

مجله‌ی پیشرفت‌های حسابداری دانشگاه شیراز

دوره‌ی هفتم، شماره‌ی دوم، پاییز و زمستان ۱۳۹۴، پیاپی ۶۹/۳، صفحه‌های ۱۱۷-۱۴۸

(مجله‌ی علوم اجتماعی و انسانی پیشین)

## رویکرد مهندسی جدید برای پیش‌بینی نوسان شاخص‌های بورس اوراق بهادار تهران

دکتر مهدی مرادی\*      دکترهادی صدوقی یزدی\*\*      جواد عبدالهیان\*\*\*

دانشگاه فردوسی مشهد

دانشگاه آزاد اسلامی واحد مشهد

### چکیده

سرمایه‌گذاران و مسؤولان بورس اوراق بهادار به منظور دستیابی به تصویر مناسبی از روند بازار بورس، عملکرد شرکت‌ها و توانایی ارزیابی گذشته جهت پیش‌بینی آینده، از شاخص‌های بورس اوراق بهادار استفاده می‌کنند. مطالعات اخیر نشان می‌دهد که همبستگی غیرخطی در شاخص‌های بازار سهام وجود دارد. به همین منظور، این مطالعه الگوریتم جدید حداقل میانگین مربعات خطا مبتنی بر کرنل که الگوریتم آنلاین و غیرخطی است را برای پیش‌بینی نوسان شاخص‌های بورس اوراق بهادار تهران پیشنهاد و به کار می‌گیرد و عملکرد آن را با الگوریتم شبکه‌های عصبی مبتنی بر حداقل میانگین مربعات خطا در افق‌های زمانی کوتاه مدت، میان مدت و بلندمدت مقایسه می‌کند. نتایج این بررسی نشان می‌دهد که الگوریتم حداقل میانگین مربعات خطا مبتنی بر کرنل دارای عملکرد مناسبی جهت پیش‌بینی نوسان شاخص‌های سهام می‌باشد. با این حال، نتایج پیش‌بینی الگوریتم بر اساس معیارهای میانگین مربعات خطا، مجذور میانگین مربعات خطا، میانگین مربعات خطای نرمالیزه شده، قدر مطلق میانگین خطا، صحت جهت پیش‌بینی، صحت روند پیش‌بینی رو به بالا و صحت روند پیش‌بینی روبه پایین در افق‌های زمانی کوتاه مدت، میان مدت و بلند مدت نشان می‌دهد که پیش‌بینی انجام شده توسط الگوریتم حداقل میانگین مربعات خطا مبتنی بر کرنل دارای عملکرد بهتری نسبت به پیش‌بینی با الگوریتم شبکه‌های عصبی مبتنی بر حداقل میانگین مربعات خطا می‌باشد.

**واژه‌های کلیدی:** شاخص سهام، پیش‌بینی، الگوریتم حداقل میانگین مربعات خطا مبتنی بر کرنل، الگوریتم شبکه‌های عصبی مبتنی بر حداقل میانگین مربعات خطا.

\* دانشیار گروه حسابداری دانشگاه فردوسی مشهد (نویسنده مسئول) Mhd\_Moradi@um.ac.ir

\*\* دانشیار گروه کامپیوتر دانشگاه فردوسی مشهد

\*\*\* کارشناس ارشد حسابداری دانشگاه آزاد اسلامی واحد مشهد

تاریخ پذیرش: ۱۳۹۳/۷/۳۰

تاریخ دریافت: ۱۳۹۲/۵/۳۰

## ۱. مقدمه

بدون شک اتخاذ هر تصمیم نیازمند اطلاعات و داده‌های مرتبط با آن است. بسیاری از این اطلاعات از طریق فرایند پیش بینی تهیه می‌شوند. پیش‌بینی یک عنصر مهم تصمیم‌گیری است که بر اساس معیارهای از قبل تعیین شده و یا داده‌های گذشته به دنبال برآورد وقایع آینده می‌باشد (اسرینیواسان<sup>۱</sup>، ۲۰۰۳؛ هیور<sup>۲</sup>، ۲۰۰۴). با وجود آن که پیش‌بینی‌ها معمولاً دقیق نیست، ولی میزان خطای پیش‌بینی به سیستم مورد استفاده برای پیش‌بینی بستگی دارد (فلاح شمس و اصغری، ۱۳۸۸). افزایش دقت در پیش‌بینی، عدم اطمینان را کاهش می‌دهد و به تصمیم‌گیرندگان کمک می‌کند تا فرایند تصمیم‌گیری خود را بهبود و خطر تصمیم‌های نادرست را کاهش دهند (مونتگومری<sup>۳</sup> و همکاران، ۱۹۹۰، حقیقت و معتمد، ۱۳۹۰).

اهمیت روز افزون پیش‌بینی برای عوامل اقتصادی از یک طرف و کاستی مدل‌های ساختاری در پیش‌بینی از طرف دیگر منجر به توسعه مدل‌های سری زمانی برای مدل‌سازی و پیش‌بینی شده است. با گسترش بازارهای مالی، سهامداران، محققان و سیاست‌گذاران برای تصمیم‌گیری بهینه و کاهش ریسک نیازمند آشنایی با مدل‌های پیش‌بینی و استفاده از آنها هستند. (مشیری و مروت، ۱۳۸۴). به طور معمول سرمایه‌گذاران به عنوان گروه اصلی استفاده‌کننده از اطلاعات مالی بیشتر علاقه مند به تداوم سودآوری شرکت‌ها و پیش‌بینی عملکرد آن‌ها از نظر شاخص‌های مختلف می‌باشند. به همین دلیل، مطالعات نسبتاً زیادی در زمینه پیش‌بینی نوسان شاخص‌های بورس اوراق بهادار، پیش‌بینی عملکرد شرکت‌ها (سودآوری)، پیش‌بینی ورشکستگی شرکت‌ها در سه دهه اخیر به انجام رسیده است. از آن جایی که شاخص‌های بورس اوراق بهادار نشان‌دهنده عملکرد بورس و شرکت‌ها می‌باشد، لذا سهامداران، سرمایه‌گذاران و مسؤولان بورس اوراق بهادار به منظور دستیابی به تصویر مناسبی از روند بازار بورس، عملکرد شرکت‌ها و توانایی ارزیابی گذشته جهت پیش‌بینی آینده، از شاخص‌ها استفاده می‌کنند. شاخص‌های بورس مهم‌ترین متغیر در زمینه تصمیمات سرمایه‌گذاری، قیمت‌گذاری اوراق بهادار، مدیریت ریسک، تدوین مقررات و سیاست‌گذاری پولی می‌باشد (تیمور محمدی و نصیری، ۱۳۸۹). بنابراین، پیش‌بینی نوسان روزانه شاخص‌های بورس اوراق بهادار در پیش‌بینی سری‌های زمانی بسیار مهم است، زیرا نشان‌دهنده وضعیت کلی مطلوب یا نامطلوب

سرمایه‌گذاری برای سهامداران و سرمایه‌گذاران است.

مره<sup>۴</sup> و همکاران (۲۰۱۱) بیان می‌کنند در شاخص‌های سهام، داده‌ها نمونه‌ای از سری زمانی است. پیش‌بینی سری‌های زمانی، تجزیه و تحلیل داده‌های تاریخی و تلاش برای تعیین ارزش تقریبی آینده یک سری زمانی به عنوان ترکیب خطی از این داده‌های تاریخی است. پیش‌بینی نوسان‌های روزانه شاخص‌های بورس اوراق بهادار در پیش‌بینی سری‌های زمانی مالی بسیار مهم است زیرا نوسان‌ها در بازار سهام ضرورتاً پویا، غیرخطی، پیچیده، ناپارامتریک، و ماهیتاً بی‌نظم هستند (کارا<sup>۵</sup> و همکاران، ۲۰۱۱؛ منیش و دنموژی<sup>۶</sup>، ۲۰۰۵؛ ابومصطفی و آتیا<sup>۷</sup>، ۱۹۹۶). با توجه به وجود ریسک در بازار بورس اوراق بهادار، در سال‌های اخیر، محققان در مطالعات خود، مدل‌ها و الگوریتم‌های مختلفی برای برآورد و پیش‌بینی نوسان شاخص‌های بورس اوراق بهادار به منظور کمک به سرمایه‌گذاران معرفی نموده‌اند (پویانما<sup>۸</sup>، ۲۰۰۶). لیکن، هیچ‌یک از این مطالعات الگوریتم آنلاین غیرخطی را برای پیش‌بینی نوسان شاخص‌ها ارائه نکرده‌اند. با به‌کارگیری رویکردهای مهندسی، سرمایه‌گذاران می‌توانند در مقابل ریسک‌های بالقوه بازار مصون بمانند و و دلالت و آربیتراژها فرصت‌هایی جهت کسب سود به وسیله پیش‌بینی شاخص سهام بدست آورند (منیش و دنموژی، ۲۰۰۵). به همین منظور در این مطالعه از رویکرد مهندسی جدید که دقت بالایی در پیش‌بینی نوسان شاخص‌های بازار سهام دارند استفاده می‌شود. این می‌تواند به کاهش ریسک سیستماتیک بازار سهام و تخصیص بهینه منابع کمک نماید.

با توجه به آنکه معمولاً عوامل و نیروهای پیچیده بسیاری باعث تغییر پدیده‌های اقتصادی می‌شوند که احتمالاً از روابط غیرخطی تبعیت می‌کنند، مدل‌های سری زمانی غیرخطی آنلاین و پویا از جایگاه مهمی در ادبیات اقتصادی مدل‌های پیش‌بینی (به ویژه در دهه اخیر) برخوردار شده‌اند و در حال حاضر نقش مهمی در مدل‌سازی روابط اقتصادی که یک مرحله اصلی در پیش‌بینی‌های کوتاه مدت و بلندمدت می‌باشند ایفا می‌نمایند (مشیری و مروت، ۱۳۸۵). در این مقاله، به دنبال ارزیابی فرضیه فوق و به طور مشخص پاسخ به این سؤال هستیم که آیا پیش‌بینی نوسان شاخص‌های بورس اوراق بهادار تهران با استفاده از الگوریتم غیرخطی آنلاین حداقل میانگین مربعات مبتنی بر کرنل (KLMS)<sup>۹</sup> دقیق‌تر از پیش‌بینی الگوریتم شبکه‌های عصبی مبتنی بر حداقل میانگین مربعات خطا (LMS)<sup>۱۰</sup> است؟

## ۲. تشریح و بیان مساله

سرمایه‌گذاران در بازار، مایل به پذیرش سطح معینی از ریسک هستند، بنابراین مهم است که قادر به سنجش و پیش‌بینی این خطر باشیم (لیو و مرلی<sup>۱۱</sup>، ۲۰۰۹). برای انتخاب مناسب اوراق بهادار نیاز به یک ابزار پیش‌بینی احساس می‌شود. چنین سیستمی نیاز به آرایه اطلاعات دقیقی از چگونگی رفتار بازار سهام در روز بعد دارد. به همین دلیل است که مطالعات زیادی در مورد پیش‌بینی نوسان‌های بازار سهام انجام شده است. از این رو است که همیشه جستجوی مداوم برای یک فرضیه که قادر به پیش‌بینی نوسان‌های مقادیر آینده باشد وجود دارد (اسرینیواسان، ۲۰۰۳). لئونگ<sup>۱۲</sup> و همکاران (۲۰۰۰) بیان می‌کنند که ابزارهای گوناگون برای پیش‌بینی شاخص‌های مالی همراه با رشد اقتصادی در چند سال گذشته افزایش یافته است، که این باعث گسترش فرصت‌های سرمایه‌گذاری جهانی برای سرمایه‌گذاران فردی و نهادی شده است. لو و مکینلی<sup>۱۳</sup> (۱۹۸۸) نیز با استفاده از داده‌های بازار سهام با ثبات مانند ایالات متحده، اروپای غربی، و ژاپن، ادعا می‌کنند که شواهد قابل توجهی وجود دارد و نشان می‌دهند که بازده سهام تا حدودی قابل پیش‌بینی هستند. بنابراین پیش‌بینی نوسان شاخص‌های سهام برای توسعه استراتژی‌های تجاری موثر در بازار بسیار مهم است (لئونگ و همکاران، ۲۰۰۰). به منظور راهنمایی سرمایه‌گذاران و بهینه‌سازی مدیریت ریسک، الگوریتم‌های مختلف اعم از خطی یا غیر خطی جهت پیش‌بینی نوسان سری‌های زمانی شاخص‌های بورس اوراق بهادار ممکن است در لحظات و فواصل زمانی مختلف استفاده شود. با این حال، بیشتر محققان دریافته‌اند مسائل مالی و اقتصادی عمدتاً با یک سری روابط غیر خطی مخصوصاً در بازار سهام سرو کار دارند. لذا پیش‌بینی وضعیت آینده بازار سهام با استفاده از مدل‌های متعارف خطی نتایج مناسبی نخواهد داشت (یو<sup>۱۴</sup> و همکاران، ۲۰۰۹). در چنین شرایطی، برخی از تلاش‌ها به منظور توسعه مدل‌های مهندسی مالی با استفاده از برخی روش‌های محاسباتی هوشمند در حال ظهور صورت گرفته است. روش‌های مهندسی ابزار مناسبی برای محققان و پژوهشگران بابت پیش‌بینی و پیدا کردن راه‌حل‌های دقیق در یک محیط مشخصی که شامل اطلاعات پیچیده و نویزی است می‌باشند.

