دوفصلنامه اقتصاد پولی، مالی(دانش وتوسعه سابق) دوره جدید، سال بیست وسوم، شمارۀ12، پاییز و زمستان 1395

توانایی مدل‌های تخمینگر بردار پشتیبان، تخمینگر حداقل درجه و شبکه عصبی فازی در پیش‌بینی سود هر سهم

محمد رسول چوپانی[1]
کارشناس ارشد حسابداری دانشگاه فردوسی مشهد
فرزانه نصیرزاده[2]
دانشیار گروه حسابداری دانشگاه فردوسی مشهد
مهدی صالحی[3]
دانشیار گروه حسابداری دانشگاه فردوسی مشهد

تاریخ دریافت: 1394/4/20 تاریخ پذیرش: 1395/5/11

چکیده

پیش‌بینی سود حسابداری و تغییرات آن به جهت استفاده در مدل‌های ارزیابی سهام، توان پرداخت، ریسک، عملکرد واحد اقتصادی و مباشرت مدیریت از دیرباز مورد علاقه سرمایه‌گذاران، مدیران، تحلیلگران مالی و اعتباردهندگان بوده است. سود هر سهم اغلب برای بررسی سودآوری و ریسک مرتبط با سود و نیز قضاوت در خصوص قیمت سهام استفاده می‌شود. در این تحقیق، عملکرد سه مدل تخمینگر بردار پشتیبان، تخمینگر حداقل درجه و شبکه عصبی فازی در پیش‌بینی سود هر سهم مورد ارزیابی قرار گرفته است. در این تحقیق از 9 متغیر مالی، 7 متغیر بنیادی و 4 متغیر کلان استفاده شده است. ابتدا متغیرهای مالی، متغیرهای بنیادی و کلان اقتصادی به صورت جداگانه و سپس به طور همزمان وارد مدل‌ها شده‌اند تا توانایی آن‌ها در هر سه حالت مورد ارزیابی قرار گیرد. نتایج تحقیق نشان می‌دهد که مدل تخمینگر حداقل درجه در هر سه حالت فوق (متغیرهای مالی، متغیرهای بنیادی و کلان اقتصادی) عملکرد بهتری نسبت به سایر مدل‌ها داشته است. همچنین استفاده از متغیرهای مالی در مدل تخمینگر حداقل درجه به تنهایی، باعث عملکرد مطلوب‌تر آن نسبت به سایر حالت‌ها شده است.

1- mrasoolchopani@gmail.com

2- نویسنده مسئول: nasirzadeh@um.ac.ir

کلیدواژه‌ها: بورس اوراق بهادار تهران، پیش‌بینی سود هر سهم، تخمینگر بردار پشتیبان، تخمینگر حداقل درجه، شبکه عصبی فازی.
طبقه‌بندی JEL: E4 , G3

1. مقدمه

در مباحث مالی و حسابداری، مفهوم سود از جایگاه ویژه‌ای برخوردار است و به عنوان معیاری برای سنجش کارایی مدیریت، پیش‌بینی سود آتی شرکت، توزیع سود سهام و مبنایی برای تشخیص مالیات استفاده می‌شود (Cllyde, 2007). در بیانیه شماره 1 هیئت استانداردهای حسابداری مالی سود ابزاری برای ارزیابی سودآوری، توان پرداخت سود سهام، پیش‌بینی سودهای آتی و ارزیابی ریسک سرمایه‌گذاری معرفی شده است. بنابراین، پیش‌بینی سود هر سهم از اهمیت ویژه‌ای برخوردار است و هدف اولیهٔ حسابداری مالی نیز فراهم آوردن اطلاعات مفید برای سرمایه‌گذاران جهت پیش‌بینی عملکرد آتی واحدهای اقتصادی است.

سود هر سهم یکی از متغیرهای مالی بسیار مهم است که اغلب برای ارزیابی سودآوری، ریسک مرتبط با نقدینگی و قیمت‌گذاری سهام استفاده می‌شود. برای به کارگیری سود هر سهم در مدل‌های ارزشیابی سهام از الگوریتم‌های مختلفی استفاده می‌شود که برخی از آن‌ها مدل‌های آماری و برخی مدل‌های هوشمند هستند، لذا برخی از این الگوریتم‌ها از مدل‌های آماری و برخی از مدل‌های هوشمند بهره می‌برند (Rahmani & Bakhtari, 2007). از آنجایی که روش‌های آماری سنتی پیش‌فرض‌هایی مانند خطی بودن، نرمال بودن و مستقل بودن متغیرهای پیش‌بینی دارند در بسیاری از موارد، از دقت کمی برخوردارند(Raei And Fallahpour, 2009). تحقیقات قبلی نیز نشان می‌دهد که مدل‌های هوشمند نسبت به مدل‌های آماری قابلیت پیش‌بینی بیشتری دارند و در صورتی که دو یا چند مدل باهم ترکیب شوند، احتمال پیش‌بینی دقیق‌تر افزایش می‌یابد (Lee And et al, 1996).

تنوع منابع پیش‌بینی سود سبب شده است که دقت پیش‌بینی این مدل‌ها بیشتر مورد توجه قرار گیرد؛ بنابراین، مدلی که درصد اشتباه کمتری داشته باشد، مطمئن‌تر خواهد بود (Mashayekh And Shahrokhi, 2008). مطالعاتی که به تازگی بر روی دقت پیش‌بینی مدل‌های هوشمند انجام گرفته است، بیانگر عملکرد مطلوب‌تر این مدل‌ها نسبت به مدل‌های آماری در طبقه‌بندی و یافتن راه حل بهینه است (Thommano, 1999, kelly 2007, Wang And Tan, 2010, Ming And

Kosaka And et al, 1991, Chi Lee, 2009). از این رو، سرمایه‌گذارانی که از این مدل‌های پیش‌بینی استفاده می‌کنند، به شکل مناسب‌تری فرصت‌های سرمایه‌گذاری را تشخیص خواهند داد. تحقیق حاضر به لحاظ استفاده همزمان متغیرهای مالی، بنیادی و کلان در مدل‌های هوشمند جهت پیش‌بینی سود هر سهم، جامع‌تر از سایر تحقیقات در این زمینه باشد.

با توجه به موارد یادشده، هدف از انجام این پژوهش، بررسی توانایی سه مدل تخمینگر بردار پشتیبان، تخمینگر حداقل درجه و شبکه عصبی فازی (که جزء مدل‌های هوشمند با کمترین میزان خطا هستند) در پیش‌بینی سود هر سهم شرکت‌های پذیرفته شده در بورس اوراق بهادار تهران طی بازه زمانی 1384 الی 1391 می‌باشد.

2. مبانی نظری

سود، مفهومی اقتصادی دارد و از مبهم‌ترین مفاهیمی است که در دنیای تجارت استفاده می‌شود و با وجود پیشرفت‌های زیاد در حسابداری، هنوز تعریف جامعی برای آن ارائه نشده است (Mojtahedzade And Nazari tanha, 2009). اهمیت سود به قدری است که به عنوان یکی از معیارهای اساسی در تعیین قیمت سهام محسوب می‌گردد و در مدل‌های ارزشیابی سهام نیز به طور گسترده استفاده می‌شود. بنابراین، سرمایه‌گذاران همواره به دنبال پیش‌بینی هر چه دقیق‌تر سود می‌باشند. پیش‌بینی بررسی احتمال وقوع حوادث بر پایهٔ اطلاعات موجود طبق اصول و قواعد علمی به منظور کاهش ریسک است (Ezati, 1998). از آنجایی که معمولاً هر پیش‌بینی با خطا همراه است باید از روش‌هایی استفاده شود که کمترین خطا را داشته باشد. سرمایه‌گذاری در بازار سهام نیز همواره با عدم اطمینان همراه است، از این رو لازم است با پیش‌بینی صحیح، عدم اطمینان را کاهش داد. در واقع پیش‌بینی خوب با کمترین خطا می‌تواند به سرمایه‌گذاران در امر تصمیم‌گیری کمک نماید و میزان ریسک آن‌ها را کاهش دهد (Hodavandi, 2012). عوامل بسیاری مانند متغیرهای درون‌سازمانی و متغیرهای برون‌سازمانی از جمله متغیرهای کلان اقتصادی بر سود تاثیرگذار هستند و همین تعدد باعث افزایش خطای پیش‌بینی می‌شود و از طرف دیگر انگیزهٔ استفاده از سیستم‌ها و الگوریتم‌های جدید را ایجاد می‌کند. بنابراین، متخصصان بازارسرمایه سال‌های متمادی به مطالعهٔ بازار و شناسایی الگوهای مختلف برای پیش‌بینی پرداخته‌اند و ترکیبی از

الگوی تشخیص و تجربهٔ مبتنی بر مشاهدهٔ روابط علی و معلولی را به کار بسته‌اند. این فرایند ممکن است به صورت مدل‌های خطی، مدل‌های غیرخطی و یا مدل‌های تصادفی صورت گیرد. این تکنیک‌ها را می‌توان براساس نوع ابزار و نوع داده‌های مورد استفاده به چهارگروه طبقه‌بندی نمود که شامل 1 - روش‌های تحلیل فنی[1]، 2 -روش‌های تحلیل بنیادی[2]، 3 - روش‌های پیش‌بینی سری-های زمانی کلاسیک[3] و 4- روش‌های هوشمند[4] است.

