

ارائه رویکرد داده کاوی به منظور جذب پرسنل کارکنان شهرداری قزوین

امین محمودی^۱، نسترن سانگرد

چکیده

موفقیت یا شکست سازمان ارتباط مستقیمی با چگونگی جذب و نگهداری منابع انسانی آن دارد. همچنین انتخاب و به کارگیری شایسته ترین افراد در هر شغل یکی از مهم ترین مسایل و تصمیم های مدیریتی محسوب می شود. اغلب داده ها و اطلاعات فراوانی در سازمان ها وجود دارد که بدون استفاده قرار می گیرند داده کاوی، به عنوان راه حل برای چنین مسایلی است. با توجه به وجود بانک داده های فراوان در سازمان ها (به ویژه در حوزه مدیریت منابع انسانی) در این پژوهش نیز سعی بر این است ضمن استفاده از فن داده کاوی به عنوان یکی از ابزارهای فناوری اطلاعات و با پرداختن به یکی از مهم ترین مسایل مدیریتی در حوزه مدیریت منابع انسانی (گزینش و انتخاب کارکنان)، عوامل و قواعد موثر بر عملکرد نیروی انسانی شناسایی شده و با کشف الگوی پنهان بین وضعیت کارکنان با کارایی و اثربخشی آن ها و اعلام آن به مدیران، بتوان آنها را برای تصمیم گیری در مورد استخدام و انتخاب صحیح کارکنان، هدایت نمود. در این پژوهش، تکنیک داده کاوی مورد استفاده درخت تصمیم گیری، شبکه عصبی و رگرسیون بوده که در نهایت مشخص گردید که متغیرهای نوع مدرک تحصیلی، سابقه سازمانی و پست سازمانی تاثیر به سزایی در ارتقاء کارکنان داشته است. و با توجه به نتایج به دست آمده، پیشنهادهایی جهت کاربردی نمودن نتایج ارائه گردید.

کلید واژه ها: داده کاوی؛ انتخاب و گزینش کارکنان؛ درخت تصمیم؛ شبکه عصبی؛ رگرسیون

¹ استادیار گروه مهندسی صنایع، دانشکده مهندسی صنایع و مدیریت، دانشگاه غیاث الدین جمشید کاشانی، قزوین، ایران

² دانش آموخته رشته مدیریت صنعتی دانشگاه رجا، قزوین، ایران

پیشرفت شگفت‌انگیز فن آوری رایانه ای و مجهز شدن بشر به این ابزار، امکان جمع آوری اطلاعات دقیق و کامل در زمینهای مختلف را فراهم ساخته است و منجر به پیدایش ساختارهای داده بسیار حجیم شده است. بنابراین دستیابی به اطلاعات نهفته در داده های حجیم که لازمه مدیریت موثر است، با بکار بردن سیستم های سنتی استفاده از پایگاه داده های داده میسر نیست. شدت رقابت ها در عرصه های علمی، اجتماعی، اقتصادی، سیاسی و حتی نظامی نیز اهمیت عامل سرعت یا زمان دسترسی به اطلاعات را دو چندان کرده است. بنابراین نیاز به طراحی سیستم هایی که قادر به اکتشاف سریع اطلاعات مورد علاقه کاربران با تاکید بر حداقل مداخله انسانی باشند از یک سو و روی آوردن به روش های آماری متناسب با حجم داده های زیاد از سوی دیگر به خوبی احساس می شود. بنابراین در دنیای امروز، داده ها و اطلاعات اولیه از اهمیت زیادی برخوردار نیستند و اغلب به صورت گسترده در اختیار همه است. اما فرآوری و بازپروری آن و تولید دانش و استخراج گزاره های آن در فرآیند غنی سازی داده ها از اهمیت زیادی برخوردار است. لذا آنچه که مهم است، استخراج اطلاعات مفید از میان انبوهی از این داده ها در جهت گزینش مناسب به منظور کاهش نرخ جابجایی است. (خائف الهی و همکاران، ۱۳۸۶: ۱)

از طرفی دیگر، معمولاً سازمان ها در برخورد با ارزشمندترین سرمایه سازمانی یعنی سرمایه انسانی دچار چالشی به نام مدیریت منابع انسانی هستند که شاید در ظاهر و در ساختارها و نمودارهای سازمانی وجود داشته باشد ولی به دلیل عدم بکارگیری های نگرش علمی در این مقوله، معمولاً شاهد ایجاد مشکلاتی از قبیل، نارضایتی، عدم کارایی، جابجایی بالا و در بعضی موارد عدم تطابق کارکنان با مشخصات و ویژگی های شغلی هستیم که نهایتاً به کاهش کارایی و اثربخشی افراد و در نتیجه به کاهش بهره وری سازمان می انجامد. منظور از منابع انسانی یک سازمان، تمام افرادی است که در سطوح مختلف سازمان مشغول به کارند. کارمندیابی و گزینش گذرگاهی برای به کارگیری افراد و ایجاد منابع انسانی یک سازمان است (عادل آذر و همکاران، ۱۳۸۸: ۹۳) مشاغل مختلف و متنوع هستند و هر یک ویژگیهای خاص خود را دارند. از این رو هرگز نمی توان مجموعه ای از شاخصهای بخصوصی را تعیین کرد که عمومیت داشته باشند و بتوان از آنها برای ارزیابی عملکرد کارکنان در هر شغلی استفاده نمود (اسفندیار سعادت، ۱۳۹۱). از آنجا که کار در قالب شغل شکل می گیرد، لازم است تا ویژگی های یک شغل در ارتباط با خصوصیات شاغل، سنجیده شده و لحاظ گردد. لذا شرکت های با تکنولوژی بالا به حفظ سرمایه انسانی متکی هستند که از جمله مزایایی رقابتی پژوهش چارچوبی برای استخراج قوانین مفید از بین روابط اطلاعات و مشخصات کارکنان با رفتار کاری آنها می باشد (چن فو چاین^۱، ۲۰۰۸)

با وجود بانک داده های فراوان در سازمانها (به ویژه در حوزه مدیریت منابع انسانی)، در این پژوهش با استفاده از روش های داده کاوی^۲، به بررسی ارتباط میان داده های موجود در بانک داده ها با متغیر (ارتقاء) اقدام شده است.

فارغ از جامعه آماری مورد استفاده در این پژوهش، به کارگیری چنین شیوه ای در گزینش منابع انسانی، به ویژه در علوم مدیریتی، نوین و حائز اهمیت تلقی می شود. که تعمیم آن در سایر سازمانها باعث گزینش بهتر و شایسته تر افراد می باشد.

