



فصلنامه مهندسی مالی و مدیریت اوراق بهادار

دوره چهارده، شماره پنجاه و شش، پائیز ۱۴۰۲

نوع مقاله : علمی پژوهشی

صفحات : ۱۹-۳۶

مقایسه مدل‌های مختلف یادگیری ماشین در پیش‌بینی شاخص بازار سهام

مریم سهرابی^۱

سید مظفر میربرگ کار^۲

تاریخ دریافت مقاله : ۱۴۰۰/۰۹/۰۱ تاریخ پذیرش مقاله : ۱۴۰۱/۰۹/۰۶ ابراهیم چیرانی^۳

سینا خردیار^۴

چکیده

پیش‌بینی سری‌های زمانی بازارهای مالی مسئله‌ای چالش‌برانگیز در حوزه مطالعات تخصصی سری‌های زمانی محسوب می‌شود و نظر بسیاری از پژوهشگران را به خود جلب نموده است. این موضوع با توجه به حضور کلان داده‌ها موجب رشد تحولات در زمینه مدل‌های یادگیری ماشین شده است. با توجه به اهمیت این موضوع، در این پژوهش با بهره‌گیری از مقایسه مدل‌های مختلف یادگیری ماشین از قبیل رویکردهای جنگل تصادفی، ماشین بردار پشتیبان، شبکه عصبی مصنوعی و شبکه عصبی بازگشتی مبتنی بر یادگیری عمیق به بررسی توانایی مدل‌های مختلف یادگیری ماشین در پیش‌بینی شاخص کل بورس اوراق بهادار تهران در طی دوره ۱۳۹۲ تا ۱۳۹۹ پرداخته شده است. نتایج پیش‌بینی دوره‌های ۱، ۳ و ۶ روزه برای دوره خارج از نمونه نشان می‌دهد که روش یادگیری ماشین مبتنی بر شبکه عصبی بازگشتی حافظه طولانی کوتاه‌مدت (LSTM) در مقایسه با سایر مدل‌های مورد بررسی نتیجه بهتری داشته است.

کلمات کلیدی

بازار سهام، پیش‌بینی، یادگیری ماشین، شبکه عصبی بازگشتی

۱- گروه مدیریت بازرگانی، واحد رشت، دانشگاه آزاد اسلامی، رشت، ایران. sohrabi.1446@gmail.com

۲- گروه مدیریت بازرگانی، واحد رشت، دانشگاه آزاد اسلامی، رشت، ایران. (نویسنده مسئول) mirbargkar@yahoo.com

۳- گروه مدیریت بازرگانی، واحد رشت، دانشگاه آزاد اسلامی، رشت، ایران. chirani@iaurasht.ac.ir

۴- گروه حسابداری، واحد رشت، دانشگاه آزاد اسلامی، رشت، ایران. sinakheradyar@gmail.com

بر اساس فرضیه بازار کارا، قیمت‌های بازار به طور کامل تمام اطلاعات موجود را نشان می‌دهد و قیمت‌ها با اطلاعات جدید تغییر می‌کنند. بازار کارا بازاری است که در آن، اطلاعات موجود بلافاصله بر قیمت اوراق بهادار تأثیر می‌گذارد. مفهوم بازار کارا بر این فرض استوار است که سرمایه‌گذاران در تصمیمات خرید و فروش خود، تمامی اطلاعات مربوط را در قیمت سهام لحاظ خواهند نمود. در چنین بازاری، قیمت اوراق بهادار عادلانه بوده و به ارزش ذاتی آن نزدیک است. مفهوم بازار کارا، یک ادعا یا توقع نیست؛ بلکه تغییرات کاملی است که در اثر انتشار اطلاعات جدید در قیمت سهام صورت می‌گیرد (تهرانی و نوربخش، ۱۳۸۲). بر اساس این فرضیه، بازده بیشتر از طریق انتخاب سهام به روش‌های هوشمندانه یا تکنیک‌های دیگر امکانپذیر نیست. تنها راهی که خریداران می‌توانند سود بیشتری کسب کنند، انتخاب سهام با ریسک بیشتر یا به صورت شانس است (فاما و مالکیل، ۱۹۷۰). در هر بازار کارا، اگر انتظارات و اطلاعات همه شرکت‌کنندگان از بازار در قیمت‌ها منعکس شود، نوسان قیمت‌ها پیش‌بینی‌ناپذیر است. نظریه دیگری که با این نظریه سازگار است، نظریه گام تصادفی نام دارد که بر اساس آن، روند تغییرات قیمت سهام به صورت تصادفی بوده و پیش‌بینی‌پذیر نیست (فاما، ۱۹۹۵). در حال حاضر مطالعات مختلفی همچون کریستین و رونالدو^۱ (۲۰۰۹)، مارکوات و همکاران^۲ (۲۰۰۹)، ژیانو و همکاران^۳ (۲۰۱۳)، سرولوروی و همکاران^۴ (۲۰۱۵)، چیانگ و همکاران^۵ (۲۰۱۶)، ژونگ و اینک^۶ (۲۰۱۷) و کیم و چانگ^۷ (۲۰۱۸) رفتار تصادفی قیمت سهام را رد کرده‌اند. همچنین کاربرد هوش مصنوعی در مسائل مالی موجب تقویت این نظریه شده که بازار ممکن است همیشه کاملاً کارا نباشد و بتوان قیمت‌های آینده را از داده‌های تاریخی با تکنیک‌های مختلف پیش‌بینی نمود.

برای پیش‌بینی بازار سهام به کمک داده‌های تاریخی، روش‌های مختلفی وجود دارد. در یک طبقه‌بندی، این روش‌ها در دو گروه خطی و غیرخطی قرار می‌گیرند و در نوع دیگر، به روش‌های یادگیری ماشین و آماری دسته‌بندی می‌شوند؛ اما دسته‌بندی مناسب، گروه‌بندی آنها به روش‌های هوشمند و کلاسیک است (پاتل و همکاران^۸، ۲۰۱۵). در روش‌های پیش‌بینی کلاسیک، فرض بر این است که مقادیر آینده قیمت، روند خطی مقادیر گذشته را دنبال می‌کند و مدل‌های رگرسیون و ARMA، GARCH و ماشین‌های بردار پشتیبان و منطق فازی در این دسته جای دارند. شبکه‌های عصبی و مدل‌های یادگیری جمعی در دسته روش‌های هوشمند قرار دارند (کوالکانت و همکاران^۹، ۲۰۱۶). نتیجه مقایسه‌ها نشان داده است که روش‌های هوشمند با غلبه بر محدودیت‌های موجود در مدل‌های خطی

مقایسه مدل‌های مختلف یادگیری ماشین در پیش‌بینی.../سهرابی، میربرگ کار، چیرانی و خردیار

در مقایسه با روش‌های کلاسیک، برای استخراج مدل از داده‌ها توانایی بهتر و برای پیش‌بینی دقت بیشتری دارند (آدبی، ادمووی و آیو، ۲۰۱۴).

