



رتبه بندی میزان تاثیر اقلام کلیدی ترازنامه ای و نسبت‌های سودآوری در انتخاب پرتفوی بهینه (با استفاده از تکنیک های داده کاوی)

امیررضا کیقبادی^۱

سمیه فتحی^۲

سمیرا سیف^۳

تاریخ دریافت: ۱۳۹۴/۰۷/۱۸ تاریخ پذیرش: ۱۳۹۴/۰۹/۱۴

چکیده

انتخاب پرتفوی بهینه، یکی از موضوعات مهم در ادبیات مالی محسوب می‌شود که هدف‌های حداکثر کردن بازده آتی و حداقل نمودن ریسک سرمایه‌گذاری را به همراه دارد. اساساً سرمایه‌گذاران برای انتخاب پرتفوی بهینه و کارا شرکت‌هایی را انتخاب می‌کنند که جزء شرکت‌های برتر بورس باشند و همچنین شرکت‌های که در رده‌های بالای این رتبه بندی قرار دارند. رتبه‌بندی می‌تواند برحسب معیارهای مختلفی انجام گیرد. یکی از سودمندترین رتبه‌بندی‌ها، می‌تواند استفاده از نسبت‌های سودآوری و اقلام کلیدی ترازنامه ای باشد. در این تحقیق نسبت‌های سودآوری و اقلام کلیدی ترازنامه‌ای ۱۵۰ شرکت برتر بورس طی ۳ سال استخراج شده است و در هر سال ۵۰ شرکت با توجه نسبت‌های مالی بالاتر به دو دسته موفق و برتر تقسیم بندی شده‌اند. سپس با استفاده از ۴ مدل مبتنی بر تکنیک داده‌کاوی و همچنین توان یادگیری هر یک از مدل‌ها به بررسی ترتیب اهمیت این اقلام و نسبت‌ها جهت انتخاب پرتفوی بهینه پرداخته شده است. نتایج حاکی از آن است که اقلام کلیدی ترازنامه ای و نسبت‌های سودآوری هر دو برای ارائه پرتفوی حائز اهمیت بوده اما ارجحیت تاثیر گذاری آنان در هر مدل ترتیبی متفاوت است؛ اما برخی از این اقلام مانند جمع کل دارایی‌ها، درصد سود به درآمد، درصد سود عملیاتی به درآمد و نسبت قیمت به سود هر سهم (P/E) در تمام مدل‌ها به عنوان یک فصل مشترک از ارجحیت بالاتری برخوردارند که می‌تواند به عنوان مبنایی مناسب مورد توجه سرمایه‌گذاران قرار گیرد.

واژه‌های کلیدی: پرتفوی، اقلام کلیدی ترازنامه ای، معیارهای سودآوری، داده کاوی.

۱- استادیار، گروه حسابداری، واحد تهران مرکزی، دانشگاه آزاد اسلامی، تهران، ایران، نویسنده اصلی و مسئول مکاتبات.
a.keyghobadi@iauctb.ac.ir

۲- کارشناس ارشد حسابداری، مدرس دانشگاه پیام نور واحد بروجرد.

۳- کارشناس ارشد حسابداری، دستیار علمی دانشگاه پیام نور واحد نهاوند.

۱- مقدمه

برای انتخاب اوراق بهادار، ضروری است که فرآیند تصمیم گیری در اوراق بهادار شامل تجزیه و تحلیل اوراق بهادار و مدیریت پرتفوی طی گردد. اساسا سرمایه گذاران برای انتخاب پرتفوی بهینه و کارا شرکتهای را انتخاب می کنند که جزء شرکتهای برتر بورس باشد و همچنین شرکت های که در رده های بالای این رتبه بندی قرار دارند.

برای انتخاب اوراق بهادار، ضروری است که فرآیند تصمیم گیری در اوراق بهادار شامل تجزیه و تحلیل اوراق بهادار و مدیریت پرتفوی طی گردد. انتخاب ابزار و تکنیک هایی که بتواند سبد سهام مناسب را تشکیل دهد، آرزوی دنیای سرمایه گذاری است. در بازار سرمایه، روشها و ابزارهای مختلف مورد تجزیه و تحلیل قرار می گیرند. افزایش پیچیدگی روزافزون سیستم های اطلاعاتی، حجم عظیم داده های مورد نیاز و همچنین پیچیدگی مبادلات و معاملات تجاری شرکتها باعث شده تا مدیران مالی و حسابرسان قادر به استفاده از روش های سنتی به منظور تصمیم گیری و اظهار نظر مناسب نسبت به اطلاعات ارائه شده نباشند. یکی از فناوری هایی که در عصر اطلاعات می تواند استفاده کنندگان از اطلاعات مالی، اعتبار دهندگان، تحلیل گران مالی و سایر ذینفعان را در امر تصمیم گیری یاری رساند، به کارگیری روش های داده کاوی است. دانش داده کاوی یکی از ده دانش در حال توسعه ای است که دهه آینده را با انقلاب تکنولوژیک مواجه خواهد ساخت به گونه ای که دیگر در دنیای توسعه یافته امروزی، مکان و موضوعی بدون بهره از دانش داده کاوی یافت نخواهد شد. در این مقاله نظر بر این است که با استفاده از مدل های منتج از داده کاوی و اقلام کلیدی ترازنامه ای و نسبت های سودآوری پرتفوی کارا و بهینه بتوان تشکیل داد.