¹ Chen-fu chien

² Data mining Methods

تعاریف

داده کاوی

در متون آکادمیک تعاریف گوناگونی برای داده کاوی ارائه شده است. در برخی از این تعاریف داده کاوی در حد ابزاری که کاربران را قادر به ارتباط مستقیم با حجم عظیم داده ها می سازد معرفی گردیده است و در برخی دیگر، تعاریف دقیق تر که در آنها به کاوش در داده ها توجه می شود موجود است. داده کاوی یعنی جستجو در یک پایگاه داده ها برای یافتن الگوهایی میان داده ها. (بری مایکل وهمکاران^۱، ۲۰۰۴)

پیش بینی می شود که داده کاوی در دهه ی آینده با توسعه ای انقلابی مواجه می شود. (دنیل و جان لاروس^۲، ۲۰۰۶)

انتخاب و گزینش کارکنان

استخدام و گزینش، گذرگاهی برای به کارگیری افراد و ایجاد منابع در سازمان است. یک محیط تجاری که به طور فزاینده ای رقابتی و جهانی است، ضمن همراهی با نیاز به کیفیت و خدمت رسانی به مشتری، اهمیت استخدام و گزینش افراد مناسب را افزایش داده و به عنوان " کارفرمای انتخاب " مورد ملاحظه قرار گرفته است. (پورتر^۳، ۲۰۰۶) شرکت هایی با تکنولوژی بالا به حفظ سرمایه انسانی متکی هستند چرا که مهم ترین سرمایه یک سازمان نیروی انسانی آن است. و در سال های اخیر توجه به نیروی انسانی در سازمان ها بخش عظیمی از زمان و سرمایه سازمان های پیشرو را به خود اختصاص داده است. (چن فو چاین و لی فی چن^۴، ۲۰۰۸)

مدیریت منابع انسانی را شناسایی، انتخاب، استخدام، تربیت و پرورش نیروی انسانی به منظور نیل به اهداف سازمان تعریف کرده اند. (فرنچ^۵، ۱۹۸۶)

درخت تصمیم

درخت های انتخاب راهی برای نمایش یک سری از قوانین که به یک کلاس یا مقدار منجر می شود می باشند. برای مثال شما ممکن است بخواهید درخواستهای وام را برحسب ریسک اعتبار خوب یا بد طبقه بندی کنید. شکل بعد یک مدل ساده از یک درخت انتخاب به همراه توضیح در مورد تمام بسته های پایه آن یعنی گره انتخاب، شاخه ها و برگهای آن که این مساله را حل می کند نشان می دهد. (ادل اشتاین^۶، ۱۹۹۹)

¹ Berry Micheal

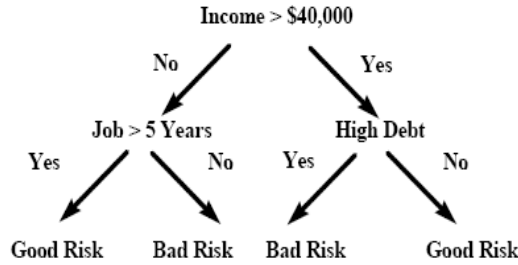
² Daniel T, John Laros

³ Porter

⁴Chen-Fu Chine ,Li-Fi Chen

⁵ French

⁶ Edelstein

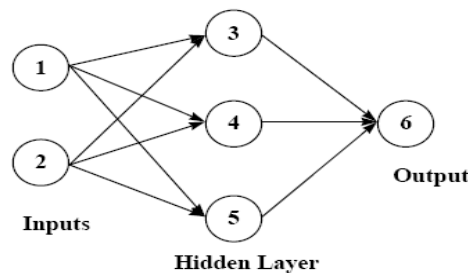


نگاره ۱.۱ درخت تصمیم

شبکه عصبی

شبکه های عصبی به طور خاصی مورد استفاده اند چرا که آنها ابزاری موثر برای مدل سازی مسائل بزرگ و پیچیده که ممکن است در آنها صدها متغیر پیش بینی کننده که فعل و انفعالات زیادی دارند وجود داشته باشد. (شبکه های عصبی زیستی بطور غیر قابل مقایسه ای پیچیده تر هستند). شبکه های عصبی می توانند در مسائل طبقه بندی یا حدس های بازگشتی (که در آنها متغیر خروجی پیوسته است) استفاده شوند. شبکه های عصبی قابلیت یادگیری دارند. این یادگیری با آزمایش روی رکوردهای خاص و با تنظیم وزن ارتباط، بین نرون ها صورت می گیرد (گوپتا^۱، ۲۰۰۶)

رایج ترین مدل شبکه های عصبی که مورد استفاده قرار میگیرد، مدل پرسپترون چند لایه یا شبکه های با پس انتشار خطا می باشد. در این مدل یک تابع که معمولاً غیرخطی می باشد بر روی جمع وزنی ورودی ها اعمال شده و خروجی مدل ساخته می شود. متغیرهای ورودی مساله به لایه اولیه یا لایه ورودی وارد می شوند و متغیرهای پاسخ در لایه نهایی یا خروجی قرار می گیرند و در این لایه ها برای هر متغیر یک گره در نظر گرفته می شود. لایه های بین این دو به لایه های پنهان موسوم هستند. روش قطعی برای تعیین تعداد گره های لازم در این لایه ها وجود ندارد و اغلب از روش سعی و خطا تعداد مناسب این گره ها تعیین می شود. در شبکه هایی که به صورت کامل به هم متصل شده اند هر گره در هر لایه با تمام گره های لایه بعدی خود در ارتباط می باشد. (زورادا^۲، ۱۹۹۲)



نگاره ۲.۲ شبکه عصبی

رگرسیون

¹ Gupta

² Zurada

این تکنیک آماری قدرتمند می تواند در کلیه ی مراحل فرایند داده کاوی داده ها مورد استفاده شود. رگرسیون خطی^۱، رگرسیون چندگانه^۲ و رگرسیون غیر خطی^۳ از انواع آن به شمار میروند.

در رگرسیون خطی دو نمونه مختلف از داده ها به صورت یک خط راست مدل سازی می شوند. خط رگرسیون ابزاری است برای پیش بینی مقدار یک متغیر بر حسب متغیری که به آن وابسته است.

رگرسیون چند گانه حالت توسعه یافته ای از رگرسیون خطی است، با این تفاوت که در آن بیش از دو متغیر وجود دارند و مقدار یکی با توجه به دیگر متغیرها تخمین زده می شود.

با تعمیم مدل خطی می توانیم برای متغیرهای طبقه بندی شده نیز تخمینی داشته باشیم. رگرسیون لجستیک نوع دیگری از رگرسیون هست که این کار را انجام می دهد.

پیشینه کاربرد داده کاوی در مدیریت منابع انسانی

داده کاوی را می توان در شناسایی ریشه ها و عوامل برخی از مشکلات منابع انسانی سازمان ها به کار برد. با توجه به اینکه برنامه ریزی منابع انسانی توسط سازمان ها برای اطمینان از اینکه آنها دارای تعداد صحیح و افراد مناسب برای موقعیت ها و مکان های مورد نیاز هستند مورد استفاده قرار می گیرند جایی که این فرایند به خوبی انجام می شود مزیت های بلند مدتی را هم برای سازمان و هم برای شخص کارمند به همراه دارد. داده کاوی برای تشخیص کارکنان غیر سودمند و خطا کار به منظور اخراج از شرکت، کشف کارکنان با استعداد و استفاده از آنها در واحدهای تجاری پیشنهاد شده است. (شرکت SPSS، ۲۰۰۴)

در این میان، تحقیقاتی که برای انتخاب کردن کارکنان با استفاده از رویکردهای داده کاوی انجام شده است به شرح جدول زیر می باشد.

