

## بررسی عملکرد روش های آماری، یادگیری ماشین و یادگیری عمیق در پیش بینی و مدیریت ریسک اعتباری

ابوالفضل روحی

کارشناسی ارشد مهندسی صنایع گرایش مهندسی مالی، دانشگاه تربیت مدرس

abolfazl.financial@gmail.com

### چکیده

ارزیابی ریسک اعتباری هسته اصلی اقتصادهای مدرن است. به طور سنتی، با روش های آماری و حسابرسی دستی اندازه گیری می شود. پیشرفت های اخیر در هوش مصنوعی مالی ناشی از موج جدیدی از مدل های ریسک اعتباری مبتنی بر یادگیری ماشین (ML) است که توجه فوق العاده ای را هم از سوی صنعت و هم از سوی دانشگاه ها به خود جلب کرد. در این مقاله، به طور سیستماتیک مجموعه ای از پژوهش های صورت گرفته را که در بازه های زمانی متفاوت با استفاده از تکنیک های آماری، یادگیری ماشین و یادگیری عمیق برای رسیدگی به مشکلات ریسک اعتباری مرور می شود. به طور خاص، یک روش طبقه بندی جدید برای الگوریتم های ریسک اعتباری مبتنی بر ML و رتبه بندی عملکرد آن ها با استفاده از مجموعه داده های عمومی پیشنهاد می گردد. در این پژوهش بیشتر در مورد چالش ها از جمله عدم تعادل داده ها، ناسازگاری مجموعه داده ها، شفافیت مدل و استفاده ناکافی از مدل های یادگیری عمیق بحث می گردد. نتایج بررسی نشان می دهد که بیشتر مدل های یادگیری عمیق از یادگیری ماشین کلاسیک و الگوریتم های آماری در برآورد ریسک اعتباری بهتر عمل می کنند و همچنین روش های ترکیبی دقت بالاتری را در مقایسه با مدل های تک ارائه می کنند. در نهایت، جداول خلاصه از نظر مجموعه داده ها و مدل های پیشنهادی را ارائه می شود.

واژگان کلیدی: ریسک اعتباری، یادگیری ماشین، یادگیری عمیق، پیش بینی ریسک



## مقدمه

پیشرفت یادگیری ماشینی<sup>۱</sup> در دهه های گذشته به شدت بر صنعت و دانشگاه تأثیر گذاشته و در نهایت زندگی روزمره مردم را متحول کرده است. هوش مصنوعی<sup>۲</sup> (AI) تقریباً برای هر فعالیت انسانی، از جمله تشخیص الگو، طبقه بندی تصویر، تجارت، کشاورزی، حمل و نقل و امور مالی استفاده شده است. (بکیو، لسمن، ۲۰۱۷)

این پژوهش بر یادگیری ماشینی که برای برآورد ریسک مالی و اعتباری اعمال می شود تمرکز دارد. سیستم های مالی مدرن بر اعتبار و اعتماد متکی هستند. ریسک اعتباری یک پارامتر اساسی است که احتمال نکول یک بدهکار را اندازه گیری و پیش بینی می کند. برآورد صحیح ریسک اعتباری برای کل سیستم بسیار مهم است. ناکامی در برآورد ریسک اعتباری می تواند منجر به شکست های سیستمی مانند بحران مالی رخ داده در سال ۲۰۰۸ شود. در نتیجه، وام دهندگان منابع زیادی را برای پیش بینی اعتبار مصرف کنندگان و شرکت ها اختصاص می دهند تا استراتژی های وام دهی مناسبی را توسعه دهند که خطرات آنها را به حداقل برساند. (شی و همکاران، ۲۰۲۲)

ریسک اعتباری امکان از دست دادن دارایی های وام دهنده به دلیل خطر نکول بدهی است که ممکن است از ناتوانی وام گیرنده در پرداخت های لازم ناشی شود. این خطر نکول همواره زیان هایی را به همراه خواهد داشت که برخی از آن ها عبارتند از:

- یک مصرف کننده ممکن است در پرداخت وام رهنی، کارت اعتباری، خط اعتباری یا سایر وام ها شکست بخورد.
- یک شرکت قادر به بازپرداخت بدهی شارژ ثابت یا شناور تحت تضمین دارایی نیست.
- یک تجارت یا مصرف کننده فاکتور تجاری را در زمان مقرر پرداخت نمی کند.
- یک کسب و کار در زمان مقرر دستمزد به دست آمده کارمند را پرداخت نمی کند.
- یک شرکت تجاری یا ناشر اوراق قرضه دولتی در زمان سررسید، پرداختی را بابت کوپن یا اصل پرداخت انجام نمی دهد.

