	0	0	0	+1	0	+1	15
RE	-1	0	0	0	0	0	0	0	0	+1	-1	+1	16
AC	0	+1	+1	+1	+1	0	0	0	0	0	0	+1	17
RE	+1	0	0	0	0	0	0	0	0	0	0	+1	18
RE	0	0	0	0	0	+1	+1	+1	+1	+1	0	-1	19
AC	0	0	0	0	0	+1	+1	+1	+1	0	0	0	20
AC	+1	+1	+1	+1	0	+1	+1	+1	+1	0	+1	+1	21
RE	+1	0	0	0	0	+1	+1	+1	+1	-1	+1	0	22

جدول 5- تکرارهای الگوریتم فروکاست سریع

تکرار اول												
فروکاست (C)	h_1	h_2	h_3	h_4	h_5	h_6	h_7	h_8	I_1	I_2	I_3	T
$\gamma_C(D)$	0	0/1094	0/0832	0/1445	0	0/1060	0	0/1891	0/1985	0/1821	0/1989	0/1261

تکرار دوم										
فروکاست (C)	$\{h_1,I_3\}$	$\{h_2\ I_3\}$	$\{h_3,\ I_3\}$	$\{h_4\ I_3\}$	$\{h_5,\ I_3\}$	$\{h_6,\ I_3\}$	$\{h_7,\ I_3\}$	$\{h_8,I_3\}$	$\{I_1,\ I_3\}$	$\{I_2,\ I_3\}$	$\{T,\ I_3\}$
$\gamma_C(D)$	0/1011	0/1410	0/1745	0/1636	0/1989	0/1978	0/1017	0/3887	0/3515	0/3787	0/2760

تکرار سوم										
فروکاست (C)	$\{h_8,I_3,h_1\}$	$\{h_8,I_3,\ h_2\ \}$	$\{h_8,I_3,h_3\}$	$\{h_8,I_3,h_4\}$	$\{h_8,I_3,h_5\}$	$\{h_8,I_3,h_6\}$	$\{h_8,I_3,h_7\}$	$\{h_8,I_3,I_1\ \}$	$\{h_8,I_3,I_2\ \}$	$\{h_8,I_3,T\}$
$\gamma_C(D)$	0/5112	0/5602	0/5234	0/5610	0/6102	0/6023	0/5176	0/7491	0/7946	0/93

با استفاده از خصیصه‌های به‌دست آمده در گام پیشین، مجموعه حداقلی از قواعد تصمیم‌گیری از پایگاه داده استخراج می‌شود. در جدول 6 بخشی از تصمیم‌های استخراج شده آورده شده است. در این قواعد مانند روش گسسته‌سازی 1- به معنای نمره کمتر از $\mu - \sigma$، 0 نمره بین $\mu - \sigma$ و $\mu + \sigma$ و نمره 1+ به معنای نمره بیشتر از $\mu + \sigma$ است.

جدول 6- بخشی از قواعد تصمیم‌گیری

ردیف	قواعد
1	اگر نمره مقاله 1+، نمره استعداد و قدرت بیان 1+ و نمره آزمون کتبی 1+ باشد، داوطلب پذیرفته می‌شود.
2	اگر نمره مقاله 1+، نمره استعداد و قدرت بیان 1+ و نمره آزمون کتبی 1- باشد، داوطلب پذیرفته نمی‌شود.
3	اگر نمره مقاله 1-، نمره استعداد و قدرت بیان 1+ و نمره آزمون کتبی 1- باشد، داوطلب پذیرفته نمی‌شود.
4	اگر نمره مقاله 1-، نمره استعداد و قدرت بیان 1- و نمره آزمون کتبی 1+ باشد، داوطلب پذیرفته نمی‌شود.
5	اگر نمره مقاله 1+، نمره استعداد و قدرت بیان 1- و نمره آزمون کتبی 1- باشد، داوطلب پذیرفته نمی‌شود.
6	اگر نمره مقاله 1+، نمره استعداد و قدرت بیان 0 و نمره آزمون کتبی 1+ باشد، داوطلب پذیرفته می‌شود.
7	اگر نمره مقاله 0، نمره استعداد و قدرت بیان 1- و نمره آزمون کتبی 1+ باشد، داوطلب پذیرفته می‌شود.
8	اگر نمره مقاله 0، نمره استعداد و قدرت بیان 1+ و نمره آزمون کتبی 0 باشد، داوطلب پذیرفته نمی‌شود.
9	اگر نمره مقاله 1-، نمره استعداد و قدرت بیان 1- و نمره آزمون کتبی 1- باشد، داوطلب پذیرفته نمی‌شود.
10	اگر نمره مقاله 0، نمره استعداد و قدرت بیان 0 و نمره آزمون کتبی 1+ باشد، داوطلب پذیرفته می‌شود.

