



بررسی ریسک اعتباری مشتریان حقوقی با استفاده از مدل ماشین بردار پشتیبان و مدل هیبریدی الگوریتم ژنتیک - مطالعه موردی بانک تجارت

امین محمدیان حاجی کرد^۱

ملیحه اصغرزاده زعفرانی^۲

مصطفی امام دوست^۳

تاریخ پذیرش: ۹۵/۲/۷

تاریخ دریافت: ۹۴/۱۲/۵

چکیده

طراحی و استقرار مدل رتبه بندی اعتباری در نظام بانکی نقش مهمی در بالا بردن کارایی تخصیص منابع به مشتریان هدف دارد. در این تحقیق با هدف تدوین مدلی جهت ارزیابی ریسک اعتباری مشتریان حقوقی بانک از ماشین بردار پشتیبان (SVM) و الگوریتم ژنتیک بهره گرفته شده است. بدین منظور، مطالعه‌ای بر روی متغیرهای مالی ۲۸۲ شرکت که طی سال‌های ۱۳۸۷ تا ۱۳۹۰ از بانک تجارت تسهیلات دریافت کرده‌اند، صورت گرفته است. در این پژوهش برای بهینه سازی ورودی های ماشین بردار پشتیبان از الگوریتم ژنتیک بهره گرفته شده است، توان بسیار بالای الگوریتم ژنتیک در انتخاب نقاط بهینه، همواره این اطمینان خاطر را برای استفاده کننده فراهم می آورد که نقاط بهینه پیشنهادی، نقاط بهینه بهتری برای مساله خواهند بود. در مدل هیبریدی GA-SVM، الگوریتم ژنتیک داده‌های ورودی مدل SVM را بهینه می سازد.

یافته‌های تحقیق نشان می دهد مدل هیبریدی GA-SVM نسبت به مدل SVM عملکرد بهتری در شناسایی مشتریان خوش حساب و بد حساب و پیش بینی ریسک اعتباری مشتریان دارد.

واژه‌های کلیدی: رتبه بندی اعتباری، ریسک اعتباری، ماشین بردار پشتیبان، الگوریتم ژنتیک.

۱- کارشناس ارشد مدیریت مالی دانشگاه تهران mohamadian_1245@ut.ac.ir
۲- کارشناس ارشد مدیریت بازرگانی دانشگاه تهران malihe.zaferani@gmail.com
۳- دانشجوی دکتری مدیریت مالی دانشگاه تهران (نویسنده مسئول) emandoost@ut.ac.ir

۱- مقدمه

بررسی عملکرد اغلب کشورها نشانگر آن است که بین سرمایه‌گذاری و سطح پیشرفت اقتصادی رابطه تنگاتنگی وجود دارد، بدین معنی که کشورهایی که دارای الگوی کارآمدی در تخصیص سرمایه به بخش‌های مختلف اقتصادی هستند، اغلب از پیشرفت‌های اقتصادی بیشتری برخوردار بوده و به تبع آن دارای رفاه اجتماعی بالاتری نیز می‌باشند.

در اقتصادهای نوظهوری چون ایران، اهمیت ایجاد نظام تامین مالی کارا، شفاف و روان جهت تشویق کارآفرینی و نیل به اهداف توسعه‌ای، دو چندان است. یکی از کلیدی‌ترین خرده سیستم‌های تامین مالی که منجر به تحقق اهداف فوق می‌شود، نظام سنجش اعتبار است. با وجود این موسسات متقاضی تامین مالی با روشی عینی و علمی و بر اساس ضوابط، غربال شده و در نتیجه، استفاده بهینه از منابع صورت گرفته و توزیع عادلانه سرمایه میسر می‌گردد.

اهمیت اعطای تسهیلات در بانکداری و نقش خطیر آن در رشد اقتصادی و افزایش اشتغال، منجر به توسعه چندین مدل گوناگون برای ارزیابی اعتباری مشتریان متقاضی این تسهیلات شده است. اما بسیاری از این مدل‌ها، مدل‌های کلاسیک بوده و توانایی ارزیابی را به طور کامل و بهینه ندارند بنابراین زمینه ورود مدل‌های پیچیده‌تر نظیر هوش مصنوعی به این حوزه فراهم گردیده است. بطور کلی تکنیک‌های گسترده‌ای در حوزه‌های آمار، ریاضی، اقتصاد سنجی و پژوهش عملیاتی در بانک‌ها و موسسات اعتباری در زمینه امتیازدهی اعتباری مورد استفاده قرار می‌گیرد که به دو دسته کلی تقسیم می‌شوند:

الف) الگوهای امتیازدهی اعتباری ناپارامتری (نظیر: برنامه ریزی ریاضی، طبقه بندی درختی، الگوی نزدیک ترین همسایه، فرایند تحلیل سلسله مراتبی، سیستم‌های خبره، شبکه‌های عصبی مصنوعی و الگوریتم ژنتیک)

ب) الگوهای امتیازدهی اعتباری پارامتری (نظیر: مدل احتمالی خطی، مدل تحلیلی ممیزی، مدل لاجیت و مدل پروبیت)

در این تحقیق از الگوریتم ژنتیک، یکی از ابزارهای دقیق و منعطف بهینه‌سازی، برای بهینه‌سازی مدل SVM (Support Vector Machine) در ارزیابی ریسک اعتباری مشتریان بانک‌ها و موسسات اعتباری استفاده شده است. توان بسیار بالای الگوریتم ژنتیک در انتخاب نقاط بهینه، همواره این اطمینان خاطر را برای استفاده کننده فراهم می‌آورد که نقاط بهینه پیشنهادی، نقاط بهینه‌تری برای مساله خواهد بود.

یکی از مهمترین تصمیمات موسسات مالی اعطای وام به مشتریان است. برای تصمیم‌گیری در این مورد، مشتریان را در دو دسته طبقه بندی می‌کنند: "مشتریان خوش حساب و مشتریان بد حساب". تا چندی پیش قضاوت در مورد اینکه کدام یک از مشتریان خوش حساب یا بد حساب هستند، توسط کارشناس اعتباری انجام می‌شد بدین صورت که کارشناسان اعتباری با بررسی فرم‌های اطلاعات مشتریان که درباره موقعیت اجتماعی و شرایط اقتصادی در اختیار داشتند، تصمیم‌گیری لازم را اتخاذ می‌کردند. طبقه

بندی متقاضیان اعتبار یکی از نیازهای انکارناپذیر موسسه‌ها و بازارهای مالی به ویژه بانک‌ها و موسسات مالی اعتباری است (بیزنس، ۲۰۰۳).

