****

****

دانشگاه آزاد اسلامي

واحد تهران شمال

دانشکده فنی مهندسی، گروه صنایع

پایان‌نامه براي دريافت درجه كارشناسي ارشد (M.Sc) / دکتری (Ph.D)

گرايش: **مهندسی مالی**

عنوان:

**تلفیق ماتریس زیان و طبقه بندی ترکیبی مبتنی بر روش بگینگ به‌منظور ارزیابی ریسک اعتباری**

استاد راهنما:

دکتر **شروین اسدزاده**

استاد مشاور:

دکتر (نام استاد مشاور)

نگارش:

**نیما عالم شناس**

**زمستان** **1403**

**A blue and black logo

Description automatically generated**

تعهدنامه اصالت رساله یا پایان‌نامه

این‌جانب نیما عالم شناس دانش‌آموخته مقطع کارشناسی ارشد ناپیوسته/دکترای تخصصی در رشته مهندسی صنایع گرایش مهندسی مالی که در تاریخ .................. از پایان‌نامه/ رساله خود تحت عنوان

"تلفیق ماتریس زیان و طبقه‌بندی ترکیبی مبتنی بر روش بگینگ به منظور ارزیابی ریسک اعتباری"

با کسب نمره ................. و درجه ..................... دفاع نموده‌ام بدین‌وسیله متعهد می‌شوم:

1. این پایان‌نامه/ رساله حاصل تحقیق و پژوهش انجام‌شده توسط این‌جانب بوده و در مواردی که از دستاوردهای علمی و پژوهشی دیگران(اعم از پایان‌نامه، کتاب،مقاله و....) استفاده نموده‌ام، مطابق ضوابط و رویه موجود، نام منبع مورداستفاده و سایر مشخصات آن را در فهرست مربوطه ذکر و درج کرده‌ام.
2. این پایان‌نامه/ رساله قبلاً برای دریافت هیچ مدرک تحصیلی(هم‌سطح،پایین‌تر یا بالاتر) در سایر دانشگاه‌ها و مؤسسات آموزشی عالی ارائه نشده است.
3. چنانچه بعد از فراغت تحصیل، قصد استفاده و هرگونه بهره‌برداری اعم از چاپ کتاب،ثبت اختراع و... از این پایان‌نامه داشته باشم،از حوزه معاونت پژوهشی واحد مجوزهای مربوطه را اخذ نمایم.
4. چنانچه در هر مقطعی زمانی خلاف موارد فوق ثابت شود،عواقب ناشی از آن را می‌پذیرم و واحد دانشگاهی مجاز است با این‌جانب مطابق ضوابط و مقررات رفتار نموده و در صورت ابطال مدرک تحصیلی‌ام هیچ‌گونه ادعایی نخواهم داشت.

نام و نام خانوادگی

تاریخ و امضاء

تقدیم به

پدرم، نماد استقامت و شرافت  
 و مادرم، سرچشمه عشق و مهر

با سپاس از

آنکه با مهر، درخت دانش را بارور کرد  
با تلاش و صبوری، مسیرمان را هموار کرد  
خانواده‌ای که پشتوانه این راه دشوار بودند  
و استادی که فانوس علم را در دستمان سپرد

فهرست مطالب

عنوان صفحه

[چکیده 1](#_Toc187755666)

[فصل اول 2](#_Toc187755667)

[بیان مسئله 3](#_Toc187755668)

[1-2 اهمیت و ضرورت انجام تحقیق 5](#_Toc187755669)

[1-3 جنبه جديد بودن و نوآوري در تحقيق 7](#_Toc187755670)

[1-4 سؤالات تحقیق 9](#_Toc187755671)

[1-5 فرضيه‏هاي تحقیق 9](#_Toc187755672)

[1-6 روش شناسی تحقیق 10](#_Toc187755673)

[1-7 جامعه آماري 11](#_Toc187755674)

[1-8 مدل مورد استفاده در تحقیق و تعریف متغیرها 12](#_Toc187755675)

[1-9 ساختار پایان‌نامه 14](#_Toc187755676)

[فصل دوم: 15](#_Toc187755677)

[مروری بر مطالعات پیشین 15](#_Toc187755678)

[15](#_Toc187755679)

[2-1 مقدمه 16](#_Toc187755680)

[2-2 مبانی نظری 16](#_Toc187755681)

[2-2-1 مفهوم ریسک و مدیریت آن 17](#_Toc187755682)

[2-2-3 مدل‌های ارزیابی ریسک اعتباری 18](#_Toc187755683)

[2-2-4 ماتریس زیان در ارزیابی ریسک 18](#_Toc187755684)

[2-2-5 روش بگینگ در یادگیری ماشین 19](#_Toc187755685)

[2-2-6 ترکیب روش‌های طبقه‌بندی 20](#_Toc187755686)

[2-3 پیشینه تحقیق 20](#_Toc187755687)

[2-3-1 تحقیقات داخلی 20](#_Toc187755688)

[2-3-2 تحقیقات خارجی 21](#_Toc187755689)

[2-4 جمع‌بندی 26](#_Toc187755690)

[فصل سوم: 27](#_Toc187755691)

[روش تحقیق 27](#_Toc187755692)

[(مراحل انجام پژوهش) 27](#_Toc187755693)

[27](#_Toc187755694)

[3-1 عنوان موردنظر 28](#_Toc187755695)

[فصل چهارم: 30](#_Toc187755696)

[شبیه‌سازی و تحلیل نتایج 30](#_Toc187755697)

[(نتایج و بحث) 30](#_Toc187755698)

[4-1 عنوان موردنظر 31](#_Toc187755699)

[فصل پنجم: 33](#_Toc187755700)

[نتیجه‌گیری و پیشنهادها 33](#_Toc187755701)

[33](#_Toc187755702)

[5-1 نتیجه‌گیری 34](#_Toc187755703)

[5-2 پيشنهادها 34](#_Toc187755704)

[پیوست الف (در صورت وجود) 35](#_Toc187755705)

[فهرست 37](#_Toc187755706)

[منابع و مأخذ 37](#_Toc187755707)

[37](#_Toc187755708)

[فهرست منابع و مأخذ 38](#_Toc187755709)

فهرست جدول‌ها

عنوان صفحه

فهرست شکل­ها

عنوان صفحه

چکیده

متن چکیده

کلمات کلیدی: متن کلمات کلیدی.

فصل اول

بیان مسئله

ریسک اعتباری به‌عنوان یکی از مهم‌ترین چالش‌های مالی، به احتمال نکول وام‌ها یا عدم بازپرداخت تعهدات مالی توسط وام‌گیرندگان اشاره دارد. این موضوع تأثیر مستقیمی بر عملکرد بانک‌ها، مؤسسات مالی و در نهایت اقتصاد کلان دارد. ناکامی در مدیریت این ریسک می‌تواند منجر به افزایش بدهی‌های غیرجاری، کاهش منابع قابل‌تخصیص بانک‌ها، و ایجاد فشار بر توانایی آن‌ها در اعطای وام‌های جدید شود. چنین وضعیتی، نه‌تنها عملکرد مالی مؤسسات را مختل می‌کند، بلکه به کاهش اعتماد مشتریان و سرمایه‌گذاران نیز منجر می‌شود. بنابراین، مدیریت کارآمد ریسک اعتباری، کلید حفظ ثبات مالی بانک‌ها و مؤسسات مالی است.

روش‌های سنتی مدیریت ریسک، عمدتاً مبتنی بر تصمیم‌گیری دو کلاسه بودند که مشتریان را به دو دسته خوش‌حساب و بدحساب تقسیم می‌کنند. این رویکرد ساده‌سازی شده، در مواجهه با داده‌های نامطمئن و ناقص، ناکارآمدی‌هایی از خود نشان می‌دهد. تصمیم‌گیری‌های نادرست در این روش‌ها منجر به افزایش خطاهای تصمیم‌گیری و در نهایت کاهش بهره‌وری بانک‌ها می‌شود. همچنین، در این روش‌ها دسته‌ای از مشتریان که اطلاعات کافی برای قضاوت قاطع درباره وضعیت آن‌ها وجود ندارد، نادیده گرفته می‌شوند. این چالش‌ها ضرورت استفاده از رویکردهای نوین را افزایش داده است .

پیشگیری از نکول مستلزم به‌کارگیری روش‌های پیشرفته ارزیابی و مدیریت ریسک اعتباری است. بانک‌ها می‌توانند با استفاده از مدل‌های پیش‌بینی مبتنی بر داده، مشتریان پرخطر را شناسایی کرده و سیاست‌های اعتباری مناسب را اعمال کنند. این راهکارها به کاهش نکول و افزایش ثبات مالی بانک کمک می‌کند.

نکول وام‌ها تأثیر مستقیمی بر نقدینگی بانک‌ها دارد و می‌تواند منجر به کاهش منابع مالی قابل‌دسترس برای عملیات جاری آن‌ها شود. با عدم بازپرداخت وام توسط مشتریان، بانک‌ها با کاهش جریان‌های نقدی ورودی مواجه شده و در نتیجه، توانایی آن‌ها در تأمین مالی سایر وام‌ها یا پرداخت به سپرده‌گذاران کاهش می‌یابد. این موضوع فشار مضاعفی بر منابع مالی بانک وارد کرده و ممکن است نیاز به تأمین مالی اضطراری از طریق بازارهای سرمایه یا استفاده از ذخایر نقدی را افزایش دهد. در شرایط حادتر، نکول می‌تواند به افزایش بدهی‌های غیرجاری در ترازنامه بانک منجر شود و سلامت مالی آن را به خطر بیندازد. این وضعیت، نه‌تنها هزینه‌های تأمین مالی بانک را افزایش می‌دهد، بلکه اعتماد مشتریان و سرمایه‌گذاران به بانک را نیز تضعیف می‌کند. مدیریت صحیح ریسک اعتباری و اجرای سیاست‌های پیشگیرانه برای کاهش نکول، کلیدی برای حفظ نقدینگی و ثبات مالی بانک‌ها است.

