



## مقایسه رگرسیون خطی چندگانه و الگوریتم‌های یادگیری ماشین در پیش‌بینی نگهداشت وجه نقد

سمیرا سیف<sup>۱</sup>

مصطفی یوسفی طرزجان<sup>۲</sup>

تاریخ دریافت مقاله: ۱۴۰۲/۰۱/۳۰ تاریخ پذیرش مقاله: ۱۴۰۲/۰۳/۱۷

### چکیده

اخیراً، در ادبیات مالی توجه به سطح نگهداشت وجه نقد در شرکت‌ها اهمیت زیادی یافته است. لذا؛ پیش‌بینی برای تعیین سطح بهینه نگهداشت وجه نقد اهمیت دارد. در این پژوهش با استفاده از روش‌های خطی و غیرخطی و ۱۳ متغیر ورودی تأثیرگذار در میزان وجه نقد در ۱۰۳ شرکت پذیرفته شده در بورس اوراق بهادار ایران طی سال‌های ۱۳۹۲ تا ۱۴۰۰ پیش‌بینی شده است. روش‌های به‌کار رفته شامل رگرسیون خطی چندگانه (MLR)، نزدیکترین k همسایه (KNN)، ماشین بردار پشتیبان (SVM) و شبکه‌های عصبی چند لایه (MLNN) می‌باشد. نتایج نشان می‌دهد که روش سنتی رگرسیون خطی چندگانه در پیش‌بینی وجه نقد موفق عمل نکرده است، ولی الگوریتم‌های یادگیری ماشین با دقت ۰/۹۹ برتر بوده‌اند. متغیرهای سود هر سهم، نسبت دارایی‌های جاری به بدهی‌های جاری و نسبت بدهی کوتاه‌مدت به کل دارایی‌ها تأثیرگذاری بیشتری در همه الگوریتم‌ها داشته‌اند. بنابراین، مدیران می‌توانند از الگوریتم‌های پیشرفته یادگیری ماشین جهت پیش‌بینی میزان وجه نقد آینده شرکت‌ها بهره بگیرند.

### کلمات کلیدی

پیش‌بینی، نگهداشت وجه نقد، رگرسیون خطی چندگانه، الگوریتم‌های یادگیری ماشین.

۱-استادیار، گروه حسابداری، دانشگاه پیام نور، تهران، ایران. (نویسنده مسئول) samiraseif@pnu.ac.ir

۲-استادیار، گروه مهندسی صنایع، دانشگاه جامع علمی کاربردی، کرج، ایران. yousofi@uast.ac.ir

همواره هدف شرکت‌ها، نگهداری مقدار مناسب وجه نقد است. از مزایای مهم نگهداشت وجه نقد این است که به شرکت اجازه می‌دهد از افزایش هزینه‌های تأمین مالی خارجی یا خروج دارایی‌های نقدی موجود در شرکت هنگام تأمین مالی برای فرصت‌های رشد پیش آمده، جلوگیری نماید (پاولینا و رنوبگ<sup>۱</sup>، ۲۰۰۵). ناکافی بودن وجه نقد در شرکت موجب چشم‌پوشی از پروژه‌های سرمایه‌گذاری سودآور و یا حمایت بیش از حد و غیرمنطقی از تأمین مالی پرهزینه، می‌گردد (لی و پوول<sup>۲</sup>، ۲۰۱۱).

از سوی دیگر انباشت وجه نقد بدون بازده موجب کاهش کارایی و در نهایت کاهش ارزش شرکت خواهد شد (پینکوویتزو<sup>۳</sup> همکاران، ۲۰۰۶). این مسئله باعث می‌شود مدیران بدنبال علایق خود بروند و در نتیجه منجر به زیان سهامداران و ایجاد بحران مالی می‌گردد. نرخ بازده نقدی شرکت نگهداری معمولاً کمتر از نرخ بهره بازار است که هزینه فرصت دارایی‌های نقدی را افزایش می‌دهد (وو و همکاران، ۲۰۲۱). بنابراین، ایجاد توازن بین نگهداری وجوه نقد و نیازهای نقدی، عامل مهمی در سلامت اقتصادی واحد تجاری و تداوم فعالیت آن است. تصمیم‌گیری برای تعیین میزان ذخایر نقدی شرکت‌ها، یکی از مسائل شایان توجه در ادبیات تأمین مالی محسوب می‌شود. برای تعیین رفتار نگهداری وجه نقد شرکت‌ها، مطالعات با متغیرهای مالی مختلفی در سطح شرکت در این ادبیات مورد استفاده قرار داده است. برخی از متغیرهای مورد استفاده به شرح زیر است: اندازه، اهرم، سود تقسیم شده، رشد فروش، سرمایه در گردش (رهنمای رودپشتی و کیایی، ۱۳۸۷؛ بیگیلی و سانچز<sup>۴</sup>، ۲۰۱۲؛ بوبکری و همکاران<sup>۵</sup>، ۲۰۱۳)، جریان‌ات نقدی (کاشانی‌پور و نقی‌نژاد، ۱۳۸۸؛ بوبکری و همکاران، ۲۰۱۳؛ دیو<sup>۶</sup>، ۲۰۲۱؛ فریرا و ویللا<sup>۷</sup>، ۲۰۰۴؛ لوزانو و یامان<sup>۸</sup>، ۲۰۲۰؛ اویار و کوزی، ۲۰۱۴)، مخارج سرمایه‌ای (بوبکری و همکاران، ۲۰۱۳؛ دیو، ۲۰۲۱؛ اویار و کوزی، ۲۰۱۴).

با استفاده از روش‌های رگرسیون کلاسیک، تاثیر بسیاری از متغیرهای مالی بر رفتار نگهداشت وجه نقد شرکت‌ها مورد بررسی قرار گرفته است. برخلاف ادبیات قبلی، ما سعی می‌کنیم رفتار نگهداری وجه نقد شرکت‌ها را با اعمال رویکردهای یادگیری ماشین پیشرفته تحت نظارت برای رفع شکاف موجود در ادبیات پیش‌بینی کنیم که تا کنون پژوهشی با این رویکرد در ایران انجام نشده است. یادگیری ماشین که امروزه یکی از رایج‌ترین روش‌های تجزیه و تحلیل داده‌ها است، الگوریتم‌هایی را شامل می‌شود که نتایج را تا حد امکان دقیق پیش‌بینی می‌کنند. یادگیری ماشین می‌تواند به شرکت‌ها در پیش‌بینی وجه نقد کمک کند یا سطح نگهداشت وجه نقد آنها را در آینده تخمین بزنند. پیش‌بینی وجه نقد به مدیران در تعیین چگونگی استفاده از وجه نقد برای ایجاد سود بیشتر کمک خواهد کرد و مدیران می‌توانند از

## مقایسه رگرسیون خطی چندگانه و الگوریتم‌های یادگیری ماشین... / سیف و یوسفی طزرجان

شرکت در برابر چالش‌های مالی محافظت کنند (دونپودی و همکاران ۲۰۲۰). علاوه بر این، تکنیک‌های یادگیری ماشین را می‌توان برای پیش‌بینی و تجزیه و تحلیل به جای گزارش صرف اعداد و آمار استفاده کرد (رافی و همکاران، ۲۰۲۰).

