Archive of SID

International Journal of Industrial Engineering & Production Management (2017)

September 2017, Volume 28, Number 1
pp 161- 174

http://IJIEPM.iust.ac.ir/



A Novel Hybrid Learning Approach to Train Hierarchical Mixture of Neural Networks Model for Credit Scoring in Bank Industry

Danial Dadmohamadi & Abbas Ahmadi *

Danial Dadmohamadi, *Department of Industrial Engineering & Management Systems, Amirkabir University of Technology.*

Abbas Ahmadi*, Department of Industrial Engineering & Management Systems, Amirkabir University of Technology.*

Keywords

Bank Industry, Credit Scoring (CS), Binary Particle Swarm Optimization (BPSO), Modular Neural Network, Hierarchical Mixture of Neural Networks (HMNNs).

ABSTRACT

Credit risk in bank industry is the probability of non-repayment of obligations by customers at specific time. It is one of the most important hazards for banks and private institutes. Due to huge bulk of banks' overdue receivables, establishment of a Credit Scoring (CS) system is one of the most important means of controlling such a risk. This paper uses the powerful Neural Networks in predicting and mixing them, which can classify customers in two groups of customers who pays their debts on time and customers who don't. The used model, which has modular based structure and training, is named Hierarchical Mixture of Neural Networks (HMNN). In mentioned model, for the decomposition of problem among networks and combining results to achieve the final prediction and also the method of training of it uses new approach. The purposed approach applies Binary Particle Swarm Optimization (BPSO) for dimension reduction and decomposing the problem among modules at first, then using the modulation of the training rules specific to each module and the general training rule of this network. Results are achieved in compersion with Multi-Layer Perceptron and Laterally Connected Neural Network.Based on observed results, the suggested model could predict customers' behaviour with punctuality.

© 2017 IUST Publication, IJIEPM Vol. 28, No. 1, All Rights Reserved

Downloaded from ijiepm.iust.ac.ir at 9:54 IRST on Sunday November 26th 2017

www.SID.ir

Archive of SID

نشریه بین المللی مهندسی صنایع و مدیریت تولید



شماره ۱، جلد ۲۸، بهار ۱۳۹۶
صفحه ۱۶۲- ۱۷۴

نشریه بین المللی مهندسی صنایع و مدیریت تولید
دانشگاه علم و صنعت ایران
ISSN: 2008-4870

http://IJIEPM.iust.ac.ir/

یک رویکرد ترکیبی جدید برای آموزش مدل ترکیب سلسله مراتبی شبکه‌های عصبی به منظور رتبه‌بندی اعتباری در صنعت بانکداری

دانیال دادمحمدی و عباس احمدی *

کلمات کلیدی

صنعت بانکداری،
رتبه‌بندی اعتباری،
الگوریتم گسسته بهینه‌سازی انبوه ذرات،
شبکه عصبی ماژولار،
مدل ترکیب سلسله مراتبی شبکه‌های عصبی.

چکیده:

ریسک اعتباری در صنعت بانکداری به عنوان خطر ناشی از احتمال عدم بازپرداخت تعهدات توسط مشتریان در سررسید بوده و یکی از مهم‌ترین ریسک‌ها در بانک‌ها و مؤسسات مالی به حساب می‌آید. استقرار نظام رتبه‌بندی اعتباری با توجه به حجم انبوه مطالبات معوق بانک‌ها، یکی از مهمترین ابزارهای کنترل این نوع ریسک است. این مقاله با استفاده از شبکه‌های عصبی توانمند در حوزه پیش‌بینی و ترکیب آنها قادر است مشتریان را در دو گروه خوش‌حساب و بدحساب دسته‌بندی کند. مدل پیشنهادی که دارای ساختار و آموزش ماژولار است، مدل ترکیب سلسله مراتبی شبکه‌های عصبی نام دارد. در مدل یادشده برای تجزیه مساله میان شبکه‌ها و ترکیب نتایج برای رسیدن به پیش‌بینی نهایی و همچنین شیوه آموزش آن از رویکردی جدید استفاده شده است. رویکرد پیشنهادی ابتدا، از الگوریتم گسسته بهینه‌سازی انبوه ذرات برای کاهش ابعادی و تجزیه مساله میان ماژول‌های مختلف استفاده می‌کند، سپس برای آموزش، از تلفیق قانون‌های مختص به هر ماژول و قانون آموزش کلی این شبکه استفاده می‌کند. نتایج در مقایسه با شبکه عصبی پرسپترون چندلایه و شبکه عصبی با اتصالات جانبی بدست آمده است. طبق نتایج بدست آمده مدل پیشنهادی توانسته با دقت بسیار بالا رفتار مشتریان را پیش‌بینی نماید.

۱. مقدمه

فرایند جهانی‌شدن و به تبع آن تشدید رقابت موسسات مالی و اقتصادی در بازارهای محلی و جهانی، روز به روز ضرورت تقویت و اصلاحات سیستمی را در بنگاه‌های مالی و سرمایه‌ای در سرتاسر دنیا افزایش می‌دهد. بانک‌ها نیز از این امر مستثنا نیستند [۲-۱]. ایجاد تعادل میان عرضه و تقاضا در منابع و تسهیلات بانکی، مدیریت و کاهش مطالبات معوق و رهایی از سیستم وثیقه‌محوری از جمله مسائلی است که ضرورت و نیاز پیاده‌سازی نظام رتبه‌بندی و اعتبارسنجی را در سیستم بانکی بیش از سایر مسائل نمایان می‌سازد. از انواع مدل‌ها و روش‌های مختلفی که برای اعتبارسنجی مشتریان بانک‌ها وجود دارد، می‌توان به مدل‌های تحلیل ممیزی، مدل‌های احتمالی خطی، لاجیت و پروبیت، مدل‌های مبتنی بر تئوری قیمت‌گذاری اختیار معامله و مدل‌های شبکه‌های عصبی مصنوعی اشاره کرد[۳]. از جمله مطالعات انجام شده در این زمینه می‌توان به کارهای دیکن [i] در زمینه طراحی مدل نمره‌دهی اعتباری، مطالعات آلتمن [ii] در طراحی مدل نمرهZ و مورگان [iii] در طراحی مدل اندازه‌گیری ریسک اشاره نمود[۴-۱].

روش‌های آماری، متداول‌ترین و پرکاربردترین روش‌ها برای ساختن مدل‌های رتبه‌بندی هستند [۳-۱]. در حوزه هوش مصنوعی، وست [iv] به مقایسه دقت طبقه‌بندی پنج مدل شبکه عصبی و چهار مدل آماری رگرسیون لجستیک، نزدیک ترین

تاریخ وصول: ۹۳/۰۸/۰۳
تاریخ تصویب: ۹۴/۰۲/۰۹

دانیال دادمحمدی، دانشکده مهندسی صنایع و مدیریت سیستم‌ها، دانشگاه صنعتی امیرکبیر
*نویسنده مسئول مقاله: **عباس احمدی**، دانشکده مهندسی صنایع و مدیریت سیستم‌ها، دانشگاه صنعتی امیرکبیر

Downloaded from ijiepm.iust.ac.ir at 9:54 IRST on Sunday November 26th 2017

www.SID.ir

Archive of SID

Downloaded from ijiepm.iust.ac.ir at 9:54 IRST on Sunday November 26th 2017

همسایه[v]، آنالیز ممیزی و چگالی کرنل پرداخته است. در این میان مدل‌های MLP[vi], RBF[vii] و MOE[viii] به عنوان مدل‌های برتر انتخاب شدند[۷]. در این مقاله از شبکه‌های عصبی، ترکیب آنها و الگوریتم بهینه‌سازی برای پیش‌بینی در این حوزه استفاده شده است. این مفاهیم با مفاهیم آماری بر پایه احتمالات با یکدیگر تلفیق می‌شوند و مدل ترکیب سلسله مراتبی شبکه‌های عصبی(خبرگان)[ix] را ایجاد می‌کنند. برای تجزیه مساله میان ماژول‌های مختلف، ابعاد داده‌های ورودی بوسیله فرم گسسته الگوریتم بهینه‌سازی انبوه ذرات[x] بر روی هر ماژول شکسته می‌شود و ویژگی‌های برتر از میان کل ویژگی‌ها به شبکه معرفی می‌شوند. این مقاله به لحاظ نوع استفاده از این شبکه ماژولار[xi] برای پیش‌بینی مسائل رتبه‌بندی و همچنین نحوه آموزش آن، دارای نوآوری است. با توجه به اینکه برای آموزش شبکه‌های سلسله مراتبی، اغلب از الگوریتم‌هایی نظیر ماکسیمم‌سازی امید ریاضی[xii] که محاسبات پیچیده و احتمالی دشواری دارد استفاده می‌شود، در این مقاله برای آموزش این نوع شبکه از ترکیب الگوریتم برخط[xiii] و الگوریتم‌های رایج آموزش هریک از ماژول‌ها استفاده شده است. این موضوع ضمن اینکه حجم انبوه محاسبات پیچیده را کاهش می‌دهد، باعث می‌شود تفسیر احتمالی مدل نیز بر قرار باشد. با استفاده از روش پیشنهادی، مقاله حاضر به تهیه و تدوین الگوی رتبه‌بندی مشتریان حقوقی یکی از بانک‌های دولتی خواهد پرداخت.

