Archive of SID

داده کاوی نسبت های مالی جهت طبقه بندی شرکت های ورشکسته وسالم با استفاده از مدل مخفی مارکوف

الهه سلیمانپور[۱]، مهناز آهنگری [۲]

[۱]دانشگاه فردوسی ، دانشکده مهندسی کامپیوتر،کارشناس ارشد هوش مصنوعی ،e.soleymanpour@yahoo.com

[۲]دانشگاه آزاد اسلامی ، واحد نیشابور ، کارشناس ارشد حسابداری ،mahnaz.ahangary@gmail.com

چکیده - یکی از مهمترین موضوع های مطرح شده در زمینه مدیریت مالی، این است که سرمایه گذاران فرصت های مطلوب سرمایه گذاری را از فرصت های نا مطلوب تشخیص دهند و منابع در اختیار را در فرصت های مناسب، سرمایه گذاری کنند. از مهم ترین روشهایی که می توان با استفاده از آن، به بهره گیری مناسب از فرصت های سرمایه گذاری و همچنین جلو گیری از هدر رفتن منابع کمک کرد، پیش بینی ورشکستگی شرکت ها می باشد. محققان امروزه با استفاده از تکنیک های مختلف داده کاوی وبازشناسایی الگو ،توانسته اند به درصد بالایی از تفکیک مناسب شرکت ها از لحاظ ورشکستگی بپردازند.دراین تحقیق با داده کاوی نسبت های مالی شرکت ها وبا استفاده از مدل مخفی مارکوف(*HMM*) ،به تعیین شرکت های ورشکسته وسالم پرداخته شده است.نتایج نشان می دهد مدل مخفی مارکوف،پس از آموزش توانسته است شرکت ها را بادقت۹۲.۲۳ درصد به طور صحیح تفکیک نماید.

کلید واژه‌ها- داده کاوی ، ورشکستگی،درخت تصمیم (CART) مدل مخفی مارکوف (HMM) ،شرکت های بورسی.

۱. مقدمه

یکی از مسائلی که می تواند به نحوه تصمیم گیری سرمایه گذاران کمک نماید ، وجود ابزارها و مدل های مناسب برای ارزیابی شرایط مالی و وضعیت شرکت ها است. یکی از ابزارهای مورد استفاده برای تصمیم به سرمایه گذاری ، مدل های پیش بینی ورشکستگی است. سرمایه گذاران همواره در پی آن هستند تا با پیش بینی امکان ورشکستگی یک شرکت از ریسک سوخت شدن اصل و فرع سرمایه خود مطلع شده و از آن جلوگیری نمایند. از این رو آن ها به دنبال یافتن روش هایی هستند تا بتوانند ورشکستگی شرکت ها را تخمین بزنند. به همین دلیل، پیش بینی ورشکستگی از دور دست به عنوان مساله ای اساسی در شرکت ها مطرح بوده و به تبع اهمیت آن ، مطالعات بسیار زیادی نیز برای دستیابی به بهترین مدل پیش بینی ورشکستگی انجام گرفته است. یکی از اولین مطالعات در این زمینه توسط ویناکر و اسمیت در سال ۱۹۳۵ ، درباره ی کفایت نسبت های مالی به عنوان عامل پیش گویی مشکلات مالی بود. این دو با مطالعه ی خود به این نتیجه رسیدند که، نسبت های مالی شاخص مطلوبی برای پیش بینی درماندگی هستند. در سال ۱۹۴۲، مروین نسبت های متعددی را در ۶ سال اول عمر شرکت های دارای تداوم فعالیت و شرکت هایی که فعالیتشان متوقف شده بود بررسی نمود و به این نتیجه رسید که ۳ نسبت سرمایه در گردش به کل دارایی ها ، خالص ارزش ویژه به بدهی و نسبت جاری، در پیش بینی موثرتر از بقیه می باشد. بیور در سال ۱۹۶۶ در مطالعه ی تجربی خود که با استفاده از تکنیک های آماری قوی تری نسبت به مروین انجام شده بود، دریافت که ارزش هر نسبت در میزان اعتبار طبقه بندی شرکت ها در گروه شرکت های درمانده مالی و سالم است و میزان خطای طبقه بندی کمتر، نشان دهنده ی ارزش بالای هر نسبت می باشد. از جمله مطالعات مطرح در پیش بینی ورشکستگی، مطالعات آلتمن و ارائه ی مدلی به نام وی بوده است[۱]و[۲]. در سال ۱۹۹۳، وی اقدام به اصلاح و بازنگری در مدل خود نمود و این بار به جای استفاده از پنج نسبت مالی از چهار نسبت مالی استفاده نمود تا قدرت پیش بینی مدل وی برای موسسات غیرتولیدی نیز افزایش یافته و اطلاعات دقیق تری برای پیش بینی درماندگی شرکت ها در اختیار استفاده کنندگان قرار دهد. پس از گذشت تقریبا چهل سال، مدل آلتمن کماکان توسط بسیاری از پژوهشگران وتحلیل گران جهت پیش بینی درماندگی مورد