یکی از الگوریتم‌های غیرخطی که در سال‌های اخیر جهت پیش‌بینی نوسان شاخص‌های بورس اوراق بهادار استفاده شده است، شبکه‌های عصبی مصنوعی است که توسط بسیاری از

محققین به عنوان تکنیک مفید برای مدل سازی سری‌های زمانی مالی به دلیل توانایی آن با در نظر گرفتن روابط کارکردی نامحسوس در میان داده‌های تجربی معرفی شده است. (چنج<sup>۱۵</sup> و همکاران، ۱۹۹۶؛ ون و رابرت<sup>۱۶</sup>، ۱۹۹۷). با این حال، برخی مطالعات نشان می‌دهند که استفاده از شبکه‌های عصبی در پیش‌بینی نوسان شاخص‌های بازار سهام داری محدودیت‌هایی است. یکی از این محدودیت‌ها، محدودیت در یادگیری الگوهاست. زیرا داده‌های بازار سهام دارای نویز بالا و ابعاد پیچیده ای است، شبکه‌های عصبی اغلب نتایج متناقضی نشان می‌دهند و عملکرد غیر قابل پیش‌بینی بر روی داده‌های دارای نویز بالا دارند (منیش و دنموژی، ۲۰۰۵). هم چنین، استفاده از شبکه‌های عصبی جهت پیش‌بینی معمولاً به تعداد زیادی پارامتر نیاز دارند (ساو<sup>۱۷</sup>، ۲۰۰۱)، به طوری که وضعیت بهینه شان با انجام یک فرایند بهینه سازی غیر خطی تکرار شونده یافت می‌شود. برای بسیاری از مسائل، این رویه آموزش کند است، و همگرایی را برای رسیدن به راه حل بهینه تضمین نمی کند (ون ورنبرق<sup>۱۸</sup>، ۲۰۰۹). لیکن، روش‌های کرنل، به طور کلی دارای یک راه حل ظریف تر هستند که ناشی از فضای هیلبرت قابل بازیابی توسط کرنل و تحذب مسئله بهینه سازی حاصل شده، می‌باشند. بنابراین، بسیاری از الگوریتم‌های مبتنی بر کرنل راه حل منحصر به فرد دارند که می‌تواند از طریق حل یک مسئله بهینه سازی محدب یافت شود. یکی از الگوریتم‌های مبتنی بر کرنل که در این مطالعه جهت پیش‌بینی نوسان شاخص‌های بورس اوراق بهادار تهران استفاده شده است، الگوریتم حداقل میانگین مربعات خطا مبتنی بر کرنل است که اخیراً معرفی شده است و در تجزیه و تحلیل سری زمانی مالی می‌تواند بابت حل مشکل پیش‌بینی مفید باشد (مدقق<sup>۱۹</sup> و همکاران، ۲۰۱۰). مزیت عمده الگوریتم KLMS در مقایسه با شبکه‌های عصبی و سایر الگوریتم‌ها این است که یک روش پیش‌بینی غیر خطی آنلاین می‌باشد که جهت داده‌های بورس اوراق بهادار که به تدریج تغییر می‌کنند بسیار مناسب است. در سناریوهای آنلاین که در آن داده‌ها به طور مداوم در حال دریافت هستند، اندازه شبکه KLMS به طور مداوم رشد خواهد کرد و چالش‌های پیاده سازی نمایان خواهد شد (ون ورنبرق، ۲۰۰۹). بنابراین، در این مطالعه، هدف معرفی یک رویکرد جدید غیرخطی بر اساس الگوریتم KLMS جهت پیش‌بینی نوسان شاخص‌های بورس اوراق بهادار تهران می‌باشد که بتواند رفتار غیر خطی شاخص‌های بورس اوراق بهادار تهران را دقیق تر پیش بینی نماید، هم چنین قابلیت انطباق با تغییر مفهوم

را داشته باشد(هایکین<sup>۲۰</sup>، ۱۹۹۵).

### ۳. پیشینه تحقیق

شبکه‌های عصبی مصنوعی، الگوریتم ژنتیک و مدل‌های گارچ<sup>۲۱</sup> مدت زمان طولانی است که توسط محققان در تحقیقات مالی، با هدف دستیابی به فرایند تصمیم‌گیری قابل اعتماد از طریق شیوه‌های علمی مورد استفاده قرار گرفته است. ولی تاکنون تحقیقی در مورد استفاده از الگوریتم KLMS جهت پیش‌بینی نوسان شاخص‌های بازار سهام انجام نشده است. نمونه‌هایی از تحقیقات انجام شده با استفاده از سایر مدل‌ها و الگوریتم‌ها به شرح زیر ارائه می‌شود.

برای بررسی بازارهای اروپایی، فرانسس و ون دیچک<sup>۲۲</sup> (۱۹۹۶) از مدل گام تصادفی، گارچ، تارچ<sup>۲۳</sup> و مدل کیو گارچ<sup>۲۴</sup> برای بررسی عملکرد پیش‌بینی نوسان‌ها در کشورهای توسعه یافته سوئد، هلند، آلمان، ایتالیا و اسپانیا استفاده نمودند. نمونه شامل شاخص هفته‌ای در طی دوره ۱۹۸۶ تا ۱۹۹۴ بود. آنها دریافتند که مدل کیو گارچ مدل برتری است گارلیاسکاس<sup>۲۵</sup> (۱۹۹۹)، دونیس و جلیوف<sup>۲۶</sup> (۲۰۰۱)، اگلی<sup>۲۷</sup> و همکاران (۲۰۰۳)، نیگرن<sup>۲۸</sup> (۲۰۰۴)، پاکدین امیری و همکاران (۱۳۸۸) در تحقیقات مختلف از مدل شبکه عصبی برای پیش‌بینی نوسان شاخص‌های بازار سهام استفاده نمودند. آن‌ها نتیجه گرفتند که مدل شبکه‌های عصبی از توانایی بالایی در پیش‌بینی شاخص قیمت سهام در بورس اوراق بهادار برخوردار است.

لی<sup>۲۹</sup> (۲۰۰۵)، آرمانو<sup>۳۰</sup> و همکاران (۲۰۰۵) در تحقیقات جداگانه از ترکیب الگوریتم ژنتیک و الگوریتم شبکه‌های عصبی برای پیش‌بینی نوسان شاخص‌های سهام استفاده کردند. نتایج، توانایی بالای این الگوریتم ترکیبی را برای پیش‌بینی نشان می‌دهد.

لو<sup>۳۱</sup> و همکاران (۲۰۰۹) در مطالعه خود، یک مدل پیش‌بینی دو مرحله‌ای را با ادغام مدل‌های تجزیه و تحلیل ترکیبی مستقل<sup>۳۲</sup> و رگرسیون بردار پشتیبان<sup>۳۳</sup> برای سری‌های زمانی مالی پیشنهاد کردند. نتایج تجربی نشان می‌دهد که مدل پیشنهادی فوق دارای خطای پیش‌بینی پایین‌تر و دقت پیش‌بینی بالاتر نسبت به مدل گام تصادفی و رگرسیون بردار پشتیبان است. با توجه به آزمون، می‌توان نتیجه گرفت که مدل پیشنهادی می‌تواند به طور موثر نویز را از داده‌های سری زمانی مالی شناسایی و حذف کند و عملکرد پیش‌بینی مدل

رگرسیون بردار پشتیبان را بهبود دهد.

فیچچانگ و وانگ<sup>۳۴</sup> (۲۰۰۹) در مطالعه خود، ده تکنیک مختلف داده کاوی را برای پیش‌بینی تغییرات قیمت شاخص‌هانگ سنگ<sup>۳۵</sup> بازار سهام هنگ کنگ شامل تجزیه و تحلیل جدا کننده خطی<sup>۳۶</sup>، تجزیه و تحلیل جدا کننده کوادراتیک<sup>۳۷</sup>، روش طبقه بندی بیزی<sup>۳۸</sup>، مدل لگیت<sup>۳۹</sup>، درخت تصمیم<sup>۴۰</sup>، شبکه‌های عصبی<sup>۴۱</sup>، طبقه بندی بیزی با فرایند گاوسی<sup>۴۲</sup>، ماشین بردار پشتیبان<sup>۴۳</sup> و مدل حداقل مربعات ماشین بردار پشتیبان<sup>۴۴</sup> مورد استفاده قرار دادند. نتایج تجربی نشان می‌دهد که ماشین بردار پشتیبان و حداقل مربعات ماشین بردار پشتیبان پیش‌بینی بهتری در میان سایر مدل‌ها دارند. هم چنین ماشین بردار پشتیبان پیش‌بینی بهتری نسبت به حداقل مربعات ماشین بردار پشتیبان دارد. هم‌چنین

لیو و مرلی (۲۰۰۹) در مطالعه خود پیش‌بینی نوسان‌های شاخص‌هانگ سنگ را با استفاده از داده‌های بازار سهام هنگ کنگ بر مبنای انواع مختلف مدل‌های گارچ<sup>۴۵</sup> و مدل‌های تاریخی گام تصادفی<sup>۴۶</sup> و میانگین تاریخی<sup>۴۷</sup> انجام دادند. آنها بدین نتیجه رسیدند که مدل‌های گارچ نوسان‌های بازار سهام را با دقت بیشتری نسبت به مدل‌های تاریخی پیش‌بینی می‌کند. خمچندنی<sup>۴۸</sup> و همکاران (۲۰۰۹) در مطالعه خود، رویکرد جدیدی بر مبنای رگرسیون بردار پشتیبان<sup>۴۹</sup> را برای پیش‌بینی مالی معرفی نمودند و عملکرد آن را با رگرسیون بردار پشتیبان مقایسه نمودند. نتایج نشان می‌دهد که رگرسیون بردار پشتیبان دارای خطای پیش‌بینی کمتری نسبت به الگوریتم ماشین بردار پشتیبان است.

فیچچانگ<sup>۵۰</sup> (۲۰۱۰) روش جدید یادگیری ماشین را برای پیش‌بینی نوسان‌های شاخص کامپوزیت شانگهای به کار گرفت. در این تحقیق، او یک الگوریتم جدید از ماشین بردار مربوطه<sup>۵۱</sup> را برای پیش‌بینی نوسان‌های شاخص کامپوزیت شانگهای بر اساس مدل گارچ پیشنهاد کرد. مدل‌های ماشین پشتیبان بردار برگشتی، ماشین پشتیبان بردار برگشتی حداقل مربعات و مدل‌های گارچ نیز برای مقایسه با مدل پیشنهادی استفاده شدند. نتایج تجربی نشان می‌دهد که عملکرد روش ماشین بردار مربوطه پیشنهادی برای پیش‌بینی بهتر از روش‌های دیگر است.

هیسوهوانگ و جیسونگ<sup>۵۲</sup> (۲۰۱۰) از چهار مدل فازی Sugeno، FMI<sup>۵۳</sup>، FMII و FMIII برای پیش‌بینی روند شاخص قیمت سهام استفاده نمودند. نتایج نشان می‌دهد که مدل

FMI و FMII برای پیش بینی دارای خطای پیش‌بینی کمتری است. هم چنین آنها نتیجه گیری کردند که مدل FMII قابل اطمینان تر و دقیق تر از FMI است.

کارا و همکاران (۲۰۱۱) خود از دو مدل شبکه‌های عصبی مصنوعی و ماشین بردار پشتیبان در پیش‌بینی جهت تغییرات روزانه شاخص National 100 بورس اوراق بهادار استانبول استفاده کرده و نتایج بدست آمده را با یکدیگر مقایسه نمودند. نتایج تجربی نشان می‌دهد که عملکرد مدل شبکه‌های عصبی مصنوعی بهتر از مدل ماشین بردار پشتیبان است. اکوموس و گنس<sup>۵۴</sup> (۲۰۱۱) از مدل SV<sup>۵۵</sup> همراه با اجرای روش کارآمد لحظه ای<sup>۵۶</sup> به منظور تعیین روند نوسان‌های شاخص بازده سهام استفاده نمودند. به منظور انجام تجزیه و تحلیل تجربی، آنها از بازده روزانه شاخص‌های ISE-100، FTSE-100، S&P-500، NIKKEI-225 و BOVESPA استفاده نمودند. آنها بیان می‌کنند که مدل ماشین بردار همراه با اجرای روش کارآمد لحظه ای مطابق با بازده شاخص‌های ISE-100 و NIKKEI-225 است، هم چنین آن می‌تواند برای اهداف پیش‌بینی استفاده شود، اما آن برای پیش‌بینی شاخص‌های S&P-500، FTSE-100 و شاخص‌های BOVESPA مناسب نیست.