ساختار مقاله به این صورت است که ابتدا مروری بر پیشینه ادبیات موضوع و مطالعات گذشته بیان می‌شود. سپس، مدل HMNN و مدل پیشنهادی مطرح و معرفی می‌شود. پس از آن الگوریتم‌های بهینه‌سازی استفاده شده توضیح داده خواهند شد و در بخش ماقبل پایانی به بررسی نتایج اجرای مدل پیشنهادی و مقایسه آن با مدل‌های دیگر پرداخته می‌شود. در انتها نیز، نتایج مقاله و تحقیقات آتی بیان خواهد شد.

۲. پیشینه تحقیق

تحقیقات انجام‌شده در حوزه رتبه‌بندی اعتباری[xiv] مشتریان بانک را می‌توان در دو گروه عمده کارها و تحقیقاتی که بر پایه روش‌های آماری انجام گرفته و کارها و پژوهش‌هایی که بر پایه هوش مصنوعی و تکنیک‌های مربوط به آن انجام گرفته است، دسته‌بندی کرد. در زمینه تحقیقات بر مبنای روش‌های آماری، شروع کار رتبه‌بندی اعتباری و توسعه آن در سال‌های اولیه و ظهور آن بوسیله روش‌های آماری مختلف بوده است. این روند به لحاظ پایش ادبیاتی موضوع سهم بیشتری در حوزه این نوع مسائل را به خود اختصاص می‌دهد. همچنین این روش‌ها بدلیل حجم محاسبات بالا و وقت گیر بودن و دقت پایین تر نسبت به برخی مدل‌های شبکه عصبی، در سالیان اخیر کمتر مورد توجه می‌باشند. از این رو تمرکز پایش ادبیات موضوع در حوزه مسائل رتبه‌بندی اعتباری در این مقاله، بر تکنیک‌های هوش مصنوعی و روش‌های مرتبط با آن است.

پژوهشی با عنوان «تحلیلی از امتیازدهی اعتباری برای وام‌های کشاورزی» در تایلند صورت گرفته است که هدف پژوهشگران از انجام این مطالعه، تخمین مدل امتیازدهی اعتباری (تصمیم‌گیری وام‌دهی) برای وام‌های کشاورزی در تایلند بوده است. آنها از مدل لاجیت و دو نوع از مدل‌هایی‌که به صورت اختصار با شبکه عصبی مصنوعی، یکی با عنوان «شبکه عصبی احتمالی[xv]» و دیگری با عنوان «شبکه عصبی چند لایه بازخوردی[xvi]»، نشان داده می‌شوند، برای برآورد مدل امتیازدهی اعتباری خود استفاده کرده‌اند. نتایج حاصله در این تحقیق نشان می‌دهد که مدل بطور کلی، در پیش‌بینی PNN دیگر صحیح برای داده‌های داخل نمونه، قوی‌تر از دو از مدل بوده است[۳].

همانطور که گفته شد، روش‌های آماری، متداول‌ترین و پرکاربردترین روش‌ها برای ساختن مدل‌های رتبه‌بندی هستند[۱۹]، که در این خصوص نیز مقالات بسیاری وجود دارد. می‌توان گفت که در ابتدا آنالیز ممیزی و رگرسیون تنها مدل‌هایی بودند که در زمینه رتبه‌بندی اعتباری استفاده می‌شدند. ویگینتون[xvii] برای اولین بار از رگرسیون لجستیک در زمینه رتبه‌بندی اعتباری استفاده کرد. او مدل رگرسیون لجستیک را با آنالیز ممیزی مقایسه‌کرد و نتیجه گرفت که رگرسیون لجستیک، برتر از آنالیز ممیزی است. گرابلوسکی و تالی[xviii] از آنالیز ممیزی و مدل پروبیت برای رتبه‌بندی متقاضیان یک فروشگاه زنجیره‌ای بزرگ در آمریکا استفاده کردند[۲۰].

از زمانی که سیستم‌های هوش مصنوعی نظیر شبکه عصبی، الگوریتم ژنتیک و سیستم‌های خبره طراحی و معرفی شدند، استفاده از آنها در تحقیقات مالی و رتبه‌بندی اعتباری مرسوم گشته و به سرعت در حال گسترش است. در این زمینه می‌توان به‌کاربرد الگوریتم ژنتیک در رتبه‌بندی اعتباری توسط شین و لی[xix] اشاره کرد که صورت‌های مالی حسابرسی شده ۵۲۸ شرکت صنعتی آمریکا را برای تخمین احتمال ورشکستگی مورد استفاده قرار دادند. نتایج نشان داد که مدل، به طور متوسط در ۸۰ درصد موارد قابلیت پیش‌بینی صحیح را دارد[۱۸]. از نتایج محاسباتی توسط تکنیک‌های جدید داده‌کاوی که بوسیله تام و کیانگ[xx] انجام گرفته، شبکه‌های عصبی با دقت‌ترین تکنیک‌ها و به دنبال آن تجزیه و تحلیل جدا کننده خطی[xxi]، رگرسیون لجستیک، درختان تصمیم[xxii] و نزدیکترین همسایه بوده‌اند. در مقایسه با روش‌های دیگر آنها نتیجه گرفتند که مدل‌های

www.SID.ir

Archive of SID

Downloaded from ijiepm.iust.ac.ir at 9:54 IRST on Sunday November 26th 2017

شبکه‌های عصبی بادقت‌تر، مقاوم‌تر و قابل پذیرش‌تر هستند[۲۱].

هنلی و هند[xxiii] در یک مطالعه جامع، دقت چهار روش مختلف در رتبه‌بندی را با استفاده از داده‌های مربوط به یک شرکت فروش پستی بزرگ مقایسه‌کردند. آنها در تحقیق خود چهار روش رگرسیون خطی، رگرسیون لجستیک، نزدیک‌ترین همسایه و درخت تصمیم استفاده کردند. نتیجه تحقیق حاکی از برتری روش نزدیک‌ترین همسایه نسبت به سه روش دیگر داشت و رگرسیون لجستیک، رگرسیون خطی و درخت تصمیم از نظر دقت طبقه‌بندی در رده‌های بعدی قرار گرفتند[۹].

مالهترا[xxiv] از یک مدل شبکه عصبی پرسپترون چند لایه برای طبقه‌بندی مشتریان ۱۲ موسسه مالی در آمریکا استفاده و نتایج آن را با روش آنالیز ممیزی مقایسه کرد که مدل شبکه عصبی دقت طبقه‌بندی بالاتری نشان داد. کیم[xxv] نیز در مقاله خود از یک روش شبکه عصبی برای طبقه‌بندی متقاضیان وام استفاده کرد. در نهایت، مدل، دقت کلی ۸۴-۷۱ درصد در داده‌های آموزش و ۸۴-۶۹ درصد در داده‌های آزمایشی را به دست آورد[۱۸].

دیسای، کروک و آوراستریت[xxvi]، شبکه عصبی پرسپترون چند لایه، شبکه عصبی ترکیب خبرگان، تحلیل جداکننده خطی و رگرسون لجستیک را برای امتیازدهی متقاضیان اعتباری در صنعت اتحادیه اعتباری بررسی کردند. روش‌شناسی آنها شامل اعتبارسنجی ضربدری ۲-بخشی[xxvii] از داده‌های زمینه‌ای بود که از سه اتحادیه اعتباری با فرض هزینه‌های برابر برای ریسک‌های اعتباری خوب و بد بدست آمده‌بودند. آنها نتیجه گرفتند که مدل‌های شبکه عصبی عملکرد بهتری نسبت به تحلیل جدا کننده خطی دارند اما فقط اندکی از مدل‌های رگرسیون لجستیک بهتر هستند[۹،۷].

هوآنگ، تی ژنگ و آنگ[xxviii]، الگوریتم ژنتیک دو مرحله‌ای را برای ارتباط با مساله رتبه‌بندی اعتباری از طریق ترکیب مزایای استفاده از قوانین اگر-آنگاه[xxix] و تابع جدا کننده ارائه نمودند. بر اساس نتایجی که آنها بدست آوردند، الگوریتم ژنتیک دو مرحله‌ای دقت بهتری نسبت به مدل‌های رگرسیون لجستیک، شبکه عصبی MLP، CART[xxx] و C4.5[xxxi] داشته است[۱۱].

۳. معرفی مجموعه داده اعتباری

مجموعه داده و جامعه آماری مورد بررسی در این مقاله، مشتریان حقوقی در سطح شعب بانک سپه در شهر تهران هستند که تسهیلات اعتباری دریافت نموده‌اند. بر این اساس، مشتریانی که طبق ظوابط بانک از اهمیت بیشتری برخوردارند، انتخاب شده‌اند و برای پیش‌بینی رفتار اعتباری آنهااز نسبت‌های مالی مربوط به آنها استفاده شده است که در ادامه توضیح داده خواهد شد. با توجه به اینکه هدف این مقاله کلاس‌بندی مشتریان به دو دسته خوش‌حساب و بدحساب است، جامعه آماری از نظر وضعیت اعتباری به دو دسته مشتریان خوش‌حساب (مشتریانی که تسهیلات دریافتی خود را در سر رسید مربوطه قبل از اینکه تسهیلات سر رسید گذشته شود، تسویه می‌نمایند) و مشتریان بدحساب (مشتریانی که پس از دریافت تسهیلات، اقدام می‌نمایند) تقسیم شده است.