www.SID.ir

Archive of SID

استفاده قرار می گیرد. پس از وی تحقیقات دیگری نیز برای پیش بینی درماندگی و ورشکستگی با استفاده از نسبت‌های مالی انجام گرفت که هر چند در آن تحقیقات به واسطه‌ی بهبود در جمع آوری داده هاو روشهای آماری، روش اصلاح شده ولی نتایج بدست آمده توسط محققان، به صورت قابل ملاحظه‌ای تغییر نیافت(به عنوان مثال: ادوارد دیکین، ۱۹۷۲ [۳]؛اهلسون،۱۹۸۰[۴]؛مک وگریستین،۲۰۰۰[۵]). در پی افزایش شرکت ها و موسسات و افزایش پیچیدگی روابط اقتصادی، محققان در پی آن هستند تا با استفاده از روش های جدیدتر، قدرت پیش بینی ورشکستگی شرکت‌ها و موسسات را افزایش دهند. در سال های اخیر، با یاری گرفتن از الگوریتم‌های معروف و روش‌های کامپیوتری نظیر شبکه های عصبی مصنوعی ، درخت تصمیم‌گیری، نزدیکترین همسایه و الگوریتم مورچگان ، دست به پیش بینی درماندگی مالی شرکت‌ها زده و یافته‌های آن‌ها حاکی از آنست که استفاده از نسبت های مالی در این الگوها در پیش بینی درماندگی مالی و ورشکستگی سودمند است[۶]و[۷]. اما نتایج پژوهش های مختلف حاکی از متفاوت بودن دقت هر یک از الگوها در پیش بینی است. استفاده از الگوهای مختلف به طور همزمان و در یک محیط اقتصادی می تواند قضاوت ما را در مورد توانایی هریک از الگوهای پیشگو بهبود بخشد. تکنیک داده کاوی می تواند در این مورد مفید باشد. داده کاوی تکنیک و ابزار بسیار متداولی است که امروزه در زمینه های مختلفی کاربرد پیدا کرده است. تشخیص و پیش بینی درماندگی شرکت‌ها، یکی از زمینه های اقتصادی پرکاربرد داده کاوی محسوب می شود که در سال های اخیر تحقیقات و مطالعات زیادی پیرامون آن انجام شده است(به عنوان مثال: شاه ومرتزا،۲۰۰۰[۸]؛مین ولی ،۲۰۰۵ [۹]،مین و جانگ ،۲۰۰۸[۱۰]). داده کاوی از جمله ابزارهایی است که قادر است با بهره گیری از تکنیک های مختلف آماری ، اطلاعات پنهان و پیشگویانه پایگاههای داده بزرگ و حجیم را استخراج نماید[۲]. در واقع داده کاوی فرایند تبدیل اطلاعات به دانش می باشد، به طوری که با استفاده از مدلی که از داده کاوی حاصل می گردد سرمایه گذاران ، اعتباردهندگان وسایر استفاده کنندگان می توانند فرصت های مطلوب را از نامطلوب تشخیص داده و منابعشان را در فرصت های مناسب سرمایه گذاری کنند. این پژوهش سعی دارد با کمک مدل مخفی مارکوف به پیش بینی شرکت های ورشکسته و سالم بپردازد. همچنین در این پژوهش از الگوریتم درخت تصمیم CART جهت انتخاب ویژگی های بهینه برای ورودی مدل مارکوف استفاده شده است که در بخش ۲.۲ به شرح آن به طور کامل پرداخته ایم.