اما پیشرفت تکنولوژی و ظهور فناوری ذخیره‌سازی داده‌ها و تسهیل توانایی موسسات مالی برای ذخیره تمام اطلاعات در رابطه با خصوصیات و رفتار بازپرداخت متقاضیان اعتباری باعث شد تا موسسات مالی به صورت خودکار و با استفاده از یادگیری الگوریتم‌های آماری بتوانند به تصمیم‌گیری درباره مشتریان متقاضی تسهیلات اعتباری بپردازند. بدین ترتیب نیاز به الگوریتم‌ها و مدل‌های ریسک اعتباری احساس شد. چرا که این مدل‌ها دانش لازم در مورد زبان‌های احتمالی اعطای وام را ارائه خواهند داد (میونگ، ۲۰۰۳). تا کنون مدل‌های زیادی برای سنجش ریسک اعتباری و رتبه‌بندی مشتریان ارائه شده است و ابزارهای زیادی از حوزه‌ی علم ریاضی، آمار، اقتصاد سنجی و تحقیق در عملیات از قبیل: برنامه ریزی ریاضی، شبیه‌سازی احتمالی و قطعی، شبکه‌های عصبی، تحلیل بقاء، نظریه بازی‌ها، تحلیل لاجیت و پروبیت در پیشرفت اندازه‌گیری دقیق ریسک اعتباری سهمیم بوده‌اند اما همواره نیاز برای پیشرفت این مدل‌های اعتبارسنجی احساس می‌شود. همانطور که "توماس" دو علت اساسی برای توسعه سیستم‌های فعلی رتبه‌بندی اعتباری را به صورت زیر ذکر می‌کند:

- به علت شرایط اقتصادی موسسات مالی و اعتباری و بانک‌ها نیازمند شناسایی تکنیک‌های پیش‌بینی ریسک مصرف‌کننده برای تطبیق خودکار با شرایط جدید هستند.
 - موسسات مالی و اعتباری و بانک‌ها به جای سعی در کاهش مشتریان بد حساب، امیدوارند بتوانند مشتریانی را شناسایی کنند که بر منفعت هستند (طورزنی، ۱۳۸۹).
- الگوریتم SVM (ماشین بردار پشتیبان)، جز الگوریتم‌های تشخیص الگو دسته‌بندی می‌شود. از الگوریتم SVM، در هر جایی که نیاز به تشخیص الگو یا دسته‌بندی اشیاء در کلاس‌های خاص باشد می‌توان استفاده کرد. ماشین‌های بردار پشتیبان دارای خواص زیر هستند:
- طراحی طبقه‌بندی‌کننده با حداکثر تعمیم
 - رسیدن به نقطه بهینه کلی تابع
 - تعیین خودکار ساختار و توپولوژی بهینه برای طبقه‌بندی‌کننده
- این روش از جمله روش‌های نسبتاً جدیدی است که در سال‌های اخیر کارایی خوبی نسبت به روش‌های قدیمی‌تر برای طبقه‌بندی از جمله شبکه‌های عصبی پرسپترون نشان داده است. مبنای کاری دسته‌بندی‌کننده SVM دسته‌بندی خطی داده‌ها است و در تقسیم خطی داده‌ها سعی می‌کنیم خطی را انتخاب کنیم که حاشیه اطمینان بیشتری داشته باشد. SVM الگوریتمی است که نوع خاصی از مدل‌های خطی را می‌یابد که حداکثر حاشیه ابر صفحه را حاصل می‌کنند. حداکثر کردن حاشیه ابر صفحه منجر به حداکثر شدن تفکیک بین طبقات می‌شود. به نزدیک‌ترین نقاط آموزشی به حداکثر حاشیه ابر صفحه، بردارهای پشتیبان اطلاق می‌گردد. تنها از این بردارها (نقاط) برای مشخص کردن مرز بین طبقات استفاده می‌شود (شین،

۲۰۰۵). در سال‌های اخیر در حوزه‌های مختلف مدیریت مالی مانند رتبه بندی اعتباری و پیش بینی سری-های زمانی، از رویکرد SVM استفاده‌های زیادی شده است.

در این تحقیق از الگوریتم ژنتیک -یکی از ابزارهای دقیق و منعطف برای بهینه‌سازی- برای بهینه‌سازی مدل SVM (Support Vector Machine) در ارزیابی ریسک اعتباری مشتریان بانک‌ها و موسسات اعتباری استفاده می‌شود. توان بسیار بالای الگوریتم ژنتیک در انتخاب نقاط بهینه، همواره این اطمینان خاطر را برای استفاده کننده فراهم می‌آورد که نقاط بهینه پیشنهادی، نقاط بهینه بهتری برای مساله خواهد بود. آنچه که باعث شده الگوریتم ژنتیک دارای این قابلیت استثنایی و ویژه باشد، نحوه جستجوی این مدل است. تفاوت اساسی که این روش با سایر روش‌های جستجوی موجود مثل شبکه عصبی دارد، این است که الگوریتم ژنتیک به طور همزمان از چند نقطه در فضای جواب، فرایند جستجو را آغاز می‌کند حال آنکه سایر روش-های بهینه‌یابی به دلیل آغاز فعالیت خود از یک نقطه همواره احتمال از دست دادن جواب‌های بهتر مواجه هستند.

در این تحقیق توان بالای الگوریتم ژنتیک در بهینه‌سازی متغیرهای ورودی و پارامترهای ماشین بردار پشتیبان مورد استفاده قرار می‌گیرد تا بتوان عملکرد الگوریتم ژنتیک را در بهینه‌سازی برای ارزیابی ریسک اعتباری مشتریان بانک‌ها را مورد سنجش قرار داد. جنبه نوآوری این تحقیق در این است که تاکنون در ایران از این روش برای بررسی ریسک اعتباری مشتریان بانک‌ها مورد استفاده قرار نگرفته است.

۲- مبانی نظری و مروری بر پیشینه پژوهش

پیشینه مدل رگرسیون خطی به زمانی که فیتز پاتریک^۱ (۱۹۳۲) مطالعاتی را در مورد پیش بینی عدم توانایی یک مشتری برای بازپرداخت وام، به عنوان یکی از مسائل مهم و مورد بحث در امور مالی انجام داد، باز می‌گردد و طی ۸۰ سال اخیر این موضوع به زمینه مهمی در تحقیقات نظری و تجربی در حوزه اقتصاد مالی تبدیل شده است.

روش تحلیل ممیزی خطی^۲ بیور (۱۹۶۶) یکی از اولین تحقیقاتی است که در موضوع عدم ایفای تعهدات مالی انجام شده و می‌توان او را در شمار پیشگامان این مبحث برشمرد.

در سال ۱۹۶۸ این روش توسط ادوارد آلتمن^۳ به یک روش چند متغیره^۴ توسعه یافت. او تلاش بسیاری برای یافتن یک رابطه معنی‌دار بین متغیرهای مالی (نسبت‌های مالی) یک شرکت و احتمال عدم توانایی در پرداخت دیون این شرکت در آینده انجام داد و رابطه ای معروف به Z-Score را ارائه داد. این روش مبتنی بر تحلیل ممیزی خطی بین شرکت‌های خوب و بد بود که بوسیله آن فقط شرکت‌های خوب از بد جدا می‌شدند. با اینکه این روش بسیار ابتدایی بود، اما تا حدودی وضعیت شرکت‌های بد را می‌توانست پیش‌بینی نماید و تا سال ۱۹۸۰ چارچوب تحلیلی اکثر مطالعات بود. در خلال دهه ۱۹۸۰ روش تحلیلی لاجستیک جایگزین آن شد که تاکنون هم در شمار پرکاربردترین روش‌های آماری مورد استفاده برای پیش‌بینی احتمال نکول است.

