چشم‌انداز مدیریت مالی
شماره ۱۷ ـ بهار ۱۳۹۶
ص‌ص ۸۶ـ۶۹

پیش‌بینی روند قیمت سهام با استفاده از ماشین بردار پشتیبان تعدیل‌یافته همراه با انتخاب ویژگی هیبرید

سعید باجلان*، سعید فلاحپور، ناهید دانا*****

چکیده

در پژوهش حاضر، مدلی برای پیش‌بینی روند قیمت سهام برپایه‌ی ماشین بردار پشتیبان وزن‌دهی‌شده توسط حجم‌های روزانه معاملات، همراه با روش انتخاب ویژگی هیبرید F-SSFS ارائه می‌شود. به منظور ارزیابی دقت پیش‌بینی، مدل پیشنهادی با مدل ماشین بردار پشتیبان ساده همراه با انتخاب ویژگی هیبرد و نیز با روش‌های انتخاب ویژگی مرسوم از جمله بهره اطلاعات[۲]، عدم قطعیت متقارن[۳] و انتخاب ویژگی بر پایه همبستگی[۴]، از طریق انجام آزمون تی زوجی، مقایسه می‌شود؛ همچنین به‌عنوان مجموعه ویژگی‌های اولیه که در واقع ورودی ماشین بردار پشتیبان تعدیل‌یافته هستند، از شاخص‌های تحلیل تکنیکال و شاخص‌های آماری که برای ۱۰ سهم محاسبه شده‌اند، استفاده می‌شود. نتیجه این پژوهش نشان می‌دهد که عملکرد ماشین‌بردار پشتیبان وزن‌دهی‌شده، در مورد مسئله پیش‌بینی روند قیمت سهام، به میزان قابل‌توجهی بهتر از ماشین بردار پشتیبان ساده است. علاوه براین، نتایج عملیاتی نشان‌می‌دهد که ماشین بردار پشتیبان وزن‌دهی‌شده همراه با انتخاب ویژگی هیبرید پیشنهادی، بالاترین میزان دقت پیش‌بینی را نسبت به سه روش انتخاب ویژگی دیگر دارد. براساس نتایج این پژوهش می‌توان ادعا کرد مدل VW-SVM همراه با انتخاب ویژگی F-SSFS عملکرد بهتری در پیش‌بینی قیمت سهم، نسبت به روش‌های موجود دارد.

کلیدواژه‌ها: پیش‌بینی روند؛ ماشین بردار پشتیبان؛ انتخاب ویژگی؛ قیمت سهم.

تاریخ دریافت مقاله: ۱۳۹۵/۰۶/۲۳، تاریخ پذیرش مقاله: ۱۳۹۵/۱۲/۱۳

* استادیارگروه مالی و بیمه، دانشکده مدیریت دانشگاه تهران.
** استادیارگروه مالی و بیمه، دانشکده مدیریت دانشگاه تهران.
*** کارشناس ارشد مهندسی مالی، دانشگاه تهران (نویسنده مسئول).
E-mail: nahiddana@gmail.com

2. Information Gain
3. Symmetrical uncertainty
4. Correlation based feature selection

۱. مقدمه

پیش‌بینی قیمت سهام و حرکات و روند آن، همواره به‌عنوان یکی از چالش برانگیزترین حوزه‌های به‌کارگیری پیش‌بینی با استفاده از سری‌های زمانی مطرح بوده‌است. هرچند که تاکنون، پژوهش‌های داخلی و خارجی بسیاری در ارتباط با پیش‌بینی قیمت سهم و شاخص انجام‌شده‌است، پژوهش‌های انجام‌شده در حوزه پیش‌بینی جهت حرکت و روند قیمت و یا شاخص بسیار اندک است. به‌خصوص در ارتباط با بازار داخلی کشور، تعداد پژوهش‌هایی که به این موضوع پرداخته‌اند بسیار محدود است و نیاز به تحقیقات بیشتر در این زمینه وجود دارد. پیش‌بینی دقیق جهت تغییرات قیمت، در ارائه استراتژی معاملاتی سودمند، بسیار با اهمیت است [۱۸].

در میان فرآیندهای پیش‌بینی سری‌های زمانی، پیش‌بینی قیمت سهام به علّت اینکه بازار سهام اساسا پویا، غیرخطی، پیچیده، غیرپارامتریک و ماهیتا تصادفی است، جزء پیچیده‌ترین فرآیندها است [۱].

