Downloaded from jmbr.mbri.ac.ir at 10:01 +0330 on Monday November 25th 2019

فصلنامه پژوهش‌های پولی_ بانکی
سال دوازدهم، شماره ۳۹، بهار ۱۳۹۸
صفحات ۱۲۹_۱۶۶

دسته بندی ریسک اعتباری مشتریان حقیقی با استفاده از یادگیری جمعی (مطالعه موردی بانک سپه)

مهدی نظرآقایی [*] **حسین غیاثی** [†]
محمد اصغر خواه چافی [‡]

تاریخ دریافت: ۱۳۹۸/۰۴/۰۱ تاریخ پذیرش: ۱۳۹۸/۰۷/۲۳

چکیده

فعالیت‌های بانکی و اعطای تسهیلات در بانک‌ها همواره با ریسک اعتباری مواجه بوده و با توجه به محدودیت منابع مالی بانک‌ها جهت ارائه تسهیلات، ارزیابی توان بازپرداخت مشتریان بانک پیش از اعطای تسهیلات، یکی از مهمترین چالش‌های پیش روی سیستم بانکی کشور است. بر این اساس در این تحقیق سعی بر آن شد که مدلی در جهت تعیین عوامل مهم موثر بر رفتار اعتباری مشتریان بانکی ارائه شود. تحقیق حاضر با استفاده از اطلاعات مشتریان حقیقی بانک سپه در سال های ۱۳۹۵ و ۱۳۹۶ صورت گرفته است و مدل_سازی تحقیق با استفاده از شبکه عصبی، درخت تصمیم فازی انجام شده است. آنچه نوآوری تحقیق می‌تواند محسوب شود استفاده از روش‌های یادگیری جمعی است که به منظور افزایش دقت در نتایج درخت تصمیم فازی در این تحقیق مدنظر قرارگرفته است. نتایج حاصل از تحقیق نشان می دهد که درآمد و تراکنش‌های مالی مشتریان از بیشترین اهمیت در تعیین ریسک اعتباری مشتریان برخوردار بوده است. هم‌چنین نتایج نشان می‌دهد که درخت تصمیم فازی با استفاده روش بگینگ دقت بالاتری نسبت به روش شبکه عصبی و درخت تصمیم فازی معمولی دارد.

واژه‌های کلیدی: ریسک اعتباری، شبکه عصبی، درخت تصمیم فازی، یادگیری جمعی.
طبقه‌بندی JEL: G32,C45.H81

[*] کارشناسی ارشد اقتصاد دانشگاه خوارزمی؛ mehdi.nazaraghaei@gmail.com
[†] کارشناسی ارشد مدیریت مالی دانشگاه خوارزمی(نویسنده مسئول)؛ Hossein.gh1413@gmail.com
[‡] کارشناسی ارشد مدیریت مالی دانشگاه خوارزمی ؛ mohammad.reinfall@yahoo.com

Downloaded from jmbr.mbri.ac.ir at 10:01 +0330 on Monday November 25th 2019

۱ مقدمه

بررسی عملکرد بیشتر کشورهای جهان نشان می‌دهد که سرمایه‌گذاری و پیشرفت اقتصادی رابطه نزدیکی با یکدیگر دارند. یعنی کشورهایی که الگوی کارآمدی در تخصیص سرمایه به بخش‌های مختلف اقتصادی دارند، اغلب از پیشرفت اقتصادی و در نتیجه رفاه اجتماعی بالاتری برخوردار هستند. تجهیز و تخصیص منابع جهت سرمایه گذاری در فعالیت‌های اقتصادی از طریق بازار مالی انجام می‌پذیرد که بازار اعتبارات بانکی، قسمتی از این بازار است. انجام این امر به عنوان اصلی‌ترین نقش بانک در بازار مالی از طریق اعطای اعتبار به مشتریان صورت می‌گیرد. بنابراین یکی از موضوعات دارای اهمیت در رابطه با این موضوع بررسی و ارزیابی ریسک اعتباری مرتبط با تسهیلات گیرندگان بانک ها است. اندازه‌گیری این ریسک در میان ریسک‌هایی که بانک در حیطه وسیع عملکرد خود با آن روبه‌روست، از جایگاه ویژه‌ای برخوردار است. کاهش و کنترل ریسک به عنوان یکی از عوامل مهم مؤثر بر بهبود فرایند اعطای اعتبار و در نتیجه بر عملکرد بانک‌ها مطرح است و نقش اساسی در تداوم ارائه تسهیلات و بقای بانک‌ها و مؤسسات مالی دارد. آنچه برای بانک اهمیت دارد این است که قبل از اعطای تسهیلات به متقاضیان، احتمال عدم بازپرداخت از سوی آنان را ارزیابی و گروهی را که از ادعای دین آن‌ها در موعد مقرر اطمینان بیشتری دارد، انتخاب نماید. انجام این امر به وسیله یک سیستم جامع، تعریف ساختار و معیارهای مناسب امکان پذیر است. امروزه بانک‌ها به شکل وسیعی از مدل‌های سنجش ریسک اعتباری برای تصویب و پرداخت وام‌های اعطایی خود استفاده می کنند و با استفاده از معیارهای عینی و اطلاعات حال و گذشته مشتری، در قالب تهیه انواع گزارشات اطلاعاتی و کارشناسی و اتخاذ تصمیم در ارکان اعتباری ذی‌صلاح، به اعتبار سنجی مشتریان می‌پردازند. در مورد وام‌های بزرگ و با توجه به تعداد اندک آن‌ها، ارزیابی دقیق متقاضی امکان‌پذیر است؛ بنابر این در مورد وام‌های متوسط و کوچک، چون تعداد متقاضیان زیاد است، ارزیابی دقیق تک تک آن‌ها پرهزینه است و از این رو نیازمند ارزیابی سیستماتیک و ایجاد مدلی است که بر اساس آن بتوان ریسک اعتباری را تعیین و کاهش داد (مهرآرا و مهرانفر، ۱۳۹۲).

علی رغم اهمیت این موضوع، در کشور ما در زمینه اعطای تسهیلات اعتباری به مشتریان، روند منسجم و منظمی به منظور تعیین ریسک اعتباری، امتیازدهی، درجه‌بندی و همچنین تعیین سقف‌های اعتباری بر اساس شاخص‌های ریسک ملاحظه نمی‌شود و شاخص‌ها بر اساس تشخیص کارشناسی و کمیته اعتباری صورت می‌پذیرد. برخورداری از یک مدل ریسک کارآمد نه‌تنها تصمیم‌گیری در زمینه اعتبار و گرفتن وثیقه را تسهیل می‌نماید،

بلکه افزون بر کاهش هزینه مبادله موجب خواهد شد که سیستم بانکی از الگوی کارآمدی در تخصیص سرمایه به بخش‌های مختلف اقتصادی برخوردار شود. هدف علمی تحقیق مشارکت در پیشبرد ادبیات تجربی موجود در زمینه ریسک اعتباری و اهداف کاربردی تحقیق: شناسایی ویژگی های اثرگذار در سیستم بانکداری مشتریان با ریسک اعتباری آنان. و ارائه چارچوبی برای دسته بندی مشتریان با استفاده از ریسک اعتباری است. فرضیه تحقیق نیز شامل این فرض است که روش بگینگ نسبت به سایر الگوها دقت بالاتری را دارا است و تراکنش های مالی مشتریان بانکی بیشترین توضیح دهندگی را در رتبه بندی ریسک اعتباری مشتریان دارد.

روش گردآوری اطلاعات در این مقاله به دو بخش بخش اصلی (بررسی ادبیات موضوعی و پیشینه تحقیق)، روش گردآوری اطلاعات به‌صورت کتابخانه‌ای است و بخش انتهایی (مورد مطالعه) از روش‌های میدانی نیز بهره گرفته می‌شود. همچنین تجزیه و تحلیل اطلاعات در مقاله به دو صورت کمی و کیفی می‌باشد. در ادامه به معرفی مبانی و تجزیه و تحلیل ریسک های موجود در نظام بانکی و معرفی به روش بگینگ و داده های مدل و در نهایت به آزمون فرضیه ها و نتایج حاصل از آن پرداخته شده است.

۲ مبانی نظری

ریسک در هر حیطه‌ای قابلیت مطرح شدن دارد که یکی از مهمترین حیطه‌ها، بانک‌ها هستند که به علت اهمیت به سزایی که در نظام اقتصادی دارند؛ به طور خاص مورد توجه قرار می گیرند. زیرا بانک‌ها هر روزه در عملیات گوناگون خود با انواع مختلف ریسک روبه رو می‌شوند. لذا دلایل وجود ریسک در بانک‌ها را با نوع کارکرد آن ها به راحتی می‌توان توجیه کرد؛ چرا که بانک‌ها از یک سو سرمایه‌های مردم را که در قبال آن مسؤلیت دارند؛ جمع‌آوری کرده و از سوی دیگر با استفاده از این سرمایه‌ها اقدام به انجام عملیات بانکی و فعالیت‌های اقتصادی می‌نمایند.

Downloaded from jmbr.mbri.ac.ir at 10:01 +0330 on Monday November 25th 2019

جدول ۱
انواع ریسک در بانک‌ها

نام	توصیف
ریسک اعتباری[۱]	ریسک عدم بازگشت وام‌ها و قراردادهای مؤسسه
ریسک نقدینگی[۲]	ریسک نقدکردن مطالبات توسط مشتریان به‌صورت آنی و نیاز بانک به تبدیل فوری دارایی به پول نقد
ریسک نرخ بهره[۳]	ریسک کاهش ارزش دارایی‌ها به علت نوسانات نرخ بهره
ریسک بازار[۴]	ریسک کاهش ارزش دارایی‌ها و پرداخت‌ها به علت تغییر شرایطی مانند نرخ‌های برابری
ریسک خارج ترازنامه[۵]	ریسک مؤسسه ناشی از نتایج فعالیت‌های مربوط به دارایی‌ها یا پرداخت‌های اقتضایی (پرداخت‌هایی که بسته به موقعیت زمانی و تغییر شرایط زمانی باید انجام شوند)
ریسک نرخ‌های برابری[۶]	ریسک ناشی از تغییر ارزش دارایی‌ها یا بدهی‌های مؤسسه در خارج از کشور به علت تغییر نرخ ارز
ریسک دولت[۷]	ریسک عدم بازپرداخت مشتریان خارجی به علت دخالت دولت‌های خارجی
ریسک عدم کفایت سرمایه[۸]	ریسک ناشی از عدم وجود سرمایه مکفی برای جبران کاهش ارزش دارایی‌ها

منبع: (رادپور و همکاران، ۱۳۸۸)

از میان ریسک‌های مالی ارایه شده در جدول ۱ چهار ریسک بازار، ریسک اعتباری، ریسک نقدینگی و ریسک نرخ بهره از مهم‌ترین عوامل توجیه‌کننده نوسانات مالی هستند که برای آن‌ها سیستم‌های مدیریت و کنترل ریسک، توسعه بیش‌تری یافته‌اند (رادپور و همکاران، ۱۳۸۸).