نام محقق	سال ارائه	موضوع تحقیق	فن هوشمند استفاده شده
خانف الهی و همکاران	۱۳۸۶	انتخاب نیروی انسانی	داده کاوی
آذر و همکاران	۱۳۸۹	نرخ جابجایی کارکنان	داده کاوی
هوپر ^۴ و همکاران	۱۹۹۸	انتخاب پرسنل	سیستم دانش محور/سیستم خبره
هوآنگ ^۵ و همکاران	۲۰۰۴	انتخاب پرسنل	داده کاوی
هوآنگ و همکاران	۲۰۰۶	انتخاب پرسنل	شبکه عصبی مصنوعی
تای و هسو ^۶	2006	انتخاب پرسنل	داده کاوی فازی
ویجایا سارادهی ^۷ و همکاران	۲۰۱۱	پیش بینی شغل	داده کاوی/ماشین بردار
نی لی ^۸ و همکاران	۲۰۱۶	عملکرد انسانی	داده کاوی

1 Linear Regression
 2 Multiple Regression
 3 Nonlinear Regression
 4 Hooper
 5 Huang
 6 Tia Hsu
 7 Vijaya Saradhi
 8 NiLi

جیسوس پرال ^۱ و همکاران	۲۰۱۶	شناسایی شاخص های عملکرد انسانی	درخت تصمیم
آندرنا ^۲ و همکاران	۲۰۱۷	نقش های شغلی و مجموعه مهارت های مورد نیاز	ماشین بردار

جدول ۱. تحقیقات در حوزه کاربرد فن های هوشمند در منابع انسانی

همان گونه که در جدول فوق مشاهده می شود فن های هوشمند مورد استفاده در تحقیقات به صورت تکی یا استفاده از دو فن می باشد در تحقیق حاضر سعی بر این است که ضمن استفاده از سه تکنیک درخت تصمیم، شبکه عصبی و رگرسیون به طور همزمان تکنیک ها با یکدیگر مقایسه شوند و بهترین تکنیک با بیشترین درصد دقت دسته بندی انتخاب گردد.

روش فرایند تحقیق

در این مقاله سعی بر این شده است که با استفاده از فنون داده کاوی، عوامل موثر در جذب پرسنل شناسایی و رابطه آن با وضعیت ارتقای کارکنان بررسی گردید.

بعد از انتخاب نمونه ها نوبت به آماده سازی داده ها می رسد که شامل کلیه مراحل است که رکوردها و متغیرها را برای ساخت مدل، آماده می نماید. آماده سازی داده ها به مراحل قبل از داده کاوی اطلاق میگردد، هر چند از این تکنیک ها می توان در حین اجرای الگوریتم های داده کاوی نیز استفاده نمود، این مرحله گاهی با نام مرحله پیش پردازش داده ها نیز شناخته می شود. آماده سازی داده ها یکی از مهم ترین گام ها در فرآیند توسعه مدل ها به شمار می رود. (مهدی اسماعیلی، ۱۳۹۱). اغلب این گام از فرایند داده کای و کشف دانش، طولانی ترین و مهم ترین مرحله محسوب می شود. به طوری که هر قدر کیفیت آماده سازی داده ها بهتر انجام شود و زمان بیشتری برای آن صرف شود، مدل سازی و نتیجه گیری نیز بهتر و آسان تر خواهد بود.

در ادامه برای آشنایی با فرآیند طی شده در این فاز، فعالیت های طی شده به صورت گام به گام ذکر شده است.

- گام ۱) با برگزاری جلسات توجیهی با مدیران و مسولان منابع انسانی شهرداری، هدف از انجام پروژه شش تشریح و کلیه داده ها و اطلاعات موجود در رابطه با پرسنل و کارکنان ارزیابی شده است.
- گام ۲) در این مرحله نوع متغیرهای مورد نیاز که شامل ۳ دسته (متغیرهای شخصی، متغیرهای شغلی و متغیر ارزشیابی) بودند و با توجه به فیلهای موجود با همکاری کارشناسان منابع انسانی شهرداری تعریف شده است.
- گام ۳) طبق آمار مدیریت منابع انسانی شهرداری، از میان ۱۳۱۱ نفرات مشغول به کار در حوزه شهرداری ۱۲۶ نفر کارمند رسمی دارای ارزشیابی سالیانه بودند. با توجه به پراکندگی و مجزا بودن ارزشیابی سالیانه و بانک داده های پرسنلی، با استفاده از انواع ابزارهای برنامه نویسی و قابلیت های برنامه Microsoft Excel این انطباق انجام و در نهایت ۱۲۶ رکورد با ۱۱ فیلد مطابقت داده شدند.
- گام ۴) از آن جا که هر فیلد نمایانگر یک متغیر است بنابراین در اولین گام پس از تشکیل بانک داده ها، اقدام به حذف یا شناسایی متغیرهای مشابه و ترکیب و تبدیل آنها به یکدیگر شده است. به طور مثال متغیر رشته تحصیلی به چهار دسته (علوم انسانی، علوم پایه، فنی، مهندسی، کاردانش و ندارد) تقسیم بندی شد.

- گام ۵) در مرحله بعد، به بررسی و بازنگری مقادیر (کمی و کیفی) هر متغیر پرداخته شده که با ترکیب، تلفیق، تعریف دسته و طبقه، حذف و کاهش همراه بوده. طی این اقدام، ۱۱ فیلد موجود، تبدیل به ۱۰ فیلد نهایی شده است.
- گام ۶) در خاتمه پس از انجام فعالیت های ذکر شده، بانک داده ها در قالب فایل Excel و با ۱۲۶ رکورد و ۱۰ متغیر جهت ساخت مدل و درخت آماده شده است. این متغیرهای عبارتند از:
 - ✓ متغیرهای شغلی (شامل: سابقه کاری، پست سازمانی، حوزه محل خدمت)
 - ✓ متغیرهای شخصی (شامل: سن، جنسیت، وضعیت تأهل، مدرک تحصیلی، رشته تحصیلی، بومی بودن)
 - ✓ متغیر ارزشیابی (شامل: ارزشیابی سالانه)
- تعریف متغیر هدف: در این مقاله متغیر هدف ارتقاء کارکنان می باشد. ارتقاء با توجه به توانمندی کارکنان در ارتقای شغل خود در مدت اشتغال می باشد.
- ارزشیابی سالیانه: این متغیر بر مبنای آیین نامه مشخصی در شهرداری، حداقل سالی ۱ بار برای هر کارمند و توسط مدیر مستقیم وی تکمیل و عدد آن بر مبنای ۱۰۰ ارایه می شود. برای آسانی اجرای داده کاوی بر اساس متغیر وابسته کلیه اعداد ارایه شده تبدیل به شاخص کیفی شده اند که در جدول ذیل به ۷ دسته طبقه بندی شده است.