پی اسرینی واسان<sup>۵۷</sup> (۲۰۱۱) تحقیقی در مورد مدل سازی و پیش‌بینی نوسان‌های بازده شاخص S&P500 در بازار سهام ایالات متحده را انجام داد. در این مطالعه مدل‌های پیش‌بینی گارچ نسبتا ساده<sup>۵۸</sup> تا مدل‌های گارچ نسبتا پیچیده<sup>۵۹</sup> به کار گرفته شده اند، نتایج تحقیق نشان می‌دهد که مدل گارچ متقارن<sup>۶۰</sup> پیش‌بینی بهتری از روند شاخص S & P 500 نسبت به مدل‌های گارچ نامتقارن<sup>۶۱</sup> انجام می‌دهد.

بت شکن (۱۳۸۰) تحقیقی در مورد مدل سازی و پیش‌بینی نوسان‌های بازده شاخص S&P500 در بازار سهام ایالات متحده را انجام داد. در این مطالعه مدل‌های پیش‌بینی گارچ نسبتا ساده تا مدل‌های گارچ نسبتا پیچیده به کار گرفته شده اند، نتایج نشان می‌دهد که مدل گارچ متقارن پیش‌بینی بهتری از روند شاخص S&P500 نسبت به مدل‌های گارچ نامتقارن انجام می‌دهد.

آذر و افسر (۱۳۸۵) از مدل شبکه‌های عصبی فازی برای پیش‌بینی قیمت سهام استفاده نمودند و آن را از لحاظ معیارهای ارزیابی عملکرد میانگین مربعات خطا، مجذور میانگین مربعات خطا، میانگین مربعات خطای نرمالیزه شده، قدر مطلق میانگین خطا و  $R^2$  با روش



آریمما (ARIMA)<sup>۶۲</sup> مقایسه نمودند. نتایج تحقیق بیانگر آن است که شبکه‌های عصبی فازی در تمامی شش معیار ارزیابی عملکرد بر روش آریمما برتری دارد و دارای ویژگی‌های منحصر به فرد همگرایی سریع، دقت بالا، و توانایی تقریب تابع قوی هستند و برای پیش‌بینی شاخص قیمت سهام مناسب می‌باشند.

مشیری و مروت (۱۳۸۵) به پیش‌بینی شاخص کل بورس اوراق بهادار تهران را با استفاده از مدل‌های خطی و غیرخطی پرداختند. در این مطالعه، آنها نوسان شاخص کل را با استفاده از داده‌های روزانه و هفتگی در بازه زمانی سال‌های ۱۳۷۷ تا ۱۳۸۲ با مدل خطی آریمما و مدل‌های غیرخطی آرفیما (ARFIMA)<sup>۶۳</sup>، گارچ و شبکه عصبی مصنوعی<sup>۶۴</sup> پیش‌بینی نمودند. مقایسه دقت پیش‌بینی مدل‌های مذکور از طریق معیارهای پیش‌بینی مانند مجذور میانگین مربعات خطا، مجذور میانگین مربعات خطا و آماره U-Thiel نشان می‌دهد که مدل شبکه‌های عصبی مصنوعی در پیش‌بینی شاخص روزانه و هفتگی عملکرد بهتری نسبت به سایر مدل‌ها دارد، اما مقایسه آماری دقت پیش‌بینی مدل‌های مختلف با استفاده از آماره دیبلد-ماریانو<sup>۶۵</sup>، تفاوت معنی‌داری بین دقت پیش‌بینی مدل‌های مذکور را نشان نمی‌دهد.

پاکدین امیری و همکاران (۱۳۸۸) از شبکه‌های عصبی مصنوعی جهت پیش‌بینی شاخص قیمت سهام در بورس اوراق بهادار استفاده نمودند. شاخص صنعت، شاخص مالی و شاخص بازده نقدی به صورت سالانه به عنوان متغیرهای ورودی در نظر گرفته شدند. نتایج نشان می‌دهد که مدل شبکه عصبی پیشنهادی، توانایی بالایی در پیش‌بینی شاخص قیمت سهام دارد.

فلاح شمس (۱۳۸۹) کارایی مدل‌های ریسک سنجی شرکت جی.پی.مورگان و مدل اقتصاد سنجی گارچ را جهت تخمین ارزش در معرض خطر در بورس اوراق بهادار تهران با استفاده از نوسان‌های شاخص کل مورد بررسی قرارداد. با انجام آزمون شکست‌های احتمالی کوپیک، در سطح اطمینان ۹۵ و ۹۹ درصد، مشخص گردید که کارایی مدل‌های اقتصاد سنجی گارچ و ریسک سنجی تفاوت معنی‌داری نداشته و هر دو مدل از کارایی مناسبی برای پیش‌بینی ریسک بازار برخوردار می‌باشند.

دموری و همکاران (۱۳۹۰) شاخص کل قیمت سهام را با استفاده از الگوریتم پرواز پرندگان پیش‌بینی نمودند و عملکرد آن را با الگوهای سنتی مقایسه نمودند. نتایج تحقیق نشان

می‌دهد که در میان الگوهای سنتی، آریما دارای بهترین پیش‌بینی است؛ هم‌چنین در مقایسه الگوریتم پرواز پرندگان با آریما مشخص شد که خطای پیش‌بینی این الگوریتم هوشمند نسبت به آریما بسیار کم است که می‌توان از آن به منظور پیش‌بینی‌های آینده شاخص کل قیمت سهام استفاده کرد.

#### ۴. فرضیه‌های تحقیق

**فرضیه ۱:** پیش‌بینی نوسان شاخص کل بورس اوراق بهادار تهران با استفاده از الگوریتم KLMS دقیق‌تر از پیش‌بینی الگوریتم شبکه‌های عصبی مبتنی بر LMS است.

**فرضیه ۲:** پیش‌بینی نوسان شاخص قیمت و بازده نقدی بورس اوراق بهادار تهران با استفاده از الگوریتم KLMS دقیق‌تر از پیش‌بینی الگوریتم شبکه‌های عصبی مبتنی بر LMS است.

**فرضیه ۳:** پیش‌بینی نوسان شاخص صنعت بورس اوراق بهادار تهران با استفاده از الگوریتم KLMS دقیق‌تر از پیش‌بینی الگوریتم شبکه‌های عصبی مبتنی بر LMS است.

**فرضیه ۴:** پیش‌بینی نوسان شاخص بازار اول (شاخص تالار اصلی) بورس اوراق بهادار تهران با استفاده از الگوریتم KLMS دقیق‌تر از پیش‌بینی الگوریتم شبکه‌های عصبی مبتنی بر LMS است.

**فرضیه ۵:** پیش‌بینی نوسان شاخص ۵۰ شرکت برتر بورس اوراق بهادار تهران با استفاده از الگوریتم KLMS دقیق‌تر از پیش‌بینی الگوریتم شبکه‌های عصبی مبتنی بر LMS است.

**فرضیه ۶:** پیش‌بینی نوسان شاخص واسطه‌گری‌های مالی و پولی بورس اوراق بهادار تهران با استفاده از الگوریتم KLMS دقیق‌تر از پیش‌بینی الگوریتم شبکه‌های عصبی مبتنی بر LMS است.

**فرضیه ۷:** پیش‌بینی نوسان شاخص شناور آزاد بورس اوراق بهادار تهران با استفاده از الگوریتم KLMS دقیق‌تر از پیش‌بینی الگوریتم شبکه‌های عصبی مبتنی بر LMS است.

## ۵. روش تحقیق

### ۵-۱. داده‌ها و متغیرهای پژوهش

در این تحقیق جهت پیش‌بینی نوسان شاخص‌های بورس اوراق بهادار تهران، از داده‌های تاریخی روزانه شاخص قیمت کل، شاخص قیمت و بازده نقدی، شاخص صنعت، شاخص بازار اول (شاخص تالار اصلی)، شاخص ۵۰ شرکت برتر، شاخص واسطه‌گری‌های مالی و پولی طی دوره زمانی از ۱۳۸۶/۴/۳۱ تا ۱۳۹۰/۵/۳۰ شامل ۱۰۰۰ مشاهده و شاخص شناور آزاد طی دوره زمانی از ۱۳۸۷/۱۲/۵ تا ۱۳۹۰/۱۲/۱۳ شامل ۷۳۵ مشاهده به عنوان مجموعه‌های آموزشی که به بهترین وجه توصیف‌کننده ویژگی‌های سیگنال‌های مربوط به خود است استفاده می‌شود. داده‌های این پژوهش از سایت بورس اوراق بهادار تهران به صورت روزانه جمع‌آوری شده است. داده‌های جمع‌آوری شده، ابتدا به صورت سری‌های زمانی در قالب بانک اطلاعاتی در نرم افزار Excel دسته‌بندی گردید و سپس با استفاده از نرم افزارهای مطلب و SPSS16 زمینه تجزیه و تحلیل این داده‌ها و نتایج حاصل از آن فراهم شده است. دلیل این که در این مطالعه ۱۰۰۰ مشاهده وارد الگوریتم شده است این است که الگوریتم KLMS یک الگوریتم آنلاین می‌باشد که با هر تعداد داده (حتی خیلی کم) اجرا می‌شود و خیلی زود به مرحله بهینه می‌رسد و از این مرحله به بعد خطای پیش‌بینی آن بسیار کم می‌شود.

### ۵-۲. ارزیابی عملکرد پیش‌بینی

به طور معمول، سرمایه‌گذاران در بازارهای مالی، منافع بالاتر و توجه بیشتری نسبت به دقت و صحت پیش‌بینی از بابت میانگین مربعات خطا و صحت روند پیش‌بینی در شاخص‌های بازار سهام دارند، زیرا یک پیش‌بینی با دقت بالا می‌تواند در تصمیم‌گیری سرمایه‌گذاری‌های صحیح و سودمند، مفید باشد. لذا در این مطالعه، عملکرد پیش‌بینی و قابلیت اطمینان الگوریتم‌های پیشنهادی با استفاده از معیارهای میانگین مربعات خطا (MSE)<sup>۶۶</sup>، مجذور میانگین مربعات خطا (RMSE)<sup>۶۷</sup>، میانگین مربعات خطای نرمالیزه شده (NMSE)<sup>۶۸</sup> قدر مطلق میانگین خطا (MAD)<sup>۶۹</sup>، صحت جهت پیش‌بینی (DS)<sup>۷۰</sup>، صحت روند رو به بالا (CP)<sup>۷۱</sup> و صحت روند روبه پایین (CD)<sup>۷۲</sup> سنجیده می‌شود. RMSE، MSE و NMSE و MAD معیارهای انحراف بین مقادیر واقعی و مقادیر پیش‌بینی شده می‌باشند. هر چه مقادیر

MSE, RMSE, NMSE و MAD کمتر باشند؛ الگوریتم، معیار مناسب‌تری برای پیش‌بینی نوسان شاخص‌ها تلقی خواهد شد. مقدار کم MSE, RMSE, NMSE و MAD نشان‌دهنده این است که مقادیر پیش‌بینی شده سری به مقادیر واقعی آن‌ها نزدیک ترند. معیار DS صحت جهت پیش‌بینی شاخص‌ها را به درصد بیان می‌کند. معیارهای CP و CD صحت روند پیش‌بینی رو به بالا و روند پیش‌بینی رو به پایین شاخص را به درصد بیان می‌کند. معیارهای DS, CP و CD را می‌توان برای ارزیابی دقت پیش‌بینی مورد استفاده قرار داد. معیارهای پیش‌بینی MSE, RMSE, NMSE, MAD, DS, CP و CD به صورت زیر تعریف می‌شوند:

$$MSE = \frac{\sum_{i=1}^n (d_i - \hat{d}_i)^2}{n}$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (d_i - \hat{d}_i)^2}{n}}$$

$$NMSE = \frac{1}{\sigma_n^2} \sum_{i=1}^n (d_i - \hat{d}_i)^2, \quad \text{where } \sigma^2 = \frac{1}{n-1} \sum_{i=1}^n (\hat{d}_i - \hat{d}_{i-1})^2$$

$$MAD = \frac{\sum_{i=1}^n |d_i - \hat{d}_i|}{n}$$

$$DS = \frac{100}{n} \sum_{i=1}^n T_i, \quad \text{where } T_i = \begin{cases} 1 & (d_i - d_{i-1})(\hat{d}_i - \hat{d}_{i-1}) \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

$$CP = \frac{100}{n_1} \sum_{i=1}^n T_i, \quad \text{where } T_i = \begin{cases} 1 & (d_i - d_{i-1}) > 0 \text{ and } (\hat{d}_i - \hat{d}_{i-1}) \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

$$CD = \frac{100}{n_2} \sum_{i=1}^n T_i, \quad \text{where } T_i = \begin{cases} 1 & (d_i - d_{i-1}) < 0 \text{ and } (\hat{d}_i - \hat{d}_{i-1}) \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

که در آن  $d_i$  و  $\hat{d}_i$  به ترتیب نشان دهنده مقادیر واقعی و پیش‌بینی شده است.  $n$  مجموع داده‌هاست،  $n_1$  تعداد نقاط داده متعلق به روند رو به بالا و  $n_2$  تعداد نقاط داده متعلق به روند رو به پایین است.