[1] Credit Risk
[2] Liquidity Risk
[3] Interest Rate Risk
[4] Market Risk
[5] Off-Balance-Sheet Risk
[6] Foreign Exchange Risk
[7] Sovereign Risk
[8] Insolvency Risk

Downloaded from jmbr.mbri.ac.ir at 10:01 +0330 on Monday November 25th 2019

Downloaded from jmbr.mbri.ac.ir at 10:01 +0330 on Monday November 25th 2019

۱.۲ ریسک اعتباری

ریسک اعتباری عبارت است از ریسک مربوط به تأخیر در پرداخت و یا عدم پرداخت اقساط وام. در هریک از این دو حالت، ارزش فعلی دارایی کاهش می‌یابد که در نتیجه توان پرداخت تسهیلات دهی بانک کم می‌شود (هفرنان[1]، ۲۰۰۳).

با توجه به محدودیت منابع مالی و تسهیلات در اختیار بانکها، ارزیابی توان بازپرداخت مشتریان پیش از اعطای تسهیلات به آنها یکی از مهمترین چالش های پیشروی سیستم بانکی کشور است. به عبارت دیگر، بانک ها درصدد اعطای تسهیلات خود به مشتریانی هستند که ضمن برخورداری از ریسک پایین دارای بازده متناسب با سود تسهیلات اعطایی باشند، لذا تعیین ریسک اعتباری هریک از متقاضیان و اتخاذ تصمیم مناسب پیش از اعطای تسهیلات ضرورتی اجتناب ناپذیر است. این مهم زمانی امکانپذیر است که بانک ها قادر به شناسایی مشتریان اعتباری خود (اعم از حقیقی و حقوقی) براساس توانایی و تمایل آنها نسبت به بازپرداخت کامل و به موقع تعهدات خود بوده و قادر به طبقه بندی آنها باشند زیرا تحت چنین سیستمی تسهیلات به متقاضیانی اعطا می شود که رتبه کمتری داشته و احتمال بازپرداخت بدهی آنها در موعد مقرر بیشتر است. در ادامه با تعریف سیستم عصبی مصنوعی و سیستم خبره فازی و بررسی پیشینه پژوهش به بررسی شبکه عصبی مصنوعی و سیستم های فازی مرتبط با آن در جهت کنترل ریسک اعتباری می پردازیم.

یکی از مهم‌ترین روش‌های مدیریت ریسک اعتباری استفاده از سیستم‌های امتیازدهی اعتباری برای کنترل ریسک انواع وام‌هاست. در مطالعات گذشته اغلب از روش‌های آماری مانند مدل‌های رگرسیونی لاجیت و پروبیت و روش تحلیل ممیزی برای امتیازدهی و رتبه‌بندی مشتریان استفاده می‌گردید، ولی در سال‌های اخیر با توسعه مدل‌های مبتنی بر هوش مصنوعی و روش‌های ابتکاری، مطالعات بسیاری در کاربرد این روش‌ها در مدل‌های امتیازدهی و رتبه‌بندی اعتباری مشاهده می‌شوند (نادعلی جلوخانی و همکاران، ۱۳۹۷).

۱.۱.۲ شبکه عصبی مصنوعی (De Wilde, 1997)

شبکه عصبی مصنوعی، یک سیستم پردازش اطلاعات است که خصوصیات عملکردی آن مشابه شبکه‌های عصبی بیولوژیکی می‌باشد (فایاست[2]، ۱۹۹۴). به عبارت دیگر هدف شبکه‌های عصبی، کوشش برای ساخت ماشین‌هایی است که همانند مغز انسان عمل

---

[1] Heffernan

[2] Fausett

می‌کنند. این ماشین‌ها متشکل از اجزایی هستند که مشابه عصب‌های بیولوژیکی رفتار می‌کنند (پیکتون[1]، ۲۰۰۰).

۲.۱.۲ سیستم خبره فازی

سیستم‌های خبره فازی نتیجه به‌کارگیری تئوری فازی در ساختن سیستم‌های خبره است. یک سیستم خبره فازی یک سیستم خبره‌ای است که از یک مجموعه از توابع عضویت فازی و قوانین فازی به‌جای منطق بولین استفاده می‌کند. به عبارت دیگر، در سیستم‌های خبره فازی، علاوه بر استانداردهای سیستم خبره معمولی از داده فازی، قوانین فازی و استنتاج فازی استفاده می‌شود.

هدف از این تحقیق ارایه یک مدل مناسب اعتبارسنجی مشتریان بانک‌ها به طور بخصوص بانک سپه برای اعطای تسهیلات اعتباری متناسب با دسته مشتریان می‌باشد. برای پاسخ به مسأله و رسیدن به هدف پژوهش این سؤال اساسی مطرح گردید که چگونه می‌توان با استفاده از تکنیک درخت فازی به اعتبارسنجی مشتریان بانک سپه پرداخت؟ این تحقیق از نظر هدف، از نوع کاربردی[2] می‌باشد؛ چرا که تحقیق کاربردی، تحقیقی است که هدف آن، توسعه دانش کاربردی در یک زمینه خاص بوده و به سمت کاربرد علمی حرکت می‌کنند.

۳ پیشینه تحقیق

اسلامی[3] و همکاران (۲۰۱۴) در مطالعه خود با عنوان «ارزیابی وضعیت اعتباری مشتریان بانکی با استفاده از DEMATEL و سیستم فازی» به بررسی عوامل مؤثر در تعیین رفتار اعتباری مشتریان پراختند. آن‌ها با مدل‌سازی باتوجه به رویکرد حاضر و مدل‌های سنتی لاجیت و برنامه‌ریزی خطی به این نتیجه رسیدند که تقسیم‌بندی با استفاده از سیتسم استنتاج فازی توانایی بیش‌تری در توضیح‌دهندگی رفتار اعتباری مشتریان بانکی دارد.

عبدالرحمن و همکاران[4] (۲۰۱۴) در مطالعه‌ای با عنوان «رویکرد منطق فازی در تأمین مالی خرد در غنا» به بررسی رفتار اعتباری متقاضیان بانکی در این کشور پرداختند. آن‌ها با استفاده از داده‌های مربوط به وام‌دهی KWIQPLUS بهره بردند. آن‌ها در تحقیق خود نتیجه گرفتند که انتخاب متغیرها و هم‌چنین تعیین حد آستانه برای متغیرها می‌تواند بر نتایج و

---

[1] Picton

[2] Applied research

[3] Eslami

[4] Abdulrahman et. al

Downloaded from jmbr.mbri.ac.ir at 10:01 +0330 on Monday November 25th 2019

Downloaded from jmbr.mbri.ac.ir at 10:01 +0330 on Monday November 25th 2019

مفید بودن آن اثرگذار باشد. هم‌چنین این محققان بیان نمودند که پس‌انداز و به‌طور خاص توانایی سپرده‌گذاری روزانه اهمیت فراوانی بر تعیین رفتار اعتباری مشتریان دارد.

هوانگ و همکارانش (۲۰۱۶) در مطالعه‌ای با عنوان «مقایسه تطبیقی الگوریتم ژنتیک در پیش بینی» طی یک برنامه‌نویسی ژنتیک دو مرحله‌ای هم از الگوریتم‌های استنتاجی به‌منظور ایجاد قوانین هوشمند برای تصمیم‌گیرندگان و از روش‌های مبتنی بر تابع به دلیل توانایی پیش‌بینی‌شان استفاده کردند که بتواند نقص الگوریتم‌های استنتاجی را در این زمینه جبران کنند. نتایج نشان داد که الگوریتم‌های دو مرحله‌ای توانایی بیشتری را در پیش رفتار مشتریان بانک‌ها دارد.

تقوی فرد (۱۳۹۶) «دسته بندی مشتریان حقوقی و پیش بینی توانایی سوددهی آنان با استفاده از ارزش طول عمر مشتری و رویکرد زنجیره مارکوف»، ابتدا مبحث ارزش طول عمر مشتری و ضرورت و انواع دسته بندی مشتریان ارائه گردیده و سپس با استفاده از مدل RFM و تکنیک AHP و نظرات خبرگان بانکی نسبت به وزن دهی متغیرهای مورد بررسی اقدام و با توجه به میانگین سه متغیر فوق، مشتریان گروه بندی گردیدند. یافته اصلی این تحقیق را می توان استخراج ماتریس احتمال که قابلیت پیش بینی جابجایی مشتریان در گروه های مختلف را دارد ذکر نمود. با استفاده از این ماتریس به عدم تمایل تعداد زیادی از مشتریان به جابجایی در بین گروه های تعیین شده دست می یابیم. از دیگر نتایج این تحقیق، تعیین ضریب C است که جهت رتبه بندی گروه های مختلف استفاده شده است.

فرانسیسکو[1] (۲۰۱۶) در مطالعه خود با عنوان «ارزیابی سیستم اعتباری برای بانک‌های اعتباری شاخص» به بررسی وضعیت اعتباری بانک‌های فعال در کشور فرانسه پرداخت. این محقق در تحقیق خود سعی در برآورد هزینه و عایدی ناشی از پرداخت وام باتوجه به سیستم اعتباری موجود پرداخت. نتایج تحقیق نشان داد که درصورتی‌که سیستم اعتباری تنها بر ویژگی خاص درآمدی افراد تکیه کند، هزینه پرداخت وام به شدت افزایش می‌یابد که این عامل می‌تواند ناشی از اریب نتایج حاصل از ارزیابی سیستم اعتباری باشد.

لیو[2] (۲۰۱۶) در مطالعه خود با عنوان «اثر سیستم اعتباری پویا بر رفتار اعتباری مشتریان با استفاده از مدل‌های فازی مارکو» به بررسی اثرات پویای سیستم اعتباری بر عملکرد مشتریان از لحاظ خوش‌حسابی و بدحسابی پرداخت. نتایج تحقیق نشان می‌دهد که اثرات

---

[1] Fransisco

[2] Liu

Downloaded from jmbr.mbri.ac.ir at 10:01 +0330 on Monday November 25th 2019

ناهمسانی در بین مشتریان در گذر زمان وجود دارد که این عامل نشان می‌دهد باید سیستم فازی و به تبع آن قواعد آن باتوجه به این تغییرات در نظر گرفته شود.

تی سای و وو[1]، (۲۰۱۷) در مطالعه‌ای با عنوان «مقایسه عملکرد الگوریتم پس انتشار با مدل‌های جایگزین» استفاده از مدل‌های ترکیبی شبکه عصبی برای پیش‌گویی ورشکستگی و اعتبارسنجی پرداختند. در این تحقیق از پرسپترون چند لایه که توسط الگوریتم پس انتشار خطا آموزش می‌بیند برای اعتبارسنجی استفاده شد و نتایج آن با طبقه‌کننده‌های ترکیبی[2] مورد مقایسه قرار گرفت. از سه زیرمجموعه داده که شامل داده‌های اعتباری کشورهای استرالیا، آلمان و ژاپن بود، استفاده شد که مدل ترکیبی تنها در یکی از این سه مجموعه نتایج بهتری را کسب نمود. از خطای نوع اول و دوم برای مقایسه مدل‌ها استفاده شد.

محمدی (۱۳۹۱) در مطالعه‌ای با عنوان «اعتبار سنجی مشتریان بانک سامان» به تدوین یک مدل رتبه‌بندی اعتباری برای مشتریان حقیقی بانک سامان پرداخت. در این تحقیق از مدل پروبیت ترتیبی برای رتبه‌بندی استفاده شد. از ۹۲ مشتری و ۷ ویژگی برای تخمین نکول وام‌های پرداختی به کمک مدل پروبیت در نهایت مشتریان به سه طبقه خوش‌حساب، رسید گذشته و معوق تفکیک شدند.