امتیاز	معیار
۱۰۰	عالی
۹۶-۹۹	خیلی خوب
۹۱-۹۵	خوب
۸۶-۹۰	معمولی
۸۱-۸۵	ضعیف
۷۶-۸۰	خیلی ضعیف
	بدون کارایی

جدول ۲. گزینه ها و طبقه بندی ارزشیابی سالیانه

جامعه آماری

منظور از جامعه آماری همان جامعه اصلی است، که از آن نمونه ای نمایا یا معروف به دست آمده است (ساروخان، ۱۳۸۸). بنا به تعریف دیگر جامعه آماری به کل گروه و افراد، وقایع یا چیزهایی اشاره دارد که پژوهشگر می خواهد به تحقیق درباره آن بپردازد. (سکاران، ۱۳۸۰)

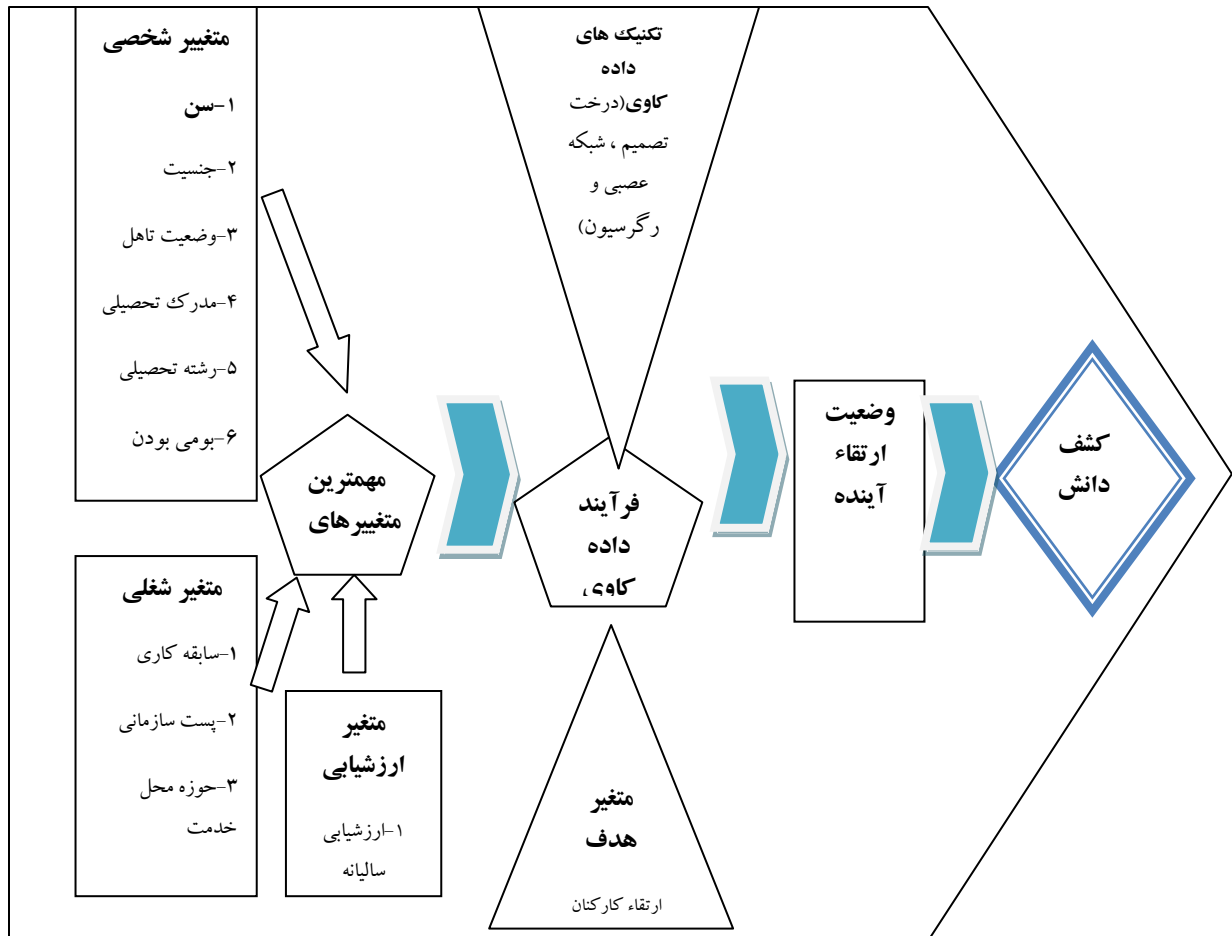
با توجه به تعاریف فوق جامعه آماری مقاله حاضر، کلیه کارمندان رسمی هستند که در شهرداری قزوین مشغول به کار می

باشند.

با توجه به این توضیحات تعداد کل پرسنل شهرداری استان قزوین ۱۳۱۱ نفر می باشند که با توجه به دسترس بودن نمره ارزشیابی مخصوص کارمندان رسمی می باشد از این میان تعداد ۱۲۶ نفر کارمند رسمی که دارای ۱۰ فیلد اطلاعاتی مفید، صحیح و کامل از اطلاعات شخصی، شغلی و ارزشیابی بوده اند به عنوان جامعه آماری نهایی انتخاب شده اند.

مدل مفهومی

با توجه به متغیرهای تعریف شده و بانک داده موجود، مدل مفهومی کشف دانش از نتایج داده های موجود به صورت نمودار زیر طراحی شده است.



نگاره ۳. مدل مفهومی کشف دانش از بانک داده

همان گونه که از این مدل نیز مشخص است، نخست، اقدام به تعریف گروه های متغیر و شناسایی مهم ترین متغیرهای تأثیر گذار خواهد شد، سپس با ورود این متغیرها و هم چنین متغیر هدف انتخاب شده به تفکیک هر الگوریتم، فرایند داده کاوی اجرا و نتایج آن به منظور پیش بینی ارتقاء شغلی در آینده به کار خواهد رفت. نتایج نهایی نیز به عنوان دانش سازمانی و به منظور افزایش کارایی و اثر بخشی در مجموعه سازمان، مورد استفاده قرار خواهد گرفت.

ساختار مسئله و تعریف متغیرها

مساله مورد بررسی در این مقاله مربوط به موضوع انتخاب پرسنل (جذب) میباشد که بصورت مطالعه موردی در شهرداری استان قزوین اجرا گردیده است. بدین منظور این مقاله به دنبال ارایه یک سیستمی است که ویژگیهای افراد را به عنوان ورودی گرفته و در خروجی در رابطه با جذب یا عدم جذب افراد تصمیم سازی کند. به منظور بررسی نتایج به دست آمده از هریک از روش ها، در خصوص هریک از مدل های تهیه شده ماتریس دقت و MatrixConfusion ارائه شده و کارایی و دقت روش های مختلف با یکدیگر مقایسه شده اند تا در نهایت بهترین روش با بیشترین دقت انتخاب شود.

مساله دارای ۹ متغیر ورودی (متغیر مستقل) میباشد و متغیر خروجی یک متغیر صفر و یک است که دسته یا طبقه هر فرد را به دو دسته طبقه بندی میکند. عنوان متغیرهای ورودی و خروجی در ذیل بیان شده است.