پس با توجه به این موارد اگر ریسک اعتباری نتواند به خوبی مدیریت شود می تواند زیان های قابل توجهی را شامل شده و در کنار آن به دلیل عدم ایفای نقش وام گیرنده و عدم بازپرداخت وام ریسک نقدینگی نیز در کنار آن افزایش خواهد یافت. از لحاظ تاریخی، رویکردهای ریسک اعتباری از روش های آماری مانند تحلیل تشخیصی خطی و رگرسیون لجستیک استفاده می کنند. با این حال، این روش ها به راحتی نمی توانند داده های بزرگ را مدیریت پیش بینی کنند. پیشرفت در قدرت محاسباتی و در دسترس بودن مجموعه داده های اعتباری بزرگ، راه را برای الگوریتم های تخمین ریسک اعتباری<sup>۳</sup> مبتنی بر هوش مصنوعی مانند یادگیری ماشین سنتی و یادگیری عمیق هموار کرد. (مسمودی و همکاران، ۲۰۱۹)

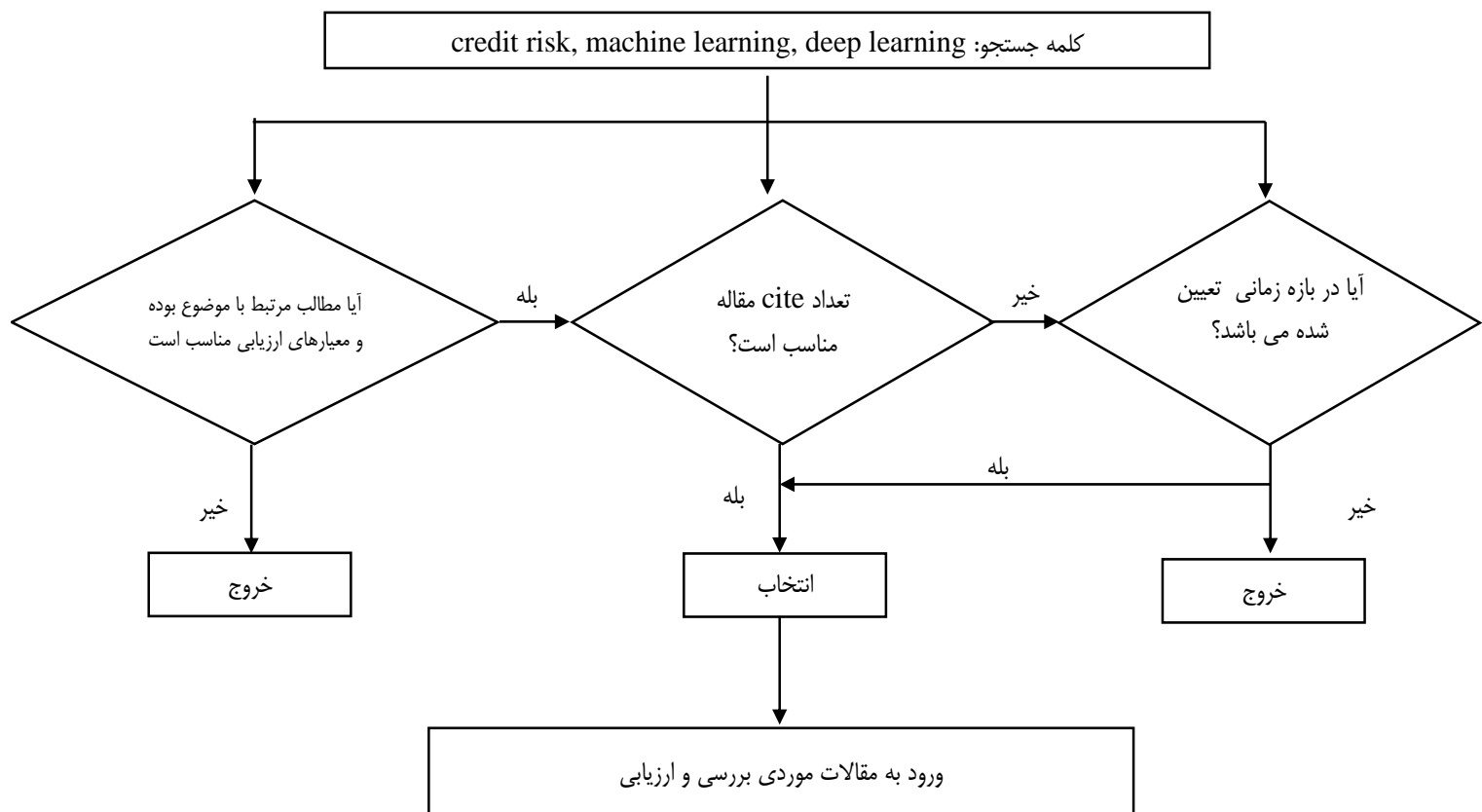
تکنیک های یادگیری ماشین معمولی، به عنوان مثال، k-نزدیک ترین همسایه<sup>۴</sup>، جنگل تصادفی<sup>۵</sup> و ماشین های بردار پشتیبان<sup>۶</sup>، نسبت به روش های آماری مؤثرتر و انعطاف پذیرتر هستند. به طور خاص، شاخه حیاتی تکنیک های یادگیری عمیق یادگیری ماشین که در دریاچه داده های ریسک اعتباری بزرگ اعمال می شود، هم از نظر دقت و هم از نظر کارایی بهتر از پیشینیان خود عمل می کنند. به طور خاص، شاخه حیاتی تکنیک های یادگیری عمیق یادگیری ماشین که در دریاچه داده های ریسک اعتباری بزرگ اعمال می شود، هم از نظر دقت و هم از نظر کارایی بهتر از پیشینیان خود عمل می کنند.