برای ارزیابی مجموعه قواعد پذیرش داوطلبان، از آنها برای پیش‌بینی نتایج آزمون سالهای 1392 و 1393 استفاده و نتیجهٔ الگوریتم در پذیرش یا عدم پذیرش داوطلبان با نتیجهٔ واقعی مقایسه شد. در این ارزیابی دقت قواعد تصمیم‌گیری حدود 71% برای سال 1392 و 68% برای سال 1393 بود. مقایسهٔ نتایج به‌دست آمده از مجموعهٔ قواعد و نتایج واقعی نشان می‌دهد که خطای تصمیم‌گیری بیشتر در خصوص داوطلبانی رخ داده است که امتیازی در سطح متوسط (بین $\mu - \sigma$ و $\mu + \sigma$) داشته‌اند. از این‌رو، برای این دسته از داوطلبان شاخص دیگری نیز باید برای تصمیم‌گیری مدنظر قرار گیرد. نتایج تکرار اول الگوریتم فروکاست سریع در جدول 5 نشان می‌دهد که شاخص I_1؛ یعنی قدرت تجزیه و تحلیل علمی دومین شاخص معنادار است. از این رو، برای تصمیم‌گیری در باره داوطلبانی که امتیاز آنها در سطح متوسط است، این شاخص نیز به مجموعه قواعد تصمیم‌گیری اضافه می‌شود. دقت تصمیم‌گیری با اضافه شدن این شاخص در جدول 7 بیان شده است.

جدول 7- ارزیابی دقت سامانه با استفاده از نتایج آزمون سال 1392 و 1393

		نتایج سامانه	
		پذیرش	رد
نتیجه واقعی سال 1392	پذیرش	%34	%6
	رد	%4	%56
نتیجه واقعی سال 1393	پذیرش	%36	%5
	رد	%10	%49

همان‌گونه که در جدول 7 مشخص است، در مجموع تطابق نتیجه سامانه در پذیرش و رد داوطلبان حدود 90% در سال 1392 و 86% در سال 1393 است. کاهش دقت سامانه در نتایج سال 1393 به دلیل کاهش ضریب آزمون کتبی در این سال است. از آنجا که این متغیر به‌عنوان عامل مؤثر در پذیرش دانشجویان شناسایی شده است، کاهش ضریب آن به کاهش دقت سامانه منجر شده است.

نتیجه‌گیری

فرایند انتخاب داوطلبان آزمون دکترای تخصصی مستلزم صرف هزینه و زمان زیادی است و شاخص‌ها و متغیرهای بسیاری در آن تأثیر دارند. این شاخصها بسیار ابهام دارند، ضمن اینکه تعداد زیادی از آنها در پذیرش داوطلبان اهمیت چندانی ندارد و به ارزیابی آنها نیازی نیست. در این پژوهش از نظریهٔ مجموعه‌های نادقیق و توسیع آن؛ یعنی مجموعه‌های نادقیق با دقت متغیر برای طراحی سامانه‌ای به‌منظور شناسایی مؤثرترین عوامل در پذیرش داوطلبان استفاده شده است. با کاربرد این سامانه روی اطلاعات آزمونهای سالهای 1387 تا 1391 دانشکده فنی و مهندسی دانشگاه تربیت مدرس، شاخص‌های «قدرت و استعداد بیان»، «تألیف مقاله» و «امتیاز آزمون کتبی» به‌عنوان مؤثرترین عوامل در پذیرش داوطلبان در این دانشگاه شناسایی شدند. پس از شناسایی این شاخصها، با بررسی پایگاه داده‌ای مسئله، قواعد تصمیم‌گیری بر مبنای این عوامل بیان شدند و از آنها برای پیش‌بینی نتایج داوطلبان آزمون سال 1392 استفاده شد. دقت این مجموعه از قواعد حدود 71% بود. از آنجا که بیشترین خطا در پیش‌بینی نتیجه در خصوص داوطلبانی رخ دارد که امتیاز آنها در سطح متوسط بود، برای این دسته از داوطلبان شاخص «قدرت تجزیه و تحلیل علمی» نیز برای تصمیم‌گیری استفاده و مجموعه قواعد دیگری برای این دسته از داوطلبان تعریف شد. با تصمیم‌گیری بر اساس این مجموعه از قواعد دقت تصمیم‌گیری حدود 91% می‌رسد.