علاوه بر این، بازار سهام تحت تاثیر عوامل اقتصاد کلان بسیاری از جمله رخدادهای سیاسی، شرایط کلی اقتصادی، انتظارات سرمایه‌گذاران، انتخاب‌های سرمایه‌گذاران نهادی، حرکات سایر بازارهای سهام و روان‌شناسی سرمایه‌گذاران است [۲۱].

پژوهش‌های اخیر در زمینه پیش‌بینی روند تغییرات قیمت سهام حاکی از برتری روش‌های جدید داده‌کاوی و مدل‌های غیرخطی نسبت به روش‌های سنتی مانند لوجیت، پروبیت و گام تصادفی دارند. ماشین بردار پشتیبان که در ۱۹۹۵ توسط وپنیک پیشنهاد شد، یک تکنیک یادگیری ماشینی نسبتا جدید است که امکان به‌کارگیری در حوزه وسیعی از مسائل را دارد. این تکنیک توانایی رسیدن به جواب بهینه با حجم داده آموزش کم را نیز دارد.

طبق مطالعات صورت‌گرفته و تجارب فعالان بازار سهام، روندهای قیمتی در بازار سهام تحت تاثیر حجم معاملات سهم هستند و دخالت‌دادن این پارامتر در طبقه‌بندی‌کننده، به میزان چشمگیری دقت پیش‌بینی تغییرات روند قیمت را بهبودمی‌دهد [۲]. نتایج پژوهش‌ها نشان‌داده است که رابطه‌ای منفی بین حجم معاملات و نوسان‌پذیری وجود دارد و استفاده از متوسط حجم معاملات می‌تواند عملکرد پیش‌بینی نوسانات و در نتیجه بهینه‌سازی بازده را بهبود دهد [۱۰]. علاوه‌براین، پژوهش بدری و همکارانش (۱۳۹۵)، نشان داد که رابطه علّی یک‌طرفه‌ای از سمت متغیرهای کلان اقتصاد به سمت ارزش جاری بازارسهام و حجم معاملات بازار سهام وجود دارد [۲].

هرچند که طبق پژوهش‌های انجام‌شده، ماشین بردار پشتیبان عملکرد قابل‌قبولی در

پیش‌بینی و طبقه‌بندی دارد؛ اما دقت عملکرد آن به میزان قابل‌توجهی تحت تاثیر تعداد متغیرهای ویژگی ورودی آن است؛ بنابراین کاهش تعداد ویژگی‌هایی که باید در آموزش ماشین بردار پشتیبان به‌کار گرفته شود، تاثیر به‌سزایی در افزایش دقت نتایج و کاهش هزینه دارد [۱۹، ۲۵]؛ لذا یکی از بهبودهایی که در مطالعات جدید، با توجه به پیشرفت‌های کاربرد روش‌های هوش مصنوعی به‌کار گرفته شده است، استفاده از انتخاب ویژگی به‌عنوان یک پیش مرحله برای مدل طبقه‌بندی کننده اصلی می‌باشد.

این پژوهش به پیش‌بینی روندهای قیمتی بازار سهام، با استفاده از ماشین بردار پشتیبان وزن‌دهی‌شده با حجم معاملات همراه با روش انتخاب ویژگی هیبرید می‌پردازد. در ادامه، نتایج آن با نتایج ماشین بردار پشتیبان ساده همراه با این روش انتخاب ویژگی و سایر تکنیک‌های انتخاب ویژگی عنوان‌شده، مقایسه می‌شود.

۲. مبانی نظری و پیشینه پژوهش

انتخاب ویژگی. به‌طورکلی الگوریتم‌های انتخاب ویژگی به دودسته تقسیم می‌شوند: اول، روش‌های فیلترکننده [۷، ۲۳] و دوم، روش‌های پوششی [۱۵]. روش‌های فیلترکننده، احتیاج به هیچ بازخوردی از طبقه‌بندی‌کننده ندارند و این روش، عملکرد طبقه‌بندی را به وسیله یک‌سری روش‌های ارزیابی غیرمستقیم، مثل شاخص‌های فاصله‌ای که نشان‌دهنده میزان تفکیک‌شدگی صحیح طبقات از یکدیگر می‌باشند، تخمین می‌زنند [۱۳]؛ اما برخلاف این‌ها، روش‌های پوششی، وابسته به طبقه‌بندی‌کننده هستند و بسته به دقت آن، میزان «خوبی» زیرمجموعه انتخاب‌شده را مستقیما می‌سنجند. در مورد این دو روش باید گفت که نتایج بسیاری از بسیاری، برتری روش‌های پوششی را بر فیلترکننده، ثابت‌کرده‌اند [۱۶، ۲۰]؛ اما با وجود عملکرد خوب روش‌های پوششی، به علّت پیچیدگی محاسباتی بالایی که به همراه دارند، کاربردهایشان محدود است. در این پژوهش به منظور بهره‌گیری از مزایای هر دو روش فیلترکننده و پوششی، از یک روش انتخاب ویژگی هیبریدی به نام روش FSSFS[1] متشکل از انتخاب ویژگی فیشر به‌عنوان بخش فیلترکننده و روش جست‌وجوی ترتیبی پیشرو پشتیبانی‌شده به‌عنوان بخش پوششی، استفاده می‌شود.