۲- پیشینه تحقیق

لوراشی^۱ (۱۹۹۵) و همکاران برای انتخاب سبد سهام بدون محدودیت، با استفاده از مدل های جزیره ای (مدل های مبتنی بر الگوریتم ژنتیک برای استفاده در جوامع جدا از یکدیگر که در آنها، داده های برانده تر در میان جوامع کوچک می نمایند) اقدام به حل مدل سبد سهام بدون محدودیت نمودند. در مطالعه جنیفر پاورز^۲ و همکاران (۲۰۰۰)، از رویکرد تحلیل پوششی داده ها جهت تشکیل پرتفوی، از یک لیست شامل ۱۸۵ شرکت استفاده شده است. از میان این تعداد، تنها ۱۴ عدد به عنوان شرکت کارآمد، ارزیابی شده اند. در مطالعه آنان، از ۸ متغیر استفاده شده است که ۵ متغیر خروجی شامل نرخ بازده یک ساله، سه ساله، پنج ساله، ده ساله و سود هر سهم هستند. متغیرهای نسبت قیمت به درآمد، ضریب بتا و ضریب سیگما به عنوان متغیرهای ورودی در این پژوهش به کار رفته است.

در مطالعه آنان میزان توانایی اوراق بهادار کارآمد را در شرایط ناپایدار بررسی کرده اند، به این معنا که اوراق بهادار کارآمد هنگام تغییرات نامطلوب تا چه مدت می توانند کارا باقی بمانند. پاورز و همکاران اعتقاد دارند که مزایای روش اتخاذ شده آنان، یعنی تحلیل پوششی داده ها، آن است که می توان به نحوی میزان کسری یا مازادی را که اوراق بهادار ناکارآمد برای تبدیل به کارآمد شدن دارند نیز محاسبه کرد. این دو

پژوهشگر، در مقاله خود تأکید دارند که برای بهره‌گیری از این شیوه به صورت عملی در یک سرمایه‌گذاری باید حتماً به ضرایب وزنی توجه شود.

سویل^۳ و همکاران (۲۰۰۶) در ترکیه نیز اقدام به تشکیل پرتفویی با استفاده از الگوهای کمی ارزیابی کارایی کردند. آنان در رویکرد خود از دارایی خالص (ارزش ویژه)، نسبت قیمت به درآمد، نسبت قیمت به جریان نقدی، نسبت ارزش بازار هر سهم به ارزش دفتری هر سهم و حجم داد و ستد به عنوان متغیرهای ورودی استفاده کردند و از سود ویژه، سیستم ارزیابی کارایی EMS به عنوان متغیرهای خروجی مسئله استفاده کردند. پس از تشکیل پرتفوی، بازده آن با شاخص بازار مقایسه شد. از سال ۲۰۰۲ تا سال ۲۰۰۵، ۱۴ پرتفوی تشکیل و بازده آن محاسبه و با شاخص بازار در همان دوره مقایسه شده است. نتیجه این بود که بازده پرتفوی تشکیل شده با استفاده از تحلیل پوششی داده بیشتر از بازده بازار بوده است.

لوپز^۴ و همکاران (۲۰۰۸) در پژوهشی اقدام به استفاده از روشهای ارزیابی کارایی به عنوان استراتژی انتخاب سهام در بازار سهام برزیل کردند. آنها از نسبت قیمت به سود هر سهم، ضریب بتا و نوسان‌پذیری بازده هر سهم به عنوان متغیر ورودی و از درآمد هر سهم، از بازده ۱۲، ۳۶ و ۶۰ ماهه به عنوان متغیرهای خروجی روش ارزیابی کارایی در یک بازه زمانی ۱۰ ساله یا ۱۲۰ ماه استفاده کردند. آنها دریافتند که پرتفوی ساخته شده با استفاده از روشهای ارزیابی کارایی نسبت به دو شاخص بازار برزیل عملکرد بهتری را ارائه کرده است.

راعی (۱۳۷۷) تحقیقی با هدف دستیابی به یک مدل سرمایه‌گذاری مناسب در سبد سهام انجام داد. مطالعات انجام شده، نشان داد که استفاده از شبکه‌های عصبی مصنوعی در تشکیل سبد سهام نسبت به مدل مارکویتز می‌تواند مؤثر باشد.