در روش تحلیل فنی سعی بر آن است که با استفاده از رفتار گذشته متغیر الگویی تهیه شود که بتوان با آن تغییرات آینده آن را پیش‌بینی کرد. در این روش اعتقاد بر آن است که رفتارهای گذشتهٔ یک متغیر در آینده‌ای با همان شرایط، قابل تکرار است(Sinaii And et al, 2005). روش‌های کلاسیک به دنبال یافتن یک رابطهٔ خطی بین مقادیر گذشته و آینده یک متغیر هستند. روش‌های هوشمند الگوهای خطی و غیرخطی موجود در داده‌ها را دنبال می‌کنند تا بدین وسیله فرایند ایجاد آن‌ها را حدس بزنند(Nassirzade and Nikravesh, 2013). در این روش سعی بر آن است که بتوان تا حد امکان متغیرهای تاثیرگذار را شناسایی و تأثیر آن را بر متغیر تحقیق در نظر گرفت. از این رو تمرکز مقاله حاضر بر روی روش‌های هوشمند و مقایسه آن‌ها است. بررسی تحقیقات انجام شده با استفاده از مدل‌های هوشمند نشان می‌دهد که سه مدل تخمینگر بردار پشتیبان، تخمینگر حداقل درجه و شبکه عصبی فازی دارای بیشترین دقت در پیش‌بینی هستند لذا تحقیقات حاضر به مقایسه دقت این سه مدل در پیش‌بینی هر سهم پرداخته است.

2-2 -1 ماشین بردار پشتیبان[5] مدلی هوشمند است که به عنوان یک روش یادگیری با نظارت شناخته می‌شود. این الگوریتم در زمینه شناسایی الگو و پیش‌بینی رگرسیون استفاده می-شود. هدف آن تشخیص و متمایز کردن الگوهای پیچیده در داده‌ها و یافتن قوانین حاکم بر آن‌ها می‌باشد . تعمیم پذیری خوب، توانایی در طبقه بندی الگوهای ورودی، رسیدن به الگوی بهینه

1- Technical Analysis

2- Fundamental Anahysis

3- Traditinal Time Series

4- Machine Learning

5- Support Vector Machine

کلی، قابلیت یادگیری (تعیین خودکار ساختار بهینه برای مجموعه داده‌های تحت آموزش) از ویژگی‌های این الگوریتم به شمار می‌رود. این مدل داده‌ها را به بردارهای یادگیری تبدیل می‌نمایدکه هر بردار با یک مقدار خروجی متناظر می‌باشد و قابلیت پیدا کردن مقدار بهینه را در فضای غیرخطی و خطی دارا است (Raei And Fallahpour, 2009).

2-2 -2 تخمینگر حداقل درجه یک روش خطی است که اساس کار آن مانند رگرسیون انتخاب رو به جلو (در این روش متغیرهای مستقل درصورتی که معیار ورود به مدل را داشته باشند تک به تک وارد معادله می‌شوند و بعد از ورود حذف نمی‌شوند) و انتخاب رو به عقب (در این روش تمامی متغیرهای مستقل ابتدا به معادله وارد می‌شوند و سپس در صورتی که معیار لازم برای باقی ماندن در مدل را نداشته باشد تک به تک از مدل حذف می‌شوند) می‌باشد. روش رگرسیون خطی انتخاب رو به جلو و انتخاب رو به عقب با توجه به تفاوت کلی در ساختار و فرمول‌ها به جواب‌های بسیار نزدیکی می‌رسند. تلفیق این دو روش سبب به وجود آمدن روشی ساده‌تر و با محاسبات کم‌تر به نام تخمینگر حداقل درجه شده‌اند. در این روش جهت پیدا کردن ضرایب، به تعداد متغیرهای مستقل مرحله وجود خواهد داشت، با این حال حداقل زمان استفاده برای مدل‌بندی و دقت بالاتر از مزایای این مدل می‌باشد (Nassirzade and Nikravesh, 2013).

2-2 -3 مدل شبکه عصبی فازی یک نظریه نسبتاً جدید ریاضی است که توانائی نشان دادن فرایندها و ترکیبات دل‌خواه غیرخطی جهت ارتباط بین ورودی‌ها و خروجی‌های هر سیستمی را داراست. این شبکه با داده‌های موجود طی فرایند یادگیری آموزش دیده و جهت پیش‌بینی در آینده مورد استفاده قرار می‌گیرد که دارای ویژگی‌های قابلیت یادگیری، قابلیت تعمیم و پردازش موازی را می‌توان از ویژگی‌های این روش ذکر کرد (Sinaii And at al, 2005).

3. پیشینه تحقیق

در ادبیات مربوط به پیش‌بینی سود هر سهم از مدل‌سازی خطی، سری‌های زمانی و یا قضاوت حرفه‌ای برای پیش‌بینی استفاده شده و به بررسی دقت این مدل‌ها نیز پرداخته شده است. قبل از سال 1997 در تحقیقات مربوط به پیش‌بینی از روش‌های خطی مانند رگرسیون چند متغیره استفاده شده است. Abarbanell and Bushee (1997) روش‌های غیرخطی را در پیش بینی خود مطرح کردند، از آن پس بیشتر تحقیقات مربوط به پیش‌بینی سود، از مدل‌های غیرخطی استفاده کرده‌اند.

متغیرهای مورد استفاده آباربانل و بوشی (1997) شامل موجودی کالا، حساب‌های دریافتنی، اموال و تجهیزات، هزینه‌های اداری و فروش، حاشیه سود ناخالص، نرخ مؤثر مالیاتی و بهره‌وری نیروی کار، که رابطه غیرخطی با سود دارند، می‌باشد. آن‌ها با استفاده از این متغیرها و مدل‌های غیرخطی به پیش‌بینی سود پرداختند.

1-3 مطالعات خارجی

Zhang (2004) در تحقیق خود 283 شرکت را در قالب 41 صنعت در نظر گرفتند. آن‌ها در این تحقیق از 4 مدل خطی تک متغیره، خطی چند متغیره، شبکه عصبی تک متغیره و شبکه عصبی چند متغیره برای پیش‌بینی سود هر سهم استفاده کردند. متغیرهای مورد استفاده در این تحقیق شامل موجودی کالا، دارایی‌های سرمایه‌ای، حساب‌های دریافتنی، حاشیه سود ناخالص، هزینه‌های اداری و فروش، نرخ مالیات مؤثر و بهره‌وری نیروی کار می‌باشد. نتایج به دست آمده از تحقیق نشان می‌دهد که شبکه عصبی دارای بهترین عملکرد در پیش‌بینی سود می‌باشد.

Kelly (2007) با استفاده از مدل تخمینگر حداقل درجه و روش‌های بیزی به پیش‌بینی حجم پول در آمریکا نموده است. متغیرهای مورد استفاده نرخ بهره بلندمدت، نرخ بهره کوتاه‌مدت، نرخ بیکاری، میزان سپرده‌گذاری و هزینه خدمات پولی است. داده‌های مربوط به متغیرها برای یک بازه زمانی 50 ساله از سال 1960 تا 2009 به صورت ماهانه جمع آوری شده است. نتایج نشان می‌دهد که مدل تخمینگر حداقل درجه نسبت به روش‌های بیزی دارای عملکرد مطلوب‌تری است.

Ming (2009) با استفاده از مدل تخمینگر بردار پشتیبان و مدل شبکه عصبی به پیش‌بینی شاخص نزدک پرداختند. مدل تخمینگر بردار پشتیبان توسط 29 شاخص فنی به عنوان مجموعه‌ای از ویژگی‌های کامل در جهت تغییر شاخص استفاده شد. داده‌های تحقیق از سال 2001 تا 2008 می‌باشد. نتایج نشان دهنده برتری مدل تخمینگر بردار پشتیبان نسبت به شبکه عصبی است.

Wang & Tan (2010) در پژوهش خود از مدل تخمینگر حداقل درجه برای استخراج عوامل تاثیرگذار بر بازده سهام استفاده نموده‌اند. ایشان اطلاعات مورد نیاز را از ترازنامه، صورت سود و زیان و صورت جریان وجوه نقد شرکت‌های فعال در سه بازار اصلی آمریکا (بورس نیویورک، بورس آمریکا و بورس نزدک) استخراج کردند که در مجموع 65 متغیر را شامل می‌شود. داده‌ها طی یک دوره 5 ساله (2003-1999) جمع‌آوری شده‌اند. نتیجه حاصل از پژوهش نشان دهنده

عملکرد مطلوب مدل تخمینگر حداقل درجه در تعیین عوامل تاثیرگذار بر پیش‌بینی نرخ بازده است. همچنین بر اساس نتایج پژوهش دو عامل سود هر سهم حاصل از عملیات مستمر و ارزش بازار بیشترین تأثیر را بر پیش‌بینی نرخ بازده داشته است.