۲. درک حوزه کاربرد

دراین قسمت به درک وشناخت حوزه کاربرد پرداخته می شود:

۲.۱. بیان علت انتخاب مساله

از انجا که بازار سرمایه درایران جوان است و به عنوان یک بازار کارا مطرح نمی باشد،ارائه راهکارهای علمی می تواند مبنایی قرار گیرد تا اطلاعات منتشرشده از سوی بورس اوراق بهادار تهران در تصمیم گیری های صاحبان سرمایه وسرمایه گذاران ،بالقوه مفید باشد.لذا باید ابزارهای مفیدی جهت تصمیم گیری،دراختیار کاربران بازار سرمایه قرارگیرد.یکی از راه هایی که می توان با استفاده از آن به بهره گیری مناسب از فرصت های سرمایه گذاری وجلوگیری از به هدررفتن منابع پرداخت ،پیش بینی درماندگی مالی یاورشکستگی است.به این ترتیب که اولا،باارائه هشدارهای لازم به شرکت های درمعرض ورشکستگی می توان آنها را نسبت به وقوع درماندگی مالی هوشیار کردرتادست به اقدام های مقتضی بزنند و دوم این که،سرمایه گذاران فرصت های مطلوب سرمایه گذاری را از فرصت های نامطلوب تشخیص داده ومنابعشان را در فرصت های مناسب سرمایه گذاری کنند.

۲.۲.جمع آوری داده ها

شرکت های پذیرفته شده دربورس اوراق بهادار تهران به عنوان جامعه مورد مطالعه انتخاب و نمونه ی مورد استفاده متشکل از دو گروه شرکت های ورشکسته وشرکت های غیرورشکسته می باشد.باتوجه به اینکه شرکت های ورشکسته دراین پژوهش شرکت هایی هستند که مشمول ماده ۱۴۱ قانون تجارت شده اند [۲]،ابتدا لیستی از شرکت های پذیرفته شده در بورس اوراق بهادار تهران بین سالهای ۱۳۸۶ تا۱۳۸۸ تهیه می شودو نسبت سود یازیان انباشته به سرمایه از طریق صورت های مالی این شرکت ها محاسبه می گردد،درصورتی که زیان انباشته بیش از نصف سرمایه شرکت باشد،شرکت ورشکسته(ماده۱۴۱ قانون تجارت)ودرغیراین صورت سالم است.باتوجه به دردسترس بودن اطلاعات،تعداد۷۲شرکت ورشکسته تعیین وسپس تعداد ۷۲ شرکت غیرورشکسته نیز درهمین بازه زمانی براساس روش نمونه گیری تصادفی انتخاب می شوند.

۲.۳.انتخاب ویژگی

متغیرهای پیش بینی کننده ورشکستگی به کاررفته دراین تحقیق،نسبت های مالی شرکت ها هستند. درابتدا برای انتخاب نسبت های مالی مورد استفاده در آزمون ، به نتایج پژوهش های صورت گرفته در کشورمان مراجعه نمودیم(به عنوان مثال سلیمانی امیری۱۳۸۱ [۱۱]و موسوی و طبرستانی ۱۳۸۸ [۱۲])