### **مبانی نظری و پیشینه پژوهش**

#### **نگهداشت وجه نقد**

نگهداشت وجه نقد شامل پول نقد در دسترس برای سرمایه‌گذاری در دارایی‌های فیزیکی و توزیع بین سرمایه‌گذاران و یا موجودی در بانک، سرمایه‌گذاری کوتاه‌مدت می‌باشد. از مزایای مهم نگهداشت وجه نقد این است که به شرکت اجازه می‌دهد از افزایش هزینه‌های تأمین مالی خارجی یا خروج دارایی‌های نقدی موجود در شرکت هنگام تأمین مالی برای فرصت‌های رشد پیش‌آمده، جلوگیری نماید (گیل و شاه<sup>۹</sup>، ۲۰۱۲).

انگیزه‌های نگهداشت وجه نقد شامل انگیزه معاملاتی که نشان‌دهنده این است که شرکت‌ها وجه نقد را به شکلی ساده برای معاملات روزانه خود نگهداری کنند و میزان وجه نقد شرکت‌ها بسته به نوع فعالیت، پیچیدگی فناوری و هزینه فرصت از دست رفته آنها متفاوت است. (اپلر و همکاران، ۱۹۹۹).

انگیزه احتیاطی: در دوره‌هایی که تأمین مالی خارجی گران است شرکت‌ها وجه نقد را برای ادامه سرمایه‌گذاری در طرح‌هایی با ارزش فعلی خالص مثبت نگهداری می‌کنند. به‌ویژه زمانی که شرکت‌ها قادر به ایجاد وجه نقد داخلی کافی برای تأمین مالی این سرمایه‌گذاری‌ها نباشند. بنابراین، نگهداری وجه نقد برای شرکت‌هایی که با محدودیت تأمین مالی روبرو هستند ایجاد ارزش مثبت می‌کند (مولیگان، ۱۹۹۷).

#### **نظریات مرتبط با وجه نقد**

#### **نظریه عدم تقارن اطلاعاتی**

طبق نظریه عدم تقارن اطلاعاتی، کاهش عدم تقارن اطلاعاتی (به‌عنوان مثال از طریق افزایش کیفیت گزارشگری مالی) می‌تواند منجر به کاهش سطح نگهداشت وجه نقد توسط شرکت‌ها شود.

#### **نظریه نمایندگی**

طبق نظریه نمایندگی، تضادهایی که منجر به شناسایی هزینه‌های نمایندگی می‌شوند می‌توانند رفتار نگهداری وجه نقد توسط مدیریت را توجیه کند. با توجه به این تئوری مدیران در شرکت‌ها با دارایی‌های نقدی بالا، انگیزه بیشتری برای استفاده از این دارایی‌ها در جهت منافع خود دارند و توجهی به میزان سطح بهینه و انجام سرمایه‌گذاری‌های مناسب ندارند (فریرا و ویللا، ۲۰۰۴).

### نظریه موازنه

طبق نظریه موازنه، شرکت‌ها میزان بهینه وجه نقد خود را با برقراری تعادل میان منافع و هزینه‌های نگهداری وجه نقد تعیین می‌کنند. نکته مهم این نظریه این است که سطح مطلوبی از وجه نقد برای شرکت‌ها وجود دارد که در آن مدیران با رویکردی فعالانه، بر اساس تحلیله هزینه-منفعت نسبت به نگهداری وجه نقد تصمیم‌گیری می‌کنند (جنی و همکاران، ۲۰۰۴).

در ایران پژوهش‌های زیادی در ارتباط با تعیین میزان بهینه نگهداشت وجه نقد و ارتباط این متغیر با سایر متغیرها انجام شده است؛ لیکن هیچگونه پژوهشی جهت پیش‌بینی میزان نگهداشت وجه نقد از الگوریتم‌های یادگیری ماشین انجام نشده است. در ادامه به بررسی سایر پژوهش‌های مرتبط می‌پردازیم: آقای و همکاران (۱۳۸۸) به بررسی متغیرهای مؤثر بر نگهداری وجه نقد در شرکت‌های پذیرفته شده بورس اوراق بهادار تهران طی ۱۳۷۹-۱۳۸۴ پرداختند. نتایج پژوهش آنها نشان می‌دهد که حساب‌های دریافتی، خالص سرمایه در گردش، موجودی‌های کالا و بدهی‌های کوتاه‌مدت، به ترتیب، از مهم‌ترین عواملی هستند که بر نگهداری موجودی‌های نقدی، تاثیر منفی دارند. اما شواهد کافی در مورد تاثیر منفی بدهی‌های بلندمدت و اندازه شرکت‌ها بر نگهداری موجودی‌های نقدی وجود ندارد. حساس یگانه و همکاران (۱۳۹۰) به بررسی عوامل مؤثر در سطح نگهداشت وجه نقد از قبیل بحران مالی، فرصت‌های رشد، اندازه شرکت، ساختار سررسید بدهی‌ها، ارتباط با موسسات مالی، اهرم مالی، درصد تقسیم سود و نسبت سرمایه‌گذاری در دارایی‌های ثابت پرداختند. نتایج بیانگر اینست که متغیرهای به کار گرفته شده می‌تواند ۸۶٪ از تغییرات در سطح نگهداشت وجه نقد را توضیح دهد. فروغی و فرزادی (۱۳۹۳) در پژوهشی رابطه بین تغییرات جریان‌های نقدی و نگهداشت وجه نقد با در نظر گرفتن محدودیت تأمین مالی شرکت‌های پذیرفته‌شده در بورس اوراق بهادار تهران را بررسی کردند. نتایج حاکی از آن است که بین جریان‌های نقدی (مثبت و منفی) و سطح نگهداشت وجه نقد شرکت‌ها رابطه معکوس و معناداری وجود دارد. از سوی دیگر، شرکت‌های دارای محدودیت مالی، بدون در نظر گرفتن نوع جریان‌های نقدی، تمایل بیشتری به نگهداشت وجه نقد دارند.

مشکی میاوقی و صنایعی (۱۳۹۵) به بررسی تاثیر انحراف از سطح بهینه نگهداشت وجه نقد بر ارزش وجه نقد پرداختند و نتایج نشان داد که ارزش یک ریال وجه نقد تحصیل شده در نزد سهامداران شرکت‌هایی که سطح نگهداشت وجوه نقد آنها از سطح بهینه وجه نقد کمتر است، بیشتر از شرکت‌هایی است که وجوه نقد آنها از سطح بهینه وجه نقد بیشتر است. همچنین نتایج پژوهش موید آن است که ارزش وجه نقد در شرکت‌هایی با نگهداشت وجه نقد بیشتر از سطح بهینه، کمتر از ارزش اسمی یک ریال