جمالی و همکاران (۱۳۹۱) تحقیقی با هدف دستیابی به یک مدل سرمایه‌گذاری مناسب در سبد سهام با استفاده از روش DIA5 روش بهبود یافته تحلیل پوششی داده‌ها انجام دادند. نتایج تحقیق حاکی از آن است که با استفاده از اقلام کلیدی ترانزنامه ای و معیارهای سودآوری جهت بدست آوردن پرتفوی و مقایسه تطبیقی آن با شاخص‌های نقدینگی می‌توان پرتفوی کارای ارائه نمود.

۳- چارچوب نظری

تخصیص منابع یکی از مهمترین مسائل روز در ایران است. در اغلب موارد پروژه‌های سرمایه‌گذاری بسیار مفیدی وجود دارند ولی امکان دسترسی به منابع مالی برای آنها وجود ندارد. در یک بازار سرمایه کارآ از بعد عملیاتی، سرمایه در اختیار بهترین گزینه‌های سرمایه‌گذاری قرار می‌گیرد و اولویت‌های بعدی، سایر منابع را به خود اختصاص می‌دهد. ورود به بازار سرمایه و استفاده از ابزارهای هوشمند جهت کسب بازدهی بیشتر گامی در راستای کارآتر شدن بازار است. کارآئی بازار، تخصیص منابع مناسب را به همراه خواهد داشت. این امر هم از بعد شخصی و هم از بعد ملی حائز اهمیت است زیرا:

(۱) استفاده از ابزارهای هوشمند، می‌تواند بازدهی بیشتر را در بازار نصیب سرمایه‌گذار نماید.

- (۲) استفاده از ابزارهای مناسب و بازدهی بیشتر برای سرمایه گذار، تخصیص منابع بهتر را به همراه دارد و این امر در سطح ملی، می تواند حافظ منافع جامعه شود.
- داده کاوی یک فعالیت میان رشته‌ای و ماحصل رویارویی و هم‌افزایی علوم مختلفی چون آمار، یادگیری ماشینی، هوش مصنوعی، تکنولوژی پایگاه داده و مصورسازی (توصیف و نمایه‌سازی) می‌باشد (نخعی‌زاده، ۱۳۸۶ به نقل از نجاری، ۱۳۸۹). در عین حال این محصول مشترک وجوه افتراق مشخصی با پدیدآورندگان خود دارد. یکی از کاربردهای داده کاوی در زمینه اوراق بهادار و کمک به سرمایه گذاران برای اخذ تصمیمات بهینه اقتصادی می باشد.
- داده کاوی در بسیاری از شاخه‌ها همچون بازاریابی، امور مالی، بانکداری، تولید، پزشکی، مدیریت ارتباط با مشتری، ردیابی، پیش بینی خرابی‌ها، آموزش سازمانی و... کاربرد دارد. که می‌توان به موارد زیر اشاره کرد:
- (۱) پایگاه داده عظیم و بسیاری وجود دارند.
 - (۲) اطلاعات تجاری ارزشمندی می‌تواند از این پایگاه داده استخراج شوند.
 - (۳) استفاده از روش‌های سنتی گذشته برای پشتیبانی تصمیم و تحلیل‌ها اجرا شدنی نیست.
 - (۴) تحلیل‌های انسانی تحت تأثیر ابعاد و حجم داده‌ها قرار می‌گیرد.
 - (۵) متدهای آماری سنتی رتبه قادر به رتبه‌بندی نیستند و نیاز به کارشناسان و تحلیلگران مهم و قابل توجه دارد.

در نتیجه ما می‌توانیم از مدل‌های مختلفی پیش‌بینی پرتفوی کارا استفاده کنیم که این مدل‌ها به دو گروه اصلی تقسیم می‌شوند که عبارتند از:

گروه اول؛ مدل‌های پارامتریک:

- (۱) مدل احتمال خطی
- (۲) مدل لاجیت و پروبیت
- (۳) مدل‌های تحلیل متمایز کننده

گروه دوم؛ مدل‌های ناپارامتریک:

- (۱) برنامه‌ریزی خطی
- (۲) شبکه‌های عصبی
- (۳) درخت‌های تصمیم
- (۴) مدل نزدیک‌ترین همسایگی
- (۵) فراگرد تجزیه و تحلیل سلسله مراتبی
- (۶) سیستم‌های خبره
- (۷) الگوریتم ژنتیک

در این مقاله نظر بر این است که با استفاده از مدل‌های منتج از داده کاوی و اقلام کلیدی ترازنامه ای و نسبت های سودآوری پرتفوی کارا و بهینه از بین شرکت های برتر و موفق بورس در طی سه سال ۱۳۸۸ و ۱۳۸۹ و ۱۳۹۰ بتوان تشکیل داد.

۴- فرضیه‌ی های پژوهش

بررسی فرضیات تحقیق در داده‌کاوی می‌باشد پژوهش‌های کاربردی از اهمیت ویژه‌ای برخوردار است. در این پژوهش هدف اصلی تحقیق، ارائه نسبت های برتر سودآوری برای پرتفوی کارا با استفاده از تکنیک‌های از این جهت فرضیات تحقیق متناسب با این هدف مورد توجه قرار گرفته و بررسی گردید.