www.SID.ir

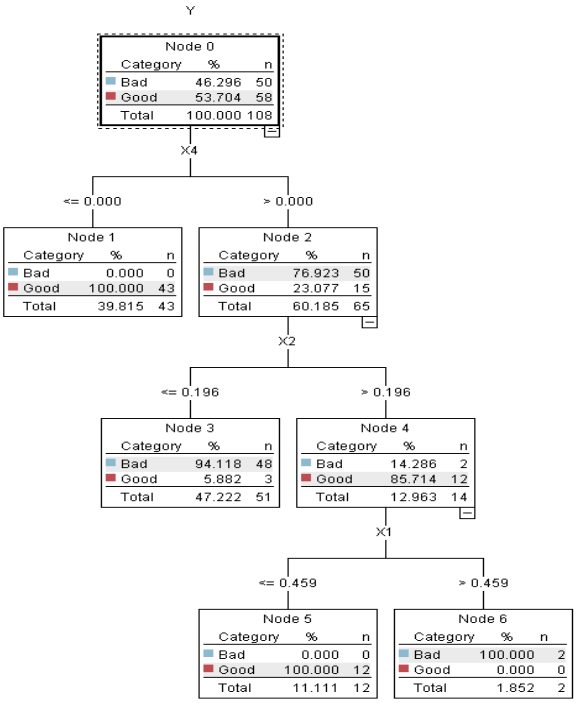
Archive of SID

با توجه به نتایج پژوهش های گذشته، از بین ۲۲ نسبت های مالی در پیش بینی در ماندگی مالی و ورشکستگی،۵ نسبت به عنوان نسبت های مالی موثرانتخاب ومورد استفاده قرار گرفتند . متغیرهای این تحقیق به شرح جدول شماره (۱) می باشد

متغیر	تعریف
X_1	نسبت سرمایه در گردش به مجموع داراییها
X_2	نسبت سود انباشته به مجموع داراییها
X_3	نسبت سود قبل از بهره و مالیات به مجموع داراییها
X_4	نسبت حقوق صاحبان سهام به مجموع بدهیها
X_5	نسبت فروش خالص به مجموع داراییها

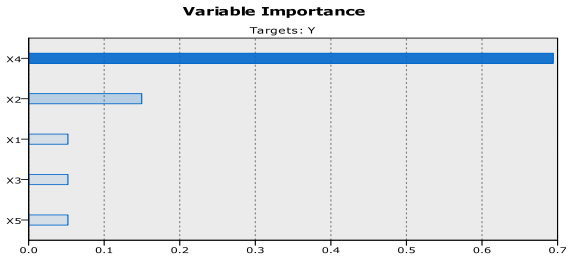
جدول شماره(۱): متغیرهای مسئله

به منظور بهینه سازی ویژگی های انتخابی جهت استفاده در مدل مارکوف، الگوریتم درخت تصمیم CART اعمال شده است درخت تصمیم گیری یکی از ابزارهای قوی و متداول برای دسته بندی و پیش بینی می باشد. درخت تصمیم گیری به تولید قانون می پردازد. الگوریتم درخت تصمیم گیری با انتخاب آزمونی شروع می شود که بهترین جداسازی را برای دسته ها انجام دهد.مهم ترین هدف از انجام دسته بندی ،به دست آوردن مدلی برای پیش بینی می باشد.بدین منظور از مجموعه ای از داده ها به نام "نسبت های مالی " که مجموعه ای از متغیرها ورکوردها است استفاده می کنیم .درمراحل بعدی همین کار برای گره های پایین تر با داده های کمتر صورت می گیرد تا بهترین قانون حاصل شود.سرانجام درخت آنقدر بزرگ می شود تادیگر نتوان جداسازی بهتری برای داده های گره انجام داد.در این مرحله باید اثربخشی درخت ایجاد شده اندازه گیری شود. برای این کار از یک مجموعه ی رکوردها یا داده های آزمایشی استفاده می شود که متفاوت با داده های اولیه ای که درخت را ایجاد کرده اند، می باشد. معیاری که اندازه گیری می شود عبارت است از:درصد داده هایی که به طور صحیح دسته بندی می شوند و دسته ی پیش بینی شده با دسته ی واقعی آن ها یکسان است[۱۳]. نتایج بکارگیری این الگوریتم بر روی ۵ نسبت مالی انتخاب شده در پژوهش های قبلی به شرح شکل شماره(۱) می باشد.



