Journal of Financial Management Strategy
Vol. 3, No. 9
Summer 2015

Alzahra University- Faculty of
Social Sciences and Economics
Received: 2015/04/04
Accepted: 2015/06/17

The Use of Feature Selection Method (HARC) in Predicting Financial Distress in Tehran Stock Exchange

Maedeh Taj Mazinani[1]
Saeid Fallah Poor[2]
Saeid Bajlan[3]

Abstract

One of the most important issues in financial distress prediction is the selection of predicting variables. The present research in contrast to previous works provides a new approach to feature selection using the classification of financial ratios based on financial semantics and a combination of statistical methods with meta-heuristic algorithms. To this end, 34 financial ratios based on the audited financial statements are collected for distressed manufacturing companies according to paragraph 141 of Iran Trade Law and healthy companies listed in Tehran stock exchange (TSE) sampled randomly between 2001 and 2011 for one and two years prior to distress. Then, using t-test and genetic algorithm respectively, best ratios from the primary feature set are selected and support vector machine is applied to predict financial distress. The experimental results showed that the proposed HARC method outperforms logistic regression and Altman model significantly for one and two years prior to distress in predicting financial distress.

Keywords: Feature selection, financial distress prediction, financial ratios, genetic algorithm

JEL: C38‹C53‹C88‹G15‹G17‹G33

1- M.A. in Management, Tehran University, (Corresponding Author), maedetaj@ut.ac.ir.
2- Assistant Professor, Management Department, Tehran University, sfallahpour@gmail.com
3- Assistant Professor, Management Department, Tehran University,saeedbajalan@gmail.com

http://jfm.alzahra.ac.ir/

راهبرد مدیریت مالی

سال سوم، شماره نهم

تابستان ۱۳۹۴

صص ۱۰۶-۷۷

دانشگاه الزهرا (س)

دانشکده علوم اجتماعی و اقتصادی

تاریخ دریافت: ۱۳۹۴/۰۱/۱۵

تاریخ تصویب: ۱۳۹۴/۰۳/۲۷

کاربرد روش انتخاب ویژگی هارک (HARC) در پیش‌بینی درماندگی مالی شرکت‌ها در بورس اوراق بهادار تهران

مائده تاج مزینانی[۱]، سعید فلاح‌پور[۲] و سعید باجلان[۳]

چکیده

یکی از مسائل مهم در پیش‌بینی درماندگی مالی، انتخاب متغیرهای پیش‌بین می‌باشد. پژوهش پیش رو به نشان رویکردی جدید برای انتخاب ویژگی با استفاده از دسته‌بندی نسبت‌های مالی بر مبنای مفاهیم مالی و ترکیب روش‌های آماری با الگوریتم‌های فراابتکاری می‌پردازد. بدین منظور ۳۴ نسبت مالی برای شرکت‌های تولیدی درمانده براساس ماده ۱۴۱ قانون تجارت و به همان تعداد شرکت سالم به صورت تصادفی از شرکت‌های پذیرفته شده در بورس اوراق بهادار تهران در بازه زمانی ۱۳۸۰ تا۱۳۹۰با استفاده از صورت‌های مالی حسابرسی شده برای یک و دو سال قبل از درماندگی جمع‌آوری شده است. سپس با استفاده از آزمون آماری تی و الگوریتم ژنتیک، بهترین نسبت‌ها انتخاب و با استفاده از ماشین بردار پشتیبان، پیش‌بینی درماندگی مالی انجام شده است. - نتایج بدست آمده از پژوهش حاکی از آن است که روش پیشنهادی هارک در یک و دو سال پیش از وقوع درماندگی به طور معناداری در پیش‌بینی درماندگی مالی نسبت به رگرسیون لجستیک و مدل آلتمن از عملکرد بهتری برخوردار است.

واژه های کلیدی: انتخاب ویژگی، پیش بینی درماندگی مالی، نسبت‌های مالی، الگوریتم ژنتیک

طبقه‌بندی موضوعی: G33،G17،G15،C88،C53،C38

---

۱. کارشناس ارشد مهندسی مالی دانشگاه تهران، نویسنده مسئول، maedetaj@ut.ac.ir

۲. استادیار دانشکده مدیریت دانشگاه تهران، sfallahpour@gmail.com

۳. استادیار دانشکده مدیریت دانشگاه تهران، saeedbajalan@gmail.com

مقدمه

درماندگی مالی به شرایطی گفته می‌شود که شرکت نتواند به تعهدات خود در قبال تامین کنندگان مالی خود عمل کند یا در عمل به این تعهدات دچار مشکل باشد. گاهی درماندگی مالی به ورشکستگی می‌انجامد. پیش‌بینی درماندگی مالی یا ورشکستگی از مسائل مهم در اقتصاد هر کشوری به شمار می‌آید. زیرا هزینه‌های زیادی را بر شرکت، اعتبار دهندگان و در سطحی کلان بر کل اقتصاد تحمیل می‌کند. در این زمینه پژوهش‌های بسیاری انجام شده است اما آنچه که کمتر مورد توجه قرارگرفته است، پارامترهای مورد بررسی برای پیش‌بینی درماندگی مالی است. در این پژوهش با رویکردی جدید به مسئله انتخاب ویژگی در پیش‌بینی درماندگی مالی پرداخته شده است.

در ادبیات مالی، تعاریف مختلفی برای درماندگی مالی داده شده است. گردون[1] در یکی از نخستین بررسی‌های آکادمیک بر روی تئوری درماندگی مالی، آن را به عنوان کاهش قدرت سودآوری شرکت تعریف می‌کند که احتمال عدم توانایی بازپرداخت بهره و اصل بدهی را افزایش می‌دهد(گوردون،۱۹۷۱).

ویتاکر[2]، درماندگی مالی را وضعیتی در نظر می‌گیرد که در آن، جریان‌های نقدی شرکت از مجموع هزینه‌های مربوط به بدهی بلندمدت کمتر است(ویتاکر، ۱۹۹۹). حالت‌های دیگری از درماندگی مالی عبارتند از زمانی که جریان‌های نقدی شرکت برای بازپرداخت اصل و فرع بدهی ناکافی باشد و همچنین زمانی که ارزش ویژه شرکت عددی منفی شود.

در این پژوهش، ملاک درماندگی مالی شرکت‌ها، مشمول ماده ۱۴۱ بودن قانون تجارت می‌باشد. در قسمتی از این ماده آمده است: اگر بر اثر زیان‌های وارده، دست کم نصف سرمایه شرکت از میان برود، هیأت مدیره مکلف است بلافاصله مجمع عمومی فوق العاده صاحبان سهام را دعوت نماید تا موضوع انحلال یا بقای شرکت مورد شور و رأی واقع شود. هرگاه مجمع مزبور به انحلال شرکت رأی ندهد، باید در همان جلسه و با رعایت مقررات ماده ۶ این قانون، سرمایه شرکت را به مبلغ سرمایه موجود کاهش دهد. باید توجه کرد که درماندگی مالی به طور حتم سبب ورشکستگی نمی‌شود، بلکه اعلام ورشکستگی یکی از پیامدهای آن است که به طور معمول آخرین راه حل است. شرکت‌ها برای مقابله با درماندگی مالی می‌توانند ساختاردهی دوباره انجام دهند.

---

1. Gordon
2. Whitaker

پیش‌بینی مالی و به‌ویژه بررسی تاثیر انتخاب ویژگی در پیش بینی درماندگی مالی باعث می‌شود تا ضعف‌های مدل‌های قبلی بر طرف شده و بتوان با مدلی بهتر و مقرون به صرفه‌تر پیش‌بینی را انجام داده و با نشان دادن هشدارهای لازم، شرکت‌ها را نسبت به وقوع آن هشیار کرد تا آن‌ها با توجه به این هشدارها دست به اقدام‌های مقتضی بزنند و همچنین از این طریق سرمایه‌گذاران و اعتباردهندگان فرصت های مطلوب سرمایه گذاری را از فرصت‌های نامطلوب تشخیص می‌دهند و منابع‌شان را در فرصت‌های مناسب سرمایه‌گذاری می کنند.

هدف پژوهش پیش رو نیز دادن مدلی برای انتخاب بهترین متغیرها در پیش‌بینی درماندگی مالی است تا از این طریق بتوان به تصمیم‌گیری درست ذینفعان و حداقل کردن ریسک ورشکستگی در سازمان‌ها کمک کرد. همچنین با دادن هشدارهای به موقع می‌توان از هدررفتن سرمایه‌های ملی جلوگیری کرد.

مروری بر مبانی نظری و پیشینه پژوهش

۱- انتخاب ویژگی

انتخاب ویژگی یکی از دو عامل مهم و تاثیرگذار در عملکرد یک مدل پیش‌بینی برای موضوع طبقه‌بندی می‌باشد. تعداد متغیرهایی که برای هر مشاهده باید اندازه گیری شود، ابعاد داده نامیده می‌شود. انتخاب ویژگی‌های مرتبط و ضروری در مرحله پیش پردازش از اهمیتی بنیادین برخوردار است. یک انتخاب ویژگی مناسب می‌تواند کارآیی یک مدل استنتاجی را افزایش دهد.

روش‌های کاهش ابعاد داده به دو دسته تقسیم می‌شوند:

۱. **روش‌های مبتنی بر استخراج ویژگی:** این روش‌ها یک فضای چند بعدی را به یک فضای با ابعاد کمتر تبدیل می‌کنند. در واقع با ترکیب مقادیر ویژگی‌های موجود، تعداد کمتری ویژگی ایجاد می‌کنند. به طوری که این ویژگی‌ها دارای تمام (یا بخش اعظمی از) داده‌های موجود در ویژگی‌های اولیه باشند. این روش‌ها به دو دسته خطی و غیر خطی تقسیم می‌شوند.