طی سالیان اخیر مدل‌های شبکه‌های عصبی که از جایگاه ویژه‌ای در مقایسه با برخی از مدل‌های کلاسیک مانند تحلیل ممیزی در برآورد و تخمین برخوردار هستند (کوپر، ۱۹۹۹). این مدل‌ها را می‌توان به نوعی تعمیمی از مدل‌های رگرسیون غیر خطی به شمار آورد. از این رو به کارگیری این مدل‌ها از سوی محققین برای تشخیص روندها و الگوها در داده‌ها و همچنین تولید دانش از داده‌ها به صورت گسترده‌ای فراگیر شده است (لیائو، ۱۹۹۲).

سالچین برگر و همکارانش از شبکه‌های عصبی MLP جهت پیش‌بینی سلامت مالی پس‌انداز و وام‌ها استفاده کرد. وی این روش را با مدل رگرسیونی لاجیت مقایسه کرد. این مطالعه بر روی اطلاعات S&L از ژانویه ۱۹۸۶ تا دسامبر ۱۹۸۷ انجام شد و برتری شبکه‌های عصبی بر مدل لاجیت نشان داده شد (سالچین برگر و همکاران، ۱۹۹۲).

دیوید وست در سال ۲۰۰۰ پنج نوع شبکه‌های عصبی $L\dot{V}Q^{\delta}$ ، FAR^{ϵ} ، RBF^{γ} ، MOE^{α} و MLP را با درخت‌های تصمیم‌گیری، تخمین چگالی کرنل^۹، نزدیکترین همسایگی، تحلیل ممیزی و رگرسیون لاجیت در امتیازدهی اعتباری بر روی اطلاعات اعتباری کشور آلمان و استرالیا مقایسه کرد و نشان داد که MLP از همه آنها دقیق‌تر می‌باشد (وست، ۲۰۰۰).

تیان شیونگ لی، چی چو چیو، چی جی لو و آی فی چین در سال ۲۰۰۲ مدل شبکه‌های عصبی BPN^{10} را با تحلیل ممیزی مقایسه کردند و نشان دادند که BPN دارای عملکرد بهتری در مقایسه با تحلیل ممیزی می‌باشد. همچنین به این نتیجه رسیدند که شبکه‌های هیبریدی دارای سرعت و دقت بیشتری هستند (شیونگ و همکاران، ۲۰۰۲).

زان هوانگ، هسین چینا، چیا جونگ هسو، ون هوا چن و سوشان وو در رساله خود، اوراق بهادار منتشره شرکت‌های موجود در بازار رقابتی در آمریکا و تایوان بر اساس متغیرهای مالی (نسبت‌های مالی) با استفاده از روش SVM و شبکه‌های عصبی BPN رتبه‌بندی کرده‌اند. در این تحقیق سعی شده است کارایی این دو روش نشان داده شود. نتایج این تحقیق نشان می‌دهد روش SVM روشی کارآتر می‌باشد (هوانگ و همکاران، ۲۰۰۴).

روش‌های دیگری که برای حل مسائل مربوط به طبقه‌بندی و ریسک اعتباری وجود دارد، روش درخت تصمیم‌گیری و ماشین بردار پشتیبان می‌باشند که لیو و همکاران از SVM بهینه شده توسط الگوریتم ژنتیک برای امتیازدهی اعتباری استفاده کردند (لیو، ۲۰۱۰).

شی چن و همکاران از SVM برای به حداقل رساندن خطای پیش‌بینی ورشکستگی شرکت‌ها استفاده کردند (شی چن و همکاران، ۲۰۱۱).

فلاح شمس (۱۳۸۴) در رساله دکتری خود با عنوان طراحی و تبیین مدل ریسک اعتباری در نظام بانکی کشور کارایی مدل‌های خطی، لجستیک و شبکه‌های عصبی مصنوعی برای پیش‌بینی ریسک اعتباری مشتریان نظام بانکی کشور مورد بررسی قرار داده است که نتایج حاصله بیانگر آن است که توابع سیگموئیدی مناسبترین مدل برای پیش‌بینی ریسک اعتباری محسوب می‌گردد و بیشترین کارایی برای

پیش بینی ریسک اعتباری به ترتیب مربوط به شبکه های عصبی مصنوعی و مدل لاجیت می باشد (شمس، ۱۳۸۴).

افشار (۱۳۸۶) در پایان نامه خود با عنوان "بررسی کاربرد استفاده از مدل K.M.V در پیش بینی ریسک اعتباری مشتریان بانک ها و موسسات اعتباری" دقت و کارایی مدل K.M.V به عنوان یکی از روش های اعتبار سنجی مشتریان، با مدل آلتمن مورد بررسی قرار داده است و نتایج حاصله نشان داده که مدل K.M.V از دقت عمل بالاتری برخوردار است.

سهندی (۱۳۸۹) در پایان نامه خود با عنوان "طراحی مدل رتبه بندی اعتباری مبتنی بر متغیرهای درون شرکتی و صنعت با استفاده از شبکه های عصبی مصنوعی" دقت مدل شبکه عصبی در شناسایی مشتریان خوش حساب و بد حساب توسط شبکه عصبی سنجیده است. نتیجه این تحقیق نشان می دهد شبکه عصبی با دسته بندی حدود ۸۹ درصد مشتریان بد حساب و ۸۵ درصد مشتریان خوش حساب از رگرسیون لجستیک با دسته بندی حدود ۸۳ درصد مشتریان بد حساب و ۸۲ درصد مشتریان خوش حساب عملکرد بهتری داشته است (سهندی، ۱۳۸۹).

۳- روش شناسی پژوهش

در این تحقیق، جامعه آماری شامل کلیه اشخاص حقوقی است که جزء مشتریان بانک بوده و در سال های مورد بررسی از بانک تسهیلات اعتباری دریافت نموده اند. دوره مورد مطالعه با توجه به شرایط موجود از سال ۱۳۸۷ الی ۱۳۹۰ است. بر این اساس در این پژوهش از اطلاعات پرونده های تسهیلاتی ۲۸۲ شرکت نزد بانک تجارت که مشتمل بر ۱۹۷ شرکت خوش حساب و ۸۵ شرکت بد حساب می باشد، استفاده شده است.

۴- فرضیه های پژوهش

با توجه به این مقدمه، پژوهش حاضر دارای یک فرضیه اصلی می باشد که به شرح زیر است:
فرضیه اصلی: دقت پیش بینی مدل هیبریدی GA-SVM در پیش بینی ریسک اعتباری اشخاص حقوقی از مدل SVM بیشتر است.

۵- متغیرهای پژوهش

متغیرها را می توان به شکل های مختلف و بر اساس معیارهای متفاوت دسته بندی کرد. در این تحقیق با توجه به داده های موجود، مجموعه ۳۱ متغیر مالی که در برگزیده نسبت های مالی مشتریان وام گیرنده می باشد به عنوان متغیرهای کاندید در نظر گرفته شده است.