متغیرهای ورودی مسئله

- ۱ - جنسیت
- ۲ - وضعیت تاهل
- ۳ - رشته تحصیلی
- ۴ - مدرک تحصیلی
- ۵ - سن
- ۶ - سابقه کار
- ۷ - بومی بودن
- ۸ - پست سازمانی
- ۹ - محل خدمت

متغیر خروجی

دسته یا طبقه مربوط به فرد که در دو دسته جذب یا عدم جذب دسته بندی می شوند.

متغیر ورودی به میزان درجه اهمیتی که سازمانهای استخدام کننده برای عوامل ذکر شده قائل می شوند، بستگی به شرایط بازار کار نیز دارند. اصولاً هنگامی انتخاب به معنای واقعی وجود خواهد داشت که برای هر شغل بخصوصی تعداد زیادی متقاضی وجود داشته باشد. (سعادت، ۱۳۹۱) بنابراین متغیر ورودی این پژوهش بنابر میزان اهمیتی که سازمان مورد نظر برای انتخاب کارکنان در نظر گرفته است انتخاب شده اند.

انتخاب مشخصه

در این مرحله از پژوهش و قبل از استفاده از تکنیکهایی جهت مدلسازی و آموزش دادهها، ابتدا از بین متغیرها (مشخصه ها) که بطور بالقوه در متغیر جذب یا عدم جذب افراد تاثیر گذار هستند باید متغیرهایی که بیشترین اثرگذاری را دارند شناسایی شوند تا در ادامه برای مدلسازی داده ها از این متغیرها استفاده شود. بدین منظور روش رگرسیون Stepwise در نرم افزار مینی تب ۱۶ مورد

استفاده قرار گرفته است. خروجی اجرای آن در جدول ۳ آمده است نشان می دهد که از بین همه متغیرها، متغیرهای نوع مدرک، سابقه و پست سازمانی در هر مرحله وارد مدل شده‌اند و از آنجائیکه تاثیر به سزایی در کارایی مدل داشته اند از مدل خارج نشده اند.

Stepwise Regression: ... تاهل, جنسیت versus دسته

Alpha-to-Enter: 0.15 Alpha-to-Remove: 0.15

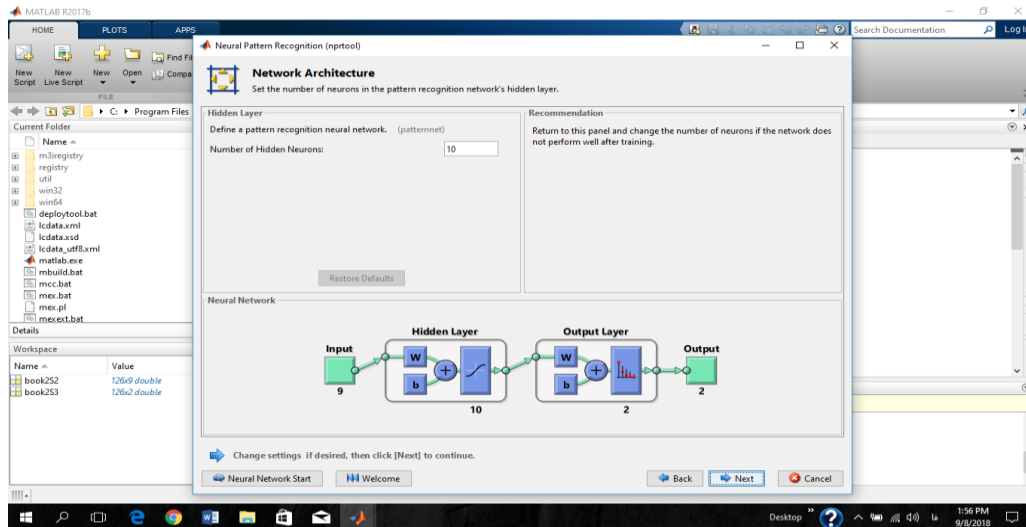
Response is دسته on 9 predictors, with N = 126

Step	1	2	3
Constant	2.094	2.277	2.428
مدرک	-0.137	-0.124	-0.139
T-Value	-4.08	-3.82	-4.15
P-Value	0.000	0.000	0.000
سابقه		-0.0146	-0.0142
T-Value		-3.35	-3.26
P-Value		0.001	0.001
پست			-0.078
T-Value			-1.65
P-Value			0.101
S	0.416	0.400	0.397
R-Sq	11.85	19.20	20.97
R-Sq (adj)	11.14	17.89	19.02
Mallows Cp	11.2	2.0	1.4

جدول ۳. اجرای رگرسیون stepwise جهت شناسایی تاثیرگذارترین متغیرها

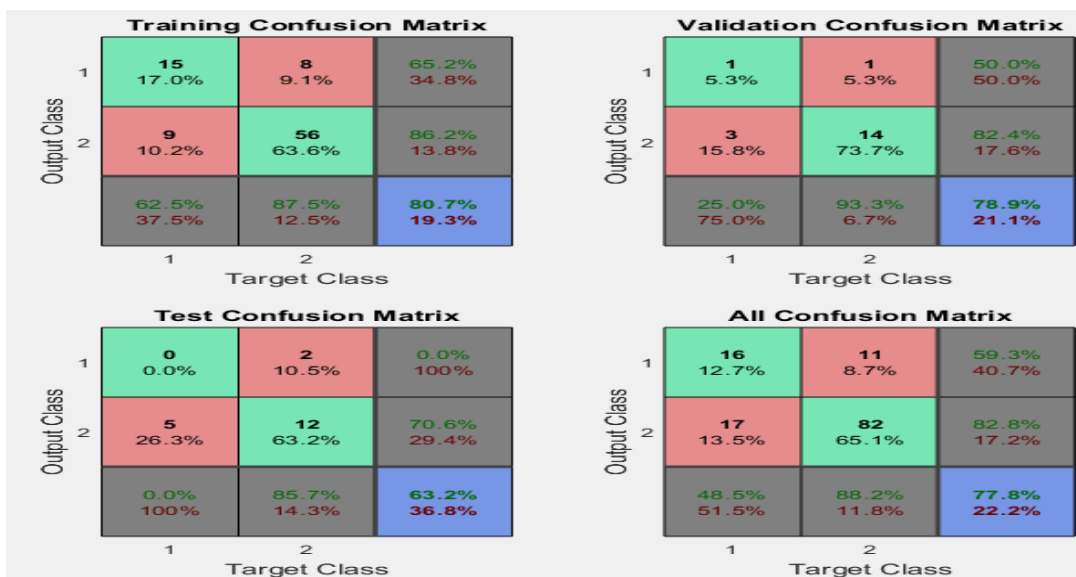
روش شبکه عصبی مصنوعی

به منظور ایجاد مدلی با استفاده از تکنیک شبکه‌های عصبی مصنوعی (ANN) از نرم افزار متلب (MATLAB) ورژن ۲۰۱۷ مورد استفاده قرار گرفته است. بدین منظور از جعبه ابزار nprtool^۱ استفاده شده است. ساختار شبکه عصبی مورد استفاده قرار گرفته در نگاره ۴ مشاهده می شود.



نگاره ۴. ساختار شبکه عصبی استفاده شده جهت آموزش داده ها

به منظور ارزیابی روش مورد نظر و سنجش میزان دقت تکنیک مورد استفاده قرار گرفته ماتریس درهم ریختگی^۱ برای داده های آموزش و داده های ارزیابی بصورت نگاره ۵ به نمایش درآمده است.