شکل شماره (۱) : نسبت های مالی بهینه حاصل از الگوریتم درخت تصمیم CART

اهمیت متغیرها (نسبت های مالی) از لحاظ این الگوریتم جهت تفکیک شرکت ها در دو طبقه ورشکسته و سالم به عنوان یک دانش مناسب می تواند مورد استفاده قرار گیرد (شکل 2).



شکل شماره (۲) : اهمیت متغیرها از لحاظ الگوریتم CART

۳ .مدل مخفی مارکوف

مدل پنهان مارکوف ،یک مدل آماری است که در آن سیستم مدل شده به صورت یک فرآیند مارکوف با حالت‌های مشاهده نشده (پنهان) فرض می‌شود. یک مدل پنهان مارکوف می‌تواند به عنوان ساده‌ترین شبکه بیزی پویا در نظر گرفته شود. در مدل عادی مارکوف، حالت به طور مستقیم توسط ناظر قابل

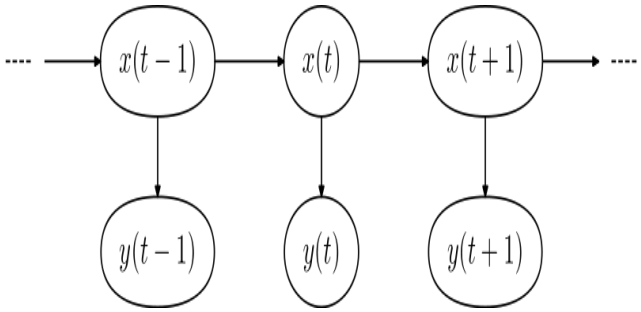
www.SID.ir

Archive of SID

مشاهده‌است و بنابراین احتمال‌های انتقال بین حالت‌ها تنها پارامترهای موجود است. در یک مدل پنهان مارکوف، حالت به طور مستقیم قابل مشاهده نیست، اما خروجی، بسته به حالت، قابل مشاهده‌است. هر حالت یک توزیع احتمال روی سمبل‌های خروجی ممکن دارد. بنابراین دنباله‌ی سمبل‌های تولید شده توسط یک مدل پنهان مارکوف اطلاعاتی درباره‌ی دنباله‌ی حالت‌ها می‌دهد[14] .

3.1. معماری مدل مخفی مارکوف

شکل شماره (۳) معماری کلی یک نمونه HMM را نشان می‌دهد.هر شکل بیضی یک متغیر تصادفی است که می‌تواند هر عددی از مقادیر را انتخاب کند.متغیر تصادفی x(t) یک حالت پنهان در زمان t و متغیر تصادفی y(t) مشاهده در زمانt است.فلش‌ها به معنای وابستگی‌های شرطی می‌باشند.از شکل مشخص است که توزیع احتمال شرطی متغیر پنهان x(t)، در همه ی زمان هایt مقداری را برایx ارائه می‌دهد که فقط به مقدار متغیر پنهان x(t-1)وابسته‌است:مقادیر در زمان‌های (t − 2)و قبل تر از آن هیچ اثری ندارند.این مشخصه ی مارکو نامیده می‌شود. به طور مشابه، مقدار متغیر مشاهده ایy(t) تنها به مقدار متغیر پنهانx(t) (هر دو در زمان خاص t) بستگی دارد. در حالت استاندارد مدل پنهان مارکوف که در اینجا در نظر گرفته شده‌است، فضای حالت متغیرهای پنهان گسسته‌است.در حالی که متغیرهای مشاهده‌ای می‌توانند گسسته یا پیوسته (از توزیع گوسین) باشند.در مدل پنهان مارکوف دو نوع پارامتر وجود دارد:احتمال جابجایی‌ها (بین حالات) و احتمال خروجی‌ها (یا مشاهدات).احتمال جابجایی نحوه ی انتقال از حالت (t-1)به t را کنترل می‌کند.فضای حالت پنهان شامل N مقدار ممکن برای حالات است.متغیر پنهان در زمان t می‌تواند هر یک از این مقادیر را اختیار کند.احتمال جابجایی احتمال این است که در زمان t در حالت k (یکی از حالات ممکن) و در زمان (t+1) در حالت k1 باشیم.بنابرین در کل N2احتمال جابجایی وجود دارد.(مجموع احتمالات جابجایی از یک حالت به تمام حالتهای دیگر ۱ است)احتمال خروجی، احتمال رخ دادن هر یک از اعضای مجموعه ی مشاهده‌ای را برای هر حالت پنهان ممکنه مشخص می‌سازد که می‌تواند از یک توزیع احتمال پیروی کند.تعداد اعضای مجموعه ی مشاهده‌ای بستگی به طبیعت متغیر مشاهده‌ای دارد.اگر تعداد اعضای مجموعه ی متغیرهای مشاهده‌ای برابر M باشد، بنابراین مجموع تعداد احتمالات خروجی NM می‌باشد[14].



