

عوامل موثر بر تمایل به ترک سازمان با استفاده از الگوریتم‌های مبتنی بر شبکه عصبی و ژنتیک چند هدفه

سهام خضیری عفاوی^۱ - محسن سرداری زارچی^۲ - سیدمحمد مهدی فاطمی بوشهری^۳

چکیده

بهبود بازدهی سرمایه انسانی از آن‌جا که می‌تواند نقش موثری در کارایی سازمان داشته باشد، همواره یکی از موضوعات پژوهش بوده است. میزان تمایل به ترک سازمان یکی از عوامل تأثیرگذار بر کارایی سرمایه انسانی است که آن را می‌توان با استفاده از الگوهای درون داده‌ای، شرایط حاکم بر سازمان و بررسی عوامل مؤثر بر آن پیش‌بینی کرد. به همین منظور، از الگوریتم‌های هوشمند مبتنی بر شبکه عصبی و الگوریتم ژنتیک چندهدفه برای پیش‌بینی تمایل کارکنان به ترک سازمان در این پژوهش بهره گرفته شده است. در این راستا، ابتدا با طراحی پرسش‌نامه‌ای، نظرات کارکنان شرکت بهره‌بردار نفت و گاز کارون در مورد رضایت مندی و تمایل به ترک سازمان، جمع‌آوری شد و بر اساس آن، یک مجموعه داده تدوین گردید. سپس با استفاده از شبکه‌های عصبی مصنوعی به عنوان طبقه بند و الگوریتم تکاملی ژنتیک چندهدفه برای انتخاب ویژگی‌های مؤثر، یک سیستم خبره طراحی شد. به منظور تست و ارزیابی الگوریتم شبکه عصبی طراحی شده با مجموعه داده استاندارد ایجاد شده، آموزش‌های لازم ارائه شد. نتایج ارزیابی سیستم پیشنهادی بیانگر آن است که با به کارگیری الگوریتم ژنتیک چندهدفه و شبکه‌های عصبی مصنوعی می‌توان مدلی ارائه کرد تا علاوه بر پیش‌بینی میزان تمایل کارکنان به ترک سازمان با دقت بالای ۸۸٪، با انتخاب ویژگی‌های مؤثر، عوامل کلیدی ترک سازمان را نیز مشخص کند.

واژگان کلیدی: جایگزینی، داده کاوی، انتخاب کارکنان، شبکه‌های عصبی مصنوعی، الگوریتم ژنتیک چندهدفه

^۱ کارشناسی ارشد مهندسی کامپیوتر- نرم افزار، دانشگاه آزاد اسلامی واحد یزد، یزد، ایران (khaziri.s@nisoc.ir)

^۲ دکتری مهندسی کامپیوتر- هوش مصنوعی، استادیار گروه کامپیوتر دانشگاه سراسری میبد، یزد، ایران، نویسنده مسئول (sardari@meybod.ac.ir)

^۳ کارشناسی ارشد مهندسی کامپیوتر- نرم افزار، مرکز فناوری اطلاعات و خدمات رایانه ای، دانشگاه یزد، یزد، ایران (fatemi@yazd.ac.ir)

مقدمه

سرمایه انسانی یکی از با ارزش‌ترین دارائی‌های یک سازمان است و به عنوان مهم‌ترین عامل کسب مزیت رقابتی برای سازمان، از جایگاه ویژه‌ای برخوردار است. هدف از این مطالعه یافتن راهکاری برای عدم ترک سازمان توسط کارکنان و حفظ رقابت‌پذیری شرکت‌ها است. تمرکز این مطالعه بر روی کارکنان عملیاتی بود که سهم بالایی از کارکنان جایگزین شده را تشکیل می‌دهند. شرکت بهره‌برداری نفت و گاز کارون از شرکت‌های تابعه شرکت ملی نفت ایران در اهواز است. این شرکت از بزرگ‌ترین شرکت‌های تولیدکننده نفت است که با در اختیار داشتن صدها حلقه چاه نفت، ده‌ها کارخانه بهره‌برداری، نمکزدائی، تزریق گاز، جمع‌آوری و تقویت فشار گاز و هزاران کیلومتر خطوط لوله اصلی و جریان‌ی نفت و گاز، در پهنه استان خوزستان گسترده است.

هدف این پژوهش ارائه و ارزیابی یک سیستم خبره با استفاده از ابزارهای یادگیری ماشین و مدل مبتنی بر شبکه عصبی مصنوعی و الگوریتم ژنتیک چندهدفه برای طبقه‌بندی میزان تمایل کارکنان به ترک سازمان است. وجود یک سیستم خبره با این قابلیت می‌تواند به سازمان‌های مشابه برای جلوگیری به از دست دادن منابع انسانی کلیدی سازمان کمک کند. همچنین تلاش شده است پارامترهای تأثیرگذار در تغییر موقعیت شغلی مورد ارزیابی قرار گرفته و انتخاب شوند تا بدین وسیله گامی مهم در جهت رفع این چالش‌ها و پیش‌بینی میزان تمایل کارکنان به ترک سازمان برداشته شود.

در این مقاله از الگوریتم‌های شبکه عصبی و الگوریتم‌های تکاملی برای طبقه‌بندی و کاهش ویژگی‌ها استفاده شده است. شبکه عصبی^۱ طبقه‌بندی‌کننده کارایی است، اما متأسفانه در آزمایش‌های مکرر دارای نتایج منحصر به فردی نیست و با هر بار اجرای شبکه، ممکن است نتایج متفاوت به دست آید، بنابراین بهتر است از ترکیب شبکه عصبی و یک الگوریتم تکاملی دیگر به منظور افزایش سرعت و کارایی مشابه استفاده گردد. از طرفی افزایش تعداد متغیرها، اگرچه می‌تواند منجر به تولید دقت بالایی در طبقه‌بندی شود، اما

^۱ Neural Network

به طور همزمان ممکن است باعث پیچیده‌تر شدن طبقه‌بندی و افزایش حجم محاسباتی آن نیز شود. بنابراین پیشنهاد می‌شود ضمن در نظر گرفتن معیار کاهش تعداد متغیرها، معیار خطای دسته‌بندی نیز در انتخاب عوامل مؤثر در نظر گرفته شود. به همین دلیل در این پژوهش برای دسته‌بندی از ماشین بردار پشتیبان و از روش تکاملی چند هدفه و به طور خاص از الگوریتم ژنتیک برای انتخاب ویژگی (عوامل مؤثر) استفاده شده است. به منظور در نظر گرفتن اهداف توأم انتخاب ویژگی، از معیارهای اطلاعات متقابل که شامل محتویات اطلاعات ویژگی‌ها و قدرت حفظ ساختار آن هاست و معیار خطای دسته‌بندی ماشین بردار در تابع برازندگی الگوریتم تکاملی استفاده شده است. الگوریتم ژنتیک چندهدفه^۲ برای ایجاد فضای جستجو یا انتخاب زیرمجموعه ویژگی‌های مؤثر و الگوریتم شبکه عصبی برای ارزیابی این زیرمجموعه استفاده می‌شود. یکی از دلایل استفاده از الگوریتم ژنتیک، بررسی مجموعه‌ای از جواب‌های ممکن و هم‌چنین حساسیت کمتر نسبت به شکل خاصی از نقاط بهینه است که باعث می‌شود برای به کارگیری در مسائل بهینه‌سازی چندهدفه مناسب گردد. مسئله بهینه‌سازی در واقع یافتن جواب یا جواب‌هایی بر روی یک مجموعه از گزینه‌های امکان‌پذیر باهدف بهینه کردن معیار یا معیارهای مسئله است. مسئله بهینه‌سازی چند هدفه زیرشاخه‌ای از مجموعه روش‌های تصمیم‌گیری چند معیاره^۳ است که در میان مجموعه نامحدودی از جواب‌های محتمل صورت می‌گیرد. هم‌چنین درخت تصمیم‌گیری یکی از ابزارهای قوی و متداول برای دسته‌بندی و پیش‌بینی است. درخت تصمیم‌گیری برخلاف شبکه‌های عصبی به تولید قانون می‌پردازد. در واقع درخت تصمیم‌گیری پیش‌بینی خود را در قالب مجموعه قوانینی توضیح می‌دهد درحالی که در شبکه‌های عصبی تنها پیش‌بینی نهایی بیان می‌شود و چگونگی آن در خود شبکه پنهان باقی می‌ماند (آذر، احمدی و سبط، ۱۳۸۶).

در مسائل بهینه‌سازی اهداف به صورت توابع ریاضی بوده و نمایانگر استفاده مناسب برای بهبود تصمیمات در بهینه‌سازی است. بر اساس مفهوم چیرگی پارتو می‌توان معیار

^۲ Non-dominated sorting genetic algorithm

^۳ Multi-criteria decision-making methods

بهینگی را ارزیابی کرد (Goldberg, ۲۰۱۱). در این مقاله دو تابع هدف در نظر گرفته شده است، که نقاط پارتو مربوط به هر جفت از توابع با استفاده از الگوریتم‌های پیشنهادی به دست می‌آیند. ساختار مقاله حاضر بدین شرح است: بخش دوم پیشینه‌ای از روش‌های به کاررفته در رابطه با موضوع موردبررسی ارائه می‌دهد. در بخش سوم داده‌ها و روش پژوهش توضیح داده می‌شود. بخش چهارم به پیاده‌سازی روش‌ها و بررسی و مقایسه نتایج می‌پردازد و بخش پنجم به نتیجه‌گیری مقاله ختم می‌شود.