۲. **روش‌های مبتنی بر انتخاب ویژگی:** این روش‌ها سعی می‌کنند با انتخاب زیرمجموعه ای از ویژگی‌های اولیه، ابعاد داده‌ها را کاهش دهند. برخلاف روش‌های مبتنی بر استخراج ویژگی، این نوع روش‌ها معنای اصلی ویژگی‌ها را بعد از کاهش حفظ می‌کنند (فودور، ۲۰۰۲).

روش‌های مختلف انتخاب ویژگی، تلاش می‌کنند تا از میان 2^N زیر مجموعه نامزد، بهترین زیرمجموعه را پیدا کنند. در تمام این روش‌ها براساس کاربرد و نوع تعریف، زیر مجموعه‌ای به عنوان جواب انتخاب می‌شود که بتواند مقدار یک تابع ارزیابی را بهینه کند(دش و لیو، ۱۹۹۷).

تکنیک‌های انتخاب ویژگی را می‌توان به دو شیوه دسته بندی کرد:

در یک شیوه، روشهای مختلف انتخاب ویژگی را براساس دو معیار تابع تولیدکننده و تابع ارزیابی طبقه بندی می‌کنیم و براساس رویکرد دیگر تکنیک های انتخاب ویژگی را به سه دسته فیلتر[۱]، بسته بندی[۲] و هیبرید[۳] یا ترکیبی تقسیم می‌کنیم(دش و لیو، ۱۹۹۷).

در این پژوهش به اختصار رویکرد دوم را معرفی می‌کنیم.

۱-۱ رویکرد فیلتر

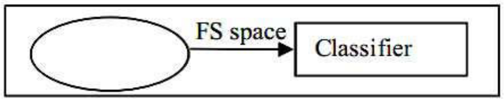
ویژگی‌ها براساس یک مرحله پیش پردازش انتخاب می‌شود که الگوریتم یادگیری را نادیده می‌گیرد. روش فیلتر براساس ویژگی‌های ذاتی داده بنا شده است نه براساس دسته بندی خاصی. ذات فیلترها برجستجوی ویژگی‌های وابسته و حذف غیر وابسته‌هاست. به راحتی برای مجموعه داده‌های با ابعاد بالا به کار می‌روند. از نظر محاسباتی خیلی ساده و سریع هستند و مستقل از الگوریتم دسته بندی می‌باشند. شیوه کار آن‌ها به این طریق است که ابتدا ویژگی‌ها به‌وسیله یک سری معیارها امتیازدهی می‌شود. سپس d تا از بهترین ویژگی‌ها انتخاب می‌شود. سپس این مجموعه به عنوان ورودی به الگوریتم دسته بندی داده می‌شود. به بیان دیگر، به تک تک ویژگی‌ها امتیازی داده می‌شود و سپس براساس آن امتیازات مرتب می‌شوند. این امتیازات می‌توانند نشان‌دهنده رابطه میان ویژگی و عنوان کلاس باشند. بدین منظور می‌توان از آزمون t یا آزمون فیشر استفاده کرد. در واقع می‌توان گفت تمامی روش‌های آماری در این گروه قرار می‌گیرند.

مشکل رویکرد فیلتر آن است که رابطه درونی میان ویژگی‌ها در نظر گرفته نشده و ممکن است مجموعه انتخابی دارای تعداد زیادی ویژگی وابسته به هم باشد، در نتیجه هدف ما یعنی کاهش ویژگی تا حد ممکن و حذف ویژگی‌های غیرضروری را محقق نمی‌سازد (سعیز، انزو و لاراناگا، ۲۰۰۷).

---

1. Filter
2. Wrapper
3. Hybrid

شکل ۱، رویکرد فیلتر به صورت شماتیک نشان داده شده است.

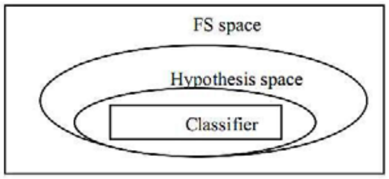


شکل ۱. رویکرد فیلتر

۱-۲ رویکرد بسته‌بندی

این رویکرد مجموعه ویژگی را براساس دسته‌بندی‌کننده، انتخاب و این مجموعه را با استفاده از دقت پیش‌بینی یا میزان خوب بودن خوشه‌ها امتیازدهی می‌کند. یعنی کلاس بندی را یک جعبه سیاه محسوب و زیر مجموعه ویژگی‌ها را براساس توان پیش‌بینی آن‌ها رتبه‌بندی می‌کند. این متد از بعضی استراتژی‌های جستجو مثل انتخاب رو به جلوی متوالی (SFS)[1] و انتخاب روبه عقب متوالی (SFS)[2] و الگوریتم ژنتیک استفاده می‌کند. در این رویکرد، تمامی زیر مجموعه‌های ممکن از ویژگی‌ها در نظر گرفته می‌شود و با ارزیابی همه حالت‌ها، بهترین آن‌ها با کمترین خطای عمومی انتخاب می‌شود. این متدها در انواع مسائل قادر به یافتن بهترین پاسخ می‌باشند.

اگرچه ممکن است روش‌های بسته‌بندی نسبت به روش‌های فیلتر، نتایج بهتری را تولید کنند، ولی اجرای این روش‌ها پرهزینه است و در بعضی مواقع با تعداد ویژگی‌های زیاد، سبب شکست می‌شوند. مشکل اصلی روش‌های بسته‌بندی، پیچیدگی بسیار بالاست که به کنترل ناپذیر شدن مسیر حل مسئله می‌انجامد(سعیز و همکاران، ۲۰۰۷). در شکل ۲، رویکرد بسته‌بندی نشان داده شده است.



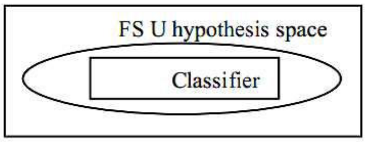
شکل ۲. رویکرد بسته‌بندی

---

1. Sequentional Forward Selection
2. Sequentional Forward Selection

۱-۳- مدل ترکیبی

این مدل از مزایای هر دو مدل قبلی به وسیله استفاده از معیارهای ارزیابی مختلف‌شان در مراحل متفاوت جستجو سود می‌برد. در روش ترکیبی، جستجو برای یک مجموعه ویژگی بهینه، درون ساختار دسته‌بندی کننده قرار دارد. در این رویکرد، کنترل انتخاب تعداد مناسب ویژگی مشکل است و به طور معمول با افزایش ویژگی‌ها مواجه می‌شویم. در اینجا نیز ارتباط میان ویژگی‌ها نادیده گرفته شده است که باعث مشکلاتی در انتخاب نهایی ویژگی‌ها می‌شود(سعیز و همکاران، ۲۰۰۷). در شکل ۳، رویکرد ترکیبی نشان داده شده است.



شکل ۳. رویکرد ترکیبی

۲- الگوریتم ژنتیک

الگوریتم ژنتیک با یک جمعیت اولیه از راه حل‌ها شروع می‌شود. هر راه حل از طریق یک کروموزوم نمایش داده می‌شود و تمامی راه حل‌های ممکن باید با استفاده از یک سیستم کدگذاری، تبدیل به کد شوند. در هر بار تکرار الگوریتم با استفاده از یک تابع ارزیابی، میزان شایستگی عناصر جمعیت فعلی مشخص و عناصر بهتر به عنوان جمعیت نسل بعد انتخاب می‌شود. مراحل الگوریتم ژنتیک استاندارد به صورت زیر می باشد:

۱. **شروع:** تولید جمعیت اولیه. در این مرحله، یک جمعیت اولیه شامل n کروموزوم تولید می شود.
۲. **برازندگی**[1]**:** مقدار برازندگی f(x) هر کروموزوم x از جمعیت ارزیابی می‌شود. مناسب بودن یا نبودن جواب، با معیار حاصل از تابع هدف سنجیده می‌شود.
۳. **جمعیت جدید:** ایجاد یک جمعیت جدید با تکرار مراحل زیر تا زمانی که جمعیت جدید کامل شود:
۳-۱ انتخاب: انتخاب دو کروموزوم از جمعیت مطابق با مقدار برازندگی آنها (افراد با برازندگی بیشتر، شانس بیشتری برای انتخاب دارند). روش‌های انتخاب عبارتند از: چرخ رولت، انتخاب رقابتی، انتخاب بالتزمن، روش تصادفی و روش رتبه‌بندی.

---

1. Fitness

۳-۲ عملگر تقاطعی[1]: با استفاده از عملگر تقاطعی احتمالی روی والدین انتخاب شده، فرزندان[2] جدید تولید می‌شوند. اگر عملگر تقاطع انجام نشود، فرزندان همان والدین خود خواهند بود. عمل تقاطع به یکی از روش‌های تقاطع تک نقطه ای، دو نقطه ای، چند نقطه ای، یکنواخت و یا تقاطع از سه والد صورت می گیرد.

۳-۳ جهش[3]: با استفاده از یک عملگر جهش احتمالی، فرزندان جدید تحت عمل جهش قرار می گیرند. در این حالت، جهش روی هر موقعیت در کروموزوم صورت می گیرد. برای جهش دو روش معکوس کردن و تعویض وجود دارد.

۳-۴ پذیرش: فرزند جدید در جمعیت جدید قرار داده می شود.

۴. **جابه جایی:** جمعیت جدید (فرزندان) یا جمعیت قبلی (والدین) یا همدیگر تشکیل یک نسل جدید را می دهند. در این حالت، بعضی از والدین حذف و بعضی از فرزندان جانشین آنها می شوند و دو جمعیت تبدیل به یک جمعیت شده و اندازه جمعیت نیز ثابت می‌ماند.