### ۳-۵. مروری بر مباحث نظری الگوریتم‌های به کار رفته:

#### ۳-۵-۱. الگوریتم شبکه‌های عصبی مبتنی بر حداقل میانگین مربعات خطا (LMS)

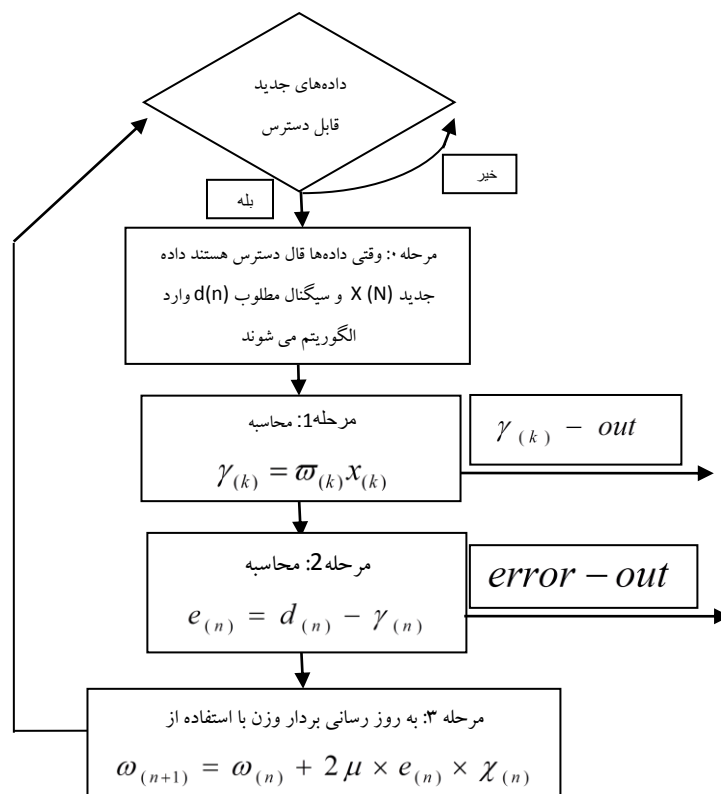
الگوریتم حداقل میانگین مربعات خطا توسط ویدرو<sup>۷۳</sup> و همکاران در سال ۱۹۵۹ به عنوان یک راه حل ساده برای آموزش یک سیستم خطی وفقی با به حداقل رساندن میانگین مربعات

خطا پیشنهاد شد. الگوریتم شبکه‌های عصبی مبتنی بر LMS به طور گسترده‌ای به عنوان الگوریتم مرجع به دلیل سادگی و کارآمدی بالای آن استفاده می‌شود (ون ویرنبرق، ۲۰۰۹)<sup>۷۴</sup>، زیرا آن نیاز به محاسبه همبستگی تابع و هم چنین نیاز به معکوس ماتریس ندارد. هم چنین پیاده سازی این الگوریتم بسیار ساده بوده و در مسائل خطی بسیار خوش رفتار است. الگوریتم شبکه‌های عصبی مبتنی بر LMS یک الگوریتم کلاسیک گرادیان تصادفی است و دارای یک پروسه تکرار شونده‌ای است که باعث اصلاحات متوالی بردار وزن در جهت منفی بردار گرادیان می‌شود و در نهایت منجر به حداقل میانگین مربعات خطا می‌شود (گیوانیس، ۲۰۱۰)<sup>۷۵</sup>. منظور از گرادیان منفی، تغییر وزن‌ها با توجه به مشتقات خطای شبکه، به گونه ای است که خطای شبکه به حداقل برسد. بنابراین در الگوریتم شبکه‌های عصبی مبتنی بر LMS از روش تکراری گرادیان تصادفی برای محاسبه بردار وزن استفاده می‌شود. اما به جای گرادیان گرفتن از میانگین مربعات خطا، از خود مربعات خطا گرادیان می‌گیرد و در نهایت ثابت می‌شود که با تعداد نمونه کافی می‌توان میانگین مربعات خطا را حداقل کرد. ساختار این الگوریتم در پیوست الف نشان داده شده است.

### ۲-۳-۵. پیاده سازی گام به گام الگوریتم شبکه‌های عصبی مبتنی بر LMS

شکل ۱ نشان دهنده روند کار الگوریتم شبکه‌های عصبی مبتنی بر LMS است، پیاده سازی الگوریتم شبکه‌های عصبی مبتنی بر LMS در ۳ مرحله بیان شده است. مرحله صفر، مرحله اولیه سیستم می‌باشد. هنگامی که داده‌ها قابل دسترس هستند، داده جدید  $x_{(n)}$  و سیگنال مطلوب  $d_{(n)}$  وارد الگوریتم می‌شوند. داده‌های وارد شده در مرحله صفر ابتدا نرمالیزه می‌شوند و سپس به مرحله ۱ فرستاده می‌شود، جایی که سیگنال خروجی  $\gamma_{(n)}$  از طریق معادله  $\gamma_{(n)} = \omega_{(n)} x_{(n)}$  طبق فرمول ۳ پیوست ب جهت پیش‌بینی مقدار شاخص برای روز بعد، یک ماه بعد و یا برای یک سال بعد محاسبه می‌شود. در مرحله ۲ سیگنال خطا با استفاده از اختلاف خروجی واقعی  $\gamma_{(n)}$  (مقدار پیش‌بینی شده با استفاده از الگوریتم برای روز بعد، یک ماه بعد و یا برای یک سال بعد) و خروجی مطلوب  $d_{(n)}$  (مقدار مطلوب شاخص) یعنی  $e_{(n)} = d_{(n)} - \gamma_{(n)}$  طبق فرمول ۲ پیوست الف محاسبه می‌شود. تا این مرحله، خروجی الگوریتم و سیگنال خطا محاسبه شدند. سپس، به روز رسانی ضرایب فیلتر و فقی از طریق معادله  $\omega_{(n+1)} = \omega_{(n)} + 2\mu \times e_{(n)} \times \chi_{(n)}$  طبق فرمول ۱ پیوست الف انجام می‌شود.

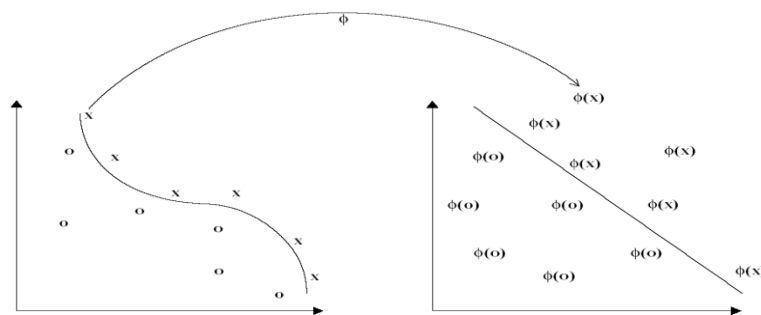
$\mu$  سرعت و پایداری الگوریتم را نشان می‌دهد. در عمل، برای رسیدن به نرخ همگرایی سریع و پایداری مناسب، ضریب اندازه گام  $\mu$  کوچکتر از یک می‌باشد. افزایش اندازه گام، سرعت الگوریتم و همچنین خطر ناپایداری را افزایش می‌دهد. در مقابل کاهش اندازه گام، موجب پایداری بیشتر الگوریتم می‌گردد. اگر  $\mu$  خیلی کوچک باشد، ممکن است پروسه خیلی طولانی شود و اگر خیلی بزرگ باشد، امکان ناپایداری وجود دارد. با توجه به اینکه الگوریتم شبکه‌های عصبی مبتنی بر LMS اساساً الگوریتمی مبتنی بر گرادیان تصادفی است، بنابراین  $\mu$  باید بگونه ای انتخاب شود که میانگین مربعات خطا حداقل بماند. و در نهایت، در مرحله ۳، ضریب  $\omega_{(n+1)}$  به مرحله اول ارسال می‌شود، که در آن مجموعه جدیدی از داده‌ها انتظار می‌رود چرخه ی جدید فرآیند مناسب را شروع نمایند.



شکل ۱: نمودار الگوریتم شبکه‌های عصبی مبتنی بر LMS

### ۳-۳-۵. بحث کرنل

در سال‌های اخیر، روش‌های کرنل بسیار مورد استقبال محققین قرار گرفته‌اند. این به دلیل پایه محکم ریاضی و کاربردهای تجربی موفق آنهاست (لیو و پرینسیپ<sup>۷۶</sup>، 2008a). روش‌های کرنل برای نگاشت داده‌های ورودی از فضای اولیه به فضای ویژگی (که معمولاً دارای بُعد خیلی بالایی است) به کار می‌روند. در فضای ویژگی تابع نگاشت  $\Phi$  برای پیدا کردن روابط خطی بین داده‌های ورودی مورد استفاده قرار می‌گیرد که در شکل ۲ نشان داده شده است. در قلب روش‌های کرنل یک تابع کرنل است. تابع کرنل تابعی است که روش‌های کرنل را برای عمل در فضای ویژگی نگاشته شده، بدون حتی محاسبه مختصات در آن فضا قادر می‌سازد، که این با کمک حقه کرنل مشهور انجام می‌شود (غفاریان<sup>۷۷</sup> و همکاران، ۲۰۱۰). برای این که گاهی تابع تبدیل مشخص نیست و یا به دلیل محاسبات زیادی که این عمل دارد از حقه کرنل استفاده می‌شود، که به موجب آن به جای تابع کرنل در فضای حاصل از نگاشت، از تابع کرنل در فضای اولیه استفاده می‌کنیم. الگوریتم‌ها بسیاری، توسعه یافته‌اند که بر مبنای روش‌های کرنل کار می‌کنند. یکی از الگوریتم‌هایی که اخیراً پیشنهاد شده است الگوریتم حداقل میانگین مربعات خطا مبتنی بر کرنل (KLMS) است. که در ادامه توضیح آن ارائه می‌شود.



شکل ۲: تابع نگاشت  $\Phi$

### ۳-۳-۴. الگوریتم حداقل میانگین مربعات خطا مبتنی بر کرنل

پیش‌بینی برخی نوسان سری‌های زمانی می‌تواند با یک رویکرد بهینه جدید، به نام الگوریتم KLMS انجام شود (پخارل<sup>۷۸</sup> و همکاران، ۲۰۰۷). اگر چه الگوریتم شبکه‌های عصبی

مبتنی بر LMS جهت الگوهای خطی مناسب است، اما آن برای الگوهای غیر خطی توسعه پیدا نکرده است (سالخورده حقیقی<sup>۹</sup> و همکاران، ۲۰۱۰). برای غلبه بر این مشکل، اخیراً، پخارل<sup>۸۰</sup> و همکاران (۲۰۰۷) و لیو<sup>۸۱</sup> و همکاران (۲۰۰۸b) نشان دادند الگوریتم شبکه‌های عصبی مبتنی بر LMS را می‌توان در فضای هیلبرت قابل بازبازی توسط کرنل برای به دست آوردن راه حل در فضای ورودی استفاده کرد. الگوریتم پیشنهادی KLMS نامیده می‌شود. هدف استفاده از KLMS نگاشت سیگنال ورودی به فضای دارای بُعد بالاست تا بتوان مسائل غیرخطی را با استفاده از آن حل کرد (مدقق و همکاران، ۲۰۱۰). ساختار این الگوریتم در پیوست ب نشان داده شده است.