شکل شماره(۳) : معماری کلی یک نمونه HMM

3.2 .مدلسازی داده ها با استفاده از مدل مخفی مارکوف

کلیه عملیات پردازشی برروی داده ها بااستفاده از نرم افزار MATLAB7.2 انجام شده است.جهت پیاده سازی سیستم شناخت ازمدل های مخفی مارکوف استفاده شده است. HMM یک مدل اماری است که امکان تحلیل سری های زمانی چندمتغیره غیرایستان را بامدل کردن توابع چگالی احتمال داده های چندمتغیره و احتمال انتقال بین حالات پایدار فراهم می کند.HMM شامل تعداد محدودی حالت گسسته است که قابل مشاهده نیست (حالت های مخفی)و احتمال هرحالت دریک زمان خاص فقط به حالت قبلی وابسته است؛وبا انتقال بین حالت ها به صورت تصادفی یکی از حالات بردار مشاهدات(برداریژگی ها) تولید می شود.به منظور استفاده از HMM به عنوان طبقه بندی کننده ازهرحالت مخفی مدل یک برداریژگی نتیجه می شود.برای تعیین مدل تعیین احتمالات زیر لازم است[۱۵]:

- .ماتریس احتمال انتقال بین حالت های مدل
- .ماتریس احتمال مشاهدات
- .برداراحتمال اولیه

برای این که مدل قادر به شناسایی باشد ابتدالازم است آموزش داده شود.آموزش مدل به مفهوم،تخمین احتمالات انتقال بین هرحالت(State) وبردارمشاهدات است.برای آموزش از الگوریتم EM(Expectation Maximization) استفاده شده است. در این تحقیق از ساختار mHMM استفاده شده است.در این ساختار ابتدا بردار مشاهدات توسط تعدادی mixture Gaussian تخمین

www.SID.ir

Archive of SID

زده می شودو سپس با الگوریتم K-means مدل می شود. جهت آموزش مدل مارکوف ، از داده های مربوط به سالهای t-1وt-2 و جهت آزمون مدل از داده های مربوط به سال t(سال tبرای شرکت های ورشکسته ،سال ورشکستگی و برای شرکت های سالم ،سال قرار گرفتن درنمونه است) استفاده شده است.