پژوهش‌ها نشان داده است که حذف همه کروموزوم‌های جمعیت مرحله قبل و انتخاب جمعیت جدید از میان فرزندان، ممکن است بسیاری از جواب‌های مناسب موجود در میان جمعیت مرحله قبل را حذف نماید.

۵. **شرط توقف:** با برآورده شدن شرط توقف کنید و بهترین راه حل در جمعیت جاری (آخرین جمعیت) را گزارش دهید.

۶. **تکرار:** در صورت عدم برآورده شدن شرط توقف، به قدم دوم یعنی برازندگی بازگردید (وفایی و ایمام، ۱۹۹۴).

۳- ماشین بردار پشتیبان

نخستین الگوریتم برای طبقه‌بندی و دسته‌بندی الگوها در سال ۱۹۳۶ توسط فیشر داده شد و معیار آن برای بهینه و کم کردن خطای طبقه‌بندی داده‌های آموزشی بوده است. در سال ۱۹۶۵ پژوهشگر روسی به نام ولادیمیروپنیک گامی بسیار مهم در طراحی طبقه‌بندی کننده‌ها برداشت.

---

1. Crossover
2. Offsprings
3. Mutation

وی نظریه آماری یادگیری را به صورت مستحکم‌تری بنانهاد و ماشین‌های بردار پشتیبان را بر این اساس عرضه کرد. ماشین‌های بردار پشتیبان دارای خواص زیر هستند:

۱. طراحی طبقه‌بندی‌کننده با بیشترین حد تعمیم

۲. رسیدن به نقطه بهینه کلی تابع

۳. تعیین خودکار ساختار و توپولوژی بهینه طبقه‌بندی کننده

۴. مدل کردن توابع تمایز غیر خطی با استفاده از هسته‌های غیر خطی و مفهوم حاصل ضرب داخلی در فضاهای هیلبرت.

ماشین بردار پشتیبان الگوریتمی است که نوع خاصی از مدل‌های خطی را می‌یابد. این ماشین بیشترین حاشیه ابر صفحه را حاصل می‌کند. حداکثر کردن حاشیه ابر صفحه به حداکثر شدن تفکیک میان طبقه‌ها می‌انجامد. به نزدیک‌ترین نقاط آموزشی به حداکثر حاشیه ابر صفحه، بردارهای پشتیبان نامیده می‌شود. تنها از این بردارها (نقاط) برای مشخص کردن مرز میان طبقات استفاده می‌شود(شین، لی و کیم، ۲۰۰۵).

اگر داده‌ها به صورت خطی مجزا از هم باشند، ماشین بردار پشتیبان به ماشین های خطی برای تولید یک سطح بهینه که داده ها را بدون خطا و با حداکثر فاصله میان صفحه و نزدیکترین نقاط آموزشی (بردارهای پشتیبان) تفکیک می نماید، آموزش می دهد. اگر نقاط آموزشی به صورت $\{x_i, y_i\}$ **و بردار ورودی** $x_i \in R^n$ **و ارزش طبقه** l ....$y_i \in \{-1, 1\}, i = 1,$ تعریف کنیم، درحالتی که داده‌ها بصورت خطی قابل تفکیک هستند، یک مرز تصمیم‌گیری خطی (ابرصفحه) در حالت کلی به صورت زیر تعریف می‌شود:

$$w \cdot x + b = 0$$

X یک نقطه روی مرز تصمیم‌گیری و w یک بردار عمود بر مرز تصمیم‌گیری است. همانطور که در فرمول نشان داده شده است، $\frac{b}{\|w\|}$ فاصله مبدا تا ابرصفحه و $(w \cdot x)$ بیانگر ضرب داخلی دو بردار W و x است. با ضرب یک ثابت در دو طرف، در معادله فوق باز هم تساوی برقرار است. در نتیجه برای تعریف یکتای مقدار b و w شرایط زیر بر روی آن‌ها اعمال می‌شود:

$$y_i(wx_i + b) = 1$$
$$y_i(wx_i + b) > 1$$

با بدست آوردن w، مقدار b به ازای بردارهای پشتیبان مختلف بدست آمده و b نهایی با میانگین‌گیری از مقادیر بدست آمده محاسبه می‌شود. طبقه بندی کننده نهایی داده‌ها از رابطه زیر بدست می‌آید:

$$Y=\text{sign}\ (\sum_{i=1}^{N} y_i \alpha_i (x.x) + b)$$

اگر داده‌ها به صورت خطی قابل تفکیک نباشند، رابطه بالا به صورت زیر تغییر می‌یابد:

$$y = sign(\sum_{i=1}^{N} y_i\, \alpha_i K(x, x_i) + b)$$

که در آن $K(x, x_i)$ تابع کرنلی است که برای ایجاد ماشین‌هایی با انواع مختلفی از سطوح تصمیم‌گیری غیرخطی در فضای داده‌ها، ضرب‌های داخلی تولید می‌کند. به عنوان مثال، سه نوع تابع کرنل در مدل ماشین بردار پشتیبان عبارتند از:

۱. ماشین چند جمله ای با تابع کرنل $K(x, x_i) = (x \cdot x_i + 1)^d$ که در آن d، درجه کرنل چندجمله ای است.

۲. ماشین تابع پایه شعاعی با تابع کرنل $K(x, x_i) = \exp\ (-1/\delta^2 (x - x_i)^2)$ که در آن δ پهنای باند کرنل تابع پایه شعاعی است.

۳. ماشین NN دو لایه با تابع کرنل

$K(x, x_i) = S\{(\mathrm{x} \cdot x_i)\} = 1/(1 + \exp\{v(x \cdot x_i) - c\})$ که در آن c و v پارامترهای توزیع زیگموییدی $S\{(\mathrm{x} \cdot x_i)\}$ هستند؛ به نحوی که نامعادله $c \geq v$ برقرار باشد(ونگ، ۲۰۰۵).

SVM در شناسایی الگو، برآورد رگرسیون، پیش‌بینی‌های سری زمانی مالی، بازاریابی، بازده‌های برآوردی تولید، طبقه‌بندی متن، تشخیص چهره با استفاده از تصویر، شناسایی دست خط و تشخیص‌های پزشکی در مقایسه با سایر تکنیک های یادگیری عملکرد بهتری دارد(کریستیانی و تیلور، ۲۰۰۰). طی سال های اخیر در حوزه‌های مختلف مدیریت مالی مانند رتبه بندی اعتباری و پیش بینی سری‌های زمانی از رویکرد SVM استفاده‌های زیادی شده است.

۴- پیشینهٔ تجربی پژوهش

تاکنون از روش های مختلفی برای پیش بینی درماندگی مالی استفاده شده است.

راعی و فلاح‌پور (۱۳۸۷) نتایج دو روش ماشین بردار پشتیبان و روش لوجیت را مقایسه کرده اند. نمونه آن‌ها متشکل از چهل شرکت ورشکسته و چهل شرکت سالم میان سال‌های ۱۳۷۵ تا ۱۳۸۰ می‌باشد. آن‌ها از ۵ نسبت مالی در مطالعه خود استفاده کرده اند. نتایج آن‌ها نشان می‌دهد که روش ماشین بردار پشتیبان به طور معنا داری دقت بیشتری از روش لوجیت دارد.

مهرانی و همکاران (۱۳۸۴) با استفاده از نسبت های مالی، دو الگوی زیمسکی و شیراتا و روش لوجیت به پیش بینی ورشکستگی پرداختند. آن ها ۳۶ شرکت ورشکسته و ۷۸ شرکت سالم میان سال های ۱۳۷۹ تا ۱۳۸۱ را در نظر گرفتند. دقت پیش بینی مدل آن ها برای الگوی شیراتا ۹۴.۷ درصد و برای الگوریتم زیمسکی ۹۷.۴ درصد بوده است.

پورزمانی و همکاران (۱۳۸۸) نیز برای پیش‌بینی ورشکستگی، یازده نسبت مالی را در دوره زمانی ۱۳۸۱ تا ۱۳۸۶ انتخاب و از روش لوجیت برای برآورد مدل استفاده کرده اند. نمونه مورد مطالعه شامل ۳۰ شرکت ورشکسته و ۳۰ شرکت سالم می باشد. دقت پیش بینی در حالت‌های مختلف میان ۷۸ و ۹۶ درصد بوده است

رهنمای رودپشتی و همکاران (۱۳۸۸) به بررسی کاربرد مدل‌های پیش‌بینی ورشکستگی آلتمن و فالمر در شرکت‌های پذیرفته بورس اوراق بهادار تهران پرداختند. نتایج این پژوهش نشان می‌دهد که مدل آلتمن در پیش‌بینی ورشکستگی محافظه کارتر از مدل فالمر عمل می‌کند. در مدل آلتمن، شرکت‌های با مقادیر $z < 1.81$ و در مدل فالمر شرکت‌هایی با مقادیر $F < 0$ ورشکسته تلقی می‌شوند.