Kara, Boyacioglu & Baykan (2011)پیش‌بینی جهت حرکت شاخص بورس استانبول را با استفاده از شبکه عصبی مصنوعی و ماشین بردار پشتیبان بررسی کردند. در این تحقیق که طی دوره 1997 تا 2007 انجام‌شده است درمجموع 1440 بار (52,7 درصد) افزایش و 1293 بار (47,3 درصد) کاهش شاخص مشاهده شده است. نتایج تحقیق نشان می‌دهد که مدل شبکه عصبی با 75,74 درصد دقت نسبت به مدل ماشین بردار پشتیبان با 71,52 درصد دقت، عملکرد مطلوب‌تری داشته است.

Svalina, Galzina, Lujić & Šimunović (2013) در پژوهش خود با استفاده از شبکه عصبی فازی به پیش‌بینی قیمت پایانی شاخص بورس زاگرب پرداختند. داده‌های تحقیق به‌صورت سری زمانی و برای دوره 2010 تا 2012 به‌صورت روزانه جمع‌آوری شده است. روش کار به این صورت بوده است که با استفاده از اطلاعات یک روز، قیمت پایانی شاخص را برای 5 روز بعد پیش‌بینی کرده‌اند. بررسی معیارهای کمی و کیفی فرضیات نشان می‌دهد که مدل شبکه عصبی فازی با دقت مناسبی قادر به پیش‌بینی قیمت پایانی شاخص زاگرب است.

Ou & Wang (2014) با استفاده از ترکیب تخمینگر بردار پشتیبان و الگوریتم ژنتیک به پیش‌بینی نوسانات بازده سهام در بورس نزدک پرداختند. متغیرها با استفاده از روش رگرسیونی باکس-جنکینز استخراج و با الگوریتم ژنتیک بهینه‌شده و نهایتاً وارد مدل شده‌اند. نتایج تحقیق بیانگر عملکرد مطلوب‌تر مدل ترکیبی در مقایسه با مدل GARCH و تخمینگر بردار پشتیبان به‌تنهایی است.

Xion, Bao, And Hu (2014) با استفاده از ترکیب مدل تخمینگر بردار پشتیبان چندمرحله‌ای و الگوریتم بهینه‌سازی کرم شب‌تاب مدلی ترکیبی را برای پیش‌بینی شاخص قیمت سهام طراحی کردند. در این تحقیق از داده‌های 500 شرکت برتر ایالات‌متحده، 100 شرکت برتر بریتانیا و 225 شرکت برتر ژاپن استفاده‌شده است. نتایج تحقیق نشان می‌دهد که استفاده از مدل ترکیبی نسبت به مدل تخمینگر بردار پشتیبان عملکرد مطلوب‌تری دارد.

2-3 مطالعات داخلی

Ghvidel, Otadi, And Mosleh (2011) بر اساس شبکه عصبی فازی، روش جدیدی برای برآورد ضرایب فازی یک تابع عرضه و تقاضای نیروی کار با ورودی‌ها و خروجی‌های فازی ارائه نمودند. ایشان به بررسی و برآورد تابع عرضه و تقاضای فازی بازار ایران پرداختند و توانایی روش شبکه عصبی فازی را با روش کاو و تاناکا مقایسه نموده‌اند که نتایج پژوهش نشان دهنده برتری روش شبکه عصبی می‌باشد.

Nassirzadeh, And Nikravesh (2012) در پژوهش خود به بررسی عملکرد سه مدل تخمینگر بردار پشتیبان، تخمینگر حداقل درجه و شبکه عصبی فازی در پیش‌بینی قیمت سهام پرداخته است. در این تحقیق متغیرهای فنی (قیمت سهم) و روانشناسی (حجم و تعداد معاملات) در سه بازه زمانی روزانه، هفتگی و ماهانه، طی یک دوره 10 ساله برای 10 شرکت نمونه استخراج شده است. بررسی میانگین مجذور خطای پیش‌بینی به دست آمده از اجرای مدل‌های تحقیق نشان می‌دهد که در هر سه سطح روزانه، هفتگی و ماهانه مدل تخمینگر بردار پشتیبان عمکلرد مطلوب‌تری نسبت به مدل‌های دیگر تحقیق داشته است.

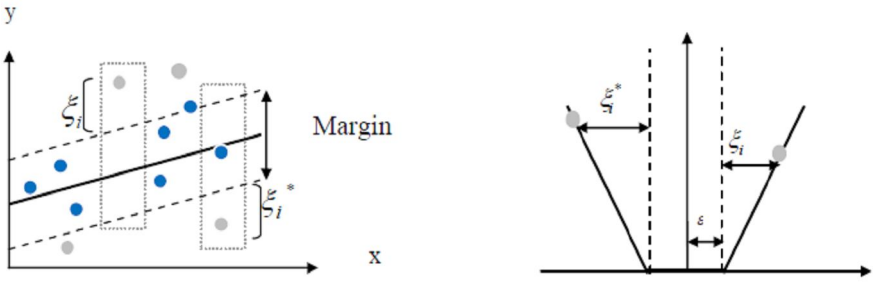
Bagheri, Mohammadi Peyhani, And Akbari (2014) از مدل ترکیبی شبکه عصبی فازی برای پیش‌بینی نرخ ارز در بازار ایران استفاده کردند. ابتدا با استفاده از تحلیل موجک‌ها به بررسی سری زمانی نرخ ارز پرداخته سپس با استفاده از سیستم عصبی فازی اقدام به پیش‌بینی نرخ ارز کردند. هم‌چنین ایشان در این مقاله از مدل ترکیبی DTW و WT برای بهینه‌سازی الگوهای ورودی استفاده کردند. نتایج تحقیق نشان می‌دهد که مدل ترکیبی ارائه‌شده در این مقاله با دقت بسیار بالایی قادر به پیش‌بینی سری‌های زمانی است. بررسی پیشینه تحقیقات انجام شده نشان می‌دهدکه مدل‌های هوشمند با خطای کمتر، بیشتر مورد استفاده قرار گرفته‌اند، همچنین در اکثر موارد این مدل‌ها نسبت به سایر مدل‌ها برتری داشته‌اند. با توجه به اهمیت پیش‌بینی سود هر سهم که در مبانی نظری مقاله ذکر شده است، در این تحقیق از مدل‌های هوشمند با خطای پیش‌بینی کم استفاده شده است. همچنین بررسی تحقیقات انجام شده نشان می‌دهد که به استفاده همزمان متغیرهای کلان اقتصادی و متغیرهای مالی در مدل‌های هوشمند توجه نگردیده، از این رو در پژوهش حاضر اثر متغیرهای کلان اقتصادی را نیز بررسی شده است.

4. مدل‌های تحقیق

همانطور که ذکر شده مدل‌هایی که دقت بالاتر دارند از مطلوبیت بیشتری برخوردارنـد از ایـن رو در تحقیق حاضر به مقایسه توان پیش‌بینی سود هر سهم با استفاده از سـه مـدل تخمینگـر بـردار پشتیبان، تخمینگر حداقل درجه و شبکه عصبی فازی پرداخته شده است؛ ابتدا به معرفی مـدل‌هـای استفاده شده در تحقیق و نحوه کارکرد آن‌ها پرداخته شده و سپس متغیرهای تحقیـق و چگـونگی فرآیند اجرای الگوریتم بیان شده است.

4-1. تخمینگر بردار پشتیبان

تخمینگر بردار پشتیبان به عنوان یک جایگزین محبوب برای ابزارهای رگرسـیون تبـدیل شـده است (Smola,2004). این روش از همان اصولی که ماشین بردار پشتیبان برای طبقه‌بنـدی اسـتفاده می‌کند (تنها با جند تفاوت جزئی) بهره می‌برد اما ایده‌ای اصلی همان به حداقل رساندن خطـا یـا ε است. اگر خطای صورت گرفته در حد فاصل خطای اپسیلون[6] باشد، خطا صفر در نظر گرفته مـی-شود. حد فاصل خطای اپسیلون فضایی را به وجود می‌آورد که اصـطلاحاً تیـوب[7] گفتـه مـی‌شـود. تیوب تخمینگر بردار پشتیبان در شکل شماره 1 نشان داده شده است.



شکل 1. تخمینگر بردار پشتیبان با خطای بیشتر از ε (Smola,2004).