عینی (۱۳۹۷) در مطالعه ای با عنوان «مدیریت بهینه ریسک اعتباری» به تبیین رویکرد مدیریت بهینه ریسک و همچنین ارائه تحلیل نظری در مقایسه با ریسک اعتباری در بانکداری اسلامی و بانکداری متداول پرداخت.

۴ روش‌شناسی پژوهش

۱.۴مدل مفهومی تحقیق

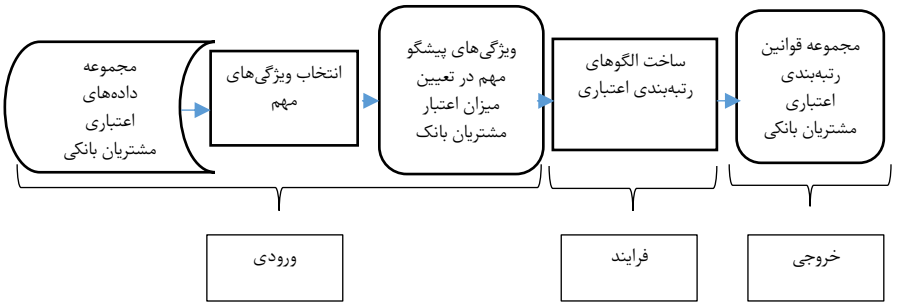
مدل مفهومی این تحقیق به شرح زیر می باشد:

---

[1] TSAI and Wu

[2] Classifier Ensembles

Downloaded from jmbr.mbri.ac.ir at 10:01 +0330 on Monday November 25th 2019



شکل ۱. مراحل تحقیق

۱.۱.۴ مدل مفهومی رتبه بندی اعتباری مشتریان بانک سپه

همان طور که در شکل ۱ شرح داده شده است، در اولین مرحله تحقیق از طریق مطالعات کتابخانه‌ای و بررسی مقالات پایان‌نامه‌های مختلف به شناسایی ویژگی‌های تأثیرگذار بر رتبه‌بندی اعتباری مشتریان حقیقی بانک‌ها پرداخته شده است. سپس با توجه اطلاعات در دسترس در فایل اطلاعات مشتریان بانک سپه ویژگی‌های مهم آن ها انتخاب و در مرحله بعد الگوهای مورد نظر در این تحقیق به منظور پیش بینی ریسک اعتباری مشتریان بانک سپه طراحی می‌شود و در انتها نتایج آنها بدست می‌آید.

۲.۴ مدل تحقیق

در این تحقیق از مدل‌های درخت تصمیم فازی با استفاده از روش بگینگ، سیستم فازی و شبکه عصبی استفاده شده است.

جامعه آماری در این تحقیق به جهت تعیین ویژگی‌های اعتباری، کلیه اشخاص حقیقی می‌باشند که در سال‌های ۱۳۹۵ تا ۱۳۹۶ از تسهیلات بانک سپه استفاده کرده‌اند.

هم‌چنین جامعه آماری به جهت تعیین توابع عضویت ویژگی‌ها و قواعد سیستم خبره فازی، ۵ نفر از کارکنان بانک سپه به عنوان خبرگان تحقیق می‌باشند، که همگی دارای تخصص‌های لازم در رابطه با بحث تحقیق می‌باشد. جهت گرداوری اطلاعات تحقیق، علاوه بر استفاده از پایگاه داده بانک سپه، در قسمت دوم مدل این تحقیق، برای طراحی سیستم خبره فازی، مصاحبه‌ای به صورت باز یا سازمان نایافته با کارشناسان اعتباری بانک سپه به عنوان خبره، به جهت تبیین متغیرهای مدل اعتبار سنجی انجام شد.

Downloaded from jmbr.mbri.ac.ir at 10:01 +0330 on Monday November 25th 2019

۳.۴ تجزیه و تحلیل اطلاعات

برای انجام تحقیق ابتدا به بررسی اسناد موجود به منظور تهیه ادبیات و پیشینه تحقیق پرداخته شده است. سپس باتوجه به هدف تحقیق، مطالعه و جمع آوری ادبیات در مورد مدل‌های هوش مصنوعی و به طور خاص یادگیری تلفیقی انجام شده است و داده‌های متغیرهای مورد استفاده در تحقیق جمع آوری می‌شود. در مرحله بعد با استفاده نرم افزارهای اکسل[1] داده ها پردازش و مدل مدنظر با بکارگیری نرم افزار متلب[2] برآورد می‌گردد و نتایج بدست آمده تفسیر خواهد شد.

در تحقیق حاضر در ابتدا با استفاده از آزمون انحنا معناداری و اثر هر یک از متغیرها در توضیح ریسک اعتباری مورد آزمون قرار می‌گیرد. بنابراین می‌توان متوجه شد که کدام یک از متغیرها بیشترین اثر را بر ریسک اعتباری دارد. سپس به طراحی شبکه عصبی پرداخته خواهد شد. در این بخش سعی خواهد شد که ۹۰ درصد از داده‌ها به منظور آموزش سیستم و ۱۰ درصد باقی مانده به‌منظور آزمون مورد استفاده قرار می‌گیرد. در مرحله بعد با استفاده منطق فازی به طراحی سیستم فازی و آموزش و آزمون بر اساس این سیستم استنتاج پرداخته خواهد شد. سپس با استفاده از الگوریتم C4.5 درخت تصمیم فازی تشکیل خواهد شد و این امر با استفاده از روش بوت کردن بگینگ انجام خواهد شد.

۵. داده‌ها

۱.۵ ویژگی‌های داده‌ها

- نمونه آماری

متشکل از داده‌های مربوط به ۸۲۶ مورد از مشتریان حقیقی بانک مورد مطالعه می‌باشد.

- نمونه‌های مورد بررسی پایگاه داده

باتوجه به این‌که داده‌های موجود در جامعه آماری از لحاظ کیفیت مورد بررسی قرارگرفته و تا حد امکان پالایش شده‌اند، بنابراین مدل‌های تحقیق فقط بر روی ۸۲۶ مورد از داده‌ها به عنوان نمونه نهایی اجرا گردیده است. این داده‌ها در یک پایگاه داده در نرم‌افزار Excel ذخیره شده‌اند. که در ادامه فیلدها و رکوردهای این پایگاه داده تشریح می‌گردند. در ضمن چارچوب کلی پایگاه داده نیز به‌صورت جدول ۲ می‌باشد.

---

[1] Excel

[2] Matlab

Downloaded from jmbr.mbri.ac.ir at 10:01 +0330 on Monday November 25th 2019

جدول ۲

چارچوب کلی پایگاه داده مورد بررسی

طبقه‌ها	A19	...	A2	A1	خصیصه‌ها
Class1					مشتریان
Class3			مقادیر		C1
...					C2
Class2					...
Class2					C120

- رکوردهای پایگاه داده

مشتریان بانک که همان نمونه مورد بررسی تحقیق می‌باشند، رکوردهای پایگاه داده نهایی را تشکیل می‌دهند.

- فیلدهای پایگاه داده

خصیصه‌های مشتریان که همان متغیرهای مستقل تحقیق و طبقه مشتری که همان متغیر وابسته تحقیق می‌باشند، فیلدهای این پایگاه داده را تشکیل می‌دهند. عناوین این فیلدها شامل شش خصیصه مشتریان و ۱ طبقه کلاس می‌باشند که در جدول ۳ نشان داده شده است.

جدول ۳

عناوین فیلدها

نام فیلد	نام لاتین
درآمد	Wage
سن	Age
سابقه اعتباری	Past-behavior
تحصیلات	Schooling
تراکنش	Transaction
واریانس تراکنش	Variance

- نوع و مقیاس داده‌ها

نوع داده‌های تحقیق (که همان نوع متغیرها می‌باشد) به‌صورت داده‌های عددی (ترتیبی) می‌باشند که در این تحقیق به داده‌های فازی تبدیل می‌شوند. مقیاس‌های داده‌های عددی در جدول ۴ قابل مشاهده می‌باشند.

جدول ۴
مقیاس‌های داده‌های عددی

مقیاس‌ها	متغیرهای عددی (قابل فازی نمودن)
سال	سن مشتری
هزار ریال	تراکنش
درصد	واریانس تراکنش
نسبت	سابقه اعتباری مشتری
هزار ریال	درآمد
مرتبه	تحصیلات

- ویژگی‌های خبرگان (کارشناسان بانک مورد مطالعه)

ویژگی‌های خبرگان در جدول ۵ آورده شده است. این ویژگی‌ها گواهی بر خبرگی و اطمینان به دقت و صحت گفته‌های آنان می‌باشد.

جدول ۵
ویژگی‌های خبرگان

ردیف	سمت	سن	میزان تحصیلات	تجربه سیستم بانکی (سال)	تجربه در بانک مورد مطالعه (سال)	تجربه در بخش اعتبارات (سال)
کارشناس ۱	مدیر بخش رتبه‌بندی اعتباری	۵۴	کارشناسی ارشد	۳۰	۳	۱۷
کارشناس ۲	ریاست کمیسیون اعتبارات	۵۱	کارشناسی	۳۳	۶	۲۷
کارشناس ۳	سرپرست کارشناسان بخش اعتبارات	۴۴	کارشناسی ارشد	۱۵	۵	۱

- آماده‌سازی داده‌ها برای مدل

تغییرات و آماده سازی داده ها در سه بخش اصلی اصلاح اشتباهات کاربر، یک شکل نمودن داده‌ها، و فازی نمودن متغیرها صورت گرفته است که در ادامه هر کدام بیان شده‌اند.

- یک شکل نمودن داده‌ها

Downloaded from jmbr.mbri.ac.ir at 10:01 +0330 on Monday November 25th 2019

از جمله فیلدهایی که در این مرحله تغییرات بر روی آن‌ها صورت گرفته در جدول ۶ قابل مشاهده می‌باشند.

جدول ۶
یک شکل نمودن داده‌ها

شکل جدید	شکل قدیم	فیلد
ماه	روز	تراکنش
سن	سال تولد	سن

- فازی نمودن متغیرها

فازی نمودن متغیرها در سه مرحله اصلی جدا نمودن متغیرها، تعیین اصطلاحات کلامی و طیف متغیرها در مقادیر کلامی، و تعیین تابع عضویت صورت گرفته است.

۲.۵ توصیف آماری داده‌ها

ویژگی آماری متغیرهای تحقیق در جدول ۷ گزارش شده است.