تحلیل جریان نقدی یکی از ابزارهای اساسی در ارزیابی سلامت مالی بانک‌ها و شرکت‌ها است که میزان ورود و خروج وجوه نقد در یک دوره زمانی مشخص را بررسی می‌کند. این تحلیل کمک می‌کند تا توانایی سازمان در تأمین تعهدات مالی، پرداخت بدهی‌ها، و سرمایه‌گذاری در پروژه‌های جدید ارزیابی شود. در بانک‌ها، جریان نقدی مثبت نشان‌دهنده عملکرد مطلوب در مدیریت وام‌ها و بازپرداخت‌ها است، در حالی که جریان نقدی منفی می‌تواند نشان‌دهنده افزایش نکول و مشکلات نقدینگی باشد. این تحلیل به مدیران کمک می‌کند تا استراتژی‌هایی برای مدیریت بهینه منابع مالی و کاهش ریسک نقدینگی تدوین کنند. همچنین، پیش‌بینی جریان نقدی با استفاده از مدل‌های پیشرفته تحلیل داده، می‌تواند تصمیم‌گیری‌های آگاهانه‌تر در تخصیص منابع و سرمایه‌گذاری‌ها را ممکن سازد. بهره‌گیری از تحلیل جریان نقدی، نقش مهمی در تقویت پایداری مالی و اعتماد سرمایه‌گذاران و مشتریان به بانک‌ها و مؤسسات مالی ایفا می‌کند.

بحران‌های بانکی اغلب ناشی از عدم مدیریت صحیح ریسک‌های اعتباری، نقدینگی، و بازار هستند و می‌توانند تأثیرات مخربی بر اقتصاد کلان داشته باشند. کاهش و کنترل این ریسک‌ها نیازمند اجرای چارچوب‌های قوی مدیریت ریسک است که شامل شناسایی، ارزیابی، و پایش مداوم ریسک‌ها می‌شود. بانک‌ها می‌توانند با به‌کارگیری مدل‌های پیش‌بینی ریسک و تحلیل داده‌های تاریخی، احتمال وقوع بحران‌ها را کاهش دهند.از سوی دیگر، تقویت شفافیت در گزارش‌دهی مالی و افزایش اعتماد عمومی از طریق ارتباطات شفاف با سهام‌داران و مشتریان، نقشی حیاتی در مدیریت و کنترل ریسک بحران‌های بانکی دارد. اجرای این رویکردها به حفظ ثبات نظام مالی و جلوگیری از بروز بحران‌های سیستمی کمک می‌کند.

یکی از راهکارهای نوین برای مدیریت بهتر ریسک اعتباری، استفاده از ماتریس زیان است. این رویکرد، به‌جای محدود کردن تصمیم‌گیری به دو گزینه ساده، داده‌ها را به سه دسته خوش‌حساب، بدحساب و مرزی تقسیم می‌کند. دسته‌بندی موارد مرزی فرصتی برای جمع‌آوری اطلاعات بیشتر و اتخاذ تصمیمات دقیق‌تر فراهم می‌آورد. این دسته‌بندی، انعطاف‌پذیری بیشتری در مواجهه با داده‌های نامطمئن ایجاد می‌کند و امکان کاهش ریسک‌های ناشی از تصمیم‌گیری‌های اشتباه را فراهم می‌سازد.

موفقیت در ارزیابی ریسک اعتباری به تعیین دقیق آستانه‌های α و β وابسته است که مرزهای این دسته‌ها را مشخص می‌کنند. اگر این آستانه‌ها به‌درستی تعیین نشوند، دسته‌بندی نادرست می‌تواند منجر به افزایش خطاها و کاهش دقت مدل شود.در این پژوهش، آستانه‌های α و β با استفاده از تحلیل ماتریس زیان و با توجه به پیامدهای تصمیم‌گیری به‌گونه‌ای تعیین می‌شوند که زیان کل سیستم کاهش یابد. این روش به مدل امکان می‌دهد تا تصمیمات مؤثرتری در ارزیابی ریسک اعتباری اتخاذ کرده و دقت آن را ارتقا دهد . برای تحلیل و پیش‌بینی دقیق‌تر دسته موارد مرزی، از رویکرد طبقه‌بندی ترکیبی با روش بگینگ استفاده می‌شود. بگینگ با اجرای چندین مدل پایه به‌صورت موازی و ترکیب خروجی آن‌ها از طریق روش‌هایی مانند رأی‌گیری یا میانگین‌گیری، دقت تحلیل و پیش‌بینی را بهبود می‌بخشد و پایداری مدل را افزایش می‌دهد.

استفاده از این روش‌ها، علاوه بر افزایش دقت پیش‌بینی، نیاز به جمع‌آوری اطلاعات اضافی و پردازش داده‌ها را کاهش می‌دهد. ترکیب ماتریس زیان و طبقه‌بندی ترکیبی، چارچوبی جامع و نوآورانه برای مدیریت ریسک اعتباری ارائه می‌دهد. این چارچوب نه‌تنها به بانک‌ها و مؤسسات مالی کمک می‌کند تا بهره‌وری خود را افزایش دهند، بلکه توانایی آن‌ها در تخصیص بهینه منابع مالی و مدیریت بهتر سرمایه‌ها را نیز تقویت می‌کند.

این تحقیق با هدف توسعه چنین چارچوبی انجام شده است و تلاش دارد با کاهش خطاهای تصمیم‌گیری و ارتقای فرایندهای پیش‌بینی، عملکرد بانک‌ها و مؤسسات مالی را بهبود بخشد. متغیرهای کلیدی این تحقیق شامل سابقه اعتباری، نرخ نکول، درآمد، نسبت بدهی به درآمد و آستانه‌های α و β است. این متغیرها، پس از تحلیل در قالب مدل پیشنهادی برای افزایش دقت پیش‌بینی و بهبود بهره‌وری سیستم‌های مالی به کار گرفته می‌شوند. این رویکرد با فراهم کردن شفافیت بیشتر در مدیریت ریسک، به ارتقای سطح اطمینان میان بانک‌ها و مشتریان کمک خواهد کرد.

1-2 اهمیت و ضرورت انجام تحقیق

ریسک اعتباری به‌عنوان یکی از چالش‌های کلیدی در نظام بانکی و مالی، تأثیری عمیق بر پایداری اقتصادی و کارایی مؤسسات مالی دارد. در محیط رقابتی فعلی بین بانک‌های کشور، جذب مشتریان جدید و حفظ مشتریان فعلی به یک اولویت استراتژیک تبدیل شده است. افزایش تعداد مؤسسات مالی و ارائه خدمات متنوع، بانک‌ها را به سمت استفاده از روش‌های نوین مدیریت ریسک و تحلیل داده‌های مشتریان سوق داده است. رقابت فشرده، بانک‌ها را ملزم می‌کند تا نه تنها دقت خود را در ارزیابی ریسک اعتباری افزایش دهند، بلکه با بهره‌گیری از فناوری‌های پیشرفته مانند یادگیری ماشین، مزیت رقابتی خود را در بازار تثبیت کنند.

ریسک اعتباری نه تنها عملکرد داخلی بانک‌ها و مؤسسات مالی را تحت تأثیر قرار می‌دهد، بلکه ابعاد گسترده‌تری مانند اشتغال، سرمایه‌گذاری و رفاه اجتماعی را نیز شامل می‌شود. عدم مدیریت صحیح این ریسک می‌تواند منجر به کاهش اعتماد عمومی، افت سرمایه‌گذاری‌های داخلی و خارجی، و کاهش پایداری اقتصاد کلان شود.

افزایش تقاضا برای وام‌ها و رشد پیچیدگی در داده‌های مالی، بانک‌ها را مجبور به استفاده از روش‌های پیشرفته‌تر برای مدیریت ریسک کرده است. روش‌های سنتی که عموماً بر پایه تصمیم‌گیری دوکلاسه بنا شده‌اند، به دلیل ناکارآمدی در تحلیل داده‌های ناقص و ناتوانی در مدیریت مشتریان مرزی، دیگر پاسخگوی نیازهای امروز نیستند. علاوه بر این، بحران‌های مالی و اقتصادی مانند رکود یا همه‌گیری‌های جهانی، ضرورت بهره‌گیری از رویکردهای انعطاف‌پذیر و پویا را افزایش داده‌اند. بانک‌ها نیازمند مدل‌هایی هستند که بتوانند در شرایط عدم قطعیت و نوسانات اقتصادی، پایداری مالی را تضمین کنند و ریسک‌های مرتبط با اعتباردهی را به حداقل برسانند.

این روش‌ها معمولاً منجر به افزایش خطاهای تصمیم‌گیری، اتلاف منابع و کاهش بهره‌وری سیستم مالی می‌شوند.یکی از خلأهای تحقیقاتی موجود در این حوزه، نبود یک چارچوب برای تعیین آستانه‌های تصمیم‌گیری آلفا و بتا می باشد. تحقیقات پیشین اغلب به دلیل تمرکز بر مدل‌های غیرتطبیقی، کاستی‌هایی در تحلیل داده‌های پیچیده و متغیر داشته‌اند. این محدودیت‌ها منجر به کاهش دقت در پیش‌بینی و افزایش خطاهای تصمیم‌گیری شده است. همچنین، در تحلیل موارد مرزی و داده‌های ناقص، این مدل‌ها نتوانسته‌اند کارایی کافی ارائه دهند.