تحقیقات در زمینه یادگیری ماشین هم‌اکنون در حال پیشرفت است و بررسی و توانایی دقت آنها در بررسی و پیش‌بینی سری‌های زمانی بیش از پیش دارای اهمیت است. بر این اساس، تحقیق در این زمینه نه تنها در حوزه بررسی پیش‌بینی قیمت سهام، بلکه در بررسی جهش‌های ناگهانی می‌تواند یکی از چالش‌ها و ضروریات در این زمینه باشد. برای غلبه بر این چالش‌ها، پس از بررسی تحقیقات انجام شده در این حوزه، در این پژوهش به تبیین و بررسی توانایی مدل‌های یادگیری عمیق و شبکه عصبی بازگشتی برای پیش‌بینی بازار سهام پرداخته شده است.

در ادامه، پس از مروری بر مبانی نظری و تجربی مبحث یاد شده، به روش‌شناسی پژوهش اشاره خواهد شد. سپس یافته‌های تجربی حاصل از پژوهش حاضر ارائه و در انتها نتیجه‌گیری و بحث موضوعی مورد بررسی قرار خواهد گرفت.

مروری بر پیشینه پژوهش

بازارهای مالی نقش حائز اهمیتی در سازماندهی تشکیلات اقتصادی و اجتماعی جامعه مدرن ایفا می‌نمایند. چنانچه رابطه منطقی بین بازار مالی و سایر بخش‌های اقتصادی وجود نداشته باشد، احتمال بروز اختلالات و نقصان‌هایی در ساز و کار اقتصاد وجود خواهد داشت. بدون شک امروزه در جهان، بیشترین مقدار سرمایه از طریق بازارهای بورس تبادل می‌شود. اقتصادهای ملی نیز به شدت متأثر از عملکرد پیش‌بینی قیمت سهام در بازار بورس هستند و همین امر باعث شده مورد علاقه فعالان این بخش در راستای کنترل ریسک و افزایش بازدهی باشد. یکی از مهمترین اطلاعات در بازار بورس اوراق بهادار برای سرمایه‌گذاران، اطلاعات قیمت سهام است. بنابراین پیش‌بینی حرکت قیمتی سهام نه فقط بسیار چالش برانگیز هست، بلکه یکی از دغدغه‌های اصلی سهامداران و فعالان بازار سرمایه است (سارنج و همکاران، ۱۳۹۶). سرمایه‌گذاری در بورس بخش مهمی از اقتصاد را تشکیل می‌دهد. لذا موضوع پیش‌بینی به ویژه در کشورهای در حال توسعه، جهت مدیریت صحیح بورس اوراق بهادار برای رسیدن به رشد اقتصادی و توسعه پایدار اهمیت به سزایی را دارا است که تصمیم‌گیری در شرایط نامطمئن را برای تصمیم‌گیران اجرایی بورس هموار می‌نماید. همچنین سرمایه‌گذاران نیز می‌توانند قیمت سهام و یا شاخص کل را پیش‌بینی نموده و نسبت به آن تصمیمات منطقی اتخاذ نمایند. پیش‌بینی قیمت یا بازده سهام کار پیچیده‌ای است؛ زیرا عوامل بازاری بسیاری در تعیین آن دخالت دارند که تمام این عوامل را نمی‌توان صرفاً در تحلیل تکنیکی که فقط داده‌های تاریخی مربوط به حرکت قیمت و حجم معاملات سهام را برای

پیش‌بینی حرکت آتی قیمت مطالعه می‌کند، در نظر گرفت (تهرانی و عباسیون، ۱۳۸۷). اخیراً بازارهای مالی شاهد افزایش روزافزون محصولات و خدمات پیشرفته هوش مصنوعی بوده است. هوش مصنوعی (AI) و فناوری یادگیری ماشینی طی سال‌های اخیر به سرعت در حال پیشرفت هستند؛ به طوری که با دسترسی بیشتر به داده‌ها و پیشرفت در قابلیت محاسبات ریاضیاتی و راه‌حل‌های ذخیره‌سازی، در حال رشد و ارائه امکانات جدید می‌باشند. الگوریتم‌های بسیار مختلفی برای یادگیری ماشین وجود دارد و هر روزه صدها الگوریتم جدید نیز تولید می‌شوند و به طور معمول توسط سبک یادگیری (learning style) و یا با توجه به شباهت‌شان در فرم و عملکرد (مانند طبقه‌بندی، برگشت، درخت تصمیم‌گیری، دسته کردن، یادگیری عمیق و ...) گروه‌بندی می‌شوند. در بخشی از ادبیات شکل گرفته پیرامون این موضوع مطالعات مختص به شبکه عصبی است. مدل‌های شبکه عصبی قادرند به طور مستقل از داده‌ها بیاموزند و بطور مداوم عملکرد و دقت در تصمیم‌گیری را در رابطه با یک کار خاص بهبود بخشند. در ادامه، به مطالعات داخلی و خارجی می‌پردازیم.

شمس و پارسائیان (۱۳۹۱)، به مقایسه مدل فاما و فرنچ و شبکه‌های عصبی مصنوعی در پیش‌بینی بازده سهام در بورس تهران پرداختند. نتایج این پژوهش نشان از برتری مدل شبکه‌های عصبی مصنوعی بر مدل فاما و فرنچ دارد.

زراءنژاد و رئوفی (۱۳۹۱)، در مقاله‌ای با عنوان "ارزیابی و مقایسه عملکرد مدل خودرگرسیون میانگین متحرک انباشته و شبکه عصبی فازی برای پیش‌بینی روزانه قیمت طلا"، از یک شبکه عصبی فازی ANFIS بر مبنای مدل تاکاگی - سوگنو و همچنین الگوریتم یادگیری ترکیبی پس انتشار و حداقل مربعات خطا در جهت بهبود دقت پیش‌بینی و افزایش سرعت همگرایی استفاده نموده‌اند. بازه زمانی به کار گرفته شده برای پیش‌بینی قیمت روزانه طلا از ۲۰۱۰/۰۷/۱۲ تا ۲۰۱۲/۰۵/۱۸ است. هدف این مطالعه ارزیابی و مقایسه عملکرد دو روش خودرگرسیون میانگین متحرک انباشته (ARIMA) و شبکه عصبی فازی ANFIS در پیش‌بینی قیمت روزانه طلا است. نتایج تحقیق با توجه به معیارهای متداول ارزیابی خطای پیش‌بینی نشان داد که مدل شبکه عصبی فازی ANFIS نسبت به مدل ARIMA پیش‌بینی دقیق‌تری ارائه داده است.

فلاح‌پور و همکاران (۱۳۹۲)، در پژوهش خود به پیش‌بینی روند حرکتی قیمت سهام با استفاده از ماشین بردار پشتیبان بر پایه الگوریتم ژنتیک در بورس اوراق بهادار تهران پرداختند. این پژوهش نشان داد که ماشین بردار پشتیبان بر پایه الگوریتم ژنتیک از دقت بالاتری نسبت به ماشین بردار پشتیبان ساده در پیش‌بینی حرکت قیمت سهام برخوردار است.

مقایسه مدل‌های مختلف یادگیری ماشین در پیش‌بینی.../سهرابی، میربرگ کار، چیرانی و خردیار

مطالعات خارجی در این زمینه در طی سال‌های گذشته روند صعودی داشته است. با توجه به حجم متفاوت مطالعات در این خصوص، در ادامه به نزدیک‌ترین مطالعات صورت گرفته اشاره شده است.