روش فیشر. روش انتخاب ویژگی فیشر که در این پژوهش به‌عنوان بخشی از فرآیند انتخاب ویژگی هیبرید پیشنهاد شده است، یک روش فیلترکننده ساده امّا پایه‌ای است که میزان تمایز و

1. F-score-supported sequential forward search

تفاوت بین دو مجموعه داده را می‌سنجد [۵]. روش انتخاب ویژگی فیشر، برای هر ویژگی در مجموعه اوّلیه از ویژگی‌ها، با توجه به تساوی زیر، یک امتیاز محاسبه می‌کند. با داشتن بردار داده‌های آموزش به‌صورت x_k، k=1,2,...,m و n_+ و n_- به ترتیب، به‌عنوان تعداد نمونه‌های مثبت و تعداد نمونه‌های منفی امتیاز فیشر برای i امین ویژگی به صورت زیر محاسبه می‌شود:

$$F(i) \equiv \frac{\left(\bar{x}_i^{(+)} - \bar{x}_i\right)^2 + \left(\bar{x}_i^{(-)} - \bar{x}_i\right)^2}{\frac{1}{n_+ - 1}\sum_{k=1}^{n_+}\left(x_{k,i}^{(+)} - \bar{x}_i^{(+)}\right)^2 + \frac{1}{n_- - 1}\sum_{k=1}^{n_-}\left(x_{k,i}^{(-)} - \bar{x}_i^{(-)}\right)^2}$$

که در این عبارت، $\bar{x}_i$، $\bar{x}_i^{(+)}$ و $\bar{x}_i^{(-)}$ به ترتیب میانگین کل و میانگین ویژگی‌های مثبت و منفی می‌باشد و $x_{k,i}^{(+)}$ و $x_{k,i}^{(-)}$ویژگی iام از k امین ویژگی به ترتیب مثبت و منفی است. صورت کسر در تساوی بالا، تمایز بین مجموعه‌های مثبت و منفی را نشان می‌دهد و مخرج کسر تفاوت‌ها را در داخل هر یک از دو مجموعه مثبت و منفی اندازه‌گیری می‌کند. هرچه میزان امتیاز فیشر محاسبه‌شده برای هر ویژگی بیشتر باشد، احتمال این‌که آن ویژگی بتواند کمک بیشتری به فرآیند طبقه‌بندی بکند، بیشتر است.

روش انتخاب ویژگی جست‌وجوی ترتیبی پیشرو پشیبانی‌شده. جست‌وجوی ترتیبی پیشرو پشتیبانی‌شده، همان جست‌وجوی ترتیبی پیشرو که نوعی از روش‌های انتخاب ویژگی پوششی است، می‌باشد که به صورت یک‌پارچه با ماشین بردار پشتیبان وزن‌دهی‌شده در آمده است. در این بخش به توضیح مختصر این فرآیند انتخاب ویژگی پرداخته می‌شود.

ابتدا ماشین بردار پشتیبان وزن‌دهی‌شده با حجم معاملات را به تعداد k بار برای هر ویژگی f_i آموزش داده، تا بهترین ویژگی انتخاب شود. از لحاظ ریاضیاتی به‌عنوان مجموعه داده‌های آموزش داریم،

$$F_1^i = f_i \, , f_i \in F$$

و

$$V_1^i = \{1,2, \dots , N\}$$

از آنجاکه بردار ورودی یک بعدی است، حجم و پیچیدگی محاسباتی بسیار اندک است. پس از انجام این آموزش، ویژگی‌ای که طبقه‌کننده به ازای آن بهترین عملکرد را داشته‌باشد، به‌عنوان «اولین، بهترین» ویژگی انتخاب می‌شود؛ بنابراین پس از اولین مرحله، برای زیر مجموعه اوّلیه از ویژگی‌ها و مجموعه آموزش مرتبط با آن خواهیم داشت $F_1 = \{f_i\}$ و $V_1 = \{v_j\}$. برای