مجموعهٔ این قواعد نشان می‌دهد در شرایطی که وضعیت داوطلبان از نظر تعداد مقالات مشابه بوده (قواعد 1 و 2)، داوطلبانی که امتیاز آزمون کتبی آنان بهتر بوده است، شانس بیشتری برای پذیرش در دورهٔ دکترای تخصصی داشته‌اند. قواعد 3 و 4 سامانه نشان می‌دهد زمانی‌که داوطلبان امتیاز مقاله را دریافت نکرده‌اند، حتی در شرایطی که امتیاز آزمون کتبی آنان بالا بوده است و امتیاز خوبی نیز در مصاحبه به دست آورده‌اند، شانس پذیرش آنان در دورهٔ دکترای تخصصی کم بوده است. مجموعهٔ قواعد نیز نشان می‌دهد داوطلبانی که فقط در یکی از محورها نمرهٔ مناسبی داشته (ردهٔ 1+ در یکی از شاخصهای آزمون کتبی، مقاله و استعداد و قدرت بیان) و در دو شاخص دیگر نمرهٔ پایینی کسب کرده‌اند (ردهٔ 1- در دو شاخص)، شانس چندانی برای پذیرش نداشته‌اند. در نقطهٔ مقابل داوطلبانی که توانسته‌اند در دو شاخص از سه شاخص شناسایی شده نمرهٔ مطلوبی را کسب کنند (ردهٔ 1+ در دو شاخص)، از شانس بالایی برای پذیرش برخوردار بودند.

در روند اجرای الگوریتم فروکاست سریع و استخراج قواعد تصمیم‌گیری، این نکته نیز قابل توجه است که شاخص «استعداد و قدرت بیان» معنادارترین شاخص شناسایی شده است و با استفاده از آن می‌توان دربارهٔ حدود 20% از داوطلبان تصمیم‌گیری کرد. ضمن اینکه با استفاده از این شاخص و همچنین «تعداد مقالات» دقت تصمیم‌گیری دربارهٔ پذیرش یا عدم پذیرش داوطلبان حدود 39% خواهد بود و با اضافه شدن شاخصهای «امتیاز آزمون کتبی» و «قدرت تجزیه و تحلیل علمی» دقت تصمیم‌گیری حدود 91% افزایش خواهد یافت. بدین ترتیب، با استفاده از چهار شاخص «استعداد و قدرت بیان»، «تعداد مقالات»، «امتیاز آزمون کتبی» و «قدرت تجزیه و تحلیل علمی» به جای استفاده از 12 شاخص معرفی شده در جدول 2، می‌توان دربارهٔ نتیجهٔ آزمون 91% از داوطلبان تصمیم‌گیری کرد.

سامانه طراحی شده امکان شناسایی عوامل مختلف مؤثر در پذیرش داوطلبان در دانشگاههای مختلف را بر مبنای داده‌های ورودی و با توجه به معیارهای پذیرش هر دانشگاه فراهم می‌کند.

References

1. Ali, P. (2008). Admission criteria and subsequent academic performance of general nursing diploma students. *Journal of the Pakistan Medical Association*, 58(3), 128-32.

۲. Carlsson, Ch., & Fuller, R. (1997). *OWA operators for doctoral student selection problem*. Kluwer Academic Publisher, 167-178.

3. Fayyad, U., & Piatetsky-Shapiro, G. (1996). *Advances in knowledge discovery and data mining*. MIT/AAAI Press.

4. Gordon, C., Williams, S., Hudson, G., & Stewart, J. (2010). Factors associated with academic performance of physical therapy students. *West Indian Medical Journal*, 59(2), 203-208.