متغیرهای کاندید برای طراحی مدل شامل ۳۱ متغیر است که از صورت های مالی استخراج شده و مربوط به نسبت های مالی می باشند. از آنجایی که از صورت های مالی و اطلاعات پایه ای آن استخراج شده

اند به صورت دو به دو با همدیگر همبستگی داشته و اگر بدون انجام تحلیل‌هایی، در مدل وارد شوند به دلیل وجود هم خطی باعث بی معنی شدن سایر ضرایب از طریق بالا رفتن کواریانس بین ضرایب و در نتیجه کاهش کارایی تخمین زنده‌های مدل می‌شوند. بدین منظور در ابتدای طراحی مدل ماتریس همبستگی برای این متغیرها تشکیل گردید که بر مبنای آن ۲۱ متغیر جهت انجام تحلیل‌هایی بر مبنای رهیافت گام به گام انتخاب گردید. رهیافت گام به گام جهت انتخاب متغیرهای مدل در میان متغیرهای انتخاب شده که از حداقل همبستگی جهت ورود به مدل ارزیابی ریسک اعتباری مشتریان برخوردارند، کاربرد دارد.

در این پژوهش به منظور تدوین بهترین مدل با کمترین متغیرهای ممکن، استفاده از رهیافت گام به گام در دستور کار قرار گرفته است. در این رویکرد در ابتدا فرایند انتخاب رو به جلو متغیر وابسته به ازای تک تک متغیرها، بهترین مدل دو متغیره و پس از آن مدل‌های سه متغیره و بیشتر برآزش می‌شوند تا جایی که معیارهای خوب بودن مدل رگرسیونی جدید از مدل گزینش شده قبلی بهتر نباشد. پس از آن در رویکردی رو به عقب سعی می‌شود تا ابعاد مدل کاهش یابد و این اقدام نیز با این شرط انجام می‌پذیرد که در معیارهای برآزش مدل تاثیر منفی نداشته باشد. بدین منظور در این پژوهش از نرم افزار MINITAB استفاده گردید. در نهایت از بین ۲۱ متغیر فوق، ۱۴ متغیر جهت مدل نهایی انتخاب شدند. نتایج نهایی این فرایند در جدول زیر نشان داده شده است:

عنوان	متغیر مالی	نحوه محاسبه متغیر
x_2	کل بدهی/کل دارایی	از حاصل تقسیم کل بدهیهای شرکت به کل داراییهای شرکت بدست میاید
x_3	حقوق صاحبان سهام/کل دارایی	حاصل تقسیم حقوق صاحبان سهام به کل داراییهای شرکت می باشد
x_4	بدهی بلند مدت/کل دارایی ها	حاصل تقسیم بدهی های بلند مدت به کل داراییهای شرکت می باشد
x_8	بدهی بلند مدت/حقوق صاحبان سهام	حاصل تقسیم بدهی بلند مدت به حقوق صاحبان سهام می باشد
x_9	نسبت جاری	حاصل تقسیم دارایی های جاری به بدهیهای جاری می باشد
x_{12}	موجودی نقد+سپرده/کل دارایی	حاصل تقسیم موجودی نقد به علاوه سپرده به کل داراییهای شرکت میباشد
x_{16}	نسبت آنی	حاصل تقسیم داراییهای جاری منهای موجودی کالا بر بدهیهای جاری میباشد.
x_{18}	ح دریافتنی/بدهی های جاری	حاصل تقسیم حسابهای دریافتنی به بدهی های جاری می باشد
x_{17}	ح دریافتنی/کل فروش	حاصل تقسیم حسابهای دریافتنی به کل فروش می باشد
x_{19}	ح پرداختنی/فروش	حاصل تقسیم حساب پرداختنی به فروش شرکت می باشد
x_{20}	فروش/دارایی ها	حاصل تقسیم فروش به دارایی های شرکت می باشد
x_{22}	فروش/سود قبل از بهره و مالیات	حاصل تقسیم فروش به سود قبل از بهره و مالیات می باشد
x_{30}	کل دارایی ها /دارایی های ثابت	حاصل تقسیم کل داراییهای شرکت به داراییهای ثابت می باشد
x_{31}	اسناد پرداختنی کوتاه مدت /سرمایه در گردش	حاصل تقسیم اسناد پرداختنی کوتاه مدت به سرمایه در گردش می باشد

روند کلی الگوریتم های ژنتیکی

قبل از این که یک الگوریتم ژنتیکی بتواند اجرا شود، ابتدا باید کدگذاری (یا نمایش) مناسبی برای مسئله مورد نظر پیدا شود. معمولی ترین شیوه نمایش کروموزومها در الگوریتم ژنتیک به شکل رشته های دودویی است. هر متغیر تصمیم گیری به صورت دودویی درآمده و سپس با کنار هم قرار گرفتن این متغیرها کروموزوم ایجاد می شود. گرچه این روش گسترده ترین شیوه کدگذاری است اما شیوه های دیگری مثل نمایش با اعداد حقیقی در حال گسترش هستند. همچنین یک تابع برازندگی نیز باید ابداع شود تا به هر راه حل کدگذاری شده ارزشی را نسبت دهد. در طی اجرا، والدین برای تولید مثل انتخاب می شوند و با استفاده از عملگرهای آمیزش و جهش با هم ترکیب می شوند تا فرزندان جدیدی تولید کنند. این فرآیند چندین بار تکرار می شود تا نسل بعدی جمعیت تولید شود. سپس این جمعیت بررسی می شود و در صورتی که ضوابط همگرایی برآورده شوند، فرآیند فوق خاتمه می یابد.

ساختار مدل SVM

فرمول بندی مدل SVM بر اساس مفاهیم کمینه کردن ساختاری ریسک^{۱۱} است که نسبت به کمینه کردن تجربی ریسک^{۱۲} برتری هایی را نشان داده است. مدل های سنتی شبکه عصبی بر اساس کمینه کردن تجربی ریسک عمل می کنند. SVM حد بالای ریسک مورد انتظار را کمینه می کند این ویژگی است که توانایی تعمیم پذیری SVM را افزایش می دهد و هدف نظریه های یادگیری آماری است.

کمینه کردن ساختاری ریسک

اگر S_{11} را به عنوان فضای فرضیه با بعد VC، h و به صورتی در نظر بگیریم که:

$$S_1 \subset S_2 \subset \dots \subset S_{\infty}$$

کمینه کردن ساختاری ریسک به معنی کمینه کردن رابطه زیر است:

$$\min R_{\text{emp}}[f] + \sqrt{\frac{h \ln\left(\frac{2l}{h} + 1\right) - \ln\left(\frac{\delta}{4}\right)}{l}}$$

بعد VC در واقع به نوعی ظرفیت یک مجموعه تابع را نشان می دهد.