## مقایسه رگرسیون خطی چندگانه و الگوریتم‌های یادگیری ماشین.../ سیف و یوسفی طزر جان

و در شرکت‌هایی با نگهداشت وجه نقد کمتر از سطح بهینه بیشتر از ارزش اسمی یک ریال است. حاجیها و رجب دری (۱۳۹۶) نیز ادعا کردند که اثربخشی کمیته حسابرسی موجب بهبود مدیریت وجه نقد، افزایش سطح نگهداشت وجه نقد و در نهایت افزایش سرعت حرکت به سمت بهینه خواهد شد. نیکبخت و همکاران (۱۴۰۰) در پژوهشی به بررسی آزمون اثربخشی تمرکز مشتری بر سرعت تعدیل نگهداشت وجه نقد در ۱۴۰ شرکت پذیرفته شده در بورس اوراق بهادار تهران پرداختند. نتایج پژوهش حاکی از تأثیر مثبت و معنادار تمرکز مشتری بر سطح نگهداشت وجه نقد و سرعت تعدیل نگهداشت وجه نقد می‌باشد. به عبارتی، در شرکت‌های با مشتری عمده، میزان سطح نگهداشت وجه نقد و سرعت تعدیل نگهداشت وجه نقد بالاتر بوده است. این نتایج مبین اهمیت وجود مشتریان عمده در شرکت و اثربخشی این مشتریان در افزایش سطح نگهداشت وجه نقد و سرعت تعدیل نگهداشت وجه نقد می‌باشد. غلامزاده و همکاران (۲۰۲۱) محدودیت‌های مالی را برای شرکت‌های پذیرفته شده در بورس اوراق بهادار تهران با استفاده از فرآیند گاوسی و شعاعی شبکه عصبی پیش‌بینی کردند. آنها تأیید کردند که روش‌های یادگیری ماشین برای پیش‌بینی محدودیت‌های مالی مناسب هستند. درصد مالکیت نهادی، بازده دارایی‌ها، اهرم مالی، جریان نقدی عملیاتی به دارایی‌ها و ارزش شرکت متغیرهای اصلی در پیش‌بینی محدودیت‌های مالی هستند.

### **داده‌ها و روش‌شناسی پژوهش**

در این پژوهش برای تجزیه تحلیل داده‌ها از الگوریتم‌های یادگیری ماشین تحت نظارت شامل ماشین بردار پشتیبان، شبکه‌های عصبی چند لایه، نزدیک‌ترین K همسایگی، رگرسیون خطی چندگانه به‌عنوان مدل‌های پیش‌بینی استفاده شده است و آزمون‌های آماری با استفاده از نرم‌افزارهای EXCEL، SPSS Modeler انجام شده است. جامعه آماری شامل داده‌های ۱۰۳ شرکت پذیرفته شده در بورس ایران از سال ۱۳۹۲ تا ۱۴۰۰ است. متغیرهای داده سالانه در سطح شرکت با استفاده از نرم افزار ره آورد نوین به دست آمده است. شرکت‌هایی مانند بانک‌ها، بیمه، لیزینگ و سایر شرکت‌هایی که مرتبط با موسسات مالی هستند به دلیل اینکه نسبت‌های حسابداری آنها با نسبت‌های حسابداری سایر شرکت‌ها قابل مقایسه نیست، حذف شدند. همچنین، داده‌ها شامل شرکت‌هایی است که حداقل هفت سال متوالی داده داشته باشند که به‌طور کلی شامل ۹۴۵ داده است. برای جلوگیری از مشکل بیش از حد برازش که باعث پیش‌بینی‌های ضعیف داده‌های دیده نشده می‌شود (۰/۷۵) از مجموعه داده برابر با ۷۱۲ نمونه را به‌عنوان داده آموزشی تقسیم کرده و (۰/۲۵) داده‌ها برابر با ۲۳۶ نمونه به‌عنوان داده‌های آزمایشی در نظر گرفته شده است.

## فصلنامه مهندسی مالی و مدیریت اوراق بهادار / دوره ۱۴ / شماره ۵۷ / زمستان ۱۴۰۲

در این پژوهش متغیرهای ورودی تاثیرگذار برای پیش‌بینی میزان وجه نقد با استفاده از رگرسیون خطی چندگانه و الگوریتم‌های یادگیری ماشین طبق پژوهش (اولزم و عمر، ۲۰۲۲) و سایر پژوهش‌های مرتبط به شرح جدول ۱ است.

**جدول ۱- تعریف و نماد متغیرهای ورودی مدل‌ها**

ردیف	متغیر	نماد	منابع
۱	نسبت دارایی‌های جاری به بدهی‌های جاری	(CR)	مانوئل و همکاران <sup>۱۰</sup> (۲۰۱۸)، اوزکان <sup>۱۱</sup> (۲۰۰۱)
۲	نسبت دارایی‌های ثابت خالص به کل دارایی‌ها	(TANG)	بهویان و هوکس (۲۰۱۹)، درویتس گرونینگر (۲۰۰۷)، اویار و کوزی (۲۰۱۴)
۳	نسبت سود نقدی پرداختی به کل دارایی‌ها	(DIV)	بیگلی و سانچز ویدال (۲۰۱۲)، بهویان و هوکس (۲۰۱۹)، سونگ و لی (۲۰۱۲)، وو و همکاران (۲۰۲۱)
۴	سود هر سهم	(EPS)	سرفریز و همکاران (۲۰۲۰)
۵	نسبت بدهی کوتاه‌مدت به کل دارایی‌ها	(STD)	بنکریم و همکاران (۲۰۲۰)، لوزانو و یامان (۲۰۲۰)
۶	لگاریتم طبیعی کل دارایی‌ها	(SIZE)	بیگلی و سانچز ویدال (۲۰۱۲)، بوبکری و همکاران (۲۰۱۳)، درویتس و گرونینگر (۲۰۰۷)، گارسیا-تروئل و مارتینز-سولانو (۲۰۰۸)، لوزانو و یامان (۲۰۲۰)، اوزکان و اوزکان (۲۰۰۴)
۷	نسبت سود عملیاتی پس از کسر مالیات به کل دارایی‌ها	(ROIC)	سرفریز و همکاران (۲۰۲۰)
۸	نسبت حساب‌های دریافتی به کل دارایی‌ها	(AR)	محمدی و همکاران (۲۰۱۸)، وو و همکاران (۲۰۱۲)
۹	نسبت سود خالص به کل دارایی‌ها	(ROA)	باتومن و همکاران (۲۰۲۱)، بهویان و هوکس (۲۰۱۹)، کای و همکاران (۲۰۱۶)، کامبرا و همکاران (۲۰۲۱)، سرفریز و همکاران (۲۰۲۰)
۱۰	نسبت سود خالص به مجموع حقوق صاحبان سهام	(ROE)	مانوئل و همکاران (۲۰۱۸)
۱۱	نسبت حساب‌های پرداختی به کل دارایی‌ها	(AP)	چن و همکاران (۲۰۱۴)، وو و همکاران (۲۰۱۲)
۱۲	تغییر سالانه در رشد فروش (%)	(SG)	بیگلی و سانچز ویدال (۲۰۱۲)، بوبکری و همکاران (۲۰۱۳)، کیم و همکاران (۲۰۲۱)، سونگ و لی (۲۰۱۲)
۱۳	نسبت هزینه بهره به کل دارایی‌ها	(IE)	شاوتن و همکاران (۲۰۱۱)
۱۴	نسبت پول نقد و معادل وجه نقد به کل دارایی‌ها	(CASH)	تامسون رویترز

منبع: یافته‌های پژوهش

اخیراً الگوریتم‌های یادگیری ماشین اغلب به‌عنوان ابزار پیش‌بینی امور مالی، به‌ویژه برای پیش‌بینی قیمت، مدیریت ریسک مالی، خدمات مالی، و تصمیم‌گیری مورد استفاده قرار گرفته‌اند.