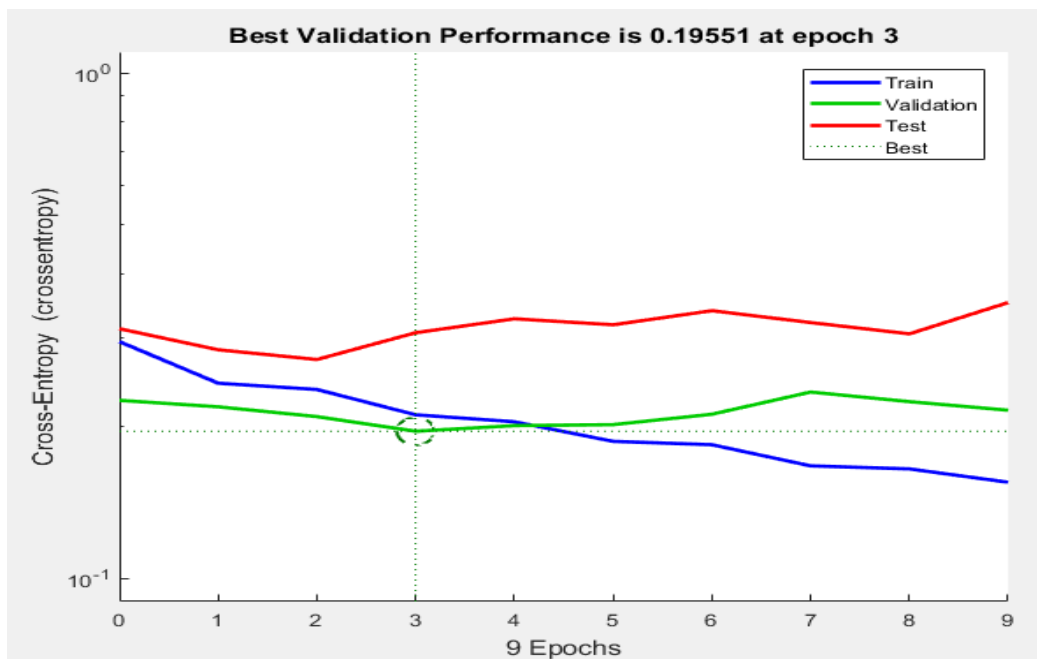


نگاره ۵. ماتریس درهم ریختگی روش شبکه های عصبی به تفکیک داده های آموزش و ارزیابی

همانگونه که در نگاره ۵ پیداست، دقت تشخیص و پیش بینی مدل شبکه های عصبی در مجموع برای داده های آموزشی و ارزیابی ۷۷/۸٪ بوده است و این مقدار فقط برای داده های ارزیابی ۷۸/۹٪ میباشد که در نهایت هدف در هر مدلی مینیمم کردن این مقدار است.

¹Confusion matrix

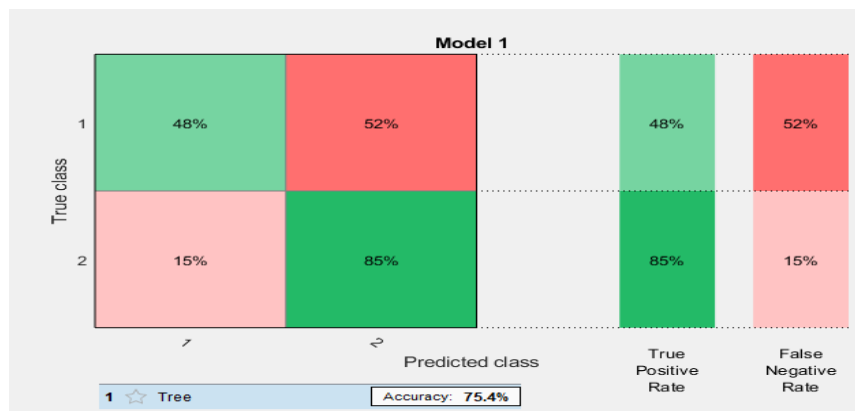
در نهایت، در روش شبکه های عصبی به منظور بررسی کارایی مدل آموزش یافته، نمودار روند کارایی آموزش مدل در تکرارهای مختلف بصورت نمودار زیر شده است بطوری که کمترین مقدار کارایی برای داده های ارزیابی در تکرار ۳ ام اتفاق افتاده است و همچنین یک روند منطقی در نمودار مشاهده میشود.



نمودار ۱. نمودار کارایی روش شبکه های عصبی

روش درخت تصمیم

یکی دیگر از تکنیکهایی که در این مقاله جهت مدلسازی داده ها مورد استفاده قرار گرفته است، تکنیک درخت تصمیم به عنوان یکی از مشهورترین روشهای دسته بندی داده ها میباشد. به منظور اجرای این تکنیک از نرم افزار متلب و از جعبه ابزار classification learner و از ابزار Fine Tree استفاده شده است. به منظور ارزیابی میزان دقت روش پیشنهادی ماتریس درهم ریختگی روش درخت تصمیم بصورت نگاره ۶ نشان داده شده است. همانگونه که در نگاره ۶ نشان داده شده میزان دقت پیش بینی روش درخت تصمیم برای داده های مساله ۷۵.۴٪ میباشد که این مقدار برای داده مقدار خوب و قابل قبولی است.

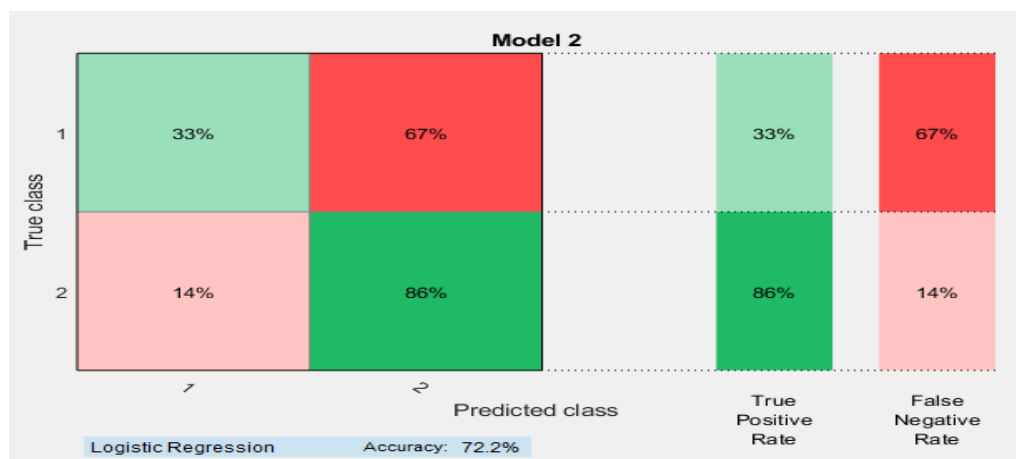


نگاره ۶. ماتریس درهم ریختگی روش درخت تصمیم

روش رگرسیون لجستیک^۱

تکنیک رگرسیون لجستیک یکی دیگر از روشهای مورد استفاده در پژوهش می باشد که برای مدلسازی داده ها مورد استفاده قرار گرفته است. این تکنیک عمدتاً برای داده هایی که خروجی آنها بصورت یک متغیر باینری میباشد خروجی بهتری از خود نشان میدهد. به منظور اجرای این تکنیک از نرم افزار متلب و از جعبه ابزار classification learner و از ابزار Logistic regression استفاده شده است.