۳-۱. تشریح متغیرهای پژوهش

متغیر وابسته: با توجه به اینکه مشتریان اعتباری به دو دسته خوش‌حساب و بدحساب تقسیم شده‌اند، این متغیر می‌تواند دو حالت صفر و یک را به خود اختصاص دهد. بر این اساس، با توجه به وضعیتی که مشتری دارد، این متغیر یکی از مقادیر صفر و یک را خواهد داشت. اگر مشتری در بازپرداخت تسهیلات نکول کرده باشد تسهیلات سررسید گذشته یا معوق داشته باشد (t=۰) در نظر گرفته می‌شود و چنانچه بازپرداخت تسهیلات در سررسید صورت گرفته باشد، (t=۱) در نظر گرفته می‌شود.

متغیرهای مستقل: شامل آن دسته از متغیرهایی می‌شوند که بر ریسک اعتباری مشتری و به عبارتی دیگر بر متغیر وابسته مدل تأثیر گذارند. بنابراین با توجه به بررسی‌های انجام‌شده، مجموعه متغیرهایی که به نوعی می‌توانند بر متغیر وابسته مدل تأثیر گذارند به شرح زیر هستند:

۱. سرمایه در گردش (بدهی‌های جاری- دارایی‌های جاری) x_1
۲. نسبت بدهی جاری (بدهی‌های جاری به دارایی‌ها)x_2
۳. نسبت مالکانه (حقوق صاحبان سرمایه یا ارزش ویژه به جمع دارایی‌ها) x_3
۴. نسبت دارایی (دارایی جاری به کل دارایی‌ها)x_4
۵. نسبت سریع(آنی) (دارایی جاری منهای موجودی کالا به بدهی جاری)x_5
۶. نسبت گردش نقد (موجودی نقدی و سپرده‌ها به بدهی جاری)x_6
۷. موجودی نقد و سپرده‌ها به کل دارایی‌ها. x_7
۸. نسبت جاری (دارایی جاری به بدهی جاری)x_8
۹. نسبت بدهی (کل بدهی‌ها به کل دارایی‌ها)x_9
۱۰. حاشیه سود (سود خالص به فروش).x_{10}
۱۱. بازده دارایی (سود خالص به کل دارایی‌ها)x_{11}
۱۲. سود انباشته به کل دارایی‌ها. x_{12}
۱۳. نسبت بدهی به حقوق صاحبان سهام. x_{13}
۱۴. کل وام‌هایی بانکی به کل دارایی‌ها. x_{14}
۱۵. کل وام‌های بانکی به کل بدهی‌ها. x_{15}
۱۶. فروش خالص به کل دارایی‌ها. x_{16}
۱۷. موجودی کالا به فروش خالص. x_{17}
۱۸. بدهی‌های جاری به فروش خالص. x_{18}

www.SID.ir

Archive of SID

Downloaded from ijiepm.iust.ac.ir at 9:54 IRST on Sunday November 26th 2017

۳-۲. جمع‌آوری و پیش‌پردازش داده‌ها

داده‌های مربوط به مشتریان مختلف در طی سال‌های مختلف جمع‌آوری و در پایگاه داده ذخیره شدند. برای این منظور داده‌های مربوط به ۱۰۸ مشتری حقوقی مختلف در طی بازه‌های زمانی معین جمع‌آوری شدند. همانطور که گفته شد،این داده‌ها مربوط به مشتریاتی است که برای بانک از اهمیت بالاتری نسبت به سایر مشتریان برخوردارند. نمای کلی از وضعیت مجموعه داده در دسترس، در جدول زیر مشاهده می‌شود:

جدول ۱. نمای کلی از وضعیت مجموعه داده

تعداد نمونه	خوش حساب	بدحساب
۱۰۸	۸۲	۲۶

کیفیت داده‌ها در پیش‌بینی همواره یک موضوع کلیدی است. برای این منظور وبرای بهبود نتایج، پیش از اینکه داده‌ها برای آموزش شبکه‌ها به مدل ارائه شوند، پردازش‌هایی بر روی آنها انجام می‌گیرد. با توجه به اینکه داده‌ها دارای مقادیر خام و در دامنه‌های مختلف قرار دارند، برای معرفی به عنوان ورودی به شبکه عصبی، باید نرمال و تبدیل به دامنه مشخصی از اعداد بین صفر و یک یا منفی یک و یک شوند. یکی از مهمترین روش‌های استانداردسازی، تبدیل داده‌ها به مجموعه جدیدی است که میانگین آنها صفر و واریانس آنها واحد باشد برای این منظور می‌توان از رابطه زیر استفاده کرد:

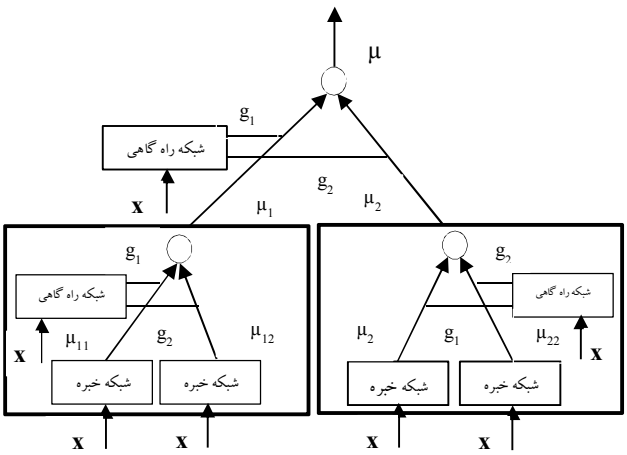
$$z=\frac{(x-\bar{x})}{s} \qquad (۲۲)$$

که در آن x مقدار اولیه (تغییر نیافته)، $\bar{x}$ میانگین داده‌ها و s انحراف معیار آنهاست. تابع توزیع نرمالی که میانگین آن صفر و انحراف معیار آن واحد باشد را تابع توزیع نرمال استاندارد می‌نامند. پس از تبدیل داده‌ها به نرمال استاندارد با توجه به اینکه داده‌ها باید در دامنه صفر و یک قرار داشته باشند، مقادیر بدست آمده به عنوان ورودی در تابع سیگموید قرار می‌گیرد تا آنچه به عنوان خروجی استخراج می‌گردد مجموعه داده‌ها با مقادیر بین صفر و یک باشد.

۴. مدل ترکیب سلسله مراتبی شبکه‌های عصبی

مدل ترکیب سلسله مراتبی شبکه‌های عصبی یکی از شبکه‌های بسیار قدرتمند در حوزه مسائل دسته‌بندی است. استفاده از مفاهیم آماری و احتمالات در آموزش این نوع شبکه باعث افزیش کارایی و دقت هر چه بیشتر آن شده است. از آنجا که هدف اصلی مقاله حاضر، دسته‌بندی مشتریان بر اساس اعتبارات آنان است، با استفاده از این شبکه که دارای ساختار و آموزش ماژولار است، دسته‌بندی مشتریان صورت می‌گیرد. شبکه‌های عصبی سنتی از جمله پرسپترون چند لایه در زمینه تخمین تابع موفق عمل کرده‌اند و در مسائل بسیاری از کلاس‌بندی الگو تا مهندسی کنترل بکار رفته‌اند. اما این در حالی است که بی‌شک کارهای ارزشمند بیشتری وجود دارد که باید در چنین ساختارهایی از جمله روش های بهبود آموزش انجام شود. بطور ایده‌ال چنین ساختارهایی از نظر آماری با انگیزه خواهند بود و دارای پارامترهایی هستند که به آسانی تفسیر می‌شوند. آنها همچنین باعث می شوند سرعت آموزش افزایش یابد، زیرا الگوریتم کاهش گرادیان که در پس انتشار [xxxi] سنتی در شبکه پرسپترون چندلایه بکار می رود، به طور معمول برای حل مسائل دنیای واقعی در زمان واقعی بسیار کند است[۲۴] یک روند معمول در چنین کارهایی، کاربرد مدل‌های جداگانه برای تخمین قسمت‌های مختلف یک مساله است. رویکرد کلی، تقسیم مساله به مجموعه‌ای از زیر مساله‌ها و تخصیص یک مجموعه از تخمین زننده‌های تابع یا خبرگان [xxxi i i] به هر زیر مساله است. رویکردهای متفاوت از روش‌های مختلف و از خروجی‌های خبرگان، برای تقسیم مساله به زیر مساله و محاسبه بهترین راه حل، استفاده می‌کنند[۲۴،۶].