## مقایسه رگرسیون خطی چندگانه و الگوریتم‌های یادگیری ماشین... / سیف و یوسفی طزرجان

طبقه‌بندهای ترکیبی از ترکیب چندین طبقه‌بند ۱۲ استفاده می‌کنند. طبقه‌بندها، هر کدام مدل خود را بر روی داده‌ها ساخته و این مدل را ذخیره می‌کنند. در نهایت جهت طبقه‌بندی یک رای‌گیری در بین این طبقه‌بندها انجام می‌شود و آن طبقه‌ای که بیشترین میزان رای را بیاورد، طبقه نهایی محسوب می‌شود. طبقه‌بندهای ترکیبی عموماً از بیش برآزش مدل یادگرفته شده توسط الگوریتم جلوگیری می‌کنند و در بسیاری از موارد نتایج بهتری نسبت به الگوریتم‌های دیگر تولید می‌کنند. الگوریتم‌های به کار رفته در این پژوهش به شرح زیر است.

### رگرسیون خطی چندگانه<sup>۱۳</sup>

مدل رگرسیون خطی چندگانه، شامل یک متغیر وابسته و چندین متغیر مستقل است و معادله آن به شکل رابطه ۱ می‌باشد:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_k x_k + \varepsilon \quad \text{رابطه ۱}$$

که در آن  $x_1 + x_2 + \dots + x_k$  متغیرهای مستقل،  $\varepsilon$  میزان اختلاف بین مقدار محاسباتی رگرسیون خطی چندگانه و مقدار واقعی متغیر وابسته، مقدار  $\beta_0$  عرض از مبدا و  $\beta_1$  ها ضرایب رگرسیونی یا ضرایب رگرسیونی جزئی را نشان می‌دهند. هر کدام از مقادیر  $\beta_1$  ها میزان تغییر در مقدار  $y$  به ازای یک واحد تغییر در مقدار  $x$  متناظر را، وقتی که سایر متغیرهای مستقل ثابت باشند، نشان می‌دهند.

به‌طور کلی، اگر چندین متغیر مستقل در تغییرپذیری متغیر وابسته مؤثر باشند، مدل رگرسیونی چندگانه، گزینه مناسبی برای تحلیل می‌باشد و قادر است متغیر وابسته را با دقت بیشتری پیش‌بینی کند. در رگرسیون خطی چندگانه فرض‌های زیر باید برقرار باشند: ۱- خطاها باید دارای توزیع نرمال با میانگین مساوی با صفر و واریانس ثابت باشد. ۲- خطاها ناهمبسته باشند. ۳- یک رابطه خطی بین متغیر وابسته و مستقل وجود داشته باشد (فرض خطی بودن). ۴- متغیرهای مستقل با همدیگر همبستگی خطی نداشته باشند (بیورانی و جباری، ۱۳۹۲).

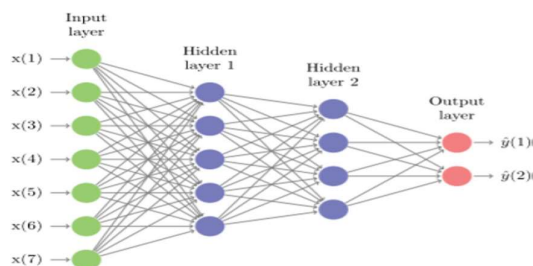
### ماشین بردار پشتیبان<sup>۱۴</sup>

ماشین بردار پشتیبان یکی از شناخته‌ترین الگوریتم‌ها می‌باشد که در حوزه یادگیری ماشین وجود دارد و اولین بار در سال ۱۹۹۵ توسط واپنیک به عنوان الگویی برای شناسایی و طبقه‌بندی مسائل مورد استفاده قرار گرفت. ماشین بردار پشتیبان دارای ویژگی تعمیم‌پذیری خوب، رسیدن به الگوی بهینه کلی، توانایی در طبقه بندی الگوهای ورودی و قابلیت یادگیری می‌باشد. در این مدل داده‌ها ابتدا به بردار

یادگیری تبدیل می‌شود و سپس هر بردار با یک بردار خروجی متناظر می‌باشد و قابلیت یافتن مقدار بهینه را در فضای غیرخطی و خطی دارا می‌باشد (واپنیک، ۱۹۹۵).

### شبکه عصبی چند لایه<sup>۱۵</sup>

روش شبکه‌های عصبی چند لایه مبنای مطالعات یادگیری عمیق را تشکیل می‌دهد و توسط روملپهارت و همکاران در سال ۱۹۸۶ توسعه داده شده است. شبکه‌ها از یک لایه ورودی، حداقل یک لایه پنهان و یک لایه خروجی تشکیل شده‌اند و هر لایه از مجموعه‌ای از واحدها (نرون‌ها) تشکیل شده است. لایه‌ها کاملاً متصل هستند به این معنی که تمام واحدهای ورودی از یک لایه به هر واحد فعال‌سازی لایه بعدی متصل هستند نمودار ۱. شبکه، پیش‌بینی را از طریق انتشار به جلو با چندین تابع فعال‌سازی محاسبه می‌کند و با اصلاح وزن‌ها و خطاهای شبکه برای تنظیم پارامترهای بهینه برای پیش‌بینی، خطا را از طریق انتشار به عقب به حداقل می‌رساند.



نمودار ۱- نمایش یک شبکه عصبی چند لایه. منبع: (دیکسون و همکاران، ۲۰۱۷)

این طبقه‌بندی مبتنی بر شبکه عصبی است که می‌تواند با بارگذاری لایه ورودی با بردار ورودی تحریک شود و سپس مقادیر را از طریق ارتباطات وزنی در کل شبکه پخش می‌کند (هورنیک و همکاران، ۱۹۸۹).

### نزدیک‌ترین K همسایگی<sup>۱۶</sup>

یک روش مدلسازی رفتار بی‌نظم، استفاده از روش نزدیک‌ترین همسایه است. در این روش در مرحله آموزش، همه نمونه‌ها در فضای هندسی به صورت بردارهای چند بعدی هستند. این فضا به برجسب‌های کلاس و موقعیت این نقاط تفکیک می‌شود. کارایی این دسته‌بندی کننده وابسته به معیار فاصله اقلیدسی است که استفاده می‌کند. روش نزدیک‌ترین K همسایه، یک گروه شامل K داده از مجموعه داده‌های آموزشی که نزدیک‌ترین داده‌ها به داده ورودی باشند را انتخاب کرده و براساس برتری دسته یا برجسب مربوط به آنها در مورد دسته داده آزمایشی مزبور تصمیم‌گیری می‌نماید. به عبارت ساده‌تر این روش



## مقایسه رگرسیون خطی چندگانه و الگوریتم‌های یادگیری ماشین... / سیف و یوسفی طزر جان

دسته‌ای را انتخاب می‌کند که در همسایگی انتخاب شده بیشترین تعداد داده، متناسب به آن دسته باشند، بنابراین دسته‌ای که از همه رده‌ها بیشتر در بین  $K$  نزدیکترین همسایه مشاهده شود، به‌عنوان دسته داده جدید در نظر گرفته می‌شود (فورد و همکاران، ۱۹۷۰).