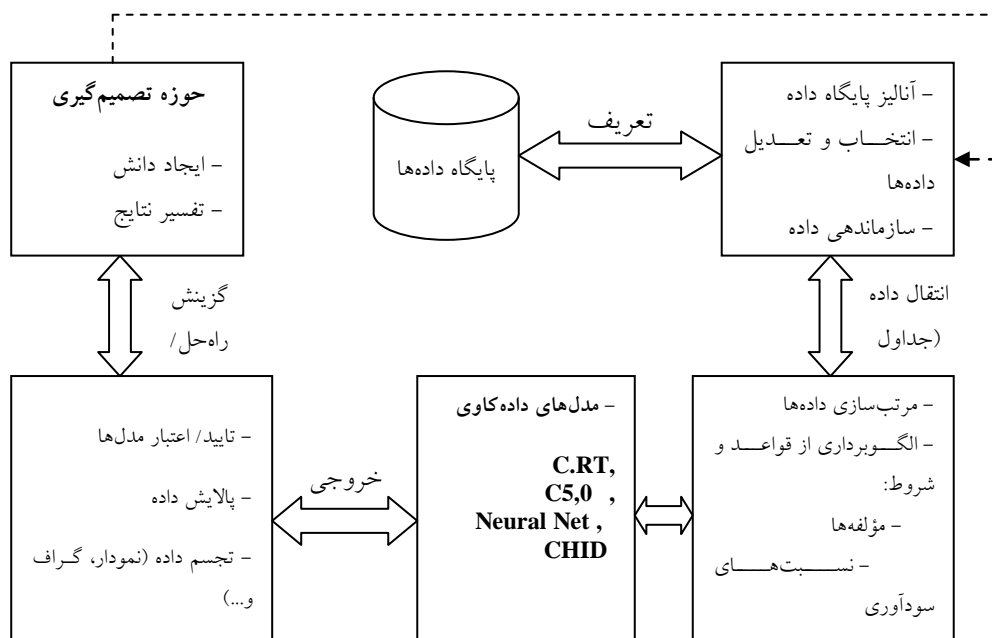
فرضیه ۱: اقلام کلیدی ترازنامه ای در مقایسه با نسبت های سودآوری در ارائه پرتفوی بهینه برتری دارند.

فرضیه ۲: نسبت های سودآوری در مقایسه با اقلام کلیدی ترازنامه ای در ارائه پرتفوی بهینه برتری دارند.

۵- روش پژوهش

این تحقیق با توجه به نتایجی که می‌تواند به همراه داشته باشد یک تحقیق بنیادی می‌باشد. زیرا درصد شناسای معیارهایی برای تهیه پرتفوی بهینه جهت سرمایه گذاری می‌باشد. از طرف دیگر با توجه به کاربرد این تحقیق در ارائه پرتفوی، یک تحقیق کاربردی می‌باشد. از نظر روش تحقیق با توجه به ماهیت پژوهش در حوزه علوم مالی، تحقیق حاضر از روش پیمایشی استفاده می‌کند. مراحل اجرایی و گام‌های اساسی در اجرای این پژوهش به صورت زیر قابل خلاصه شدن می‌باشد:

- ۱) تعیین اقلام کلیدی ترازنامه ای و نسبت های سودآوری
- ۲) جمع‌آوری داده از پایگاه داده‌های موجود (سازمان بورس اوراق بهادار)
- ۳) تعیین ۲۰ شرکت برتر هر سال به عنوان شرکت برتر و ۳۰ شرکت مابقی به عنوان شرکت موفق بر اساس نسبت های سودآوری بالاتر
- ۴) تقسیم داده‌های نمونه به دو مجموعه داده‌های آموزشی و داده‌های تست تعیین
- ۵) ساخت مدل‌ها با استفاده از داده‌های آموزشی
- ۶) آزمون مدل‌ها با مجموعه داده‌های تست
- ۷) بررسی دقت و سنجش اعتبار مدل‌ها در تعیین پرتفوی کارا
- ۸) ارائه بهترین الگو جهت تعیین پرتفوی کارا بر اساس تحقیق صورت گرفته مدل مفهومی تحقیق را می‌توان به صورت شکل ۱ در نظر گرفت.



شکل شماره ۱ - مدل مفهومی تحقیق (منبع یافته‌های پژوهشگر)

۶- گردآوری داده‌ها

برای اقلام کلیدی ترانزنامه ای از پرسشنامه استفاده شده است. سطح سنجش متغیرهای مورد بررسی عمدتاً رتبه‌ای (شبه فاصله‌ای) است. سؤال‌های پرسشنامه این پژوهش از نوع پنج گزینه‌ای خیلی کم تا خیلی زیاد (طیف لیکرت) است، که به ترتیب امتیازی از یک تا پنج به گزینه‌های بسیار کم، کم، متوسط، زیاد، بسیار زیاد داده شده است. بنابراین وزن‌های ۱ تا ۳ تأیید کننده فرض صفر، یعنی ناقص روابط بین متغیرها هستند و وزن‌های ۴ و ۵ رد کننده فرض صفر، یعنی تأیید کننده روابط بین متغیرها می‌باشند. برای بررسی فرضیه‌ها میانگین پاسخ‌هایی که توسط اعضای نمونه به سؤالات داده شده است محاسبه و با مطالب بیان شده مقایسه می‌شود.