ساختارHMNN شامل شبکه‌های عصبی انباشته و ماژولار، بطور سلسله مراتبی است. HMNN بر پایه مفهوم «تجزیه کن غلبه کن [xxxi v]» بنا نهاده شده است. ساختار شبکه HMNN در شکل ۱ نشان داده شده است. این ساختار یک درخت است که در آن شبکه‌های راه‌گاهی [xxxv] در نقاط غیرترمینال درخت قرار دارند. این شبکه‌ها بردار ورودی $\mathbf{x}$ را دریافت می کنند و خروجی‌های اسکالر تولید می کنند که در هر نقطه در فضای ورودی کسری از واحد هستند[۱۲].



شکل ۱. طرح کلی از شبکه HMNN

شبکه‌های خبره در قسمت برگ‌های درخت قرار می گیرند. هر شبکه خبره بردار خروجی μ_{ij} را برای هر بردار ورودی تولید می‌کند که این بردارهای خروجی تا بالای درخت ادامه

www.SID.ir

Archive of SID

Downloaded from ijiepm.iust.ac.ir at 9:54 IRST on Sunday November 26th 2017

پیدا می کنند و بوسیله خروجی‌های شبکه راه‌گاهی ترکیب می‌شود.

شبکه خبره(i,j) خروجی μ_{ij} را به عنوان تابع خطی تعمیم یافته از ورودی $\mathbf{x}$ تولید می‌کند:

$$\mu_{ij} = f(U_{ij}\mathbf{x}) \quad (۱)$$

که در آن U_{ij}، ماتریس وزن و f تابع غیرخطی پیوسته ثابت است.

برای مسائل رگرسیون، $f(\cdot)$ معمولا به عنوان تابع شناسایی و برای مسائل دسته‌بندی معمولا به عنوان تابع لجستیک [xxxvi] انتخاب می‌شود که در اینصورت خروجی‌های شبکه خبره به عنوان لگاریتم شانس موفقیت [xxxvii] تحت توزیع برنولی تفسیر می شود.

شبکه‌های راه‌گاهی، خطی تعمیم یافته [xxxviii] هستند. متغیرهای واسطه ξ_i بصورت زیر هستند:

$$\xi_i = V_i^T X \quad (۲)$$

که در معادله بالا، V_i بردار وزن است. i امین خروجی شبکه راه‌گاهی سطح بالا، تابع سافتمکس [xxxix] ξ_i است[۱۲].

$$g_i = \frac{e^{\xi_i}}{\sum_k e^{\xi_k}} \quad (۳)$$

توجه داشته باشید که g_i ها مثبت و جمعشان برای هر $\mathbf{x}$ یک است. آنها می توانند به عنوان یک تقسیم بندی نرم از فضای ورودی تفسیر شوند. به طور مشابه شبکه‌های راه‌گاهی در سطوح بالاتر هم سیستم‌های خطی تعمیم یافته هستند. ξ_{ij} بصورت زیر تعریف می شود:

$$\xi_{ij} = V_{ij}^T \quad (۴)$$

سپس

$$g_{j|i} = \frac{e^{\xi_{ij}}}{\sum_k e^{\xi_{ik}}} \quad (۵)$$

که $g_{j|i}$ ، خروجی j امین واحد در i امین شبکه راه‌گاهی در سطح دوم از ساختار شبکه است. $g_{j|i}$ ها نیز، مثبت و جمعشان برای هر $\mathbf{x}$، یک است.

بردار خروجی درهر نقطه غیرترمینال درخت، خروجی موزون شبکه‌های خبره زیر آن نقاط هستند. خروجی در i امین نقطه غیرترمینال در لایه دوم یک درخت رو سطحی برابر است با:

$$\mu_i = \sum_j g_{j|i}\mu_{ij} \quad (۶)$$

و خروجی در سطح بالایی درخت :

$$\mu = \sum_i g_i \mu_i \quad (۷)$$

توجه شود که هر دو پارامتر g و μ به ورودی $\mathbf{x}$ بستگی دارند، اگرچه خروجی نهایی یک تابع غیر خطی از ورودی است[۱۲].

مساله آموزش HMNN را می‌توان به عنوان مساله برآورد درستنمایی ماکسیمم [xl] تعبیر کرد. بر این اساس احتمال کل تولید y از $\mathbf{x}$، ترکیب احتمالات تولید y از هر یک از توزیع‌های جزئی است که در آن جزءهای ترکیبی احتمالات چند جمله‌ای هستند:

$$P(y|X,\theta) = \sum_i g_i(X,V_i) \sum_j g_{j|i}(X,V_{ij}) \quad (۸)$$
$$P(y|X,\theta_{ij})$$

در مساله دسته‌بندی باینری، خروجی y یک متغیر تصادفی گسسته است که دارای نتایج ممکن شکست و پیروزی است. بر این اساس $P(y|X,\theta_{ij})$ در رابطه بالا معمولا توزیع برنولی در نظر گرفته می شود. در این حالت میانگین μ، احتمال شرطی دسته بندی ورودی به عنوان موفقیت است. رابطه ۸ با جایگذاری توزیع برنولی بصورت زیر خواهد شد:

$$P(y|X,\theta) = \sum_i g_i \sum_j g_{j|i} \mu_{ij}^{y}(1-\mu_{ij})^{1-y} \quad (۹)$$

بنابراین لگاریتم احتمال مجموعه داده $\{[x^{(t)}, y^{(t)}]\}_1^N$ بوسیله لگاریتم گرفتن از حاصلضرب N توزیع از معادله بالا بدست می آید که لگاریتم احتمال زیر را نتیجه می دهد:

$$l(\theta;x) = \sum_t \ln \sum_i g_i^{(t)} \sum_j g_{j|i}^{(t)} P_{ij}(y^{(t)}) \quad (۱۰)$$

اکنون هدف حداقل کردن این تابع است. روش‌های مختلفی برای این کار وجود دارد که از آن جمله می توان به روش ماکزیمم سازی امید ریاضی، الگوریتم حداقل مربعات موزون [xli] و الگوریتم آنلاین(بر اساس کاهش گرادیان [xlii]) اشاره کرد. در این مقاله از روش کاهش گرادیان برای آموزش شبکه استفاده شده است. جردن و جاکوب [xliii] الگوریتم یادگیری کاهش گرادیان را برای ساختار سلسله مراتبی ارائه کردند. در این قسمت معادلات الگوریتم آنلاین ارائه می‌شود که این معادلات به‌روز رسانی پارامترهای [xliv] هر شبکه را در بر می‌گیرند.

با دیفرانسل گرفتن از $l(\theta;\mathbf{x})$ با توجه به پارامترها، قانون یادگیری کاهش گرادیان زیر برای ماتریس اوزان U_{ij} بدست خواهد آمد:

$$\Delta U_{ij} = \rho \sum_t h_i^{(t)} h_{j|i}^{(t)} (y^{(t)} - \mu^{(t)}) X^{(t)T} \quad (۱۱)$$

که در آن ρ نرخ یادگیری است. قانون یادگیری کاهش گرادیان برای i امین بردار وزن در شبکه راه‌گاهی سطح بالایی برابر است با:

$$\Delta v_i = \rho \sum_t (h_i^{(t)} - g_i^{(t)}) X^{(t)} \quad (۱۲)$$

و قانون کاهش گرادیان برای j امین بردار وزن شبکه راه‌گاهی در سطح پایین تر برابر است با:

$$\Delta v_{ij} = \rho \sum_t (h_{j|i}^{(t)} - g_{j|i}^{(t)}) X^{(t)} \quad (۱۳)$$

www.SID.ir

Archive of SID

Downloaded from ijiepm.iust.ac.ir at 9:54 IRST on Sunday November 26th 2017

که در روابط ۱۱، ۱۲ و۱۳، h_i و $h_{j|i}$ با استفاده از قانون بیز، احتمالات پسین[xl v] در گره های درخت هستند:

$$h_i=\frac{g_i \sum_j g_{j|i} P_{ij}(y)}{\sum_i g_i \sum_j g_{j|i} P_{ij}(y)} \quad (۱۴)$$

و

$$h_{j|i}=\frac{g_{j|i} P_{ij(y)}}{\sum_j g_{j|i} P_{ij}(y)} \quad (۱۵)$$

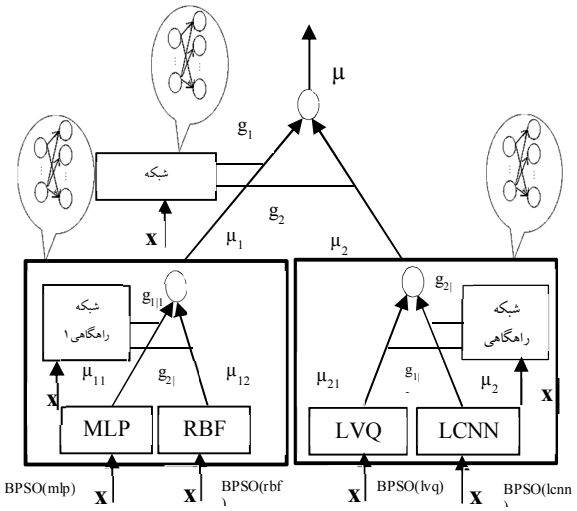
لازم به ذکر است که g_i و $g_{j|i}$ احتمالات پیشین[xl vi] هستند، زیرا بر اساس ورودی $\mathbf{x}$ بدون دانشی از خروجی هدف مربوطه y محاسبه می شوند. احتمالات پسین نیز یک بار، بر اساس ورودی و خروجی هدف که مشخص هستند، تعریف می شوند[۶].