5. Graham, L.D. (1991). Predicating academic success of students in a master of business administration program. *Educational Psychology Measure*, 4, 721-727.

6. Hong, T., Wang, T., & Wang, S. (2007). Minimizing fuzzy β-certain and β-possible rules from quantitative data based on the variable precision rough set model. *Expert Systems with Applications*, 32, 223-232.

7. Hu, X. (1995). Knowledge discovery in databases: An attribute oriented rough set approach. (Doctoral dissertation). Regina University.

8. Jensen, R., & Shen, Q. (2001). A rough set-aided system for sorting WWW bookmarks. *Web Intelligence*, 95-105.

9. Jensen, R. (2005). Combining rough and fuzzy sets for feature selection. (Doctoral dissertation). School of Informatics, University of Edinburgh.

10. Kusiak, A. (2001). Rough set theory: A data mining tool for semiconductor manufacturing. *IEEE Transaction on Electronics Packing Manufacturing*, 24(1), 44-50.

11. Lopez, L. (2005). Multi criteria decision aid application to a student selection problem. *Pesquisa Operational*, 25(1), 45-68.

12. Matthew, J. L., & Nydick, R. L. (1997). Group decision making in higher education using the analytic hierarchy process. *Research in Higher Education*, 38(5), 593-614.

13. Pal, S.K., & Skowron, A. (1999). *Rough fuzzy hybridization; A new trend in decision- making*. Springer.

14. Pawlak, Z. (1998). Rough set theory and its application to data analysis. *Cybernetics and Systems: An International Journal*, 29, 661- 688.

15. Pawlak, Z. (2002). Rough sets and intelligent data analysis. *Information Sciences*, 147, 1-12.

16. Ponce, J., & Karahoca, A. (2009). *Data mining and knowledge discovery in real life applications*. Vienna, Austria.

17. Roy, B. (1996). *Multi criteria methodology for decision aiding*. Kluwer, 1996.

18. Salvatore Greco, S., Matarazzo, B., & Slowinski, R. (2001). Rough sets theory for multi criteria decision analysis. *European Journal of Operational Research*, 129, 1-47.

19. Shusaku, T. (1988). Automated extraction of medical expert system rules from clinical databases based on rough set theory. *Information Sciences*, 112, 67-84.

20. Sinh Hoa, N., & Hung Son, N. (1988). Pattern extraction from data. *Fundamental Informaticae*, 34, 1-16.

21. Swiniarski, R., Hunt, F., Chalvet, D., & Pearson, D. (1995, August). Intelligent data processing and dynamic process discovery using rough sets, statistical reasoning and neural networks in a highly automated production systems. In: Proceedings of the Firs European Conference on Application of Neural Networks in Industry, Helsinki, Finland.

22. Tuan-Fang, F., Duen-Ren, L., & Gwo-Hshiung, T. (2007). Rough set-based logics for multi criteria decision analysis. *European Journal of Operational Research,* 182(1), 340-355.

23. Vandamme, J.P., Meskens, N., & Superby, J.F. (2007). Predicting academic performance by data mining methods. *Education Economics*, 15(4), 405-419.

24. Walczak, B., & Massart, D. L. (1991). Rough sets theory (Tutorial). *Chemometrics and Intelligent Laboratory Systems*. 47, 1-16.

25. Wilson, T. (1999). A student selection method and predictors of success in a graduate nursing program. *Journal of Nursing Education*, 38, 183-187.

26. Yao, Y.Y. (1996). Two views of the theory of rough sets in finite universes. *International Journal of Approximation Reasoning*. 15, 291-317.

27. Leung, Y., Wu, W.Z., & Zhang, W.X. (2006). Knowledge acquisition in incomplete information systems: A rough set approach. *European Journal of Operational Research*, 168(1), 164–180.

28. Ziarko, W. (1993). Variable precision rough sets model. *Journal of Computer and Systems Sciences*, 46(1), 39–59. http:// www.sanjesh.org. September 2013.

29. http://isna.ir/fa/news/92101507522.

30. http://www.sanjesh.org.