مروری بر مبانی نظری و پیشینه تحقیق

داده کاوی را می‌توان درشناسایی ریشه‌ها و عوامل برخی از مشکلات منابع انسانی سازمان‌ها به کاربرد. هم‌چنین از داده کاوی به منظور کاهش زمان، کاهش هزینه و گسترش دستیابی به داوطلبان دارای بالاترین کیفیت در استخدام نیروهای جدید استفاده شده است. خائف الهی، متقی و سبط (۱۳۸۵)، تأثیر به کارگیری مدل استخدام مبتنی بر داده کاوی را بر نرخ جابه جایی کارکنان موردبررسی قرار داده‌اند. این مطالعه با هدف طراحی یک مدل داده کاوی مبتنی بر درخت تصمیم و ایجاد الگویی مناسب به منظور استفاده کاربردی از فن داده کاوی، تلاش نموده است تا ساختاری مبتنی بر الگوریتم CHAID برای داده کاوی منابع انسانی ارائه نماید تا امکان استخراج ارتباط بین پروفایل کارکنان و پیشگویی رفتار آن‌ها فراهم آورد. نتایج حاصل از این تحقیق نشان داد اگر کارمندی از کارش خارج شود، برای جایگزینی نیروی جدید به دوره‌ای سه ماهه (به عنوان دوران آزمایشی) نیاز است که این مسئله در واقع نوعی خسارت محسوب می‌شود. از سوی دیگر خروج افراد حدوداً یکسال بعد از استخدام آن‌ها، تبدیل به یک مسئله مدیریتی خواهد شد، لذا مشخصات کسانی که در دوران آزمایشی رد می‌شوند یا کسانی که در طول یک سال از کار اخراج شده یا استعفا می‌دهند، به طور کامل ارتباط بین پروفایل افراد و ویژگی‌های شغلی آن‌ها را نشان می‌دهد.

چن فا-چین و لو-فا-چن (۲۰۰۸) پژوهشی را با استفاده از تجزیه و تحلیل درخت تصمیم و استخراج قوانین انجمنی بر روی داده‌های منابع انسانی به منظور استخراج روابط بین پروفایل‌های کارکنان و رفتارهای سازمانی انجام دادند. در این پژوهش عمده دلایل

جابه جایی فشار، کشش عوامل خارجی، مسائل کارکنان و طرح‌های شرکت عنوان شد. با توجه به استقرار نظام مدیریت ارزیابی عملکرد در شرکت، کارکنان به سه دسته برجسته ۱۰٪، موفق ۸۵٪ و نیاز به بهبود ۱۵٪ تقسیم‌بندی شدند و درنهایت ۳۰ قانون معنی‌دار مربوط به استراتژی استخدام، استخراج شد.

آذر، احمدی و سبط (۱۳۸۹)، مدلی را با استفاده از داده کاوی برای انتخاب نیروی انسانی ارائه دادند. تکنیک داده کاوی مورد استفاده در این پژوهش، درخت تصمیم‌گیری است و استخراج قواعد نیز با استفاده از الگوریتم‌های QUEST، CHAID، C5، CART و QUEST انجام شده است. درنهایت ضمن ارائه مدلی جهت انتخاب متغیرهای تأثیرگذار، متغیرهدف والگوریتم‌های مناسب، از بین قواعد به دست آمده، قواعد غیر بدیهی مشخص و علت وجود این قواعد با کمک خبرگان تبیین شد. در این پژوهش ۲۶ متغیر در چهار دسته متغیرهای شخصی، متغیرهای آزمون، متغیرهای شغلی و متغیرهای وابسته به عملکرد و دو متغیر هدف شامل امتیاز ارتقا و ارزیابی عملکرد با شش نوع الگوریتم مطرح برای درخت تصمیم (QUEST، CART، C5، CHAID در حالت Towing و CART برای حالت Ordered) مورد بررسی قرار گرفتند.

الگوریتم درخت تصمیم‌گیری با انتخاب آزمونی شروع می‌شود که بهترین جداسازی را برای دسته‌ها انجام دهد. مهم‌ترین هدف از انجام دسته بندی، به دست آوردن مدلی برای پیش‌بینی است. بدین منظور از مجموعه‌ای از داده‌ها به نام «داده‌های آموزشی» که مجموعه‌ای از متغیرها و رکوردهاست استفاده می‌شود. در مراحل بعدی، همین کار برای گره‌های پایین‌تر با داده‌های کمتر صورت می‌گیرد تا بهترین قانون حاصل شود. سرانجام درخت آن قدر بزرگ می‌شود تا دیگر نتوان جداسازی بهتری برای داده‌های گره انجام داد.

چن-یانگ فان، پی-شو فان، تی-وای چن و شو-هاو چانگ (۲۰۱۲)، روند جایگزینی متخصصان حوزه تکنولوژی در شرکت‌های مطرح تایوانی را مورد بررسی قرار دادند. نتایج حاصل از این مطالعه نشان می‌دهد که جایگزینی کارکنان بیش‌تر به خاطر دلایلی چون عدم وجود وفاداری و رهبری و مدیریت مناسب صورت می‌گرفت. با استفاده از اعتبارسنجی متقابل مشخص شد که دقت خوشه‌بندی برابر با ۹۲٫۷٪ است. این مطالعه از تحلیل خوشه‌ای برای داده کاوی و یادگیری ماشین به منظور پیش‌بینی روند جایگزینی

متخصصان حوزه فناوری استفاده می‌کند که شامل شبکه عصبی مصنوعی هیبرید و تحلیل خوشه‌ای می‌شود که به عنوان یکی از الگوریتم معروف در شبکه عصبی و تحت عنوان نقشه خود سازمان‌دهی^۴ نیز شناخته می‌شوند. از این روش خوشه‌بندی هیبرید برای مطالعه مشخصات فردی روند خوشه‌های جایگزینی استفاده می‌شود. آن‌ها با استفاده از پرسش‌نامه اقدام به مطالعه دوره اوج جایگزینی که بعد از سال نوبی چینی روی می‌دهد، پرداختند و افراد را در گروه‌های سنی مختلف قرار دادند. در این مطالعه از نرم‌افزار Statistical V,۰ برای به کارگیری ابزارهای اجرایی شبکه عصبی خود سازمان دهنده و از روش خوشه‌بندی شبکه عصبی خود سازمان دهنده برای طبقه‌بندی اولیه گروه‌های خوشه‌بندی استفاده شده است. بعد از شناسایی بهترین خوشه‌ها، روش خوشه بندی k میانگین^۵ را در مورد تمام مجموعه داده‌ها اعمال کردند. این تحقیق بر ۲۸ متغیر متمرکز بود. آن‌ها برای افزایش بازدهی، تصمیم گرفتند از خطای مقدار مؤثر^۶ برای شاخص اصلی خوشه‌بندی استفاده کنند لذا مشخص شد که در صورت قرار داشتن دامنه خوشه‌بندی بین دو تا شش، بهترین نتیجه حاصل خواهد شد. بعد از استخراج گروه خوشه‌بندی اولیه از شبکه عصبی خود سازمان دهنده، از ANN^۷ و BPN^۸ نیز برای خوشه‌بندی تمام داده‌ها استفاده کردند. این رویکرد از طریق خوشه‌بندی دو مرحله ای SOM+BPN تمام داده‌ها را با موفقیت در چهار خوشه (گروه) قرار داد: خوشه گرایش بالا به جایگزینی، خوشه گرایش متوسط به جایگزینی، خوشه گرایش نسبتاً پایین به جایگزینی و خوشه پایین‌ترین گرایش به جایگزینی.

اسماعیلی، قوسی و سیکارودی^۹ (۱۳۹۳)، تحقیقی بر روی پیش‌بینی جابه‌جایی شغلی کارکنان شرکت خودروسازی اراک با رویکرد داده کاوی انجام داده‌اند. این شرکت تامین‌کننده قطعات خودرو است که از چهار بخش مونتاژ، پرس، جوشکاری و رنگ آمیزی

^۴ Self-Organizing-map

^۵ k-means

^۶ RMSE

^۷ Artificial neural network

^۸ Back propagation network

^۹ Amin Mohammad Esmaeeli-Sikaroudi

تشکیل شده است. آموزش کارکنان مجرب در این شرکت بسیار پرهزینه است. بنابراین داشتن یک استراتژی مناسب جهت حفظ بهترین کارکنان این شرکت ضروری است. فقدان چنین استراتژی‌ای یکی از مشکلات عمده این شرکت است. پژوهش مذکور با هدف پیدا کردن دلایل این مشکل و ساخت یک سیستم پشتیبانی تصمیم‌گیری^{۱۰} با استفاده از الگوریتم‌های داده کاوی^{۱۱} انجام گرفته که قادر است کارکنانی که احتمال ترک شغل آن‌ها بسیار زیاد است را تشخیص دهد. این سیستم پشتیبانی تصمیم‌گیری کمک زیادی در تصمیم‌گیری‌های مدیریتی خواهد داشت. آن‌ها در این تحقیق ویژگی‌های کارکنان از جمله سن، مهارت فنی و تجربه کاری را مدنظر قرار دادند و دقت و زمان محاسبه پیش‌بینی را با روش‌های مختلف داده کاوی مورد آزمایش قراردادند و سپس ویژگی‌های مؤثر را با روش Pearson's Chi-Square ارزیابی کردند که در نهایت متغیرهای دارای ارزش نه به عنوان ویژگی‌های مهم و مؤثر انتخاب شدند. سپس از مدل مبتنی بر شبکه عصبی پرسپترون چند لایه^{۱۲} با ۱۰ نورون، ۱ لایه پنهان و ۱۰۰۰ بار تکرار استفاده کردند. آن‌ها از نوع شبکه عصبی PNN^{۱۳} و SVM^{۱۴}، RF^{۱۵} و NB^{۱۶} و تکنیک CART^{۱۷} استفاده کردند. هم‌چنین برای به دست آوردن بهترین همسایگی از روش KNN^{۱۸}، برای استخراج قوانین از درخت تصمیم و برای ارزیابی نتایج از روش K-fold cross بهره جستند. روش K-fold cross داده‌ها را به K مجموعه به شکل تصادفی تقسیم می‌کند که یک قسمت برای آموزش و K-۱ قسمت برای تست داده‌ها استفاده می‌شود. در این تحقیق آن‌ها از روش ۱۰-fold cross validation استفاده کرده‌اند.