اضافه‌کردن ویژگی به این زیرمجموعه اوّلیه ویژگی، به‌صورت زیر عمل می‌شود:

$$F_{n+1}^{i} = F_n \cup \{f_i\}\, for\, f_i \in F_n^{av}$$
$$V_{n+1}^{i} = V_n \cup \{v_i\}$$

و برای هریک از ویژگی‌های باقی‌مانده این عملیات یک بار تکرار می‌شود و سپس ویژگی‌ای که به ازای اضاف‌شدن آن به مجموعه قبلی بهترین عملکرد برای طبقه‌بندی‌کننده حاصل‌شود، به‌عنوان ویژگی جدید به زیرمجموعه ویژگی‌ها اضافه می‌شود و این فرآیند تا زمانی ادامه می‌یابد که هیچ بهبود قابل‌توجهی در عملکرد مدل حاصل نشود.

سایر روش‌های انتخاب ویژگی مرتبط. در این پژوهش از سه روش انتخاب ویژگی برای مقایسه با روش انتخاب ویژگی هیبرید پیشنهادی استفاده می‌شود. این سه روش که روش‌های آماری برای انتخاب ویژگی هستند، به شرح زیر می‌باشند [۱۳]:

بهره اطلاعات: این روش، اهمیت هر ویژگی را با استفاده از اندازه‌گیری اینکه هر ویژگی، به چه میزان توان تقسیم‌بندی داده‌ها به طبقات هدف را دارد، اندازه‌گیری می‌کند. بهره اطلاعات به‌صورت زیر محاسبه می‌شود:

$$InfoGain = H(Y) - H(Y|X)$$

که در آن X و Y ویژگی‌ها هستند و داریم:

$$H(Y) = -\sum_{y \epsilon Y} p(y) log_2\big(p(y)\big)$$
$$H(Y) = -\sum_{x \epsilon X} p(y) \sum_{y \epsilon Y} p(y|x) log_2\big(p(y|x)\big)$$

عدم قطعیت متقارن: از این تابع برای سنجش میزان توانایی تفکیک‌کنندگی هر ویژگی استفاده می‌شود. ویژگی با بیشترین مقدار تابع، دارای بیشترین توان تفکیک‌کنندگی است.

$$SU = 2.0 * \frac{InfoGain}{H(Y) + H(X)}$$

انتخاب ویژگی بر پایه همبستگی: این روش، یک زیرمجموعه از ویژگی‌ها را با در نظرگرفتن توانای پیش‌بینی فردی آن ویژگی همراه با درجات افزونگی میان آن‌ها، ارزیابی می‌کند:

$$CFS_s = \frac{k\,\bar{r}_{cf}}{\sqrt{k + k(k-1)\bar{r}_{ff}}}$$

که در آن CFS_s میزان امتیاز هر زیرمجموعه ویژگی، S، شامل k ویژگی می‌باشد و $\bar{r}_{cf}$ میانگین همبستگی ویژگی کلاس ($f\epsilon S$) و $\bar{r}_{ff}$ میانگبن همبستگی متقابل ویژگی- ویژگی است. تفاوت بین الگوریتم‌های فیلتری معمول و CFS این است که الگوریتم‌های معمولی برای هر ویژگی یک امتیاز مجزا محاسبه می‌کنند؛ در حالی‌که CFS یک «مریت[1]» هیوریستیک برای هر زیرمجموعه محاسبه می‌کند و مشخص می‌کند که کدام زیرمجموعه از همه بهتر است.

ماشین بردار پشتیبان وزن‌دهی‌شده. در ارتباط با ماشین بردار پشتیبان در منابع داخلی و خارجی بسیاری از[۸، ۱۳] و بسیاری مقالات و منابع دیگری توضیحات لازم ارائه شده است که در اینجا به منظور جلوگیری از اطناب تنها به توضیح مدل ماشین بردار پشتیبان وزن‌دهی‌شده پرداخته می‌شود.