دسته‌بندی دقیق موارد مرزی که در این رویکرد نقش مهمی دارند به دلیل نبود روش‌های مؤثر تحلیل، همچنان یک چالش باقی مانده است. تحقیقات پیشین کمتر بر استفاده از ماتریس زیان و طبقه بندی ترکیبی مبتنی بر روش بگینگ متمرکز شده‌اند.این تحقیق با هدف ارائه یک مدل ترکیبی مبتنی بر ماتریس زیان و طبقه‌بندی ترکیبی، به دنبال رفع این چالش‌ها است. در این مدل، آستانه‌های تصمیم‌گیری با استفاده از ماتریس زیان تعیین می‌شوند. این رویکرد امکان سازگاری بیشتر با داده‌های متغیر و شرایط پویا را فراهم کرده و به بهبود دقت و کارایی تصمیم‌گیری کمک می‌کند. علاوه بر این، موارد مرزی که در روش‌های سنتی نادیده گرفته می‌شدند، در این تحقیق با استفاده از تکنیک بگینگ تحلیل می‌شوند.

بگینگ، با اجرای چندین مدل پایه به‌صورت موازی و ترکیب خروجی آن‌ها از طریق روش‌هایی مانند رأی‌گیری یا میانگین‌گیری، دقت تحلیل و پیش‌بینی را بهبود می‌بخشد و پایداری مدل را افزایش می‌دهد.این نوآوری نه‌تنها از نظر عملی به بانک‌ها و مؤسسات مالی در کاهش خطاهای تصمیم‌گیری و بهبود فرایندها کمک می‌کند، بلکه از منظر نظری نیز به توسعه دانش در حوزه مدیریت ریسک اعتباری و ترکیب روش‌های پیشرفته کمک می‌کند.

علاوه بر این، تحقیق حاضر با تأکید بر شفافیت و رعایت اصول اخلاقی، تلاش می‌کند تا اعتماد عمومی به نظام مالی را افزایش دهد. استفاده از مدل‌های پیشرفته مانند ماتریس زیان و بگینگ می‌تواند از بروز رفتارهای تبعیض‌آمیز یا غیراخلاقی در تصمیم‌گیری جلوگیری کند و به بهبود شفافیت در تخصیص منابع مالی منجر شود.این رویکرد با کاهش نیاز به منابع عملیاتی، افزایش دقت پیش‌بینی و بهبود تخصیص منابع، می‌تواند نقش مهمی در تقویت سیستم‌های مالی ایفا کند. همچنین، استفاده از تکنیک‌های پیشرفته یادگیری ماشین در کنار ماتریس زیان ، بانک‌ها را قادر می‌سازد تا با چالش‌های ناشی از داده‌های ناقص و نامطمئن بهتر مقابله کنند و عملکرد خود را در بازار رقابتی بهبود بخشند.

1-3 جنبه جديد بودن و نوآوري در تحقيق

در این تحقیق، تعیین آستانه‌ها با بهره‌گیری از ماتریس زیان صورت می‌پذیرد. این رویکرد، روشی نظام‌مند برای تنظیم دقیق آستانه‌ها ارائه می‌دهد. این روش با بهره‌گیری از ماتریس زیان و افزودن دسته‌ای تحت عنوان موارد مرزی به رویکرد سنتی دو‌کلاسه، شرایطی را برای مدیریت بهینه داده‌های نامطمئن فراهم می‌آورد. این روش با ارائه سه دسته خوش‌حساب، بدحساب و مرزی، نه‌تنها خطاهای تصمیم‌گیری را کاهش می‌دهد، بلکه فرصت‌هایی برای جمع‌آوری اطلاعات بیشتر درباره موارد مرزی فراهم می‌آورد. بااین‌حال، موفقیت این رویکرد به تعیین دقیق آستانه‌های تصمیم‌گیری آلفا و بتا وابسته است که امکان تطبیق آستانه‌ها با شرایط متغیر بازار و ویژگی‌های داده‌های ورودی را فراهم می‌کند. برخلاف روش‌های دستی، که به‌طور ثابت و بدون انعطاف تعیین می‌شوند، این رویکرد به مدل اجازه می‌دهد که با داده‌های نامطمئن یا نوسانات بازار بهتر سازگار شود. تعیین این آستانه‌ها نقش اساسی در افزایش دقت طبقه‌بندی موارد مرزی و کاهش خطاهای پیش‌بینی ایفا می‌کند.

این رویکرد ترکیبی نه‌تنها نقاط قوت ماتریس زیان را در مدیریت خطاهای تصمیم‌گیری به‌کار می‌گیرد، بلکه با استفاده از ساختار بگینگ، یکپارچگی مدل‌ها را تضمین می‌کند. هم‌افزایی میان این دو روش باعث افزایش تطبیق‌پذیری با شرایط متغیر داده‌ها و کاهش خطاهای ناشی از سوگیری مدل‌ها می‌شود. به‌ویژه، رویکرد تلفیقی ارائه‌شده امکان استفاده از چندین مدل پایه را فراهم می‌آورد که در نهایت دقت پیش‌بینی را بهبود بخشیده و توانایی سیستم در مدیریت ریسک‌های پیچیده را تقویت می‌کند.

توانایی این مدل در تحلیل داده‌های نامطمئن یکی از مهم‌ترین نوآوری‌های آن به شمار می‌رود. در شرایطی که داده‌ها دارای عدم قطعیت بالا یا تنوع گسترده هستند، این رویکرد با ارائه راهکاری ساختاریافته، امکان تحلیل دقیق‌تری را فراهم می‌کند. توانایی تعیین آستانه‌های پویا و سازگار با تغییرات محیطی، این مدل را به ابزاری قابل اعتماد برای پیش‌بینی ریسک در محیط‌های پرچالش تبدیل کرده است. همچنین، مدل قادر است عدم قطعیت داده‌ها را به اطلاعات معنادار تبدیل کرده و دقت تصمیم‌گیری را افزایش دهد.

یکی از ویژگی‌های برجسته این تحقیق، بهره‌گیری از مدل‌های پیشرفته در ساختار بگینگ است. به‌کارگیری مدل‌هایی نظیر شبکه‌های عصبی، درخت تصمیم و نزدیک‌ترین همسایه باعث کاهش واریانس و بهبود پایداری پیش‌بینی می‌شود. این مدل‌ها با رأی‌گیری وزنی، خروجی نهایی را با دقت بالاتری ارائه می‌دهند، که این امر منجر به بهبود عملکرد سیستم در شرایط پیچیده می‌گردد.

علاوه بر دقت بالای پیش‌بینی، رویکرد پیشنهادی از نظر کاربرد عملیاتی نیز برای مؤسسات مالی حائز اهمیت است. این روش، منابع پردازشی کمتری را مصرف کرده و زمان تحلیل را کاهش می‌دهد. به‌این‌ترتیب، بانک‌ها و مؤسسات مالی می‌توانند با استفاده از این رویکرد، بهره‌وری خود را افزایش داده و خطاهای پرهزینه را به حداقل برسانند. این ویژگی به‌ویژه در محیط‌های رقابتی و تحت فشار زمانی، ارزش افزوده قابل توجهی ایجاد می‌کند.

یکی از دستاوردهای کلیدی این تحقیق، پایداری پیش‌بینی‌ها حتی در مواجهه با تغییرات ناگهانی در داده‌های ورودی است. این ویژگی به مدل اجازه می‌دهد که در شرایط ناپایدار، مانند تغییرات اقتصادی یا نوسانات بازار، عملکرد دقیقی داشته باشد. همچنین، خطاهای پیش‌بینی به دلیل استفاده از آستانه‌های تطبیق‌پذیر و مدل‌های ترکیبی کاهش یافته و دقت تصمیم‌گیری بهبود می‌یابد.

برای ارزیابی عملکرد این مدل، مقایسه‌ای میان روش پیشنهادی و سایر رویکردهای موجود انجام شد. نتایج نشان داد که این مدل با کاهش خطاهای پیش‌بینی و افزایش دقت در تحلیل موارد مرزی، عملکرد بهتری نسبت به روش‌های رایج مانند شبکه‌های عصبی مستقل یا درخت تصمیم سنتی دارد. این ارزیابی تطبیقی برتری روش پیشنهادی در مدیریت ریسک اعتباری را به‌وضوح نشان می‌دهد.

این ساختار از نقاط قوت هر دو بخش بهره می‌برد؛ ماتریس زیان و تعیین آستانه‌ها برای دسته‌بندی اولیه داده‌ها، و روش طبقه‌بندی ترکیبی با رویکرد بگینگ برای تحلیل دقیق‌تر موارد مرزی. این هماهنگی بین دو مرحله به بهبود دقت پیش‌بینی، کاهش خطاهای تصمیم‌گیری و افزایش کارایی مدیریت ریسک اعتباری منجر می‌شود. این رویکرد نه‌تنها قابلیت تحلیل دقیق‌تر را فراهم می‌کند، بلکه با کاهش منابع و زمان پردازش، یک راه‌حل عملی و مؤثر برای بانک‌ها و مؤسسات مالی ارائه می‌دهد.

1-4 سؤالات تحقیق

آیا تلفیق ماتریس زیان و طبقه‌بندی ترکیبی مبتنی بر روش بگینگ می‌تواند منجر به افزایش دقت در پیش‌بینی ریسک اعتباری و بهبود تخصیص منابع مالی شود؟

آیا طبقه‌بندی ترکیبی مبتنی بر روش بگینگ در مدیریت و ارزیابی ریسک اعتباری عملکرد بهتری نسبت به روش‌های طبقه‌بندی پایه ارائه می‌دهد؟

1-5 فرضيه‏هاي تحقیق

تعیین آستانه‌های α و β با استفاده از ماتریس زیان می‌تواند دقت پیش‌بینی ریسک اعتباری را افزایش داده و به بهبود کارایی مدل در مدیریت مشتریان خوش‌حساب، بدحساب و موارد مرزی منجر می شود.