6- Error-epsilon

7- Tube

بنابراین مسئله بهینه سازی به صورت زیر مطرح است:

$$Min_{w,b}\ P = \frac{1}{2}\|w\|^2 + C\sum_{i=1}^{l}(\S_i + \S_i^*) \qquad (1)$$

$$s.t.\begin{cases} y_i - w\emptyset(x_i) - b \leq \varepsilon + \S_i \\ w\emptyset(x_i) + b - y_i \leq \varepsilon + \S_i^* \\ \S_i, \S_i^* \geq 0 \qquad i = 1,2,\ldots,l \end{cases}$$

که در آن W، حاشیه و (x_i, y_i) مجموعه آموزش هستند. محدودیت‌ها نیز در واقع بیانگر این موضوع هستند که کلیه داده‌ها در تیوب واقع شده‌اند و مجموع فاصله نقاط خارج از تیوب حداقل گردد. در اینجا C عبارت جریمه[8] است که تعادل بین حداقل کردن خطا و بزرگ شدن حاشیه یا تیوب را کنترل می‌کند. اگر C زیاد باشد به این معناست که به کم کردن خطا اهمیت بیشتری می‌دهیم در نتیجه تیوب بزرگتر می‌شود و اگر C کوچک باشد به این معناست که به کوچک کردن تیوب اهمیت می‌دهیم و در نتیجه خطا بیشتر خواهد بود. با انجام عملیات ریاضی و بهره‌گیری از قضیه کوهن-تاکر می‌توانیم تابع تخمین شده را به شکل زیر بنویسیم (Smola,2004):

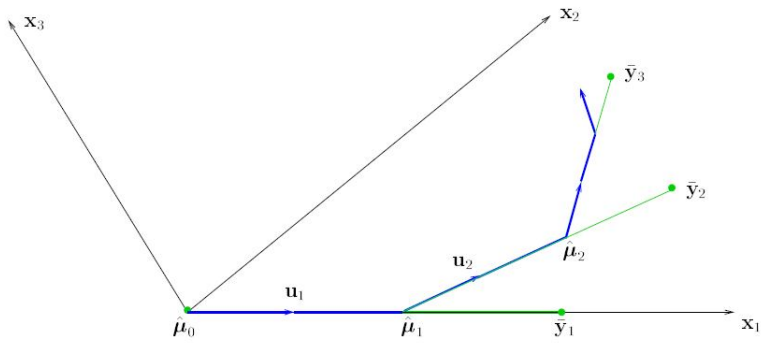
$$f(x) = \sum_{i=1}^{n}(\alpha_i - \alpha_i^*)\,K(x_i, x_j) + b \qquad (2)$$

2-4. تخمینگر حداقل درجه

تخمینگر حداقل درجه مدلی جدید بر اساس رگرسیون انتخاب رو به جلو و رو به عقب است. برای به دست آوردن این مدل، روش‌هایی مانند انتخاب رو به جلو[9]، حذف رو به عقب[10] و یا ترکیبی از این دو روش وجود دارد. معیاری که برای برتری یک روش مورد بررسی قرار می‌گیرد علاوه بر خطا و دقت تخمین، صرفه جوئی[11] نیز می‌باشد. در این روش نیز همانند روش انتخاب رو به جلو ابتدا تمامی ضرایب را برابر صفر قرار داده و متغیری که بیشترین همبستگی را با هدف دارد

8- Penalty Term
9- Forward Selection
10- Backward Elimination
11- Parsimony

(x_{j1}) انتخاب می‌کنیم. سپس بیشترین طول گامی که در جهت این متغیر می‌توان برداشت را، تا زمانی که متغیر دیگری مثل x_{j2} به همین میزان همبستگی با باقیمانده فعلی داشته باشد، بر می‌داریم. در این هنگام مدل تخمینگر حداقل درجه به جای اینکه در جهت x_{j1} ادامه دهد، در جهتی که زاویه برابر با هر دو متغیر داشته باشد - تا زمانی که متغیر سوم x_{j3} وارد "مجموعه بیشترین همبستگی" شود - ادامه می‌دهد. سپس در جهت زاویه برابر بین سه متغیر x_{j1} و x_{j2} و x_{j3} ادامه می‌دهیم که به آن حداقل زاویه همسویی[12] می‌گویند. شکل‌های شماره 2 نحوه عمل مدل تخمینگر حداقل درجه نشان می‌دهد (Efron and et al,2004).



شکل 2. روال هندسی الگوریتم تخمینگر حداقل درجه (Efron and et al,2004)

در این الگوریتم تنها به m گام نیاز داریم که m برابر با تعداد متغیرها یا همان ویژگی‌های[13] مسئله است. بنابراین مرتبه زمانی این روش وابسته به زمان محاسبه زاویه برابر در هر یک از m گام خواهد بود.

در این روش با داشتن ماتریس زیر و فرض نرمال بودن داده‌ها بدنبال β ای هستیم که در نهایت داشته باشیم:

$$X\beta = y \qquad (3)$$

12- Least Angle Direction

13- Features

$$X=\begin{bmatrix} x_{11} & x_{21} & x_{31} & \dots & x_{d1} \\ x_{12} & x_{22} & x_{32} & \dots & x_{d2} \\ \vdots & & & & \vdots \\ x_{1n} & x_{2n} & x_{3n} & \dots & x_{dn} \end{bmatrix}, \quad Y=\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$$

X_1

میزان تغییر به اندازه γ خواهد بود که با توجه به همبستگی بین متغیرها به دست می‌آید. با به دست آوردن γ، الگوریتم تخمینگر حداقل درجه به صورت همزمان با محاسبه γ برای اولین متغیر انتخابی برای اضافه کردن به مدل، دومین متغیری که باید به مدل اضافه شود را نیز مشخص می‌کند. به محض این که بیش از یک متغیر فعال داشته باشیم، الگوریتم تخمینگر حداقل درجه تقریب فعلی را در جهت زاویه برابر تغییر می‌دهد (جهتی که زاویه (همبستگی) برابر با همه متغیرهای فعال دارد). حرکت در این جهت تضمین می‌کند که همبستگی فعلی هر متغیر فعال با باقیمانده به مقدار مساوی کاهش می‌یابد (Berghen,2005).

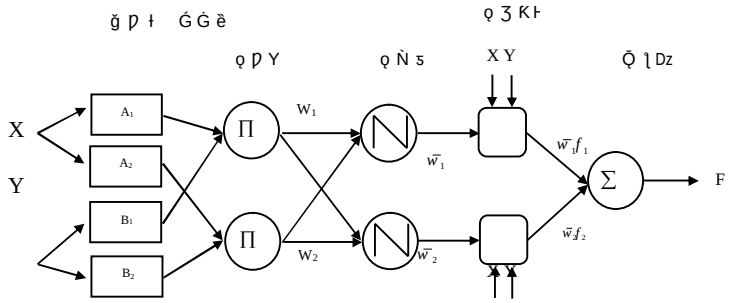
3-4. شبکه عصبی فازی

به طور خلاصه دو انگیزه اصلی برای استفاده از شبکه‌های عصبی فازی وجود دارد که عبارتند از (Nelles,2000):

1- استخراج و استفاده از دانش مقدماتی: این مدل امکان افزایش دقت را فراهم می‌کند. به منظور کاهش نیاز به تعداد و کیفیت داده‌ها و به منظور تعریف رفتار یک مدل در یک سیستم خاص و شرایطی که امکان به دست آوردن داده دیگری وجود ندارد این مدل‌ها مناسب هستند.

2- بهبود فهم فرآیند: فهم فرآیند این امکان را به وجود می‌آورد که از مدل با اطمینان بالایی استفاده شود. این خاصیت به خصوص در کاربردهایی که نیاز حیاتی به اطمینان و امنیت دارند و یا حتی در سیستم‌هایی که از اهداف گوناگون موجود در سیستم چشم‌پوشی شده مفید است.

همان‌طور که در شکل شماره 3 مشاهده می‌شود ساختار مدل عصبی فازی از پنج لایه تشکیل شده است:



شکل 3: ساختار مدل شبکهٔ عصبی فازی (Nelles,2000)

1- گره‌های ورودی: در این لایه درجه عضویت گره‌های ورودی به بازه‌های مختلف فازی با استفاده از تابع عضویت مشخص می‌گردد

$$O_i^1 = \mu_{A_i}(x) \qquad i = 1,2 \qquad (4)$$

2- لایهٔ قوانین: گره‌هایی که در این لایه قرار می‌گیرند به عنوان گره‌های $\prod$ در نظر گرفته می‌شوند که درجه فعالیت یک قانون را محاسبه می‌کند و خروجی آن‌ها به صورت زیر محاسبه می‌شود:

$$O_i^2 = w_i = \mu_{A_i}(x).\mu_{B_i}(y) \qquad i = 1,2 \qquad (5)$$

3- لایهٔ نرمالیزاسیون: هر گره در این لایه، یک گره ثابت به نام N می‌باشد و i امین گره، نسبت i امین قاعده شدت برانگیختگی را برای تمام قاعده‌های شدت برانگیختگی به صورت زیر محاسبه می‌کند:

$$O_i^3 = \bar{w}_i = \frac{w_i}{w_1 + w_2} \qquad i = 1, 2 \qquad (6)$$

4- گره‌های نتیجه: هر گره i در این لایه، یک گره منطبق با تابع گره، به صورت زیر می‌باشد:

$$O_i^4 = \bar{w}_i f_i = \bar{w}_i (p_i x + q_i y + r_i) \qquad i = 1,2 \qquad (7)$$