برای تعیین پایایی از روش آزمون مجدد استفاده شد. به عبارت دیگر، بین اعضای نمونه مقدماتی دوبار پرسشنامه در زمان‌های متفاوت توزیع و از مقایسه نتایج آن مشخص شد که انحراف قابل ملاحظه‌ای وجود ندارد. همچنین در پژوهش انجام شده با توجه به نمونه مقدماتی، میزان آلفای کرونباخ 0.85 بدست آمد که به پایایی بالای پرسشنامه به کار برده شده اشاره دارد.

که در سطح اطمینان 0.99 و $0.1 = \alpha$ مقدار محاسبه شده $17.25 = \chi^2$ و مقدار استخراج شده جدول $49.6 = \chi^2$ می‌باشد.

برای تعیین روایی پرسشنامه ابتدا پرسشنامه را در سطح محدود توزیع کرده و سپس با انجام مصاحبه در نمونه کوچک پیش آزمون شده است. همچنین برای بالا بردن روایی نظرات استادان دانشگاه و برخی صاحب‌نظران دیگر اخذ گردید و پیشنهادهای آنان در مورد سؤالات پرسشنامه اعمال گردید. در نتیجه برخی از سؤالات حذف، برخی تعدیل و یا به پرسشنامه اضافه شد.

لازم به ذکر است که نسبت‌های سودآوری حسابداری که عبارتند از سود هر سهم، سود تقسیمی هر سهم، درصد سود تقسیمی، حاشیه سود، نرخ بازده سرمایه‌گذاری و نرخ بازده حقوق صاحبان سهام و همچنین نسبت قیمت به سود هر سهم انتخاب شده‌اند. در جدول شماره (۱) خلاصه شده است.

جدول شماره ۱ - متغیرهای پژوهش

نسبت‌های سودآوری II		اقدام کلیدی تراز نامه ای BI	
سود(زیان) عملیاتی	۱	حسابها و اسناد دریافتی تجاری	۱
سود(زیان) ویژه پس از کسر مالیات	2	جمع دارایی‌های جاری به‌استثناء موجودی کالا	۲
درصد سود به درآمد	3	موجودی مواد و کالا	3
درصد سود به سود ناویژه	4	خالص دارایی‌های ثابت	4
درصد سود عملیاتی به درآمد	5	جمع کل دارایی‌ها	5
درصد سود ناویژه به درآمد	6	حساب‌ها و اسناد پرداختی تجاری	6
نسبت قیمت به سود هر سهم (P/E)	7	جمع بدهی‌های جاری	7
درصد بازده مجموع دارایی‌ها	8	حساب‌ها و اسناد پرداختی بلندمدت	8
درصد بازده سرمایه	9	حقوق صاحبان سهام در پایان سال مالی	9
		جمع کل بدهی‌ها	10

منبع یافته‌های پژوهشگر

برای ساخت مدل لازم است ابتدا تکنیک مدل‌سازی و نرم‌افزارهای لازم برای اعمال تکنیک‌ها انتخاب شود. برای اعمال تکنیک‌ها در تحقیق حاضر از نرم‌افزارهای Microsoft SQL Server V 2008 و Microsoft Excel V 2007 و SPSS Clementine V 12 استفاده شده است.

۷- تجزیه و تحلیل داده‌ها

برای آزمون فرضیه از نرم‌افزار SPSS Clementine استفاده کردیم. بدین صورت که ابتدا داده‌های خام را جمع‌آوری و طبقه‌بندی کردیم سپس در قالب SQL وارد نرم‌افزار کردیم. در این تحقیق ۲۵٪ داده‌ها، به عنوان داده‌های تست و ۷۵٪ به عنوان داده‌های آموزش انتخاب می‌نماییم. در این تحقیق از مدل C.R.T و CHAID, Neural Net, C5.0 استفاده شده است.