بسیاری از تحلیل‌های اولیه در زمینه گرایش به جایگزینی یا سایر مباحث تحقیقاتی

^{۱۰} Detection-support-System

^{۱۱} Data Mining

^{۱۲} Multi-Layer Perceptron

^{۱۳} Probabilistic neural network

^{۱۴} Support vector machine

^{۱۵} Random forest

^{۱۶} Naive Bayes

^{۱۷} Classification and regression tree

^{۱۸} K-nearest neighbor

مرتبط بیش تر از آمار توصیفی به همراه تحلیل همبستگی، تحلیل متغیرها، تحلیل واریانس و تحلیل رگرسیون، تحلیل رگرسیون با چند سلسله مراتب و تحلیل ساختاری با مدل سازی معادله استفاده می کردند (Fan, ۲۰۱۲).

ساموئل^{۱۹} و چیپونزا^{۲۰}، هائو^{۲۱}، جونگ^{۲۲} و ینهوی^{۲۳} و برخی دیگر از پژوهشگران از شبکه های انتشار بازگشتی، رگرسیون لجستیک و تحلیل داده های سبد بازار برای پیش بینی روند جایگزینی استفاده کرده اند ولی این دسته از تحلیل ها ایرادهای متعددی دارند. شبکه های انتشار بازگشتی بر پایگاه های داده های داخلی یک شرکت که از اطلاعات فرم های جایگزینی به دست می آیند، تمرکز دارند. این فرم ها حاوی نظرات کارکنان شرکت هستند که اکثر آن ها تمایلی به ارائه دلایل اصلی استعفای خود ندارند که خود اعتبار BPN را سؤال برانگیز می سازد. افزون بر این، سنجش اعتبار رگرسیون لجستیک برای پیش بینی امکان پذیر نیست. اگرچه ایده و نتایج تحقیقاتی حاصل از تحلیل داده های سبد بازار قابل پذیرش هستند، ولی موارد مشابه برای مقایسه وجود ندارد. از این رو تحلیل داده های سبد بازار نمی تواند کمک چندانی به پیش بینی روند جایگزینی کند.

روش شناسی

با بررسی روش های ارائه شده و همچنین با در نظر گرفتن نقاط ضعف و قوت روش های به کاررفته در تحقیقات گذشته، این نتیجه به دست آمده که سیستمی می تواند برای بررسی میزان گرایش به جایگزینی یا ترک سازمان مناسب باشد که درحین سادگی ازدقت بالایی برخوردار باشد و بتواند همزمان چندین معیار مهم را بررسی کند. روش تحقیق مورد استفاده در این پژوهش از نوع مطالعه موردی، میدانی، توصیفی و کاربردی است. در این پژوهش از ترکیب شبکه عصبی و الگوریتم ژنتیک چندهدفه با مرتب سازی نامغلوب برای طبقه بندی عوامل مؤثر بر میزان تمایل به ترک سازمان توسط کارکنان سازمان هدف

^{۱۹}Samuel

^{۲۰}Chypvnza

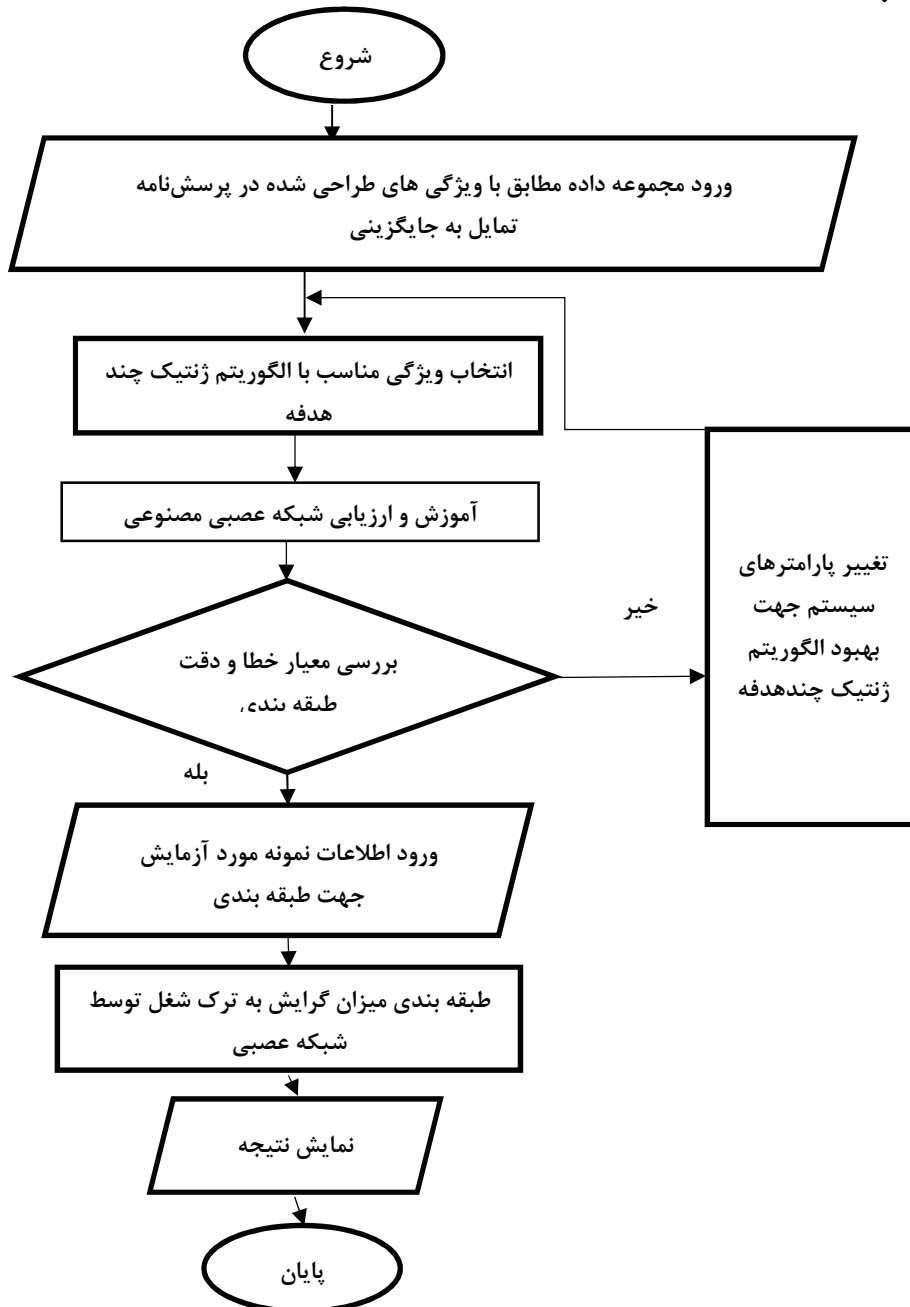
^{۲۱}Hao

^{۲۲}Jung

^{۲۳}Ynhvy

استفاده شده است. نحوه عملکرد مدل پیشنهادی در شکل (۱) نشان داده شده است. الگوریتم‌های ژنتیک به علت بررسی مجموعه‌ای از جواب‌های ممکن و هم‌چنین حساسیت کمتر نسبت به شکل خاصی از نقاط بهینه، برای بهینه‌سازی چندهدفه مناسب هستند. اهداف به صورت توابع ریاضی بوده و نمایانگر استفاده مناسب برای بهبود تصمیمات در بهینه‌سازی است. با توجه به عدم وابستگی به تجهیزات سخت افزاری خاص و هم‌چنین با در نظر گرفتن ویژگی‌های مشترک، این مدل به راحتی می‌تواند در سایر شرکت‌های وابسته به شرکت ملی نفت پیاده‌سازی و استفاده شود.

در این مقاله سعی بر این است که با استفاده از فنون الگوریتم‌های هوشمند در داده کاوی، روابط میان داده‌ها و تابع هدف شناسایی و عوامل تاثیرگذار در ترک سازمان استخراج گردد. بدین منظور ابتدا با مطالعه سوابق تحقیقاتی انجام شده در این زمینه، داده‌های جمع‌آوری شده از واحد بهبود روابط کار و امور اجتماعی که متولی این امر در سازمان هدف است و نظرات کارشناسان و خبرگان سازمان و هم‌چنین استفاده از پرسش‌نامه‌های معتبر در این زمینه، مؤلفه‌ها شناسایی و متناسب با مؤلفه‌های شناسایی شده، پرسش‌ها طراحی گردید. در نتیجه پرسش‌نامه‌ای با ۱۲ مؤلفه و ۶۵ گویه طراحی گردید. علاوه بر آن نه ویژگی مشخصات فردی/سازمانی نیز به آن اضافه شد.