۵. مدل پیشنهادی

این مدل همانطور که گفته شد دارای ساختار ماژولار است و از شبکه‌های عصبی که در ادامه معرفی می‌شوند تشکیل می شود. به طور کلی ساختار ماژولار یا بر روی داده‌ها شکل می‌گیرد یا بر روی ویژگی‌های مجموعه داده، بدین معنا که در یک رویکرد، داده‌ها خرد می شوند و به ماژول‌های مختلف شبکه وارد می‌شوند و در رویکردی دیگر، ویژگی‌ها شکسته می‌شوند و به ماژول‌های مختلف معرفی می‌شوند. در این مقاله از رویکرد دوم برای ساخت شبکه ماژولار استفاده شده است. سه شبکه راه‌گاهی که وظیفه تجمیع خروجی‌های هر سطح را دارند، خطی تعمیم یافته و شبکه‌های خبره، شبکه‌های عصبی رو به جلو هستند. این شبکه‌ها به شرح ذیل می‌باشند:

۱. شبکه عصبی MLP [۲۳،۲۱،۱۳]
۲. شبکه عصبی RBF [۱۰،۷]
۳. شبکه عصبی LCNN [xl vi i] [۱۷]
۴. شبکه عصبی LVQ [xl vi i i] [۲۲،۱۳]

بنابراین ساختار نهایی مدل پیشنهادی مطابق شکل زیر است:



شکل ۲. معماری مدل پیشنهادی HMNN همراه با ساختار ماژول‌های آن

نحوه قرار گرفتن شبکه‌های خبره در کنار یکدیگر بر اساس میزان قدرت و دقت آنها در برخورد با مساله پیش رو است. این فرض باعث می‌شود که نوعی تعادل در خروجی ماژول‌های سمت چپ و راست بر قرار شود. همه شبکه‌ها اعم از راه‌گاهی و خبره، به تعداد ویژگی‌ها، واحدهای ورودی دارند و تعداد واحدهای خروجی در شبکه‌های خبره، برابر تعداد کلاس‌ها است، درحالی که این رقم در شبکه‌های راه‌گاهی، به تعداد شبکه‌های خبره در هر ماژول است[۶]. با مشخص شدن ساختار شبکه‌های خبره و راه‌گاهی، فرایند معرفی داده‌ها به شبکه آغاز می‌گردد. در این مدل، ویژگی‌های مجموعه داده مورد نظر، بر مبنای روش «تقسیم نرم[xl i x]» میان چهار شبکه خبره توزیع می‌شوند. برای این منظور، همانطور که در شکل مشخص است، ابتدا توسط الگوریتم BPSO برای هر شبکه، فرایند کاهش ویژگی اجرا و ویژگی‌های بدست امده، به HMNN معرفی می‌شوند. در واقع برای هر یک از ماژول‌ها الگوریتم گسسته بهینه سازی انبوه ذرات بطور جداگانه بکار گرفته می‌شود و مجموعه ویژگی‌ها متناسب با عملکرد الگوریتم بهینه‌سازی، میان شبکه‌ها تقسیم می‌شوند.

۵-۱. آموزش ترکیبی-گروهی شبکه

بردار ورودی x، همانطور که در شکل ۲ مشخص است، برای تمامی شبکه‌های راه‌گاهی و خبره یکسان است، اما خروجی‌های شبکه‌های راه‌گاهی به عنوان وزن‌های شبکه‌های خبره در نظر گرفته می‌شوند. در واقع وظیفه شبکه‌های راه‌گاهی، یادگیری ترکیب موزون مناسب شبکه‌های خبره برای هر بردار ورودی است[۱۲]. آموزش برای تمامی داده‌ها و بر اساس الگوریتم اعتبار سنجی ضربدری ۵- بخشی انجام می‌شود. شبکه‌های

www.SID.ir

Archive of SID

Downloaded from ijiepm.iust.ac.ir at 9:54 IRST on Sunday November 26th 2017

خبره و راه‌گاهی بطور همزمان آموزش می‌بینند. روند آموزش بر اساس رویکردی ترکیبی انجام می‌گیرد بدین صورت که هر یک از شبکه‌های خبره بر اساس روش‌های مرسوم و مربوط به هر شبکه خبره، و شبکه‌های راه‌گاهی طبق الگوریتم آموزش برخط شبکه HMNN آموزش می‌بینند و در نهایت با تجمیع دو رویکرد آموزش، یادگیری مدل پیشنهادی انجام می‌گیرد و پارامترهای مربوط به هر شبکه بدست می‌آید. در واقع شبکه‌های راه‌گاهی، همزمان با یادگیری هر یک از خبره‌ها، یاد می‌گیرند که به هر خبره چه وزنی اختصاص داده شود تا نرخ خطا کاهش پیدا کند. تابع خطایی هم که در فرایند آموزش حداقل می‌گردد، تابع میانگین مربعات خطا[I] است.

الگوریتم ۱: الگوریتم ترکیبی یادگیری مدل HMNN
1. بردار ورودی x را به ازای هر داده و بر اساس تعداد ویژگی‌های بدست آمده مربوط به آن به شبکه‌های خبره و راه‌گاهی ارائه کنید.
2. خروجی شبکه‌های خبره و راه‌گاهی ۱ و ۲ را بر اساس قوانین مربوط به هر شبکه بدست آورید.
3. مقادیر μ_1 و μ_2 بر اساس رابطه ۶ محاسبه کنید.
4. خروجی شبکه راه‌گاهی صفر را بر اساس قوانین مربوطه محاسبه کنید.
5. خروجی نهایی شبکه را طبق رابطه ۷ محاسبه کنید.
6. مقدار خطا را بر اساس خروجی نهایی و تابع MSE محاسبه کنید.
7. برای هر داده مقدار احتمالات پسین را با استفاده از مقادیر پارامتهای جاری محاسبه کنید.
8. اوزان شبکه های راه‌گاهی را با استفاده از روابط ۱۲ و ۱۳ به روز رسانی کنید.
9. اوزان شبکه‌های خبره با استفاده قوانین یادگیری هر شبکه به روز رسانی کنید.
10. گام های ۱ تا ۹ برای سایر نمونه ها تکرار کنید.
11. چک کنید اگر خطای تجمعی از مقدار معینی کمتر شده است، شبکه آموزش دیده در غیر اینصورت کل فرایند آموزش برای یک دوره آموزشی دیگر تکرار کنید.

۵-۲. الگوریتم بهینه‌سازی انبوه ذرات[II]

این الگوریتم که از نوع الگوریتم‌های تکاملی می‌باشد، در سال ۱۹۹۵ توسط کندی و ابرهارت[III] مطرح شد و از رفتار هجومی و دسته جمعی حیوانات مانند پرندگان و ماهی‌ها، اقتباس شده است.عملکرد یک الگوریتم PSO به این صورت است که دسته‌ای از ذرات (به عنوان متغیرهای مسئله بهینه‌سازی) در محیط جستجو پخش می‌شوند. واضح است که بعضی از ذرات، موقعیت بهتری نسبت به ذرات دیگر خواهند داشت. در نتیجه بر طبق رفتار ذرات هجومی بقیه ذرات سعی می‌کنند موقعیت خود را به موقعیت ذرات برتر برسانند، در عین حال که موقعیت ذرات برتر نیز در حال تغییر می‌باشد. در واقع هر ذره از برتری یا عدم برتری خود نسبت به ذرات همسایه و همچنین نسبت به کل گروه آگاه است به همین دلیل در صورتی که اشتباهی رخ دهد و تصمیم بدی گرفته شود، به زودی آن را جبران می‌کنند. به این ترتیب ذرات می توانند به راحتی محدوده اطراف خود را جستجو کنند، بدون آنکه نگران خراب‌تر شدن نتیجه باشند[۲،۸،۱۴].