استفاده از طبقه‌بندی ترکیبی با روش بگینگ می‌تواند دقت پیش‌بینی و تحلیل موارد مرزی در مدیریت ریسک اعتباری را به‌طور قابل‌توجهی افزایش می دهد.

مدل تلفیقی مبتنی بر ماتریس زیان و طبقه‌بندی ترکیبی می‌تواند به بهبود عملکرد سیستم‌های مدیریت ریسک اعتباری و کاهش خطاهای تصمیم‌گیری در بانک‌ها و مؤسسات مالی منجرمی شود.

1-6 روش شناسی تحقیق

تحقیق حاضر با هدف توسعه یک مدل ترکیبی برای مدیریت ریسک اعتباری طراحی شده است که از ماتریس زیان وطبقه‌بندی ترکیبی بهره می‌برد. این روش در قالب چندین مرحله اجرایی و با تکیه بر داده‌های مشتریان اعتباری از یک پایگاه داده معتبر متعلق به یکی از بانک‌ها پیاده‌سازی می‌شود. نوع داده‌ها شامل اطلاعات مالی، سابقه اعتباری، نرخ نکول و سایر ویژگی‌های مشتریان است که از پایگاه اطلاعاتی این مؤسسه جمع‌آوری می‌شود. داده‌ها به‌صورت ساختارمند و عددی بوده و برای تحلیل و پردازش به فرمت‌های استاندارد تبدیل می‌شوند.

در مرحله اول، داده‌های اولیه از طریق همکاری با بانک جمع‌آوری می‌شوند. این داده‌ها شامل متغیرهای کلیدی مانند درآمد مشتری، سابقه بازپرداخت وام و تاریخچه اعتباری هستند. پس از گردآوری، داده‌ها با استفاده از روش‌های استاندارد پاک‌سازی شده و مقادیر گمشده حذف می‌شوند تا کیفیت و انسجام داده‌ها برای تحلیل تضمین شود.

در مرحله بعد، ماتریس زیان برای تعیین آستانه‌های آلفت و بتا طراحی و پیاده‌سازی می‌شود. ماتریس زیان به‌طور خاص برای تعیین استراتژی بهینه در مدل ترکیبی به کار گرفته می‌شود. این ماتریس وزن‌های متفاوتی را به خطاهای نوع اول و دوم تخصیص می‌دهد تا زیان کلی سیستم کاهش یابد. از طریق ماتریس زیان، تصمیم‌گیری در مورد طبقه‌بندی نمونه‌ها با توجه به اهمیت نسبی هر نوع خطا انجام می‌شود. این فرآیند به مدل اجازه می‌دهد تا تمرکز بیشتری بر کاهش زیان‌های پرهزینه مانند عدم شناسایی نکول‌های پرریسک داشته باشد. همچنین، ماتریس زیان در تنظیم آستانه‌های تصمیم‌گیری و بهینه‌سازی عملکرد کلی مدل نقش کلیدی ایفا می‌کند.

این الگوریتم‌ها با بهره‌گیری از ابزارهای پیشرفته یادگیری ماشین و کتابخانه‌های برنامه‌نویسی زبان پایتون با استفاده از بسته‌هایی نظیر Scikit-learn و TensorFlow توسعه داده می‌شوند. تعیین این آستانه‌ها با هدف بهبود دقت تفکیک داده‌ها و کاهش خطاهای طبقه‌بندی صورت می‌گیرد.

برای تحلیل دقیق‌تر موارد مرزی، از روش طبقه‌بندی ترکیبی با رویکرد بگینگ استفاده می‌شود. روش بگینگ به دلیل توانایی در کاهش واریانس مدل و بهبود پایداری آن انتخاب شده است. این روش با ایجاد چندین مدل پایه و ترکیب پیش‌بینی‌های آن‌ها، عملکرد کلی مدل را ارتقا می‌دهد. بگینگ به‌ویژه در مقابله با داده‌های پر نویز و جلوگیری از بیش‌برازش مؤثر است. در مقایسه با بوستینگ، بگینگ پیچیدگی کمتری دارد و احتمال بیش‌برازش کمتری را به همراه دارد، زیرا به بهبود تدریجی خطاها متکی نیست. این مزایا باعث می‌شود بگینگ انتخابی مناسب برای مسئله حاضر باشد که نیازمند دقت و پایداری بالاست.

در این رویکرد، مدل‌های پایه‌ای مانند درخت‌های تصمیم یا جنگل تصادفی، بر روی نمونه‌هایی از داده‌ها آموزش داده می‌شوند. پیش‌بینی‌های این مدل‌ها به‌طور همزمان ترکیب می‌شوند تا دقت نهایی افزایش یابد. تمامی این مدل‌ها با استفاده از داده‌های آموزشی و آزمایشی اعتبارسنجی شده و با استفاده از معیارهایی نظیر دقت، نرخ خطا ارزیابی می‌شوند.

در نهایت، خروجی مدل به‌صورت جداول و نمودارهای تحلیل‌پذیر ارائه شده و نتایج تحقیق در قالب یک چارچوب مقایسه‌ای با روش‌های موجود بررسی می‌شوند.

مدیریت ریسک‌های احتمالی ناشی از خطای الگوریتم یا محدودیت‌های داده‌ای یکی از چالش‌های اساسی این تحقیق بوده است. برای کاهش اثر خطاهای الگوریتمی، مدل‌ها به‌طور مکرر اعتبارسنجی شده و پارامترهای حساس با استفاده از تحلیل حساسیت تنظیم شده‌اند. استفاده از معیارهای ارزیابی متنوع و تحلیل دقیق نتایج به شناسایی و کاهش ریسک‌های بالقوه کمک کرده است. این فرآیندها با رعایت استانداردهای علمی و نظارتی حوزه مدیریت ریسک اعتباری طراحی شده‌اند تا قابلیت پیاده‌سازی در محیط واقعی را فراهم کنند.

1-7 جامعه آماري

در این تحقیق، جامعه آماری شامل داده‌های اعتباری مربوط به مشتریان یکی از بانک‌ کشور است که اطلاعات جامعی در زمینه سابقه اعتباری مشتریان فراهم می‌کند. این جامعه شامل تمامی مشتریانی است که در بازه زمانی مشخصی درخواست وام یا خدمات مالی داشته‌اند و اطلاعات مربوط به وضعیت بازپرداخت وام‌های آن‌ها در دسترس است. داده‌ها بر اساس ویژگی‌های کلیدی مانند میزان وام، مدت بازپرداخت، درآمد ماهانه و سابقه نکول انتخاب می‌شوند تا تحلیل‌های موردنظر تحقیق را پشتیبانی کنند.

روش نمونه‌گیری در این پژوهش به‌صورت تصادفی طبقه‌بندی‌شده انجام می‌شود. این روش امکان انتخاب نمونه‌ای متنوع و نماینده از تمامی زیرگروه‌های جامعه آماری را فراهم می‌کند. به‌این‌ترتیب، داده‌ها به‌گونه‌ای انتخاب می‌شوند که از نظر توزیع ویژگی‌های کلیدی مانند وضعیت اعتباری (ریسک پایین، ریسک بالا و موارد مرزی) تعادل وجود داشته باشد. این تعادل به بهبود دقت تحلیل‌ها و نتایج مدل پیشنهادی کمک می‌کند.

داده‌ها به دو بخش آموزشی و آزمایشی تقسیم می‌شوند تا امکان ارزیابی عملکرد مدل در شرایط مختلف فراهم شود. این تقسیم‌بندی به بهبود دقت پیش‌بینی و کاهش خطاهای تصمیم‌گیری کمک می‌کند.

در نهایت، ترکیب جامعه آماری گسترده، روش نمونه‌گیری تصادفی طبقه‌بندی‌شده و حجم نمونه مناسب، شرایط لازم برای اجرای موفقیت‌آمیز تحقیق و دستیابی به اهداف پژوهش را فراهم می‌آورد.

1-8 مدل مورد استفاده در تحقیق و تعریف متغیرها

در این تحقیق، هدف اصلی طراحی و استفاده از یک مدل مفهومی است که بتواند با دقت بالا روابط میان متغیرها را تحلیل کرده و ریسک اعتباری را پیش‌بینی کند. این مدل بر اساس روش طبقه‌بندی ترکیبی توسعه یافته است تا تحلیل جامع‌تری از عوامل مؤثر بر ریسک اعتباری ارائه کند. متغیر اصلی تحقیق، ریسک اعتباری است که نشان‌دهنده احتمال نکول یا عدم بازپرداخت تعهدات مالی مشتریان است. برای اندازه‌گیری این متغیر از شاخص‌هایی مانند سابقه نکول، میزان تأخیر در بازپرداخت و نرخ بازپرداخت استفاده شده است. داده‌های مربوط به این شاخص‌ها از پایگاه اطلاعاتی بانک‌ها گردآوری شده و پس از آماده‌سازی به صورت کمی تحلیل می‌شوند.