### سوالات پژوهش

- ۱- آیا با استفاده از متغیرهای مالی و روش رگرسیون چندگانه خطی امکان پیش‌بینی میزان نگهداشت وجه نقد در شرکت‌های پذیرفته شده در بورس اوراق بهادار ایران وجود دارد؟
- ۲- آیا با استفاده از متغیرهای مالی و روش شبکه‌های عصبی امکان پیش‌بینی میزان نگهداشت وجه نقد در شرکت‌های پذیرفته شده در بورس اوراق بهادار ایران وجود دارد؟
- ۳- آیا با استفاده از متغیرهای مالی و روش نزدیکترین همسایه امکان پیش‌بینی میزان نگهداشت وجه نقد در شرکت‌های پذیرفته شده در بورس اوراق بهادار ایران وجود دارد؟
- ۴- آیا با استفاده از متغیرهای مالی و درخت تصمیم امکان پیش‌بینی میزان نگهداشت وجه نقد در شرکت‌های پذیرفته شده در بورس اوراق بهادار ایران وجود دارد؟
- ۵- آیا با استفاده از متغیرهای مالی و روش ماشین بردار پشتیبان امکان پیش‌بینی میزان نگهداشت وجه نقد در شرکت‌های پذیرفته شده در بورس اوراق بهادار ایران وجود دارد؟

### یافته‌های پژوهش

در این پژوهش سعی شده است با استفاده از روش‌های یادگیری ماشین تحت نظارت، میزان نگهداری وجه نقد شرکت‌ها پیش‌بینی شود. همچنین، با روش‌های رگرسیون چندمتغیره برای پیش‌بینی مقدار بازیابی اقدام سپس دقت مدل‌ها مقایسه گردید. در روش‌شناسی مبتنی بر متریک خطا  $RMSE$  و ضریب تعیین  $R^2$  تابعی از تفاوت بین مقادیر مشاهده شده و پیش‌بینی شده است. در همین حال،  $R^2$  نشان می‌دهد که چقدر خوب مدل رگرسیون با مقادیر مشاهده شده متغیر وابسته مطابقت دارد. از این رو،  $R^2$  مقادیر بالاتر مورد نظر است.  $RMSE$  معیار خطای جذر میانگین مربعات است و مقادیر پایین‌تر از ۵٪ نشان‌دهنده مقدار مطلوب است.

آمار توصیفی داده‌ها جدول ۲ آمار توصیفی داده‌ها را نشان می‌دهد. چون شرکت‌های استفاده شده در مدل‌های پژوهش یکسان‌اند، آمار توصیفی براساس مدل‌های پژوهش تفکیک نشده و به‌صورت یکجا ارائه شده است.

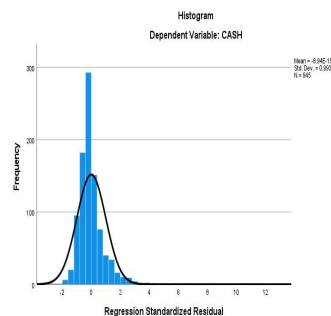
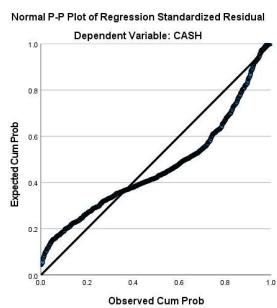
جدول ۲- آمار توصیفی داده‌های کمی پژوهش

ردیف	نماد متغیر	میانگین	انحراف معیار	تعداد
۱	CASH	۰/۰۴۳۸	۰/۰۵۲۴۱	۹۴۵
۲	CR	۱/۶۲۷۳	۱/۰۸۰۰۹	۹۴۵
۳	TANG	۰/۲۴۳۲	۰/۱۷۲۹۴	۹۴۵
۴	DIV	۰/۱۰۴۷	۰/۱۵۱۶۷	۹۴۵
۵	EPS	۱۳۰۵/۱۱۵۷	۳۰۴۲/۰۲۴۲۰	۹۴۵
۶	STD	۰/۴۹۹۷	۰/۱۹۴۷۵	۹۴۵
۷	SIZE	۱۴/۷۲۳۵	۱/۵۸۳۸۷	۹۴۵
۸	ROIC	۰/۱۵۲۰	۰/۱۹۳۹۹	۹۴۵
۹	AR	۰/۲۷۸۴	۰/۱۹۱۶۴	۹۴۵
۱۰	ROA	۰/۱۳۲۵	۰/۱۹۵۶۷	۹۴۵
۱۱	ROE	۰/۳۳۶۸	۰/۶۵۶۳۱	۹۴۵
۱۲	AP	۰/۱۸۸۸	۰/۱۵۹۹۹	۹۴۵
۱۳	SG	۶۰۰۷۷۴/۸۵۸۲	۶۱۷۹۵۱۲۶/۳۷۲۰۲	۹۴۵
۱۴	IE	۰/۰۳۳۴	۰/۰۳۳۶۱	۹۴۵

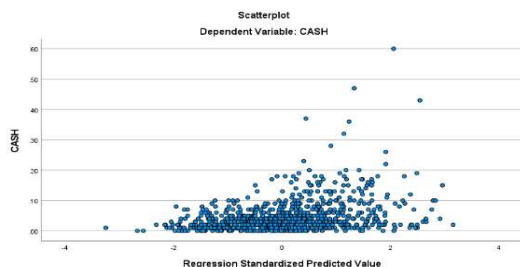
منبع: یافته‌های پژوهش

نتایج حاصل از روش رگرسیون چندگانه

ابتدا از روش رگرسیون چندگانه برای پیش‌بینی وجه نقد با استفاده از متغیرهای پیش‌بینی کننده مختلف انجام می‌شود. نمودار ۲ نشان می‌دهد که خطاها به‌طور معمول توزیع می‌شوند و رابطه خطی است.



## مقایسه رگرسیون خطی چندگانه و الگوریتم‌های یادگیری ماشین... / سیف و یوسفی طزرجان



### نمودار ۲- نمایش توزیع خطاها. منبع: یافته‌های پژوهش

نمودار ۲ مفروضات رگرسیون را برای مدل نشان می‌دهد. خطاها تقریباً عادی توزیع شده‌اند و میانگین خطاها تقریباً صفر است. این نشان می‌دهد که همسویی خطاها تقریباً دارای واریانس‌های مساوی هستند و نقاط پرت ناچیز است.

اگر درجه همبستگی بین متغیرها به اندازه کافی بالا باشد، هم‌خطی در تحلیل رگرسیون اتفاق می‌افتد که می‌تواند در برازش و تفسیر مدل رگرسیون مشکلاتی ایجاد کند. یکی از راه‌های تشخیص هم‌خطی بودن، استفاده از متریک معروف به ضریب تورم واریانس<sup>۱۷</sup> (VIF) است که همبستگی و قدرت همبستگی بین متغیرهای پیش‌بینی‌کننده را در یک مدل رگرسیونی اندازه‌گیری می‌کند. مقدار VIF از ۱ شروع می‌شود و حد بالایی ندارد. یک قانون کلی برای تفسیر ضریب تورم واریانس به شرح زیر است:

مقدار ۱ نشان می‌دهد که هیچ ارتباطی بین یک متغیر پیش‌بینی‌کننده معین و سایر متغیرهای پیش‌بینی‌کننده در مدل وجود ندارد. مقدار بین ۱ و ۵ نشان‌دهنده همبستگی متوسط بین یک متغیر پیش‌بینی‌کننده معین و سایر متغیرهای پیش‌بینی‌کننده در مدل است، اما این اغلب آنقدر شدید نیست که نیاز به توجه داشته باشد. مقدار بیشتر از ۵ نشان‌دهنده همبستگی شدید بالقوه بین یک متغیر پیش‌بینی‌کننده معین و سایر متغیرهای پیش‌بینی‌کننده در مدل است. در این مورد، تخمین ضرایب و مقادیر  $p$  در خروجی رگرسیون به احتمال زیاد غیر قابل اعتماد هستند (گویزانی<sup>۱۸</sup>، ۲۰۱۷). ضرایب همبستگی زوجی و نتایج ضریب تورم واریانس در جدول ۳ ارائه شده است که نشان‌دهنده عدم وجود مشکل چندخطی در میان متغیرها است.