هدف اصلی الگوریتم ماشین بردار پشتیبان، یافتن ابرصفحه جداکننده‌ای است که با حل مسئله بهینه‌سازی زیر، حاشیه‌ی $||w||$ را بیشینه کند. در این راستا از اصول کمینه‌سازی ریسک که [۶] مورد بحث قرار گرفته‌است، استفاده می‌کند.

$$\min \frac{1}{2||w||^2} + \sum_{i=1}^{m} C_i \xi_i,$$

با محدودیت:

$$y_i(wx_i - b) \geq 1 - \xi_i$$
$$\xi_i \geq 0$$

که در آن ξ_i متغیر اسلک غیر صفر است و C یک پارامتر جریمه برای هر ورودی مشخص x_i و y_i مقدار هدف مربوط به آن است. در ماشین بردار پشتیبان تعدیل‌شده با وزن‌دار کردن تابع جریمه

1. Merit

توسط حجم معاملات که در این پژوهش از آن استفاده می‌کنیم، عبارت جریمه به‌صورت زیر تعریف می‌شود:

$$C_i = V_i C$$

که در آن داریم، $V_i = \frac{V_t}{\sum_{i=0}^{d} W_{t-i}}$، که در آن V_t نشان‌دهنده حجم معامله شده از سهم در روز t ام و d طول پنجره زمانی معاملاتی است. با استفاده از ضرایب لاگرانژ α_i و μ_i می‌توان مسئله بهینه‌سازی به‌صورت زیر بازنویسی می‌شود:

$$\max L(w, b, \alpha) = \frac{1}{2||w||^2} + C\sum_{i=1}^{m} \xi_i - \sum_{i=1}^{m} \alpha_i (y_i(wx_i - b) - 1 + \xi_i) - \sum_{i=1}^{m} \mu_i \xi_i$$

با محاسبه مشتقات جزئی، می‌توان دوگان این مسئله را به‌صورت زیر نوشت:

$$Q(\alpha) = \sum_{i=1}^{m} \alpha_i - 1/2 \sum_{i=1}^{m}\sum_{j=1}^{m} \alpha_i \alpha_j y_i y_j \varphi(x_i)\varphi(x_j)$$

که این مسئله باید با درنظرگرفتن محدودیت‌های زیر بیشینه‌شود:

$$\sum_{i=1}^{m} \alpha_i y_i = 0$$
$$\alpha_i \geq 0$$

پیشینه پژوهش. با وجود اینکه مدل‌های آماری توانسته‌اند پیش‌بینی‌های نسبتا خوبی را در مورد قیمت ارائه‌دهند؛ اما مفروضات محدودکننده‌ای که در برخی از این مدل‌ها وجود دارد، براثر بخشی آن‌ها موثر بوده و در نتیجه، به تدریج روش‌های دیگری برای مقابله با این محدودیت‌ها و بهبود عملکرد پیش‌بینی‌ها ارائه‌شد. با توجه به نتایج پژوهش‌های داخلی و خارجی بسیاری، برتری روش‌های ماشین بردار پشتیبان بر روش‌هایی از جمله روش شبکه‌های عصبی مصنوعی و نیز روش‌های آماری ثابت شده‌است [۸، ۱۹]؛ بنابراین در ادامه به ذکر پژوهش‌هایی که بیشترین ارتباط را با موضوع مورد بررسی دارند پرداخته می‌شود.

بریو و باتر به منظور مقایسه عملکرد مدل‌های پیش‌بینی‌کننده ماشین بردار پشتیبان و شبکه عصبی انتشار پسرو، اقدام به پیش‌بینی شاخص نزدک با استفاده از انتخاب ویژگی با روش آنالیز مولفه اصلی و آنالیز عاملی[1]، بر روی شاخص‌های تکنیکال کردند که نتایج نشان‌دهنده برتری ماشین بردار پشتیبان بود.

استفاده از ترکیب ماشین بردار پشتیبان و الگوریتم ژنتیک به منظور پیش‌بینی روند حرکت قیمت سهام نشان داد که دقت مدل ترکیبی از مدل ماشین بردار پشتیبان به تنهایی بیشتر است [۶].

هوانگ و همکارانش (۲۰۱۱)، اقدام به پیش‌بینی قیمت سهم با استفاده از روش انتخاب ویژگی پوششی همراه با یک سیستم مرکب از چندین طبقه‌بندی‌کننده، شامل ماشین بردار پشتیبان و شبکه عصبی کردند. نتایج آن‌ها نشان داد که دقت پیش‌بینی‌ها با استفاده از روش انتخاب ویژگی پوششی نسبت به روش‌های فیلترکننده مختلف بالاتر است [۱۳].