جدول ۷
توصیف آماری داده ها

نام متغیر	میانگین	میانه	مد	حداقل	حداکثر	انحراف معیار
Variance	۰/۲۳	۰/۲۴	۰/۰۱۳	۰/۰۱۳	۰/۵۴	۰/۰۸۲
Schooling	۱۰/۶۹	۱۱	۱۰	۱	۲۲	۳/۲۷
Wage	۴۰۰۰	۴۰۱۳	۳۳۱۲	۱۳۱۶	۷۰۰۷	۸۷۱/۵
Transaction	۲۴۶۰۰	۲۴۶۵۵	۲۲۵۲۳	۲۷۴۱	۴۹۶۱۴	۶۵۲۰
Age	۴۵/۳	۴۵	۴۳	۱۶	۸۷	۱۰/۵
Past-behavior	۱/۲۹۲	۱/۲۹۷	۰/۸۶	۰/۸۶	۱/۷۷	۰/۱۳
Fraud	۰/۵۱	۱	۱	۰	۱	۰/۵۰

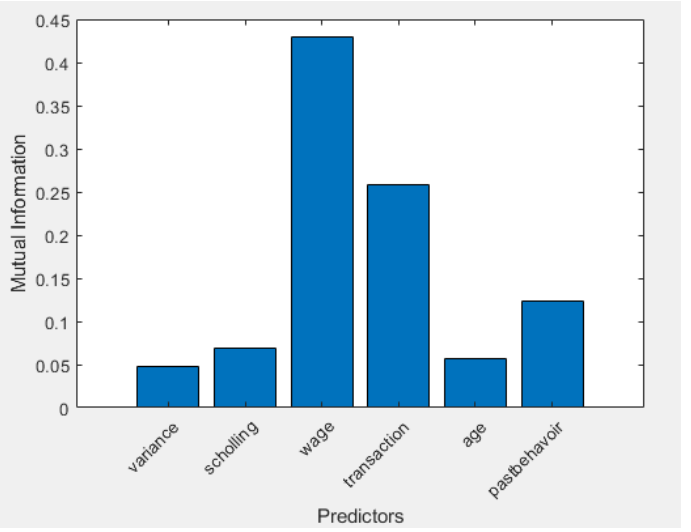
Downloaded from jmbr.mbri.ac.ir at 10:01 +0330 on Monday November 25th 2019

Downloaded from jmbr.mbri.ac.ir at 10:01 +0330 on Monday November 25th 2019

۶ نتایج مدل

ابتدا جهت پاسخ به پرسش اول تحقیق درخت تصمیم معمولی اجرا شده و به کمک آنتروپی[1] در درخت تصمیم که در آن فرضیه صفر، غیرمرتبط[2] بودن متغیرهای توضیحی است، اهمیت هر یک از متغیرهای توضیحی مورد بررسی و تقریب قرارگرفت که نتایج آن به شرح ذیل است:

۱.۶ تعیین اهمیت متغیرها



شکل ۲. درجه اهمیت متغیرهای مدل

در این بخش با استفاده از آزمون انحنا مشخص می‌شود که در بین متغیرهای مورد بررسی در تحقیق درآمد شخص از درجه اهمیت بالاتری نسبت به سایر متغیرها برخوردار است.

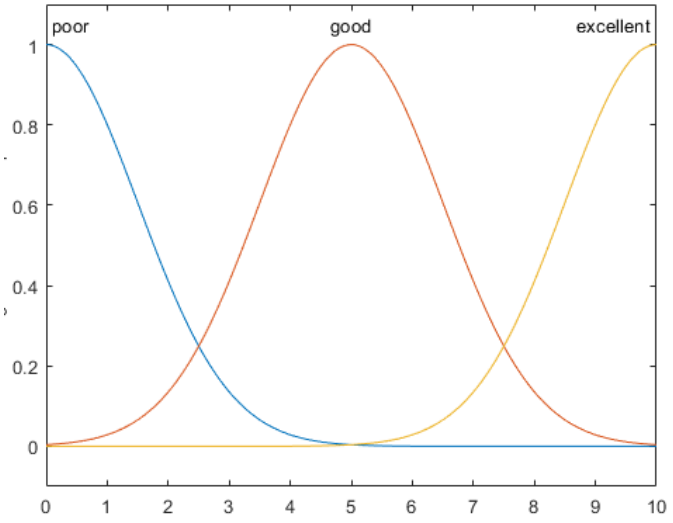
۲.۶ تعیین تابع عضویت

تابع عضویتی که در این تحقیق برای تعیین درجات عضویت بکارگرفته شده است، «گوسی» می‌باشد. بر همین اساس با کمک خبرگان بانک، مقادیر متغیرکه متناسب با درجات عضویت

---

[1] Antropy
[2] Unassociated

صفر و یک هستند را برای نقاط به ترتیب پایین‌ترین و بالاترین درجات عضویت تعیین نموده و بقیه مقادیر دیگر، بین صفر و یک در نظر گرفته می‌شوند.



شکل ۳. تابع عضویت گوسی

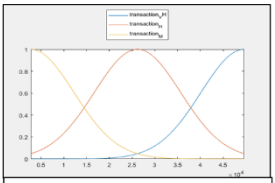
۳.۶ فازی‌سازی طبقات

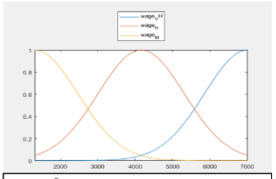
باتوجه به میزان «نمره اعتباری تعیین شده برای مشتریان متقلب» (به عنوان مبنا) در هر رکورد (برای هریک از مشتریان)، یک درجه تعلق به کلاس‌های هر یک (متقلب، نامتقلب) در نظر گرفته می‌شود. بدین ترتیب طبقات نیز حالت قطعی نداشته و فازی تعیین می‌گردند.

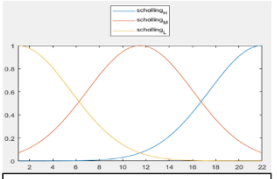
اکنون باتوجه به فازی نمودن متغیرها و طبقات، همه داده‌ها به فرمت‌های مشخصی برای ورودی به نرم‌افزار تبدیل می‌شوند که در نرم‌افزار مورد استفاده (متلب) موجود می‌باشد. فازی شده هر یک از متغیرها در اشکال ذیل نشان داده شده‌اند.

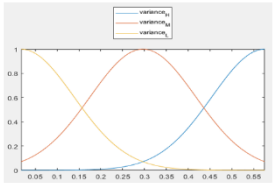
در ابتدا متغیرهای مدل به متغیرهای فازی با سه تابع عضویت کم، متوسط و زیاد تبدیل می‌شوند:

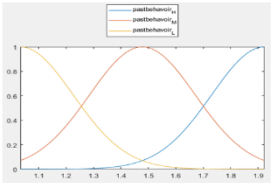
Downloaded from jmbr.mbri.ac.ir at 10:01 +0330 on Monday November 25th 2019

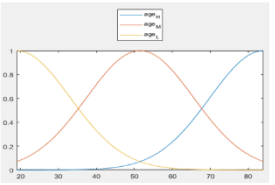

تابع عضویت فازی و بازه تغییرات تراکنش

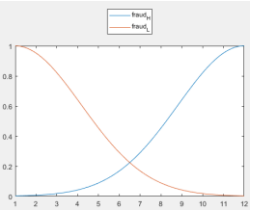

تابع عضویت فازی و بازه تغییرات درآمد


تابع عضویت فازی و بازه تغییرات تحصیلات


تابع عضویت فازی واریانس تراکنش


تابع عضویت فازی و بازه تغییرات سابقه اعتباری


تابع عضویت فازی و بازه تغییرات سن


تابع عضویت فازی نمره اعتباری

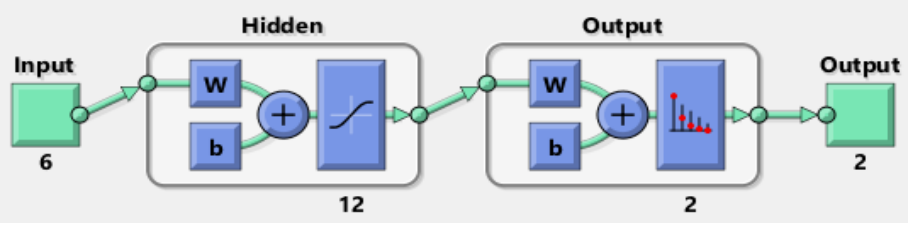
شکل ۴. نمایش فازی شده هر یک از توابع متغیرهای مدل

باتوجه به هدف تحقیق، در ابتدا متغیر وابسته تحقیق (نمره اعتباری) با استفاده از تابع گوسی فازی‌سازی شده، سپس در ادامه تحقیق از داده‌های فازی در فرایند مدلسازی استفاده شده است به نحوی که شش متغیر توضیحی هر یک دارای ۳ تابع عضویت و یک متغیر هدف دارای دو تابع عضویت گوسی می‌باشد.

۴.۶ برآورد مدل شبکه عصبی فازی

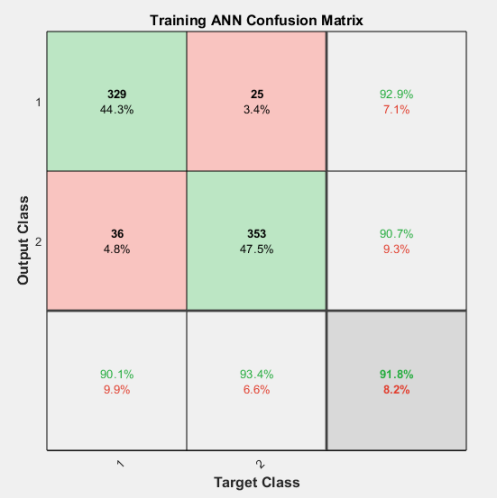
به منظور مقایسه کارایی الگوی درخت تصمیم فازی با الگوی رقیب به برآورد شبکه عصبی فازی با داده های مورد استفاده در درخت تصمیم پرداخته می‌شود. بدین منظور یک شبکه عصبی پیش‌خور با داده‌های فازی ایجاد و به کمک الگوریتم آموزش لونبرگ مارکوارت به آموزش آن پرداخته شد، تعداد نرون‌های لایه پنهان براساس آزمون و خطا ۱۲ نرون انتخاب شد. شکل زیر ساختار شبکه را نشان می‌دهد:

Downloaded from jmbr.mbri.ac.ir at 10:01 +0330 on Monday November 25th 2019



شکل ۵. ساختار شبکه عصبی دسته بندی کننده مشتریان

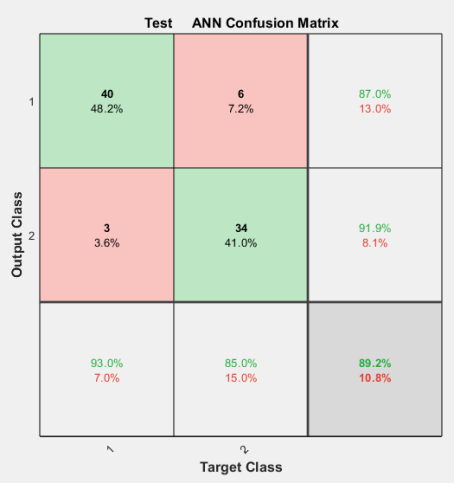
همچنین ماتریس ارزیابی کارایی شبکه عصبی پس از آموزش و ارزیابی بصورت زیر حاصل شد:



شکل ۶. ماتریس درهم ریختگی شبکه عصبی برای داده های آموزش

شکل (۶) با توجه به ماتریس درهم‌ریختگی را برای شبکه عصبی می‌توان متوجه شد که دقت در مرحله آموزش برابر با ۹۱/۸ درصد است و الگو توانسته است گروه ۱ و ۲ به ترتیب با دقت ۹۲/۹ و ۹۰/۷ درصد شناسایی نماید. هم‌چنین نتایج نشان می‌دهد که الگو ۳۲۹ مورد از ۳۵۴ مورد گروه ۱ را به درستی شناسایی کرده است و در ۲۵ مورد خطا داشته است. ماتریس درهم‌ریختگی نشان می‌دهد که با توجه به الگوی طراحی شده ۳۵۳ مورد از گروه ۲ به درستی و ۳۶ مورد با خطا شناسایی شده است.