5- گره‌های خروجی: تنها گره این لایه یک گره ثابت به نام $\sum$ می‌باشد که تمامی خروجی‌ها را به عنوان مجموع همهٔ سیگنال‌های ورودی، به صورت زیر محاسبه می‌کند (تعداد گره‌ها برابر تعداد خروجی‌هاست):

$$O_i^5 = \sum_{i=1}^{R} \bar{w}_i f_i = \frac{\sum_{i=1}^{R} w_i f_i}{\sum_{i=1}^{R} w_i} \qquad (8)$$

در مدل عصبی فازی زمانی عمل شبیه سازی به‌درستی انجام شده است که مجموعه پارامترهای تطبیقی S_1 و مجموعه پارامترهای متعاقب S_2 طوری تخمین زده شوند که مقدار تابع خطای مدل در بخش آموزش و آزمایش به حداقل برسد. به دست آوردن مقدار این پارامترها معمولاً در دو گام صورت می‌گیرد. در گام اول که تا لایهٔ چهارم پیش می‌رود و گام رو به جلو نامیده می‌شود، مجموعه پارامترهای S_1 ثابت فرض شده و مجموعه پارامترهای S_2 با استفاده از الگوریتم حداقل مجذور خطاها محاسبه می‌شوند. در گام دوم که گام رو به عقب نامیده می‌شود، مجموعه پارامترهای S_2 ثابت فرض شده و مجموعه پارامترهای S_1 به دست می‌آیند. انتخاب تابع عضویت بر اساس آزمایش انواع مختلف توابع عضویت صورت می‌گیرد. به این معنی که توابع عضویت، هر کدام جداگانه مورد بررسی قرار می‌گیرند و مدل عصبی فازی برای هر کدام از این توابع عضویت به صورت جداگانه آموزش می‌بینند. در پایان میزان خطای مدل‌های حاصل با هم مقایسه می‌شود و تابعی که کمترین میزان خطا را در کمترین زمان آموزش حاصل کند، به عنوان تابع عضویت برگزیده خواهد شد (Jang, 1993).

5. روش تحقیق

در این تحقیق از 19 متغیر مستقل در سه گروه مالی، بنیادی و کلان استفاده شده که در جدول شماره 1 ارائه گردیده است. در بسیاری از تحقیقات تأثیر این متغیرها بر سود هر سهم بررسی شده است.

متغیر وابسته در این تحقیق سود هر سهم در سطح سالانه می‌باشد.

پس از بررسی مسئله و اهداف تحقیق و مطالعات مقدماتی درباره پاسخ‌های احتمالی، فرضیه-های تحقیق به شرح زیر قابل طرح است:

1- دقت مدل تخمینگر بردار پشتیبان نسبت به مدل تخمینگر حداقل درجه در پیش‌بینی سود هر سهم بیشتر است.

2- دقت مدل تخمینگر بردار پشتیبان نسبت به مدل شبکه عصبی فازی در پیش‌بینی سود هر سهم بیشتر است.

3- دقت مدل تخمینگر حداقل درجه نسبت به مدل شبکه عصبی فازی در پیش‌بینی سود هر سهم بیشتر است.

جدول 1. متغیرهای مورد استفاده در تحقیق

			متغیر	توضیحات
گروه اول- متغیرهای بنیادی	Zhang, Cao AND Schniederijans (2004)		INV	ارزش موجودی کالا تقسیم بر تعداد سهام عادی
			AR	حساب‌های دریافتنی تقسیم بر تعداد سهام عادی
			CAPX	اموال، ماشین‌آلات و تجهیزات تقسیم بر تعداد سهام عادی
			SA	هزینه‌های اداری و فروش تقسیم بر تعداد سهام عادی
			GM	فروش منهای بهای تمام‌شده کالای فروش رفته تقسیم بر تعداد سهام عادی
			ETR	نرخ مالیاتی عبارت است از مالیات بر درآمد تقسیم بر سود قبل از مالیات
			LFP	لگاریتم نسبت فروش به تعداد کارکنان
گروه دوم- متغیرهای مالی	Lgzian, Baqayi & Homayoni rad (2011)		ROA	سود پس از کسر مالیات تقسیم بر کل دارایی‌ها
			ΔROA	تغییرات سود پس از کسر مالیات تقسیم بر کل دارایی‌ها
			ΔATR	تغییرات فروش خالص تقسیم بر کل دارایی‌ها
			ΔITR	تغییرات فروش خالص تقسیم بر متوسط موجودی کالا
			ΔDER	تغییرات کل بدهی‌ها تقسیم بر حقوق صاحبان سهام
			ΔDA	تغییرات کل بدهی‌ها تقسیم بر کل دارایی‌ها
			ΔCR	تغییرات دارایی‌های جاری تقسیم بر بدهی‌های جاری
			ΔQR	تغییرات دارایی‌های جاری به جز موجودی‌ها تقسیم بر بدهی‌های جاری
			EPS_{-1}	سود خالص پس از کسر ادعای مقدم بر سود بر تعداد سهام عادی منتشره
گروه سوم- متغیرهای کلان	Brid (2001)		GDP	ارزش ریالی محصولات نهایی تولیدشده توسط واحدهای اقتصادی مقیم کشور در دوره زمانی معین
			INF	افزایش دائم و بی‌رویه سطح عمومی قیمت کالاها و خدمات
			M	جمع سکه و اسکناس در دست مردم به اضافه سپرده‌های دیداری بخش خصوصی نزد بانک‌ها
			EX	با مقدار واحدهای پول ملی که برای به دست آوردن یک واحد پول خارجی باید پرداخت شود

منبع: یافته‌های پژوهشگر

جامعه آماری تحقیق کلیه شرکت‌های پذیرفته‌شده در بورس اوراق بهادار تهران است که طی سال‌های 1384 تا 1391 در بورس فعال بوده‌اند. نمونه انتخابی در این تحقیق 171 شرکت (1368

سال-شرکت) در قالب 27 صنعت فعال طی دوره زمانی 1384 تا 1391 است که به صورت تصادفی و با استفاده از روش نمونه‌گیری خوشه‌ای انتخاب شده‌اند. نمونه‌های انتخابی باید حداقل دارای ویژگی‌های زیر باشند:

1- قبل از سال 1384 در بورس اوراق بهادار تهران پذیرفته‌شده باشند.

2- دارای فعالیت‌های غیر مالی باشند و همچنین فعالیت‌های آن‌ها فصلی نباشد.

3- در دوره مورد بررسی تغییر سال مالی نداشته باشند.

4- صورت‌های مالی اساسی برای سال‌های 1384 تا 1391 به صورت کامل ارائه شده باشد.

داده‌ها پس از جمع آوری، با استفاده از فرمول زیر نرمال شده‌اند تا تأثیر اعداد بزرگ کاهش یابد. در این روش داده‌ها به مجموعه جدیدی که میانگین آن‌ها صفر و واریانس آن‌ها واحد باشد تبدیل می‌شوند. برای این کار داده‌ها را از میانگین داده‌ها کم و بر انحراف معیار تقسیم کرده‌ایم.

$$X_i = \frac{(X_i - X_{max})}{\sigma}, \qquad i = 1, \dots, 9 \tag{9}$$

پس از انتخاب نمونه، برای پیاده‌سازی الگوریتم‌های تحقیق از جعبه‌ابزار مربوطه در نرم‌افزار متلب[1] استفاده شده است. برای استفاده از جعبه‌ابزار متلب نیاز است تا داده‌ها با استفاده از روش K-Fold Cross-Validation به داده‌های آموزشی و داده‌های آزمایشی تقسیم گردند. پیش‌فرض درصد ترکیب داده‌های آموزشی و آزمایشی به ترتیب 70-30 یا 80-20 درصد است که در این پژوهش از ترکیب 80-20 استفاده شده است. در این تحقیق مقدار K، 10 در نظر گرفته شده است یعنی در هر مرحله مدل 10 با اجرا می‌شود و خطای هر مرحله برابر است با میانگین خطاها در هر بار اجرای مدل. پس از آن، فرایند اصلی مدل‌سازی انجام می‌شود به این صورت که با استفاده از تکنیک‌های تخمینگر بردار پشتیبان، تخمینگر حداقل درجه و شبکه عصبی فازی الگوها و روابط بین داده‌ها (متغیرهای مستقل و متغیر وابسته) استخراج می‌گردد. در این مرحله برای مدل‌سازی از داده‌های آموزشی استفاده می‌شود. پس از به دست آوردن الگوی بین داده‌ها، با استفاده از داده-های آزمایشی دقت مدل ساخته شده تخمین زده می‌شود و در نهایت برای بررسی دقت مدل‌ها با

1- MATLAB Software

یکدیگر از معیارهای سنجش خطا استفاده می‌گردد.