مینگ (۲۰۰۹)، با هدف پیش‌بینی روند تغییرات قیمت شاخص نزدک، اقدام به انجام پژوهشی با ترکیب مدل ماشین بردار و انتخاب ویژگی هیبرید حاصل از ترکیب انتخاب ویژگی فیشر و جست‌وجوی ترتیبی پیشروی پشتیبانی‌شده، نمود. به‌عنوان ویژگی‌های ورودی مدل وی از ۳۰ سهم موجود در این شاخص استفاده نمود، نتایج پژوهش مینگ نشان داد که استفاده از ماشین بردار پشتیبان همراه با انتخاب ویژگی هیبرید دقت بیشتری در پیش‌بینی روند نسبت به شبکه عصبی انتشار پسرو و نیز استفاده از سایر روش‌های انتخاب ویژگی دارد [۱۷].

در پژوهش دیگری که منظور ارزیابی و مقایسه روش انتخاب ویژگی فرکتال[2] با سایر روش‌ها، از جمله رلیف[3]، سی‌اف‌اس[4] و چند روش‌دیگر، با استفاده از طبقه‌بندی‌کننده ماشین بردار پشتیبان توسط لیونگ انجام‌شد، برتری روش فرکتال بر سایر روش‌های فلتری انتخاب ویژگی اثبات‌شد [۱۸].

در پژوهشی که توسط زیبکوسکی (۲۰۱۴) انجام شد، از ماشین بردار پشتیبان وزن‌دهی‌شده با حجم معاملات واقعی روزانه، همراه با انتخاب ویژگی فیشر، به منظور پیش‌بینی روند استفاده شد. وی به‌عنوان ورودی از ۷ شاخص تحلیل تکنیکال به‌منظور پیش‌بینی روند قیمت سهام استفاده نمود. وی با نشان‌دادن برتری عملکرد مدل پیشنهادی خود، برتری این مدل را بر مدل ماشین بردار پشتیبان ساده و ماشین بردار پشتیبان بدون انتخاب ویژگی را اثبات نمود [۲۴].

1.Factor analysis
1. Fractal
2. Relief
3. CFS

در ایران در حوزه مطالعات پیش‌بینی روند، می‌توان به پژوهش فلاح‌پور، گل‌ارضی و فتوره‌چیان (۱۳۹۲)، اشاره نمود. آن‌ها اقدام به پیش‌بینی روند حرکت و تغییرات قیمت سهام در بازار بورس اوراق بهادار تهران نمودند. آن‌ها متغیرهای ورودی ماشین بردار پشتیبان را توسط الگوریتم ژنتیک، بهینه‌سازی نمودند. نتایج پژوهش آن‌ها نشان داد که ماشین بردار پشتیبان برپایه الگوریتم ژنتیک، دقت بسیار بیشتری در پیش‌بینی نسبت به ماشین بردار پشتیبان ساده دارد [۱۱].

در پژوهشی که توسط منجمی، ابرازی و رعیتی (۱۳۸۸)، انجام‌شد، قیمت سهام در بازار بورس اوراق بهادار، توسط شبکه عصبی فازی و الگوریتم ژنتیک، پیش‌بینی شد. نتایج پژوهش آن‌ها نشان داد که مدل ترکیبی پیشنهادی هم از لحاظ دقت و هم از لحاظ سرعت عملکرد بهتری دارد [۲۰].

در پژوهشی که توسط عبادی (۱۳۸۸)، برای پیش‌بینی شاخص قیمت سهام بازار بورس تهران، با استفاده از شبکه‌های عصبی مصنوعی انجام‌شد، مشخص‌شد که شبکه عصبی عملکرد خوبی در پیش‌بینی این شاخص دارد [۹]. همچنین در پژوهش دیگری که به‌منظور مطالعه تاثیر عوامل رفتاری بر قیمت سهام توسط هاشمی (۱۳۸۹) بر روی سهام ۱۰ شرکت از شرکت‌های شاخص داوجونز انجام‌شد، مشخص‌شد که عوامل رفتاری در قیمت ۹ سهم از ۱۰ سهم موثر بوده و مدنظر قرار دادن آن‌ها منجر به افزایش دقت پیش‌بینی می‌شود [۱۴].

۳. روش‌شناسی پژوهش

در مجموع مراحل کار در این پژوهش را از انتخاب ویژگی با روش پیشنهادی تا طبقه‌بندی با استفاده از مدل ماشین بردار پشتیبان وزن‌دهی‌شده، می‌توان به‌صورت زیر جمع‌بندی و خلاصه کرد. ابتدار مقدار امتیاز فیشر برای هر ویژگی f_i محاسبه می‌شود.