SVM الگوریتمی است که نوع خاصی از مدل های خطی را می یابد که حداکثر حاشیه ابر صفحه را حاصل می کنند. حداکثر کردن حاشیه ابر صفحه منجر به حداکثر شدن تفکیک بین طبقات می شود. به نزدیک ترین نقاط آموزشی به حداکثر حاشیه ابر صفحه، بردارهای پشتیبان اطلاق می گردد. تنها از این بردارها (نقاط) برای مشخص کردن مرز بین طبقات استفاده می شود (شین، ۲۰۰۵)

اگر داده ها به صورت خطی مجزا از هم باشند، SVM 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 را به صورت $[x_i, y_i]$ و بردار ورودی $x_i \in \mathbb{R}^n$ و ارزش

طبقه $y_i \in \{-1, 1\}$, $i = 1, \dots, n$ تعریف کنیم، آنگاه در حالتی که داده‌ها بصورت خطی قابل تفکیک هستند، قواعد تصمیم‌گیری که تعریف می‌شود و توسط یک صفحه بهینه که طبقات تصمیم‌گیری باینری را تفکیک می‌کند، به صورت معادله زیر است:

$$y = \text{sign} \left(\sum_{i=1}^N y_i \alpha_i (X_i \bullet X_i) + b \right)$$

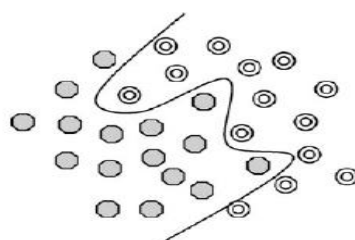
که در آن y خروجی معادله، α_i ارزش طبقه نمونه آموزشی X_i و \bullet نشان دهنده ضرب داخلی است. بردار $x = (x_1, x_2, \dots, x_n)$ نشان دهنده یک داده ورودی و بردارهای X_i ، $i = 1, \dots, N$ بردارهای پشتیبان هستند. در معادله فوق پارامترهای b و α_i تعیین‌کننده ابر صفحه هستند. اگر داده‌ها به صورت خطی قابل تفکیک نباشند، معادله فوق به معادله زیر تغییر می‌یابد:

$$y = \text{sign} \left(\sum_{i=1}^N y_i \alpha_i K(X, X_i) + b \right)$$

تابع $K(X, X_i)$ تابع کرنلی است که برای ایجاد ماشین‌هایی با انواع مختلفی از سطوح تصمیم‌گیری غیر خطی در فضای داده‌ها، ضرب‌های داخلی تولید می‌کند.

تابع کرنل

تابع کرنل، تابع وزنی است که در تکنیک‌های پیش‌بینی غیرپارامتریک استفاده شده و دارای دو شرط $\int_{-\infty}^{+\infty} k(u) du = 1$ و $\int_{-\infty}^{+\infty} k(u) du = 1$ است و برای تمام مقادیر u . اگر دو کلاس به نحوی باشند که نتوان آنها را با یک صفحه جدا کرد، از روش‌های غیر خطی و تعریف تابع تصمیم‌گیری و توابع کرنل استفاده می‌کنیم.



تفکیک غیر خطی دو کلاس

در SVM ها از کرنلها برای تعریف شباهت ورودی ها و از تابع تلفات برای سنجش میزان شباهت خروجی‌ها بهره می‌گیرند. در واقع کرنلها ابزارهایی برای تعیین شباهت بین ورودی‌ها هستند که دارای کلاس‌های متفاوتی می‌باشند. هر کلاس از کرنلها توانایی تشخیص نوع خاصی از شباهت را دارد و بنابراین نوع خاصی از فضای ویژگی را بازسازی می‌کند (گان، ۱۹۹۸)

روش هیبریدی SVM و GA

در این مدل پس از انتخاب متغیر های درگیر در مدل و نرمال سازی متغیر ها بطور تصادفی قسمتی از داده ها را به عنوان داده تست در نظر گرفته و از بین مابقی داده ها به عنوان داده های ورودی در الگوریتم ژنتیک استفاده می گردد. این کار به این دلیل صورت می گیرد که مدل ترکیبی SVM-GA دچار بیش برآزشی داده های تست نشده و کم بودن میزان خطای پیش بینی در داده های تست به دلیل ورود داده های تست در حین آموزش مدل نباشد. در این روش با استفاده از الگوریتم ژنتیک بهترین مجموعه از داده های آموزش از بین داده های ورودی الگوریتم ژنتیک که انتخاب آنها در ابتدای مدل بصورت تصادفی انجام گرفته است، انتخاب می شود.

به دلیل حساس بودن SVM به داده های نویز، وجود نویز می تواند عملکرد SVM را تحت تاثیر قرار دهد. در این ایده با استفاده از روش GA حالت های مختلف انتخاب داده های آموزش در نظر گرفته می شود و تلاش می شود تا بهینه ترین مجموعه برای آموزش داده ها استخراج شود. در این روش طبقه بند SVM خود به عنوان تابع شایستگی الگوریتم ژنتیک می باشد و هدف کمینه کردن خطای طبقه بندی با استفاده از GA می باشد. کروموزم ها در این روش باینری می باشد و طول کروموزم به اندازه داده های ورودی است. جمعیت اولیه الگوریتم ژنتیک بصورت تصادفی تولید می گردد بدین صورت که ابتدا کروموزم هایی بطول داده های ورودی که تمام بیت های آنها ۱ می باشد تولید کرده و بطور تصادفی حداکثر ۲۰ درصد بیت ها را به صفر تبدیل کرده و جمعیت تولید شده را در الگوریتم ژنتیک استفاده می کنیم در خروجی الگوریتم ژنتیک اگر بیت مربوطه ۱ باشد داده به عنوان داده آموزش در SVM در نظر گرفته می شود و اگر ۰ باشد به همراه داده های تست ابتدای مدل به عنوان داده های تست در انتهای مدل آزمون می شود.

۶- نتایج پژوهش

جداول زیر نتایج پیش بینی مدل های SVM و مدل هیبریدی GA-SVM را در پانزده بار اجرای مدل را نشان می دهند. در این پژوهش برای اجرای مدل های مورد نظر از نرم افزار MATLAB و توابع موجود در این نرم افزار استفاده شده است.

جدول ۱. نتایج پیش بینی مدل SVM

	آموزشی			آزمایشی			مجموع		
	خوش حساب	بد حساب	مجموع	خوش حساب	بد حساب	مجموع	خوش حساب	بد حساب	مجموع
run1	۱۵۵	۷۱	۲۲۶	۴۲	۱۴	۵۶	۱۹۷	۸۵	۲۸۲
	۰,۷۲	۰,۷۲	۰,۷۲	۰,۶۰	۰,۳۶	۰,۵۴	۰,۷۰	۰,۶۶	۰,۶۸
run2	۱۵۷	۶۹	۲۲۶	۴۰	۱۶	۵۶	۱۹۷	۸۵	۲۸۲
	۰,۷۱	۰,۷۱	۰	۰,۶۵	۰,۵۶	۰,۶۳	۰,۷۰	۰,۶۸	۰,۶۹

بررسی ریسک اعتباری مشتریان حقوقی با استفاده از مدل ماشین بردار ... / امین محمدیان حاجی کرد، ملیحه اصغرزاده زعفرانی و مصطفی امامدوست