مرادی، شفیعی سردشت و ابراهیم‌پور (۱۳۹۱) در پژوهشی با عنوان پیش‌بینی درماندگی مالی شرکت‌ها بوسیله مدل ماشین بردار پشتیبان و تحلیل ممیزی چندگانه بر هر دو بخش مهم در فرایند پیش‌بینی درماندگی مالی یعنی انتخاب ویژگی و پیش‌بینی تاکید داشته اند. در این پژوهش به منظور برآورد متغیرهای تاثیرگذار و همچنین تسریع عملیات پیش بینی با استفاده از آزمون تی از میان ۲۱ متغیر مورد نظر، تعداد ۱۰ نسبت مالی به همراه متغیر غیر مالی کارایی برای ورود به مدل‌های پیش‌بینی انتخاب شدند. نتایج پژوهش آن‌ها نشان می‌دهد که عملکرد ماشین بردار

پشتیبان به عنوان یک مدل هوش مصنوعی به طور محسوسی از مدل سنتی و آماری تحلیل ممیزی چندگانه بهتر بوده است، اما متغیر کارایی محاسبه شده در عمل تاثیری روی دقت کلی مدل‌های پیش‌بینی کننده نداشته است.

موسوی شیری، بافنده ایمان دوست و بلندرفتار پسیخانی (۱۳۹۲) به بررسی قدرت روش k، نزدیکترین همسایه در پیش‌بینی درماندگی مالی شرکت‌های پذیرفته شده بورس اوراق بهادار تهران طی سال‌های ۱۳۸۴ تا۱۳۸۶پرداخته‌اند. در این پژوهش برای انتخاب نسبت‌های مالی به عنوان متغیرهای مستقل پژوهش از مدل چهار متغیره موسوی شیری و طبرستانی (۱۳۸۸)، مدل سه متغیره کمیجانی و سعادت‌فر (۱۳۸۵) و نیز مدل شش متغیره نیکبخت و شریفی (۱۳۸۹) استفاده شده است. نتایج نشان می‌دهد که روش مذکور با دقت خوبی توان پیش‌بینی درماندگی مالی شرکت‌ها را دارد.

محسنی، اقابابایی و محمدقربانی (۱۳۹۲) از کارایی به عنوان یک متغیر پیش‌بینی کننده در کنار نسبت‌های مالی مهم پیش‌بینی درماندگی مالی استفاده کرده اند. آن‌ها از روش تحلیل پوششی داده‌ها برای محاسبه شاخص کارایی شرکت‌های تولیدی پذیرفته شده در بورس اوراق بهادار تهران طی دوره ۱۳۸۸-۱۳۸۴ استفاده نموده‌اند و سپس با استفاده از رگرسیون لجستیک در قالب دو الگو (با متغیرکارایی و بدون آن) به پیش‌بینی درماندگی مالی پرداخته اند. یافته‌های پژوهش حاکی از آن است که متغیر کارایی، دقت مدل پیش‌بینی را به طورمعناداری افزایش می‌دهد.

بولو و کرمی (۱۳۹۳) در پژوهشی با عنوان ارزیابی میزان اثربخشی الگوهای جریان وجوه نقد و مدل برنامه ریزی ژنتیک در پیش‌بینی درماندگی مالی شرکت‌ها، ۸۲ شرکت را در بازه زمانی سال‌های ۱۳۸۱ تا ۱۳۸۶ مورد بررسی قرار دادند. آن‌ها از الگوهای جریان نقدی در مرحله نزول برای پیش‌بینی درماندگی مالی استفاده نموده و نتایج آن را با مدل برنامه‌ریزی ژنتیک مقایسه کرده‌اند. الگوی جریان نقدی، نشان‌دهنده سیاست‌های شرکت برای تامین منابع، تخصیص منابع و ظرفیت عملیاتی می‌باشد. نتایج پژوهش مربوطه حاکی از آن است که الگوی جریان نقدی دارای قدرت پیش‌بینی درماندگی مالی می‌باشد. همچنین اثربخشی برنامه‌ریزی ژنتیک در مقایسه با الگوی جریان نقدی بیشتر است.

پژوهش‌های اولیه بر مدل های یک متغیره مانند یک نسبت مالی تمرکز داشتند. یکی از قدیمی‌ترین نسبت‌های مالی برای ارزیابی وضعیت اعتباری در سال ۱۸۷۰، نسبت جاری بود(آلتمن، ۱۹۶۸).

هوریگان (۱۹۶۸)، مطالعه‌ای انجام داد که در آن، نسبت‌های مالی را برای شرکت‌های ورشکسته و غیر ورشکسته در یک دوره شش ساله بررسی کرد. او اظهار کرد که سه نسبت سرمایه در گردش تقسیم بر کل دارایی‌ها، ارزش ویژه تقسیم بر کل بدهی و نسبت جاری، پیش‌بینی کننده خوبی برای ورشکستگی می‌باشند. مروین به این نتیجه رسید که از میان این نسبت‌ها، نسبت سرمایه در گردش به کل دارایی‌ها، بهترین شاخص برای ورشکستگی شرکت است.

ادوارد آلتمن، در سال ۱۹۶۸ برای نخستین بار اثر ترکیبات مختلف نسبت‌های مالی را برای پیش‌بینی درماندگی مالی شرکت‌های تولیدی آمریکا بررسی کرد. آلتمن در این مطالعه با استفاده از تحلیل تشخیصی چندگانه ۵ نسبت مالی را از میان ۲۲ نسبت انتخاب کرد. او در این مطالعه به دقت پیش‌بینی حدود ۹۵ درصد برای یک سال قبل از ورشکستگی دست یافت. مدل او که به «Z-Score» معروف است، هنوز به عنوان شاخصی برای سلامت مالی شر کت‌ها مورد استفاده قرار می گیرد(آلتمن، ۱۹۶۸).

آتیا (۲۰۰۱)، دقت دو گروه از متغیرها در پیش بینی ورشکستگی برای شرکت‌های آمریکایی را با یکدیگر مقایسه کرد. او یک گروه را تنها نسبت‌های مالی در نظر گرفت و در گروه دیگر از متغیرهای مستخرج از قیمت سهام به همراه نسبت‌های مالی استفاده نمود. وی از میان ۱۲۰ متغیر مختلف با استفاده از ماتریس همبستگی، ۵ متغیر را برای گروه نخست و ۶ متغیر را برای گروه دوم در نظر گرفت و سپس با استفاده از شبکه عصبی مصنوعی به پیش‌بینی ورشکستگی پرداخت. دقت پیش بینی مدل او برای گروه نخست ۸۱.۴۶ درصد و برای گروه دوم ۸۵.۵۰ درصد می‌باشد. وی همچنین اظهار داشت که استفاده از متغیرهای مربوط به قیمت سهام برای افق‌های زمانی بلند بهتر است. زیرا اطلاعات صورت‌های مالی از آنجا که تنها بر اساس ارزش دفتری گزارش می‌شوند به طور تقریبی عقب مانده می‌باشند. متغیرهای مرتبط با قیمت سهامی که وی از آن‌ها در پژوهش خود استفاده کرد به شرح زیر می‌باشند:

قیمت به نسبت جریان نقدی، نرخ تغییر قیمت سهام، نوسان‌پذیری قیمت سهام.

مین و لی (۲۰۰۵) در مقاله‌ای با عنوان پیش‌بینی ورشکستگی با استفاده از ماشین بردار پشتیبان با انتخاب بهینه پارامترهای تابع کرنل، نسبت‌های مالی را برای شرکت های کره ای در بازه سال‌های ۲۰۰۰ تا ۲۰۰۲ ابتدا با استفاده از آزمون تی و تحلیل جز اساسی به ۳۸ نسبت رساندند و سپس با استفاده از رگرسیون لجستیک گام به گام، ۱۱ نسبت مالی را در انتها انتخاب کردند. آن‌ها برای پیش‌بینی ورشکستگی مدل ماشین بردار پشتیبان را با شبکه عصبی، تحلیل ممیزی چندگانه و لوجیت مقایسه کردند. آن ها اظهار داشتند که ماشین‌بردار پشتیبان از سایر روش‌ها عملکرد بهتری دارد.

تسای (۲۰۰۹) در پژوهشی به انتخاب بهترین روش انتخاب ویژگی برای پیش‌بینی ورشکستگی پرداخته است. وی ۵ روش آزمون تی، رگرسیون لجستیک گام به گام، تحلیل عاملی، تحلیل جز اساسی و ماتریس همبستگی را با یکدیگر مقایسه کرده است. او سه فاکتور خطای نوع نخست، نوع دوم و میانگین دقت را برای مقایسه مد نظر قرار داده است همچنین از ۵ مجموعه داده استفاده نموده است. در این پژوهش پس از انتخاب مجموعه ویژگی‌ها، شبکه عصبی به عنوان مدل پیش‌بینی ورشکستگی به کار برده شده است. نتایج پژوهش وی حاکی از آن است که روش انتخاب ویژگی آزمون تی در دو معیار عملکرد بهتر از سایر روش‌ها بوده است.

چو، کیم و بائی (۲۰۰۹) با استفاده از آزمون تی و رگرسیون لجستیک به انتخاب نسبت‌های مالی برای پیش‌بینی ورشکستگی در شرکت‌های تولیدی کره پرداخته اند. مدل وی برای پیش‌بینی ورشکستگی دارای دقت ۷۸.۹۲ درصد می‌باشد.

لیانگ، تسای و ویو (۲۰۱۵) در پژوهش خود با استفاده ازچند روش به انتخاب ویژگی پرداخته‌اند. این روش‌ها عبارتند از: تحلیل ممیزی خطی (LDA)، آزمون تی، رگرسیون لجستیک، الگوریتم ژنتیک و الگوریتم ازدحام ذرات. برای پیش بینی نیز شش روش طبقه بندی ماشین بردار پشتیبان خطی، ماشین بردار پشتیبان با تابع RBF، K نزدیکترین همسایه، شبکه های بیزی نیو، درخت تصمیم و شبکه عصبی با یکدیگر مقایسه شده اند. آن‌ها بیان کردند که هیچ پاسخ دقیقی برای بهترین ترکیب روش انتخاب ویژگی و الگوریتم طبقه بندی ندارند اما به طور میانگین می‌توان گفت که انتخاب ویژگی باعث بهبود پیش‌بینی می‌شود.