### ۵-۳-۵. پیاده سازی گام به گام الگوریتم KLMS در این تحقیق

هنگامی که داده‌ها قابل دسترس هستند، داده‌های جدید  $x_{(n)}$  و سیگنال مطلوب  $d_{(n)}$  وارد الگوریتم KLMS می‌شوند. با ورود داده‌های بیشتر، اندازه شبکه KLMS به طور مداوم رشد خواهد کرد و عملکرد پیش‌بینی بهتر خواهد شد. مزیت الگوریتم‌های آنلاین این است زمانی که داده جدید می‌رسند نیاز به یادگیری مجدد ندارند. در این الگوریتم داده‌های وارد شده ابتدا نرمالیزه می‌شوند. سپس بایستی تابع کرنل مورد استفاده جهت پیش‌بینی مشخص شود. اگرچه که چندین انتخاب برای تابع کرنل قابل دسترس هستند، ولی گسترده ترین تابع کرنلی که استفاده می‌شود کرنل گوسی است که به شکل

$$k(x_{(n)}, x_{(n-1)}) = e^{-\frac{\|x_{(n)} - x_{(n-1)}\|^2}{2\sigma^2}}$$

تعریف می‌شود. سپس با استفاده از تابع کرنل گوسی طبق فرمول شماره ۵ پیوست ب فاصله داده  $x_{(n)}$  را با هر یک از داده‌های قبلی آن به شرح ذیل محاسبه می‌کنیم:

$$k(x_{(n)}, x_{(n-1)}) = e^{-\frac{\|x_{(n)} - x_{(1)}\|^2}{2\sigma^2}} + e^{-\frac{\|x_{(n)} - x_{(2)}\|^2}{2\sigma^2}} + e^{-\frac{\|x_{(n)} - x_{(3)}\|^2}{2\sigma^2}} + \dots + e^{-\frac{\|x_{(n)} - x_{(n-1)}\|^2}{2\sigma^2}}$$

که در آن  $e$  نشان دهنده مقدار ۲/۷۱۸۲۸۱۸ و  $\sigma^2$  نشان دهنده عرض کرنل گوسی است که مقدار آن بسته به نوع داده‌ها توسط کاربر تعیین می‌شود را محاسبه نموده و در مقدار خطای پیش‌بینی محاسبه شده در مراحل قبل (یعنی اگر تفاوت مقدار  $x_{(n)} - x_{(3)}$  را محاسبه می‌کنیم تفاوت به دست آمده را در خطای پیش‌بینی شاخص در روز سوم ( $d_{(3)} - x_{(3)}$ ) ضرب



نموده و مقدار  $\sum_{i=0}^{n-1} e_{(i)} k(x_{(n)}), (x_{(i)})$  بدست می‌آوریم. سپس طبق فرمول شماره ۶ پیوست ب، حاصل فوق در مقدار ضریب اندازه گام  $\mu$  ضرب می‌شود تا خروجی الگوریتم KLMS محاسبه شود. یعنی  $\gamma_{(n)} = \mu \sum_{i=0}^{n-1} e_{(i)} k(x_{(n)}), (x_{(i)})$  به دست می‌آید. که همان خروجی الگوریتم یعنی مقدار پیش‌بینی شده برای روز بعد، ماه بعد و یا سال بعد می‌باشد.  $\mu$  ضریب اندازه گام است که مقدار آن توسط کاربر تعیین می‌شود و معمولاً بین صفر و ۱ می‌باشد.

الگوریتم‌های ارائه شده چون از نوع سری زمانی است نحوه کنترل مسائل سری‌های زمانی از جمله خود همبستگی به این صورت می‌باشد که: الف) پیش‌بینی سری زمانی بر خط توسط فیلترهای وفقی صورت می‌گیرد. الگوریتم‌های KLMS و شبکه‌های عصبی مبتنی بر LMS از جمله الگوریتم‌های شناخته شده در این حیطه هستند. ب) همان طور که در روابط ریاضی الگوریتم شبکه‌های عصبی مبتنی بر LMS و KLMS نیز مشاهده می‌گردد به روز رسانی فیلتر شامل مفهوم همبستگی و همبستگی تقاطعی می‌باشد. لذا مفهوم همبستگی در ذات این فیلتر نهفته است. الگوریتم شبکه‌های عصبی مبتنی بر LMS و KLMS برگرفته از فیلتر وینر است که مفهوم همبستگی و همبستگی تقاطعی را در پیش‌بینی وارد کرده است.

## ۶. تجزیه و تحلیل نتایج و آزمون فرضیه‌ها

### ۶-۱. آمار توصیفی

برای ارزیابی خصوصیات توزیعی سری‌های بازدهی شاخص‌ها، آمار توصیفی برای خلاصه کردن یا توصیف مجموعه داده‌ها در نگاره ۱ ارائه شده است. همان طور که مشاهده می‌شود میانگین و میانه بازده روزانه شاخص‌ها مثبت و به میزان قابل توجهی متفاوت از صفر است. این نشان می‌دهد که قیمت سهام در طی زمان افزایش کمی دارد (هیین<sup>۸۲</sup>، ۲۰۰۸). انحراف معیار دارای مقدار بالایی می‌باشد، نشان دهنده سطح بالایی از نوسان‌های بازدهی روزانه شاخص‌های بورس اوراق بهادار است (پی اسرینوآسان، ۲۰۱۱). ضریب چولگی تمامی شاخص‌ها به استثنای شاخص شناور آزاد مثبت است. لذا جامعه دارای چوله به راست است و توزیع نامتقارن است. ضریب کشیدگی سری‌ها در تمامی موارد کمتر از ۳ است. این مساله نشانگر این است که توزیع سری‌ها در اطراف میانگین نیست. به عبارت دیگر ضریب کشیدگی نشان می‌دهد مقادیر شاخص‌ها دارای کشیدگی کمتری نسبت به مقدار نرمال است.

## نگاره ۱: آمار توصیفی

شاخص	تعداد	میانگین	میانه	انحراف معیار	ضریب چولگی	ضریب کشیدگی	کمترین مقدار	بیشترین مقدار
کل	۱۰۰۰	۱۳۵۸۴.۳	۱۱۶۷۸.۸	۵۲۰۹.۱۴	۱.۱۷۵	۰.۱۵	۷۹۵۵.۴	۲۶۶۷۵.۹
قیمت و بازده نقدی	۱۰۰۰	۵۲۲۱۵.۶	۴۶۷۹۱	۲۱۹۳۰.۵۵	۱.۰۵۶	۰.۱۶	۱۰۱۸۰.۹	۱۰۸۲۶۵
شناور آزاد	۷۳۵	۲۱۸۱۷۷.۶	۲۲۱۵۹.۲	۸۸۹۱.۶۲	-۰.۰۷۷	-۱.۱۱	۱۱۴۰.۳	۳۵۵۹۶.۱
بازار اول	۱۰۰۰	۱۱۴۲۱.۸	۹۶۹۶.۴	۴۴۱۱.۷۱	۱.۱۹۵	۰.۲۹	۶۳۷۳.۵	۲۲۶۳۶.۸
صنعت	۱۰۰۰	۱۰۵۲۶.۸	۹۱۶۸.۶۹	۳۹۵۱.۵	۱.۳۱۳	۰.۳۳	۶۱۵۷	۲۰۷۹۶.۶
۵۰ شرکت برتر	۱۰۰۰	۶۲۴.۵۶۱	۵۵۷.۵۶	۲۵۹.۴۴	۱.۱۴۵	۰.۵۱	۲۸۹.۸	۱۲۹۰.۴
واسطه گری‌های مالی و پولی	۱۰۰۰	۳۳۷۸۱.۷	۲۵۶۸۵.۶	۱۵۷۶۲.۹۷	۰.۹۵۱	-۰.۵۸	۱۸۶۳۹.۲	۷۰۱۹۵.۷

## ۲-۶. تجزیه و تحلیل نتایج و آزمون فرضیه‌ها

جهت بررسی فرضیه‌های تحقیق، نتایج پیش‌بینی مقدار روز بعد، ماه بعد و سال بعد شاخص‌های بورس اوراق بهادار تهران با استفاده از الگوریتم‌های شبکه‌های عصبی مبتنی بر LMS و KLMS از لحاظ معیار میانگین مربعات خطا و صحت روند پیش‌بینی در زیر ارائه می‌شود. الگوریتم شبکه‌های عصبی مبتنی بر LMS از قیمت روز قبل و بر مبنای به روز رسانی بردار وزن برای پیش‌بینی قیمت امروز استفاده می‌کند. به همین دلیل دارای سرعت بالایی برای پیش‌بینی است. اما الگوریتم KLMS از تمامی قیمت روزهای قبل تا مرحله آموزش برای پیش‌بینی قیمت امروز استفاده می‌کند. باید توجه داشت مجموعه داده‌های اصلی ابتدا نرمالیزه شدند. هدف از انجام این کار این است که اطمینان حاصل شود اثر متغیرهای ورودی با مقدار زیاد در مقادیر ورودی کوچکتر کاهش پیدا نمی‌کند، بنابراین به کاهش خطاهای پیش‌بینی کمک می‌کند (لو و همکاران، ۲۰۰۹). هم‌چنین میانگین مربعات خطا، مجذور میانگین مربعات خطا، میانگین مربعات خطای نرمالیزه شده، قدر مطلق میانگین خطا، صحت جهت پیش‌بینی، صحت روند پیش‌بینی رو به بالا و صحت روند پیش‌بینی روبه پایین الگوریتم‌های KLMS و شبکه‌های عصبی مبتنی بر LMS از زمان همگرا شدن آن‌ها محاسبه و ارائه شده است زیرا این الگوریتم‌ها نیاز به زمان یادگیری دارند و بعد از یادگیری (همگرا شدن) میزان خطا محاسبه و با یکدیگر مقایسه می‌شوند.

۱-۲-۶. نتایج پیش‌بینی نوسان شاخص‌های بورس اوراق بهادار تهران با استفاده از الگوریتم آنلاین KLMS و الگوریتم شبکه‌های عصبی مبتنی LMS (در افق‌های زمانی کوتاه مدت، میان مدت و بلندمدت):

نتایج پیش‌بینی نوسان شاخص‌های بورس اوراق بهادار تهران با استفاده از الگوریتم آنلاین KLMS پیشنهادی با نتایج پیش‌بینی الگوریتم شبکه‌های عصبی مبتنی بر LMS از لحاظ معیارهای میانگین مربعات خطا و صحت روند پیش‌بینی مقایسه و در نگاره‌های ۲، ۳ و ۴، ارائه شده است. از نگاره‌های ۲، ۳ و ۴ مشاهده می‌شود که معیار میانگین مربعات خطا (MSE)، مجذور میانگین مربعات خطا (RMSE)، میانگین مربعات خطای نرمالیزه شده (NMSE) و قدر مطلق میانگین خطای (MAD) الگوریتم KLMS، در تمامی شاخص‌ها به میزان قابل توجهی کمتر از الگوریتم شبکه‌های عصبی مبتنی بر LMS است. این نشان دهنده این است که میزان انحراف پایین تری بین مقادیر واقعی و مقادیر پیش‌بینی شده با استفاده از الگوریتم پیشنهادی KLMS وجود دارد. هم‌چنین مقادیر صحت روند پیش‌بینی (DS)، صحت روند پیش‌بینی رو به بالا (CP) و صحت روند پیش‌بینی رو به پایین (CD) الگوریتم KLMS دارای مقدار بالاتری نسبت به الگوریتم شبکه‌های عصبی مبتنی بر LMS می‌باشد.

نگاره ۲: پیش‌بینی نوسان‌های روز بعد شاخص‌های بورس اوراق بهادار تهران با استفاده از

الگوریتم‌های KLMS و شبکه‌های عصبی مبتنی بر LMS

الگوریتم	MSE SUM	RMSE	NMSE	MAD	DS(%)	CD(%)	CP(%)
شاخص کل	KLMS	۰.۰۰۰۰۱۷	۰.۰۰۰۴۱	۰.۴۲۹۱	۰.۰۰۲۶	۶۸.۷۹	۶۵.۳۷
	LMS	۰.۰۰۰۰۱۸	۰.۰۰۰۴۳	۰.۴۶۴۹	۰.۰۰۲۶	۶۸.۰۴	۶۶.۲۸
شاخص قیمت و بازده نقدی	KLMS	۰.۰۰۱۵۷۸	۰.۰۰۳۹۷	۳۸.۳۷۵۶	۰.۰۰۴۵	۶۸.۳۸	۶۴.۹۱
	LMS	۰.۰۰۲۱۱۲	۰.۰۰۴۶	۴۹.۳۰۸۸	۰.۰۰۵۹	۶۸.۲۵	۶۶.۷۴
شاخص شناور آزاد	KLMS	۰.۰۰۰۰۵۷	۰.۰۰۲۳۹	۶.۶۹۹۴	۰.۰۰۵۵	۶۵.۱	۵۷.۵
	LMS	۰.۰۰۰۰۶۰۵	۰.۰۰۲۴۶	۷.۰۲۱۵	۰.۰۰۶۵	۶۵.۹۶	۵۹.۳۴
شاخص بازار اول	KLMS	۰.۰۰۰۰۱۹	۰.۰۰۰۴۳	۰.۴۹۲۳	۰.۰۰۲۷	۶۸.۹۹	۶۶.۸۲
	LMS	۰.۰۰۰۰۰۲۱	۰.۰۰۰۴۶	۰.۵۲۸۵	۰.۰۰۲۷	۶۸.۶۶	۶۷.۷۳
شاخص صنعت	KLMS	۰.۰۰۰۰۰۱۹	۰.۰۰۰۴۳	۰.۵۱۴۳	۰.۰۰۲۷	۶۸.۷۹	۶۷.۰۵
	LMS	۰.۰۰۰۰۰۲۱	۰.۰۰۰۴۵	۰.۵۴۸۱	۰.۰۰۲۷	۶۷.۸۴	۶۸.۰۶