به منظور ارزیابی میزان دقت روش پیشنهادی ماتریس درهم ریختگی روش رگرسیون لجستیک بصورت نگاره ۷ نشان داده شده است. همانگونه که در نگاره ۷ نشان داده شده میزان دقت پیش بینی روش رگرسیون لجستیک برای داده های مساله ۷۲/۲٪ می باشد.



نگاره ۷. ماتریس درهم ریختگی روش رگرسیون لجستیک

مقایسه کارایی روش های مورد استفاده

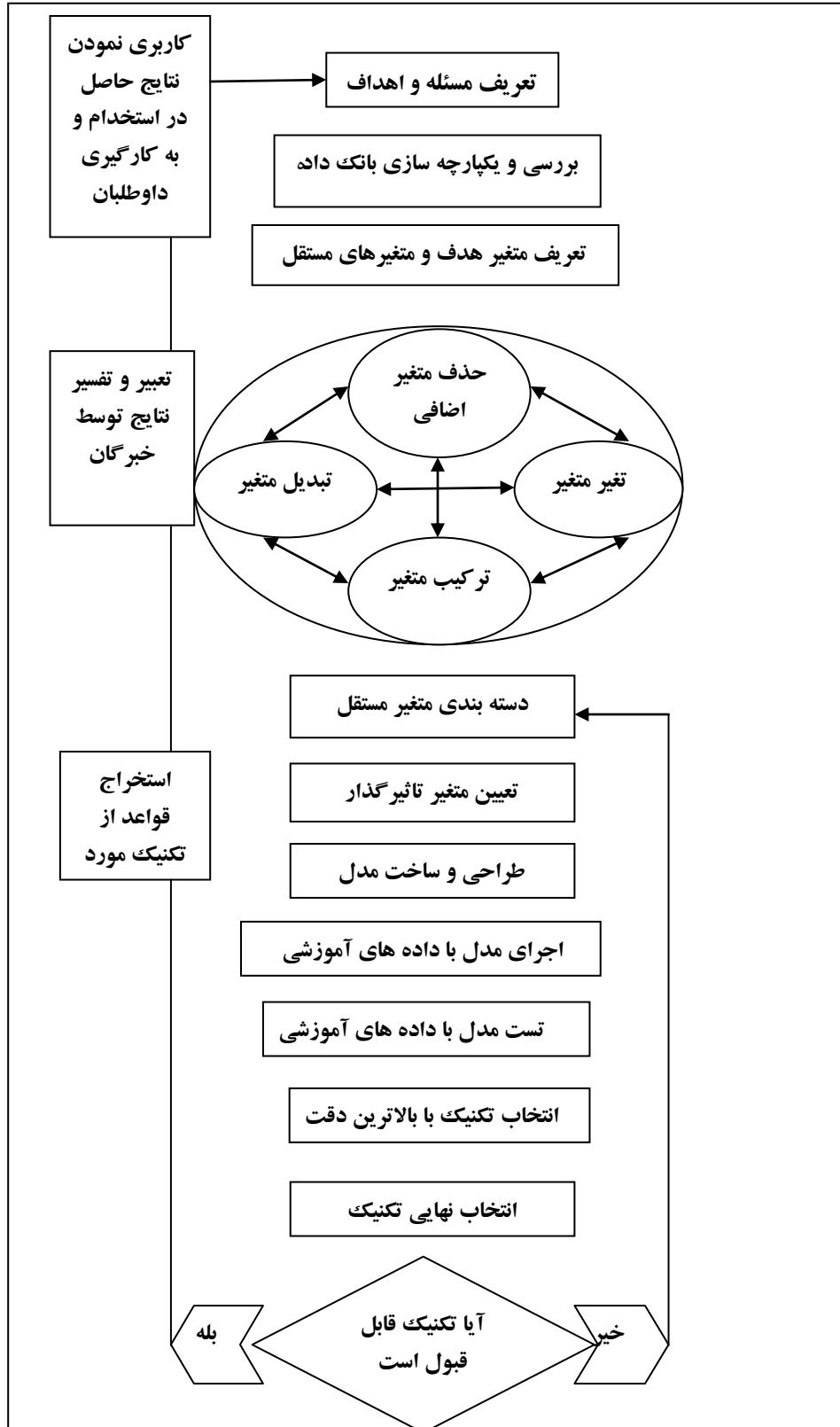
در این بخش مقایسه بین کارایی روشهای اجرا شده در جدول ۴ نشان داده شده است. همانگونه که مشاهده میشود از بین روش های شبکه های عصبی، درخت تصمیم و رگرسیون لجستیک روشی که بیشترین دقت را در شناسایی صحیح دسته ها دارد در این مساله روش شبکه های عصبی میباشد که با دقت ۷۸/۹٪ دقت مناسب تری نسبت به بقیه روش ها در این مساله دارد.

نام روش	دقت دسته بندی
شبکه های عصبی	۷۸/۹٪
درخت تصمیم	۷۵/۴٪
رگرسیون لجستیک	۷۲/۲٪

جدول ۴.مقایسه کارایی روش های اجرا شده

مدل نهایی پژوهش

با توجه به مراحل طی شده در پژوهش و نتایج ، مدل نهایی این پژوهش به صورت نمودار ۲ ارائه شده است.



نمودار ۲. مدل نهایی کشف دانش حاصل از متغیرهای ورودی با روش داده کاوی

نتیجه گیری و پیشنهادات

به کارگیری صحیح و علمی کارکنان به ویژه در سازمان های خدمات محور ، دارای اهمیت خاصی است که استفاده از علم داده کاوی و کشف دانش نهفته در این زمینه بسیار موثر خواهد بود. در این پژوهش ضمن بیان رویکردهای پیشین به حوزه استفاده از داده کاوی در مدیریت منابع انسانی، تلاش شد تا مدل های پیشین توسعه داده شود و مدل جدیدی با نگرش پویا و تاکید بر انتخاب بهترین تکنیک مناسب با بانک داده ها ارائه شود. همچنین با رویکرد مفهومی در خصوص متغیرهای تاثیر گذار با متغیر هدف مشخص گردید که از بین ۹ متغیر تاثیر گذار فقط سه متغیر: "مدرک ، سابقه کاری و پست سازمانی" تاثیر گذار هستند.