فرضیه پژوهش

در مورد فرضیه پژوهش باید گفت که با توجه به نتایج پژوهش‌های انجام شده، فرضیه اصلی این پژوهش به این صورت است:

دقت کلی پیش‌بینی روش عرضه شده در پیش‌بینی درماندگی مالی از سایر روش‌های رقیب بیشتر است. این فرضیه پژوهشی، در قالب دو فرضیه آماری برای سال‌های t-1 و t-2 مورد آزمون قرار گرفته است.

روش‌شناسی پژوهش

در پژوهش حاضر تلاش می‌شود مدلی برای انتخاب ویژگی با استفاده از روش‌های آماری و هوش مصنوعی عرضه شود. نخست روش پیشنهادی هارک و مدل آلتمن را بررسی می‌کنیم.

۱- روش هارک (HARC)

در روش پیشنهادی ابتدا متغیرها بر اساس مفاهیم مالی در گروه‌های نقدینگی، عملیاتی، سودآوری و ساختار سرمایه دسته‌بندی می‌شوند، سپس در هرگروه میان هر دو متغیر دو به دو به صورت زوجی در کل داده‌ها یعنی هردو گروه درمانده و سالم توام در سطح اطمینان ۹۵ درصد آزمون تی انجام می‌شود تا وجود یا عدم وجود تفاوت معنادار میان دو نسبت مشخص شد و پس از انجام آزمون ها در هر گروه نسبت‌ها از میان متغیرهایی که تفاوت معناداری با یکدیگر ندارند بایستی نماینده انتخاب کنیم. در اینجا نسبت‌هایی را به عنوان نماینده انتخاب می‌کنیم که با تعداد بیشتری از نسبت های هم گروه تفاوت معنادار ندارد. سپس نسبت‌های انتخاب شده در این مرحله وارد نرم افزار متلب ۲۰۱۲ شده و با استفاده از کدنویسی الگوریتم ژنتیک و ماشین بردار پشتیبان انتخاب ویژگی انجام می‌گیرد (لین و همکاران، ۲۰۱۴).

۲- مدل آلتمن

همانطور که در بخش پیشینه تجربی پژوهش نیز بیان شد، ادوارد آلتمن در ۱۹۶۸ مدلی را عرضه داد که به z-score مشهور است و همچنان به عنوان شاخصی برای ارزیابی سلامت شرکت‌ها به کار می‌رود. وی در مدل خود از ۵ نسبت مالی استفاده نمود که به نظر او بهترین

نسبت‌ها برای پیش‌بینی ورشکستگی بود. نسبت‌های مورد استفاده در الگوریتم آلتمن به شرح زیر می‌باشند: سرمایه در گردش به کل دارایی‌ها، کل دارایی‌ها به سود انباشته، کل دارایی‌ها به سود قبل از بهره و مالیات، ارزش دفتری کل بدهی‌ها به ارزش دفتری حقوق صاحبان سهام، کل دارایی‌ها به فروش.

پس از مطالعه جامع ادبیات پژوهش، نسبت‌های مالی‌ای به‌عنوان متغیر در نظر گرفته شدند که در پژوهش‌های پیشین بیشتر مورد استفاده قرار گرفته بودند و اطلاعات آن‌ها در ایران قابل دسترسی بود. لین و همکاران (۲۰۱۴)، هوانگ و همکاران (۲۰۰۴)، آتیا (۲۰۰۱)، گنگ و همکاران (۲۰۱۵)، (چن و هیاسو، ۲۰۰۸)، سان و لی (۲۰۰۹). این نسبت‌ها در جدول داده شده‌اند.

جدول ۱. متغیرهای پژوهش

نام متغیر	مفهوم	نام متغیر	مفهوم
X1	دارایی جاری / بدهی جاری	X9	فروش / حساب های دریافتنی
X2	(موجودی مواد کالاو − دارایی جاری) / بدهی جاری	X10	بهای تمام شده / حساب های پرداختنی
X3	دارایی سریع / دارایی کل	X11	فروش / سرمایه در گردش خالص
X4	وجه نقد / دارایی کل	X12	فروش / دارایی های ثابت
X5	دارایی جاری / دارایی کل	X13	فروش / دارایی ها
X6	سرمایه در گردش خالص / دارایی کل	X14	فروش / دارایی های جاری
X7	بدهی جاری / دارایی کل	X15	سود خالص / فروش

ادامه جدول ۱

X8	بهای تمام شده / موجودی کالا	X16	سود خالص / دارایی ها
X17	سود خالص / دارایی های جاری	X26	کل بدهی ها / حقوق صاحبان سهام
X18	سود خالص / حقوق صاحبان سهام	X27	حقوق صاحبان سهام / دارایی کل
X19	سود خالص / دارایی های ثابت	X28	دارایی های ثابت / (بدهی های بلند مدت + حقوق صاحبان سهام)
X20	سود ناخالص / فروش	X29	اگر کل بدهی ها از کل دارایی ها بیشتر شود یک و در غیر اینصورت صفر
X21	($EBIT -$ مالیات) / حقوق صاحبان سهام	X30	EBIT / هزینه مالی
X22	EBIT / دارایی های کل	X31	هزینه مالی / حقوق صاحبان سهام
X23	($EBIT -$ مالیات) / دارایی های کل	X32	هزینه مالی / فروش
X24	سود انباشته / دارایی های کل	X33	حقوق صاحبان سهام / دارایی های ثابت
X25	کل بدهی ها / دارایی کل	X34	دارایی های ثابت / دارایی های کل

جامعه مورد بررسی در این پژوهش عبارت از شرکت‌های تولیدی پذیرفته شده در بورس اوراق بهادار تهران است. نمونه مورد استفاده متشکل از ۱۳۸ شرکت تولیدی است که به شرح زیر انتخاب شده‌اند:

نخست فهرستی از شرکت‌های تولیدی در مانده مالی بین سال‌های ۱۳۸۰ تا ۱۳۹۰ تهیه شد. گفتنی است با توجه به اینکه در مورد هر شرکت از اطلاعات یک و دو سال قبل از درماندگی نیز

استفاده شده است، در مجموع باید گفت که از اطلاعات شرکت‌ها میان سال های ۱۳۷۸ تا ۱۳۹۰ استفاده شده است. پس از اینکه ۶۹ شرکت مالی به این ترتیب انتخاب شدند، ۶۹ شرکت سالم نیز به صورت تصادفی به‌عنوان گروه دوم انتخاب شدند.

به منظور کاهش اثر تغییرپذیری نمونه‌ها، نمونه اصلی به دو بخش تقسیم شده است. بدین صورت که بر روی یک بخش مدلسازی انجام می‌شود و روی بخش دیگر آزمایش می‌شود. برای غنی‌تر شدن مدل‌ها، ۵ دسته نمونه فرعی ایجاد شده است. در هر نمونه فرعی ۱۱۱نمونه برای آزمون مدل و ۲۷ نمونه برای آزمایش قدرت مدل استفاده شده است.

به منظور تلخیص داده‌ها، ابتدا نسبت‌های مورد نظر با استفاده از داده‌های جمع‌آوری شده برای تمامی شرکت‌ها و هر یک از سال‌های t-1 و t-2 محاسبه شدند. همه فعالیت‌های مربوط به عملیات تلخیص با استفاده از نرم‌افزار اکسل انجام پذیرفته است.

در این پژوهش از کدینگ باینری در الگوریتم ژنتیک استفاده شده است. جمعیت اولیه ۶۰ شرکت در نظر گرفته شده است. میانگین دقت پیش‌بینی صحیح به‌عنوان تابع ارزیابی درنظر گرفته شده است. روش انتخاب رقابتی برای انتخاب ۳۰ جفت کروموزوم یعنی نصف اندازه جمعیت جاری به کار گرفته شده است. تقاطع یکنواخت برای ایجاد کروموزوم‌های جدید با احتمال ۰.۷ استفاده شده است. عملیات جهش به دنبال تقاطع انجام می‌شود و جهش یکنواخت در این پژوهش مورد استفاده قرار گرفته است. هر کروموزم در جمعیت با احتمال ۰.۰۵ وارد جهش می‌شود. در فرآیند تکامل ۴۰ بار تکرار یا عدم بهبود در پاسخ و واگرایی معیار توقف در نظر گرفته شده است.

در مورد مدل SVM، پارامترهای مختلفی وجود دارند که تغییر در آن‌ها بر عملکرد مدل بسیار اثرگذار است. یکی از مهمترین این پارامترها، تعیین نوع تابع کرنل است. در این پژوهش تابع کرنل RBF برای مدلسازی استفاده شد. تمامی فعالیت‌های مربوط به الگوریتم ژنتیک و مدل SVM با استفاده از نرم‌افزار MATLAB انجام شده است.

پس از انتخاب ویژگی با روش پیشنهادی هارک، انتخاب ویژگی با رگرسیون لجستیک نیز انجام گرفته و سپس نتایج پیش‌بینی درماندگی مالی با استفاده از نسبت‌های مالی انتخابی این دو روش و نیز نسبت‌های مالی مدل آلتمن مورد مقایسه قرار می‌گیرند.

یافته‌های پژوهش

در این قسمت یافته‌های پژوهش و تجزیه و تحلیل آن‌ها آورده می‌شود.