سپس، امتیازهای محاسبه‌شده را به ترتیب نزولی مرتب می‌کنیم و تعداد k ویژگی اول را که k عبارت است از $K = [i * \frac{|F|}{4}]$ برای $\{1,2, \dots, m\} \in$ که در این‌جا، |F| تعداد کل ویژگی‌ها می‌باشد و m بزرگترین عدد صحیحی می‌باشد که در رابطه‌ی $|F| /2^m \geq 1$ صدق کند.

در مرحله بعد، ویژگی با بیشترین امتیاز را برای ساخت F_n^{av} انتخاب می‌کنیم.

پس از انتخاب اولین ویژگی، ویژگی بعدی را که باید به F_n^{av} اضافه کنیم، با توجه به تساوی $F_{n+1}^i = F_n \cup \{f_i\}$ ، برای $f_i \in F_n^{av}$ و $V_{n+1}^i = V_n \cup \{v_i\}$ ، همان‌طور که در زیربخش توضیح روش انتخاب ویژگی جست‌وجوی ترتیبی پیشرو پشیبانی‌شده، توضیح دادیم انتخاب می‌کنیم.

پس از آن مجموعه داده‌های آموزش را با استفاده از $F_{n+1} = F_{n+1}^{j}$ و $V_{n+1} = SV_{n+1}^{j}$ بروزرسانی می‌کنیم و متوسط دقت طبقه‌بندی را محاسبه می‌کنیم.

این مراحل را تا جایی تکرار می‌کنیم که هیچ افزایش قابل‌توجهی در دقت طبقه‌بندی ایجاد نشود.

پس از انتخاب زیرمجموعه بهینه از ویژگی‌ها، مدل ماشین بردار پشتیبان وزن‌دهی‌شده با حجم معاملات را با استفاده از زیرمجموعه انتخاب‌شده، آموزش می‌دهیم، تا طبقه‌بندی‌کننده ماشین بردار پشتیبان وزن‌دهی‌شده را به‌دست آوریم؛ سپس دقت طبقه‌بندی را برآورد می‌کنیم.

درجدول ۱، به‌صورت خلاصه، شاخص‌های تحلیل تکنیکالی که به‌عنوان بردارهای ورودی ماشین بردار پشتیبان، استفاده می‌کنیم، ارائه‌شده‌است. لازم به ذکر است که در عباراتی که در ادامه می‌آید، P_t^o، P_t^H، P_t^L و P_t^c به ترتیب نشان‌دهنده مقادیر مربوط به قیمت شروع، بالاترین قیمت، پایین‌ترین قیمت و قیمت پایان روز می‌باشند.

جدول ۱. شاخص‌های تحلیل تکنیکال به‌عنوان ویژگی‌های ورودی فرآیند پیشنهادی

نام شاخص	نحوه محاسبه	توضیحات
حجم جاری معاملات OBV	$OBV_t = OBV_{t-1} + V_t$ **آنگاه**، $P_t^C > P_{t-1}^C$ **اگر** $OBV_t = OBV_{t-1} - V_t$ **آنگاه**، $P_t^C < P_{t-1}^C$ **اگر** $OBV_t = OBV_{t-1}$ **آنگاه**، $P_t^C = P_{t-1}^C$ **اگر**	حجم جاری معاملات، یک شاخص آنی است که میزان افزایش و کاهش حجم معاملات را اندازه‌گیری می‌کند و روندهای حجم و قیمت را به یکدیگر مرتبط می‌سازد
دامنه متوسط واقعی ATR	$TR_t = \max\{P_t^H - P_t^L, P_t^H - P_t^{c-1}, P_t^L - P_t^{c-1}\}$ **و** $ATR_t^n = \frac{1}{n\sum_{i=0}^{n} TR_{t-i}}$	شاخصی است که میزان نوسانات را برای دوره مورد نظر محاسبه می‌کند و براساس یک شاخص کمکی به نام دامنه واقعی (TR) به دست می‌آید
شاخص قدرت نسبی RSI	$RSI_t^n = 100 - \frac{100}{1 - RS_t^n}$ $RS_t^n = \frac{\sum_{t=0}^{n} U_t}{\sum_{t=0}^{n} D_t}$ **که در آن:** $U_t = P_t^C - P_{t-1}^C$ **آنگاه**، $P_t^C > P_{t-1}^C$ **اگر** $D_t = P_{t-1}^C - P_t^C$ **آنگاه**، $P_t^C > P_{t-1}^C$ **اگر**.	معامله‌گران از شاخص قدرت نسبی، به منظور تعیین وضعیت‌های اشباع خرید و یا فروش استفاده می‌کنند