Downloaded from jmbr.mbri.ac.ir at 10:01 +0330 on Monday November 25th 2019



شکل ۷. ماتریس درهم‌ریختگی شبکه عصبی برای داده‌های آموزش

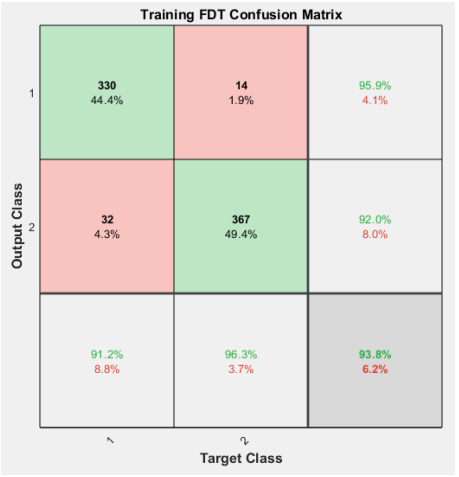
شکل (۷) ماتریس درهم‌ریختگی برای مجموعه آموزش بیانگر این امر است که دقت تست برابر با ۸۹/۲ درصد است و دقت پیش‌بینی گروه ۱ و ۲ به ترتیب برابر با ۸۷ و و ۹۱/۹ درصد است. نتایج نشان می‌دهد که الگو ۴۰ مورد از بین ۴۶ مورد گروه اول را به درستی پیش‌بینی نموده است و در گروه دوم ۳۴ مورد از ۳۷ مورد تشخصی درست گزارش داده است.

۵.۶ سیستم فازی

در این بخش به طراحی سیستم فازی پرداخته شده است.

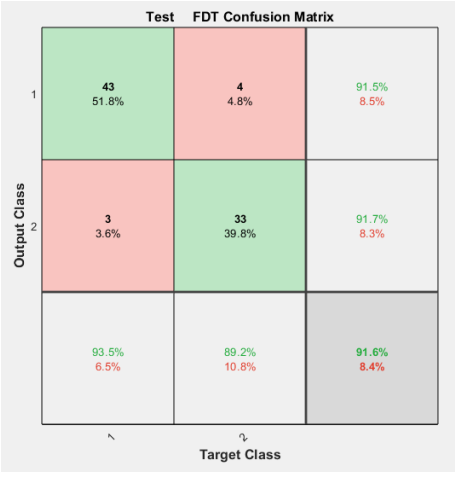
Downloaded from jmbr.mbri.ac.ir at 10:01 +0330 on Monday November 25th 2019

Downloaded from jmbr.mbri.ac.ir at 10:01 +0330 on Monday November 25th 2019



شکل ۸. ماتریس درهم‌ریختگی دقت درخت تصمیم فازی در داده‌های آموزش

شکل (۸) نتایج ماتریس آموزش را برای ۹۰ درصد از داده‌ها در یک سیستم فازی نشان می‌دهد نشان می‌دهد که این الگو توانسته است ۹۳/۸ درصد را به درستی شناسایی نماید و باتوجه به نتایج می‌توان مشاهده نمود که دقت در شناسایی گروه ۱ و ۲ به ترتیب برابر با ۹۵/۹ و ۹۲ درصد بوده است.



شکل ۹. ماتریس درهم ریختگی دقت درخت تصمیم فازی در داده های آزمایش

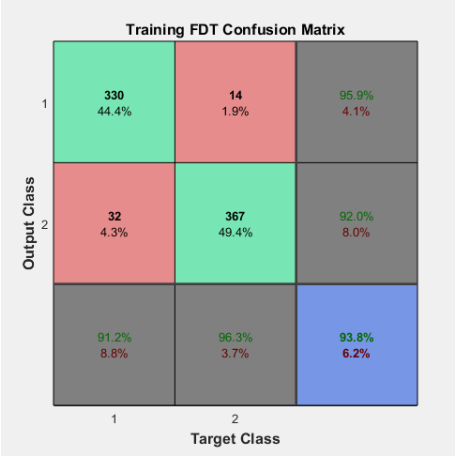
با توجه ماتریس درهم‌ریختگی برای داده‌های تست در شکل (۹) می‌توان مشاهده نمود که الگو توانسته است ۴۳ مورد از گروه ۱ را به درستی گزارش نماید و ۴ مورد را خطا نماید. هم‌چنین نتایج نشان می‌دهد که الگو در شناسایی گروه ۲ در ۳۳ مورد موفق بوده است و ۳ مورد را خطا کرده است. ماتریس مذکور نشان می‌دهد که دقت در شناسایی گروه ۱ و ۲ به ترتیب برابر با ۹۱/۵ و ۹۱/۷ درصد بوده است و دقت الگو در داده‌های آموزش برابر با ۹۱/۶ درصد گزارش شده است.

۶.۶ روش بگینگ

از این روش در ایجاد تخمینهای مختلف استفاده می شود. اصولاً میتوان از این تکنیک به منظور ارزیابی دقت تخمینهای بکارگرفته شده در روشهای داده کاوی از طریق نمونه برداری با جایگزینی از داده‌های آموزشی استفاده نمود. در این تکنیک فرض بر آنست که مجموعه‌داده های آموزشی نماینده جامعه تحت بررسی بوده و انواع حالات تحقق یافته جامعه را میتوان از این مجموعه داده شبیه سازی نمود. این مفهوم برای ترکیب رده بندی های پیش بینی شده از چند مدل به کار می رود. در این روش به طور کلی برای نمونه های مختلف به درخت های متفاوتی خواهید رسید. سپس برای پیش بینی با کمک درخت های متفاوت به دست آمده از نمونه ها، یک رای گیری ساده انجام دهید. رده بندی نهایی، رده بندی ای خواهد بود که درخت های مختلف آنرا پیش بینی کرده اند.

در این بخش به بررسی نتایج بدست آمده از طراحی درخت تصمیم فازی با استفاده از تکنیک بگینگ پرداخته شده است. در این بخش داده‌های تحقیق به ۱۰ بخش تقسیم‌بندی شده‌اند و با استفاده از الگوی بگینگ ده بار درخت تصمیم مورد آموزش و تست قرار گرفته است و نتایج آن به صورت ماتریس‌های درهم‌ریختگی تفسیر شده است در مرحله بعدی با استفاده از تکنیک رای گیری از بین نتایج بدست آمده یک نتیجه گزارش شده است.

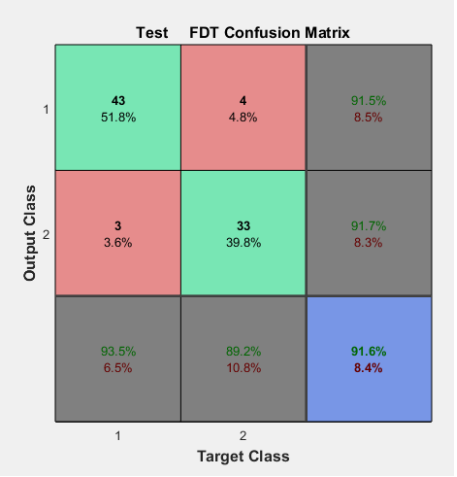
Downloaded from jmbr.mbri.ac.ir at 10:01 +0330 on Monday November 25th 2019



شکل ۱۰. ماتریس درهم‌ریختگی دقت درخت تصمیم فازی در داده‌های آموزش در فولد اول

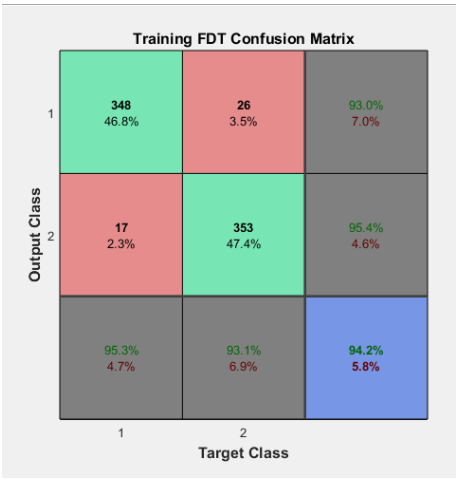
ماتریس فوق نتایج حاصل از آموزش درخت تصمیم فازی را در فولد اول نشان می‌دهد. در مرحله فولد اول که ۲۰ درصد اول مشاهدات است به عنوان مجموعه تست کنار گذاشته می‌شود و با سایر مشاهدات درخت تصمیم فازی در مرحله آموزش قرار می‌گیرد. نتایج نشان می‌دهد که الگوی طراحی شده در فولد اول ۹۵/۹ درصد در شناسایی افراد خوش حساب دقت داشته است و تنها در ۴/۱ درصد خطا داشته است در حالی‌که این درصد برای گروه ۲ به ترتیب ۹۲ درصد و ۸ درصد بوده است. با توجه به نتایج می‌توان متوجه شد که دقت درخت تصمیم در فولد اول برابر با ۹۳/۸ در صد بوده است.

Downloaded from jmbr.mbri.ac.ir at 10:01 +0330 on Monday November 25th 2019



شکل ۱۱. ماتریس درهم ریختگی دقت درخت تصمیم فازی در داده های تست در فولد اول

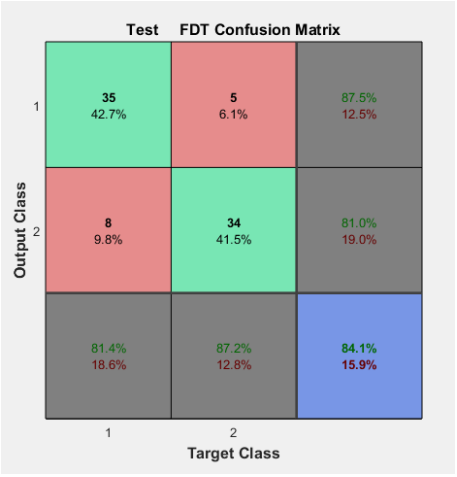
شکل (۱۱) ماتریس درهم‌ریختگی را برای فولد اول نشان می دهد. نتایج نشان می‌دهد که الگو توانسته است دقتی برابر با ۹۱/۶ درصد را در این فولد داشته باشد.



شکل ۱۲. ماتریس درهم ریختگی دقت درخت تصمیم فازی در داده های آموزش در فولد دوم

Downloaded from jmbr.mbri.ac.ir at 10:01 +0330 on Monday November 25th 2019

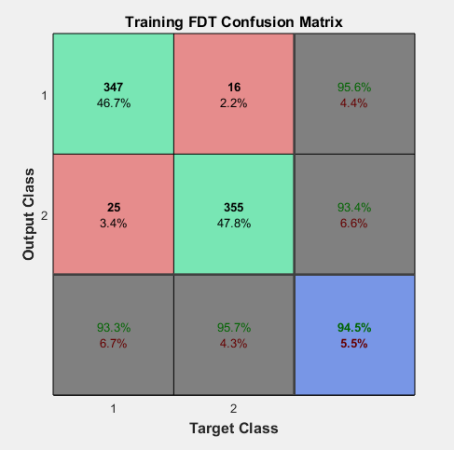
شکل (۱۲) نیز نتایج حاصل از آموزش الگو را در فولد ۲ نشان می‌دهد که دقت شناسایی گروه یک برابر با ۹۳ درصد و برای گروه ۲ برابر با ۹۵/۴ درصد بوده است. هم‌چنین نتایج نشان می‌دهد که دقت درخت در ایان مرحله با با ۹۴/۲ درصد است.