متغیرهای مستقل شامل مجموعه‌ای از عوامل مالی، اعتباری و جمعیتی هستند که به طور مستقیم یا غیرمستقیم بر ریسک اعتباری تأثیر می‌گذارند. این متغیرها شامل سابقه اعتباری مشتریان، درآمد ماهانه، نوع وام درخواستی، میزان وام، مدت زمان بازپرداخت و نسبت استفاده از اعتبار هستند. هر یک از این متغیرها با دقت بالا از منابع اطلاعاتی معتبر استخراج شده و در قالب داده‌های عددی تحلیل می‌شوند. برای مثال، سابقه اعتباری مشتریان بر اساس تعداد وام‌های دریافت‌شده قبلی و وضعیت بازپرداخت آن‌ها ارزیابی می‌شود. اطلاعات درآمد ماهانه نیز از اسناد مالی ثبت‌شده مشتریان استخراج شده و تأثیر آن بر ریسک اعتباری بررسی می‌شود. در مرحله پیش‌پردازش، داده‌های گمشده با روش‌های تکمیلی جایگزین و مقادیر نادرست تصحیح می‌شوند.

در این مدل، آستانه‌های تصمیم‌گیری مانند α و β نقش کلیدی دارند و مرز میان طبقه‌های مختلف ریسک را مشخص می‌کنند. این مقادیر بر اساس تحلیل ماتریس زیان تنظیم شده و هدف آن‌ها کاهش خطاهای طبقه‌بندی و بهبود عملکرد مدل است. ماتریس زیان هزینه‌های ناشی از طبقه‌بندی اشتباه مشتریان را در بر می‌گیرد و امکان تصمیم‌گیری دقیق‌تر را برای بانک‌ها فراهم می‌کند. داده‌ها در فرآیند پیش‌پردازش به‌گونه‌ای تنظیم می‌شوند که متغیرهای کیفی به صورت عددی کدگذاری شده و داده‌های پرت حذف شوند.

برای تحلیل و پیش‌بینی، روش‌های طبقه‌بندی ترکیبی به‌کار گرفته شده‌اند که از ترکیب چندین مدل پایه برای بهبود دقت و کاهش خطاهای احتمالی استفاده می‌کنند. این روش به کاهش خطاهای طبقه‌بندی و افزایش ثبات نتایج کمک کرده و باعث می‌شود تأثیر داده‌های پرت کاهش یابد. عملکرد مدل با معیارهایی نظیر دقت، حساسیت، ویژگی و میانگین مربع خطا ارزیابی شده تا نتایج به دست آمده اعتبار بالایی داشته باشند.

مدل پیشنهادی با تحلیل دقیق داده‌ها و روابط میان متغیرها، نه تنها دقت پیش‌بینی ریسک اعتباری را افزایش می‌دهد بلکه فرآیند تصمیم‌گیری در بانک‌ها و مؤسسات مالی را نیز بهبود می‌بخشد. این مدل امکان شناسایی دقیق مشتریان با ریسک بالا را فراهم کرده و ابزارهای مؤثری برای مدیریت بهتر ریسک ارائه می‌دهد. بانک‌ها می‌توانند بر اساس خروجی‌های این مدل، استراتژی‌های خود را برای اعطای وام و مدیریت اعتبار بهینه کنند. این تحلیل‌ها همچنین می‌توانند به بهبود تخصیص منابع مالی و کاهش هزینه‌های مرتبط با نکول کمک کنند. در نهایت، این مدل به‌عنوان ابزاری کارآمد، نقش مؤثری در ارتقای بهره‌وری و کاهش ریسک در سیستم بانکی ایفا می‌کند.

1-9 ساختار پایان‌نامه

تحقیق حاضر شامل 5 فصل می‌باشد ، در فصل اول به کلیات پژوهش شامل بیان مسأله، اهمیت و ضرورت تحقیق، اهداف و سؤالات پژوهش، و همچنین ساختار کلی پایان‌نامه پرداخته شده است.  
در فصل دوم ادبیات نظری و پیشینه تحقیق مورد بررسی قرار گرفته است. در این فصل، ابتدا مفاهیم مربوط به ارزیابی ریسک اعتباری، ماتریس زیان و روش‌های ترکیبی مبتنی بر بگینگ تشریح شده و سپس مطالعات انجام‌شده در این زمینه در دو بخش داخلی و خارجی ارائه شده است.  
فصل سوم به روش‌شناسی تحقیق اختصاص دارد. در این فصل، روش‌های جمع‌آوری داده‌ها، جامعه آماری، ابزارهای تحقیق، و روش‌های تجزیه و تحلیل داده‌ها توضیح داده شده است.  
در فصل چهارم، داده‌های گردآوری‌شده تحلیل شده و الگوی پیشنهادی تحقیق که شامل تلفیق ماتریس زیان و روش طبقه‌بندی ترکیبی مبتنی بر بگینگ است، مورد آزمون قرار گرفته است. همچنین، نتایج آزمون فرضیات و تحلیل‌های مربوط به سؤالات پژوهش ارائه شده است.  
در نهایت، فصل پنجم شامل نتایج کلی تحقیق، بحث و بررسی یافته‌ها، محدودیت‌های پژوهش، و ارائه پیشنهادهای کاربردی و پژوهشی برای تحقیقات آتی می‌باشد.

فصل دوم:

مروری بر مطالعات پیشین

(مبانی نظری و پیشینه‌ی تحقیق)

2-1 مقدمه

بانک‌ها، به‌عنوان نهادهای اصلی واسطه‌گری مالی، منابع مالی را از سپرده‌گذاران جمع‌آوری کرده و به بخش‌های مختلف اقتصاد تخصیص می‌دهند. این فرآیند باعث رشد تولید، افزایش اشتغال و تقویت زیرساخت‌های اقتصادی می‌شود. در ایران، به دلیل ساختار بانک‌محور نظام مالی، بانک‌ها نقشی حیاتی در تأمین مالی پروژه‌ها و بخش‌های مختلف اقتصادی ایفا می‌کنند. نظام بانکی، در عین حال که محرک اصلی رشد اقتصادی است، می‌تواند در معرض ریسک‌های مختلف، به‌ویژه ریسک اعتباری قرار بگیرد. ضرورت مدیریت ریسک، به‌ویژه در کشورهایی با اقتصاد بانک‌محور مانند ایران، از اهمیت ویژه‌ای برخوردار است. اگر بانک‌ها در ارزیابی و مدیریت ریسک‌های اعتباری و دیگر ریسک‌های بانکی کوتاهی کنند، پیامدهای منفی آن می‌تواند کل اقتصاد را تحت تأثیر قرار دهد. افزایش نکول و کاهش نقدینگی، از جمله پیامدهای عدم مدیریت صحیح ریسک است که می‌تواند باعث بی‌ثباتی مالی و کاهش اعتماد عمومی به نظام بانکی شود.{Aboutorabi, 2021 #1}

مدل‌های مختلفی برای طبقه‌بندی ریسک در نظام بانکی ارائه شده است. همچنین، در بررسی عقود اسلامی، این ریسک‌ها به‌صورت خاص‌تری در دو دسته «ریسک‌های مرتبط با عقود مبادله‌ای» و «ریسک‌های مرتبط با عقود مشارکتی» تحلیل می‌شوند.در این میان برخی ریسک‌ها مانند ریسک اعتباری و ریسک عملیاتی، مشترک بین بانکداری متعارف و اسلامی هستند، اما به‌دلیل ماهیت عقود اسلامی، مدیریت این ریسک‌ها در بانکداری بدون ربا پیچیدگی بیشتری دارد .{ابوالحسنی, 2008 #13} .

با توجه به این نکته که ماهیت فعالیت‌های بانکی همراه با پذیرش ریسک است و کاهش آن می‌تواند کارایی و ثبات مالی بانک‌ها را بهبود بخشد . عدم مدیریت صحیح ریسک ممکن است منجر به کاهش درآمد، از دست دادن سرمایه و ایجاد محدودیت در دستیابی به اهداف اقتصادی بانک‌ها شود .{پندار, 2020 #3}

2-2 مبانی نظری

متن موردنظر

2-2-1 مفهوم ریسک و مدیریت آن

ریسک را میتوان به نحو های گوناگون تعربف کرد در این میان میتوان ریسک را به‌عنوان احتمال وقوع پیامدهای ناخواسته و نامطلوب که ممکن است بر افراد، اموال یا محیط زیست تأثیر منفی بگذارد، تعریف کرد. ریسک به‌طور معمول به‌عنوان حاصل‌ضرب احتمال وقوع یک رخداد منفی و شدت پیامدهای آن تعریف می‌شود. این تعریف، ریسک را مفهومی چندبعدی می‌داند که نه‌تنها بر احتمال وقوع رویدادهای نامطلوب تأکید دارد، بلکه میزان زیان حاصل از آن‌ها را نیز در نظر می‌گیرد. ریسک معمولاً به‌صورت کمی از طریق حاصل‌ضرب احتمال وقوع یک رویداد و شدت پیامدهای آن ارزیابی می‌شود. به این ترتیب، مفهوم ریسک نه‌تنها شامل شناسایی پیامدهای بالقوه است، بلکه میزان و شدت این پیامدها را نیز در بر می‌گیرد.{Gratt, 1989 #15}

با توجه به اهمیت ریسک مفهوم مدیریت مطرح میشود . مدیریت ریسک به‌عنوان یکی از مؤلفه‌های اساسی تصمیم‌گیری‌های مالی شرکتی مطرح شده که هدف آن تنظیم و کنترل نوسانات ذخایر نقدی و ایجاد تعادل میان هزینه‌های سرمایه‌گذاری و پوشش‌های مالی است. مدیریت ریسک نه‌تنها شامل کاهش نوسانات ذخایر نقدی بلکه شامل تدوین سیاست‌های مالی پویا برای شرکت‌ها هستند. مدیریت ریسک نقش مهمی در تشویق شرکت‌ها به اتخاذ فناوری‌های جدید و سرمایه‌گذاری در پروژه‌های نوآورانه ایفا می‌کند. بااین‌حال، این فرآیند با چالش‌هایی همراه است، ازجمله افزایش هزینه‌های ناشی از نوسانات بالاتر ذخایر نقدی و ریسک‌های مرتبط با تأمین مالی خارجی پرهزینه. {Luo, 2024 #16}