شکل ۱: نحوه عملکرد مدل پیشنهادی

در طراحی پرسش‌نامه دقت لازم بعمل آمد تا پرسش‌ها از سادگی و وضوح کافی برخوردار باشند. برای حصول اطمینان از روایی پرسش‌نامه، هم از اعتبار محتوا و هم از اعتبار عاملی استفاده شد. برای سنجش اعتبار محتوای پرسش‌نامه، از نظرات متخصصان، و کارشناسان خبره سازمان هدف استفاده شد. بدین صورت که با انجام مصاحبه‌های مختلف و کسب نظرات افراد یاد شده، اصلاحات لازم به عمل آمده و بدین ترتیب اطمینان حاصل شد که پرسش‌نامه همان خصیصه مورد نظر پژوهشگران را می‌سنجد عنوان مؤلفه‌ها و تعداد گویه‌های هر مؤلفه در جدول (۲) نشان داده شده است. برای اندازه‌گیری قابلیت اعتماد یا پایایی ابزار تحقیق، از یک مطالعه مقدماتی یا پایلوت استفاده شده است. بدین منظور تعداد ۳۰ نفر از جامعه آماری انتخاب و پرسش‌نامه اولیه بین آن‌ها توزیع گردید و سپس داده‌های به دست آمده استخراج و به کمک نرم‌افزار SPSS آلفای کرونباخ اندازه‌گیری شد که می‌توان گفت با توجه به این که این مقدار باید بیش از ۰٫۷ به دست می‌آید، این پرسش‌نامه از قابلیت اعتماد بالایی برخوردار است.

نظرسنجی از ۱۲۰ نفر از کارکنان عملیاتی صورت گرفت، سپس انبار داده تشکیل شد. بر اساس فنون داده کاوی، داده‌های اولیه پیش پردازش و پاکسازی شدند. پردازش داده‌ها شامل تغییر و تبدیل داده‌ها، حذف و اصلاح داده‌ها جهت آماده‌سازی جداول ورودی و خروجی الگوریتم موردنظر است.

بنابراین جدول ورودی الگوریتم شامل ۱۲۰ رکورد با ۷۲ فیلد اطلاعاتی می‌باشد که مشتمل بر نتایج ۶۳ گویه و نه ویژگی مربوط به مشخصات فردی/سازمانی کارکنان است. با توجه به این که این پژوهش میزان گرایش به ترک شغل را در سه گروه (۱) گرایش پایین به ترک شغل، (۲) گرایش متوسط به ترک شغل و (۳) گرایش بالا به ترک شغل طبقه‌بندی می‌کند و با توجه به این که نمره‌گذاری گویه‌ها بر اساس درجه‌بندی پنج گانه لیکرت است، در جدول مربوط به تابع هدف به پاسخ‌های یک و دو ارزش یک، به پاسخ سه ارزش دو و به پاسخ‌های چهار و پنج ارزش سه در بانک اطلاعاتی داده شد. به همین منظور گویه ۶۴ «من اغلب به ترک شغل فکر می‌کنم.» به عنوان متغیر هدف در نظر گرفته شده است. گویه شماره ۵۷ «از کارم رضایت کلی دارم» نیز می‌تواند به عنوان دیگر خروجی میزان رضایت کارکنان را بسنجد.

جدول ۲: مؤلفه‌های پرسش‌نامه عوامل مؤثر بر ترک سازمان (تنظیم از پژوهشگران)

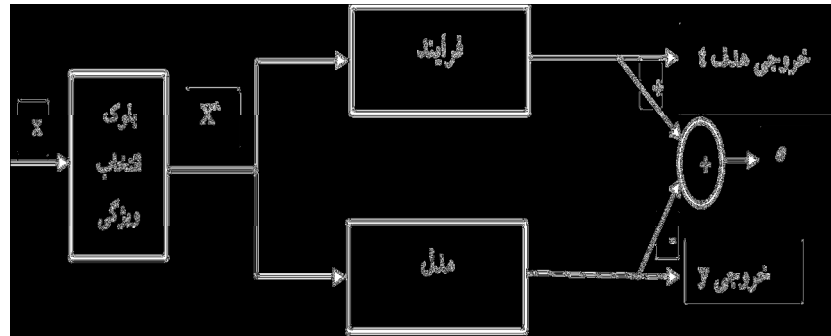
ردیف	عنوان مؤلفه	تعداد گویه‌ها
۱	احتمال پیشرفت	منظور از مؤلفه احتمال پیشرفت، بررسی ابتکار عمل، امکان یادگیری، تنوع کار و میزان ارتقای شغلی است که گویه‌های یک تا شش را به شکل مستقیم و لیکرت بررسی می‌کند.
۲	نفوذ در محل کار	مؤلفه نفوذ در محل کار میزان استقلال رأی و عمل و توانایی برنامه‌ریزی در نحوه انجام وظایف شغلی را به صورت مستقیم مورد بررسی قرار می‌دهد. این مؤلفه توسط چهار گویه به شکل مستقیم و لیکرت مورد بررسی قرار می‌گیرد.
۳	معنی‌دار بودن کار	هدف از مؤلفه معنی‌دار بودن کار، بررسی میزان رشد و توسعه فردی و میزان احساس لذت و پاسخ به نیازهای روحی از طریق کار می‌باشد. تعداد گویه‌های این مؤلفه هفت گویه به شکل مستقیم و لیکرت می‌باشد.
۴	حس در جمع و گروه بودن	منظور از این مؤلفه بررسی تجربه کارگروهی در محل کار، عضویت در گروه، بررسی روابط شغلی و حل تعارضات به شیوه مثبت می‌باشد. این مؤلفه نه گویه را به شکل مستقیم و لیکرت در بر دارد.
۵	تضاد کار و خانواده	مؤلفه تضاد کار و خانه شامل پنج گویه با مضمون میزان تداخل کار و زندگی و بررسی میزان رضایت شغلی است. این مؤلفه شامل پنج گویه به شکل مستقیم می‌باشد.
۶	تعهد شغلی	شامل سه گویه به صورت مستقیم می‌باشد.
۷	شفافیت شغلی	شامل سه گویه به صورت مستقیم می‌باشد.
۸	تعارض نقش	این مؤلفه شامل سه گویه با مضمون میزان مواجهه با درخواست‌های متفاوت کاری و تفاوت در وظایف محوله می‌باشد و دارای سه گویه مستقیم است.
۹	خواسته‌های کمی	سه گویه مستقیم
۱۰	رهبری منصفانه	سه گویه مستقیم
۱۱	فرسودگی شغلی	بررسی میزان فرسودگی شغلی و شامل ۶ گویه مستقیم

ردیف	عنوان مؤلفه	تعداد گویه ها
۱۲	رضایت شغلی	۱۱ گویه مستقیم
۱۳	تمایل به ترک	دو گویه مستقیم
۱۴	مشخصات فردی/ سازمانی	شامل جنسیت، وضعیت تاهل، میزان تحصیلات، رشته تحصیلی، سن، سابقه، نوع شغل، پایه سمت سازمانی و رتبه شخصی می‌باشد.

انتخاب ویژگی (متغیرهایی با بیش‌ترین اثربخشی) از مهم‌ترین روش‌ها و تکنیک‌ها جهت پیش پردازش داده برای کاوش داده‌هاست. به دلیل ارائه برنامه‌های جدید برای مقادیر زیاد داده مثل: داده کاوی، بازیابی اطلاعات چندرسانه‌ای و پردازش داده‌های حجیم پزشکی، محدود کردن تعداد ویژگی اهمیت زیادی پیدا کرده است. انتخاب ویژگی فرآیندی است که زیرمجموعه‌ای از ویژگی‌ها را بر اساس معیارهای بهینه‌سازی انتخاب می‌کند. روش‌های آماری سنتی به دو دلیل کارایی خود را از دست داده‌اند که یکی افزایش تعداد مشاهده و دیگری افزایش تعداد ویژگی‌های مربوط به یک مشاهده است. روش‌های انتخاب ویژگی روشی برای کاهش زمان محاسبه، بهبود پیش‌بینی عملکرد و فهم بهتر داده‌ها در یادگیری ماشین و برنامه‌های تشخیص الگوست. تکنیک‌های متعددی برای حذف ویژگی‌های بی‌ربط و زائدی که باری بر دوش وظایف هستند، ارائه شده است. تعریف ایده آل برای حل مسئله انتخاب ویژگی این است که زیرمجموعه‌ای با حداقل ویژگی تعیین شود به طوری که برای هدف مورد نظر اطلاعات لازم و کافی را داشته باشد. فرآیند انتخاب ویژگی شامل سه گام اساسی است: روش جستجو، ارزیابی زیرمجموعه و یک معیار توقف.

در این پژوهش ۷۲ متغیر مستقل و یک متغیر وابسته داریم. با توجه به شکل (۲) لازم است MSE^{۲۴} کمینه شود.

^{۲۴}Mean Squared Error



شکل ۲: محاسبه MSE

با توجه به رابطه (۱) در مسئله انتخاب ویژگی هدف ایناست که به جای این که y تابعی از x باشد، y را از زیرمجموعه‌ای از ورودی‌ها یعنی x^{\wedge} تقریب زد.

$$y = f(x) \sim f(x^{\wedge}) \quad (1)$$

طبق رابطه (۲) هدف کمینه کردن مقدار e است.

$$e = \min |t - f(x^{\wedge})| \quad (2)$$

و درعین حال باید علاوه بر کمینه بودن میزان خطا، تعداد ویژگی‌های مؤثر نیز کاهش یابد.

$$\min nf = |x^{\wedge}| \quad (3)$$

بنابراین درواقع این یک مسئله بهینه‌سازی چندهدفه است.