اکنون فرض کنید ذره j ام دارای بعد g باشد که به‌صورت زیر بیان شود:

$$X_j=[X_{j,1}\ X_{j,2}.....X_{j,g}] \quad (۱۶)$$

و هر ذره دارای یک Pbest و تمام ذرات دارای یک Gbest به صورت زیر می‌باشند:

$$Pbest_j = [Pbest_{j,1}\ Pbest_{j,2} ... Pbest_{j,g}] \quad (۱۷)$$

آنگاه تغییر موقعیت ذره بر اساس مقدار لغزش نیز به‌صورت زیر می‌باشد:

$$v_{j,g}^{(t+1)}=wv_{j,g}+c_1.Rand()\left(Pbest_{j,g}-x_{j,g}^{(t)}\right)c_2.rand()\left(Gbest_{j,g}-x_{j,g}^{(t)}\right),\ v_{min}\le v_{j,g}^{(t)}\le v_{max} \quad (۱۸)$$

$$x_{j,g}^{(t+1)}=x_{j,g}^{(t)}+v_{j,g}^{(t+1)};\ j=1,2,...,n,\ g=1,2,...,m \quad (۱۹)$$

در روابط بالا مقدار x بیانگر موقعیت ذره، n تعداد ذرات گروه و m تعداد اعضای تشکیل دهنده ذره و توابع Rand() و rand() تولید کننده یک مقدار تصادفی بین صفر و یک، Pbest بهترین موقعیت هر ذره، Gbest بهترین موقعیت ذرات در طول اجرای الگوریتم، w ضریب لختی، v پارامتر لغزش و سرعت ذرات، c_1 و c_2 پارامترهای شتاب هستند.مقدار Vmax معمولاً بین ۱۰ تا ۲۰ درصد محدوده متغیرها انتخاب می‌شود [۵].

www.SID.ir

Archive of SID

Downloaded from ijiepm.iust.ac.ir at 9:54 IRST on Sunday November 26th 2017

۵-۲-۱. الگوریتم گسسته بهینه‌سازی انبوه ذرات

برای فضای جستجوی گسسته باینری، کندی و ابرهارت [۱۴] PSO را به جستجو در فضای باینری با بکارگیری تبدیل سیگموید برای سرعت ذره به منظور تبدیل سرعت‌ها به بازه [۰،۱] و اجبار موقعیت‌های ذرات برای اختیارکردن مقادیر صفر و یک تطبیق داده‌اند[۱۵]. در رشته انتخابی، بردار موقعیت های ذره توسط رشته‌ای از مقادیر باینری بطول n مشخص می‌شود که n تعداد کل ویژگی هاست. هر بردار یک زیرمجموعه از رشته را نشان می‌دهد. اگر مقدار بیت(مقادیر صفر و یک) یک باشد، به این معناست که رشته مربوطه انتخاب شده است، در غیر اینصورت رشته مربوطه انتخاب نشده است. هر ذره در جامعه موقعیت و سرعت خود را بر اساس روابط زیر به روز رسانی می کند[۱۶]. بنابراین رابطه به‌روز رسانی موقعیت در PSO پیوسته با روابط ادامه جایگزین می‌شود:

$$\text{sigmoid}(v_{id}^{k}) = \frac{1}{1+e^{-v_{id}^{k}}} \qquad (۲۰)$$

$$x_{id}^{k} = \begin{cases} 1, & \text{if rand} < \text{sigmoid}(v_{id}^{k}) \\ 0, & \text{otherwise} \end{cases} \qquad (۲۱)$$

لازم به ذکر است که پارامترهای روابط مذکور و سایر مراحل الگوریتم، شبیه آن چیزی است که در PSO پیوسته توضیح داده شد. تحلیل‌گران عموما مجموعه‌ای از متغیرهای مستقل پیشنهادی برای شرکت دادن در مدل دارند که احتمال زیاد شامل همه متغیرهای مهم است، اما مطمئن نیستند که همه این متغیرها به منظور مدل‌سازی مناسب، در مورد متغیر وابسته لازم باشند. در چنین وضعیتی تمایل به غربال کردن متغیرهای پیشنهادی به منظور بدست آوردن مدلی داریم که «بهترین» زیر مجموعه از متغیرهای مستقل را در بر داشته باشد. روشی که در این مقاله از آن برای کاهش و انتخاب ویژگی‌ها استفاده شده، الگوریتم باینری بهینه‌سازی انبوه ذرات است.

۶. اجرای مدل HMNN

پس از ساخت مدل، در این بخش، مدل مورد نظر بر روی مجموعه داده موجود اجرا و نتایج آن با مدل شبکه عصبی پرسپترون چند لایه و شبکه عصبی با اتصالات جانبی مقایسه می‌شود. نتایج مدل‌ها بر اساس کارایی و دقت پیش‌بینی رتبه‌بندی می‌شوند و در نهایت مدل برتر انتخاب می‌گردد. با توجه به ماهیت ماژولار بودن مدل، ویژگی‌های منتخب برای هر ماژول شبکه عصبی در گام اول بدست می‌آید. این نتایج به تفکیک هر شبکه به صورت زیر است. ویژگی‌های حذف شده با علامت × در جدول ۲ مشخص شده‌اند.

جدول۲. ویژگی‌های انتخاب شده در هر ماژول شبکه عصبی

	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9	x_{10}	x_{11}	x_{12}	x_{13}	x_{14}	x_{15}	x_{16}	x_{17}	x_{18}
MLP		×									×	×	×	×	×		×	×
RBF										×	×		×	×	×	×		
LCNN	×			×		×		×	×					×	×			
LVQ								×	×	×	×			×		×		

در گام بعدی و قبل از ارائه نتایج نهایی، احتمالات پیشین بدست آمده که در واقع میزان تاثیرگذاری هریک از شبکه‌ها یا احتمال انتخاب و تاثیر آنها در فرایند آموزش است و همچنین مقادیراحتمالات پسین که مهمترین پارامترهای یادگیری و همگرایی شبکه هستند، در جدول ۳ نشان داده می‌شوند[۶،۱۲].

بنابراین با توجه به جدول ۳:

- ✓ مقادیر احتمالات پیشین، در واقع همان خروجی‌های شبکه‌های راه‌گاهی هستند.
- ✓ علت تعبیر احتمالات g_i و $g_{j|i}$ به عنون احتمال پیشین، این است که آنها فقط بر اساس ورودی **x** بدون دانشی از مقدار هدف مربوط به خروجی y محاسبه می شوند[۶].
- ✓ اعداد مربوط به احتمالات پیشین نشان می‌دهند که کدام شبکه بیشترین سهم را در یادگیری ماژولار شبکه HMNN داشته است.
- ✓ مقادیر احتمالات پسین در گره‌های درخت به ترتیب از چپ به راست مربوط به شبکه راه‌گاهی صفر، شبکه راه‌گاهی یک و دو هستند.

www.SID.ir

Archive of SID

جدول ۳. مقادیر احتمالات پیشین و پسین مدل پیشنهادی HMNN

احتمالات پیشین						احتمالات پسین			
MLP	RBF	LCNN	LVQ			$h_{j\|i}$		h_i	
۰/۹۲۴۳	۰/۰۷۵۷	۰/۰۲۵۰	۰/۹۷۵۰	۰	۱	۰/۰۲۵۱	۰/۹۷۴	۰/۰۰۶۶	۰/۹۹۳۶

با توجه به حصول ویژگی‌های هر شبکه، داده‌ها متناسب با ویژگی‌های منتخب، به شبکه معرفی می‌شوند. در ادامه و با آموزش شبکه ماژولار HMNN، نتایج پیش‌بینی این مدل و دقت آن ارائه در جدول ۴ ارائه شده است:

جدول ۴. نتایج اجرای مدل پیشنهادی

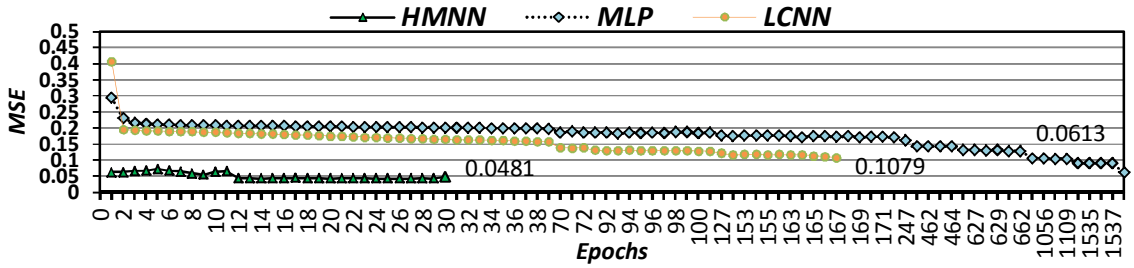
دقت پیش بینی آموزش	نرخ خطا اموزش	دقت پیش بینی تست	نرخ خطا تست
٪۹۵/۱۲۲۰	۰/۰۴۸۱	٪۸۴/۶۱۵۴	۰/۱۳۱۹

۶-۱. نتایج اجرای مدل MLP و LCNN

جدول ۵ نتایج اجرای مدل شبکه عصبی پرسپترون چند لایه و مدل شبکه عصبی با اتصالات جانبی با ۸ نرون در لایه پنهان را نشان می‌دهد. در این جدول مقادیر حاصل از پیش‌بینی نهایی شبکه‌ها بر اساس دو مجموعه داده آموزش و تست درج شده است.