۴ .نتیجه‌گیری

در تحقیق حاضر با بکارگیری تکنیک داده کاوی مدل مخفی مارکوف در زمینه ی طبقه بندی ،سعی بر آن شد که صحت پیش بینی درماندگی مالی مورد بررسی قرار گیرد چون اعتقاد ما بر آنست که روش های سنتی و قدیمی با توجه به پیشرفت فناوری اطلاعات و نیازهای استفاده کنندگان ، کارایی خود را از دست داده ، ولازم است هم راستا با این تغییر و تحول تغییراتی در مدلها ی مورد استفاده نیز ایجاد گردد.با توجه به نتایج فوق می توان گفت مدل مارکوف با صحت متوسط ۹۲.۲۳درصد در سال t قادر به پیش بینی شرکت های ورشکسته وسالم می باشد. انچه قابل مشاهده است اینست که در مجموع الگوریتم های آماری ،الگوریتم های ضعیف تری نسبت به الگوریتم های هوش مصنوعی می باشند و قدرت پیش بینی شرکت های ورشکسته و سالم در بین الگوریتم های هوش مصنوعی بالاتر میباشد.نکته قابل ذکر دیگر آنست که شاخص نسبت مالی نسبت حقوق صاحبان سهام به مجموع بدهیها باتوجه به نتایج حاصله از بکارگیری الگوریتم درخت تصمیم CART، نسبت به سایر متغیرهای مالی مورد استفاده از اهمیت فوق العاده بیشتری در تعیین وضعیت شرکت ها به لحاظ ورشکستگی برخوردار می باشد.

مراجع

[1] E.I ,Altman,".Financial ratios , discriminant analysis and prediction of corporate bankruptcy ",Journal of Finance, , 23,589-60, 1968 .

[2] E.I ,Altman. R .G,Haldeman , P , Narayanan." ZETA ANALYSIS ,a new model to identify bankruptcy risk of corporations" ,Journal of Banking Finance , 1,29-54.1977 .

[3],E.B ,Deakin,"A discriminant analysis of predictors of business failure",journal of Accounting Research,10,167-179,1972.

[4] I Ohlson,"Financial ratios and the probabilistic prediction of bankruptcy ",journal of Accounting Research ,18(1),109-131, .,1980

[5]T.E, Mckee,.,& ,M, Greenstein,"predicting bankruptcy using recursive partitioning and a realistically proportioned data set ".journal .of forecasting,2000.

[6] Zne-jungle.,chou-Yuanl,"A hybrid search Algorithm with heuristics for Resource Allocation problem",information sciences ,2005.

[7]L.M ,lopuki. J.W,,Doherty" The Determinants of professional fee in large Bankruptcy Reorganization cases" ,journal of Empirical legal studies , vol .1,Issue 1,March ,111-141,2004.

[8]J ,Shah,.and M, Murtaza "A neural network based clustering procedure for bankruptcy prediction "American business Review,2000.

[9]J.H, Min, Y.C..,Lee, ,"Bankruptcy prediction using support vector machine(SVM) with optimal choice of kernel function parameters" Expert Systems with Application:,28,603-614, 2005 .

[10]J.H ,Min , C,Jong ."A binary classification method for bankruptcy prediction",Expert systems with Application ,2008.

[۱۱] غلامرضا امیری سلیمانی"بررسی شاخص های پیش بینی کننده ورشکستگی در شرایط محیطی ایران"،رساله دکتری حسابداری،دانشگاه تهران،دانشکده مدیریت،۱۳۸۱.

[۱۲]محمود موسوی و محمدرضا طبرستانی ؛"پیش یبنی درماندگی مالی با استفاده از تحلیل پوششی داده ها "،تحقیقات حسابداری،شماره دوم ،تابستان ۱۳۸۸.

[۱۳]عادل آذر ،پرویز احمدی ومحمد وحید سیط؛"طراحی مدل انتخاب نیروی انسانی بارویکرد داده کاوی "نشریه مدیریت فناوری اطلاعات ،دوره ۲،شماره ۴،بهاروتابستان ۱۳۸۹.

[۱۴]Satish L, Gururaj BI "Use of hidden Markov models for partial discharge pattern classification". IEEE Transactions on Dielectrics and Electrical Insulation. (April 2003).

[۱۵] Ghahramani, Zoubin; Jordan, Michael I. "Factorial Hidden Markov Models". Machine Learning (1997),29 (2/3): 245–273.

www.SID.ir