<sup>1</sup> Machine learning<sup>2</sup> Artificial Intelligence<sup>3</sup> credit risk<sup>4</sup> k-Nearest Neighbor<sup>5</sup> Random Forest<sup>6</sup> Support Vector Machines



این مقاله مروری سیستمایک از الگوریتم های برآورد ریسک اعتباری را ارائه می کند. هم رویکردهای آماری اصلی و هم تکنیک های مبتنی بر هوش مصنوعی را تجزیه و تحلیل می کند. هدف این است که یک نمای کلی جامع از فناوری برآورد ریسک اعتباری پیشرو فعلی، ارائه توجیه و ارتباط بین کارهای گذشته و حال ارائه شود. این کار یک طبقه بندی جدید را پیشنهاد می کند که امور مالی را با تکنیک های یادگیری ماشین ترکیب می کند.

در این مقاله، سه معیار ورود را انتخاب می کنیم: (۱) مرتبط بودن موضوع تحقیق، (۲) دقت معیارهای ارزیابی، (۳) سال انتشار و استنادها. علاوه بر این، اگر مقالات تکراری، ناقص، خیلی زود، کم ارتباط با موضوع، نداشتن معیارهای واضح یا استناد نسبتاً کم باشند، حذف می شوند. گردش کلی آن در نمودار ۱ به نمایش در آمده است.



نمودار ۱. گردش کار انتخاب مقالات

پس از آن نیز طبقه بندی به صورت شکل ۱ نشان داده شده است. ما می توانیم آن را به دو بخش تقسیم کنیم: اولی مربوط به فناوری محاسباتی است و دومی حوزه کاربرد ریسک اعتباری است. این دو بخش بیشتر در بخش های فرعی طبقه بندی می شوند. این دو قسمت به هم متصل شده و با هم ترکیب شده اند. همه زیر دامنه های سمت راست شامل تکنیک های سمت چپ هستند و همه تکنیک ها را می توان در حوزه های مالی اعمال کرد.