در تحقیق حاضر از سه معیار رایج برای ارزیابی توان پیش‌بینی مدل‌ها استفاده شده است که عبارت‌اند از:

میانگین مجذور خطا[2]

$$MSE = \frac{\sum_{i=1}^{n}(e_i)^2}{n} \quad (10)$$

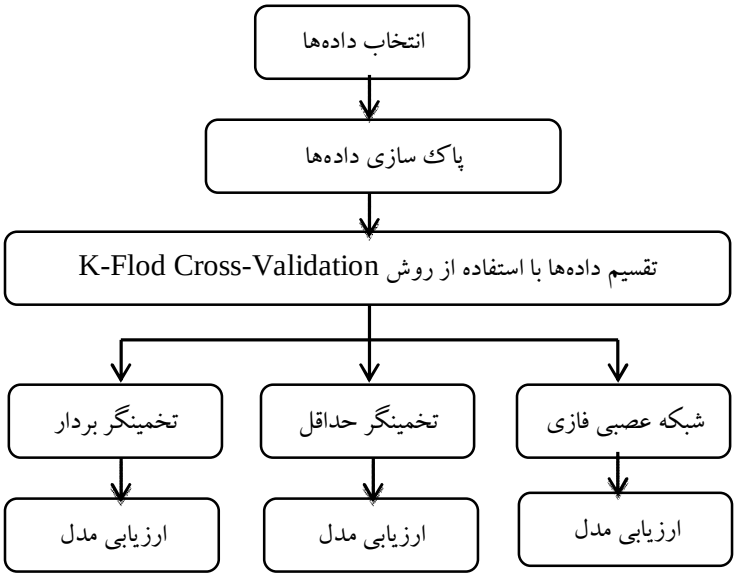
میانگین انحراف معیار خطای پیش-بینی شده[3]

$$MAD = \frac{1}{n}\sum_{i}(\hat{y}_i - \bar{\hat{y}}_i)^2 \quad (11)$$

ضریب تعیین

$$R^2 = 1 - \frac{SSE}{SSy} \quad (12)$$

که در معادلات بالا $SSE = \sum(y - \hat{y})^2$ و $SSy = \sum(y - \bar{y})^2$ است و y مقدار واقعی، $\hat{y}$ مقدار پیش‌بینی شده توسط مدل و $\bar{y}$ میانگین مقادیر واقعی y است.



شکل شماره 4. فرآیند اجرای الگوریتم‌های تحقیق

2- Mean Square Error (MSE)

3- Median Absolute Deviation (MAD)

MSE و MAD معیارهای سنجش خطا هستند که مقادیر کمتر آن‌ها نشان‌دهندهٔ قدرت بیشتر مدل خواهد بود. هم‌چنین R^2 ضریب تعیین مدل است و بین 0 تا 1 مقدار می‌پذیرد. اگر R^2 برابر 1 باشد به این معنی است که هیچ خطایی وجود ندارد که این بهترین حالت ممکن است و اگر R^2 برابر صفر باشد یعنی متغیرهای مستقل مورد مطالعه هیچ تأثیری بر برآورد متغیر وابسته ندارند. نمودار زیر فرایند کامل روش مورد استفاده را نشان می‌دهد.

جدول 2: آمار توصیفی متغیرهای تحقیق (ارقام به میلیون ریال)

	متغیر	میانه	میانگین	انحراف معیار
بنیادی	AR	1141/4	1989/40	3625/8
	INV	1243	1939	2541/7
	CAPX	1316/8	1974	2836/5
	SA	249/7	404	469/2
	GM	937/2	1370	1479/4
	ETR	0/0969	0/1753	1/75
	LFP	6/9467	7/11	1/15
مالی	ROA	0/0854	0/1016	0/14
	Δ ROA	-0/007	-0/007	0/12
	Δ ATR	0/0006	-0/004	0/25
	Δ ITR	-4/3	-20/4	2941/9
	Δ DER	0/038	0/1	20/1
	Δ QR	0/0107	0/02	0/485
	Δ CR	-0/0016	0/02	0/82
	Δ DA	0/0062	0/002	0/138
	EPS_{-1}	521/6	766/2	1126/2
کلان	GDP	3403893	3354646	13776
	INF	13/6	16/25	5/1
	M	1901366	2063704	86355
	EX	9717	9912	1019

منبع: یافته‌های پژوهشگر

6. یافته‌های تحقیق

پس از انجام مراحل اجرای روش‌ها، نتایج بدست آمده از مدل‌های تحقیق به صورت مقایسه‌ای مورد ارزیابی قرار گرفته‌اند. ابتدا آمار توصیفی برای درک مشخصات متغیرها بیان شده است سپس نتایج بررسی فرضیه‌های تحقیق ارائه گردیده است.

6-1. آمار توصیفی

برای درک بهتر مشخصات متغیرها و درک کلی جامعه، مطالعه آماره‌های توصیفی متغیرها ضروری است. خلاصهٔ ویژگی‌های آمار توصیفی مربوط به متغیرهای مورد استفاده در این تحقیق را در جدول زیر مشاهده می‌کنید.

شرکت‌های مورد مطالعه 171 شرکت می‌باشند که اطلاعات توصیفی آن‌ها برای 8 سال متوالی به شرح ذیل است.

همانگونه که در جدول 2 مشاهده می‌شود، میانگین سود هر سهم سال قبل شرکت‌های نمونه مورد بررسی برابر 766/2 ریال، میانگین فروش برابر 404 میلیون ریال و میانگین حساب‌های دریافتنی 1989/4 میلیون ریال است. این عوامل حاکی از آن است که شرکت‌های نمونه به طور متوسط در این بازده زمانی دارای وضعیت نسبتاً مناسب اقتصادی بوده‌اند. انحراف معیار سود هر سهم 1126/2، فروش 469/2 و حساب‌های دریافتنی3625/8 است که به معنای پراکندگی بسیار زیاد اقلام فوق در قلمرو زمانی تحقیق است.

6-2. آمار استنباطی و آزمون فرضیه‌ها

مدل‌های تحقیق در سه حالت اجرا و نتایج آن با استفاده از معیارهای ارزیابی عملکرد باهم مقایسه شده است. ابتدا مدل‌های تحقیق در هر حالت برای هر سال اجرا شده است. نتایج نهایی برابر با میانگین معیارهای به دست آمده طی سال‌های 1384 تا 1391 می‌باشد. داده‌های نهایی در خصوص معیارهای خطا به تفکیک متغیرهای بنیادی، مالی و مجموع متغیرها در جدول ذیل ارائه شده است.

بررسی جدول شماره 3 نشان می‌دهد که بیشترین مقدار ضریب تعیین مربوط به مدل تخمینگر حداقل درجه و در حالت استفاده از متغیرهای بنیادی است. هم‌چنین کمترین مقدار MSE و MAD مربوط به مدل تخمینگر حداقل درجه و در حالت استفاده از متغیرهای مالی می‌باشد.

جدول 3: نتایج اجرای الگوریتم‌ها

	مدل	LARS	SVR	ANFIS
متغیر	معیار	آزمون	آزمون	آزمون
بنیادی	R^2	0/877	0/679	0/814
	$MSE(10^5)$	3/506	5/684	4/373
	MAD	306/301	462/504	418/683
مالی	R^2	0/88	0/565	0/794
	$MSE(10^5)$	0/921	5/251	2/811
	MAD	206/669	495/513	325/486
مجموع متغیرها	R^2	0/789	0/661	0/582
	$MSE(10^5)$	3/414	5/642	10/013
	MAD	392/081	412/006	564/973

منبع: یافته‌های پژوهشگر

6-2-1. آزمون فرضیه اول:

دقت مدل تخمینگر بردار پشتیبان نسبت به مدل تخمینگر حداقل درجه در پیش‌بینی سود هر سهم بیشتر است.

نتایج به دست آمده از اجرای مدل‌ها به شرح زیر است:

جدول4: نتایج مدل‌های SVR و LARS

متغیر	معیار	مدل	
		SVR	LARS
بنیادی	MSE	5/684	3/506
	MAD	462/504	306/301
مالی	MSE	5/251	0/921
	MAD	495/513	206/669
مجموع	MSE	5/642	3/414
	MAD	412/006	392/081

منبع: یافته‌های پژوهشگر

6-2-1-1. حالت اول با استفاده از متغیرهای بنیادی

نتایج به دست آمده برای میانگین مجذور خطا و میانگین انحراف معیار خطای پیش‌بینی با استفاده از متغیرهای بنیادی بیانگر دقت بالاتر مدل تخمینگر حداقل درجه نسبت به مدل تخمینگر بردار پشتیبان است. میانگین مجذور خطا برای مدل تخمینگر حداقل درجه 3/506 و برای مدل تخمینگر بردار پشتیبان 5/684 و انحراف معیار خطای پیش‌بینی مدل‌ها به ترتیب 306/301 و 462/504 به دست آمده است. بنابراین فرضیه اول رد می‌شود.