جدول ۵. نتایج اجرای مدل MLP و LCNN

نام مدل	دقت پیش‌بینی آموزش	دقت پیش‌بینی تست	نرخ خطاآموزش	نرخ خطاتست
MLP	٪۹۳/۶۵۳۹	٪۷۳/۸۶۳۲	۰/۰۶۱۳	۰/۲۲۳۸
LCNN	٪۸۶/۷۶۳۱	٪۷۲/۳۳۵۴	۰/۱۰۷۹	۰/۱۹۵۲



شکل ۳. نمودار مقایسه نرخ خطا و روند همگرایی سه مدل

کارایی مدل پیشنهادی و دو مدل دیگر بر اساس نرخ خطای آموزش و همچنین تعداد دوره های آموزش در شکل ۳ و به تفکیک هر مدل ارائه شده است. روند نزولی این نمودار نشان می دهد که شبکه HMNN در هر دو معیار نسبت به سایر شبکه ها برتری محسوسی دارد و توانسته پیش بینی را تنها در طول ۳۰ دوره آموزش انجام دهد.

۶-۲. مقایسه کارایی سه مدل در پیش‌بینی

برای مقایسه مدل‌های مورد بررسی بر اساس صحت و کارایی مدل از نمودار راک [iii] استفاده شده است.این منحنی یک دیاگرام دوبعدی است که در هر بُعد آن، یک مفهوم نشان داده شده است. دیاگرام ارتباط متقابل بین نرخ TPR [iv] و نرخ FPR [v] را نشان می‌دهد. برای مقایسه چند دسته‌بندی‌کننده،سطح زیر منحنی ارزیابی می‌شود و اگر برای یک دسته‌بندی‌کننده بزرگتر از دیگری باشد(به عدد ۱ نزدیک‌تر باشد)، عملکرد دسته‌بندی آن دسته‌بندی‌کننده بهتر خواهد بود.

بنابراین، همانطور که توضیح داده شد، شکل ۴ و ۵ سطح زیر منحنی راک رابرای سه مدل موجود و با توجه به دو مجموعه آموزش و تست نشان می دهد. طبق این نمودارها سطح زیر منحنی‌های راک برای مجموعه آموزش مدل‌های HMNN، LCNN و MLP به ترتیب ۰/۹۹۹۳، ۰/۸۶۸۰ و ۰/۹۱۰۵ و برای مجموعه تست به ترتیب ۰/۶۳۱۶، ۰/۵۹۲۱ و ۰/۶۲۲۲ است که در جدول ۹ ارائه شده است.

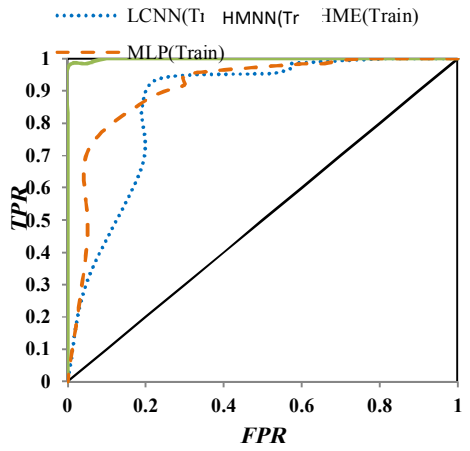
جدول ۶. مساحت زیر منحنی راک در سه مدل بررسی شده

مدل	مساحت زیر منحنی راک	
	آموزش	تست
LCNN	۰/۸۶۸۰	۰/۵۹۲۱
HMNN	۰/۹۹۹۳	۰/۶۳۱۶
MLP	۰/۹۱۰۵	۰/۶۲۲۲

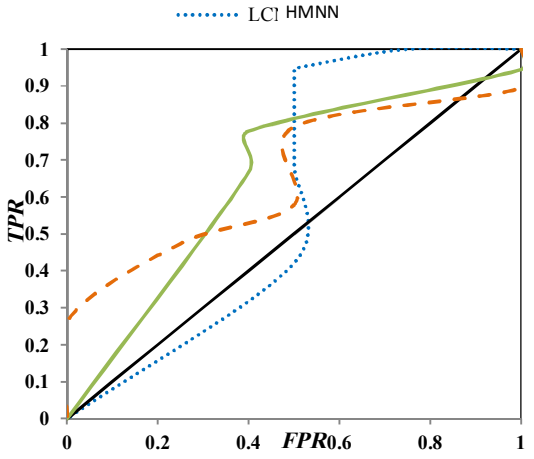
Downloaded from ijiepm.iust.ac.ir at 9:54 IRST on Sunday November 26th 2017
www.SID.ir

Archive of SID

Downloaded from ijiepm.iust.ac.ir at 9:54 IRST on Sunday November 26th 2017

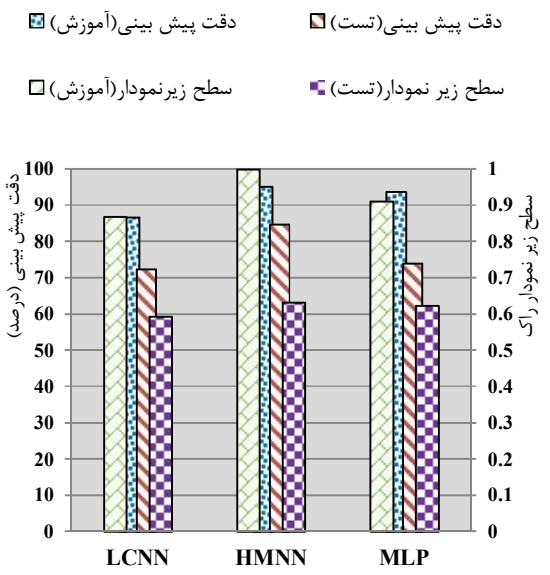


شکل ۳. منحنی ROC مجموعه آموزش شبکه‌های عصبی طراحی‌شده



شکل ۴. منحنی ROC مجموعه تست شبکه‌های عصبی طراحی‌شده

پس از اجرا، بر اساس دقت، کارایی و نرخ خطای هریک از مدل‌ها، به رتبه‌بندی آنها می‌پردازیم تا از میان مدل‌های موجود، مدل برتر انتخاب و معرفی گردد. برای مقایسه دقیق تر کارایی سه مدل از منحنی های راک و سطح زیر آنها استفاده کردیم. بر این اساس،مدل HMNN بهترین عملکرد را در میان سه مدل به خود اختصاص داده است.



شکل ۵. نمودار مقایسه نهایی مدل‌ها با یکدیگر

به منظور ارائه یک دید کلی بر اساس معیارهای مقایسه و انتخاب مدل نهایی و برتر، با استناد به مقادیر سطح زیر نمودار منحنی راک برای هر دو مجموعه آموزش و تست و همچنین مقادیر صحت کلی هر شبکه،شکل ۶ ارائه شده است که نمایانگر برتری مدل پیشنهادی در هر ۴ معیار مقایسه است. روند این انتخاب در شکل بالا مشهود است.

۷. نتیجه گیری و تحقیقات آتی

در این مقاله مدلی جدید برای رتبه‌بندی اعتباری مشتریان بانک به کار گرفته شد. مدل پیشنهادی از یک طرف به دلیل کارایی بالا و از طرفی دیگر بدلیل نرخ خطای بسیار کم نسبت به سایر مدل‌های در این حوزه برجستگی در خور توجهی دارد. برای بررسی کارایی مدل‌ها در تفکیک مشتریان به خوش‌حساب و بدحساب، از نمودار مقایسه نرخ خطای هر شبکه استفاده شد. علاوه بر این از تحلیل منحنی راک نیز سود برده شد. نتایج، مبین این است که مدل شبکه عصبی HMNN نسبت به دو مدل دیگر در این حوزه کاراتر است. این امر با توجه به حساسیت موضوع رتبه‌بندی مشتریان، برای بانک‌ها از اهمیت بسزایی برخوردار است چرا که اشتباه در این سیستم باعث بروز صدماتی جبران ناپذیر می شود که نه تنها می تواند اعتبار بانک را خدشه دار کند بلکه باعث سو استفاده های مالی عظیمی در این مسیر گردد. بدیهی است مدل برتر و منتخب در این تحقیق به عنوان یک پیشنهاد دهنده به بانک عمل می‌کند و تصمیم نهایی بر عهده مسئولین و ارزیابان بانک است. در این مقاله توانستیم با اتخاذ رویکردهایی از جمله ترکیب شبکه‌ها، تغییر در

www.SID.ir

Archive of SID

Downloaded from ijiepm.iust.ac.ir at 9:54 IRST on Sunday November 26th 2017

ساختار و یادگیری آنها و همچنین استفاده از الگوریتم های فرابتکاری، نتایج خوبی بدست آوریم که در بهترین حالت، مدل شبکه عصبی سلسله مراتبی توانست پیش‌بینی را با دقتی بیش از ۹۵٪ بهمراه خطایی کمتر از ۰/۰۶۸ انجام دهد. باتوجه به یافته‌های تحقیق که در بالا بیان شد، می‌توان نتیجه گرفت که رفتار اعتباری مشتریان تسهیلات از طریق مدل رتبه‌بندی اعتباری مشتریان قابل اجرا است. یعنی با کمک این روش می‌توان رفتار مشتریان رادر قالب مشتریان خوش‌حساب و بدحساب رتبه‌بندی کرد، که این کار سبب کاهش ریسک اعتباری و باعث افزایش سودآوری برای بانک و جذب مشتریان خوب می‌شود.