	آموزشی			آزمایشی			مجموع		
	خوش حساب	بد حساب	مجموع	خوش حساب	بد حساب	مجموع	خوش حساب	بد حساب	مجموع
run۳	۱۵۲	۷۴	۲۲۶	۴۵	۱۱	۵۶	۱۹۷	۸۵	۲۸۲
	۰,۷۱	۰,۷۴	۰,۷۲	۰,۷۳	۰,۱۸	۰,۶۳	۰,۷۲	۰,۶۷	۰,۷۰
run۴	۱۵۴	۷۲	۲۲۶	۴۳	۱۳	۵۶	۱۹۷	۸۵	۲۸۲
	۰,۷۴	۰,۶۷	۰,۷۲	۰,۸۱	۰,۵۴	۰,۷۵	۰,۷۶	۰,۶۵	۰,۷۲
run۵	۱۵۹	۶۷	۲۲۶	۳۸	۱۸	۵۶	۱۹۷	۸۵	۲۸۲
	۰,۷۰	۰,۷۶	۰,۷۲	۰,۴۵	۰,۴۴	۰,۴۵	۰,۶۵	۰,۶۹	۰,۶۷
run۶	۱۵۹	۶۷	۲۲۶	۳۸	۱۸	۵۶	۱۹۷	۸۵	۲۸۲
	۰,۷۱	۰,۷۸	۰,۷۳	۰,۵۰	۰,۶۷	۰,۵۵	۰,۶۷	۰,۷۵	۰,۷۰
run۷	۱۵۸	۶۸	۲۲۶	۳۹	۱۷	۵۶	۱۹۷	۸۵	۲۸۲
	۰,۸۱	۰,۶۵	۰,۷۶	۰,۷۲	۰,۵۳	۰,۶۶	۰,۷۹	۰,۶۲	۰,۷۴
run۸	۱۵۷	۶۹	۲۲۶	۴۰	۱۶	۵۶	۱۹۷	۸۵	۲۸۲
	۰,۷۲	۰,۷۸	۰,۷۴	۰,۴۰	۰,۶۳	۰,۴۶	۰,۶۵	۰,۷۵	۰,۶۸
run۹	۱۵۸	۶۸	۲۲۶	۳۹	۱۷	۵۶	۱۹۷	۸۵	۲۸۲
	۰,۶۸	۰,۷۶	۰,۷۱	۰,۴۶	۰,۵۳	۰,۴۸	۰,۶۴	۰,۷۲	۰,۶۶
run۱۰	۱۶۰	۶۶	۲۲۶	۳۷	۱۹	۵۶	۱۹۷	۸۵	۲۸۲
	۰,۶۳	۰,۷۷	۰,۶۷	۰,۴۱	۰,۷۴	۰,۵۲	۰,۵۸	۰,۷۶	۰,۶۴
run۱۱	۱۵۹	۶۷	۲۲۶	۳۸	۱۸	۵۶	۱۹۷	۸۵	۲۸۲
	۰,۷۴	۰,۶۹	۰,۷۳	۰,۶۶	۰,۲۲	۰,۵۲	۰,۷۳	۰,۵۹	۰,۶۸
run۱۲	۱۶۲	۶۴	۲۲۶	۳۵	۲۱	۵۶	۱۹۷	۸۵	۲۸۲
	۰,۶۵	۰,۷۷	۰,۶۹	۰,۶۹	۰,۵۲	۰,۶۳	۰,۶۶	۰,۷۱	۰,۶۷
run۱۳	۱۵۴	۷۲	۲۲۶	۴۳	۱۳	۵۶	۱۹۷	۸۵	۲۸۲
	۰,۶۸	۰,۷۵	۰,۷۰	۰,۶۵	۰,۶۲	۰,۶۴	۰,۶۸	۰,۷۳	۰,۶۹
run۱۴	۱۵۷	۶۹	۲۲۶	۴۰	۱۶	۵۶	۱۹۷	۸۵	۲۸۲
	۰,۶۸	۰,۷۴	۰,۷۰	۰,۶۵	۰,۵۶	۰,۶۳	۰,۶۸	۰,۷۱	۰,۶۸
run۱۵	۱۶۱	۶۵	۲۲۶	۳۶	۲۰	۵۶	۱۹۷	۸۵	۲۸۲
	۰,۶۹	۰,۷۵	۰,۷۱	۰,۵۶	۰,۵۰	۰,۵۴	۰,۶۶	۰,۶۹	۰,۶۷

جدول ۲. نتایج پیش بینی مدل SVM هیبریدی

	آموزشی			آزمایشی			مجموع		
	خوش حساب	بد حساب	مجموع	خوش حساب	بد حساب	مجموع	خوش حساب	بد حساب	مجموع
run۱	۱۳۴	۵۸	۱۹۲	۶۳	۲۷	۹۰	۱۹۷	۸۵	۲۸۲
	۰,۷۷	۰,۶۲	۰,۷۲	۰,۹۸	۰,۲۶	۰,۷۷	۰,۸۴	۰,۵۱	۰,۷۴
run۲	۱۳۹	۶۱	۲۰۰	۵۸	۲۴	۸۲	۱۹۷	۸۵	۲۸۲
	۰,۶۵	۰,۷۴	۰,۶۸	۰,۸۸	۰,۵۸	۰,۷۹	۰,۷۲	۰,۶۹	۰,۷۱
run۳	۱۲۵	۶۷	۱۹۲	۷۲	۱۸	۹۰	۱۹۷	۸۵	۲۸۲
	۰,۶۹	۰,۶۷	۰,۶۸	۰,۹۶	۰,۴۴	۰,۸۶	۰,۷۹	۰,۶۲	۰,۷۴
run۴	۱۳۱	۵۸	۱۸۹	۶۶	۲۷	۹۳	۱۹۷	۸۵	۲۸۲
	۰,۸۸	۰,۵۳	۰,۷۷	۰,۹۷	۰,۱۹	۰,۷۴	۰,۹۱	۰,۴۲	۰,۷۶
run۵	۱۳۵	۵۶	۱۹۱	۶۲	۲۹	۹۱	۱۹۷	۸۵	۲۸۲
	۰,۸۳	۰,۶۳	۰,۷۷	۰,۹۷	۰,۲۴	۰,۷۴	۰,۸۷	۰,۴۹	۰,۷۶
run۶	۱۳۳	۵۷	۱۹۰	۶۴	۲۸	۹۲	۱۹۷	۸۵	۲۸۲
	۰,۸۸	۰,۵۴	۰,۷۸	۰,۵۰	۰,۶۴	۰,۵۴	۰,۷۶	۰,۵۸	۰,۷۰
run۷	۱۲۴	۶۴	۱۸۸	۷۳	۲۱	۹۴	۱۹۷	۸۵	۲۸۲
	۰,۹۱	۰,۵۵	۰,۷۹	۰,۹۶	۰,۰۵	۰,۷۶	۰,۹۳	۰,۴۲	۰,۷۸
run۸	۱۳۷	۵۱	۱۸۸	۶۰	۳۴	۹۴	۱۹۷	۸۵	۲۸۲
	۰,۸۹	۰,۶۳	۰,۸۲	۰,۸۸	۰,۲۱	۰,۶۴	۰,۸۹	۰,۴۶	۰,۷۶
run۹	۱۳۰	۶۴	۱۹۴	۶۷	۲۱	۸۸	۱۹۷	۸۵	۲۸۲
	۰,۸۸	۰,۵۶	۰,۷۷	۰,۹۹	۰,۱۹	۰,۸۰	۰,۹۱	۰,۴۷	۰,۷۸
run۱۰	۱۲۹	۶۴	۱۹۳	۶۸	۲۱	۸۹	۱۹۷	۸۵	۲۸۲
	۰,۷۲	۰,۷۰	۰,۷۲	۰,۹۷	۰,۳۸	۰,۸۳	۰,۸۱	۰,۶۲	۰,۷۵
run۱۱	۱۳۹	۵۹	۱۹۸	۵۸	۲۶	۸۴	۱۹۷	۸۵	۲۸۲
	۰,۶۸	۰,۶۹	۰,۶۹	۰,۹۳	۰,۴۶	۰,۷۹	۰,۷۶	۰,۶۲	۰,۷۲
run۱۲	۱۳۵	۶۱	۱۹۶	۶۲	۲۴	۸۶	۱۹۷	۸۵	۲۸۲
	۰,۷۹	۰,۷۴	۰,۷۷	۰,۹۷	۰,۲۹	۰,۷۸	۰,۸۴	۰,۶۱	۰,۷۷
run۱۳	۱۲۷	۶۴	۱۹۱	۷۰	۲۱	۹۱	۱۹۷	۸۵	۲۸۲
	۰,۸۸	۰,۵۶	۰,۷۷	۰,۹۹	۰,۱۴	۰,۷۹	۰,۹۲	۰,۴۶	۰,۷۸
run۱۴	۱۳۹	۶۲	۲۰۱	۵۸	۲۳	۸۱	۱۹۷	۸۵	۲۸۲
	۰,۹۱	۰,۴۴	۰,۷۷	۰,۴۵	۰,۷۸	۰,۵۴	۰,۷۸	۰,۵۳	۰,۷۰
run۱۵	۱۳۵	۶۰	۱۹۵	۶۲	۲۵	۸۷	۱۹۷	۸۵	۲۸۲
	۰,۷۶	۰,۶۵	۰,۷۲	۰,۹۵	۰,۲۴	۰,۷۵	۰,۸۲	۰,۵۳	۰,۷۳