در مدل‌سازی با الگوریتم ژنتیک تک هدفه با استفاده از روش مجموع وزن‌دار مقادیر تابع هدف با یکدیگر ترکیب شده و به عنوان تابع برازندگی^{۲۵} در نظر گرفته می‌شود.

$$y = f(x) \sim f(x^{\wedge}) \quad (4)$$

β : هزینه افزودنی ویژگی

در مدل‌سازی توابع چندهدفه با روش مجموع وزن دار، β باید به گونه‌ای تعیین گردد که با کاهش تعداد ویژگی‌ها خطا کمینه باقی بماند. β در واقع نقطه تعادل است، با افزایش مقدار β ، nf کاهش می‌یابد و بالعکس.

^{۲۵} fitness

اشکالات موجود در روش‌های قدیمی:

۱. عدم پیدا کردن چندین جواب بهینه در طی یک بار اجرای الگوریتم
۲. عدم تضمین برای یافتن جواب‌های بهینه مختلف و متفاوت
۳. این روش را نمی‌توان برای مسائلی با متغیرهای گسسته و دارای چندین جواب بهینه به کار برد.

به علت این که در این مسئله هدف انتخاب ویژگی تعیین‌عوامل موثر همزمان با کاهش خطا است، دو معیار مورد بررسی قرار گرفته است.

$$\begin{cases} \min z_1 = E \\ \min z_2 = n_f \end{cases} \quad (5)$$

E : عملکرد مدل

N_f : تعداد ویژگی‌ها

درواقع این دو هدف نمی‌توانند همزمان کمینه شوند. طبق رابطه ۵ با افزایش تعداد ویژگی‌ها خطا کم می‌شود و با کاهش تعداد ویژگی‌ها، خطا افزایش می‌یابد. با تبدیل کردن این مسئله به دو تابع، این مسئله به مسئله بهینه‌سازی دو هدفه تبدیل شده است و با توجه به پارتو فرانت به دست آمده می‌توان مجموعه‌ای از پاسخ‌ها را به دست آورد و سپس با توجه به نحوه توزیع پاسخ‌ها و با توجه به هدف مسئله یکی از مقادیر به دست آمده را انتخاب کرد.

بعد از ورود مجموعه داده، مدل پیشنهادی با به کارگیری الگوریتم ژنتیک تک هدفه و چندهدفه وارد فرآیند انتخاب مجموعه‌ای از ویژگی‌ها که بیش‌ترین تأثیر را در طبقه‌بندی دارند، می‌شود. الگوریتم ژنتیک چندهدفه با اضافه شدن و عملگر ضروری به الگوریتم ژنتیک تک هدفه معمولی، به یک الگوریتم چندهدفه تبدیل شده است که به جای یافتن بهترین جواب، دسته‌ای از بهترین جواب‌ها را ارائه می‌دهد که با نام پارتو فرانت شناخته می‌شوند. این دو عملگر عبارتند از: (۱) عملگری که یک معیار برتری (رتبه) را بر اساس مرتب‌سازی نامغلوب به اعضای جمعیت اختصاص می‌دهد و (۲) عملگری که تنوع جواب را در میان جواب‌های با رتبه برابر نگه می‌دارد. به عبارت دیگر تفاوت این الگوریتم با الگوریتم ژنتیک تک هدفه، در مرحله انتخاب جمعیتی از میان والدین و فرزندان، برای ثابت نگه داشتن تعداد جمعیت در آغاز هر سیکل است لذا پیاده‌سازی الگوریتم ژنتیک چند هدفه با مرتب‌سازی نامغلوب برای این مسئله، حاصل از پژوهش است.

یافته‌ها

ابتدا در فضای جواب‌های ممکن برای ویژگی‌ها، تعدادی از زیرمجموعه‌های ویژگی به عنوان جمعیت اولیه انتخاب می‌شود. طول جمعیت برابر تعداد کل ویژگی‌های استفاده شده در پرسش‌نامه است که در این پژوهش برابر ۷۲ ویژگی است که به صورت (۰ و ۱) وارد الگوریتم می‌شود. ۱ نشان دهنده انتخاب یک ویژگی و ۰ نشان دهنده عدم انتخاب آن ویژگی است.

در مرحله دوم عدد برازش یا مقدار هزینه مربوط به هر کدام از کروموزوم‌ها با توجه به توابع هدف در نظر گرفته شده محاسبه می‌شود. تابع برازندگی استفاده شده در روش پیشنهادی، طبق رابطه (۵) معادل تعداد ویژگی‌ها و دقت کلی یا مقدار خطای کل حاصل از طبقه‌بندی با الگوریتم شبکه عصبی چندلایه و از نوع تشخیص الگو روی نمونه‌های آزمایشی است که برای زیرمجموعه ویژگی بهینه انتخاب شده، هر دو مقدار بایستی کمینه شود. به دلیل این که انتخاب ویژگی‌ها به صورت کاملاً تصادفی انجام می‌شود، می‌توان برای اطمینان بیشتر و بهبود نتیجه مقدار خطای به دست آمده برای هر کروموزوم، میانگین خطاهای به دست آمده را طی چندین بار اجرای شبکه عصبی محاسبه نمود و مقدار خطا را به دست آورد. با توجه به این که در هر مرحله، دو تابع هدف مدنظر است، لذا برای هر کروموزوم، بعد از هر انتخاب ویژگی، دو عدد به دست می‌آید که بیانگر میزان شایستگی کروموزوم نسبت به دو تابع هدف در نظر گرفته شده است.

پس از مشخص شدن عدد برازندگی برای تمام کروموزوم‌ها، به تعداد $n_{Crossover}$ تا از آن‌ها که در این پژوهش ۰٫۷ درصد جمعیت را به خود اختصاص داده است، به طور تصادفی برای تولید فرزندان انتخاب می‌شوند. سپس برای انجام عمل ادغام، باید دو تا از والدین در هر مرحله، جهت عمل ادغام انتخاب شوند. انتخاب در این مرحله نیز به این دلیل که هر کروموزوم دارای دو مقدار میزان شایستگی است، با الگوریتم ژنتیک تک هدفه متفاوت است. در این مرحله، از تابع Dominates باینری برای انتخاب هر والد استفاده می‌شود. این روش به این صورت انجام می‌شود که برای انتخاب هر والد، ابتدا دو کروموزوم α و β به طور تصادفی و با تولید رشته‌ای از صفر و یک‌ها انتخاب می‌شوند. طول این رشته برابر تعداد کل ویژگی‌هاست و صفر نشان دهنده عدم انتخاب ویژگی و یک نشان دهنده انتخاب آن

ویژگی است. سپس باقاعده زیر، کروموزوم x به عنوان والد انتخاب می‌شود به شرط آنکه رتبه x از رتبه y کمتر باشد و اگر رتبه هر دو برابر باشد از عملگر فاصله ازدحامی استفاده می‌شود.

$$X \leq_{\text{distance}} Y \quad (۶)$$

اگر فاصله ازدحامی x بیش‌تر از فاصله ازدحامی y باشد، این عمل به تعداد والدین مورد نیاز جهت ادغام تکرار می‌شود. همچنین برای عمل ادغام، ابتدا به طور تصادفی، یکی از دو روش ادغام تک نقطه‌ای^{۲۶} یا ادغام دو نقطه‌ای^{۲۷} انتخاب می‌شود، سپس نقطه یا نقاط مورد نظر نیز از بین ژن اول تا هفتاد و دوم به طور تصادفی انتخاب شده و دو کروموزوم از محل یا محل‌های مورد نظر با یکدیگر ادغام می‌شوند. ادغام بر روی یک زوج کروموزوم از نسل مولد عمل کرده و یک زوج کروموزوم جدید تولید می‌کند.

در ادغام تک نقطه‌ای، یک موقعیت تصادفی بین دو ژن در نظر گرفته می‌شود، سپس تمامی ژن‌های طرف راست یا طرف چپ این موقعیت در کروموزوم‌های والد با یکدیگر جابه جا می‌شوند تا کروموزوم‌های جدید به دست آیند. در ادغام دو نقطه‌ای، دو موقعیت به صورت تصادفی انتخاب می‌شود و تمامی ژن‌های بین این دو موقعیت در کروموزوم‌های والد با یکدیگر جابه جا می‌شوند. شایان ذکر است که ادغام معمولاً بر روی همه زوج کروموزوم‌های انتخاب شده برای جفت‌گیری به کار برده نمی‌شود. معمولاً احتمال ادغام برای هر زوج کروموزوم بین $0/6$ تا $0/95$ در نظر گرفته می‌شود که به این عدد نرخ ادغام^{۲۸} یا احتمال ادغام^{۲۹} گفته می‌شود و با P_c نمایش داده می‌شود. در صورتی که بر روی یک زوج کروموزوم عمل ادغام صورت نگیرد، فرزندان با تکرار نمودن والدین تولید می‌شوند.