شکل ۱۳. ماتریس درهم ریختگی دقت درخت تصمیم فازی در داده های تست در فولد دوم

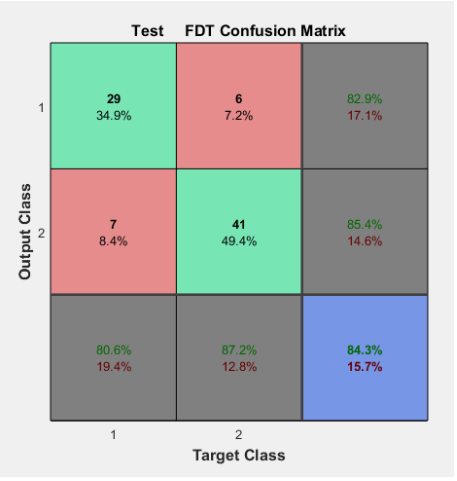
شکل (۱۳) نیز نتایج حاصل از تست در فولد دوم را نشان می‌دهد که دقت تست در این فول برابر با ۸۴/۱ درصد بوده است که در مقایسه با فولد قبلی کاهش یافته است.

Downloaded from jmbr.mbri.ac.ir at 10:01 +0330 on Monday November 25th 2019



شکل ۱۴. ماتریس درهم ریختگی دقت درخت تصمیم فازی در داده های آموزش در فولد سوم

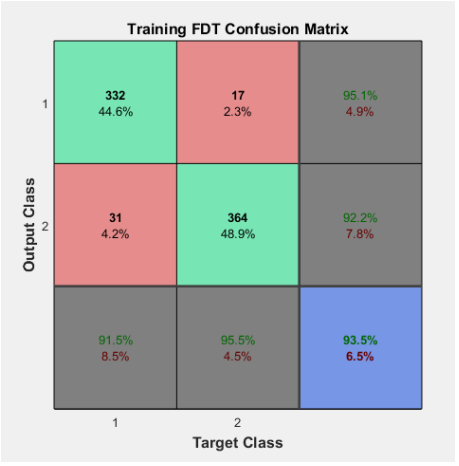
شکل (۱۴) ماتریس درهم‌ریختگی را برای فولد سوم نشان می‌دهد الگو در شناسایی گروه اول ۹۶/۶ درصد دقت داشته است در حالی‌که این دقت برای گروه دوم برابر با ۹۳/۴ درصد بوده است. هم‌چنین نتایج حاصل از الگو در فولد نشان می‌دهد که دقت آموزش برابر با ۹۴/۵ درصد بوده است.



شکل ۱۵. ماتریس درهم ریختگی دقت درخت تصمیم فازی در داده های تست در فولد سوم

Downloaded from jmbr.mbri.ac.ir at 10:01 +0330 on Monday November 25th 2019

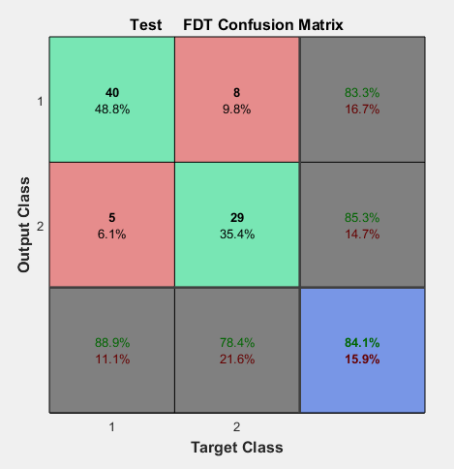
شکل (۱۵) ماتریس درهم‌ریختگی را برای فولد سوم در مرحله تست نشان می‌دهد. نتایج نشان می‌دهد که الگوی طراحی شده توانسته است در پیش بینی گروه یک ۸۲/۹ درصد موفق باشد. هم‌چنین الگوی در فولد دوم توانسته است که ۴۱ نفر از ۴۸ نفر را که به گره دوم متعلق بوده‌اند را به خوبی پیش‌بینی نماید. نتایج نشان می‌دهد که دقت پیش‌بینی در این مرحله برابر با ۸۴/۳ درصد بوده است.



شکل ۱۶. ماتریس درهم ریختگی دقت درخت تصمیم فازی در داده های آموزش در فولد چهارم

شکل (۱۶) نتایج حاصل از آموزش درخت تصمیم فازی را در فولد چهارم نشان می‌دهد. با توجه به ماتریس درهم‌ریختگی می‌توان متوجه شد که دقت شناسایی گروه یک برابر با ۹۵/۱ درصد بوده است در حالی‌که برای گروه ۲ دقت برابر با ۹۲/۲ درصد بوده است. ماتریس درهم‌ریختگی نشان می‌دهد که دقت در این مرحله برابر با ۹۳/۵ درصد بوده است.

Downloaded from jmbr.mbri.ac.ir at 10:01 +0330 on Monday November 25th 2019

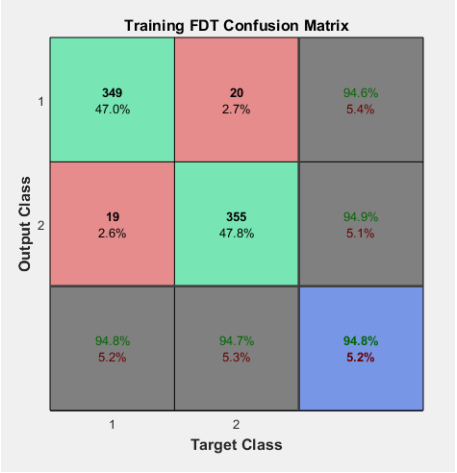


شکل ۱۷. ماتریس درهم ریختگی دقت درخت تصمیم فازی در داده های تست در فولد چهارم

شکل (۱۷) نیز ماتریس درهم‌ریختگی تست را برای فولد چهارم نشان می‌دهد. نتایج نشان می‌دهد الگو در این مرحله توانسته است که ۸۳ درصد از افراد را به خوبی پیش‌بینی نماید و تنها در ۱۶/۷ درصد خطا داشته است. باتوجه نتایج تست در فولد سوم می‌توان متوجه شد که دقت پیش‌بینی الگو برای گروه دوم نسبت به گروه اول بیشتر بوده است و الگو توانسته است که ۸۵/۳ درصد از افراد در گروه دوم را به خوبی پیش بینی نماید. در نهایت می‌توان مشاهده نمود که دقت تست در این مرحله برابر با ۸۴/۱ درصد بوده است.

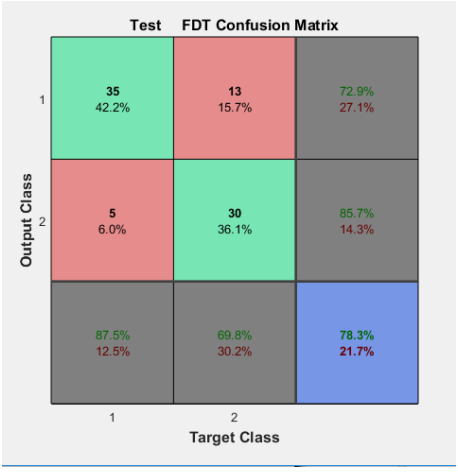
Downloaded from jmbr.mbri.ac.ir at 10:01 +0330 on Monday November 25th 2019

Downloaded from jmbr.mbri.ac.ir at 10:01 +0330 on Monday November 25th 2019



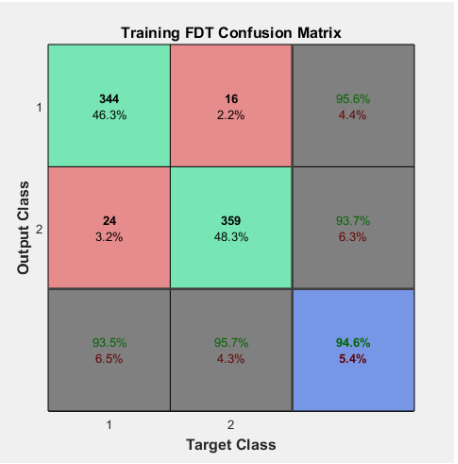
شکل ۱۸. ماتریس درهم ریختگی دقت درخت تصمیم فازی در داده های آموزش در فولد پنجم

شکل (۱۸) نیز ماتریس درهم‌ریختگی را برای مجموعه آموزش در فولد پنجم نشان می‌دهد. باتوجه به ماتریس می‌توان متوجه شد که دقت شناسایی گروه ۱ و ۲ به ترتیب برابر با ۹۴/۶ و ۹۴/۲ است. هم‌چنین نتایج نشان می‌دهد که دقت در این مرحله برابر با ۹۴/۸ درصد بوده است.



شکل ۱۹. ماتریس درهم ریختگی دقت درخت تصمیم فازی در داده های تست در فولد پنجم

شکل (۱۹) نتایج تست برای فولد پنجم نشان می‌دهد. باتوجه ماتریس درهم‌ریختگی می‌توان متوجه شد که الگو طراحی شده در این فولد توانسته است که ۳۵ مورد از ۴۸ مورد را برای گروه ۱ به خوبی پیش‌بینی نماید و در ۱۳ مورد خطا داشته است. با توجه به نتایج می‌توان متوجه شد که دقت پیش بینی برای گروه ۱ برابر با ۷۲/۹ درصد بوده است در حالی‌که این دقت برای گروه ۲ برابر با ۸۵/۷ درصد بوده است. باتوجه ماتریس درهم‌ریختگی می‌توان مشاهده نمود که دقت پیش‌بینی در این مرحله برابر با ۷۸/۳ درصد بوده است.

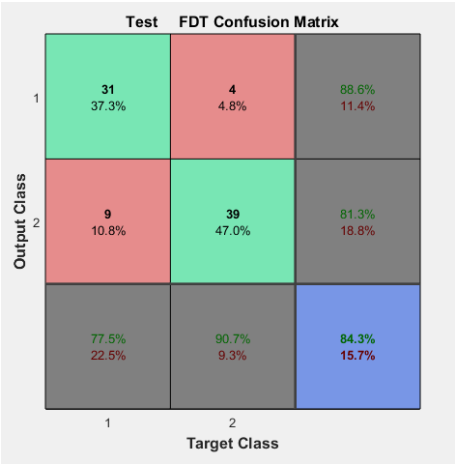


شکل ۲۰. ماتریس درهم ریختگی دقت درخت تصمیم فازی در داده های آموزش در فولد ششم

شکل (۲۰) نتایج آموزش درخت تصمیم را در فولد ششم نشان می‌دهد که دقت آموزش در این فولد برابر با ۹۴/۶ درصد بوده است. باتوجه به ماتریس می‌توان مشاهده نمود که الگو در شناسایی گروه ۱ و ۲ به ترتیب ۱۶ و ۲۴ مورد خطا داشته است.