شکل ۱. طبقه بندی انواع روش های محاسبه

## روش تحقیق

در این در این پژوهش با بررسی مقالات گذشته و پارامترهای مورد بررسی در آن ها جهت بررسی کارایی آن ها اقدام خواهد شد. این بخش در ابتدا به طور خلاصه سه تکنیک محاسباتی اصلی مورد استفاده برای تجزیه و تحلیل اعتبار را معرفی می کند، یعنی یادگیری آماری، یادگیری ماشینی و یادگیری عمیق، که هر کدام ویژگی ها و اصول مشابه خود را دارند. رویکردهای آماری<sup>۷</sup> روش های سنتی برای طبقه بندی رفتار اعتباری مشتری یا شرکت هستند. با این حال، با پیشرفت سریع هوش مصنوعی، یادگیری ماشینی و یادگیری عمیق<sup>۸</sup> به تدریج جای تحلیل های آماری را گرفتند. با توجه با تحقیقات صورت گرفته رویکردهای آماری را به سه زیربخش تجزیه و تحلیل متمایز، رگرسیون لجستیک و مدل مرتبط بیزی تقسیم می کنیم. (لو و همکاران، ۲۰۱۵)

- تجزیه و تحلیل متمایز<sup>۹</sup> یک تکنیک کلاسیک برای پیش بینی گروه های نمونه است. هدف آن تولید ویژگی هایی است که می توانند متغیرهای باینری را از هم جدا کنند.

<sup>7</sup> Statistical approaches<sup>8</sup> deep learning<sup>9</sup> Linear Discriminant Analysis



- رگرسیون لجستیک<sup>۱۰</sup> یک الگوریتم طبقه بندی است که از تابع سیگموئید لجستیک برای له کردن خروجی تابع خطی در بازه (۰، ۱) و تفسیر آن مقدار به عنوان یک احتمال استفاده می کند.
- روش های Naive Bayes الگوریتم های یادگیری آماری هستند که قضیه بیز را با فرض استقلال شرطی بین هر جفت ویژگی در صورتی که متغیر کلاس داده شود، اعمال می کنند.

پس از آن نیز ما یک سری از الگوریتم های یادگیری ماشین را بررسی می کنیم که می توانند به خوبی در حوزه ریسک اعتباری اعمال شوند. این روش ها در شکل ۱ نیز به نمایش درآمده است که برخی از آن ها عبارتند از:

- ماشین بردار پشتیبانی<sup>۱۱</sup> یک ابر صفحه (مرز تصمیم) را پیاده سازی می کند که می تواند کلاس ها را در یک فضای ویژگی با ابعاد بالا جدا کند.
- Boosting یک روش مجموعه ای است که مدل های فردی را برای به دست آوردن ظرفیت بالاتر ترکیب می کند.
- KNN متعلق به روش های طبقه بندی است که کلاس اکثر k نزدیک ترین همسایگان یک متغیر ورودی x را در یک داده به آن اختصاص می دهند.

یادگیری عمیق در مقایسه با یادگیری ماشین سنتی دارای لایه های عمیق تر و واحدهای بیشتری در یک لایه است. می تواند توابع افزایش پیچیدگی را نشان دهد. در ادامه، برخی از روش های یادگیری عمیق مورد استفاده در ریسک اعتباری را بررسی می کنیم (داستایل و همکاران، ۲۰۲۰) که آن ها عبارتند از:

- شبکه های عصبی بازگشتی (RNNs) خانواده ای از شبکه های عصبی برای پردازش داده های متوالی هستند. آنها بهتر می توانند اطلاعات متوالی را به جای داده های مکانی که شبکه های عصبی کانولوشنال (CNN) می توانند به طور موثر پردازش کنند، مدیریت کنند. RNN ها متغیرهای حالت را برای ذخیره اطلاعات گذشته و همچنین ورودی های فعلی معرفی می کنند که هر دو خروجی های فعلی را تعیین می کنند.
- LSTM برای اولین بار برای تولید مسیریایی که در آن گرادیان برای مدت طولانی جریان دارد، توسعه یافت. این نوع شبکه های عصبی بازگشتی (RNN) است. در مقایسه با RNN های سنتی، می تواند ناپدید شدن گرادیان و انفجار را در فرآیند توالی طولانی مدت حل کند.
- CNN به طور گسترده در پردازش تصویر، تشخیص صدا، سیستم های QA خودکار و بسیاری از زمینه های محاسباتی دیگر به کار گرفته شدند. CNN ها از یک لایه ورودی، لایه های کانولوشن، لایه های ادغام و لایه های کاملاً متصل تشکیل شده اند. در شبکه عصبی CNN مهمترین لایه، لایه کانولوشن می باشد که فرمول آن به صورت زیر است:

$$s(t) = \int_{-\infty}^{\infty} x(a)w(t-a)da \quad \text{رابطه (۱)}$$

در دهه های گذشته، بسیاری از محققان از الگوریتم ها و مدل های محاسباتی مختلفی برای حل پیش بینی و ارزیابی ریسک اعتباری استفاده کرده اند.