بر اساس جداول فوق مدل هیبریدی GA-SVM حدود ۷۸ درصد مشتریان خوش حساب و ۵۰ درصد مشتریان بد حساب را صحیح دسته‌بندی کرده است. در مقابل دقت پیش‌بینی مدل SVM برای مشتریان بد حساب حدود ۶۱ درصد و برای مشتریان خوش حساب ۶۳ درصد می‌باشد. بر این اساس مدل هیبریدی GA-SVM نسبت به مدل SVM عملکرد بهتری داشته است. در صورتی که بخواهیم بر اساس داده‌های آموزشی، آزمایشی به سنجش دقت مدل‌های SVM، هیبریدی GA-SVM بپردازیم، نتایج ذیل بدست می‌آید.

جدول ۳. نتایج پیش‌بینی کلی مدل‌ها

مدل	آموزشی	آزمایشی
SVM	۰,۷۱	۰,۵۷
GA-SVM هیبریدی	۰,۷۰	۰,۶۹

با بررسی جدول فوق به این نتیجه می‌رسیم که مدل SVM توانسته ۵۷ درصد از داده‌ها آزمایشی را درست پیش‌بینی نماید ولی در مقابل مدل GA-SVM هیبریدی ۶۹ درصد داده‌های آزمایشی مدل را درست پیش‌بینی کرده است، بنابراین نتیجه می‌گیریم که دقت پیش‌بینی مدل GA-SVM هیبریدی از SVM بالاتر می‌باشد.

جهت آزمون فرضیه‌ها از آزمون مقایسات زوجی استفاده شده است. این آزمون بر روی نتایج گروه آزمایشی صورت گرفته است. برای تست این آزمون از جعبه ابزار تحلیل اطلاعات نرم افزار Excel انجام شده است. فرضیه تحقیق:

- H_0 : بین دقت پیش‌بینی مدل هیبریدی GA-SVM و مدل SVM در پیش‌بینی ریسک اعتباری اشخاص حقوقی تفاوتی وجود ندارد.
- H_1 : دقت کلی پیش‌بینی مدل هیبریدی GA-SVM از دقت پیش‌بینی مدل SVM متفاوت است.

خروجی این آزمون آماری برای فرضیه پژوهش در جدول ۴ نمایش داده شده است. این آزمون بصورت دوطرفه انجام شده است، بنابراین اگر در چنین حالتی تفاوت میانگین‌ها معنی‌دار باشد، برای حالت یک طرفه نیز صادق می‌باشد. همانطور که از اعداد و ارقام جدول مشخص است، تفاوت میانگین دو گروه در سطح ۵٪ معنی‌دار است. این، بدین معنی است که فرض H_0 تایید نشده و می‌توان نتیجه گرفت که دقت پیش‌بینی مدل هیبریدی GA-SVM از مدل SVM بیشتر است.

جدول ۴. نتایج آزمون فرضیه آماری

	hybrid GASVM Rtest	SVM Rtest
Mean	۰.۷۴۰۲	۰.۵۷۳۸
Variance	۰.۰۰۰۸۷۵	۰.۰۰۰۷۱۵
Observations	۱۵	
Pearson Correlation	۰.۰۶۵۸۸۷۲۱۳	
Hypothesized Mean Difference	۰	
Df	۱۴	
t Stat	۵	
P(T<=t) one-tail	۵.۷۳E-۰.۵	
t Critical one-tail	۲.۶۲۴۴۹۴	
P(T<=t) two-tail	۰.۰۰۰۰۱۱۵	
t Critical two-tail	۲.۹۷۵۸۴۳	

۷- نتیجه گیری و بحث

یکی از موضوعات مهم در خصوص اعطای تسهیلات بانکی احتمال عدم بازپرداخت وام توسط وام گیرندگان می باشد. برآورد اینکه شرکتی در آینده دچار ورشکستگی می شود برای تسهیلات دهندگان و طلبکاران از اهمیت بالایی برخوردار است به همین دلیل یافتن مدلی که بتواند بهترین دسته بندی را در خصوص شرکتها ارائه نماید همواره مورد توجه بوده است بنابراین لزوم شناخت عواملی که باعث عدم بازپرداخت وامها می شود، در زمینه کاهش و کنترل ریسک اعتباری ضروری به نظر می رسد و باعث بهبود در فرایند اعطای اعتبارات می گردد. در این پژوهش به ارزیابی ریسک اعتباری مشتریان حقوقی بانک تجارت با بکار گیری ماشین بردار پشتیبان و الگوریتم ژنتیک پرداخته شده است. مدل های آماری در ارزیابی ریسک اعتباری در برخی شرایط (در صورتی که پیش فرض های آماری آنها فراهم باشد) در حل مسائل به خوبی جواب می دهند به همین دلیل در این پژوهش به کمک الگوریتم ژنتیک برای بهینه سازی ماشین بردار پشتیبان به ارزیابی ریسک مشتریانی که از بانک تجارت تسهیلات مالی دریافت کرده اند، پرداخته شده است. در این پژوهش یک فرضیه اصلی مطرح گردیده است که به مقایسه دقت دو مدل SVM و مدل هیبریدی GA-SVM پرداخته شده است. نتایج آزمون های آماری حاکی از این موضوع می باشد که مدل هیبریدی GA-SVM از دقت بالاتری در پیش بینی ریسک اعتباری مشتریان برخوردار می باشد. در مدل