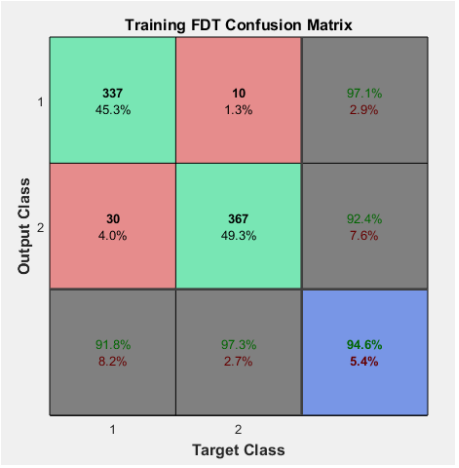
Downloaded from jmbr.mbri.ac.ir at 10:01 +0330 on Monday November 25th 2019

Downloaded from jmbr.mbri.ac.ir at 10:01 +0330 on Monday November 25th 2019



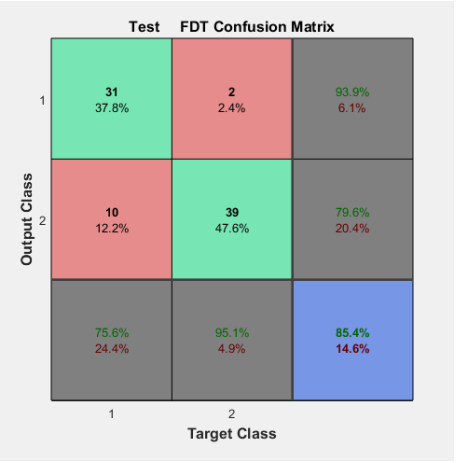
شکل ۲۱. ماتریس درهم ریختگی دقت درخت تصمیم فازی در داده های تست در فولد ششم

ماتریس درهم‌ریختگی برای مجموعه تست در فولد ششم را می‌توان در شکل (۲۱) مشاهده نمود. نتایج نشان می‌دهد که دقت پیش‌بینی گروه ۱ برابر با ۸۸/۶ درصد است. هم‌چنین با توجه به نتایج می‌توان متوجه شد که دقت پیش‌بینی گروه اول برابر با ۸۱/۳ درصد است و الگو توانسته است که ۳۹ مورد از ۴۸ مورد در گروه ۲ را به خوبی پیش‌بینی نماید. نتایج نشان می‌دهد که دقت پیش‌بیتی در این مرحله برابر با ۸۴/۳ درصد است.



شکل ۲۲. ماتریس درهم ریختگی دقت درخت تصمیم فازی در داده های آموزش در فولد هفتم

ماتریس درهم‌ریختگی آموزش را برای فولد هفتم می‌توان در شکل (۲۲) نتایج این ماتریس نشان می‌دهد که الگو در این مرحله ۳۳۷ مورد را در گروه ۱ را به درستی شناسایی نموده است و در ۱۰ مورد خطا مرتکب شده است. هم‌چنین از نتایج بدست آمده می‌توان متوجه شد که دقت شناسایی گروه ۱ و ۲ به ترتیب برابر با ۹۷/۱ و ۹۲/۲ درصد بوده است. در نهایت می‌توان بیان نمود که دقت آموزش در این مرحله برابر با ۹۴/۶ درصد بوده است.

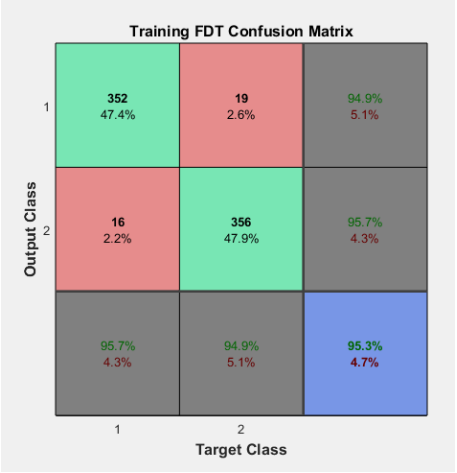


شکل ۲۳. ماتریس درهم ریختگی دقت درخت تصمیم فازی در داده های تست در فولد هفتم

شکل (۲۳) ماتریس درهم‌ریختگی تست را برای فولد هفتم می‌توان بیان نمود که الگو توانسته است ۳۱ نفر از ۳۳ نفر را در گروه ۱ به خوبی پیش‌بینی نماید و در ۲ مورد خطا داشته است و برهمین اساس دقت پیش‌بینی گروه ۱ برابر با ۹۳/۹ درصد شده است. هم‌چنین نتایج نشان می‌دهد که الگو در شناسایی گروه ۲ دقتی برابر با ۷۹/۶ درصد را داشته است و توانسته است ۳۹ مورد از ۴۹ مورد را به درستی پیش‌بینی نماید و در ۱۰ مورد خطا داشته است. در نهایت ماتریس نشان می‌دهد که دقت تست در فولد ششم برابر با ۸۵/۴ درصد است.

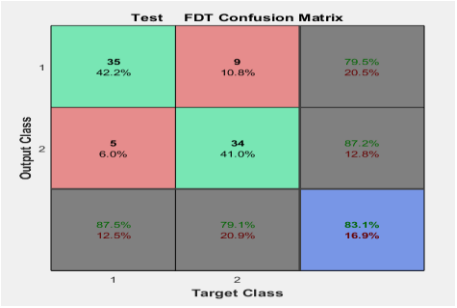
Downloaded from jmbr.mbri.ac.ir at 10:01 +0330 on Monday November 25th 2019

Downloaded from jmbr.mbri.ac.ir at 10:01 +0330 on Monday November 25th 2019



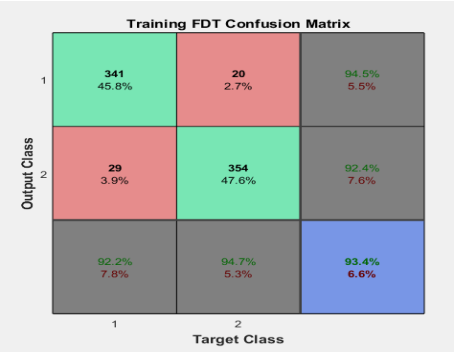
شکل ۲۴. ماتریس درهم ریختگی دقت درخت تصمیم فازی در داده های آموزش در فولد هشتم

شکل (۲۴) نتایج بدست آمده از آموزش در فولد هشتم نشان می‌دهد که الگو در این مرحله توانسته است که ۳۵۲ مورد از گروه اول را به درستی شناسایی نماید و در ۱۹ مورد خطا مرتکب شده است. باتوجه به ماتریس مذکور می‌توان مشاهده نمود که دقت شناسایی در گروه ۱ برابر با ۹۴/۹ درصد است. هم‌چنین نتایج نشان می‌دهد که الگو در شناسایی گروه ۲ دقتی برابر با ۹۵/۷ درصدی را دارا است همچنین دقت آموزش در این فولد برابر با ۹۵/۳ درصد است.



شکل ۲۵. ماتریس درهم ریختگی دقت درخت تصمیم فازی در داده های تست در فولد هشتم

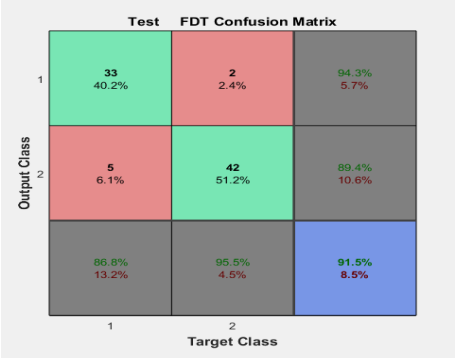
شکل (۲۵) نتایج بدست آمده از تست در فولد هشتم نشان می‌دهد که الگو توانسته است که با دقت ۷۹/۵ درصدی شناسایی گروه اول ۳۵ مرود را به درستی گزارش کند و در ۹ مورد خطا مرتکب شده است. هم‌چنین نتایج نشان می‌دهد که الگو توانسته است که ۳۴ مورد گروه ۲ را به درستی پیش‌بینی نماید و در ۵ مورد خطا داشته است و بر همین اساس دقت پیش‌بینی گروه ۲ برابر با ۸۷/۲ درصد بوده است. هم‌چنین نتایج نشان می‌دهد که دقت تست در فولد هفتم برابر با ۸۳/۱ درصد است.



شکل ۲۶. ماتریس درهم ریختگی دقت درخت تصمیم فازی در داده های آموزش در فولد نهم

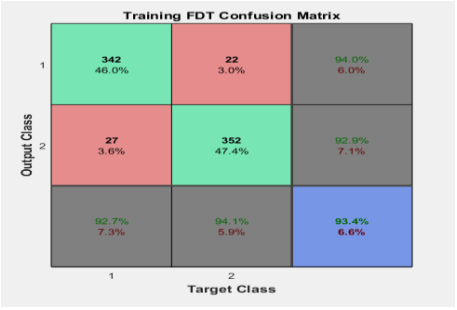
شکل (۲۶) نتایج بدست آمده از آموزش الگو می‌توان مشاهده نمود که دقت آموزش در این مرحله برابر با ۹۳/۴ درصد بوده است. همچنین الگو در فولد نهم توانسته است ۳۴۱ مورد از گروه ۱ را به درستی گزارش نماید و در ۲۰ مورد خطا داشته است و در شناسایی گروه ۱ دقت ۹۴/۵ درصدی را گزارش نموده است. نتایج برای گروه ۲ نشان می‌دهد که الگو توانسته است ۳۵۴ مورد از ۳۸۳ مورد را به درستی گزارش نماید و دقتی ۹۲/۴ درصدی را درشناسایی گروه ۲ نشان دهد.

Downloaded from jmbr.mbri.ac.ir at 10:01 +0330 on Monday November 25th 2019



شکل ۲۷. ماتریس درهم ریختگی دقت درخت تصمیم فازی در داده های تست در فولد نهم

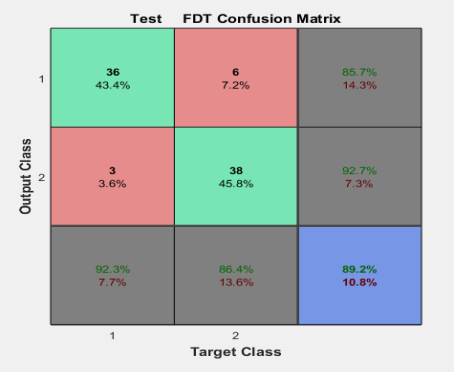
شکل (۲۷) ماتریس درهم‌ریختگی را برای مجموعه آموزش در فولد نهم نشان می‌دهد که الگو توانسته است که ۳۳ مورد در پیش‌بینی گروه ۱ به درستی عمل نماید و در ۲ مورد خطا مرتکب شده است و بر همین اساس دقتی برابر ۹۴/۳ درصد را برای شناسایی گروه ۱ گزارش داده است. هم‌چنین با توجه به ماتریس می‌توان متوجه شد که الگو توانسته است در پیش‌بینی گروه ۲ دقت ۸۹/۴ درصدی را نشان دهد و ۴۲ مورد از ۴۷ مورد را به درستی پیش‌بینی نموده است. باتوجه نتایج می‌توان مشاهده نمود که دقت تست در این مرحله برابر با ۹۱/۵ درصد بوده است.