در مرحله بعدی از بین فرزندان حاصل از ادغام، تعدادی برای عمل جهش انتخاب می‌شوند. عمل جهش برای فرار از به دام افتادن الگوریتم در بهینه محلی بسیار مفید است. عمل جهش به این ترتیب انجام می‌شود که ابتدا کروموزومی برای عمل جهش انتخاب

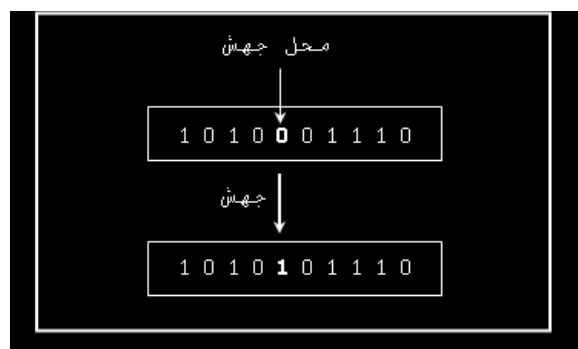
^{۲۶} One-point Crossover

^{۲۷} Two-point Crossover

^{۲۸} Crossover Rate

^{۲۹} Crossover Probability

می‌گردد سپس به تصادف یکی از ژن‌ها انتخاب می‌شود. این عملگر یک ژن از یک کروموزوم را به طور تصادفی انتخاب نموده و سپس محتوای آن ژن را تغییر می‌دهد. اگر ژن از جنس اعداد دودویی باشد، آن را به وارونش تبدیل می‌کند و چنانچه متعلق به یک مجموعه باشد، مقدار یا عنصر دیگری از آن مجموعه را به جای آن ژن قرار می‌دهد. در شکل ۳ چگونگی جهش یافتن پنجمین ژن یک کروموزوم نشان داده شده است. هریک از این ژن‌ها مربوط به یک ویژگی مشخص است و عمل جهش بر روی یکی از ویژگی‌ها انجام می‌شود. در صورت عدم برقراری شرایط عمل جهش مجدداً تکرار می‌گردد. البته لازم است که نرخ جهش به درستی انتخاب شود. در این پژوهش نرخ جهش ۰,۴ در نظر گرفته شده است. پس از اتمام عمل جهش، کروموزوم‌های تولید شده به عنوان نسل جدید شناخته شده و برای دور بعد اجرای الگوریتم ارسال می‌شوند.



شکل ۳: یک کروموزوم قبل و بعد از عمل جهش

بعد از انجام مراحل ادغام و جهش، فرزندان تولیدشده وارد شبکه عصبی شده و میزان برازندگی آن‌ها نیز برای دو تابع در نظر گرفته شده محاسبه می‌شود. اگر شرایط خاتمه فراهم شود الگوریتم پایان می‌یابد و در غیر این صورت وارد سیکل بعدی می‌شویم. در این مرحله از بین جمعیت والدین و فرزندان باید تعدادی حذف شوند تا تعداد جمعیت اصلی ثابت بماند. لذا در این مرحله ابتدا براساس رتبه و سپس براساس فاصله ازدحامی که در بخش قبل توضیح داده شد، جمعیت کل با استفاده از تابع SortPopulation مرتب می‌شوند و سپس تعدادی از این جمعیت که بر اساس رتبه و شاخص فاصله ازدحامی، وضعیت بهتری

دارند انتخاب و بقیه به میزانی که جمعیت اصلی ثابت بماند، حذف می‌شوند و الگوریتم به همان شکل قبل ادامه می‌یابد.

در پایان اجرای الگوریتم در هر بار اجرا، پارتو فرانت حاصل از جواب‌های بهینه توابع هدف بر اساس شاخص‌های تعیین شده در نمودار قابل مشاهده خواهند بود. نقاط قابل مشاهده در نمودار پارتو فرانت به معنای این است که این نقاط توسط هیچ نقطه دیگری مغلوب نشده‌اند، اما سایر نقاط که در این مجموعه نیستند قطعاً حداقل یک بار توسط سایر اعضای جمعیت مغلوب شده‌اند. در انتها نیز مقادیر هدف مورد بررسی قرار گرفته و بهینه‌ترین پاسخ از نظر ویژگی‌ها و خطای کمینه انتخاب می‌شود.

تحلیل جاری در محیط متلب و با استفاده از شبکه عصبی پرسپترون چندلایه و از نوع تشخیص الگو و الگوریتم ژنتیک تک هدفه و چندهدفه انجام شده است. جهت ارزیابی سیستم پیشنهادی سه آزمون مورد بررسی قرار گرفته است.

در آزمون مرحله اول از یک شبکه عصبی چندلایه و از نوع تشخیص الگو استفاده شد. در این آزمون ۷۰ درصد کل مجموعه داده برای آموزش، ۱۵ درصد برای ارزیابی و ۱۵ درصد باقی مانده برای تست شبکه عصبی استفاده شد. هدف از این آزمون دستیابی به بالاترین دقت طبقه‌بندی بود. نتایج آزمون‌ها در جدول (۳) نشان داده شده است. با بررسی میانگین میزان موفقیت در تمامی مراحل آموزش، ارزیابی و آزمون مشخص می‌شود که بهترین دقت با به کارگیری ۹۰ نرون برای لایه میانی و دقت ۸۸٫۹ برای مجموعه تست است.

جدول ۳: نتایج آزمون مرحله اول (منبع: یافته‌های پژوهش)

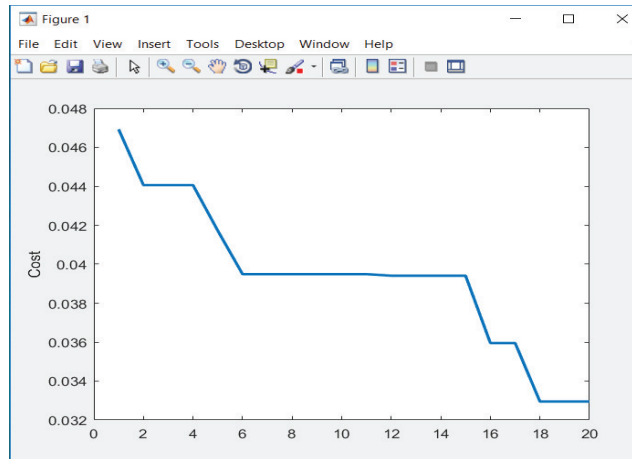
شماره آزمون	تعداد نورون‌های لایه میانی	درصد موفقیت	تعداد ویژگی‌ها
۱-۱	۵۰	۶۶٫۷	۷۲
۱-۲	۶۰	۷۲٫۲	۷۲
۱-۳	۹۰	۸۸٫۹	۷۲
۱-۴	۱۰۰	۸۳٫۳	۷۲

مرحله دوم: هدف از این آزمون دستیابی به تعداد ویژگی‌های کمینه با حداقل خطای ممکن با استفاده از شبکه عصبی و الگوریتم ژنتیک تک هدفه‌است. تعداد جمعیت اولیه ۵۰، تعداد تکرار ۵۰۰ و تعداد اجراها برای تولید اعداد تصادفی به منظور انتخاب ویژگی در هر مرحله ۱۰۰ در نظر گرفته شد. در این آزمایش می‌بایست مقدار پارامتر β (هزینه افزونگی ویژگی) به درستی تنظیم شود زیرا با افزایش مقدار آن، تعداد ویژگی‌ها کاهش و با کاهش آن تعداد ویژگی‌ها افزایش می‌یابد. در این آزمایش مقدار β ۰,۵۲ در نظر گرفته شد. تعداد نورون‌های لایه میانی در ۶ مرحله تغییر داده شد و در نهایت نیز میزان دقت طبقه‌بندی برای مجموعه تست و هم‌چنین میانگین دقت طبقه‌بندی مراحل آموزش، ارزیابی و آزمون جمع‌آوری گردید. نتایج به دست آمده از این آزمایش در جدول (۴) قابل مشاهده است. نتایج به دست آمده از این آزمون نشان می‌دهد که پس از ۱۰۰ تکرار تعداد ۳۴ ویژگی با خطای ۰/۰۳۳ بهترین پاسخ است.

جدول ۴: نتایج آزمون مرحله دوم بخش اول (منبع: یافته‌های پژوهش)

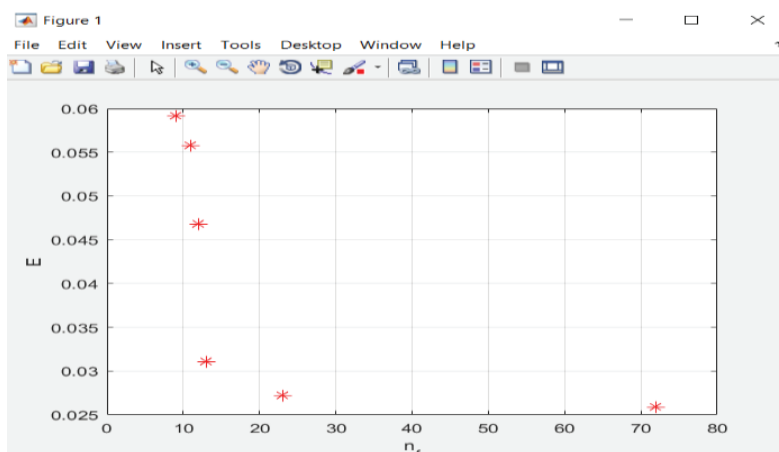
میزان خطا Z	میزان عملکرد E	درصد ویژگی‌های انتخاب شده	تعداد ویژگی‌های انتخاب شده
۰,۰۳۳	۰,۰۲۶	۰,۴۷	۳۴

همان‌گونه که اطلاعات جدول (۴) نشان می‌دهد تعداد ۳۴ عضو به عنوان ویژگی‌هایی که بیش‌ترین تأثیر را در طبقه‌بندی دارند و با خطای ۰,۳۳ بهترین پاسخ ممکن از نظر تعداد ویژگی و خطای کمینه است. شکل (۴) کارایی الگوریتم ژنتیک مقدار Cost را نشان می‌دهد که حاصل مجموع وزنی درصد خطا و تعداد ویژگی هاست.