درخت تصمیم یکی از عمومی‌ترین تکنیک‌ها داده کاوی برای حل مسائل از نوع طبقه بندی می-

باشد این الگوریتم با ایجاد یکسری گره‌ها و شعبه‌ها بهم متصل سعی در طبقه بندی متغیرهای اسمی و تخمین متغیرهای عددی دارد گره‌های یک درخت تصمیم نشان دهنده نقاط آزمون متغیر است بر مبنای نتایج آزمون صورت گرفته درخت تصمیم به گره‌ها یا برگ‌های جدید می‌رسد. ریشه درخت تصمیم اولین و اساسی ترین گره درخت و برگ‌های آن آخرین نقطه تفکیک و به عبارتی نشان دهنده طبقه بندی است. مدل‌های حاصل از درخت تصمیم دارای الگوریتم‌های متفاوتی مثل تکنیک‌های C5.0, CART, CHAID که نتایج اجرای این مدل‌ها به شرح زیر می‌باشد:

۷-۱- مدل C.RT

الگوریتم C.RT را می‌سازیم. سپس الگوریتم را اجرا نموده تا مدل C.RT ساخته شود. سایر مشخصات پارامترهای این الگوریتم عبارت است از:

میزان صحت مدل C5.0 در مجموعه داده‌های آموزش ۸۱/۰۸٪ می‌باشد. درصد‌های مربوط به تست، در این تکنیک ۸۳/۹۶٪ می‌باشد. همچنین تعداد داده‌های آموزش مدل ۱۱۳ تراکنش و تعداد داده‌های تست در این مدل ۳۷ تراکنش برآورد شد.

در این مدل از مجموعه داده‌های آموزش از ۱۱۳ داده ۹۶ داده رادرست و ۱۷ داده رانادرست تشخیص داده شده اند از مجموعه داده‌های تست از ۳۷ داده ۳۰ داده رادرست و ۷ داده رانادرست تشخیص داده شده اند بنابراین این مدل دارای دقت پیش بینی ۸۱/۰۸٪ می‌باشد.

باتوجه به دقت بالای این مدل و جدول متغیرهای مهم آن می‌توان نتیجه گرفت که اقلام کلیدی ترازنامه ای و نسبت‌های سودآوری هر دو جهت ارائه پرتفوی مناسب می‌باشد.

جدول شماره ۲- نتایج اجرای مدل C.RT

ترتیب	۱	۲	۳	۴	۵	۶	۷	۸	۹	۱۰	۱۱	۱۲	۱۳	۱۴	۱۵	۱۶	۱۷	۱۸
متغیرها	BI10	II7	II6	BI5	BI9	II4	BI3	II3	BI7	II7	II6	BI6	BI2	II9	BI8	II1	II5	BI1

منبع یافته‌های پژوهشگر

۷-۲- مدل درخت تصمیم (C5.0)

انتخاب ویژگی مبتنی بر درخت تصمیم گیری به علت اینکه این الگوریتم طبقه بندی همچون درخت تصمیم زنتیکی برای انتخاب ویژگی‌های مناسب استفاده می‌کند مبتنی بر رویکرد طرح جدا سازی شده در انتخاب ویژگی‌ها است.

میزان صحت مدل درخت تصمیم (C5.0) در مجموعه داده‌های آموزش ۸۱,۳۳٪ می‌باشد. درصد‌های مربوط به تست مربوط به صحت طبقه بندی و معیاری برای ارزیابی اعتبار و صحت مدل می‌باشد که در این تکنیک ۸۶,۶۷٪ می‌باشد. همچنین تعداد داده‌های آموزش مدل ۱۲۰ تراکنش و تعداد داده‌های تست در

این مدل ۳۰ تراکنش برآورد شد.

در این مدل از مجموعه داده های آموزش از ۱۲۰ داده ۱۰۶ داده رادرست و ۱۴ داده رانادرست تشخیص داده شده اند از مجموعه داده های تست از ۳۰ داده ۲۶ داده رادرست و ۴ داده رانادرست تشخیص داده شده اند بنابراین این مدل دارای دقت پیش بینی ۸۶,۶۷% میباشد. باتوجه به دقت بالای این مدل و جدول متغیرهای مهم آن می توان نتیجه گرفت که ارقام کلیدی ترانامه ای و نسبت های سودآوری هر دو جهت ارائه پرتفوی مناسب می باشد.

جدول شماره ۳ - نتایج اجرای مدل (C5.0)

ترتیب	۱	۲	۳	۴	۵	۶	۷
متغیرها	B18	II7	IN2	BI2	II6	BI10	II5

منبع یافته های پژوهشگر

۳-۷- مدل شبکه های عصبی^۷

مدل انتخابی از نوع پرسپترون چند لایه است این مدل ها کاربرد موفقیت آمیزی در حل برخی از مسائل مخصوصاً مسائل رتبه بندی داشته اند. این مدل جزء شبکه های پیشخور چند لایه است که مجموعه ای از واحدهای گیرنده، لایه ورودی را تشکیل می دهند و یک یا چند لایه پنهانی از گره های محاسباتی وجود دارد و همچنین یک لایه خروجی گره محاسباتی در انتها قرار دارد. میزان صحت مدل شبکه عصبی (MLP) در مجموعه داده های آموزش ۶۱,۶۷% می باشد. درصدهای مربوط به تست مربوط به صحت طبقه بندی و معیاری برای ارزیابی اعتبار و صحت مدل می باشد که در این تکنیک ۷۳,۱۷% می باشد. همچنین تعداد داده های آموزش مدل ۱۰۹ تراکنش و تعداد داده های تست در این مدل ۴۱ تراکنش برآورد شد. در این مدل از مجموعه داده های آموزش از ۱۰۹ داده ۶۷ داده رادرست و ۴۲ داده رانادرست تشخیص داده شده اند از مجموعه داده های تست از ۴۱ داده ۳۰ داده رادرست و ۱۱ داده رانادرست تشخیص داده شده اند بنابراین این مدل دارای دقت پیش بینی ۷۳,۱۷% می باشد. باتوجه به دقت بالای این مدل و جدول متغیرهای مهم آن می توان نتیجه گرفت که ارقام کلیدی ترانامه ای و نسبت های سودآوری هر دو جهت ارائه پرتفوی مناسب می باشد.