2-2-2 ریسک اعتباری

در حوزه بانکداری و امور مالی، ریسک اعتباری یکی از مهم‌ترین انواع ریسک‌هایی است که مؤسسات مالی با آن مواجه می‌شوند. ریسک اعتباری به احتمال عدم بازپرداخت اصل و سود وام از سوی وام‌گیرنده اشاره دارد که می‌تواند منجر به زیان‌های مالی و اختلال در جریان‌های نقدی بانک‌ها شود .{Kamara, 2024 #17}

به‌طور کلی، این ریسک در انواع مختلفی از وام‌ها و اعتبارها وجود دارد، از جمله وام‌های مصرفی، وام‌های تجاری، و حتی اوراق قرضه منتشرشده توسط دولت‌ها و شرکت‌ها. یکی از عوامل کلیدی در مدیریت ریسک اعتباری، ارزیابی اعتبار وام‌گیرنده بر اساس شاخص‌هایی مانند سابقه اعتباری، ظرفیت بازپرداخت، سرمایه، شرایط وام و وثیقه مرتبط است.{Graham, 2000 #18}

مقاله مذکور تأکید دارد که در دهه گذشته، بانک‌ها به سمت استفاده از مدل‌های پیچیده ارزیابی ریسک اعتباری رفته‌اند که اغلب شامل روش‌های مبتنی بر الگوریتم و داده‌کاوی برای تحلیل ریسک‌های مرتبط هستند. این مدل‌ها به بانک‌ها کمک می‌کنند تا ریسک اعتباری را به‌طور دقیق‌تر کمی‌سازی، تجمیع و مدیریت کنند و به‌ویژه در زمینه شناسایی و کاهش وام‌های غیرجاری مؤثر باشند . در این راستا، بانک‌ها باید به‌صورت منظم داده‌های مرتبط با ریسک اعتباری را تحلیل کرده و از ابزارهای پیشرفته‌ای برای پیش‌بینی اثرات احتمالی ریسک استفاده کنند. {Kamara, 2024 #17}

2-2-3 مدل‌های ارزیابی ریسک اعتباری

متن موردنظر

2-2-4 ماتریس زیان در ارزیابی ریسک

در روش های گذشته ، تمام سوابق اعتباری تحت شرایط اطلاعاتی یکسان و با استفاده از یک مدل واحد پیش‌بینی می‌شوند و به هر نمونه یکی از دو برچسب "بله" (قبول) یا "خیر" (رد) اختصاص داده می‌شود. این فرآیند به عنوان تصمیم‌گیری دودویی شناخته می‌شود. با این حال، همیشه برخی از نمونه‌ها وجود دارند که با اطلاعات موجود نمی‌توان به‌خوبی پیش‌بینی کرد. روش تصمیم‌گیری سه‌گانه یک مکانیزم تأخیر در تصمیم‌گیری به تصمیم‌گیری دودویی سنتی اضافه می‌کند. به این معنا که برای برخی از نمونه‌ها نیازی نیست بلافاصله تصمیم نهایی (قبول یا رد) گرفته شود؛ بلکه پس از دریافت اطلاعات بیشتر، تصمیم‌گیری نهایی انجام می‌شود. این رویکرد ریسک تصمیم‌گیری اشتباه در مورد نمونه‌های دشوار را کاهش می‌دهد. تصمیم‌های مختلف روی یک نمونه می‌تواند منجر به زیان‌های متفاوتی شود. این مقادیر زیان، مبنای تعیین آستانه‌ها هستند.

با این حال، مقادیر زیان (Loss Values) به‌سادگی قابل اندازه‌گیری نیستند. به‌عنوان مثال، تعیین میزان زیان در صورتی که یک نمونه به منطقه مرزی (BND) اختصاص یابد، کار دشواری است. یکی از روش های موجود آن است که مقادیر زیان را به‌صورت تصادفی تعیین گردند تا آستانه‌ها را محاسبه کنند.{Xu, 2023 #21}

از جمله روش های دیگر میتوان به روش هایی اشاره کرد که در آن‌ها به‌جای مقادیر زیان، پارامترهای جایگزین میشوند می‌شوند.یک از روش های موجود استفاده از نظریه پشیمانی (Regret Theory) میباشد و مقادیر پشیمانی، مقادیر خشنودی و میزان ادراک روانی را برای تعیین آستانه‌های تصمیم‌گیری سه‌گانه به کار بردند.{Zhu, 2022 #22}

2-2-5 روش بگینگ در یادگیری ماشین

ایده اصلی یادگیری ترکیبی [[1]](#footnote-1) این است که چند الگوریتم ضعیف [[2]](#footnote-2) را با هم ترکیب کرده و یک مدل قوی [[3]](#footnote-3) با عملکرد بهتر ایجاد کنیم.

یکی از رویکردهای مؤثر در افزایش قابلیت تعمیم‌پذیری مدل‌های یادگیری ماشین و کاهش ریسک‌های ناشی از بیش‌برازش [[4]](#footnote-4)، استفاده از الگوریتم بگینگ [[5]](#footnote-5) است. این روش که توسط 1996 معرفی شده است، با ایجاد زیرمجموعه‌های تصادفی از داده‌های آموزشی و آموزش مدل‌های پایه [[6]](#footnote-6) روی این زیرمجموعه‌ها، تلاش می‌کند واریانس مدل را کاهش داده و پایداری پیش‌بینی‌ها را بهبود بخشد. روش بگینگ می‌تواند در مسائل مرتبط با داده‌های نامتوازن، به‌ویژه در محیط‌های پرریسک مانند مالی و بانکی، به بهبود عملکرد طبقه‌بندی کمک کند. این الگوریتم با ترکیب پیش‌بینی‌های مدل‌های پایه از طریق رأی‌گیری اکثریت [[7]](#footnote-7)، باعث افزایش دقت مدل‌ها و کاهش خطاهای ناشی از تصمیم‌گیری‌های اشتباه می‌شود .{Breiman, 1996 #20}

در مسائل مرتبط با ارزیابی ریسک اعتباری، یکی از چالش‌های اساسی، وجود عدم توازن طبقات [[8]](#footnote-8) است. بیشتر داده‌های موجود متعلق به مشتریان خوش‌حساب (طبقه اکثریت) هستند و تعداد کمی از داده‌ها مربوط به مشتریان پرریسک (طبقه اقلیت) است. در چنین شرایطی، روش‌های یادگیری سنتی معمولاً تمایل به طبقه‌بندی داده‌ها در طبقه اکثریت دارند و از شناسایی دقیق طبقه اقلیت بازمی‌مانند. روش بگینگ با ایجاد زیرمجموعه‌های متوازن از داده‌ها می‌تواند به افزایش دقت در شناسایی طبقه اقلیت کمک کند.

استفاده از نسخه‌های تغییریافته از روش بگینگ، به‌ویژه در محیط‌های داده‌ای ناپایدار و نامتوازن، می‌تواند ریسک‌های مالی را کاهش دهد و به بهبود مدیریت ریسک اعتباری منجر شود. این نسخه تغییریافته از بگینگ به‌جای نمونه‌برداری تصادفی کامل، از نمونه‌برداری زیرمجموعه‌های متوازن [[9]](#footnote-9) برای ایجاد داده‌های آموزشی متوازن استفاده می‌کند. در این روش، داده‌های مربوط به مشتریان پرریسک به‌گونه‌ای انتخاب می‌شوند که مدل بتواند الگوهای رفتاری آن‌ها را بهتر شناسایی کند. نتایج این پژوهش نشان می‌دهد که کاهش واریانس مدل‌ها و ترکیب پیش‌بینی‌های مدل‌های پایه به افزایش دقت پیش‌بینی‌ها کمک کرده و می‌تواند در تصمیم‌گیری‌های مالی مؤسسات تأثیرگذار باشد .{Abdoli, 2023 #19}

2-2-6 ترکیب روش‌های طبقه‌بندی

متن موردنظر

2-3 پیشینه تحقیق

2-3-1 تحقیقات داخلی

متن موردنظر

2-3-2 تحقیقات خارجی

در حوزه امتیازدهی اعتباری، همواره انتخاب الگوریتم‌های مناسب برای بهبود دقت پیش‌بینی و کاهش ریسک‌های مالی، یک چالش اساسی بوده است. پیش از این، مدل‌های طبقه‌بندی پایه به‌تنهایی قادر به ارائه دقت مطلوب در پیش‌بینی امتیاز اعتباری نبودند و نیاز به روشی مؤثرتر احساس می‌شد. در این راستا، پژوهشی با استفاده از مجموعه داده‌های اعتباری استرالیا، پس از پیش‌پردازش داده‌ها و جایگزینی مقادیر گمشده، اقدام به کاهش ابعاد داده‌ها با تحلیل مؤلفه‌های اصلی نمود. سپس، الگوریتم‌های طبقه‌بندی پایه مانند رگرسیون لجستیک، درخت تصمیم، ماشین بردار پشتیبان، نزدیک‌ترین همسایه، شبکه عصبی چندلایه و بیز ساده به‌کار گرفته شدند. علاوه بر آن، الگوریتم‌های گروهی نظیر جنگل تصادفی، درخت تصمیم بسته‌بندی‌شده، طبقه‌بندی هایی مبنتی بر روش های تقویتی، و رأی‌گیری گروهی مورد ارزیابی قرار گرفتند. نتایج نشان داد که الگوریتم‌های گروهی، به‌ویژه جنگل تصادفی و طبقه‌بندی هایی مبنتی بر روش های تقویتی ، با دقت بالاتر و خطای کمتر، عملکرد بهتری در پیش‌بینی امتیاز اعتباری دارند. {Parvin, 2020 #1}