۱- انتخاب ویژگی با استفاده از آزمون t

در این قسمت ابتدا نسبت‌های مالی بر اساس مفاهیم آن در گروه‌ها دسته‌بندی می‌شوند. سپس همانطور که در بخش روش پژوهش ذکر شد، در هر گروه از نسبت‌های مالی آزمون زوجی دو به دو میان متغیرها انجام می‌گیرد. هر دو نسبتی که تفاوت معناداری میان آن‌ها وجود نداشته باشد، در واقع به هم وابستگی دارند و باید از میان آن‌ها نماینده‌ای برای ورود به الگوریتم ژنتیک انتخاب شود. در جدول ۲، نسبت‌های انتخاب شده به وسیله آزمون آماری آورده شده است.

جدول ۲. نسبت‌های انتخاب شده بوسیله آزمون آماری

نام متغیر	مفهوم	نام متغیر	مفهوم
X1	دارایی جاری / بدهی جاری	X22	EBIT / دارایی های کل
X3	دارایی سریع / دارایی کل	X26	کل بدهی ها / حقوق صاحبان سهام
X12	فروش / دارایی های ثابت	X29	اگر کل بدهی ها از کل دارایی ها بیشتر شود یک و در غیر اینصورت صفر
X13	فروش / دارایی ها	X32	هزینه مالی / فروش
X18	سود خالص / حقوق صاحبان سهام		

۲- انتخاب ویژگی با استفاده از الگوریتم ژنتیک

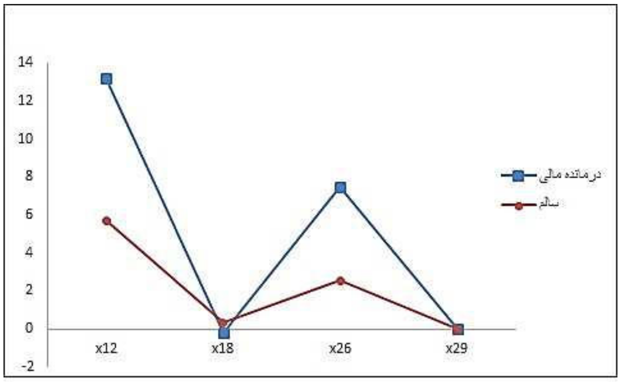
در این قسمت نسبت‌های انتخاب شده از قسمت قبل وارد الگوریتم ژنتیک شده و از میان آن‌ها متغیرهایی انتخاب می‌شوند که بیشترین قدرت را در پیش‌بینی درماندگی مالی شرکت‌ها با ماشین

بردار پشتیبان دارند. بدین ترتیب نسبت‌های مالی از دو فیلتر عبور می‌کنند. در جدول ۳، نسبت‌های منتخب روش پیشنهادی آورده شده است:

در جدول ۳. نسبت‌های انتخاب شده با الگوریتم ژنتیک

نام متغیر	مفهوم
X12	فروش / دارایی های ثابت
X18	سود خالص / حقوق صاحبان سهام
X26	کل بدهی ها / حقوق صاحبان سهام
X29	اگر کل بدهی ها از کل دارایی‌ها بیشتر شود یک و در غیر اینصورت صفر

در نمودار ۱، میانگین نسبت‌های انتخاب شده در دو گروه درمانده و سالم آورده شده‌اند.



نمودار ۱. میانگین نسبت‌های انتخاب شده در دو گروه درمانده و سالم به روش هارک

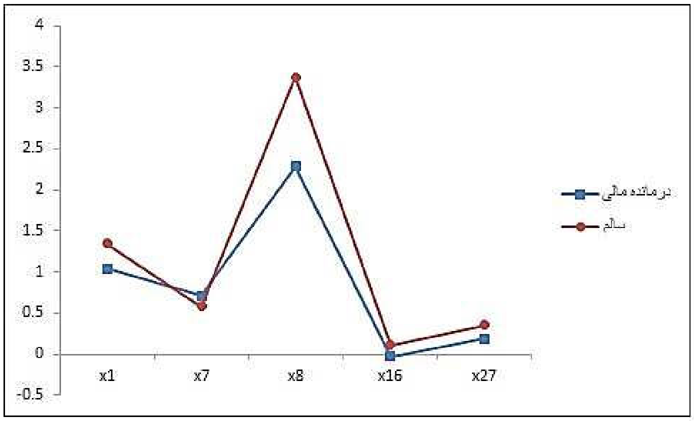
۳- انتخاب ویژگی با رگرسیون لجستیک

در این قسمت با استفاده از روش رگرسیون لجستیک به انتخاب ویژگی می‌پردازیم تا در مرحله بعد، نتایج آن را با روش پیشنهادی مقایسه کنیم. نسبت‌های انتخاب شده بوسیله رگرسیون لجستیک در جدول ۴ آورده شده است.

جدول ٤. نسبت‌های منتخب روش رگرسیون لجستیک

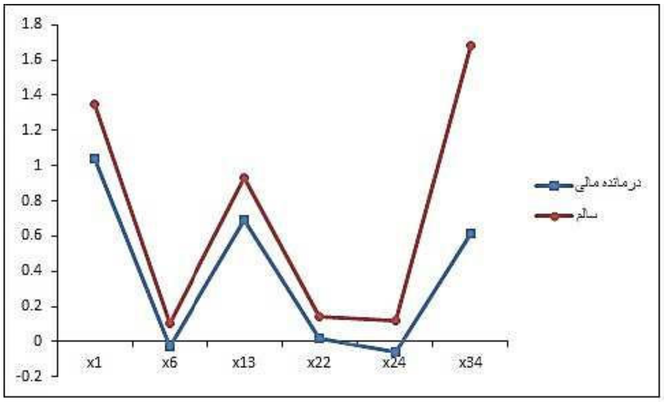
نام متغیر	مفهوم
X1	دارایی جاری / بدهی جاری
X7	بدهی جاری / دارایی کل
X8	بهای تمام شده / موجودی کالا
X16	سود خالص / دارایی ها
X27	حقوق صاحبان سهام / دارایی کل

در نمودار ۲، میانگین‌های نسبت‌های انتخاب شده به وسیله رگرسیون لجستیک را در دو گروه شرکت‌های درمانده و سالم نشان می‌دهد.



نمودار ۲. میانگین نسبت‌های انتخاب شده در دو گروه درمانده و سالم به روش رگرسیون لجستیک

همچنین در این قسمت، میانگین نسبت‌های مالی مدل آلتمن در دو گروه درمانده مالی و سالم در نمودار ۳ آورده شده است.



نمودار۳. میانگین نسبت‌های مدل آلتمن در دو گروه درمانده و سالم

جدول ۵. نتایج آزمون دقت مدل‌ها برای سال t-1

	روش هارک		رگرسیون لجستیک		مدل آلتمن	
نمونه فرعی	درمانده	سالم	درمانده	سالم	درمانده	سالم
۱	۵۳	٦٤	۵۷	٥٦	٤٨	۵۲
	٧٦/٨١درصد	٧٥/٩٢درصد	٨٢/٦٠درصد	٨١/١٥درصد	٦٩/٥٦درصد	٧٥/٣٦درصد
۲	۵۳	٦٣	۵۵	۵۷	٤٧	۵۵
	٧٦/٨١درصد	٩١/٣٠درصد	٧٩/٧١درصد	٨٢/٦٠درصد	٦٨/١١درصد	٧٩/٧١درصد
۳	٥٤	٦٥	۵۵	۵۹	٤٩	۵۵
	٧٨/٢٦درصد	٩٤/٢٠درصد	٧٩/٧١درصد	٨٥/٥٠درصد	٧١/٠١درصد	٧٩/٧١درصد
٤	٥٤	٦٤	۵۵	۵۸	۵۰	۵۵
	٧٨/٢٦درصد	٩٢/٧٥درصد	٧٩/٧١درصد	٨٤/٠٥درصد	٧٢/٤٦درصد	٧٩/٧١درصد
۵	۵۳	٦٤	٥٤	۵۷	۵۰	۵۳
	٧٦/٨١درصد	٩٢/٧٥درصد	٧٨/٢٦درصد	٨٢/٦٠درصد	٧٢/٤٦درصد	٨١/٧٦درصد

۴- دقت پیش‌بینی روش‌های مختلف با استفاده از ماشین بردار پشتیبان

جدول ۵ نتایج پیش‌بینی روش هارک، رگرسیون لجستیک و آلتمن را با استفاده از ماشین بردار پشتیبان برای سال t-1 نشان می‌دهد.

در هر ستون، عدد نخست نشان‌دهنده تعداد پیش‌بینی صحیح و عدد پایینی آن، درصد پیش‌بینی صحیح را نشان می‌دهد. دوباره یادآوری می‌شود که تعداد نمونه‌ها در هر گروه ۶۹ شرکت است.

جدول ٦. دقت کلی مدل‌ها برای سال t-1

	روش هارک	رگرسیون لجستیک	مدل آلتمن
نمونه فرعی ۱	۱۱۷	۱۱۳	۱۰۰
	۸٤/۷۸ درصد	۸۱/۸۸ درصد	۷۲/۹۹ درصد
نمونه فرعی ۲	۱۱٦	۱۱۲	۱۰۲
	۸٤/۰٦ درصد	۸۱/۱٦ درصد	۷٤/٤٥ درصد
نمونه فرعی ۳	۱۱۹	۱۱٤	۱۰٤
	۸٦/۲۳ درصد	۸۲/٦۱ درصد	۷۵/۹۱ درصد
نمونه فرعی ٤	۱۱۸	۱۱۳	۱۰۵
	۸۵/۵۱ درصد	۸۱/۸۸ درصد	۷٦/٦٤ درصد
نمونه فرعی ۵	۱۱۷	۱۱۱	۱۰۳
	۸٤/۷۸ درصد	۸۰/٤۳ درصد	۷۵/۱۸ درصد

نسبت دقت کلی حاصل‌جمع پیش‌بینی‌های صحیح در هر دو گروه است. همانطور که جدول ۶ نشان می‌دهد، دقت کلی پیش‌بینی روش هارک در سال t-1 برای نمونه‌های فرعی ۱ تا ۵، به ترتیب ۲/۹، ۲/۹، ۳/۶۲، ۳/۶۳ و ۴/۳۵ درصد بیشتر از رگرسیون لجستیک و به ترتیب ۱۱/۷۹، ۹/۶۱، ۱۰/۳۲، ۸/۸۷ و ۹/۶ درصد بیشتر از مدل آلتمن می‌باشد. این امر قدرت بالای روش هارک را در سال t-1 نشان می‌دهد.