جدول شماره ۴ - نحوه ساخت و اجرای مدل شبکه عصبی

ترتیب	۱	۲	۳	۴	۵	۶	۷	۸	۹	۱۰
متغیرها	II6	II7	II1	II4	BI5	II5	II3	BI9	BI7	BI2
ترتیب	۱۱	۱۲	۱۳	۱۴	۱۵	۱۶	۱۷	۱۸	۱۹	
متغیرها	II2	BI6	BI4	BI3	BI10	BI8	II9	BI1	II8	

منبع یافته های پژوهشگر

۴-۷- مدل CHAID

میزان صحت مدل CHAID در مجموعه داده‌های آموزش ۷۰٫۵۹٪ می‌باشد. درصدهای مربوط به تست مربوط به صحت طبقه‌بندی و معیاری برای ارزیابی اعتبار و صحت مدل می‌باشد که در این تکنیک ۷۹٫۱۷٪ می‌باشد. همچنین تعداد داده‌های آموزش مدل ۱۰۲ تراکنش و تعداد داده‌های تست در این مدل ۴۸ تراکنش برآورد شد.

در این مدل از مجموعه داده‌های آموزش از ۱۰۲ داده ۷۲ داده را درست و ۳۰ داده را نادرست تشخیص داده شده اند از مجموعه داده‌های تست از ۴۸ داده ۳۸ داده را درست و ۱۰ داده را نادرست تشخیص داده شده اند بنابراین این مدل دارای دقت پیش بینی ۷۹٫۱۷٪ می‌باشد. باتوجه به دقت بالای این مدل و جدول متغیرهای مهم آن می‌توان نتیجه گرفت که اقلام کلیدی ترازنامه ای و نسبت‌های سودآوری هر دو جهت ارائه پرتفوی مناسب می‌باشد.

جدول شماره ۵ - نتایج اجرای مدل CHAID

ترتیب	۱	۲	۳	۴	۵	۶
متغیرها	BI5	BI3	II7	II6	II4	II5

منبع یافته‌های پژوهشگر

نتایج حاصل از بررسی چهار مدل فوق، درصد بالای صحت پیش‌بینی و اعتبار و کارایی این تکنیک‌ها را در میزان سودمندی معیارهای سودآوری نشان داد. بنابراین فرضیه اصلی تحقیق تایید گردید. بدین معنی اقلام کلیدی ترازنامه‌ای و نسبت‌های سودآوری هر دو جهت ارائه پرتفوی بهینه با استفاده از داده کاوی مناسب می‌باشد.

۸- بحث و نتیجه‌گیری

با توجه به اهمیت تعیین شرکت‌های برتر بورس و راهی برای سرمایه‌گذاری مطمئن، دسترسی به مدل‌هایی که با دقت بالا بتواند این رخداد را پیش‌بینی نماید، نقش مهمی به لحاظ اقتصادی و عملی دارد. در این پژوهش چهار مدل مختلف برای پیش‌بینی معیارهای برتر برای ارائه پرتفوی اجرا گردید و یافته‌های حاصل مقایسه شدند.

در حالی که نتایج پژوهش حاضر نشان دهنده‌ی این واقعیت است که امکان پیش‌بینی شرکت‌های برتر بورس به عنوان پرتفوی کارا از راه اقلام کلیدی ترازنامه ای و نسبت‌های سودآوری وجود دارد. نتایج حاکی از این است که اقلام کلیدی ترازنامه ای و نسبت‌های سودآوری هر دو حائز اهمیت می‌باشد. این نتیجه با توجه به اهمیت اقلام کلیدی ترازنامه ای و نسبت‌های سودآوری در بررسی وضعیت و عملکرد کلی شرکت‌ها و اطمینان از صحت عملیات تحصیل سود آنها و مدیریت معقول منابع و مصارف شرکت‌ها می‌باشد. سرمایه