مطالعات پیشین بر روی ارزیابی اعتبار مشتریان نشان می‌دهد که مدل‌های رایج همچون رگرسیون لجستیک و تحلیل تمایز خطی به دلیل فرضیات غیرواقعی مانند تفکیک خطی و توزیع نرمال داده‌ها عملکرد ضعیفی داشتند. به کارگیری روش‌هایی مانند شبکه عصبی و ماشین بردار پشتیبان در تلاش بود تا دقت مدل‌های ارزیابی اعتبار را بهبود بخشد، اما این مدل‌های تک‌طبقه‌بند در شناسایی ویژگی‌های خاص مشتریان و کاهش تأثیر نمونه‌های نادر به مشکل برخوردند. برای رفع این محدودیت‌ها، برخی تحقیقات از روش‌های ترکیبی شامل طبقه‌بندی هایی با رای‌گیری های وزن‌دار استفاده کردند، اما این روش‌ها همچنان در مدیریت داده‌های نامتوازن با مشکل مواجه بودند. در این پژوهش، از نقشه‌های خودسازمان‌ده برای خوشه‌بندی چندلایه داده‌ها بهره گرفته شده و سپس از ترکیب طبقه‌بندهایی چون کا-نزدیک‌ترین همسایگی و ماشین بردار پشتیبان با رای‌گیری نرم استفاده شده است. به منظور افزایش دقت و کاهش تأثیر داده‌های نامتوازن، بهینه‌سازی وزن‌ها با الگوریتم بهینه‌سازی سیلفیش[[10]](#footnote-10) انجام شده که الگوریتمی متاهیوریستیک با سرعت همگرایی بالا و قابلیت جستجوی فراگیر است. این ترکیب از خوشه‌بندی خودسازمان‌ده و الگوریتم بهینه‌سازی سیلفیش نه تنها به طبقه‌بندی دقیق‌تر و شناسایی مؤثر نمونه‌های نادر کمک می‌کند، بلکه پردازش کارآمدتری نیز در مقایسه با روش‌های پیشین دارد. {Singh, 2024 #2}

در تحقیقات گذشته، مسئله‌ مهمی که در مدل‌سازی ریسک اعتباری با روش خوشه‌بندی وجود داشت، انتخاب متغیرهای مناسب برای ایجاد خوشه‌های دقیق بود، زیرا انتخاب نادرست متغیرها می‌توانست منجر به خوشه‌های غیرقابل‌تشخیص و ناموثر در پیش‌بینی شود. برخی از پژوهش‌ها با استفاده از تحلیل مؤلفه‌های اصلی به کاهش بعد متغیرها پرداختند اما این روش پیچیدگی خوشه‌بندی را افزایش می‌داد و قابلیت تفسیر مدل را کاهش می‌داد. روش‌های دیگر نظیر حذف تدریجی متغیرها با استفاده از رگرسیون خطی نیز به کار گرفته شده‌اند، اما این روش‌ها به دلیل انتخاب‌های تصادفی، دقت خوشه‌بندی و کیفیت پیش‌بینی را تحت تأثیر قرار می‌دادند. مقاله حاضر با معرفی یک روش مبتنی بر یادگیری ماشین، از مقادیر شاپلی[[11]](#footnote-11) برای سنجش اهمیت متغیرها استفاده می‌کند و با انتخاب متغیرهای اصلی در مرحله خوشه‌بندی، سعی دارد دقت پیش‌بینی مدل را افزایش دهد. این رویکرد، از الگوریتم‌های جنگل تصادفی و درخت‌های تقویت‌شده بهره برده و با استفاده از الگوریتم کاپروتوتایپ‌ها [[12]](#footnote-12)و تکنیک‌های اعتبارسنجی، خوشه‌بندی را طوری انجام می‌دهد که نتایج هم از نظر دقت پیش‌بینی و هم از نظر تفسیرپذیری مورد تایید قرار گیرد.{Zöllner, 2023 #3}

مطالعات پیشین در زمینه ارزیابی ریسک اعتباری عمدتاً بر استفاده از روش‌های خوشه‌بندی و طبقه‌بندی داده‌ها تمرکز داشته‌اند. روش‌های رایج مانند کا میانگین[[13]](#footnote-13) و سلسله‌مراتبی[[14]](#footnote-14)، علی‌رغم سادگی و تفسیرپذیری، در داده‌های چندبعدی و غیرخطی به محدودیت‌هایی از جمله کاهش دقت در خوشه‌بندی و ناکارآمدی در جداسازی داده‌های نویزی مواجه شده . در این مقاله، یک رویکرد ترکیبی ارائه شده است که خوشه‌بندی کا میانگین و سلسله‌مراتبی را با استفاده از مختصات فاکتوریل استخراج‌شده از تحلیل مکاتبات چندگانه[[15]](#footnote-15) تلفیق می‌کند. این روش، ابتدا با پیش‌پردازش داده‌ها شامل نرمال‌سازی و کاهش نویز داده‌ها ، ساختار داده ها را برای تحلیل آماده می‌سازد. سپس، تحلیل مکاتبات چندگانه به منظور کاهش ابعاد و استخراج ویژگی‌های اساسی داده‌ها به کار گرفته می‌شود. در ادامه، خوشه‌بندی کا میانگین و سلسله‌مراتبی به صورت ترکیبی اعمال شده و کیفیت خوشه‌بندی بر اساس تجزیه اینرسی ارزیابی می‌شود. این روش امکان مدیریت دقیق‌تر داده‌ها، کاهش نویز و بهبود دقت مدل‌های یادگیری ماشین در تحلیل ریسک اعتباری را فراهم کرده است. {Jadwal, 2022 #5}

در زمینه ارزیابی ریسک اعتباری، انتخاب ویژگی به عنوان یکی از چالش‌های اصلی در بهبود دقت پیش‌بینی و کاهش پیچیدگی مدل‌ها شناخته شده است. پیش از این، روش‌های غربالگری مانند رتبه‌بندی همبستگی[[16]](#footnote-16) یا اطلاعات متقابل[[17]](#footnote-17) به دلیل سوگیری در انتخاب ویژگی‌ها و کاهش دقت پیش‌بینی، محدودیت‌هایی داشتند. همچنین، روش‌های ترکیبی دو‌مرحله‌ای مانند استفاده از فیلترهای اولیه و سپس الگوریتم‌های جستجوی تکاملی، اگرچه در برخی موارد به بهبود دقت کمک می‌کردند، اما باعث افزایش پیچیدگی محاسبات و کاهش پایداری مدل‌ها می‌شدند. در این مقاله، رویکردهای یک‌مرحله‌ای مانند حذف بازگشتی ویژگی [[18]](#footnote-18) با مدل‌های جنگل تصادفی [[19]](#footnote-19) و ماشین بردار پشتیبان [[20]](#footnote-20) بررسی شده‌اند. این روش‌ها ویژگی‌ها را به طور پویا در فرآیند یادگیری انتخاب کرده و امکان پایداری بیشتر انتخاب ویژگی را فراهم می‌کنند. همچنین، استفاده از رگرسیون لجستیک جریمه‌ای [[21]](#footnote-21) به منظور کاهش ابعاد داده‌ها و کنترل همبستگی متغیرها بررسی شده است. روش حذف بازگشتی ویژگی با جنگل تصادفی در معیارهایی مانند دقت پیش‌بینی و پایداری انتخاب ویژگی عملکرد بهتری نسبت به روش‌های دیگر داشته و الگوریتم حذف بازگشتی ویژگی با ماشین بردار پشتیبان با پایداری بالاتر، گزینه‌ای مناسب برای مجموعه داده‌های با ابعاد بزرگ است. {Atif, 2022 #6}

در زمینه ارزیابی ریسک اعتباری رویکردهای اولیه معمولاً از مدل‌های پایه استفاده می‌کردند که به دلیل نبود ترکیب مناسب الگوریتم‌ها و عدم استفاده از تکنیک‌های پیش‌پردازش، دقت پایین‌تری داشتند. برخی تحقیقات نیز از روش‌های ترکیبی مانند ادغام یادگیری نظارت‌شده و بدون نظارت یا استفاده از شبکه‌های عصبی مصنوعی بهره بردند، اما این روش‌ها به دلیل نیاز به محاسبات با حجم بالا یا عدم مدیریت مناسب ویژگی‌های غیرمؤثر، در محیط‌های واقعی کارایی مطلوبی نداشتند. استفاده از روش‌های انتخاب ویژگی بر اساس همبستگی یا مدل‌های خطی نیز در برخی مطالعات مطرح شده، اما این روش‌ها قادر به شناسایی روابط غیرخطی بین ویژگی‌ها نبودند. در این میان، مطالعاتی که از روش‌های ترکیبی چندلایه استفاده کردند، توانستند با ادغام خروجی مدل‌های مختلف، عملکرد بهتری ارائه دهند. این روش‌ها به‌ویژه در مدیریت عدم توازن داده‌ها و بهبود معیارهایی نظیر دقت و امتیاز F1 نقش مؤثری ایفا کردند. با این حال، چالش انتخاب ویژگی‌های بهینه و کاهش پیچیدگی مدل همچنان به قوت خود باقی بود. پژوهش حاضر با ارائه رویکردی چندلایه و ترکیب آن با انتخاب ویژگی مبتنی بر اطلاعات متقابل، توانست با بهینه‌سازی ویژگی‌های ورودی و کاهش پیچیدگی محاسباتی، نتایج قابل توجهی در پیش‌بینی ریسک اعتباری ارائه دهد. {Emmanuel, 2024 #7}