هیبریدی GA-SVM به بهینه‌سازی داده‌های ورودی مدل SVM از طریق الگوریتم ژنتیک پرداخته شده است که این موضوع سبب عملکرد بهتر مدل گردیده است. مقایسه نتایج این مقاله با یافته‌های سه‌گانه (۱۳۸۹) با عنوان "طراحی مدل رتبه‌بندی اعتباری مبتنی بر متغیرهای درون شرکتی و صنعت با استفاده از شبکه‌های عصبی مصنوعی" نشان می‌دهد، مدل ترکیبی GA-SVM در خصوص پیش‌بینی ریسک اعتباری مشتریان حقوقی بهتر از مدل رگرسیون لجستیک و شبکه عصبی عمل می‌کند. در خصوص مشکلات و محدودیت‌های تحقیق می‌توان به عدم تمرکز اطلاعات مربوط به مشتریان بانک و فرآیند سخت و زمان‌بر دستیابی به اطلاعات اعتباری مشتریان اشاره کرد. محققان در تحقیقات آتی می‌توانند به مقایسه مدل ترکیبی GA-SVM با سایر مدل‌ها از جمله مدل تعدیل شده پارامترها با استفاده از الگوریتم ژنتیک SVM و شبکه عصبی بپردازند آنها همچنین می‌توانند از سایر مدل‌های بهینه‌ساز نظیر الگوریتم مورچگان و مدل‌های ترکیبی آن برای پیش‌بینی ریسک اعتباری مشتریان بهره‌گیرند.

فهرست منابع

- * جلال سه‌گانه، سه‌گانه (۱۳۸۹). طراحی مدل رتبه‌بندی اعتباری مبتنی بر متغیرهای درون شرکتی و صنعت با استفاده از شبکه‌های عصبی مصنوعی. پایان‌نامه کارشناسی ارشد دانشکده مدیریت دانشگاه تهران
- * طورزنی کوروش (۱۳۸۹). طراحی مدلی رتبه‌بندی ریسک اعتباری اشخاص حقیقی. پایان‌نامه کارشناسی ارشد دانشکده مدیریت دانشگاه تهران.
- * میرفیض فلاح شمس (۱۳۸۴). طراحی و تبیین مدل ریسک اعتبارات در نظام بانکی کشور. پایان‌نامه کارشناسی ارشد
- * منیژه افشار (۱۳۸۶). بررسی کاربرد استفاده از مدل K.M.V در پیش‌بینی ریسک اعتباری مشتریان بانک‌ها و موسسات اعتباری. پایان‌نامه کارشناسی ارشد دانشکده مدیریت دانشگاه تهران
- * Altman, Edward I. (1968). Financial ratios, discriminant analysis and the prediction of corporate bankruptcy. *Journal of Finance*. Vol. 23 (4), pp. 589-609
- * Bart Baesens, Rudy Setiono, Christophe Mues, Jan Vanthienen (2003). Using Neural Network Rule Extraction and Decision Tables for Credit-Risk Evaluation. *Management Science*, Vol 49, No 3, pp. 312-329
- * Beaver (1966). Financial ratios predictors of failure. *Journal of Accounting Research*, Vol. 4, pp. 71-111.
- * Fitzpatrick, P.J (1931). *Symptoms of Industrial failures*. Washington: Catholic University of America Press.
- * Gunn, S.R (1998). *Support Vector Machines for Classification and Regression*. Technical Report, School Of Electronics and Computer.
- * John C.B. Cooper (1999). Artificial neural network versus Multivariate statistics: an application from economics. *Journal of Applied Statistics*, Vole. 26, pp. 909-921

- * Liao, T. FA۱۹۹۲). Modified GMDH Approach for Social Science Research: Exploring Patterns of Relationships in the data. *Quality and Quantity*, Vol 26 ,pp. 19- 38
- * . Salchenberge, Linda M. E. Mine Cinar, Nicholas A. Lash. (1992). *Neural Networks: A New Tool for Predicting Thrift Failures*, *Decision Sciences*, Volume 23, Issue 4, pp. 899-916
- * Myoung-Jong Kim, Ingoo Han (2003). The discovery of experts' decision rules from qualitative bankruptcy data using genetic algorithms. *Expert Systems with Applications*, Vol 25, Issue 4, pp. 637-646
- * Shin S. Kyung, Lee S. Taik, & Kim J. Hyun (2005). An application of support vector machines in bankruptcy prediction model. *Expert Systems with Applications*, Vol 28,pp. 127-135
- * Shin S. Kyung, Lee S. Taik, & Kim J. Hyun (2005). An application of support vector machines in bankruptcy prediction model", *Expert Systems with Applications*, Vol 28, pp. 127-135
- * Shiyi Chen, W. K. Härdle & R. A. Moro (2011). Modeling default risk with support vector machines. *Quantitative Finance*, Vol 11,pp. 135-154
- * T.Kavzoglu, I.Colkesen (2009). A kernel functions analysis for support vector machines for land cover classification. *International Journal of Applied Earth Observation and Geoinformation*, pp. 352-359.
- * Tian-Shyug Leea, Chih-Chou Chiu, Chi-Jie Lu, I-Fei Chen (2002). Credit scoring using the Hybrid Neural Discriminant technique. *Expert System Applications* .Vol. 23, pp.245-254.
- * West David. (2000). Neural network credit scoring. *Computer and Operation Research*, pp. 1131-1152
- * Xiaoyong Liu, Hui Fu & Weiwei Lin (2010). A Modified Support Vector Machine model for Credit Scoring. *International Journal of Computational Intelligence Systems*, vol 3, pp. 797-804
- * Zan Huang, Hsinchun Chena, Chia-Jung Hsu, Wun- Hwa Chen, Soushan Wu. (2004). Credit rating analysis with support vector machines and neural networks: a market comparative stud

یادداشت‌ها

- ¹ Fitzpatrick
- ² Discriminate Analysis
- ³ Edward Altman
- ⁴ Multivariate Discriminate Analysis
- ⁵ Learning Vector Quantization
- ⁶ Fuzzy Adaptive Resonance
- ⁷ Radial Basis Function
- ⁸ Mixture Of Experts
- ⁹ Kernel density Estimation
- ¹⁰ Back Propagation Neural Networks
- ¹¹ Structural Risk Minimization (SRM)
- ¹² Empirical Risk Minimization (ERM)