گذاران معمولا تمایل به سرمایه گذاری در شرکت های سوآور و معتبر را دارند. این نتیجه با نتایج تحقیقاتی از جمله تحقیق جمالی و همکاران که اقلام کلیدی ترازنامه ای و نسبت های سوآوری را با شاخص های بورس مقایسه و برای ارائه پرتفوی استفاده کردند مطابقت دارد. با توجه به این نتایج سرمایه گذاران می توانند میان شرکت های برتر بورس پرتفوی بهینه با اقلام کلیدی ترازنامه ای و نسبت های سوآوری تشکیل دهند. هرچند ارجحیت تاثیر گذاری آنان در هر مدل ترتیبی متفاوت اما برخی از آنان مانند جمع کل دارایی ها، درصد سود به درآمد، درصد سود عملیاتی به درآمد و نسبت قیمت به سود هر سهم (P/E) در تمام مدل ها به عنوان یک فصل مشترک از ارجحیت بالاتری برخوردارند که می توانند به عنوان مبنایی مناسب مورد توجه سرمایه گذاران قرار گیرد. پژوهشگران آتی می توانند شاخص های دیگر تشکیل پرتفوی را نیز با داده کاوی بررسی کرده و همچنین کلیه شرکتهای بورس را مورد بررسی قرار بدهند.

فهرست منابع

- ۱) پورزندى، محمدابراهيم و هدايت‌اله منصورى، (۱۳۷۹)، "نقش گزارشگرى مالى در بازار سرمايه"، حسابدار، شماره ۱۳۸.
- ۲) تقوى فرد، محمدتقى، طاها منصورى، سيدمحمدرضا ناصرزاده و عليرضا فراست، (۱۳۸۶)، "داده كاوى و كاربرد آن در تصميم‌گيرى‌ها"، دانش مديريت، سال بيستم، شماره ۷۹.
- ۳) جمالى، على، نرگس متقى نيا، سميه فتحى و سميرا سيف (۱۳۹۱)، "ارائه پرتفوى مبتنى بر شاخص هاى سوآورى و نقدينگى باروش DIA"، کنفرانس ملى حسابدارى، مديريت مالى و سرمايه گذارى ۲۶ بهمن، ۱۳۹۱، گرگان.
- ۴) جهانخانى، على و على پارسائيان، (۱۳۷۶)، "مديريت سرمايه‌گذارى و ارزيايى اوراق بهادار"، تهران، انتشارات دانشكده مديريت دانشگاه تهران.
- ۵) راعى، رضا، (۱۳۷۷)، "طراحي مدل سرمايه‌گذارى مناسب در سبد سهام با استفاده از شبكه‌هاى عصبى مصنوعى"، رساله دكتورى، دانشكده مديريت دانشگاه تهران.
- ۶) كيقبادى، اميررضا و وحيد خدامى، (۱۳۹۲)، "داده كاوى صورت هاى مالى جهت اعطاي تسهيلات مالى"، فصلنامه پژوهش هاى حسابدارى مالى و حسابرسي، شماره ۱۷.
- 7) Abd Rahman, Mohd Shaari, (2008), "Utilization of Data Mining Technology within the Accounting Information System in the Public Sector: A Country Study- Malaysia Degree of Philosophy, University of Tasmania.
- 8) Dashtbayaz, M. L. et al., (2015), "Data Search and Discovery Process for Financial Statement Fraud", Research Journal of Finance and Accounting, Vol.6, No.3, PP. 75-80.
- 9) G.L. Gray, R.S. Debreceeny, (2014), A Taxonomy to Guide Research on the Application of Data mining to Fraud Detection in Financial Statement, Audits International Journal of Accounting Information Systems, Vol. 15, Iss. 4, December, PP. 357-380
- 10) Kirkos, S., & Manolopoulos, Y., (2004), "Data Mining in Finance and Accounting: A Review of Current Research Trends". In Proceedings of The 1st International Conference on Enterprise Systems and Accounting (ICESAcc), Thessaloniki, Greece, PP. 63-78.
- 11) Larose, D. T., (2005), "Discovering Knowledge in Data, John Wiley".

- 12) M.M. Nasrizar,(2014)," Application of Data Mining Techniques to Detect and Predict Accounting Fraud: A Comparison of Neural Networks and Discriminant Analysis, *Advances in Economics and Business Management (AEBM)*, Vol. 1, No. 3; November, PP. 226-228.
- 13) Moorthi M., Nagarathna B.,(2015)," A Review on Data Mining and its Techniques in Finance Sector", *International Journal of Advance Research in Computer Science and Management Studies* ,Vol. 3, Iss. 3, March, PP. 29-37
- 14) Sowjanya et al., (2013), " Application of Data Mining Techniques for Financial Accounting Fraud Detection Scheme", *International Journal of Advanced Research in Computer Science and Software Engineering* 3(11), November, PP. 717-724.
- 15) Tofallis, C., (1994), "Input Efficiency Profiling: An Application to Airlines", *Computer Ops Res.* Vol. 24, No. 3, PP. 253-258.

یادداشت‌ها

- ¹. Loras chi
- ². Powers Jenifer
- ³. Souiel
- ⁴. Lopes
- ⁵. AHP&IEP&DEA
- ⁶. Variable Importance
- ⁷. Neural Net