برای بررسی فرضیه پژوهشی، عملکرد بهتر روش هارک در سال t-1، از آزمون مقایسه زوجی استفاده شده است. با استفاده از این آزمون، عملکرد بهتر روش هارک از نظر آماری بررسی می‌شود. جداول ۷ و ۸ نتایج آزمون مقایسه زوجی را برای سال t-1 نشان می‌دهند.

جدول ۷. نتایج آزمون مقایسه زوجی برای سال t-1

	هارک	رگرسیون لجستیک
میانگین دقت کلی	۸۵/۰۷ درصد	۸۱/۵۹ درصد
آمارهt	۱۲/۸۲۸	
p-value	۰/۰۰۰۲۱۳	

همانطور که جدول ۷ نشان می‌دهد، می‌توان گفت که دقت کلی پیش‌بینی روش هارک با سطح اطمینان ۹۵ درصد، از دقت کلی پیش‌بینی رگرسیون لجستیک بیشتر است. بنابراین، فرضیه پژوهشی برای سال t-1 تایید می‌شود. جدول ۸ نیز نتایج آزمون مقایسه زوجی را برای سال t-1 میان روش هارک و مدل آلتمن نشان می‌دهد.

جدول ۸ نتایج آزمون مقایسه زوجی برای سال t-1

	هارک	مدل آلتمن
میانگین دقت کلی	۸۵/۰۷ درصد	۷۵/۰۳ درصد
آماره t	۲۰/۳۰٤	
p-value	۰/۰۰۰۰۳۵	

همانطور که جدول ۸ نشان می‌دهد، می‌توان گفت که دقت کلی پیش‌بینی روش هارک با سطح اطمینان ۹۵ درصد، از دقت کلی پیش‌بینی مدل آلتمن بیشتر است. بنابراین، فرضیه پژوهشی برای سال t-1 میان این دو روش تایید می‌شود. همچنین در این دو جدول مشاهده می‌شود که روش هارک تا یک سال قبل از درماندگی به‌طور متوسط بالغ بر ۸۵ درصد موارد را صحیح پیش‌بینی می‌نماید. جدول ۹، نتایج پیش‌بینی برای سال t-2 را نشان می‌دهد.

جدول ۹. نتایج آزمون دقت مدل‌ها برای سال t-2

	روش هارک		رگرسیون لجستیک		مدل آلتمن	
نمونه فرعی	درمانده	سالم	درمانده	سالم	درمانده	سالم
۱	۵٤	٤٦	٤۸	٤٤	۵۰	٤۱
	۷۸/۲٦درصد	٦٦/٦درصد	٦۹/۵٦درصد	٦۳/۷٦درصد	۷۲/٤٦درصد	۵۹/٤۲درصد
۲	۵۵	٤٦	۵۰	٤۷	٤۸	٤۲
	۷۹/۷۱درصد	٦٦/٦درصد	۷۲/٤٦درصد	٦۸/۱۱درصد	٦۹/۵٦درصد	٦۰/۸٦درصد
۳	۵٤	٤۵	٤٦	٤٤	۵۰	٤۳
	۷۸/۲٦درصد	٦۵/۲۱درصد	٦٦/٦درصد	٦۳/۷٦درصد	۷۲/٤٦درصد	٦۲/۳۱درصد
٤	۵۵	٤٦	٤۷	٤۷	٤۸	٤۳
	۷۹/۷۱درصد	٦٦/٦درصد	٦۸/۱۱درصد	٦۸/۱۱درصد	۷۲/٤٦درصد	٦۲/۳۱درصد
۵	۵۰	٤۸	٤۸	٤۷	۵۰	٤۲
	۷۲/٤٦درصد	٦۹/۵٦درصد	٦۹/۵٦درصد	٦۸/۱۱درصد	۷۲/٤٦درصد	٦۰/۸٦درصد

جدول ۱۰. دقت کلی مدل‌ها برای سال t-2

	روش هارک	رگرسیون لجستیک	مدل آلتمن
نمونه فرعی ۱	۱۰۰	۹۲	۹۱
	۷۲/٤٦درصد	٦٦/٦درصد	٦۵/۹٤درصد
نمونه فرعی ۲	۱۰۱	۹۷	۹۰
	۷۳/۱۸درصد	۷۰/۲۸درصد	٦۵/۲۱درصد
نمونه فرعی ۳	۹۹	۹۰	۹۳
	۷۱/۷۳درصد	٦۵/۲۱درصد	٦۷/۳۹درصد
نمونه فرعی ٤	۱۰۱	۹٤	۹۱
	۷۳/۱۸درصد	٦۸/۱۱درصد	٦۵/۹٤درصد
نمونه فرعی ۵	۹۸	۹۵	۹۲
	۷۱/۰۱درصد	٦۸/۸٤درصد	٦٦/٦درصد

جدول ۱۰ نیز دقت کلی مدل‌ها را برای سال t-2 نشان می‌دهد.

همانطور که جدول ۱۰ نشان می‌دهد، دقت کلی روش هارک در سال t-2 برای تمامی نمونه‌های فرعی از هر دو روش رگرسیون لجستیک و مدل آلتمن بیشتر است.

جدول ۱۱ نتایج بررسی فرضیه پژوهشی مبنی بر عملکرد بهتر روش هارک از رگرسیون لجستیک برای سال t-2 را نشان می‌دهد.

جدول ۱۱. نتایج آزمون مقایسه زوجی برای سال t-2

	هارک	رگرسیون لجستیک
میانگین دقت کلی	۷۲/۳۱درصد	۶۷/۸۰درصد
آماره t	۵/۳۳۷	
p-value	۰/۰۰۶	

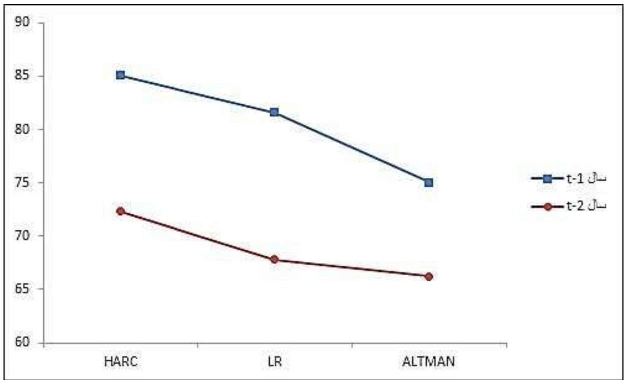
همانگونه که جدول ۱۱ نشان می‌دهد، برای سال t-2 نیز می‌توان گفت که دقت کلی پیش‌بینی روش هارک با سطح اطمینان ۹۵ درصد از دقت کلی پیش‌بینی روش رگرسیون لجستیک بیشتر است. بنابراین، فرضیه پژوهشی برای سال t-2 نیز تایید می‌گردد. جدول ۱۱ همچنین نشان می‌دهد که روش هارک تا ۲ سال قبل از درماندگی مالی، به‌طور متوسط حدود ۷۲ درصد موارد را صحیح پیش‌بینی می‌نماید. جدول ۱۲ نیز نتایج آزمون مقایسه زوجی را برای سال t-2 میان روش روش هارک و مدل آلتمن نشان می‌دهد.

جدول ۱۲. نتایج آزمون مقایسه زوجی برای سال t-2

	هارک	مدل آلتمن
میانگین دقت کلی	۷۲/۳۱درصد	٦٦/۲۱درصد
آماره t	۸/۲٤۷	
p-value	۰/۰۰۱	

همانطور که جدول ۱۲ نشان می‌دهد، می‌توان گفت که دقت کلی پیش‌بینی روش هارک با سطح اطمینان ۹۵ درصد، از دقت کلی پیش‌بینی مدل آلتمن بیشتر است. بنابراین، فرضیه پژوهشی برای سال t-2 میان این دو روش تایید می‌شود.

در انتها برای مقایسه بهتر، دقت کلی پیش‌بینی سه روش برای یک و دو سال پیش از درماندگی در نمودار ۴ آورده شده است.



نمودار ٤. میانگین دقت کلی پیش‌بینی

بحث و نتیجه‌گیری

پیش‌بینی درماندگی مالی شرکت‌ها یکی از پژوهش‌های بااهمیت در حوزه مالی است. اما آنچه که کمتر بدان پرداخته شده است، اهمیت انتخاب ویژگی در پیش‌بینی درماندگی مالی می‌باشد.

در این پژوهش، انتخاب ویژگی با استفاده از روش هارک، رگرسیون لجستیک و مدل آلتمن صورت گرفت. همچنین برای پیش‌بینی از ماشین بردار پشتیبان استفاده شد. نتایج این پژوهش نشان داد که روش هارک عملکرد بهتری از روش‌های آماری دارد. در واقع یافته‌های پژوهش نشان داد که دقت کلی روش هارک در هر دو سال از دقت کلی رگرسیون لجستیک و مدل آلتمن به طور معناداری بیشتر است. همچنین لازم به یادآوری است که در این پژوهش از معتبرسازی مقطعی استفاده شده است و در هربار اجرای الگوریتم از ۸۰ درصد داده‌ها برای آموزش و ۲۰ درصد برای آزمون استفاده شده است. با توجه به این نکته می‌توان ادعا کرد که دقت‌های گذاشته شده برای پیش‌بینی روش‌های مختلف جداول فوق، قابل اطمینان می‌باشد.