<sup>10</sup> Logistic regression

<sup>11</sup> Support Vector Machine





در بررسی های تحلیل آماری در تحقیقات انجام شده توسط گالیندو و تامایو (۲۰۰۰) جهت برآورد دقیق ریسک، و استفاده از آن در مدل های ریسک مالی جهانی شرکت ها، یک تحلیل مقایسه ای از روش های مختلف مدل سازی آماری و یادگیری ماشینی طبقه بندی بر روی مجموعه داده های وام مسکن با انگیزه درک محدودیت ها و پتانسیل آنها انجام دادند. نتایج نشان می دهد که مدل های درخت تصمیم CART بهترین تخمین را برای پیش فرض با میانگین نرخ خطای ۸,۳۱٪ برای یک نمونه آموزشی از ۲۰۰۰ رکورد ارائه می دهند.

در ادامه در مقاله سیریگنانو و همکارانش (۲۰۱۶) یک مدل یادگیری عمیق از ریسک وام مسکن چند دوره ای توسعه داده می شود و از آن برای تجزیه و تحلیل مجموعه داده های مربوط به سوابق عملکرد ماهانه برای بیش از ۱۲۰ میلیون وام مسکن بین سال های ۱۹۹۵ و ۲۰۱۴ استفاده می گردد. نتایج حاصل شده از مدل نشان داد که یادگیری عمیق در اندازه گیری ریسک های وام مسکن موثر است.

مقاله نوشته شده توسط فنگ و همکارانش (۲۰۲۰)، آن ها تکنیک های یادگیری ماشین را برای مقابله با رتبه بندی اعتباری شرکت ها پیاده سازی می کنند. با این حال، توانایی این مدل ها توسط مقادیر عظیمی از داده های گزارش های صورت های مالی محدود شده است. برای استفاده از شبکه های عصبی کانولوشنال قدرتمند و داده های مالی عظیم، آن ها یک روش جدید پایان به انتها، رتبه بندی اعتباری شرکتی از طریق شبکه های عصبی پیچیده، CCR-CNN را پیشنهاد می دهند. در مدل پیشنهادی، هر شرکت به یک تصویر تبدیل می شود. نتایج حاصل شده ثابت می کند که CCR-CNN به طور مداوم از روش های پیشرفته برتر عمل می کند.

در تحقیقات انجام شده توسط خاتری و همکارانش (۲۰۲۰) آن ها روش های LR، SVM، k-NN، NB، RF، DT، MLP را پیاده سازی کردند و دریافتند که همه آنها عملکرد مطلوبی را از خود نشان دادند اما مدل های مربوط به درخت بهترین عملکرد را دارند. در ادامه تحقیقات صورت گرفته ژا و هو (۲۰۲۰) یک شبکه عمیق بر اساس ماشین محدود بولتزمن و طبقه بندی کننده SOFTMAX ایجاد می کنند. برای ورودی مدل نیاز به داده های بانکی بوده است که این مجموعه داده از گزارش های مالی سالانه شرکت های برجسته چینی تهیه شده است. این مدل دقتی را نشان می دهد که بسیار فراتر از SVM و رگرسیون لجستیک است. در پژوهش صورت گرفته توسط دانگو همکارانش (۲۰۲۱) آن ها یک مدل پیش بینی ثقلی برای زنجیره تامین مالی را بر اساس مدل طبقه بندی SVM پیشنهاد می کنند. با توجه به حجم زیاد داده ها، ابتدا مهندسی ویژگی را روی داده ها انجام می دهند تا داده های پردازش شده ای را که می توان برای مدل سازی استفاده کرد، به دست آورد و سپس از الگوریتم مدل طبقه بندی SVM برای طبقه بندی و رگرسیون داده ها استفاده کرد. آزمایشات نشان می دهد که دقت مدل طبقه بندی SVM 98.61% است. در مقایسه با مدل رگرسیون لجستیک و مدل ساده بیز، قابلیت طبقه بندی داده ها و رگرسیون بهتری دارد.