به طور خاص باقری‌های و همکاران^{۱۱} (۲۰۱۴)، یک سیستم ترکیبی مبتنی بر شبکه عصبی را برای پیش‌بینی سری‌های زمانی نرخ ارز پیشنهاد می‌کنند که ثابت می‌کند در پیش‌بینی قیمت مفید و کارآمد است. در این راستا مطالعات سرولو روی و همکاران^{۱۲} (۲۰۱۵) با اتخاذ چشم‌انداز متفاوت، الگوهای معاملاتی را شناسایی می‌کنند که شواهدی مبنی بر عدم تأیید فرضیه بازار کارآمد را ارائه می‌دهند. بعلاوه، چیانگ و همکاران^{۱۳} (۲۰۱۶) پیشنهاد می‌کنند که معامله‌گران می‌توانند هنگام استفاده از تصمیمات پیشنهادی خود از سیستم پشتیبانی تصمیم سازگار مبتنی بر رویکرد ماشین بردار پشتیبان، بازده بالاتری را تجربه نمایند. بعلاوه نتایج چیانگ و همکاران (۲۰۱۶) حاکی از آن است که الگوریتم‌های طبقه‌بندی مبتنی بر شبکه‌های عصبی، به دلیل ظرفیت آنها در به دست آوردن رابط غیرخطی، بازدهی‌های بیشتری را با توجه به ریسک تنظیم می‌کنند.

در رشته‌ای دیگر از ادبیات موجود در اینگونه مطالعات، به ترکیب رویکردهای پارامتریک و ناپارامتریک پرداخته شده است؛ به طوری که ژونگ و اینک^{۱۴} (۲۰۱۷) تصدیق می‌کنند که ترکیبی از روش‌ها همچون شبکه عصبی و PCA می‌تواند نتایج بهبود یافته‌ای را به همراه داشته باشد. به همین ترتیب، کیم و چانگ^{۱۵} (۲۰۱۸) برای پیش‌بینی نوسان قیمت سهام، یک سری مدل‌های ترکیبی از حافظه کوتاه‌مدت (LSTM) و GARCH را پیشنهاد داده و بررسی می‌کنند که طوری که نتایج گویای این واقعیت است که این مدل‌ها منجر به کمترین خطاهای پیش‌بینی می‌شوند.

گروه بزرگی از مطالعات مربوط به سرایت مالی بر پیش‌بینی واکنش بازارها در برابر اخبار غیرمنتظره و شوک‌های برون‌زا با استفاده از مدل‌های عاملی متمرکز است. این مدل‌ها یک سیستم معادلات است که از شاخص‌های بازار مالی به عنوان متغیر درون‌زا و از فاکتورهای جهانی و برداری از عوامل خارج از کشور به عنوان متغیرهای برون‌زا استفاده می‌نماید. اثر سرایت در دوره‌های بحران با آزمایش اینکه آیا پارامترهای برون‌زا تأثیر قابل توجه بر بازار داخلی دارد مشخص شده و مدل‌سازی می‌شود (دانجی و مارتین^{۱۶}، ۲۰۰۶). به این ترتیب، پس از یک رویداد بحرانی، پیوندهای اقتصادی کلان بین دو کشور از قبل ایجاد شده می‌تواند تقویت شود (افزایش همبستگی)، شکسته شود (کاهش همبستگی) یا ثابت بماند (فوربس و ریگوبون^{۱۷}، ۲۰۰۲). اولین تلاش برای استفاده از روش‌های انتخاب گسسته چند جمله‌ای برای مدل‌سازی سرایت مالی در دنباله توزیع را می‌توان در مطالعات بانه و همکاران^{۱۸} (۲۰۰۳) و همچنین کریستین و رونالدو^{۱۹} (۲۰۰۹) یافت. آنها از یک چارچوب اقتصادسنجی مشابه برای تجزیه و تحلیل ادغام

مالی بازارهای سهام کشورهای عضو اتحادیه اروپا جدید در مقایسه با کشورهای عضو اتحادیه اروپا قدیمی استفاده کردند. سرانجام، مارکوات و همکاران^{۲۰} (۲۰۰۹) برای طبقه‌بندی سقوط در بازار سهام برای یافتن دلایل داخلی، منطقه‌ای یا جهانی، به بررسی تأثیرات دومینوی آنها با استفاده از رگرسیون لاجیت پرداختند. ژائو و همکاران^{۲۱} (۲۰۱۳) نیز بر اساس شبکه‌های عصبی مصنوعی (ANNs)، ماشین‌های بردار پشتیبانی (SVM) و جنگل‌های تصادفی (RFs)، به پیش‌بینی بحران مالی پرداختند. نتایج حاصله نشان داد که رویکرد شبکه عصبی مصنوعی در شرایط بهتری به نسبت سایر مدل‌های مورد بررسی است. استفاده از تکنیک‌های یادگیری ماشین در توسعه سیستم‌های هشدار سریع برای بحران مالی در ادبیات موجود محدود است. در این راستا، کانیت و همکاران^{۲۲} (۲۰۱۴) سه سیستم هشدار زودهنگام مبتنی بر شبکه‌های عصبی مصنوعی (ANN)، درختان تصمیم‌گیری و رگرسیون لاجیت را توسعه دادند و آنها را بر روی ترکیه آزمایش کردند. شبکه‌های عصبی مصنوعی بهترین عملکرد را در تجزیه و تحلیل خود داشتند. آتسالاکیس و همکاران^{۲۳} (۲۰۱۶)، بر پیش‌بینی یک روزه بازار سهام در دوره‌های استرس تمرکز نمودند و از یک روش مدل‌سازی عصبی - فازی استفاده کردند. ازدکین و همکاران^{۲۴} (۲۰۱۶) نیز بر پیش‌بینی قیمت روز سهام تمرکز کردند. آنها در مطالعه خود سیستم‌های استنتاج عصبی فازی تطبیقی، شبکه‌های عصبی مصنوعی و ماشین‌های بردار پشتیبانی را بررسی نمودند. دوبکه و همکاران^{۲۵} (۲۰۱۷)، رویکرد درخت تصمیم‌گیری رگرسیون را برای پیش‌بینی رکود اقتصادی به کار گرفتند. سرانجام، دابروسکی و همکاران^{۲۶} (۲۰۱۶)، مدل‌های شبکه پویای بیزین را بررسی کرده و نشان دادند که می‌توانند نسبت به رگرسیون لجستیک به طور قابل توجهی هشدارهای دقیق‌تری ارائه دهند. در نهایت در این زمینه چاتزیس و همکاران^{۲۷} (۲۰۱۸)، به بررسی و پیش‌بینی بحران مالی در بازار سهام با استفاده از رویکردهای مختلف یادگیری ماشین و همچنین یادگیری عمیق پرداختند. نتایج این تحقیق نشان می‌دهد که رویکرد شبکه‌های عصبی مصنوعی عمیق به نسبت سایر روش‌ها در وضعیت بهتری در پیش‌بینی قرار دارد.