شکل ۴: نتیجه آزمون انتخاب ویژگی با شبکه عصبی و الگوریتم ژنتیک

مرحله سوم: در این مرحله تنظیمات اولیه جهت انتخاب ویژگی کمینه با حداقل خطای ممکن با استفاده از شبکه عصبی و الگوریتم ژنتیک چندهدفه انجام شد. تعداد جمعیت اولیه ۵۰، تعداد تکرار ۱۰۰ و تعداد اجراها برای تولید اعداد تصادفی به منظور انتخاب ویژگی در هر مرحله ۵۰ در نظر گرفته شد. هرچه تعداد اجراهای تولید اعداد تصادفی بیشتر باشد دقت افزایش خواهد یافت. در هر اجرا مجموعه داده تولیدشده وارد الگوریتم ژنتیک چندهدفه گردیده و فرآیند انتخاب ویژگی انجام شده است. سپس با توجه به تعداد ویژگی‌های انتخاب شده و همچنین با توجه به نتیجه به دست آمده در مرحله اول آزمون‌ها، این ویژگی‌ها به یک شبکه عصبی چندلایه با تعداد نورون‌های لایه ورودی برابر تعداد ویژگی‌های انتخاب شده و تعداد نورون‌های لایه میانی برابر $1/3$ و ویژگی‌های انتخابی و ۳ نورون برای لایه خروجی داده شد. در نهایت میزان دقت طبقه‌بندی برای مجموعه تست و همچنین میانگین دقت طبقه‌بندی مراحل آموزش، ارزیابی و آزمون جمع‌آوری شد. شکل (۵) پارتو فرانت حاصل از نتیجه آزمون مرحله سوم را نشان می‌دهد. همچنین در شکل (۴) نتایج حاصل از آزمون مرحله سوم نشان داده شده است.



شکل ۵: پارتو فرانت حاصل از نتیجه آزمون مرحله سوم

در پارتو فرانت به دست آمده ۶ حالت قابل مشاهده است.



شکل ۶: ماتریس درهم ریختگی آزمون مرحله سوم

نتایج به دست آمده از به کارگیری شبکه عصبی به همراه الگوریتم ژنتیک چندهدفه در جدول (۵) آورده شده است. بررسی جدول (۵) نشان می‌دهد با کاهش تعداد ویژگی‌ها، میزان خطا افزایش می‌یابد. با توجه به توابع هدف موردنظر در این پژوهش تعداد ۲۳ عضو به عنوان ویژگی‌هایی که بیشترین تأثیر را در طبقه‌بندی دارند و با خطای ۰,۰۲۷ بهترین پاسخ ممکن از نظر تعداد ویژگی و خطای کمینه است. جدول (۶) نشان دهنده شماره ویژگی‌هایی است که در این مرحله انتخاب شده‌اند.

جدول ۵: نتایج آزمون به دست آمده از شبکه عصبی و الگوریتم ژنتیک چندهدفه (منبع:

یافته‌های پژوهش)

آزمون	۱	۲	۳	۴	۵	۶
تعداد ویژگی‌های منتخب	۷۲	۹	۲۳	۱۲	۱۳	۱۱
سرعت پیش‌بینی	۱	۰,۱۲۵	۰,۳۱۹	۰,۱۶	۰,۱۸	۰,۱۵
خطای پیش‌بینی	۰,۰۲۵	۰,۰۵۹	۰,۰۲۷	۰,۰۴۶	۰,۰۳۱	۰,۰۵۵
MSE	۰	۰,۳۳	۰,۰۲	۰,۱۹	۰,۰۶	۰,۳

جدول ۶: ویژگی‌های انتخاب شده توسط الگوریتم ژنتیک چندهدفه در آزمون سوم (منبع:

یافته‌های پژوهش)

شماره ویژگی	عنوان ویژگی	شماره ویژگی	عنوان ویژگی	شماره ویژگی	عنوان ویژگی	شماره ویژگی	عنوان ویژگی
۱	ابتکار عمل	۱۱	لذت از کار	۳۷	تعریف دقیق وظایف	۵۹	فرصت پیشرفت
۲	آموزش	۱۳	پاسخ به نیازهای روحی	۴۰	درخواست‌های ناسازگار	۶۲	رضایت از مافوق
۴	تنوع در کار	۱۹	رشد فردی	۴۴	توزیع منصفانه وظایف	۶۶	جنسیت
۶	ارتقای شغلی	۲۷	تداخل کار و زندگی	۴۹	خستگی عاطفی	۷۰	سن
۷	نوع وظایف	۲۹	رسیدگی به خانواده	۵۵	خستگی شدید	۷۱	سابقه
۸	نحوه انجام وظایف	۳۳	رضایت شغلی	۵۷	رضایت کلی از کار		

در ادامه آزمون قبلی مجموعه ویژگی‌های انتخاب شده از الگوریتم ژنتیک چندهدفه وارد شبکه عصبی شد و با شرایط مساوی، یعنی در ۶ مرحله و با تعداد متفاوت نورون‌های لایه‌های میانی تکرار شد. تعداد نورون‌های لایه ورودی برابر ۲۳ عدد و معادل مجموعه ویژگی‌های حاصل از الگوریتم ژنتیک چندهدفه تعیین گردید. هم‌چنین تعداد ۳ نورون برای لایه خروجی معادل تعداد کلاس‌های طبقه‌بندی در نظر گرفته شد. بهترین نتایج به دست آمده از این ۶ آزمون در جدول (۷) نمایش داده شده است.

جدول ۷: نتایج آزمون تست شبکه عصبی و الگوریتم ژنتیک چندهدفه

آزمون	۱	۲	۳	۴	۵	۶
تعداد ویژگی‌ها	۲۳	۲۳	۲۳	۲۳	۲۳	۲۳
تعداد لایه‌های پنهان	۱۰	۲۰	۴۰	۷۰	۸۰	۹۰
دقت پیش‌بینی (تست)	۷۶,۵	۷۷,۸	۷۷,۸	۶۶,۷	۸۸,۲	۷۰,۶
دقت پیش‌بینی (میانگین کل)	۸۷,۲	۸۴,۶	۸۲,۹	۸۸,۰	۹۳,۲	۹۴,۰

در این مرحله نیز مجموعه داده به صورت تصادفی به سه مجموعه داده با اندازه‌های ۷۰، ۱۵ و ۱۵ درصد به ترتیب به مجموعه داده آموزشی، ارزیابی و آزمون تقسیم شده است. بررسی جدول (۷) نشان می‌دهد که با توجه به نتایج حاصل از به کارگیری الگوریتم ژنتیک چندهدفه با کاهش تعداد ویژگی‌ها به ۲۳ عدد با حداقل خطا به میزان ۰,۰۲ درصد، دقت طبقه‌بندی مجموعه تست به ۸۸,۲ و میانگین دقت مراحل آموزش، ارزیابی و تست نسبت به آزمایش ۱ افزایش یافته و به ۹۴,۰ درصد رسیده است. نتیجه این که با استفاده از الگوریتم ژنتیک چندهدفه علاوه بر کاهش ویژگی‌ها به ۲۳ عدد با کمترین درصد خطا، می‌توانیم میانگین دقت طبقه‌بندی را نیز بهبود دهیم.

بحث و نتیجه‌گیری

مدیران و محققان، کناره‌گیری از کار را پرهزینه، مشکل آفرین و بیش‌تر همراه با عواقب تهدیدکننده برای کارایی سازمان می‌دانند. ترک سازمان می‌تواند نشان دهنده از دست

دادن هزینه زیادی در زمینه تامین، آموزش، سرمایه‌گذاری اجتماعی، جایگزینی نیروی انسانی و صرف هزینه‌های غیرمستقیم بسیاری باشد و به دنبال آن تاثیراتی منفی نیز در روحیه و بازدهی کار افرادی که در سازمان باقیمانده‌اند، داشته باشد. لذا سازمان‌هایی که بتوانند دلایل و عوامل مؤثر در ترک سازمان کارکنان را درک نمایند، می‌توانند با به کارگیری اطلاعات صحیح و با پیش‌بینی‌های لازم و به موقع، قبل از کناره‌گیری کارکنان با آگاهی از میزان تمایلات به ترک خدمت و با استفاده از سیاست‌ها و روش‌های مطمئن، مدیریت مؤثر را در فرآیند حفظ و نگهداری نیروی انسانی کارآمد به کار گیرند. مطالعات قبلی در حوزه ترک خدمت کارکنان بیش‌تر با هدف بررسی دلایل ترک سازمان انجام می‌شدند و به ندرت میزان گرایش به آن را پیش‌بینی می‌کردند. از این رو، استفاده از آن‌ها برای اجتناب از ترک سازمان توسط کارکنان کلیدی به شکلی فوری و مؤثر امکان‌پذیر نبود. این پژوهش با استفاده از طبقه‌بندی در داده کاوی و استفاده از الگوریتم‌های هوشمند به این مشکل پرداخته است. این مطالعه توانست با استفاده از الگوریتم شبکه عصبی پرسپترون میزان گرایش به ترک سازمان توسط کارکنان را با دقت ۸۸٫۹ درصد در مرحله تست شبکه عصبی پیش‌بینی کند. هم‌چنین با استفاده از ویژگی‌های استخراج شده توسط الگوریتم ژنتیک چندهدفه از میان ۱۲ عامل تاثیرگذار در ترک سازمان که موردبررسی قرار داده شد، مشخص شد که کارکنان به عوامل مشخصی چون احتمال پیشرفت در کار، معنی‌دار بودن کار، تداخل کار با خانه و زندگی، فرسودگی شغلی و رهبری منصفانه اهمیت می‌دهند. تأثیر مؤلفه احتمال پیشرفت بیش از سایر مؤلفه هاست و از اهمیت بیش‌تری برای کارکنان برخوردار می‌باشد.