روش‌های سنتی ارزیابی ریسک اعتباری مانند رگرسیون لجستیک و ماشین بردار پشتیبان با وجود دقت قابل‌قبول، در مدیریت نمونه‌های پیچیده و اطلاعات ناقص ضعف دارند و معمولاً با تصمیم‌گیری دوکلاسه، ریسک خطا را افزایش می‌دهند. رویکردهای ترکیبی مانند تقویت گرادیانی پیشرفته [[22]](#footnote-22) با وجود بهبود نسبی، به دلیل استفاده از آستانه‌های ثابت و ناتوانی در مدیریت داده‌های با هزینه بالا، محدودیت‌هایی داشته‌اند. تصمیم‌گیری سه‌ کلاسه با افزودن تأخیر در تصمیم‌گیری، به کاهش خطا کمک کرده است، اما تنظیم آستانه‌ها در روش‌های پیشین مبتنی بر فرضیات کلی بوده و داده‌های مالی واقعی را نادیده گرفته‌اند. مقاله حاضر با ترکیب تصمیم‌گیری سه کلاسه و استفاده از داده‌های جریان نقدی برای تنظیم آستانه‌ها و بهره‌گیری از مدل بهینه‌سازی چندهدفه، این محدودیت‌ها را برطرف کرده و با ترکیب الگوریتم ماشین تقویت گرادیانی سبک [[23]](#footnote-23) امکان تصمیم‌گیری دقیق‌تر برای نمونه‌های مرزی و پرهزینه را فراهم ساخته و کارایی مدیریت ریسک اعتباری را افزایش داده است. {Li, 2024 #8}

پیش‌بینی ریسک اعتباری عمدتاً بر روش‌های یادگیری ماشین متکی بوده‌اند که شامل تکنیک‌هایی مانند شبکه‌های عصبی، ماشین بردار پشتیبان و رگرسیون لجستیک می‌شوند. این مدل‌ها با وجود توانایی بالا در تحلیل داده‌ها، به دلیل اتخاذ تصمیم‌گیری دو کلاسه، دچار خطاهایی در دسته‌بندی داده‌های مرزی شده‌اند. برای حل این مشکل، رویکردهای سه‌کلاسه معرفی شده‌اند که با تقسیم داده‌ها به سه منطقه مثبت، منفی و مرزی، تصمیم‌گیری برای داده‌های نامطمئن را به تأخیر می‌اندازند. روش‌های مختلفی برای یادگیری آستانه در تصمیم‌گیری سه‌کلاسه استفاده شده است، از جمله استفاده از بهینه‌سازی چندهدفه و مدل‌های مبتنی بر کاهش هزینه‌های تصمیم‌گیری. با این حال، اکثر این رویکردها نیازمند تنظیم دستی پارامترها بوده‌اند که قابلیت کاربرد آن‌ها را محدود کرده است. مقاله حاضر با بهره‌گیری از یک چارچوب دو مرحله‌ای و استفاده از الگوریتم بهینه‌سازی ازدحام ذرات برای یادگیری خودکار آستانه‌ها، رویکردی جامع‌تر ارائه کرده است. این روش با ترکیب بهینه اطلاعات و کاهش هزینه تصمیم‌گیری، ضمن افزایش دقت در دسته‌بندی داده‌های مرزی، امکان استفاده از داده‌های تکمیلی در مراحل بعدی را فراهم می‌کند و محدودیت‌های رویکردهای پیشین را برطرف می‌سازد. {Li, 2024 #9}

در ارزیابی ریسک اعتباری، روش‌های سنتی مانند رگرسیون لجستیک و تحلیل تمایز خطی به دلیل سادگی، محبوبیت بالایی داشتند، اما عدم توانایی در مدیریت داده‌های پیچیده و نامطمئن از معایب اصلی آنها بود. مدل‌های پیشرفته‌تر مانند شبکه‌های عصبی مصنوعی و درخت‌های تصمیم‌گیری تقویتی با قدرت تشخیص بالاتر، توانستند این محدودیت را کاهش دهند، اما بیشتر این مدل‌ها تنها از طبقه‌بندی دوتایی استفاده کرده و در مواجهه با اطلاعات ناکافی یا مبهم دچار ضعف می‌شوند. مدل سه‌ کلاسه تصمیم‌گیری با ارائه دسته مرزی برای موارد نامطمئن گامی در مدیریت بهتر این چالش‌ها برداشت، اما این روش نیز به دلیل وابستگی زیاد به تنظیمات آن انعطاف‌پذیری کمی داشت. مقاله حاضر با ترکیب مدل سه‌کلاسه تصمیم‌گیری و روش‌های یادگیری ماشین مانند تقویت‌کننده گرادیان، و بهینه‌سازی با الگوریتم ژنتیک، یک چارچوب بهینه‌سازی ترتیبی ارائه داده است که بدون نیاز به تنظیمات دستی پارامترها آستانه‌های تصمیم‌گیری را بهینه می‌کند و با افزودن تدریجی اطلاعات به موارد مرزی، عدم قطعیت را کاهش داده و عملکرد کلی سیستم را بهبود می‌بخشد. {Shen, 2022 #10}

تصمیم‌گیری سه‌کلاسه ، بر اساس نظریه مجموعه‌های راف احتمالی، به‌عنوان راهکاری مؤثر برای مدیریت عدم قطعیت معرفی شده است. این رویکرد، به‌جای تصمیم‌گیری قطعی در مدل‌های دو‌کلاسه ، با افزودن یک منطقه مرزی، امکان تصمیم‌گیری منعطف‌تر را فراهم می‌کند. پیش از این، مدل‌های دو‌کلاسه مانند رگرسیون لجستیک یا درخت تصمیم برای ارزیابی اعتباری استفاده می‌شدند که مستلزم جمع‌آوری اطلاعات جامع از تمامی متقاضیان و افزایش منابع پردازشی بود. مقاله حاضر با استفاده از مجموعه‌های راف احتمالی، یک فرآیند دو مرحله‌ای را پیشنهاد می‌کند. در مرحله اول، با اطلاعات اولیه مانند فرم‌های درخواست، مشتریان به سه دسته تأیید، رد و مرزی تقسیم می‌شوند. مرحله دوم، که مختص موارد مرزی است، شامل جمع‌آوری اطلاعات تکمیلی از طریق مصاحبه و تحلیل مالی است. تحلیل دقیق در این مرحله با استفاده از رگرسیون لجستیک انجام می‌شود که احتمال بازپرداخت را بر اساس داده‌های تکمیلی پیش‌بینی می‌کند. این فرآیند با مدیریت بهینه منابع ، از جمع‌آوری اطلاعات غیرضروری جلوگیری کرده و دقت و کارایی تصمیم‌گیری را بهبود می‌بخشد. {Maldonado, 2020 #11}

2-4 جمع‌بندی

متن موردنظر

فصل سوم:

روش تحقیق

(مراحل انجام پژوهش)

بیان مقدمه‌ای کوتاه برای آغاز بحث [از درج تیتر با عنوان «مقدمه» در ابتدای هر فصل خودداری نمایید.].

3-1 عنوان موردنظر

متن موردنظر

فصل چهارم:

شبیه‌سازی و تحلیل نتایج

(نتایج و بحث)

بیان مقدمه‌ای کوتاه برای آغاز بحث [از درج تیتر با عنوان «مقدمه» در ابتدای هر فصل خودداری نمایید.]

4-1 عنوان موردنظر

متن موردنظر

فصل پنجم:

نتیجه‌گیری و پیشنهادها

بیان مقدمه‌ای کوتاه برای آغاز بحث [از درج تیتر با عنوان «مقدمه» در ابتدای هر فصل خودداری نمایید.]

5-1 نتیجه‌گیری

متن موردنظر

5-2 پيشنهادها

متن موردنظر

پیوست الف (در صورت وجود)

متن پیوست الف

فهرست

منابع و مأخذ

فهرست منابع و مأخذ

[1] منبع فارسی

[2] English Reference

Abstract

Abstract text.

Keywords: Keywords text.

****

ISLAMIC AZAD UNIVERSITY

North Tehran Branch

Faculty of ………….... – Department of ……………

"M.Sc" Thesis

On ………………………..

Research Title:

………………………

Supervisor:

………………………

Consulting Supervisor:

……………………..

By:

…………………

Summer 2023

1. Ensemble Learning [↑](#footnote-ref-1)
2. Weak Learners [↑](#footnote-ref-2)
3. Strong Learner [↑](#footnote-ref-3)
4. Overfitting [↑](#footnote-ref-4)
5. Bagging [↑](#footnote-ref-5)
6. Base Learners [↑](#footnote-ref-6)
7. Majority Voting [↑](#footnote-ref-7)
8. Class Imbalance [↑](#footnote-ref-8)
9. Subsampling [↑](#footnote-ref-9)
10. Sailfish Optimizer [↑](#footnote-ref-10)
11. Shapley Values [↑](#footnote-ref-11)
12. k-prototypes [↑](#footnote-ref-12)
13. k-means [↑](#footnote-ref-13)
14. Hierarchical clustering [↑](#footnote-ref-14)
15. Multiple Correspondence Analysis [↑](#footnote-ref-15)
16. Neighborhood Ranking [↑](#footnote-ref-16)
17. Mutual Information [↑](#footnote-ref-17)
18. Recursive Feature Elimination [↑](#footnote-ref-18)
19. RFE-RF [↑](#footnote-ref-19)
20. RFE-SVM [↑](#footnote-ref-20)
21. Penalized Logistic Regression [↑](#footnote-ref-21)
22. XGBoost [↑](#footnote-ref-22)
23. LightGBM [↑](#footnote-ref-23)