در همین رابطه پیشنهادهایی به سازمان مورد مطالعه ارائه شده است که نتیجه تجزیه و تحلیل خروجی مدل طراحی شده بوده است. از جمله این پیشنهادات می توان به موارد زیر اشاره نمود:

- با توجه به اینکه، از بین همه متغیرها، نوع مدرک بیشترین تاثیر را داشته به طوری که در هر مرحله وارد مدل شده و از آنجائیکه تاثیر به سزایی در کارایی مدل داشته از مدل خارج نشده اند، بنابراین ضروری است که در بانک داده ها متغیرهای مانند معدل و همچنین دانشگاه محل تحصیل با هدف اندازه گیری پتانسیل های آینده اضافه شود.
- می توان به بررسی و در صورت امکان پذیر بودن انجام مصاحبه قبل از ورود داوطلب به سازمان از طریق روش ها و ابزارهای آموزشی امتیاز دهی به عمل آورد.
- با توجه به اینکه در شهرداری قزوین فقط کارمندان رسمی دارای ارزشیابی سالیانه هستند و با توجه به فرم مورد نظر سنجیده می شوند می توان در صورت امکان پذیر بودن و به منظور افزایش کارایی برای سایر قراردادها اعم از رسمی و غیررسمی فرم ارزشیابی را تکمیل کرد.
- در نهایت آنچه بیشتر حائز اهمیت است، استفاده بهینه و گسترده از علم داده کاوی در سایر حوزه های مدیریتی است که می تواند ضمن استفاده ی کم هزینه از داده ه ای موجود در سازمان ها، افق جدیدی را از دانش های نهفته در سازمان برای مدیران روشن نماید.

پیشنهادات برای تحقیقات آینده

در راستای تکمیل و بهبود نتایج این پژوهش و همچنین با هدف ایجاد انگیزه جهت تحقیق در حوزه مدیریت منابع انسانی، موضوعات زیر به عنوان تحقیقات آینده به سایر پژوهشگران، پیشنهاد می شود:

- در نظر گرفت سایر متغیرهای تاثیر گذار در جذب پرسنل از جمله توانایی، مهارت، علاقه و...
- بررسی تفاوت نتایج داده کاوی مدل هوشمند جذب پرسنل توسط مجموعه های فازی و شبکه عصبی

منابع

- آذر، عادل، احمدی، پرویز، سبط، محمد وحید، (۱۳۸۸)، "تعیین شاخص های مؤثر بر گزینش منابع انسانی با رویکرد داده کاوی"، فصلنامه تعلیم و تربیت، شماره ۱۰۰
- آذر، عادل، احمدی، پرویز، سبط، محمد وحید، (۱۳۸۹)، "طراحی مدل انتخاب نیروی انسانی با رویکرد داده کاوی"، نشریه مدیریت فناوری اطلاعات، دوره ۲، شماره ۴
- خائف الهی، احمدعلی؛ متقی، ایمان و سبط، محمد وحید (۱۳۸۶)، "بررسی تاثیر و به کارگیری مدل استخدام مبتنی بر داده کاوی بر نرخ جابه جایی کارکنان"، دانشگاه امیر کبیر، کنفرانس IDMC07
- رونق، یوسف (۱۳۸۰) "مطالعه کار و استاندارد شغل"، تهران، انتشارات مرکز آموزش مدیریت دولتی
- ساروخانی، باقر، (۱۳۸۸)، "روشهای تحقیق در علوم اجتماعی"، پژوهشگاه علوم انسانی و مطالعات فرهنگی، چاپ پانزدهم
- سعادت، اسفندیار، (۱۳۹۱)، "مدیریت منابع انسانی"، سازمان مطالعه و تدوین کتب علوم انسانی دانشگاهها (سمت)، مرکز تحقیق و توسعه علوم انسانی
- سکاران، اوما، (۱۳۸۰)، "روشهای تحقیق در مدیریت"، مترجم: صیائی، محمد، شیرازی، محمد، تهران، مرکز آموزش مدیریت دولتی، چاپ اول
- میشلین کمبر، ژان پی، ژیاوی هان، (۱۳۹۱)، "مفاهیم و تکنیک های داده کاوی"، مترجم: مهدی اسماعیلی، انتشارات نیاز دانش، چاپ سوم
- Andrea De Mauro, Marco Greco, Michele Grimaldi , Paavo Ritala (2017).” **Human resources for Big Data professions: A systematic classification of job roles and required skill sets** . information processing management
- Berry Michael J. A., Gordon S. Linoff(2004). **Data Mining Techniques for Marketing Sales and Customer Relationship Management**. 2 Ed. John.Wiley.&.Sons. Publishing Inc.;
- Chen-Fu chine , li-fei chen (2008), “**data mining to improve personnel selection and enhance human capital: A case study in high-technology industry**” Expert system with Applications 34.
- Daniel, T, and John Larose (2006). **Data Mining Methods And Models**. Wiley & Sons Publishing Inc.

Edelstein, Herber A (1999). **Introduction To Data Mining And Knowledge Discovery**. Two Crows Corporation.

French, Wendell l(1986).; **Human Resources Management**; boston; Houghton Mifflin co.,

Gupta, G.K. (2006). **Introduction to Data Mining with case studies**. New Delhi :Prentice Hall of India,Inc

Hooper, R.S., et al.,(1998), **Use an Expert System in a personnel selection process**, Expert system and application, 14(4):p. 425-432

Huang, L.C., et al.,(2004), **Applying fuzzy neural network in human resource selection system**, In Proceeding NAFIPS'04, IEEE Annual Meeting of the fuzzy information 2004.

Huang, M.J., Y.L., Tsou and S.C. Lee,(2006), **integrating fuzzy data mining and fuzzy artificial neural network for discovering implicit knowledge**, Knowledge-Based Systems,19(6) pp.396-403

Jesús Peral, Alejandro Maté, Manuel Marco,(2016),” **Application of Data Mining techniques to identify relevant Key Performance Indicators**”, Author’s Accepted Manuscript, S0920-5489(16)30092-7

Ni Li, Haipeng Kong , Yaofei Ma, Guanghong Gongl , Wenqing Huai,(2016), “**Human performance modeling for manufacturing based on an improved KNN algorithm**”, Int J Adv Manuf Technol 84:473–483

Porter K., P. Smith, R. Fagg(2006),. **Leadership and Management for HR Professionals**. Oxford: Butterworth Heinemann;.

Spss Inc. **Effectively Manage Human Resource Using Data Mining**. 2004.

http://www.Spss.Com/Applications/Government/Human_Resource_Planning.Htm

Abstract:

The success and failure of the organization has a direct link with how to attract and maintain human resources.

Also, selecting and employing the most qualified people in each job one of the most important issues and management decisions is considered.

Often there are a lot of data and information in organizations that are not used. Data mining is a problem for the fetus. Due to the existence of numerous data banks in organization. in this research, we try to use the data mining fan, As one of the it tools by addressing one of the most important issues in HR management (selection and selection of employ), Factors and rules that affect the performance.

Human resources are detected by discovering the hidden pattern between the status of employ with their efficiency and effectiveness and declaring it to managers. They can be taken to decide on recruitment and the right guided by choice staff.

In this study , the data mining technique was used, decision tree, nerve network and regression.

Finally, it was determined that the variables type of degree, organizational and organizational history has had a significant impact on employee upgrades. According to the results, suggestions were made to use the results.

It is hoped that this research will be a step towards improvement and revision in current trends, the selection of human resources.

Keywords: Data mining, Selection and selection of employ, Decision tree, Neural network, Regression