پس از آن در مقاله شن و همکارانش (۲۰۲۱)، یک مدل ارزیابی ریسک اعتباری مجموعه یادگیری عمیق جدید را برای مقابله با داده های اعتباری نامتعادل ایجاد کردند. ابتدا، یک روش بهبود یافته روش نمونه گیری اقلیت مصنوعی (SMOTE) برای غلبه بر کاستی های شناخته شده SMOTE ایجاد شد، پس از آن یک روش طبقه بندی مجموعه یادگیری عمیق جدید همراه با شبکه حافظه بلندمدت (LSTM) و تقویت تطبیقی (AdaBoost) الگوریتم برای آموزش و یادگیری داده های اعتباری پردازش شده توسعه داده شد. نتایج آزمون تجربی نشان داد که مدل مجموعه یادگیری عمیق پیشنهادی به طور کلی در هنگام پرداختن به مشکلات ارزیابی ریسک اعتباری نامتعادل نسبت به مدل های دیگر رقابتی تر است.

در ادامه نیز در همان سال ژائو و شیائو (۲۰۲۱) در پژوهش خود به مطالعه تحقیق در مورد مدیریت ریسک اعتباری مالی مصرف کننده توسط کلان داده می پردازد. نتایج تجربی این مقاله نشان می دهد که این مدل توانایی پیش بینی خوبی دارد، می تواند بین

مشتریان وام عادی و مشتریان وام پیش فرض تمایز قائل شود و برای تجارت کنترل ریسک اعتباری شخصی مناسب است. دقت پیش بینی مدل پیش فرض مدل تلفیقی ۹۷,۱۴ درصد است و نرخ پیش فرض مربوط به تجارت واقعی ۲,۸۶ درصد است

#### یافته ها

پس از بررسی پژوهش های صورت گرفته، چهار چالش عمده در تحقیق ریسک اعتباری مبتنی بر یادگیری ماشین را خلاصه می توان اشاره کرد. اول، عدم تعادل داده ها در ریسک اعتباری بسیار شدید است. اگرچه روش های متعددی مانند نمونه گیری بیش از حد و نمونه گیری کمتر (معمولاً برای کم نمونه گیری اکثریت انتخاب می شوند) برای حل این مشکل پیشنهاد شده اند، نتایج هنوز هم از نظر اثربخشی و هم از نظر کارایی رضایت بخش نیستند. دوم، کمبود مجموعه داده های معیار جدی است. اکثر کارهای موجود از مجموعه داده های خصوصی استفاده می کنند، بنابراین نتایج مقایسه عملکرد نمی تواند به اندازه کافی منصفانه باشد. سوم، بیشتر مدل های یادگیری ماشینی واضح نمی باشند زیرا معمولاً شفاف نیستند. شفافیت اطلاعات باید مورد توجه قرار گیرد. . چهارم، استفاده از مدل های یادگیری عمیق هنوز در ریسک اعتباری محدود است.

منبع	روش								LR	Bayesian
	SVM	Random Forest	KNN	XGBoost	CNN	MPL	ANN	LSTM		
گلپندو و تامایو (۲۰۰۰)		✓	✓				✓			
سپرینگلو و همکارانش (۲۰۱۶)							✓		✓	
فنگ و همکارانش (۲۰۲۰)	✓	✓	✓	✓	✓					✓
مارسو و همکاران (۲۰۲۰)	✓	✓							✓	
خاتری و همکاران (۲۰۲۰)	✓	✓	✓						✓	✓
زا و هو (۲۰۲۰)	✓					✓			✓	
دانگو همکارانش (۲۰۲۱)	✓								✓	✓
شن و همکارانش (۲۰۲۱)				✓			✓	✓		
ژائو و شیائو (۲۰۲۱)									✓	

#### جدول ۱. روش های مورد استفاده در مقالات مورد مطالعه

همانطور که مشاهده می شود دو روش لجستیک رگرسیون و SVM از پراستفاده ترین روش ها جهت ارزیابی ریسک بوده که حتی در سال های اخیر نیز مورد بررسی و ارزیابی قرار گرفته است.



روش	تعداد
SVM	۵
Random Forest	۴
KNN	۳
XGBoost	۲
CNN	۱
MPL	۱
ANN	۳
LSTM	۱
LR	۶
Bayesian	۳

جدول ۲. تعداد روش های استفاده شده در مقالات مورد مطالعه

همانطور که مشاهده می شود دو روش لجستیک رگرسیون و SVM از پراستفاده ترین روش ها جهت ارزیابی ریسک بوده که حتی در سال های اخیر نیز مورد بررسی و ارزیابی قرار گرفته است.