شکل ۲۸. ماتریس درهم ریختگی دقت درخت تصمیم فازی در داده های آموزش در فولد دهم

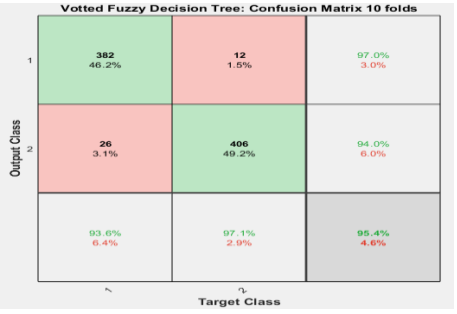
با توجه به شکل (۲۸) برای فولد دهم می‌توان متوجه شد که الگو در فولد نهم گروه ۱ را با دقت ۹۴ درصد شناسایی نموده است نتایج بدست آمده نشان می‌دهد که دقت آموزش در این مرحله برابر با ۹۳/۴ درصد است.

Downloaded from jmbr.mbri.ac.ir at 10:01 +0330 on Monday November 25th 2019



شکل ۲۹. ماتریس درهم ریختگی دقت درخت تصمیم فازی در داده های تست در فولد دهم

شکل (۲۹) ماتریس درهم‌ریختگی تست را برای فولد دهم گزارش می‌دهد که می‌توان متوجه شد که دقت تست در این فولد برابر با ۸۹/۲ درصد بوده است که ۸۵/۷ درصد در پیش‌بینی گروه ۱ و ۹۲/۷ درصد در پیش‌بینی گروه ۲ دقت داشته است. الگو در این مرحله توانسته است ۳۶ مورد را در گروه ۱ به درستی پیش‌بینی نماید و ۶ مورد را مرتکب خطا شود. در حالی‌که این نتایج برای گروه ۲ نشان می‌دهد که الگو از ۴۱ مورد ۳۸ مورد را برای گروه ۲ به درستی پیش‌بینی نموده است.



شکل ۳۰. ماتریس درهم ریختگی دقت درخت تصمیم فازی حاصل از رای گیری

شکل (۳۰) نتیجه بدست آمده از رای‌گیری بین نتایج بدست آمده از فولدهای ۱۰ گانه را نشان می‌دهد. همان طور که از ماتریس فوق قابل مشاهده است دقت درخت تصمیم فازی در این مرحله از دقت تمام فولدها بیشتر است. نتایج نشان می‌دهد که در این حالت درخت

Downloaded from jmbr.mbri.ac.ir at 10:01 +0330 on Monday November 25th 2019

تصمیم توانسته است دقتی برابر با ۹۵/۴ درصد در تعیین ریسک اعتباری مشتریان نشان دهد. نتاج ماتریس نشان می‌دهد که الگو توانسته است در شناسایی گروه ۱ دقتی برابر با ۹۷ درصد نشان دهد و از ۳۹۴ مورد ۳۸۲ نفر به درستی شناسایی نماید. هم‌چنین نتایج نشان می‌دهد دقت شناسایی گروه ۲ برابر ۹۴ درصد است و ۴۰۶ مورد را به درستی و ۲۶ مورد را همراه با خطا گزارش نموده است.

۷.۶ ارزیابی اعتبار و کارایی مدل

در این بخش شیوه ارزیابی مدل و هم‌چنین سنجش کارایی مدل به لحاظ قدرت تفکیک مشتریان در مقایسه با الگوریتم‌های مختلف مورد مطالعه تشریح شده است.

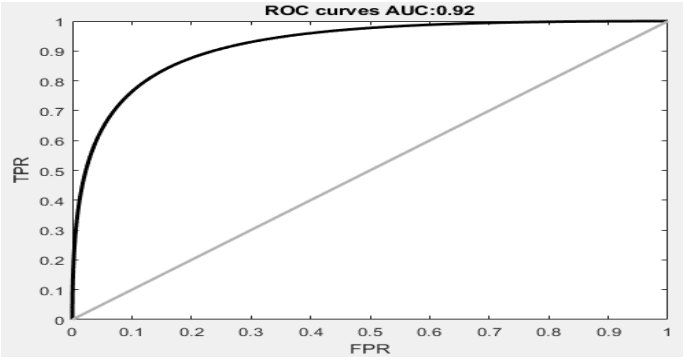
جدول ۸

مقایسه دقت درخت تصمیم فازی با الگوریتم شبکه عصبی

الگو	شبکه عصبی	درخت تصمیم فازی	درخت تصمیم فازی به روش بگینگ
نتایج آموزش%	۶۶/۴	۹۳/۸	۹۵/۴
نتایج تست %	۵۰/۶	۹۱/۶	۹۵/۴

همانگونه که نتایج جدول ۸ نشان می‌دهد بطور کلی در میان کلیه روش‌ها تکنیک درخت تصمیم فازی نتایج بسیار بهتری از سایر روش‌ها داشته است. در جدول ۸ نحوه محاسبه دقت پیش بینی به این صورت است که تعداد کل مشاهدات محاسبه می‌شود و سپس نسبت تعداد مشاهداتی که الگو توانسته است به درستی پیش بینی کند به کل مشاهدات بدست آمده است و این نسبت را درصد پیش بینی نامیده می‌شود.

Downloaded from jmbr.mbri.ac.ir at 10:01 +0330 on Monday November 25th 2019



شکل ۳۱. منحنی ROC

روش دیگر مقایسه دقت پیش‌بینی مدل‌های مورد بررسی در این مطالعه استفاده از منحنی دقت فراخوانی (Recall-Precision) می‌باشد. این منحنی که با نام منحنی ROC[1] شناخته می‌شود، امکان مقایسه تصویری مجموعه‌ای روش‌های دسته‌بندی را فراهم می‌سازد. هرچه این منحنی به سمت بالا و چپ کشیده باشد، دقت مدل بالاتر و هرچه به خط قطری نزدیکتر باشد دقت مدل اندک خواهد بود. همانگونه که تصویر ROC نشان می‌دهد[2] نتایج درخت تصمیم فازی با تکنیک بگینگ با بهتر از سایر روش‌ها بوده است. نمودار فوق نشان می‌دهد که درخت تصمیم فازی توانسته صحت ۹۲ درصدی را در گزارشات به خود اختصاص دهد.

۷ نتیجه‌گیری و پیشنهادات

به‌منظور بررسی فرضیه اصلی تحقیق مبنی بر این که روش بگینگ نسبت به سایر الگوها دقت بالاتری را دارا است یا خیر باید خاطر نشان نمود که الگوی استفاده شده در این تحقیق به‌منظور پیش‌بینی رفتار مشتریان استفاده می‌شود؛ اما باتوجه نتایج گزارش شده تحقیق می‌توان ادعا کرد که درخت تصمیم فازی از توانایی بالاتری نسب به الگوریتم شبکه عصبی در پیش‌بینی رفتار مشتریان برخوردار است و این فرضیه رد نمی شود.

همچنین فرضیه دوم تحقیق مبنی بر این که تراکنش های مالی مشتریان بانکی بیشترین توضیح دهندگی را در رتبه بندی ریسک اعتباری مشتریان دارد نیز با توجه به نتایج حاصل از

[1] Receiver Operating Characteristic

[2] TPR = true positive rate and FPR = false positive rate

Downloaded from jmbr.mbri.ac.ir at 10:01 +0330 on Monday November 25th 2019

Downloaded from jmbr.mbri.ac.ir at 10:01 +0330 on Monday November 25th 2019

تحقیق در دوره زمانی و مکانی تحقیق قابل تایید نمی‌باشد و نتایج نشان داد که درآمد مشتریان از اهمیت بالاتری برخوردار است.

به‌طور کلی نتایج مستقیم حاصل از اجرای تحقیق را می‌توان به‌صورت زیر خلاصه نمود:

درخت تصمیم فازی میزان سهم نمونه‌ها در هر طبقه را، در هر برگ مشخص می نماید و در نهایت با این دانش هر مشتری می‌تواند به هر دو طبقه متقلب و نامتقلب (با درجات متفاوت) تعلق داشته باشد. که مشتری های متقلب جهت پرداخا وام مناسب نیستند و مشتری های نا متقلب جهت پرداخت وام از نظر ریسک اعتباری مطمعن تر هستند.

استفاده از داده‌کاوی و منطق فازی در قالب مدل درخت تصمیم فازی، نتایجی به مراتب بهتر از عملکرد کارشناسان اعتبارسنجی بانک مورد مطالعه را حاصل نمود که این نشان‌دهنده کارایی بالای مدل مورد استفاده تحقیق در مقایسه با کارشناسان اعتبارسنج بانک می‌باشد که براساس تجربه و به روش قضاوتی به پیش‌بینی در خصوص نوع مشتریان می‌پردازند.

به طور خلاصه نتایج حاصل از تحقیق نشان داد که درآمد و تراکنش مالی مشتریان به ترتیب دارای بیشترین اهمیت را در تعیین ریسک اعتباری مشتریان دارد. هم‌چنین نشان داد که درخت تصمیم فازی با استفاده از تکنیک بگینگ دقت بالاتری به نسبت به الگوهای شبکه عصبی و درخت تصمیم فازی به روش معمول دارد.

فهرست منابع

تقوی‌فرد م. ت. (۱۳۹۶) دسته بندی مشتریان حقوقی و پیش بینی توانایی سوددهی آنان با استفاده از ارزش طول عمر مشتری و رویکرد زنجیره مارکوف (مورد مطالعه: مشتریان یک بانک خصوصی)، دوره ۱۵، شماره ٤٥، صفحه ٤٥ــ ٧٤.

رادپور م. ، رفیعی ا. ، رسولی‌زاده ع. ، لهراسبی ع.(۱۳۸۸) مدیریت ریسک بازار رویکرد ارزش در معرض خطر،تهران ناشر آتی نگر.

عینی، آ. (۱۳۹۷)مدیریت بهینه ریسک اعتباری، اقتصاد و بانکداری اسلامی، شماره ۲۵ صفحه۹۶-۶۷.

مهرآرا،م.، مهران‌فر، م.(۱۳۹۲)، عملکرد بانکی و عوامل کلان اقتصادی در مدیریت ریسک، *فصلنامه مدلسازی اقتصادی*، دوره۷، شماره ۲۱، صفحه ۲۱ــ۳۷.

نادعلی جلوخانی، ا.، آقاداوود، س. ر.،کرباسیان، م.،(۱۳۹۷)، الگوی سنجش تأثیر ساختار شکست ریسک بر اهداف پروژه‌های عمرانی شهرداری اصفهان با رویکرد معادلات ساختاری،*فصلنامه اقتصاد شهری*، دوره ۳، شماره ۱، صفحه ۹۷ــ۱۱۶.

Abdulrahman, U. F. I., Panford, J. K., & Hayfron-acquah, J. B. (2014). Fuzzy logic approach to credit scoring for micro finance in ghana: a case study

of KWIQPLUS money lending. International Journal of Computer Applications, 94(8).

De Wilde, P. (1997). The magnitude of the diagonal elements in neural networks. Neural Networks, 10(3), 499-504.

Eslami, A. S., Rijks, D., Fernandes, J., & Aarninkhof, S. (2014). Design and Construction Projects, Challenges and Opportunities in Management of Mega-Scale Dredging Projects, 11th International Conference on Coasts, Ports and Marine Structures (Icopmas) Ports & Maritime Organization.

Fausett, L. (1994). Fundamentals of neural networks: architectures, algorithms, and applications. Prentice-Hall, Inc..

Heffernan, S. (2003). 14. The causes of bank failures. Handbook of International Banking, 366.

Picton P. (2000). Neural Networks. 2nd ed. Palgrave, New York, NY.

Downloaded from jmbr.mbri.ac.ir at 10:01 +0330 on Monday November 25th 2019