با توجه به این خلاصه از ادبیات موجود، به راحتی می‌توان دریافت که موضوع پیش‌بینی بر اساس یادگیری ماشین و یادگیری عمیق در بازار سهام از درجه اهمیت بالایی برخوردار باشد؛ لذا در این پژوهش برای اولین بار در داخل کشور به پیش‌بینی بازار سهام و تبیین و گسترش مدل‌سازی بر اساس یادگیری عمیق پرداخته خواهد شد.

روش‌شناسی پژوهش

در این بخش، روش‌های یادگیری ماشین و رویکردهای یادگیری عمیق مورد استفاده در پژوهش

مقایسه مدل‌های مختلف یادگیری ماشین در پیش‌بینی... /سهرابی، میربرگ کار، چیرانی و خریدار

شامل دسته‌بندی‌کننده‌های مختلف، روش کاهش ابعاد مورد استفاده و روش‌هایی که به منظور تصمیم‌گیری نهایی در مورد خروجی دسته‌بندی‌کننده‌ها استفاده شده‌اند، به‌طور خلاصه معرفی شده است.

مدل جنگل تصادفی^{۲۸} (RF)

روش جنگل تصادفی جزئی از روش‌های طبقه‌بندی تحت نظارت گروهی در حوزه داده‌کاوی است که به بخش یادگیری ماشین افزوده شده است. جنگل تصادفی (RF)، پتانسیل بسیار زیادی برای تبدیل به یک روش محبوب برای طبقه‌بندی‌های آینده دارد؛ زیرا کارایی آن قابل مقایسه با دیگر روش‌های گروهی است. به عنوان یک الگوریتم گروهی، رویکرد جنگل تصادفی چندین درخت تصمیم مختلف به عنوان طبقه‌بندی‌های پایه تولید می‌کند و رأی اکثریت را برای ترکیب با نتایج درختان اصلی اعمال می‌نماید. جنگل تصادفی هم توانایی رگرسیون و هم توانایی کلاس‌بندی را با هم دارد. روش کار در یک جنگل تصادفی چنین است که داده‌های نمونه‌ای، با عمل جای‌گذاری به تعداد n مجموعه نمونه تقسیم می‌شود و از هر دسته، نمونه‌ای برای آموزش یک درخت استفاده می‌شود. تمام درخت‌ها عمق مشخصی دارند و در هر گره تقسیم یک ویژگی یا فیچر به تصادف از بین مجموعه ویژگی‌ها انتخاب می‌شود و تقسیم یا شاخه‌بندی بر اساس آن صورت می‌گیرد. به علت استفاده از چندین دسته نمونه این روش مشکل داده‌های پرت و داده‌های گم شده را ندارد.

رویکرد ماشین بردار پشتیبان (SVMs)^{۲۹}

ماشین بردار پشتیبان یکی از روش‌های یادگیری با نظارت است که هم برای دسته‌بندی و هم رگرسیون قابل استفاده است. ماشین بردار پشتیبان در اصل یک دسته‌بندی‌کننده دو کلاسی است که کلاس‌ها را توسط یک مرز خطی از هم جدا می‌نمایند. در این روش، نزدیکترین نمونه‌ها به مرز تصمیم‌گیری را بردارهای پشتیبان می‌نامند. این بردارها معادله مرز تصمیم‌گیری را مشخص می‌کنند. این روش به دلیل استفاده از اصل کمینه‌سازی ریسک ساختاری که از طریق بیشینه کردن فاصله بین دو ابر صفحه گذرا از بردارهای پشتیبان هر دو کلاس اعمال می‌شود، بر خلاف حالت کمینه‌سازی ریسک تجربی که سعی در کمینه کردن خطای آموزش را دارد، عملکرد بهتری بر روی داده‌هایی که مدل با آنها ساخته نشده است، از خود نشان می‌دهد. به منظور سادگی در فهم، تئوری ماشین بردار پشتیبان از ساده‌ترین حالت ممکن یعنی دسته‌بندی دو کلاسی در حالت جدایی‌پذیر به صورت خطی توضیح داده شده است. براین اساس، اگر فرض کنیم که X_i متغیر ورودی در یک رگرسیون وجود دارد که دارای تعداد i متغیر متناظر پاسخ‌دهنده است، ماشین بردار پشتیبان را می‌توان به صورت رگرسیون خطی زیر نشان داد:

$$f(xw) = w^T x + b \quad (1)$$

که در آن w بردار نرمال ابرصفحه جداکننده است که هدف آن، حداقل‌سازی تابع w است. برای یافتن مرز تصمیم‌گیری بهینه از روش حاشیه بیشینه استفاده می‌شود. بنابراین مرز تصمیم‌گیری علاوه بر اینکه بایستی تمام نمونه‌های هر دو کلاس را به درستی به دو دسته تقسیم نماید، باید مرز تصمیم‌گیری (بر صفحه) را پیدا کند که بیشترین فاصله از همه بردارهای پشتیبان را داشته باشد. لذا می‌توان یک مسأله بهینه‌سازی به صورت زیر را تعریف نمود:

$$\begin{aligned} \min \quad & \frac{1}{2} \|w\|^2 \\ \text{s. t} \quad & y_i (w^T x + b) \geq 1 \end{aligned} \quad (2)$$

برای حل مسأله بهینه‌سازی، از روش ضرایب لاگرانژ استفاده می‌شود. بر اساس شرایط کان - تاکر می‌توان مقادیر بهینه w بر اساس ضرایب لاگرانژ را به صورت زیر برآورد نمود:

$$w = \sum \alpha_i x_i y_i \quad (3)$$

که در آن α_i ضرایب لاگرانژین تابع بهینه‌سازی است. مقدار b نیز به صورت زیر قابل برآورد است:

$$w = \frac{1}{N_{sv}} \sum y_i - w x_i \quad (4)$$

که در آن N_{sv} تعداد بردارهای پشتیبان است. بر این اساس، تابع تصمیم‌گیری با رویکرد ماشین بردار پشتیبان به صورت زیر است:

$$f(x) = \text{sgn}(w^T x + b) \quad (5)$$

روش شبکه عصبی مصنوعی (ANN) ۳۰

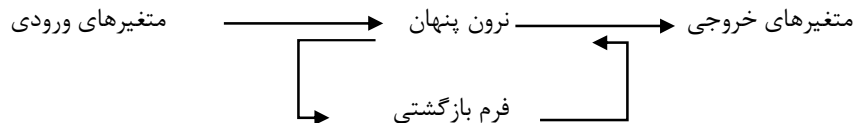
شبکه‌های عصبی از دهه ۵۰ شناخته شده بودند؛ اما تنها در اواسط دهه ۸۰ بود که الگوریتم‌ها و روش‌های مربوط به شبکه‌های عصبی مصنوعی به درجه‌ای از پیشرفت رسیدند که در حل مسائل واقعی از آنها استفاده می‌شد. شکل ساده شبکه‌های عصبی مصنوعی (ANN) یک لایه ورودی و یک لایه خروجی دارد. لایه ورودی شامل متغیرهای ورودی با n تعداد متغیر است. لایه خروجی نیز شامل متغیرهای خروجی زیادی است که هر ورودی x به یک خروجی y متصل است. در حقیقت شبکه عصبی مصنوعی، مجموعه‌ای از نرون‌های بهم متصل در لایه‌های مختلف هستند که اطلاعاتی را برای یکدیگر ارسال می‌نمایند. نرون‌های مصنوعی واحدهای ساده پردازش اطلاعات هستند؛ بنابراین تعداد زیادی از این نرون‌ها یک شبکه عصبی را می‌سازند.