6-2-1-2. حالت دوم با استفاده از متغیرهای مالی

نتایج به دست آمده برای میانگین مجذور خطا و میانگین انحراف معیار خطای پیش‌بینی با استفاده از متغیرهای مالی بیانگر دقت بالاتر مدل تخمینگر حداقل درجه نسبت به مدل تخمینگر بردار پشتیبان است. میانگین مجذور خطا برای مدل تخمینگر حداقل درجه 0/921 و برای مدل تخمینگر بردار پشتیبان 5/251 و انحراف معیار خطای پیش‌بینی مدل‌ها به ترتیب 206/669 و 495/513 به دست آمده است. بنابراین فرضیه اول رد می‌شود.

6-2-1-3. حالت سوم با استفاده از مجموع متغیرها

نتایج به دست آمده برای میانگین مجذور خطا و میانگین انحراف معیار خطای پیش‌بینی با استفاده از مجموع متغیرهای تحقیق بیانگر دقت بالاتر مدل تخمینگر حداقل درجه نسبت به مدل تخمینگر بردار پشتیبان است. میانگین مجذور خطا برای مدل تخمینگر حداقل درجه 3/414 و برای مدل تخمینگر بردار پشتیبان 5/642 و انحراف معیار خطای پیش‌بینی مدل‌ها به ترتیب 392/081 و 412/006 به دست آمده است. بنابراین فرضیه اول رد می‌شود.

در هر سه حالت فوق مشاهده می‌گردد که مدل تخمینگر حداقل درجه عملکرد بهتری نسبت به مدل تخمینگر بردار پشتیبان داشته است. همچنین، بررسی میانگین مجذور خطا برای مدل تخمینگر حداقل درجه نشان می‌دهد که در حالت استفاده از متغیرهای مالی عملکرد مدل به صورت چشم‌گیری افزایش می‌یابد.

6-2-2. آزمون فرضیه دوم:

دقت مدل تخمینگر بردار پشتیبان نسبت به مدل شبکه عصبی فازی در پیش‌بینی سود هر سهم

بیشتر است. نتایج به دست آمده از اجرای مدل‌ها به شرح زیر است:

جدول 5: نتایج مدل‌های SVR و ANFIS

متغیر	معیار	مدل	
		SVR	ANFIS
بنیادی	MSE	5/684	4/373
	MAD	462/504	418/683
مالی	MSE	5/251	2/811
	MAD	495/513	325/486
مجموع	MSE	5/642	10/013
	MAD	412/006	564/973

منبع: یافته‌های پژوهشگر

6-2-2-1. حالت اول با استفاده از متغیرهای بنیادی

نتایج به دست آمده برای میانگین مجذور خطا و میانگین انحراف معیار خطای پیش‌بینی با استفاده از متغیرهای بنیادی بیانگر دقت بالاتر مدل شبکه عصبی فازی نسبت به مدل تخمینگر بردار پشتیبان است. میانگین مجذور خطا برای مدل شبکه عصبی فازی 4/373 و برای مدل تخمینگر بردار پشتیبان 5/684 و انحراف معیار خطای پیش‌بینی مدل‌ها به ترتیب 418/683 و 562/504 به دست آمده است. بنابراین فرضیه دوم رد می‌شود.

6-2-2-2. حالت دوم با استفاده از متغیرهای مالی

نتایج به دست آمده برای میانگین مجذور خطا و میانگین انحراف معیار خطای پیش‌بینی با استفاده از متغیرهای مالی بیانگر دقت بالاتر مدل شبکه عصبی فازی نسبت به مدل تخمینگر بردار پشتیبان است. میانگین مجذور خطا برای مدل شبکه عصبی فازی 2/811 و برای مدل تخمینگر بردار پشتیبان 5/251 و انحراف معیار خطای پیش‌بینی مدل‌ها به ترتیب 325/486 و 495/513 به دست آمده است. بنابراین فرضیه دوم رد می‌شود.

6-2-2-3. حالت سوم با استفاده از مجموع متغیرها

نتایج به دست آمده برای میانگین مجذور خطا و میانگین انحراف معیار خطای پیش‌بینی با

استفاده از مجموع متغیرهای تحقیق بیان‌گر دقت بالاتر مدل تخمینگر بردار پشتیبان نسبت به مدل شبکه عصبی فازی است. میانگین مجذور خطا برای مدل تخمینگر بردار پشتیبان 5/642 و برای مدل شبکه عصبی فازی 10/013 و انحراف معیار خطای پیش‌بینی مدل‌ها به ترتیب 412/006 و 564/973 به دست آمده است. بنابراین فرضیه پذیرفته می‌شود.

بررسی فرضیه‌های فوق نشان می‌دهد که عملکرد مدل شبکه عصبی فازی با استفاده از متغیرهای مالی و بنیادی بهتر از مدل تخمینگر پشتیبان است. واردکردن همزمان متغیرهای تحقیق به دو مدل فوق موجب عملکرد بهتر مدل تخمینگر بردار پشتیبان شده است.

6-2-3. آزمون فرضیه سوم:

دقت مدل تخمینگر حداقل درجه نسبت به مدل شبکه عصبی فازی در پیش‌بینی سود هر سهم بیشتر است. نتایج به دست آمده از اجرای مدل‌ها به شرح زیر است:

6-2-3-1. حالت اول با استفاده از متغیرهای بنیادی

نتایج به دست آمده برای میانگین مجذور خطا و میانگین انحراف معیار خطای پیش‌بینی با استفاده از متغیرهای بنیادی بیان‌گر دقت بالاتر مدل تخمینگر حداقل درجه نسبت به مدل شبکه عصبی فازی است. میانگین مجذور خطا برای مدل تخمینگر حداقل درجه 3/506 و برای مدل شبکه عصبی فازی 4/373 و انحراف معیار خطای پیش‌بینی مدل‌ها به ترتیب 306/301 و 418/683 به دست آمده است. بنابراین فرضیه سوم پذیرفته می‌شود.

جدول 6: نتایج اجرای مدل‌های LARS و ANFIS

متغیر	معیار	مدل	
		LARS	ANFIS
بنیادی	MSE	3/506	4/373
	MAD	306/301	418/683
مالی	MSE	0/921	2/811
	MAD	206/669	325/486
مجموع	MSE	3/414	10/013
	MAD	392/081	564/973

منبع: یافته‌های پژوهشگر

6-2-3-2. حالت دوم با استفاده از متغیرهای مالی

نتایج به دست آمده برای میانگین مجذور خطا و میانگین انحراف معیار خطای پیش‌بینی با استفاده از متغیرهای مالی بیانگر دقت بالاتر مدل تخمینگر حداقل درجه نسبت به مدل شبکه عصبی فازی است. میانگین مجذور خطا برای مدل تخمینگر حداقل درجه 0/921 و برای مدل شبکه عصبی فازی 2/811 و انحراف معیار خطای پیش‌بینی مدل‌ها به ترتیب 206/669 و 325/486 به دست آمده است. بنابراین فرضیه سوم پذیرفته می‌شود.

6-2-3-3. حالت سوم با استفاده از مجموع متغیرهای تحقیق

نتایج به دست آمده برای میانگین مجذور خطا و میانگین انحراف معیار خطای پیش‌بینی با استفاده از مجموع متغیرهای تحقیق بیانگر دقت بالاتر مدل تخمینگر حداقل درجه نسبت به مدل شبکه عصبی فازی است. میانگین مجذور خطا برای مدل تخمینگر حداقل درجه 3/414 و برای شبکه عصبی فازی 10/013 و انحراف معیار خطای پیش‌بینی مدل‌ها به ترتیب 392/081 و 564/973 به دست آمده است. بنابراین فرضیه سوم پذیرفته می‌شود.

نتایج بدست آمده از بررسی فرضیه‌های تحقیق به صورت خلاصه در جدول شماره 7 ارائه گردیده است.

جدول 7: خطای نهایی الگوریتم‌های تحقیق

متغیر	معیار	مدل		
		SVR	LARS	ANFIS
بنیادی	MSE	5/684	3/506	4/373
	MAD	462/504	306/301	418/683
مالی	MSE	5/251	0/921	2/811
	MAD	495/513	206/669	325/486
مجموع	MSE	5/642	3/414	10/013
	MAD	412/006	392/081	564/973

منبع: یافته‌های پژوهشگر

7. نتیجه‌گیری

در این تحقیق مدل‌ها به صورت سالانه مورد آزمون قرارگرفته‌اند در هر سال مدل‌ها در سه حالت اجرا و نتایج مقایسه شده‌اند. نهایتاً میانگین خطاها در هر سال مبنای تعیین مدل دقیق‌تر در هر

حالت قرار گرفته است. بررسی خطای نهایی مدل‌ها در جدول 7 نشان می‌دهد که مدل تخمینگر حداقل درجه توان پیش‌بینی بهتری نسبت به مدل شبکه عصبی فازی و تخمینگر بردار پشتیبان دارد. همچنین بررسی میانگین خطا، برای مدل تخمینگر حداقل درجه نشان می‌دهد که استفاده از متغیرهای مالی باعث افزایش توانایی پیش‌بینی مدل شده است.