### بحث و نتیجه گیری

پیشرفت در قدرت محاسباتی و در دسترس بودن مجموعه داده های اعتباری بزرگ، راه را برای الگوریتم های تخمین ریسک اعتباری مبتنی بر هوش مصنوعی مانند یادگیری ماشین و یادگیری عمیق هموار کرد و این امکان را داده تا به وسیله آن بتوان ریسک اعتباری پیش بینی کرده و موارد موثر بر آن استخراج شود. با توجه به مطالعات صورت گرفته در تحقیقات صورت گرفته در دهه گذشته تمرکز اصلی بر روی روش های آماری و یادگیری ماشینی بوده که به مرور زمان توجه بیشتری به روش های یادگیری عمیق شده است. همچنین تلاش شده تا از روش های تک حالت خارج شده و روش ها با یکدیگر ترکیب گردد. نتایج نشان می دهد که روش های یادگیری عمیق قوی تر از روش های یادگیری ماشینی و آماری سنتی هستند، اگرچه به طور کامل به کار گرفته نشده اند. همچنین، این نتیجه که مجموعه ای از چندین روش بهتر از یک روش واحد عمل می کنند، در برخی از تحقیقات مرتبط به اثبات رسیده است.

### منابع

Bequé, A., & Lessmann, S. (2017). Extreme learning machines for credit scoring: An empirical evaluation. Expert Systems with Applications, 86, 42-53.

Dastile, X., Celik, T., & Potsane, M. (2020). Statistical and machine learning models in credit scoring: A systematic literature survey. Applied Soft Computing, 91, 106263.

Dong, Yiyang, Keyu Xie, Zhang Bohan, and Lanchuan Lin. "A machine learning model for product fraud detection based on svm." In 2021 2nd International Conference on Education, Knowledge and Information Management (ICEKIM), pp. 385-388. IEEE, 2021.



Feng, B., Xue, W., Xue, B., & Liu, Z. (2020, December). Every corporation owns its image: Corporate credit ratings via convolutional neural networks. In 2020 IEEE 6th International Conference on Computer and Communications (ICCC) (pp. 1578-1583). IEEE.

Galindo, J., & Tamayo, P. (2000). Credit risk assessment using statistical and machine learning: basic methodology and risk modeling applications. *Computational economics*, 15, 107-143.

Gao, L., & Xiao, J. (2021). Big data credit report in credit risk management of consumer finance. *Wireless Communications and Mobile Computing*, 2021, 1-7.

Khatri, S., Arora, A., & Agrawal, A. P. (2020, January). Supervised machine learning algorithms for credit card fraud detection: a comparison. In 2020 10th international conference on cloud computing, data science & engineering (confluence) (pp. 680-683). IEEE.

Liu, S., McGree, J., Ge, Z., & Xie, Y. (2015). Computational and statistical methods for analysing big data with applications. Academic Press.

Lopez, J. A., & Saidenberg, M. R. (2000). Evaluating credit risk models. *Journal of Banking & Finance*, 24(1-2), 151-165.

Marceau, L., Qiu, L., Vandewiele, N., & Charton, E. (2019). A comparison of Deep Learning performances with other machine learning algorithms on credit scoring unbalanced data. *arXiv preprint arXiv:1907.12363*.

Masmoudi, K., Abid, L., & Masmoudi, A. (2019). Credit risk modeling using Bayesian network with a latent variable. *Expert Systems with Applications*, 127, 157-166.

Shen, F., Zhao, X., Kou, G., & Alsaadi, F. E. (2021). A new deep learning ensemble credit risk evaluation model with an improved synthetic minority oversampling technique. *Applied Soft Computing*, 98, 106852.

Shi, S., Tse, R., Luo, W., D'Addona, S., & Pau, G. (2022). Machine learning-driven credit risk: a systemic review. *Neural Computing and Applications*, 34(17), 14327-14339.

Sirignano, J., Sadhwani, A., & Giesecke, K. (2016). Deep learning for mortgage risk. *arXiv preprint arXiv:1607.02470*.

Xu, R. Z., & He, M. K. (2020, April). Application of deep learning neural network in online supply chain financial credit risk assessment. In 2020 international conference on computer information and big data applications (CIBDA) (pp. 224-232). IEEE.