مقایسه مدل‌های مختلف یادگیری ماشین در پیش‌بینی... /سهرابی، میربرگ کار، چیرانی و خردیار

روش شبکه عصبی بازگشتی مبتنی بر یادگیری عمیق

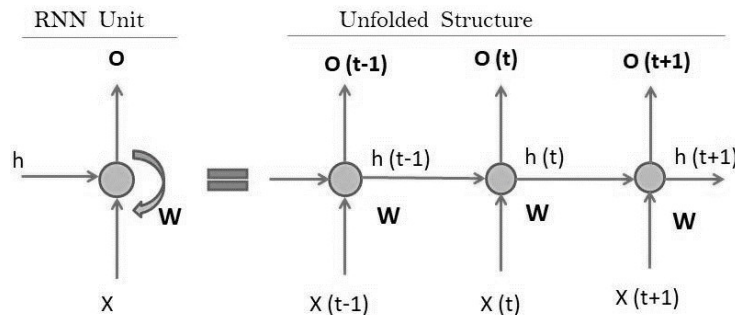
یک شبکه عصبی بازگشتی (ANN)^{۳۱} یک کلاس از شبکه‌های عصبی مصنوعی است که در آن ارتباط بین گره یک گراف هدایت شده در طول سری زمانی را ایجاد می‌نماید. همانطور که مشخص است، این رویکرد به صورت بازگشتی عمل می‌کند. یعنی یک عملیات برای تمامی سری تصادفی انجام می‌گیرد و خروجی آن وابسته به ورودی فعلی و عملیات‌های قبلی است. این مهم از طریق تکرار یک خروجی از شبکه در زمان t و ورودی شبکه در زمان $t+1$ انجام می‌شود (یعنی خروجی از مرحله قبل با ورودی تازه در مرحله جدید ترکیب می‌شوند). بدین ترتیب، این رویکرد می‌تواند وقایع مختلف تکرارشونده و تأثیرگذار را شناسایی و در پیش‌بینی‌های بعدی مورد استفاده قرار دهد.

همانطور که اشاره شد بخش اصلی این رویکرد بازگشت است. بر این اساس، بر نرون‌های مخفی بازگشتی را خواهد داشت که به صورت زیر قابل بیان است:



$$\begin{aligned} h(t) &= f_H(W_{IH}x(t) + W_{HH}h(t-1)) \\ y(t) &= f_o(W_{Ho}h(t)) \end{aligned} \quad (۶)$$

که در آن h نرون‌های مخفی، W وزن لایه‌های شبکه عصبی برای لایه ورودی به پنهان (IH)، پنهان به پنهان (HH) و پنهان به خروجی (HO) است. همچنین f_H تابع فعال‌سازی بخش پنهان و f_o تابع فعال‌سازی بخش خروجی است. شمای کلی این ارتباط را می‌توان به صورت زیر مشاهده نمود:



داشتن یک حافظه بلندمدت‌تر دارای اثر تثبیت‌کننده است؛ چرا که حتی اگر شبکه نتواند از تاریخچه اخیر خود درک صحیحی پیدا نماید، باز با این وجود قادر است با نگاه به گذشته، پیش‌بینی خود را کامل نماید. یک راه‌حل که برای مدل‌های شرطی مطرح شده است، اضافه نمودن نویز به پیش‌بینی‌های

صورت گرفته توسط شبکه، قبل از تغذیه آنها به گام زمانی بعدی است. این کار باعث تقویت شبکه در قبال ورودی‌های غیرمنتظره می‌شود. با این وجود، یک حافظه بهتر، راه‌حل به مراتب بهتر و تأثیرگذارتری است. حافظه طولانی کوتاه‌مدت (LSTM)، یک معماری شبکه عصبی بازگشتی است که برای ذخیره‌سازی و دسترسی بهتر به اطلاعات نسبت به نسخه سنتی آن طراحی شده است.

بررسی ارزیابی مدل‌های مختلف

برای ارزیابی مدل‌های پیش‌بینی‌کننده، از مدل‌هایی که خطاهای مدل‌های پیش‌بینی‌کننده را اندازه‌گیری می‌نمایند استفاده خواهد شد. این مدل‌ها بر چند نوع عمده تقسیم‌بندی می‌شوند:

$$\text{RMSE} = \frac{1}{n} \sum_i \frac{(\hat{y}_i - y_i)^2}{\hat{y}_i \bar{y}_i} \quad \text{مربع میانگین خطای استاندارد نرمال شده}$$

$$\text{SAMPE} = \frac{1}{n} \sum_i (\hat{y}_i - \bar{y}_i)^2 \quad \text{میانگین انحراف معیار خطای پیش‌بینی}$$

$$\text{MAPE} = \sqrt{n^{-1} \sum_{t=1}^n (\sigma_{a,t}^y - \sigma_{a,t}^{\hat{y}})^2} \quad \text{درصد خطای متوسط مطلق}$$

$$\text{MAE} = n^{-1} \sum_{t=1}^n |\sigma_{a,t}^y - \sigma_{a,t}^{\hat{y}}| \quad \text{متوسط مطلق خطا}$$

که در آنها n تعداد پیش‌بینی‌های انجام شده، $\sigma_{a,t}^y, \sigma_{a,t}^{\hat{y}}$ به ترتیب مقدار اصلی پارامتر و پیش‌بینی پارامتر است که با استفاده از روش خارج از نمونه صورت می‌گیرد.

یافته‌های پژوهش

جامعه آماری و متغیرهای پژوهش

متغیرهای مورد مطالعه در این پژوهش شامل بازدهی شاخص کل بورس اوراق بهادار تهران و متغیرهای به ظاهر مؤثر بر این شاخص به عنوان متغیر برون‌زا است.

متغیرهای برون‌زا بر اساس ادبیات تجربی پیرامون این موضوع و همچنین مطابق با مطالعات چاتزیس و همکاران (۲۰۱۸) و مقصود و همکاران^{۳۲} (۲۰۲۱) انتخاب شده است؛ لذا این انتخاب بر اساس طیف متغیرهای اقتصادی، شاخص بازار کالا، قیمت نفت، قیمت طلا، نرخ ارز و روند سایر بازارهای مالی و بین‌المللی است. لیست متغیرهای مورد استفاده در این پژوهش در جدول (۱) خلاصه شده است. تمامی این داده‌ها به صورت روزانه با حذف روزهای تعطیل و روزهای کاری غیرهمزمان بوده است. دوره مورد

مقایسه مدل های مختلف یادگیری ماشین در پیش بینی... /سهرابی، میربرگ کار، چیرانی و خردیار

بررسی از مرداد ۱۳۹۲ تا بهمن ۱۳۹۹ شامل ۷۵۰ روز کاری مشترک با بازارهای جهانی بوده است.