در مطالعات صورت گرفته بر مقایسه قدرت پیش‌بینی مدل‌های آماری و هوشمند تقریباً در اکثر مطالعات، روش‌های هوشمند در مقایسه با روش‌های آماری عملکرد بهتری داشته‌اند. مطالعات Raei, And Fallahpour. (2009)، Zhang, Cao AND Schniederijans (2004)، Chang and Wang (2007) که به مقایسه مدل‌های هوشمند و آماری پرداخته‌اند نشان می‌دهد که مدل‌های هوشمند نسبت به مدل‌های آماری برتری دارند. نتایج این تحقیق حاکی از برتری مدل تخمینگر بردار پشتیبان نسبت به مدل‌های هوشمند تخمینگر بردار پشتیبان و شبکه عصبی فازی است که مغایر با تحقیقات قبلی است. همچنین مطالعات Kara, Melek & Omer (2011) ، و Chang and Wang (2007) که به مقایسه مدل‌های هوشمند پرداخته‌اند. نتایج مطالعات Chang and Wang (2007)با نتایج این تحقیق همخوانی دارد و بیانگر عملکرد بهتر مدل تخمینگر بردار پشتیبان است ولی نتیجه مطالعه Kara, Melek & Omer (2011) نشان می‌دهد که مدل شبکه عصبی عملکرد بهتری دارد که با تحقیق حاضر همخوانی ندارد. نتایج تحقیق Nassirzadeh, F. And Nikravesh (2013)که به مقایسه توان پیش‌بینی سه مدل تخمینگر بردار پشتیبان، تخمینگر حداقل درجه و شبکه عصبی فازی پرداخته است نشان دهنده توان پیش‌بینی بهتر مدل تخمینگر بردار پشتیبان نسبت به سایر مدل‌های تحقیق است، نتایج به دست آمده مغایر با تحقیق حاضر است.

8. پیشنهادهای مطالعاتی

1- استفاده از سایر مدل‌ها مانند موجک‌ها در پیش‌بینی سود هر سهم، همچنین استفاده از الگوریتم‌های بهینه‌سازی مانند الگوریتم ژنتیک، الگوریتم مورچگان و الگوریتم زنبورعسل جهت بهینه‌سازی ورودی‌های سیستم هوشمند، می‌تواند موردتوجه قرار گیرد.

References

[1] Bagheri, A., Mohammadi Peyhani, H. And Akbari , M. (2014). Financial

forecasting using ANFIS networks with Quantum-behaved Particle Swarm Optimization. Expert Systems with Applications. 41(14). 6235–6250.

[2] Berghen, F. (2005). LARS Library: Least Angle Regression Stagewise Library. IRIDIA, Universit Libre de Bruxelles. November 22.

[3] Bird, G. (2001). "IMF Programs: Do They Work? Can They be Made to Work Better?", World Development, Volume 29, Issue 11, November 2001, Pages 1849–1865.

[4] Chang, P.C. and Wang, Y.W. (2007). "Fuzzy Delphi and back-propagation model for sales forecasting in PCB industry," Expert Systems with Applications, Vol. 30.

[5] Efron, B. Hastie, T. Johnstone, I. AND Tibshirani, R. (2004). Least angle regression. The Annals of Statistics, 32 (2). 407-499.

[6] Ertimur,Y. Livnat, J. and Martkainen, M. (2003). Differential markel reactions to revenue and expense surprises. Riverview of Accounting Studies, 2(3).185-211.

[7] Ghvidel, S. Otadi, M. And Mosleh, M. (2012). Estimation of Efficient market using ANFIS networks. Journal of Economic Modeling Research, 3 (4). 83-101. In Persian.

[8] Haghighat, H. Bakhtiary, M. And Beheshtipour, M. (2012). Setting the priority of the factors influencing the amount of accuracy of profit prediction of accepted companies in Tehran Stock Exchange in the time of capital growth. Journal of The Accounting AND Auditing Review, 18 (65). 41-62. In Persian.

[9] Jang, J. (1993). ANFIS: Adaptive-Network-Based Fuzzy Inference System. IEEE Transactions on Systems, Man, and Cybernetics, 23(3). 665-685.

[10] Kara, Y., Melek A, Omer K. (2011). Predicting direction of stock price index movement using artificial neuralnetworks and support vector machines: The sample of the Istanbul Stock Exchange. Expert Systems with Applications, 38(5).5311–5319.

[11] Kelly, L, J. (2007). Measuring the Economic Stock of Money. Munich Personal RePEc Archive. Online at http://mpra.ub.uni-muenchen.de/4914/.

[12] Kosaka, M., Mizuno, H., Sasaki, T., Someya, R & Hamada, N. (1991). Application of fuzzy logic/neural network to securities trading decision support system. Proceedings of the IEEE conference on systems, Man and Cybernetics 3:13-16 October. 1913-1918.

[13] Lagzian, m,. Baghaei, J & Homayoni Rad, M. (2011). The Effect on Financial Ratios to Predict Company Profits and Stock Returns Master of Financial Management. Knowledge And development. Vol 1.

[14] Lee ،I. ،Taylor ،N. ،Yee ،C. ،& Yee ،M. (1996). Prospectus Forecast Earning; Evidence and Explanations. Australian Accounting Review, 3. 21-32.

[15] Ming, C. (2009). Using support vector machine with a hybrid feature selection method to the stock trend prediction. Expert Systems with Applications, 36(8). 10896-10904.
[16] Mashayekh SH. And Shahrokhi, S, S. (2008). Evaluation the affecting factors on accuracy of earnings forecast by management. Journal of The Accounting And Auditing review, 14 (50). 65-82. In Persian.
[17] Mojtahedzade, V. And Nazari tanha, N. (2009). Review of the Relationship between predicted earnings per share Variations with Variations stock price. Journal Of Humanities And Social Sciences, 28. 117-134. In Persian.
[18] Nassirzadeh, F. And Nikravesh, Z. (2013). A Comparison of models Support Vector Regression, Least Angle Regression and Adaptive Neuro Fuzzy Inference System for Stock price forecasting. Unpublished Master, Islamic Azad University, Iran, Mashhad. In Persian.
[19] Nassirzade, F and Nikravesh, Z. (2013). Examintion the accuracy of Data Mining models to forecast the stock price. The 11th iranian academic accounting conferance, 9-10 october, mashhad. http://www.civilica.com/Paper-IAAC11-IAAC11_024.html. In Persian.
[20] Nelles, O. (2001). Non Linear System Identification, From Classical Approaches to Neural Networks and Fuzzy Models. (7rd ed). US: Springer.
[21] Ou, P. and Wang, H. (2014). Volatility Modelling and Prediction by Hybrid Support Vector Regression with Chaotic Genetic Algorithms. The International Arab Journal of Information Technology. 11(3). 287-292.
[22] Raei, R. And Fallahpour, S. (2009). Predicting financial distress of manufacturing firms using artificial neural networks. Journal of Financial Research, 17. 39-69. In Persian.
[23] Rahmani, A. And Bakhtaki, B. (2007). EPS: challenges for calculation and opportunities for improvement. Journal of Hesabdar, 21 and 175. 9-18 and 61-68. In Persian.
[24] Vakilian Aghoei, M. Vadiei, M, H. And Hoseini maasoom, M. (2010). The Relationship between Economic Value Added (EVA) and Residual Income (RI) in the Predicting Future Earning Per Share (EPS). Jornal of Financial Research, 11 (27). 111-122. In Persian.
[25] Smola, A, J. AND Schölkop, f. (2004). A tutorial on support vector regression. Statistics and omputing. 14. 199-222.
[26] Svalina, I., Galzina, V., Lujić, R. and Šimunović, G. (2013). An adaptive network-based fuzzy inference system (ANFIS) for the forecasting: The case of close price indices. Expert Systems with Applications, 40(15). 6055–6063.
[27] Sinaii, H., Mortazavi, S. And Teymori assl, Y. (2005). Predicting financial Tehran Stock Exchange Index using artificial neural networks. Journal of The Accounting AND Auditing Review, 41. 59-83. In Persian.
[28] Thammano, A. (1999). Neuro-fuzzy mode for stock market prediction. In proceedings of the artificial neural networks in engineering conference

(AVVIE 99). New York, ASME press, 587-591.

[29] Wang, Z. AND Tan, S. (2010). Identifying idiosyncratic stock return indicators from large financial factor set via least angle regression. Expert Systems with Applications, 36. 8350–8355.

[30] Xion, T., Bao, Y. And Hu, Z. (2014). Multiple-output support vector regression with a firefly algorithm for interval-valued stock price index forecasting. Knowledge-Based Systems, 55. 87–100.

[31] Zhang, W, Cao, Q AND Schniederijans, M.J. (2004). Neural network earnings per share forecasting models: A comparative analysis of alternative methods. Decision Sciences, 35 (2). 205–237.