جدول ۳- نتایج ضریب تورم واریانس و همبستگی میان متغیرهای ورودی و وجه نقد

متغیرها	ضریب تورم واریانس	ضریب همبستگی
CR	۲/۱۴۲	۰/۱۶۸
TANG	۱/۷۳۰	-۰/۱۴۳
DIV	۴/۳۹۵	-۰/۱۲۰
EPS	۱/۳۰۶	۰/۰۴۳
STD	۴/۱۹۴	-۰/۱۴۵
SIZE	۱/۲۷۷	-۰/۰۳۵
ROIC	۵/۸۹۰	۰/۱۴۴
AR	۱/۴۴۱	-۰/۱۴۱
ROA	۱/۳۰۷	۰/۰۸۲
ROE	۱/۹۳۷	۰/۰۳۲
AP	۲/۰۰۲	-۰/۰۳۱
SG	۱/۰۵۱	۰/۰۲۹
IE	۱/۹۷۶	-۰/۱۵۴

منبع: یافته‌های پژوهش

طبق جدول ۴ در الگوریتم‌های استفاده شده مقدار خطای جذر میانگین مربعات<sup>۱۹</sup> (RMSE) در همه موارد کمتر از ۰,۰۵ بوده است و اختلاف آنها بسیار کم بوده است، اما با توجه به معیارهای عملکرد پس از اعمال رگرسیون خطی چندگانه  $R^2$  کم است؛ لذا نتایج هنوز ناموفق هستند. بنابراین، می‌توان نتیجه گرفت که رگرسیون خطی چندگانه در پیش‌بینی مقادیر وجه نقد خوب نیست.

جدول ۴- خلاصه مدل رگرسیون خطی چندگانه

شماره مدل	پیش‌بینی کننده‌های مدل	$R^2$	RMSE
۱	SG, ROA, AR, CR, ROIC	۰/۰۵۶	۰/۰۵۱
۲	AP, SIZE, TANG, ROE	۰/۰۲۵	۰/۰۵۱
۳	IE, SG, ROE, EPS, STD	۰/۰۳۱	۰/۰۵۱
۴	AP, SG, TANG, DIV, ROIC	۰/۰۳۷	۰/۰۵۱

منبع: یافته‌های پژوهش

نتایج حاصل از شبکه عصبی چندلایه

ابتدا با استفاده از شبکه عصبی چندلایه مدل آموزش داده شد. در این مدل ۱۰ بار با تنظیمات

## مقایسه رگرسیون خطی چندگانه و الگوریتم‌های یادگیری ماشین... / سیف و یوسفی طزرجان

مختلف شبکه عصبی آموزش داده شد. ضریب تعیین  $R^2$  در شبکه عصبی با یک لایه پنهان  $0/0409$  می‌باشد که بیانگر عملکرد نامطلوب و قدرت پیش‌بینی‌کنندگی ضعیف مدل است. در شبکه عصبی با دو لایه پنهان ضریب تعیین  $R^2$   $0/99$  است و مقدار خطای جذر میانگین مربعات (RMSE)  $0/0017$  است؛ بنابراین عملکرد این مدل خوب بوده است. سه متغیر مهم و تاثیرگذار در شبکه عصبی با یک لایه پنهان و دو لایه پنهان سود هر سهم، تغییر سالانه در رشد فروش، سائز شرکت است.

### نتایج حاصل از نزدیک‌ترین همسایگی

در روش k نزدیک‌ترین همسایگی میزان ضریب تعیین  $R^2$  برابر با  $0/99$  است و مقدار خطای جذر میانگین مربعات (RMSE)  $0/00097$  است. نتایج نشان می‌دهد این مدل در پیش‌بینی میزان نگهداشت وجه نقد با دقت بالایی عمل می‌کند.

### نتایج حاصل از الگوریتم ماشین بردار پشتیبان

در روش الگوریتم ماشین بردار پشتیبان میزان ضریب تعیین  $R^2$  برابر با  $0/99$  است و مقدار خطای جذر میانگین مربعات (RMSE)  $0/0027$  است. نتایج نشان می‌دهد این مدل در پیش‌بینی میزان نگهداشت وجه نقد با دقت بالایی عمل می‌کند.

جدول ۵ خلاصه نتایج حاصل از الگوریتم‌های یادگیری ماشین جهت پیش‌بینی میزان نگهداشت وجه نقد در شرکت‌های پذیرفته شده در بورس را نشان می‌دهد که بیانگر عملکرد خوب و دقت بالای پیش‌بینی و میزان خطای کمتر از  $0/01$  در الگوریتم‌های K نزدیک‌ترین همسایه، ماشین بردار پشتیبانی و شبکه عصبی با دو لایه پنهان است.

جدول ۵- خلاصه یافته‌های الگوریتم‌های یادگیری ماشین به کار رفته در پژوهش

الگوریتم K نزدیک ترین همسایه	مدل ماشین بردار پشتیبانی	شبکه عصبی با یک لایه پنهان	شبکه عصبی با دو لایه پنهان	
$0/99$	$0/99$	$0/0409$	$0/99$	$R^2$
$0/00097$	$0/0027$	$0/00071$	$0/0017$	RMSE

برای ارزیابی روایی و پایایی روش‌های بکار گرفته شده مانند درخت تصمیم و شبکه عصبی و ...، می‌توان از روش‌های زیر استفاده کرد.

(۱) روایی (Validity) درخت تصمیم و شبکه عصبی و ... به معنای توانایی آنها در پیش‌بینی دقیق نتایج است. برای ارزیابی روایی، می‌توان از روش‌های مختلفی مانند مقایسه پیش‌بینی‌های درخت تصمیم و شبکه عصبی با داده‌های واقعی با استفاده از روش‌های آماری مانند ضریب همبستگی استفاده کرد.

۲) پایایی (Reliability) درخت تصمیم و شبکه عصبی به معنای پایداری و قابل اعتماد بودن آنها است. برای ارزیابی پایایی، می‌توان از روش‌های مختلفی مانند تکرار آزمایش‌ها و بررسی تفاوت نتایج استفاده کرد.

۳) برای ارزیابی روایی و پایایی درخت تصمیم و شبکه عصبی، می‌توان از داده‌های آموزش، داده‌های اعتبارسنجی و داده‌های آزمون استفاده کرد و نتایج را با یکدیگر مقایسه کرد. در این پژوهش نتایج روش‌های مختلف یادگیری ماشین استفاده شد که نتایج یکدیگر را تایید کردند. همچنین داده‌ها به دو بخش یادگیری و آزمون تقسیم شدند که دقت قابل قبول در داده‌های آزمون بیانگر پایایی مدل‌ها است. نتایج به تفصیل در هر یک از روش‌ها و در جدول ۴ و ۵ ارائه شده است.