جدول ۱. لیست متغیرهای پژوهش

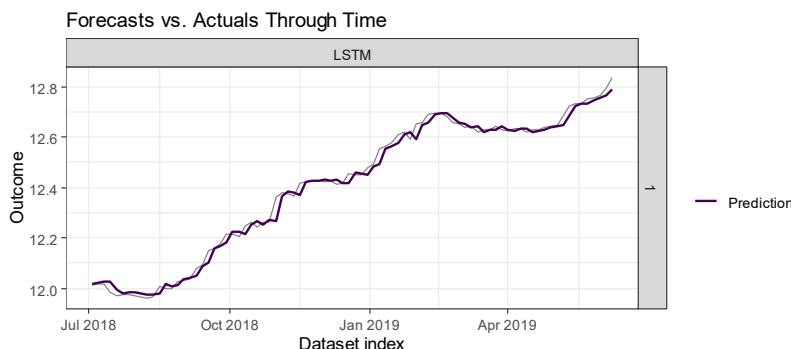
منبع	متغیر
مستخرج از سایت بورس تهران	شاخص کل بورس تهران
سایت یاهو و فایننس	لگاریتم شاخص S&P500
سایت یاهو و فایننس	لگاریتم شاخص DAXI
سایت یاهو و فایننس	لگاریتم شاخص FCHI
سایت یاهو و فایننس	لگاریتم شاخص NIKKI
سایت یاهو و فایننس	لگاریتم شاخص بورس امارات
سایت یاهو و فایننس	لگاریتم شاخص بورس چین
سایت یاهو و فایننس	لگاریتم شاخص ترکیه
سایت اوپک	لگاریتم قیمت نفت خام اوپک
بورس کالای شیکاگو	لگاریتم قیمت نفت خام برنت
آرشیو نرخ آزاد ارز داخلی	لگاریتم نرخ آزاد ارز (دلار امریکا)
آرشیو نرخ آزاد ارز داخلی	لگاریتم نرخ آزاد ارز (یورو)
آرشیو نرخ آزاد ارز داخلی	لگاریتم نرخ آزاد ارز (پوند)
آرشیو نرخ آزاد ارز داخلی	لگاریتم نرخ آزاد ارز (یوان)
آرشیو نرخ آزاد ارز داخلی	لگاریتم نرخ آزاد ارز (درهم امارات)
بورس کالای لندن	لگاریتم قیمت اونس طلا
آرشیو قیمت داخلی سکه طلا	لگاریتم قیمت سکه بهار آزادی

در فرآیند پیش بینی سری های زمانی به روش یادگیری ماشین، به جای آنکه از مدل برای طولانی مدت برای پیش بینی دوره پیش رو استفاده شود، با پیش بینی دوره های گذشته محاسبه شده از مدل برآورد شده با داده های واقعی می توان عملکرد مدل را سنجید. بدین منظور، در این پژوهش از دو دوره استفاده شده است. دوره اول دوره یادگیری (درون نمونه) و دوره آزمون (دوره خارج از نمونه) است. در این رهیافت لازم است یک دوره برازش^{۳۳} در نظر گرفته شود که تخمین پارامترهای مدل را تعریف می کند و متغیر مستقل و تأثیرگذار بر شاخص جهش های بورس تهران تبیین می گردد. این نمونه برازش^{۳۴}، در سراسر کل دوره داده ها به صورت یک دوره یادگیری طراحی می شود. با ثابت در نظر گرفتن دوره یادگیری، مدل های مختلف با استفاده از روش های الگوریتمی خود به الگوریتم مناسب جهت پیش بینی خواهند

رسید و در نهایت اقدام به پیش‌بینی می‌نمایند.

در این پژوهش برای بررسی دقت پیش‌بینی‌های انجام شده، دوره آزمون (خارج از نمونه) با در نظر گرفتن ۲۰ درصد از کل داده‌ها به صورت یک روزه، سه روزه و شش روزه پیش‌بینی شده است. این موضوع علاوه بر نشان دادن توانایی مدل‌های مختلف جهت پیش‌بینی دوره‌های مختلف، توانایی هر مدل در این دوره‌ها را نیز خواهد سنجید.

نتایج پیش‌بینی شاخص بورس اوراق بهادار که هدف اصلی این پژوهش است، برای روش شبکه عصبی بازگشتی LSTM و پیش‌بینی یک روزه در نمودار (۱) برای دوره آزمون نمایش داده شده است. همانطور که مشخص است، روش شبکه عصبی بازگشتی به روش یادگیری عمیق، روند مناسبی داشته و پیش‌بینی به مقادیر واقعی بسیار نزدیک است.



نمودار ۱. مقادیر واقعی و پیش‌بینی شده شاخص بورس بر اساس روش LSTM برای دوره آزمون

نتایج مقایسه پیش‌بینی سری زمانی شاخص بورس اوراق بهادار برای روش‌های مختلف برای پیش‌بینی یک روزه، سه روزه و شش روزه برای دوره آزمون در جدول (۲) خلاصه شده است. همانطور که مشخص است، روش شبکه عصبی بازگشتی به روش یادگیری عمیق عملکرد بهتری را داشته است و پیش‌بینی به مقادیر واقعی بسیار نزدیک است؛ به نحوی که دارای مقادیر خطای کمتری است.

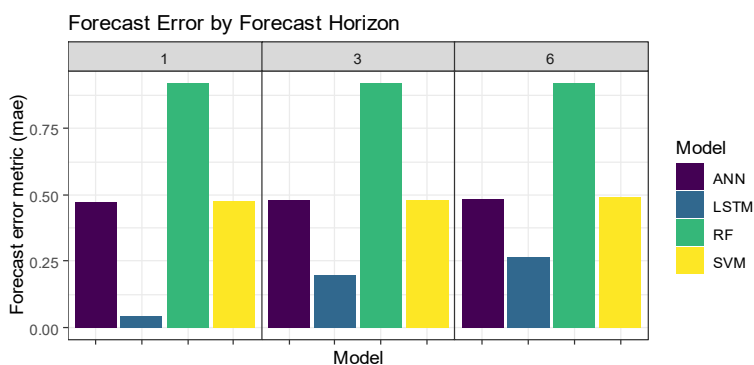
بدین ترتیب می‌توان بر اساس نتایج بیان نمود که هدف اصلی این پژوهش و سؤال اصلی در این بررسی به خوبی اثبات شده است. همچنین اذعان داشت که روش شبکه عصبی بازگشتی مبتنی بر روش یادگیری عمیق نتایج بهتری را به همراه داشته است.