این مقاله می‌تواند علاوه بر مساله رتبه‌بندی اعتباری مشتریان بانک‌ها و موسسات مالی، رتبه بندی اعتباری متقاضیان دسته چک را که قوانین جدید مربوط به صدور آن به تازگی وضع گردیده و مساله مبتلا به بانک‌ها و موسسات مالی می‌باشد را نیز در برگیرد.همچنین می توان از روش های موجود در آن برای طبقه‌بندی مشتریان در ۴ کلاس مشتری خوب، سر رسید گذشته، معوق و مشکوک الوصول استفاده کرد. لازم به ذکر است که تحلیل کیفی شاخص‌های داده‌های ورودی و در نظر گرفتن آنها با توجه به وضعیت‌های مختلف می‌تواند باعث ایجاد داده‌هایی با ابعاد کمتر یا دقیق‌تر گردد.

پی نوشت

1. Deakin
2. Altman
3. Morgan
4. West
5. K-NN
6. Multi-Layer Perceptron(MLP) Neural Network
7. Radial Basis Function (RBF) Neural Network
8. Mixture Of Experts(MOE)
9. Hierarchical Mixture of Neural Networks(HMNNs) or Hierarchical Mixture of Experts(HME)
10. Binary Particle Swarm Optimization(BPSO)
11. Modular Network
12. Expectation-Maximization(EM)
13. On Line
14. Credit Scoring
15. Probabilistic Neural Network(PNN)
16. Multi-Layer Feedforward Neural Network
17. Winginton
18. Grabllowsky & Talley
19. Shin & Lee
20. Tam and Kiang
21. Linear Discriminant Analysis
22. Decision trees
23. Henley & Hand
24. Malhotra
25. Kim
26. Desai, Crook & Overstreet
27. 2-Fold Cross Validation
28. Jih-Jeng Huang, Gwo-Hshiung Tzeng & Chorng-Shyong Ong
29. If-Then Rules
30. Classification and Regression Tree(CART)
31. C4.5 is an algorithm used to generate a decision tree.
32. Back Propagation

۳۳. منظور از خبره هریک از ماژول‌های مدل است که خود یک شبکه عصبی است.

34. Divide & Conquer
35. Gating networks
36. Logistic Function
37. Log odds of "Success"
38. Generalized linear
39. Softmax Function
40. Maximum likelihood estimation
41. A least-squares algorithm
42. Gradient descent
43. Jordan & Jacobs

۴۴. توجه شود که واژه «پارامتر» برای متغیرهایی که به طور معمول در ادبیات شبکه های عصبی «وزن» نامیده می شوند، استفاده شده است.

۴۵. احتمال پسین دانشی است که بر اساس تجربیات پیشین برای رویدادهای غیرقابل تکرار محاسبه شده است.

۴۶. احتمال پیشین دانشی است که برای رویدادهای قابل تکرار قابل محاسبه است.

47. Laterally Connected Neural Network (LCNN)
48. Learning Vector Quantizer(LVQ) Neural Network
49. Soft Split
50. MSE
51. Particle Swarm Optimization(PSO)
52. Kennedy & Eberhart
53. ROC Curve
54. True Positive Rate
55. False Positive Rate

مراجع

[۱] اخباری، مهدیه، مخاطب رفیعی، فریماه. «ارائه یک الگوریتم فراابتکاری چندهدفه فازی سیمپلکس-ژنتیک برای پیش‌بینی رتبه اعتباری بانک‌ها»، نشریه بین المللی مهندسی صنایع و مدیریت تولید، (۱۳۹۱)، جلد ۲۳، شماره ۱، صفحه ۹۱-۸۰.

www.SID.ir

Archive of SID

[۲] بهروزی، مهدی، عشقی، کوروش. «بکارگیری الگوریتم ترکیبی بهینه‌سازی دسته ذرات برای حل مساله سنتی زمانبندی کار کارگاهی»، نشریه بین المللی مهندسی صنایع و مدیریت تولید، (۱۳۸۸)، جلد ۲۰، شماره ۲.

[۳] تقوی، مهدی، لطفی، علی اصغر و سهرابی، عبدالرضا. «مدل ریسک اعتباری و رتبه بندی مشتریان حقوقی بانک کشاورزی»، مجله پژوهشنامه اقتصادی، (۱۳۸۹)، صفحات ۹۹-۱۲۸.

[۴] فردحریری، علیرضا. «مدل سازی ریسک و رتبه‌بندی اعتباری مشتریان حقوقی بانک رفاه»، پایان نامه کارشناسی ارشد علوم اقتصادی، دانشکده اقتصاد دانشگاه تهران، (۱۳۸۷).

[۵] کلامی هریس، س. م، پریز، ن. «تاثیر میزان جمعیت و ضرایب یادگیری در عملکرد الگوریتم PSO»، سومین کنفرانس فناوری اطلاعات و دانش، دانشگاه فردوسی مشهد، (۱۳۸۶).

[6] Brown, G., "Ensemble Learning. Encyclopedia of Machine Learning", Chapter No: 00400, (2010).

[7] Chen, S. C.,"Orthogonal Least Squares Learning Algorithm for Radial Basis Function Networks", IEEE Trans on Neural Networks. (1991), Vol. 2, , pp. 302-309.

[8] Chuang, L. Y., Hsiao, C. J., Yang, C. H., "An Improved Binary Particle Swarm Optimization with Complementary Distribution Strategy for Feature Selection", International Conference on Machine Learning and Computing IPCSIT, (2011), Vol.3.

[9] Desai, V. S., Crook, J. N., & Overstreet, G.A., "Acomparison of neural networks and linear scoring models in the credit union environment", European Journal of Operational Research, (2003), Vol. 95, No. 1, pp. 24–37.

[10] Golbabai, A., Mammadov, M., Seifollahi, S., "A new strategy for training RBF network with applications to nonlinear integral equations", IUST International Journal of Engineering Science, (2008), Vol. 19, No. 5-2, pp. 1–7.

[11] Huang, J. J., Tzeng, G. H., Ong, C. S., "Two-stage genetic programming (2SGP) for the credit scoring model", Applied Mathematics and Computation, (2006), Vol. 174, pp. 1039–1053.

[12] Jordan, M. I., Jacobs, R. A., "Hierarchical Mixtures of Experts and the EM Algorithm", Neural Computation, (1994), Vol. 6, No. 2, pp. 181–214.

[13] Karray, F. O., "Soft computing and intelligent systems design", Harlow, England: Pearson Education Limited, (2004).

[14] Kennedy, J., Eberhart, R. C. and Yuhui, Sh., Swarm Intelligence. Academic Press, (2001).

[15] Kennedy, J. and Eberhart, R. C., "A Discrete Binary Version of the Particle Swarm Algorithm", Proc. of the conference on Systems, Man, and Cybernetics SMC97, (1997), pp. 4104-4109.

[16] Khalil, T. M., Hosam K.M. Y., Abdel Aziz, M. M., "Optimal Capacitor Placement on Radial Distribution Feeders in Presence of Nonlinear Loads Using Binary Particle Swarm Optimizatuion", 19th International Conference on Electricity Distribution, (2006), No. 180.

[17] Kothari, R., Agyepong, K., "On Lateral Connections in Feed-Forward Neural Networks", Neural Networks, IEEE International Conference, (1996), pp.13 – 18.

[18] Malhotra, R., & Malhotra, D. K., "Evaluating consumer loans using neural networks", Omega, (2005), Vol. 31, No. 2, pp. 83–96.

[19] Nan-chen, H., "Hybrid mining approach in the design of credit scoring Models". Expert Systems with applications, (2005), Vol. 5, pp. 655-665.

[20] Ong.C, Huang, .J, Tzng .G, "Building credit scoring models using genetic programming", Expert Systems with applications (2005), Vol. 29, No. 5, pp. 41-47.

Downloaded from ijiepm.iust.ac.ir at 9:54 IRST on Sunday November 26th 2017

www.SID.ir

Archive of SID

[21] Oreski, S., Oreski, D., Oreski. G., "Hybrid system with genetic algorithm and artificial neural networks and its application to retail credit risk assessmentExpert", Expert Systems with Applications, (2012).

[22] Qiu-Hua, T., Bao-Hua, L., Yong-Qi, C., Xing-Hua, Z., Ji-Sheng, D., "Application of LVQ neural network combined with the genetic algorithm in acoustic seafloor classification", Chinese Journal of Geophysics, (2007), Vol.50, No. 1, pp. 291-298.

[23] Valdovinos, R.M., S´anchez, J.S., "Ensembles of Multilayer Perceptron and Modular Neural Networks for Fast and Accurate Learning", Proceedings of the Fifth Mexican International Conference on Artificial Intelligence (MICAI'06), 2006, pp. 229 – 236.

[24] Waterhouse, S.R., Robinson, A.J., "Classification using hierarchical mixture of experts", In Proc. IEEE Workshop on Neural Networks for Signal Processing IV, (1994), pp. 177-186.

Downloaded from ijiepm.iust.ac.ir at 9:54 IRST on Sunday November 26th 2017

www.SID.ir