### نتیجه‌گیری و پیشنهادها

تصمیم شرکت‌ها برای نگهداری وجه نقد موضوعی رایج در امور مالی شرکت‌ها است. مدیریت وجه نقد، از مهم‌ترین وظایف در فرایند مدیریت مالی است؛ زیرا از یک سو کمبود وجه نقد موجب می‌شود که شرکت در برآورده کردن نیازهای روزمره خود با مشکل روبه‌رو شود و از سوی دیگر، نگهداری سطح بالایی از وجه نقد برای شرکت هزینه فرصت به همراه دارد (گارسیا تروئل و همکاران، ۲۰۰۸).

این پژوهش شامل ۹۴۵ داده شرکت‌های پذیرفته شده در بورس اوراق بهادار ایران از سال ۱۳۹۲ تا ۱۴۰۰ است. با استفاده از ۱۳ متغیر ورودی و الگوریتم‌های شبکه عصبی، نزدیک‌ترین همسایه و ماشین بردار پشتیبان و مدل رگرسیون خطی چندگانه به پیش‌بینی میزان وجه نقد شرکت‌ها پرداخته است. یافته‌ها نشان می‌دهد که الگوریتم‌های شبکه عصبی با دو لایه پنهان، نزدیک‌ترین همسایه و ماشین بردار پشتیبان با ضریب تعیین ۰/۹۹ و خطای زیر ۰/۰۱ به خوبی قادر به پیش‌بینی میزان نگهداشت وجه نقد در شرکت‌ها بوده‌اند ولی روش خطی رگرسیون چندگانه عملکرد ضعیفی داشته است. همچنین، سه متغیر سود هر سهم، نسبت دارایی‌های جاری به بدهی‌های جاری و نسبت بدهی کوتاه‌مدت به کل دارایی‌ها تاثیرگذاری و اهمیت بیشتری در همه الگوریتم‌ها داشته‌اند. بر اساس یافته‌های ما، این پژوهش پیامدهای قابل توجهی برای مدیران شرکت‌ها و محققان دارد. مدیران می‌توانند از این اطلاعات برای تعیین میزان نگهداری وجه نقد شرکت‌ها و تعیین سیاست‌های شرکتی استفاده کنند. در همین حال، محققان می‌توانند از مدل‌های رگرسیون و سایر الگوریتم‌های یادگیری استفاده شود یا می‌توان دوره مطالعه را گسترش داد و یا متغیرهای کلان اقتصادی، مانند رشد تولید ناخالص داخلی، نرخ بهره و قیمت نفت را می‌توان به مطالعات اضافه کرد.

## منابع

- ۱) آقایی، محمدعلی. نظافت، احمدرضا. ناظمی اردکانی، مهدی. جوان، علی‌اکبر (۱۳۸۸) بررسی عوامل موثر بر نگهداری موجودی‌های نقدی در شرکت‌های پذیرفته شده در بورس اوراق بهادار تهران، پژوهش‌های حسابداری مالی سال اول پاییز و زمستان، شماره ۱ و ۲.
- ۲) بیورانی، جباری خامنه‌ای (۱۳۹۲) “تحلیل آماری به کمک نرم افزار SPSS و Minitab” چاپ دوم، انتشارات دانشگاه کردستان، سنندج.
- ۳) حساس یگانه، یحیی، جعفری، علی و رسائیان، امیر (۱۳۹۰) عوامل تعیین کننده سطح نگهداشت وجه نقد شرکت‌ها در بورس اوراق بهادار تهران، فصلنامه علمی پژوهشی حسابداری مالی، ۳۹، ۶۶-۶۹.
- ۴) فروغی، داریوش، فرزادی، سعید (۱۳۹۳) بررسی تاثیر تغییرات جریان‌های نقدی بر سطح نگهداشت وجه نقد، فصلنامه علمی - پژوهشی مدیریت دارایی و تامین مالی، سال دوم شماره چهارم، شماره پیاپی (۴) بهار، صص ۲۱-۳۶.
- ۵) رهنمای رودپشتی، فریدون و کیایی علی (۱۳۸۷) بررسی و تبیین سرمایه در گردش در شرکت‌های پذیرفته شده در بورس اوراق بهادار تهران، دانش و پژوهش حسابداری.
- ۶) حاجیها، زهره. رجب‌دري، حسین (۱۳۹۶) بررسی تأثیر ویژگی‌های کمیته حسابرسی بر سطح نگهداری وجه نقد، دانش حسابداری، ۲۹ رتبه ب (وزارت علوم 21) ISC/صفحه - ۶۳ تا ۸۳.
- ۷) کاشانی‌پور، محمد. تقی زاده (۱۳۸۸) بررسی اثر محدودیت‌های تامین مالی بر حساسیت جریان‌های نقدی، تحقیقات حسابداری، ش ۲ صص ۷۲-۹۳.
- ۸) مشکی میاوقی، مهدی و صنایعی ماسوله، مهیار (۱۳۹۵) بررسی تاثیر انحراف از سطح بهینه نگهداشت وجه نقد بر ارزش وجه نقد. فصلنامه راهبرد مدیریت مالی، سال چهارم، شماره ۱۳.
- ۹) نیک‌بخت، محمدرضا، صابرمهانی، معصومه، و دلدار، مصطفی (۱۴۰۰) آزمون اثربخشی تمرکز مشتری بر سرعت تعدیل نگهداشت وجه نقد در بورس اوراق بهادار تهران. راهبرد مدیریت مالی، ۹(۳۲)، ۷۸-۵۵.
- 10) Benkraiem R, Lakhel F, Zopounidis C (2020) International diversification and corporate cash holding behavior: What happens during economic downturns? *J Econ Behav Organ* 170:362–371. <https://doi.org/10.1016/j.jebo.2019.12.016>.
- 11) Batuman B, Yildiz Y, Karan MB (2021) The impact of global financial crisis on corporate cash holdings: evidence from Eastern European countries. *Borsa Istanbul Rev.* <https://doi.org/10.1016>.
- 12) Bigelli M, Sanchez-Vidal J (2012) Cash holdings in private firms. *J Bank Finance* 36(1):26–35. <https://doi.org/10.1016/j.jbankfin.2011.06>.