مقایسه مدل های مختلف یادگیری ماشین در پیش بینی... /سهرابی، میربرگ کار، چیرانی و خردیار

جدول ۲. مقایسه مدل های پیش بینی کننده شاخص بورس اوراق بهادار

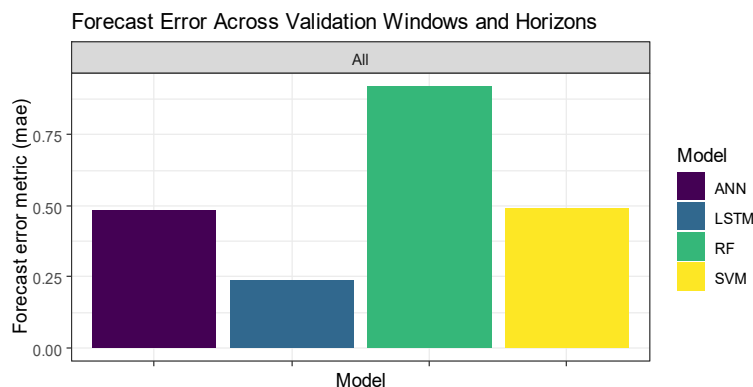
MAE	MAPE	MADPE	SAMPE	RSME	دوره پیش بینی	مدل
۰,۴۷۳	۳,۹۳۸	۳,۶۴۳	۴,۰۲۱	۰,۵۴۶	یک روزه	ANN
۰,۴۸۱	۴,۰۱۱	۳,۷۱۶	۴,۰۹۳	۰,۵۵۴	سه روزه	ANN
۰,۴۸۶	۴,۰۴۶	۳,۷۴۷	۴,۱۲۹	۰,۵۵۸	شش روزه	ANN
۰,۰۴۰	۰,۳۴۴	۰,۳۲۲	۰,۳۴۵	۰,۰۴۷	یک روزه	LSTM
۰,۱۹۵	۱,۷۰۰	۱,۵۷۰	۱,۷۲۴	۰,۲۲۶	سه روزه	LSTM
۰,۲۶۷	۲,۳۰۱	۲,۲۲۴	۲,۳۳۹	۰,۳۰۵	شش روزه	LSTM
۰,۹۲۲	۷,۸۴۴	۷,۵۷۶	۷,۸۵۹	۰,۹۷۲	یک روزه	RF
۰,۹۲۱	۷,۸۴۲	۷,۵۷۴	۷,۸۵۷	۰,۹۷۲	سه روزه	RF
۰,۹۲۱	۷,۸۴۰	۷,۵۷۲	۷,۸۵۶	۰,۹۷۲	شش روزه	RF
۰,۴۷۵	۳,۹۵۴	۳,۶۶۶	۴,۰۳۷	۰,۵۴۸	یک روزه	SVM
۰,۴۸۰	۳,۹۹۶	۳,۷۱۱	۴,۰۷۸	۰,۵۵۲	سه روزه	SVM
۰,۴۹۰	۴,۰۸۶	۳,۷۷۹	۴,۱۶۷	۰,۵۶۳	شش روزه	SVM

مأخذ: نتایج پژوهش



نمودار ۲. مقایسه تخطی مدل های مختلف پیش بینی شاخص بورس

مقایسه خطاهای کلی مدل های مختلف برای هر سه دوره پیش بینی در نمودار (۳) نیز نشان دهنده این واقعیت است که مدل شبکه عصبی بازگشتی به روش یادگیری عمیق عملکرد بهتری به نسبت سایر مدل ها داشته و خطای کمتری نشان می دهد. علاوه بر این، جدول (۳) به خوبی گویای این واقعیت است که در تمامی معیارهای خطای مدل پیش بینی، رویکرد شبکه عصبی بازگشتی یادگیری عمیق در شرایط بهتری است.



نمودار ۳. مقایسه کلی تخطی پیش‌بینی شاخص کل بورس

جدول ۳. مقایسه کلی مدل‌های پیش‌بینی‌کننده شاخص بورس اوراق بهادار

MAE	MAPE	MADPE	SAMPE	RSME	مدل پیش‌بینی
۴۸۴.۰	۰.۳۴.۴	۳,۷۴۸	۴,۱۱۴	۰,۵۵۷	ANN
۲۳۹.۰	۰.۵۶.۲	۱,۵۸۷	۲,۰۸۷	۰,۲۷۸	LSTM
۹۲۲.۰	۸۳۶.۷	۷,۵۶۸	۷,۸۵۹	۰,۹۷۲	RF
۴۸۸.۰	۰.۶۸.۴	۳,۷۷۸	۴,۱۴۸	۰,۵۶۱	SVM

مأخذ: نتایج پژوهش

نتیجه‌گیری

هدف از این پژوهش، بررسی توانایی مدل‌های مختلف پیش‌بینی شاخص کل بورس اوراق بهادار در ایران در بازه زمانی ۱۳۹۲ لغایت ۱۳۹۹ است. نتایج بررسی خطای دوره آزمون به روش‌های مختلف شامل جنگل تصادفی، ماشین بردار پشتیبان، شبکه عصبی مصنوعی و شبکه عصبی بازگشتی LSTM برای دوره‌های یک روزه، سه روزه و شش روزه برای دوره خارج از نمونه، نشان می‌دهد که رویکرد شبکه عصبی بازگشتی LSTM عملکرد بهتری برای پیش‌بینی شاخص داشته است. این موضوع گویای این واقعیت است که تبیین یادگیری عمیق و طولانی کردن یادگیری شبکه عصبی در دوره یادگیری موجب بهتر شدن عملکرد پیش‌بینی شده است.

این نتایج با پژوهش چاتزیس و همکاران (۲۰۱۸) هم‌سو می‌باشد. بر این اساس، با توجه به توانایی مناسب مدل LSTM پیشنهاد می‌گردد از این مدل برای پیش‌بینی ریسک بازارهای مالی نیز استفاده

مقایسه مدل های مختلف یادگیری ماشین در پیش بینی... /سهرابی، میربرگ کار، چیرانی و خردیار

گردد به نحوی که برای تبیین سیستم هشدار بحران و تبیین نرخ پوشش بهینه نیز استفاده مؤثری برای مؤسسات مالی و فعالان بازار سرمایه داشته باشد. در نهایت با توجه به نتایج پیش بینی شاخص و دقت مناسب شبکه عصبی بازگشتی در این مسیر، پیشنهاد می گردد موضوع سیگنال دهی ایجاد شده و وجود استراتژی سودآور بین بازارهای موازی مورد توجه بازیگران بازارهای مالی و سرمایه گذاران قرار گیرد تا بتوانند با دانش از تغییرات بین بازارها و پیش بینی آنها، به کنترل ریسک و بازدهی مالی بپردازند.