CP(%)	CD(%)	DS(%)	MAD	NMSE	RMSE	MSE SUM	الگوریتم	
۶۸.۰۳	۶۹.۳۲	۶۶.۹۷	۰.۰۰۲۹	۰.۵۴۵۸	۰.۰۰۴۷	۰.۰۰۰۰۲۲	KLMS	شاخص پنجاه
۶۹.۵	۶۷.۹	۶۶.۹۱	۰.۰۰۲۹	۰.۵۸۱۲	۰.۰۰۰۵	۰.۰۰۰۰۲۵	LMS	شرکت برتر
۶۰.۶۳	۷۰.۶۸	۶۵.۰۵	۰.۰۰۲۷	۰.۳۵۳۳	۰.۰۰۴۲	۰.۰۰۰۰۱۸	KLMS	شاخص واسطه
۶۱.۵۶	۶۹.۰۲	۶۴.۴۳	۰.۰۰۰۳	۰.۴۷۶۳	۰.۰۰۰۵	۰.۰۰۰۰۲۵	LMS	گرهای پولی و مالی

نگاره ۳: پیش‌بینی نوسان‌های یک ماه بعد شاخص‌های بورس اوراق بهادار تهران با استفاده از

الگوریتم‌های KLMS و شبکه‌های عصبی مبتنی بر LMS

CP(%)	CD(%)	DS(%)	MAD	NMSE	RMSE	MSE	الگوریتم	
۶۵.۸۹	۷۲.۸۵	۶۹.۲۷	۰.۰۰۰۲۵	۰.۴۱۶۱۵	۰.۰۰۴۰۸	۰.۰۰۰۰۱۷	KLMS	شاخص کل
۵۴.۹۹	۶۵.۵۸	۶۰.۴۲	۰.۰۰۰۶۲	۱.۷۷۹۵۶	۰.۰۰۸۴۹	۰.۰۰۰۰۷۲	LMS	
۶۵.۱۳	۷۱.۱۳	۶۸.۱۳	۰.۰۰۴۹۸	۳۸.۶۲۲۷۹	۰.۰۴۰۷۱	۰.۰۰۱۶۵۷	KLMS	شاخص قیمت و بازده نقدی
۵۳.۳۵	۶۳.۲۹	۵۸.۵۴	۰.۰۲۵۸	۳۳۴.۹۵۳۳۸	۰.۱۳۰۷۶	۰.۰۱۷۰۹۹	LMS	
۶۰.۵۲	۷۰.۴۱	۶۶.۰۴	۰.۰۰۵۴	۶.۸۵۴۱۴	۰.۰۲۴۱۷	۰.۰۰۰۵۸۴	KLMS	شاخص شناور آزاد
۵۰.۵۵	۶۶.۱۱	۵۹.۵۷	۰.۰۲۱۲۵	۴۵.۴۹۷۷۲	۰.۰۶۵۶۴	۰.۰۰۴۳۰۸	LMS	
۶۷.۴۴	۷۳.۶۷	۶۹.۴۸	۰.۰۰۰۲۷	۰.۴۷۵۹۲	۰.۰۰۴۳۸	۰.۰۰۰۰۱۹	KLMS	شاخص بازار اول
۵۳.۱۲	۶۶.۶	۵۹.۲۷	۰.۰۰۰۶۹	۲.۱۴۵۱۸	۰.۰۰۹۳۴	۰.۰۰۰۰۸۷	LMS	
۶۶.۸۲	۷۱.۷۳	۶۸.۱۳	۰.۰۰۰۲۷	۰.۴۹۶۰۹	۰.۰۰۴۳۵	۰.۰۰۰۰۱۹	KLMS	شاخص صنعت
۵۵.۶۱	۶۵.۵	۵۹.۷۹	۰.۰۰۰۶۵	۲.۰۸۱۵۷	۰.۰۰۸۹۶	۰.۰۰۰۰۰۸	LMS	
۶۶.۲۳	۶۹.۷۱	۶۶.۲۵	۰.۰۰۰۲۹	۰.۵۱۹۵۷	۰.۰۰۴۷۶	۰.۰۰۰۰۲۳	KLMS	شاخص پنجاه شرکت برتر
۵۳.۸۶	۶۴.۳۲	۵۷.۷۱	۰.۰۰۰۰۷	۲.۰۳۰۳۹	۰.۰۰۹۴۷	۰.۰۰۰۰۰۹	LMS	
۶۱.۳۳	۷۰.۷۹	۶۵.۳۱	۰.۰۰۰۲۷	۰.۳۵۱۲۵	۰.۰۰۴۳	۰.۰۰۰۰۱۸	KLMS	شاخص واسطه‌گرهای پولی و مالی
۵۰.۲۵	۶۵.۳۶	۵۷.۶	۰.۰۰۰۰۷	۱.۷۳۴۳۹	۰.۰۰۹۶	۰.۰۰۰۰۹۲	LMS	

نگاره ۴: پیش‌بینی نوسان‌های یک سال بعد شاخص‌های بورس اوراق بهادار تهران با استفاده از

الگوریتم‌های آنالین KLMS و شبکه‌های عصبی مبتنی بر LMS

الگوریتم	MSE	RMSE	NMSE	MAD	DS(%)	CD(%)	CP(%)	
KLMS	۰.۰۰۰۰۳	۰.۰۰۵۸۹	۰.۴۷۲۱۸	۰.۰۰۳۷۹	۶۶.۶۱	۷۴.۱۹	۵۶.۳۸	شاخص کل
LMS	۰.۰۰۰۰۲	۰.۰۱۴۲۴	۲.۷۳۹۴۷	۰.۰۱۰۴۳	۵۵.۹۷	۶۷.۷۴	۳۹.۰۹	
KLMS	۰.۰۰۲۴۸	۰.۰۴۹۸۱	۳۲.۲۸۰۵۱	۰.۰۰۶۸۹	۶۶.۱۳	۷۱.۵۵	۵۹.۴۵	شاخص قیمت و بازده نقدی
LMS	۰.۰۰۴۲۸	۰.۰۶۵۴۲	۵۱.۵۹۴۷۲	۰.۰۱۸۴۹	۵۶.۴۵	۶۸.۲۳	۴۰.۵۵	
KLMS	۰.۰۰۰۰۵	۰.۰۰۷۳۷	۱.۴۰۷۶۵	۰.۰۰۵۴۴	۶۰.۲۸	۶۳.۴۵	۵۶.۶۹	شاخص شناور آزاد
LMS	۰.۰۲۰۱۸	۰.۱۴۲۰۴	۱۷۶.۹۳۶۸۷	۰.۰۴۵۳۷	۵۴.۹۳	۶۰.۴۱	۴۸.۴۱	
KLMS	۰.۰۰۰۰۳	۰.۰۰۵۴۵	۰.۳۸۸۲	۰.۰۰۳۵۳	۶۷.۴۲	۷۴.۳۳	۵۸.۳۳	شاخص بازار اول
LMS	۰.۰۰۰۰۲۵	۰.۰۱۵۸۴	۳.۳۵۶۶۶	۰.۰۱۱۵۲	۵۸.۵۵	۶۹.۷۹	۴۲.۵	
KLMS	۰.۰۰۰۰۳	۰.۰۰۵۵۱	۰.۴۲۹۸۵	۰.۰۰۳۵۳	۶۸.۳۹	۷۵.۳۴	۵۹.۵۹	شاخص صنعت
LMS	۰.۰۰۰۰۲۲	۰.۰۱۴۹۹	۳.۱۵۵۳۷	۰.۰۱۰۹۲	۵۷.۷۴	۶۹.۶۵	۴۱.۲۲	
KLMS	۰.۰۰۰۰۳	۰.۰۰۵۸۵	۰.۳۸۵۰۵	۰.۰۰۳۸	۶۵.۳۲	۷۱.۳۵	۶۰.۲۳	شاخص پنجاه شرکت برتر
LMS	۰.۰۰۰۰۴	۰.۰۱۹۹۶	۴.۴۱۸۱	۰.۰۱۴۲۶	۵۸.۲۳	۷۱.۰۶	۴۳.۶۳	
KLMS	۰.۰۰۰۰۳	۰.۰۰۵۴۳	۰.۴۶۳۲۶	۰.۰۰۳۶۵	۶۵	۷۲.۷	۵۵.۳۷	شاخص واسطه گرهای پولی و مالی
LMS	۰.۰۰۰۰۲۴	۰.۰۱۵۳۴	۲.۸۵۹۶۶	۰.۰۱۱۶۶	۵۳.۷۱	۶۶.۴۹	۳۵.۹۵	

به طور کلی، جهت پیش‌بینی نوسان شاخص‌های بورس اوراق بهادار در افق‌های زمانی کوتاه مدت (پیش‌بینی نوسان‌های روز بعد)، میان مدت (پیش‌بینی نوسان‌های ماه بعد) و بلندمدت (پیش‌بینی نوسان‌های روز بعد) با استفاده از الگوریتم‌های KLMS و شبکه‌های عصبی مبتنی بر LMS می‌توان نتیجه‌گیری نمود که عملکرد الگوریتم KLMS بر اساس معیار میانگین مربعات خطا (MSE)، مجذور میانگین مربعات خطا (RMSE)، میانگین مربعات خطای نرمالیزه شده (NMSE) و قدر مطلق میانگین خطا (MAD)، در تمامی شاخص‌ها کمتر از الگوریتم شبکه‌های عصبی مبتنی بر LMS است. این نشان دهنده این است که میزان انحراف پایین تری بین مقادیر واقعی و مقادیر پیش‌بینی شده با استفاده از الگوریتم پیشنهادی KLMS وجود دارد. هم چنین مقادیر صحت روند پیش‌بینی (DS)، صحت روند پیش‌بینی رو به بالا (CP) و صحت روند پیش‌بینی رو به پایین (CD) الگوریتم KLMS در افق‌های زمانی کوتاه مدت (پیش‌بینی نوسان‌های روز بعد)، میان مدت (پیش‌بینی نوسان‌های ماه بعد) و بلندمدت (پیش‌بینی نوسان‌های روز بعد) دارای مقدار بالاتری نسبت به الگوریتم شبکه‌های

عصبی مبتنی بر LMS می‌باشد. این نشان دهنده این است که معیار مناسبی از صحت روند پیش‌بینی شاخص‌ها ارائه می‌کند. بنابراین، می‌توان نتیجه گرفت که الگوریتم KLMS پیشنهادی نتیجه پیش‌بینی بهتری نسبت به الگوریتم شبکه‌های عصبی مبتنی بر LMS از لحاظ خطای پیش‌بینی و دقت پیش‌بینی فراهم می‌کند و به طور موثر می‌تواند نویز را در داده‌های سری زمانی مالی شناسایی و حذف نماید و عملکرد پیش‌بینی را بهبود دهد. از این رو می‌توان نتیجه‌گیری کرد که آن به میزان قابل توجهی بهتر از الگوریتم شبکه‌های عصبی مبتنی بر LMS در پیش‌بینی سری‌های زمانی مالی است. لذا می‌توان گفت که الگوریتم KLMS برآزش یافته با به حداقل رساندن خطای پیش‌بینی یک الگوریتم مناسب جهت پیش‌بینی نوسان شاخص‌های بورس اوراق بهادار تهران می‌باشد.

#### ۷. نتیجه‌گیری

پیش‌بینی نوسان شاخص‌های بورس اوراق بهادار برای سرمایه‌گذاران که به دنبال به حداکثر رساندن منافع خود هستند بسیار حائز اهمیت است. آن معمولاً بر تصمیم سرمایه‌گذاران برای سرمایه‌گذاری در بازار سهام و خرید یا فروش سهام اثرگذار است. پیش‌بینی موفق شاخص‌های سهام منجر به منافع مهمی برای سرمایه‌گذاران می‌گردد. بنابراین، جهت معرفی الگوریتم‌هایی به منظور پیش‌بینی نوسان شاخص‌های بورس تلاش‌های بسیار زیادی شده است. با توجه به اهمیت پیش‌بینی نوسان‌ها در بورس اوراق بهادار، در این مطالعه سعی شده است، الگوریتمی مناسب برای پیش‌بینی نوسان شاخص‌های منتشر شده در بورس اوراق بهادار تهران ارائه شود. به همین دلیل، نوسان شاخص‌های بورس اوراق بهادار تهران با استفاده از الگوریتم غیرخطی و آنلاین KLMS در افق‌های زمانی کوتاه مدت (پیش‌بینی نوسان‌های روز بعد)، میان مدت (پیش‌بینی نوسان‌های یک ماه بعد) و بلندمدت (پیش‌بینی نوسان‌های روز بعد) مورد بررسی و نتایج آن با الگوریتم شبکه‌های عصبی مبتنی بر LMS مقایسه گردید. مقایسه نتایج عملکرد پیش‌بینی الگوریتم‌های KLMS و شبکه‌های عصبی مبتنی بر LMS بر اساس معیارهای میانگین مربعات خطا، مجذور میانگین مربعات خطا، میانگین مربعات خطای نرمالیزه شده و قدر مطلق میانگین خطا، صحت روند پیش‌بینی، صحت روند پیش‌بینی رو به بالا و صحت روند پیش‌بینی رو به پایین نشان می‌دهد که پیش‌بینی انجام شده توسط الگوریتم

KLMS نسبت به پیش‌بینی انجام شده توسط الگوریتم شبکه‌های عصبی مبتنی بر LMS بسیار دقیق‌تر است. دلیل برتری پیش‌بینی انجام شده توسط الگوریتم KLMS نسبت به پیش‌بینی الگوریتم شبکه‌های عصبی مبتنی بر LMS این است که الگوریتم KLMS، الگوریتم غیرخطی می‌باشد در حالیکه الگوریتم شبکه‌های عصبی مبتنی بر LMS خطی می‌باشد زیرا داده‌های پیش‌بینی قیمت سهام ماهیت غیرخطی دارند. از طرف دیگر الگوریتم KLMS یک الگوریتم آنلاین می‌باشد. در سناریوهای آنلاین که در آن داده‌ها به طور مداوم در حال دریافت هستند، اندازه شبکه KLMS به طور مداوم رشد خواهد کرد و چالش‌های پیاده‌سازی نمایان می‌شود. بنابراین می‌توان گفت که الگوریتم KLMS برازش یافته با به حداقل رساندن خطای پیش‌بینی یک الگوریتم مناسب جهت پیش‌بینی نوسان شاخص‌های بازار سهام می‌باشد. به طور کلی نتایج نشان دهنده توان بالای پیش‌بینی کنندگی الگوریتم KLMS است.