- 13) Bhuiyan MBU, Hooks J (2019) Cash holding and over-investment behavior in firms with problem directors. *Int Rev Econ Financial* 61:35–51.
- 14) Boubakri N, Ghouli S, Saffar W (2013) Cash holdings of politically connected firms. *J Multinatl Finance Manag* 23(4):338–355.
- 15) Cai W, Zeng C, Lee E, Ozkan N (2016) Do business groups affect corporate cash holdings? Evidence from a transition economy. *China J Acc Res* 9:1–24.
- 16) Cambrea DR, Calabro A, Rocca M, Paolone F (2021) The impact of boards of directors' characteristics cash holdings in uncertain times. *J Manag Gover.*
- 17) Chen D, Li S, Xiao JZ, Zou H (2014) The effect of government quality on corporate cash holdings. *J Corp Finance* 27:384–400. [https:// doi. org/ 10. 1016/j. jcorp fin. 2014. 05. 008.](https://doi.org/10.1016/j.jcorpfin.2014.05.008)
- 18) Diaw A (2021) Corporate cash holdings in emerging markets. *Borsa Istanbul Rev* 21(2) 139–148. [https:// doi. org/ 10. 1016/j. bir. 2020. 09. 005](https://doi.org/10.1016/j.bir.2020.09.005)
- 19) Donepudi PK, Banu MH, Khan W, Neogy TP, Asadullah ABM, Ahmed AAA (2020) Artificial intelligence and machine learning in treasury management: a systematic literature review. *Int J Manag* 11(11):13–22.
- 20) Drobetz W, Gruninger MC (2007) Corporate cash holdings: evidence from Switzerland. *Fin Mark Portfolio Mgmt* 21(3):293–324.
- 21) Ferreira, M. A., and A. Vilela. (2004). Why Do Firms Hold Cash? Evidence from EMU Countries, *European Financial Management*, Vol. 10, No. 2, 295.
- 22) Garcia-Teruel PJ, Martinez-Solano P (2008) On the determinants of SME cash holdings: evidence from Spain. *J Bus Financial Acc* 35(1–2):127–149.
- 23) Gholamzadeh Mohammadreza, Faghani Mahdi, Pifeh Ahmad (2021). Implementing machine learning methods in the prediction of the financial constraints of the companies listed on Tehran's stock exchange. *International Journal of Finance and Managerial Accounting*. Vol. 6, No.20: 131-147.
- 24) Gill, A., & Shah, C. (2012). Determinants of corporate cash holdings: Evidence from Canada. *International journal of economics and finance*, 4(1), 70-79.
- 25) Guizani M (2017) the financial determinants of corporate cash holdings in an oil rich country: evidence from Kingdom of Saudi Arabia. *Borsa Istanbul Rev* 17(3):133-143.
- 26) Hornik, K., Stinchcombe, M., and White, H. 1989. Multilayer feed forward networks are universal approximators. *Neur. Net.* 2: 5. 359-366.
- 27) Jani, E, Hoesli, M, Bender, A, Wd. (2004). Corporate Cash Holdings and Agency Conflicts, available: [atwww.ssrn.com,id=563863](http://www.ssrn.com/id=563863).
- 28) Kim HJ, Han SH, Mun S (2021) Analyzing the effects of terrorist attacks on the value of cash holdings. *Financial Res Lett*. [https:// doi. org/ 10. 1016/j. frl. 2021.](https://doi.org/10.1016/j.frl.2021)



- 29) Lee, E., & Powell, R. (2011). Excess cash holdings and shareholder value. *Accounting & finance*, 51(2), 549-574.
- 30) Lozano MB, Yaman S (2020) The European financial crisis and firms' cash holding policy: an analysis of the precautionary motive. *Glob Pol* 11(S1):84–94. <https://doi.org/10.1111/1758-5899.12768>.
- 31) Manoel AAS, Moraes MBC, Santos DFL, Neves MF (2018) Determinants of corporate cash holdings in times of crisis: insights from Brazilian sugarcane industry private firms. *Int Food Agribus Manag Rev* 21(2):201–217.
- 32) Mohammadi M, Kardan B, Salehi M (2018) The relationship between cash holdings, investment opportunities and financial constraint with audit fees. *Asian J Account Res* 3(1):15–27.
- 33) Ozlem Şirin , Faruk Omer (2022) Predicting cash holdings using supervised machine learning algorithms, *Financial Innovation*, 8:44 <https://doi.org/10.1186/s40854-022-00351-8>.
- 34) Ozkan A (2001) Determinants of capital structure and adjustment to long-run target: evidence from UK company panel data. *J Bus Financial Acc* 28(1):175–198.
- 35) Ozkan A, Ozkan N (2004) Corporate cash holdings: an empirical investigation of UK companies. *J Bank Finance* 28(9):2103–2134. <https://doi.org/10.1016/j.jbankfin.2003.08.003>.
- 36) Opler T, Pinkowitz L, Stulz H, Williamson R (1999) The determinants and implications of corporate cash holdings. *J Financial Econ* 52:3–46.
- 37) Rafi MM, Wahab, MT, Khan, MB, Raza H (2020) ATM cash prediction using time series approach. In: Paper presented at the 3<sup>rd</sup> international conference on computing, mathematics and engineering technologies (iCoMET), Sukkur IBA University, Pakistan, 29–30.
- 38) Sarfraz M, Shah SGM, Ivascu M, Quereshi MAA (2020) Explicating the impact of hierarchical CEO succession on small-medium enterprises' performance and cash holdings. *Int J Financial Econ*. <https://doi.org/10.1002/ijfe.2289>.
- 39) Schauten MB, Dijk D, van der Wall JP (2011) Corporate governance and the value of excess cash holdings of large European firms. *Eur Financial Manag* 19(5):991–1016.
- 40) Song K, Lee Y (2012) Long-term effects of a financial crisis: Evidence from cash holdings of East Asian firms. *J Financial Quant Anal* 47(3):617–641. <https://doi.org/10.1017/S0022109012000142>.

- 41) Uyar A, Kuzey C (2014) Determinants of corporate cash holdings: Evidence from the emerging market of Turkey. *Appl Econ* 46(9):1035–1048. <https://doi.org/10.1080/00036846.2013.866203>.
- 42) Vapnik, V. N. (1995). *The Nature of Statistical Learning Theory*. Springer Verlag New York.
- 43) Wu H, Chen J, Wang P (2021) Cash holdings prediction using decision tree algorithms and comparison with logistic regression model. *Cybern Syst* ÷52 (8) 689–704. <https://doi.org/10.1080/01969722.2021.1976988>.

یادداشت‌ها:

- 
- 1 Pawlina & Renneboog
  - 2 Lee & Powell
  - 3 Pinkowitz
  - 4 Bigelli and Sánchez-Vidal
  - 5 Boubakri, Ghoul, Saffar
  - 6 Diaw
  - 7 Ferreira and Vilela
  - 8 Lozano and Yaman
  - 9 Gill & Shah
  - 10 Manoel et al
  - 11 Ozkan
  - 12 Classifier
  - 13 Logistic regression
  - 14 Support Vector Machine
  - 15 Multi-layer neural networks
  - 16 K-Nearest Neighbors
  - 17 Variance inflation factor
  - 18 Guizani
  - 19 Root Mean Square Error

## Comparison of multiple linear regression and machine learning algorithms in predicting cash holdings

Samira Seif<sup>1</sup>

Receipt: 07/06/2023    Acceptance: 19/04/2023    Mostafa Yousofi Tezerjan<sup>2</sup>

### Abstract

In recent years, in the financial literature, more attention has been paid to the level of cash holding of companies. So; Forecasting is important to determine the optimal level of cash holding. In this research, using linear and non-linear methods and 13 influential input variables, the amount of cash in 103 companies admitted to the Iran Stock Exchange during the years 2013 to 2021 has been predicted. The methods used include multiple linear regression (MLR), k nearest neighbor (KNN), support vector machine (SVM) and multi-layer neural networks (MLNN) for prediction. The results show that the traditional method of multiple linear regression has not been successful in predicting cash, but machine learning algorithms have been superior with an accuracy of 0.99. The variables of profit per share, the ratio of current assets to current liabilities and the ratio of short-term debt to total assets have had a greater impact in all algorithms. Therefore, managers can use advanced machine learning algorithms to predict the future cash flow of companies.

### Key words

Forecasting, Cash Holdings, Multiple Linear Regression, Machine Learning Algorithms.

1-Assistant Professor, Department of Accounting, Payame Noor University, Tehran, Iran (Corresponding Author). samiraseif@pnu.ac.ir

2-Assistant Professor, Department of Industrial Engineering, University of Applied Science and Technology (UAST), Karaj, Iran. yousofi@uast.ac.ir