یکی از راهبردهای توسعه منابع انسانی و ایجاد فرصت پیشرفت، گردش شغلی است. گردش شغلی سیستمی است که در آن کارکنان برای دوره‌های کوتاهی تحت آموزش و توانمندسازی قرار می‌گیرند تا به تجربیات کاری جدیدی دست یابند. گردش شغلی باعث افزایش رضایتمندی، انگیزه کارکنان و کاهش فرسودگی شغلی می‌گردد. بنابراین پیشنهاد می‌شود برنامه جامع و منسجمی توسط مدیران و متولیان امر در جهت توسعه منابع انسانی و ایجاد رضایتمندی به ویژه برای کارکنان عملیاتی تدوین گردد. هم‌چنین با توجه به وجود شرایط سخت کاری در مناطق عملیاتی، تامین شرایط مناسب زندگی و توجه به ایجاد

تسهیلات رفاهی قطعاً می‌تواند در افزایش میزان رضایتمندی کارکنان مؤثر باشد. با توجه به نتایج به دست آمده پیشنهادهایی در دو سطح سازمانی و فنی به صورت زیر ارائه شده است که عبارتند از:

۱. با توجه به تایید رابطه میان مؤلفه‌های ابتکار عمل، نحوه انجام وظایف در سازمان، رضایت شغلی و همچنین لذت از کار با کاهش تمایل کارکنان به ترک خدمت، پیشنهاد می‌شود شرایطی ایجاد شود که کارکنان بتوانند در انجام وظایف محوله ابتکار عمل بیشتری داشته باشند و از ایده‌های شخصی و البته مفید و مورد تایید خود برای ایجاد تغییرات مناسب در شغلشان بهره ببرند. این کار می‌تواند به وابستگی بیشتر میان فرد و شغل منجر شده و با افزایش سطح تعهد فرد به شغل و سازمان، از تمایل فرد به ترک شغل بکاهد.

۲. با توجه به تایید رابطه میان آموزش و رشد فردی و فرصت پیشرفت با کاهش تمایل کارکنان به ترک خدمت، پیشنهاد می‌شود بر اهمیت امر آموزش در سازمان تاکید ویژه‌ای کرده و از این مقوله برای ارتقای توانمندی‌های فنی و حتی روانی افراد استفاده نماییم. شخص می‌تواند از طریق شرکت در دوره‌های آموزشی، تسلط بیشتری بر شغل داشته باشد، با خطرات احتمالی شغل آشنا شده و شیوه‌های پیشگیری از آن‌ها را فراگیرد و به نظر نگارنده، این امر می‌تواند مسبب تمایل بیشتر او به ماندن در شغل فعلی و چشم پوشی از ترک آن گردد.

۳. با توجه به تایید رابطه میان ارتقا و کاهش تمایل کارکنان به ترک خدمت، پیشنهاد می‌شود قوانین مناسبی برای ارتقای کارکنان تدوین شده و در صورت وجود چنین قوانینی نیز، نظارت بیشتری بر اجرای آن‌ها اعمال گردد تا بتوان با ایجاد امید به رشد و شکوفایی در نزد کارکنان، آنان را به ادامه کار در سازمان کنونی ترغیب نمود.

۴. با توجه به تایید رابطه میان نوع وظایف و همچنین پاسخ به نیازهای روحی با کاهش تمایل کارکنان به ترک خدمت، پیشنهاد می‌شود به هنگام استخدام کارکنان، از طریق انجام مصاحبه‌ها و آزمون‌های شخصیت‌شناسی، افراد را برای مشاغل و مسئولیت‌های متناسب با روحیات و شخصیت‌های آنان به کار بگیریم و اگر

- اکنون نیز افرادی هستند که در محیطی نامتناسب مشغول به کارند، می‌توان از طریق جابه‌جایی به چیدمان صحیح‌تری از آنان دست یافت.
۵. با توجه به تایید رابطه میان تداخل کار و زندگی و همچنین رسیدگی به خانواده با کاهش تمایل کارکنان به ترک خدمت، پیشنهاد می‌شود که تا حد امکان تمهیداتی برای پیشگیری از تداخل کار و زندگی کارکنان اندیشیده شود. یکی از راهکارهای این امر استفاده از ساعت‌های کاری انعطاف‌پذیر است که طی آن افراد بتوانند با توجه به شرایط خاص زندگی خود در ساعت‌های توافق شده، در سازمان حضور یافته و در کنار آن به زندگی خود نیز بپردازند.
۶. با توجه به تایید رابطه میان تعریف دقیق وظایف و همچنین توزیع منصفانه وظایف با کاهش تمایل کارکنان به ترک خدمت، پیشنهاد می‌شود وظایف سازمانی به درستی و بر اساس شرح وظایف، بین افراد مختلف سازمان تقسیم شده و وظیفه هر کس به درستی به او اعلام و انتظار سازمان از فرد به او تفهیم گردد.
۷. با توجه به تایید رابطه میان رضایت کلی از کار و همچنین رضایت از مافوق با کاهش تمایل کارکنان به ترک خدمت، پیشنهاد می‌شود با انجام آموزش‌هایی از قبیل مهارت‌های ارتباطی، رابطه میان سرپرستان و زیردستان مورد تقویت قرار گیرد تا هم سرپرستان بتوانند حمایت خود را به زیردستان نشان دهند و هم زیردستان بتوانند در ادراک این حمایت موفق باشند.
۸. به خاطر ملاحظات زمانی و هزینه‌ای، این مطالعه را می‌توان از لحاظ دقت پیش‌بینی و کاربرد تحقیقاتی با استفاده از روش‌های دیگر بهبود بخشید. از این رو پیشنهاد می‌کنیم از سایر الگوریتم‌های تکاملی برای انتخاب ویژگی استفاده شود و نتایج حاصل از آن با نتایج به دست آمده از الگوریتم ژنتیک تک هدفه و چند هدفه مقایسه گردد. با این اوصاف، می‌توان از نظریه فازی و تصمیم‌گیری چند معیاری برای پیدا کردن سایر عوامل مؤثر در ترک سازمان استفاده کرد.

منابع

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، صص ۲۲-۳۰.

میرحسینی، سید حسن، حسینی سیدمهدی، گندمکار مجید، قنبری مهدی (۱۳۹۱)، بازآرایی شبکه توزیع با استفاده از الگوریتم ژنتیک چندهدفه با مرتب‌سازی نامغلوب نسخه ۲ (NSGA-II)، اولین کنفرانس ایده‌های نو در مهندسی برق، دانشگاه آزاد اسلامی واحد خوراسگان، اصفهان، آذر ۱۳۹۱.

Adejuwon, A; Mosavi, A (۲۰۱۰) Domain Driven Data Mining – Application to Business, IJCSI International Journal of Computer Science Issues, July ۲۰۱۰, vol. ۷, no. ۴, pp. ۲-۵, ۱۶۹۴-۰۷۸۴.

Beardwell, J. and Claydon, T. (۲۰۰۷). Human resource management: A contemporary approach”, ۲۰۰۷.

Chin-Yuan, Fan; Pei-Shu, Fan; Te-Yi, Chan (۲۰۱۲), Using hybrid data mining and machine learning clustering analysis to predict the turnover rate for technology professionals”, Expert Syst. Appl, ۲۰۱۲, vol. ۳۹, no. ۱۰, pp. ۸۸۴۴-۸۸۵۱

Esmiaeeli Sikaroudi, Amir Mohammad ; Ghousi, Rouzbeh; Esmiaeeli Sikaroudi, Ali (۲۰۱۵), A data mining approach to employee turnover prediction (case study: Arak automotive parts manufacturing), JISE, ۲۰۱۵, Vol. ۸, No. ۴

Goldberg, D., Addison-Wesely, (۱۹۸۹), Genetic Algorithm in search, Optimization and machine Learning ,

Khosravi, I. and Mohammad-Beigi, M. (۲۰۱۴). Multiple classifier systems for hyperspectral remote sensing data classification, J. Indian Soc. Remote ,

Maghsoudi, Y; Collins, M. (۲۰۱۳), Radarsat-۲ polarimetric SAR data for boreal forest classification using SVM and a wrapper feature selector, IEEE J. Sel.

- Niknam, T; Azadfarsani, E ; Jabbari,M. (۲۰۱۲). A new hybridevolutionary algorithm based on new fuzzy adaptive PSO and NMAgorithms for Distribution Feeder ReconfigurationEnergy Conversion and Management, February ۲۰۱۲; ۷-۱۶
- Olamaei, J. et al (۲۰۱۰), A hybridevolutionary algorithm based on ACO and SA for distribution feederreconfiguration, IEEE, ۲۶۵-۲۶۹
- Xu,R. et al, (۲۰۱۰) An estimation of distribution algorithm based portfolio selection approach, Appl. Artif.