## ۸. پیشنهادها برای کارهای تحقیقاتی آتی

با توجه به نتایج تحقیق

۱. الگوریتم حداقل میانگین مربعات مبتنی بر کرنل، روش دقیقی برای پیش‌بینی موضوعات مالی است. بنابراین پیشنهاد می‌شود در سایر موضوعات مانند پیش‌بینی جریان‌های نقدی آینده، مقدار و مبلغ فروش، قیمت سهام شرکت‌ها و غیره مورد استفاده قرار گیرد.
۲. پیشنهاد می‌گردد که از الگوریتم وفقی آنلاین حداقل مربعات برگشتی<sup>۸۳</sup> جهت پیش‌بینی نوسان شاخص‌های بازار سهام استفاده و نتایج پیش‌بینی آن با الگوریتم آنلاین KLMS و الگوریتم شبکه‌های عصبی مبتنی بر LMS که در این تحقیق جهت پیش‌بینی استفاده شده است مورد مقایسه قرار گیرد.

### یادداشت‌ها

- |                            |                                    |
|----------------------------|------------------------------------|
| 1. Srinivasan              | 8. Pu Yunma                        |
| 2. Haur                    | 9. Kernel Least Mean Square (KLMS) |
| 3. Montgomery              | 10. Least Mean Square (LMS)        |
| 4. Merh                    | 11. Liu & Morley                   |
| 5. Kara                    | 12. Leung                          |
| 6. Manish, K., & Thenmozhi | 13. Lo & MacKinlay                 |
| 7. Abu-Mostafa & Atiya     | 14. Yu                             |

15. Cheng
16. Van and Robert
17. Cao
18. Van Vaerenbergh
19. Modaghegh
20. Haykin
21. GARCH
22. Franses and van Dijk
23. TARCH
24. QGarch
25. Garliuskas
26. Dunis & Jalilov.
27. Egeli
28. Nygren
29. Lee
30. Armano
31. LU and et all
32. Independent component analysis
33. Support Vector Regression
34. Phichhang Ou & Wang
35. Hang Seng
36. Linear Discriminant Analysis (LDA)
37. Quadratic Discriminant Analysis (QDA)
38. Bayesian Classification
39. Logit Model
40. Tree Based Classification
41. Neural Network
42. Bayesian Classification with Gaussian Process
43. Support Vector Machine (SVM)
44. Least Squares Support Vector Machine (LS-SVM)
45. Generalized Autoregressive Conditional Heteroskedasticity
46. Random walk
47. Historical average
48. Khemchandani
49. Regularized Least Squares Fuzzy SVR (RLFSVR)
50. Phichhang Ou
51. Relevant Vector Machin(RVM)
52. Heesoo Hwang & Jinsung Oh
53. Fuzzy Model I
54. Okumus & Genc
55. Stochastic Volatility (SV) Model
56. Efficient Method of Moments (EMM)
57. P Srinivasan
58. Relative Simple GARCH
59. Relatively Complex GARCH
60. Symmetric GARCH model
61. Asymmetric GARCH model
62. Auto Regressive Integrated Moving Average
63. Auto Regressive Fractionally Integrated Moving Average
64. Artificial Neural Networks
65. Diebold-Mariano Test
66. Mean Square Error(MSE)
67. Root Mean Square Error(RMSE)
68. Normalized Mean Square Error (NMSE)
69. Mean Absolute Difference(MAD)
70. Mean Absolute Difference(MAD)
71. Correct Up Trend(CP)
72. Correct Down Trend(CD)
73. Widrow
74. Van Vaerenbergh
75. Giovanis
76. Liu & Príncipe
77. Ghafarian
78. Pokharel
79. Salkhordeh Haghighi
80. Pokharel
81. Liu
82. Hien
83. Recursive Least Square(RLS)



### پیوست

#### پیوست الف:

ساختار الگوریتم شبکه‌های عصبی مبتنی بر LMS در شکل ۱ نشان داده شده است. در این شکل  $\omega_{(k)}$  و  $\gamma_{(k)}$  به ترتیب سیگنال ورودی، بردار وزن و سیگنال خروجی سیستم در لحظه  $k$ ام است. که سیگنال خروجی به فرم برداری معادل  $\omega_{(k)}\chi_{(k)}$  است. در این الگوریتم یک سیستم ناشناخته  $\gamma_{(k)}$  شناسایی می‌شود. الگوریتم شبکه‌های عصبی مبتنی بر حداقل میانگین مربعات خطا (LMS) تلاش می‌کند برای انطباق فیلتر  $\gamma_{(k)}$  تا جایی که امکان دارد به  $d_{(k)}$  نزدیک باشد. هم چنین الگوریتم از  $d_{(k)}$  به عنوان خروجی مطلوب و  $e_{(k)}$  به عنوان خطای محاسبه شده استفاده می‌کند. الگوریتم حداقل میانگین مربعات خطا استفاده می‌کند الگوریتم کاهش شیب را برای به روز رسانی بردار وزن به طوری که بردار وزن همگرایی به راه حل بهینه وینر<sup>۱</sup> داشته باشد (مدقق و همکاران، ۲۰۱۰). به روز رسانی بردار وزن بر اساس قاعده زیر انجام می‌شود.

$$\omega_{(k+1)} = \omega_{(k)} + 2\mu \times e_{(k)} \times \chi_{(k)} \quad (1)$$

که در آن  $\omega_{(k)}$  بردار وزن،  $\chi_{(k)}$  بردار ورودی،  $e_{(k)}$  سیگنال خطا می‌باشد و  $\mu$  ضریب اندازه گام، پایداری و سرعت همگرایی LMS را تعیین می‌کند.

خطای پیش‌بینی از طریق مقایسه خروجی واقعی با خروجی مطلوب سیستم به شرح زیر

بدست می‌آید:

$$e_{(k)} = d_{(k)} - \gamma_{(k)} \quad (2)$$

و در آن  $d_{(k)}$  خروجی مطلوب سیستم و  $\gamma_{(k)}$  خروجی واقعی فیلتر در نمونه  $k$ ام می‌باشد

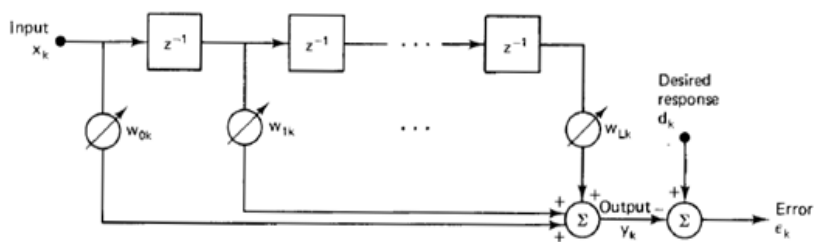
که محاسبه می‌شود به شرح:

$$\gamma_{(k)} = \omega_{(k)}\chi_{(k)} \quad (3)$$

که در آن

$$\chi_{(k)} = [\chi_{1(k)}, \chi_{2(k)}, \dots, \chi_{N(k)}]^T \quad \text{و} \quad \omega_{(k)} = [w_{1(k)}, w_{2(k)}, \dots, w_{N(k)}]^T$$

طول بافر داده است. اصلاح‌های پیاپی بردار وزنی در نهایت منجر به حداقل میانگین مجذورات خطا می‌شود.



شکل ۱: شکل یک فیلتر وقتی بر روی سیگنال ورودی

در عمل، برای رسیدن به نرخ همگرایی سریع و پایداری مناسب، ضریب اندازه گام  $\mu$  کوچکتر از یک می‌باشد. افزایش اندازه گام، سرعت الگوریتم و همچنین خطر ناپایداری را افزایش می‌دهد. در مقابل کاهش اندازه گام، موجب پایداری بیشتر الگوریتم می‌گردد. اگر  $\mu$  خیلی کوچک باشد، ممکن است پروسه خیلی طولانی شود و اگر خیلی بزرگ باشد، امکان ناپایداری وجود دارد. با توجه به اینکه LMS اساساً الگوریتمی مبتنی بر گرادیان تصادفی است، بنابراین  $\mu$  باید بگونه‌ای انتخاب شود که میانگین مربعات خطا حداقل بماند. پس اگر  $\mu$  در شرایط زیر صدق کند الگوریتم همگراست:

$$0 < \mu < \frac{1}{\lambda_{\max}} \quad (۴)$$

که  $\lambda_{\max}$  بزرگترین مقدار ویژه<sup>۱</sup> ماتریس کواریانس سیگنال ورودی است. با توجه به اینکه سرعت وقتی سازی و رفع نویز در بردار وزن،  $\mu$  باید به گونه ای تعیین شود که  $\lambda_{\max}$  از  $W^*$  (بردار وزن وینر) بیشتر نشود. پس در حالت کلی شرط لازم و کافی همگرایی الگوریتم LMS معادل:

$$0 < \mu < \frac{1}{tr[R]} \quad (۵)$$

است.

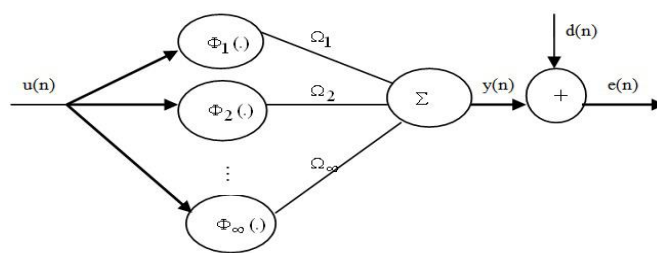
پیوست ب:

در این قسمت الگوریتم KLMS ارائه می‌شود. هدف اصلی، انجام الگوریتم شبکه‌های عصبی مبتنی بر LMS خطی در فضای ویژگی کرنل است. به همین دلیل فرض می‌شود که نگاشت‌های  $\Phi$  نقطه  $x(n)$  در فضای ورودی به  $\Phi(x(n))$  در فضای ویژگی کرنل است با  $\langle \Phi(x(n)), \Phi(x(m)) \rangle = \kappa(x(n), x(m))$  که نشان دهنده ضرب داخلی در فضای هیلبرت کرنل است. کرنل گوسی که در اینجا استفاده می‌شود به یک فضای هیلبرت بُعدی وابسته خواهد بود. زیرا این فضای ویژگی، خطی است،  $\Phi(x(n))$  می‌تواند یک بردار ستونی بُعدی را با ضرب‌های داخلی بردار معمول در نظر بگیرد.  $\Omega$  بردار وزن در این فضاست همان طور که  $\gamma(n) = \langle \Omega(n), \Phi(x(n)) \rangle$  خروجی است.  $\Omega$  در زمان  $n$  است.

$$\Omega_{(n+1)} = \Omega_{(n)} + 2\mu e(n) \Phi(u_{(n)}) \quad (1)$$

که در آن  $\Omega_{(n)}$  بردار وزن در فضای بُعدی،  $\mu$  ضریب اندازه گام است، که پایداری و سرعت همگرایی الگوریتم را کنترل می‌کند.  $\gamma(n)$  خروجی برآورد شده از طریق معادله (۲) محاسبه خواهد شد:

$$y(n) = \Omega_{(n)}, \Phi(u_{(n)}) \quad (2)$$



شکل ۲: دیاگرام یک سیستم کرنل برآوردی ساده

شکل ۲ نشان می‌دهد بردار ورودی  $x(n)$  را که تبدیل به بردار ویژگی بی‌نهایت  $\Phi(x(n))$  می‌شود، سپس مولفه‌ها به صورت خطی توسط بردار وزن بینهایت ترکیب می‌شوند. نوع برگشت ناپذیر معادله (۱) می‌تواند به شکل زیر نوشته شود:

$$\Omega_{(n)} = \Omega_{(0)} + 2\mu \sum_{i=0}^{n-1} e_{(i)} \Phi(u_{(i)}) \quad (3)$$

برای همگرایی انتخاب می‌شود  $\Omega_{(0)} = 0$  (بنابراین  $e_{(0)} = d_{(0)}$ ) و در نهایت  $\Omega_{(n)}$  می‌شود:

$$\Omega_{(n)} = 2\mu \sum_{i=0}^{n-1} e_{(i)} \Phi(u_{(i)}) \quad (۴)$$

بر مبنای معادله‌های (۲) و (۴):

$$\begin{aligned} \gamma_{(n)} &= \langle \Omega_{(n)}, \Phi(u_{(n)}) \rangle = \langle 2\mu \sum_{i=0}^{n-1} e_{(i)} \Phi(u_{(i)}), \Phi(u_{(n)}) \rangle \quad (۵) \\ &= 2\mu \sum_{i=0}^{n-1} e_{(i)} \langle \Phi(u_{(i)}), \Phi(u_{(n)}) \rangle \end{aligned}$$

می‌توان اکنون از حقه کرنل برای محاسبه  $\gamma_{(n)}$  از طریق معادله (۶) استفاده نمود:

$$\gamma_{(n)} = \mu \sum_{i=0}^{n-1} e_{(i)} k(u_{(i)}, u_{(n)}) \quad (۶)$$

معادله (۶) الگوریتم KLMS نامیده می‌شود.