منابع

- ۱) تهرانی رضا، عباسیون وحید. کاربرد شبکه های عصبی مصنوعی در زمان بندی معاملات سهام: با رویکرد تحلیل تکنیکی. پژوهش های اقتصادی. سال ۱۳۸۷. شماره ۸. ۱۷۷-۱۵۱.
- ۲) تهرانی رضا، نوربخش عسکر. مدیریت سرمایه گذاری. چاپ اول. تهران. نشر نگاه دانش. سال ۱۳۸۲.
- ۳) زراء نژاد منصور، رئوفی علی، کیانی پویان. ارزیابی و مقایسه عملکرد مدل خودرگرسیون میانگین متحرک انباشته و شبکه عصبی فازی برای پیش بینی روزانه قیمت طلا. اولین همایش بین المللی اقتصاد سنجی، روش ها و کاربردها. سنندج. سال ۱۳۹۱.
- ۴) شمس ناصر، پارسائیان سمیرا. مقایسه عملکرد مدل فاما و فرنچ و شبکه های عصبی مصنوعی در پیش بینی بازده سهام در بورس تهران. فصلنامه مهندسی مالی و مدیریت اوراق بهادار. سال ۱۳۹۱. شماره ۱۸. صص ۱۱۸-۱۰۳.
- ۵) فلاح پور سعید، گلارضا غلامحسین و فتوره چیان ناصر. پیش بینی روند حرکتی قیمت سهام با استفاده از ماشین بردار پشتیبان بر پایه الگوریتم ژنتیک در بورس اوراق بهادار تهران. فصلنامه تحقیقات مالی. سال ۱۳۹۲. شماره ۱۵. صص ۲۸۸-۲۶۹.
- 6) Adebisi, A. A., Adewumi, A.O., & Ayo, C. K. (2014). "Comparison of ARIMA and artificial neural networks models for stock price prediction". *Journal of Applied Mathematics*, 1-7.
- 7) Atsalakis, G., Protopapadakis, E. and Valavanis, K. (2016). "Stock trend forecasting in turbulent market periods using neuro-fuzzy systems". *Operational Research* 16(2): 245-269.
- 8) Bae, K., Karolyi, A., Stulz, R. (2003) "A new approach to measuring financial contagion." *Review of Financial Studies* 16: 717-763.
- 9) Bagheria, A., Mohammadi, H., Akbaric, M., P. (2014) "Financial forecasting using ANFIS networks with Quantum-behaved Particle Swarm Optimization", *Expert Systems with Applications* 41: 6235-6250.
- 10) Cavalcante, R. C., Brasileiro, R. C., Souza, V. L.F., Nobrega, J. P., & Oliveira, A. L.I. (2016). "Computational Intelligence and Financial Markets: A Survey and Future Directions". *Expert Systems with Applications*, 55, 194-211
- 11) Cervelló-Royo, R., Guijarroa, F., Michniukab, K. (2015) "Stock market trading rule based on pattern recognition and technical analysis: Forecasting the DJIA index with intraday data", *Expert Systems with Applications* 42: 5963-5975.
- 12) Chatzis, S. P., Siakoulis, V., Petropoulos, A., Stavroulakis, E., & Vlachogiannakis, N. (2018). "Forecasting stock market crisis events using deep and statistical machine learning techniques". *Expert Systems with Applications*, 112, 353-371.

- 13) Chiang, W.-C., Enke, D., Wu, T., Wang, R. (2016) "An adaptive stock index trading decision support system", *Expert Systems with Applications* 59: 195- 207.
- 14) Christiansen, C. and Rinaldo, A. (2009) "Extreme coexceedances in new EU member states' stock markets." *Journal of banking & finance* 33(6): 1048-1057.
- 15) Cuneyt, S., Oztekin, A., Ozkan, B., Serkan, G. and Erkam, G. (2014) "Developing an early warning system to predict currency crises". *European Journal of Operational Research* 237:1095–1104.
- 16) Dabrowski, J., Beyers, C. and De Villiers, DP. (2016) "Systemic Banking Crisis Early Warning Systems Using Dynamic Bayesian Networks" *Expert Systems with Applications*.
- 17) Döpke, J., Fritsche, U. and Pierdzioch, C. (2017) "Predicting recessions with boosted regression trees." *International Journal of Forecasting* 33(4): 745-759.
- 18) Dungey, M. and Martin, V. (2006) "Unravelling financial market linkages during crises". *Journal of Applied Econometrics* 22(1): 89-119.
- 19) Fama, E. F. (1995). Random walks in stock market prices. *Financial analysts journal*, 51(1), 75-80.
- 20) Fama, E. F., & Malkiel, B. G. (1970). Efficient capital markets: A review of theory and empirical work. *The journal of finance*, 25(2), 383-417.
- 21) Forbes, K. and Rigobon, R. (2002) "No contagion, only interdependence: measuring stock market comovements." *The journal of Finance* 57(5): 2223- 2261.
- 22) Kim, Y., H., Chang, W., H. (2018) "Forecasting the volatility of stock price index: A hybrid model integrating LSTM with multiple GARCH-type models", *Expert Systems with Applications* 103: 25-37.
- 23) Markwat, T., Kole, E. and Van Dijk, D. (2009) "Contagion as a domino effect in global stock markets." *Journal of Banking & Finance* 33(11): 1996-2012.
- 24) Maqsood, H., Mehmood, I., Maqsood, M., Yasir, M., Afzal, S., Aadil, F., Selim, M. M., Muhammad, K. (2021). "A local and global event sentiment based efficient stock exchange forecasting using deep learning". *International Journal of Information Management*, 50, 432-451.
- 25) Oztekin, A., Kizilaslan, R., Freund, S. and Iseri, A. (2016) "A data analytic approach to forecasting daily stock returns in an emerging market" *European Journal of Operational Research*, 253(3): 697-710.
- 26) Patel, J., Shah, S., Thakkar, P., & Kotecha, K. (2015). "Predicting stock market index using fusion of machine learning techniques". *Expert Systems with Applications*, 42, 4, 2162-2172.
- 27) Zhong, X., Enke, D. (2017) "Forecasting daily stock market return using dimensionality reduction." *Expert Systems with Applications* 67: 126-139.
- 28) Xiao, Y., Xiao, J., Lu, F., & Wang, Sh. (2013). "Ensemble ANNs-PSO-GA Approach for Dayahead Stock E-exchange Prices Forecasting". *International Journal of Computational Intelligence Systems*, 6(1), 96-114

- 1 Christiansen and Rinaldo (2009)
- 2 Markwat et al. (2009)
- 3 Xiao, Y., Xiao, J., Lu, F., & Wang, Sh. (2013)
- 4 Cervelló- Royo, Guijarroa, and Michniukab (2015)
- 5 Chiang, Enke, Wu, and Wang (2016)
- 6 Zhong and Enke (2017)
- 7 Kim and Chang (2018)
- 8 Patel, J., Shah, S., Thakkar, P., & Kotecha, K. (2015)
- 9 Cavalcante, R.C., Brasileiro, R. C., Souza, V. L.F., Nobrega, J.P., & Oliveira, A.L.I.(2016)
- 10 Adebisi, A. A., Adewumi, A.O., & Ayo, C. K. (2014)
- 11 Bagheria, A., Mohammadi, H., Akbaric, M., P. (2014)
- 12 Cervelló- Royo, Guijarroa, and Michniukab (2015)
- 13 Chiang, Enke, Wu, and Wang (2016)
- 14 Zhong and Enke (2017)
- 15 Kim and Chang (2018)
- 16 Dungey and Martin (2006)
- 17 Forbes and Rigobon (2002)
- 18 Bae, Karolyi, and Stulz (2003)
- 19 Christiansen and Rinaldo (2009)
- 20 Markwat et al. (2009)
- 21 Xiao, Y., Xiao, J., Lu, F., & Wang, Sh. (2013)
- 22 Cuneyt, S., Oztekin, A., Ozkan, B., Serkan, G. and Erkam, G. (2014)
- 23 Atsalakis, Protopapadakis, and Valavanis (2016)
- 24 Oztekin, Kizilaslan, Freund, and Iseri (2016)
- 25 Döpke, Fritsche, and Pierdzioch (2017)
- 26 Dabrowski, Bey- ers, and De Villiers (2016)
- 27 Chatzis et al. (2018)
- 28 Random Forests (RF)
- 29 Support Vector Machines (SVMs)
- 30 Artificial Neural Networks (ANN)
- 31 Recurrent Neural Networks(RNN)
- 32 Maqsood (2021)
- 33 Estimation period
- 34 Estimation sample