دقت پیش‌بینی روش هارک و رگرسیون لجستیک و مدل آلتمن در پژوهش((Lin et al., 2014)) به ترتیب ۸۱/۳۶، ۷۹/۹۶ و ۷۵/۳۴ درصد می‌باشد. در این پژوهش دقت پیش‌بینی این روش‌ها در یک سال قبل از وقوع درماندگی مالی به ترتیب ۸۵/۰۷، ۸۱/۶۰ و ۷۵/۰۳ می‌باشد. بدین ترتیب

ملاحظه می‌شود که دقت روش‌های مطرح شده در پیش‌بینی درماندگی مالی شرکت‌های پذیرفته شده بورس اوراق بهادار تهران قابل مقایسه با دقت این روش‌ها در مطالعات بین المللی است. همچنین یافته‌های این پژوهش موید پژوهش ((Liang et al., 2015)) است که درآن الگوریتم ژنتیک و رگرسیون لجستیک به ترتیب بهترین روش‌های انتخاب ویژگی می‌باشند.

پیشنهادهای پژوهش

پیشنهادهایی که از این پژوهش حاصل می‌شود به شرح زیر می‌باشد:

۱- پیشنهادهای کاربردی پژوهش

- توصیه می‌شود حسابرسان برای اظهار نظر درباره تداوم فعالیت شرکت‌ها از این مدل استفاده نمایند.
- به سرمایه‌گذاران توصیه می‌شود برای اتخاذ تصمیمات سرمایه‌گذاری بهتر از مدل ارائه شده در این پژوهش استفاده نمایند.
- بانک‌ها و موسسات رتبه‌بندی اعتباری می‌توانند این روش و نسبت‌های مالی داده شده در فرایند رتبه‌بندی شرکت‌ها و اعطای تسهیلات را به کارگیرند.

۲- پیشنهادهایی برای پژوهش های بعدی

- استفاده از الگوریتم‌های بهینه‌سازی همچون ازدحام ذرات برای بهینه‌سازی پارامترهای الگوریتم ژنتیک و ماشین بردار پشتیبان.
- استفاده از مدل‌های دیگر به جای ماشین بردار پشتیبان برای طبقه‌بندی و پیش‌بینی درماندگی مالی شرکت‌ها.
- استفاده از سایر روش‌های آماری مانند ضریب همبستگی در گام نخست روش هارک.
- بررسی و آزمایش کاربرد نسبت‌های مالی دیگر به خصوص نسبت‌های برگرفته از اقلام صورت جریان وجوه نقد.
- توجه به گزارش حسابرس و اثر دادن آن در تحلیل و پیش‌بینی درماندگی مالی.
- انتخاب زوجی شرکت‌های درمانده مالی و سالم از صنایع یا طبقات مشابه.

منابع

- بولو، قاسم، کرمی، اصغر. (۱۳۹۳). ارزیابی میزان اثربخشی الگوهای جریان وجوه نقد و مدل برنامه‌ریزی ژنتیک در پیش‌بینی درماندگی مالی شرکت‌ها. فصلنامه تحقیقات حسابداری و حسابرسی، ۲۲، ۹۳-۷۰
- پورزمانی، زهرا، جهانشاد، آزیتا، عین قلایی شهرام. (۱۳۸۸). « پیش‌بینی وضعیت مالی و اقتصادی شرکت‌ها با استفاده از نسبت‌های مالی مبتنی بر سود، جریان‌های نقدی و رشد». پژوهش نامه حسابداری مالی و حسابرسی. ۱(۳)، ۱۱۵-۹۳.
- راعی رضا، فلاح پور سعید. (۱۳۸۷). «کاربرد ماشین بردار پشتیبان در پیش‌بینی درماندگی مالی شرکتها با استفاده از نسبتهای مالی». فصلنامه بررسیهای حسابداری و حسابرسی،۵۳، ۳۴-۱۷.
- رهنمای رودپشتی فریدون، علیخانی راضیه، مران‌جوری مهدی. (۱۳۸۸). «بررسی کاربرد مدل‌های پیش‌بینی ورشکستگی آلتمن و فالمر در شرکت‌های پذیرفته شده در بورس اوراق بهادار تهران» فصلنامه بررسی‌های حسابداری و حسابرسی، ۵۵، ۳۴-۱۹.
- محسنی، رضا، آقابابایی، رضا، محمدقربانی، وحید. (۱۳۹۲). پیش‌بینی درماندگی مالی با بکار بردن کارایی به عنوان یک متغیر پیش‌بینی کننده. فصلنامه پژوهش‌ها و سیاست‌های اقتصادی، ۶۵، ۱۴۶-۱۲۳.
- مرادی، محسن، شفیعی سردشت، مرتضی، ابراهیم پور، ملیحه. (۱۳۹۱). « پیش بینی درماندگی مالی شرکت‌ها بوسیله‌ی مدل‌های ماشین بردار پشتیبان و تحلیل ممیزی چندگانه». فصلنامه بورس اوراق بهادار ،۱۸، ۱۳۶-۱۱۸.
- موسوی شیری، محمد، بافنده ایماندوست، صادق، بلندرفتار پسیخانی، محمد. (۱۳۹۲). کاربرد روش k-نزدیک‌ترین همسایه در پیش‌بینی درماندگی مالی شرکت‌های پذیرفته شده در بورس اوراق بهادار تهران. دو فصلنامه اقتصاد پولی و مالی، ۶، ۶۶-۴۸.
- مهرانی ساسان، مهرانی کاوه، منصفی یاشار، کرمی غلامرضا.(۱۳۸۴)« بررسی کاربردی الگوهای پیش‌بینیورشکستگی زمیسکی و شیراتا در شرکتهای پذیرفته شده در بورس اوراق بهادار تهران». فصلنامه بررسی‌های حسابداری و حسابرسی،۴۱، ۱۳۱ -۱۰۵.
- Altman, E. I. (1968). Financial ratios, discriminant analysis and the prediction of corporate bankruptcy. *The journal of finance, 23*(4), 589-609.

- Atiya, A. F. (2001). Bankruptcy prediction for credit risk using neural networks: A survey and new results. *Neural Networks, IEEE Transactions on, 12*(4), 929-935.
- Chen, L.-H., & Hsiao, H.-D. (2008). Feature selection to diagnose a business crisis by using a real GA-based support vector machine: An empirical study. *Expert Systems with Applications, 35*(3), 1145-1155.
- Cho, S., Kim, J., & Bae, J. K. (2009). An integrative model with subject weight based on neural network learning for bankruptcy prediction. *Expert Systems with Applications, 36*(1), 403-410.
- Cristianini, N., & Shawe-Taylor, J. (2000). *An introduction to support vector machines and other kernel-based learning methods*: Cambridge university press.
- Dash, M., & Liu, H. (1997). Feature selection for classification. *Intelligent data analysis, 1*(1), 131-156.
- Fodor, I. A Survey of Dimension Reduction Techniques. Lawrence Livermore National Laboratory (LLNL) Technical Report, June 2002: UCRL-ID-148494.
- Geng, R., Bose, I., & Chen, X. (2015). Prediction of financial distress: An empirical study of listed Chinese companies using data mining. *European Journal of Operational Research, 241*(1), 236-247.
- Gordon, M. J. (1971). Towards a theory of financial distress. *The journal of finance, 26*(2), 347-356.
- Horrigan, J. O. (1968). A short history of financial ratio analysis. *Accounting Review*, 284-294.
- Huang, Z., Chen, H., Hsu, C.-J., Chen, W.-H., & Wu, S. (2004). Credit rating analysis with support vector machines and neural networks: a market comparative study. *Decision support systems, 37*(4), 543-558.

- Liang, D., Tsai, C.-F., & Wu, H.-T. (2015). The effect of feature selection on financial distress prediction. *Knowledge-Based Systems, 73*, 289-297.
- Lin, F., Liang, D., Yeh, C.-C., & Huang, J.-C. (2014). Novel feature selection methods to financial distress prediction. *Expert Systems with Applications, 41*(5), 2472-2483.
- Min, J. H., & Lee, Y.-C. (2005). Bankruptcy prediction using support vector machine with optimal choice of kernel function parameters. *Expert Systems with Applications, 28*(4), 603-614.
- Saeys, Y., Inza, I., & Larrañaga, P. (2007). A review of feature selection techniques in bioinformatics. *bioinformatics, 23*(19), 2507-2517.
- Shin, K.-S., Lee, T. S., & Kim, H.-j. (2005). An application of support vector machines in bankruptcy prediction model. *Expert Systems with Applications, 28*(1), 127-135.
- Sun, J., & Li, H. (2009). Financial distress prediction based on serial combination of multiple classifiers. *Expert Systems with Applications, 36*(4), 8659-8666.
- Tsai, C.-F. (2009). Feature selection in bankruptcy prediction. *Knowledge-Based Systems, 22*(2), 120-127.
- Vafaie, H., & Imam, I. F. (1994). *Feature selection methods: genetic algorithms vs. greedy-like search.* Paper presented at the Proceedings of international conference on fuzzy and intelligent control systems.
- Wang, L. (2005). *Support Vector Machines: theory and applications* (Vol. 177): Springer Science & Business Media.
- Whitaker, R. B. (1999). The early stages of financial distress. *Journal of Economics and Finance, 23*(2), 123-132.