در این تحقیق انواع مختلف کرنل‌های زیر جهت پیش‌بینی توسط الگوریتم KLMS مورد استفاده قرار گرفت که نتایج بهترین کرنل بعد از مقایسه ارائه شده است: کرنل گوسی (Gaussian kernel): برای  $\sigma > 0$  داریم:

$$k(x_{(n)}, x_{(n-1)}) = e^{-\frac{\|x_{(n)} - x_{(n-1)}\|^2}{2\sigma^2}}$$

کرنل خطی (Linear Kernel):

$$k(x_{(n)}, x_{(n-1)}) = (x_{(n)} \times x_{(n-1)})$$

کرنل چند جمله‌ای (Polynomial Kernel):

$$k(x_{(n)}, x_{(n-1)}) = (x_{(n)} \times x_{(n-1)} + 1)^\sigma$$

## منابع

### الف. فارسی

- آذر، عادل؛ افسر، امیر(۱۳۸۵). مدل سازی پیش‌بینی قیمت سهام با رویکرد شبکه‌های عصبی فازی، فصلنامه پژوهش‌های بازرگانی، ۴۰(پاییز)، ۳۳-۵۲.
- بت شکن، محمود (۱۳۸۰). پیش‌بینی قیمت سهام با استفاده از شبکه‌های عصبی فازی و مقایسه آن با الگوهای خطی پیش‌بینی، پایان نامه کارشناسی ارشد، دانشگاه تهران.
- حقیقت، حمید؛ معتمد، محمد (۱۳۹۰). بررسی رابطه بین نوسان پذیری و قابلیت پیش‌بینی سود، مجله پیشرفت‌های حسابداری دانشگاه شیراز، ۳(۲)، ۶۱/۳: ۶۵-۸۷.
- پاکدین امیری، علیرضا؛ پاکدین امیری، مجتبی؛ پاکدین امیری، مرتضی(۱۳۸۸). ارائه مدل پیش بینی شاخص کل قیمت سهام با رویکرد شبکه‌های عصبی (مطالعه موردی بورس اوراق بهادار تهران)، فصلنامه علمی - پژوهشی جستارهای اقتصادی، ۱۱ (بهار و تابستان)، ۸۳-۱۰۸.
- دموری، داریوش؛ فرید، داریوش؛ اشهر، مرتضی(۱۳۹۰). پیش‌بینی شاخص کل قیمت سهام با استفاده از الگوریتم پرواز پرندگان و مقایسه آن با الگوهای سنتی، مجله دانش حسابداری، ۲ (۵)، ۷-۳۰.
- فلاح شمس، میر فیض؛ دلنواز اصغری، بیتا (۱۳۸۸). پیش بینی شاخص بورس اوراق بهادار تهران با استفاده از شبکه‌های عصبی. فراسوی مدیریت، ۳ (۹)، (تابستان)، ۱۹۱-۲۱۲.
- فلاح شمس، میرفیض (۱۳۸۹). بررسی مقایسه ای کارایی مدل ریسک سنجی و مدل اقتصادسنجی گارچ در پیش بینی ریسک بازار در بورس اوراق بهادار تهران، مجله مهندسی مالی و مدیریت پرتفوی، ۵، ۱۳۷-۱۵۹.
- محمدی، تیمور؛ نصیری، سمیه(۱۳۸۹). مقایسه مدل‌های Riskmetric و GARCH در پیش‌بینی نوسان‌های شاخص بازده کل بورس اوراق بهادار تهران، مجله مطالعات مالی، ۶ (تابستان)، ۹۵-۱۱۸.
- مشیری، سعید؛ مروت، حبیب (۱۳۸۵). پیش بینی شاخص کل بازدهی سهام تهران با استفاده از مدل‌های خطی و غیرخطی، فصلنامه پژوهشنامه بازرگانی، ۴۱ (زمستان)، ۲۴۵-۲۷۵.

## ب. انگلیسی

- Abu-Mostafa, Y. S., & Atiya, A. F. (1996). Introduction to financial forecasting. *Applied intelligence*, 6(3), 205–213.
- Armano, G., Marchesi, A., & Muru, A. (2005). A hybrid genetic-neural architecture for stock index forecasting. *Information Sciences*, 170, 3-33.
- Cao, L. J., & Tay, F. E. H. (2001). Financial forecasting using support vector machines. *Neural Computing & Applications*, 10, 84–192.
- Cheng, W., Wanger, L. & Lin Ch. (1996). Forecasting the 30-year US Treasury bond with a system of neural networks, *Journal of Computational Intelligence in Finance*, 4, 10–6.
- Dunis, C. and Jalilov, J. (2001). Neural Network Regression and Alternative Forecasting Techniques for Predicting Financial Variables. Liverpool Business School, [www.cibef.com](http://www.cibef.com)
- Egeli, Birgale, et al. (2003). Stock Market Prediction Using Artificial Neural Networks. Web: [www.hicbusiness.org / BIZ 2003 proceedings](http://www.hicbusiness.org/BIZ2003proceedings).
- Ghafarian, S. H., Sadoghi Yazdi, H., Baradaran Kashani, H. (2010). Kernel least mean square features for HMM-based signal recognition. *International Journal of Computer Theory and Engineering*, 2(2), 1793-8201.
- Garliuskas, A. (1999). Neural network chaos and computational algorithms of forecast in finance. Proceedings of the IEEE SMC Conference, Man and Cybernetic cs2, 638-643.
- Franses, P. H. and van Dijk, D. (1996). Forecasting stock market volatility using (nonlinear) GARCH models, *Journal of Forecasting*, 15, 229-235.
- Giovanis, E. (2010). Applications of Least Mean Square (LMS) Algorithm Regression in Time-series Analysis. Available at: <http://ssrn.com/abstract=1667440>.
- Haur, Koh CHung (2004). Expert Systems for Forecasting Stock Market Indices. MSc Thesis. National University of Singapore.
- Haykin, S. (1995). Adaptive Filter Theory. 3rd Edition, Englewood Cliffs, NJ: Prentice Hall, 702-704.
- Heesoo Hwang and Jinsung Oh (2010). Fuzzy models for predicting time series stock price index. *International Journal of Control, Automation and Systems*, 8(3), 702-706.
- Hien, Mai Thi Thanh (2008). Modelling and Forecasting Volatility by Garch-Type Models, the case of Vietnam stock exchange. MA Thesis.
- Kara, Y., Acar Boyacioglu, M., & Kaan Baykan, O. (2011). Predicting direction of stock price index movement using artificial neural networks and support vector machines: The sample of the Istanbul Stock Exchange,

- Expert Systems with Applications*, 38, 5311–5319.
- Khemchandani, Reshma Jayadeva, & Chandra, Suresh (2009). Regularized least squares fuzzy support vector regression for financial time series forecasting. *Expert Systems with Applications*, 36, 132-138.
- Lee, W. C. (2005). Forecasting high-frequency financial data volatility via nonparametric algorithms-Evidence from Taiwan financial market. 4th International Conference on Computational Intelligence in Economics and Finance (CIEF). Salt Lake City, Utah, 1031-1034.
- Leung, Mark T., Daouk, H., & Chen, An-Sing (2000). Forecasting stock indices: A comparison of classification and level estimation models. *International Journal of Forecasting*, 16, 173-190.
- Liu, Weifeng & Príncipe, J. C. (2008a). Kernel affine projection algorithms. *EURASIP Journal on Advances in Signal Processing*, 1-12.
- Liu .W., Pokharel, P., & Principe, J. C. (2008b). The kernel least mean square algorithm. *IEEE Transactions on Signal Processing*, 543–554.
- Liu, Wei, & Morley, Bruce (2009). Volatility forecasting in the Hang Seng Index using the GARCH approach. *Asia-Pacific Finan Markets*, 4, 51–63.
- Lu, Chi-Jie, Lee, Tian-Shyug, Chiu, Chih-Chou (2009). Financial time series forecasting using independent component analysis and support vector regression, *Decision Support Systems*, 47, 115–125.
- Lo, A. W., & MacKinlay, A. C. (1988). Stock market prices do not follow random walks: Evidence from a simple specification test, *Review of Financial Studies*, 1, 41-66.
- Manish, K., & Thenmozhi, M. (2005). Forecasting Stock Index Movement: A Comparison of Support Vector Machines and Random Forest. In Proceedings of ninth Indian institute of capital markets conference, Mumbai, India. <<http://ssrn.com/abstract=876544>>.
- Merh, N., Saxena, V. P., & Pardasani, Kamal R. (2011). Next day stock market forecasting: An application of ANN and ARIMA. *The IUP Journal of Applied Finance*, 17 (1), 70-84.
- Modagheh, H., Sadoghi Yazdi, H., Javidi, M., & Pourreza, H. R. (2010). Learning of relevance feedback using a novel kernel based neural network. *Australian Journal of Basic and Applied Sciences*, 4(2), 171-186, SSN 1991-8178, 171–186.
- Montgomery, C. D., Lynwood, A. I., & Gardiner, J. S. (1990). *Forecasting Time Series Analysis*, Second Ed., NJ: Mc Grow-Hill.
- Nygren, Karl. (2004). *Stock Prediction – A Neural Network Approach*, Master Thesis, Royal Institute of Technology, KTH, Sweden.

- Okumus, H. Saduman & Guneren Genc, Elif (2011). Stock indices return forecast: Efficient method of moments estimation of a stochastic volatility model. *The Business Review*, 18 (1), 185-191.
- Phichhang Ou, & Wang, Hengshan (2009). Prediction of stock market index movement by ten data mining techniques. *Modern Applied Science*, 3 (12), 28-42.
- Phichhang, Ou (2010). Predict GARCH Based Volatility of Shanghai Composite Index by Recurrent Relevant Vector Machines and Recurrent Least Square Support Vector Machines, *Journal of Mathematics Research*, 2(2), 51-69.
- Pokharel. P., Liu, W., & Principe, J. C. (2007). Kernel lms. In Proceedings of International conference on acoustics, speech and signal processing.
- Pu Yun Ma (2006). A fresh engineering approach for the forecast of financial index volatility and hedging strategies. PhD Dissertation. Universite du Quebec.
- Salkhordeh Haghghi, M., Vahedian, A., Sadoghi Yazdi, H., & Modaghegh, H. (2010). Designing kernel scheme for classifiers fusion. *International Journal of Computer Science and Information Security*, 6 (2), 239-248.
- Srinivasan, A. (2003). Financial Prediction using Time Series. Master of Science. Florida Atlantic University.
- Srinivasan, P. (2011). Modeling and forecasting the stock market volatility of S&P 500 index using GARCH models. *The IUP Journal of Behavioral Finance*, VIII (1), 51-69.
- Van Vaerenbergh, S. (2009). Kernel Methods for Nonlinear Identification, Equalization and Separation of Signals. PhD Dissertation, Universidad de Cantabria.
- Van, E., & Robert, J. (1997). *The Application of Neural Networks in the Forecasting of Share Prices*, Haymarket, VA: Finance & Technology Publishing.
- Yu, Lean, Wang, Shouyang, & Lai, K. K. (2009). Intelligent computational methods for financial engineering. *Journal of Applied Mathematics and Decision Sciences*, Article ID 394731, doi:10.1155/2009/394731, 2 pages.