نوسان‌گر R% ویلیام	$\%R = 100 * \frac{\max P_t^n - P_t}{maxP_t^n - minP_t^n}$ که در آن داریم، $\text{Max}\, P_t^n = \max\{P_t, P_{t-1}, \dots, P_{t-n}\}$ $\text{Min}\, P_t^n = \max\{P_t, P_{t-1}, \dots, P_{t-n}\}$	از این شاخص برای تعیین میزان قدرت و پایداری روند فعلی و از آن مهم‌تر برای زمانی‌که روند آماده تغییر جهت استفاده می‌شود.
شاخص ورتکس	$V_t^{n(+)} = \frac{\sum_{k=0}^{d}(P_{t-k}^H - P_{t-k}^L)}{\sum_{k=0}^{d} TR_{t-k}}$ $V_t^{n(+)} = \frac{\sum_{k=0}^{d}(P_{t-k}^L - P_{t-k}^H)}{\sum_{k=0}^{d} TR_{t-k}}$	شاخص ورتکس نشان‌دهنده تغییر جهت‌ها است. ایده اصلی آن است که اعلانات متوالی از بازار می‌تواند حاوی اطلاعاتی در مورد جهت روند باشد
نرخ بازده و انحراف استاندارد	$\mathbf{R_t^n} = \ln \frac{P_t^C}{P_{t-n}^C}$ و $\sigma_{\mathrm{t}}^{\mathrm{n}} = \sqrt{\frac{\sum_{\mathrm{i=1}}^{\mathrm{n}}(P_{t-i}^C - \bar{P})^2}{\mathrm{n}}}$	به‌طور ساده، عایدی یک سرمایه‌گذاری، نرخ بازده n روزه $\mathrm{R_t^n}$ است و ریسک مرتبط با آن انحراف استاندارد این بازده‌ها در طول یک دوره n روزه است

پس از محاسبه و به‌دست آوردن مقادیر مربوط به هر یک از بردارهای ورودی، با توجه به اینکه مقادیر مربوط به دامنه تغییرات هریک از شاخص‌ها متفاوت است، برای افزایش دقت پیش‌بینی، احتیاج به نرمال‌سازی داده‌ها داریم، نرمال‌سازی داده‌ها دقت طبقه‌بندی‌کننده SVM را افزایش می‌دهد [۲۲]. نرمال‌سازی را به‌صورت زیر انجام می‌دهیم [۲۲]:

$$x'_i = L + (U - L)\frac{x_i - m_i}{M_i - m_i}$$

که در آن $L,U \in R$ و $L<U$ و M_i و m_i، به ترتیب نشان‌دهنده مقادیر بیشینه و کمینه هر ویژگی می‌باشد و به‌صورت زیر تعریف می‌شود:

$$M^f = \max\{x_i\}$$
$$m_i = \min\{x_i\}$$

پژوهشگاه علوم انسانی و مطالعات فرهنگی
پرتال جامع علوم انسانی

فرضیه پژوهش. فرضیه پژوهش حاضر، به‌صورت زیر بیان می‌شود:

دقت پیش‌بینی قیمت سهام، با استفاده از روش انتخاب ویژگی هیبرید، همراه با ماشین بردار پشتیبان وزن‌دهی‌شده با حجم روزانه معاملات از سایر روش‌های رقیب بیشتر می‌باشد.

٤. یافته‌های پژوهش

در ادامه نتایج پژوهش در دو بخش که اولی مربوط به انتخاب ویژگی با روش پیشنهادی و سپس مقایسه دقت پیش‌بینی روش پیشنهادی با روش‌های رقیب می‌باشد، آورده شده است. داده‌های مورد استفاده در این پژوهش، مربوط به ۱۰ سهم انتخاب‌شده، از بین سهام شرکت‌های پذیرفته‌شده در بورس اوراق بهادار تهران، می‌باشد. داده‌های روزانه هر سهم از بازه ۱۳۸۴/۰۸/۰۱ تا ۱۳۹۴/۰۸/۳۰جمع‌آوری شده است؛ همچنین سهم‌ها از بین سهام شرکت‌های بورسی منهای بانک‌ها و سرمایه‌گذاری‌ها انتخاب‌شده‌اند. از داده‌های روزانه ۸ سال نخست برای آموزش شبکه و از داده‌های ۲ سال آخر برای تست استفاده شده است.

نتایج مربوط به دقت انتخاب ویژگی با استفاده از الگوریتم هیبرید فیشر و جست‌وجوی ترتیبی پیش‌رو بر پایه ماشین بردار پشتیبان. همان‌طور که در بخش ۳ در مورد فرآیند انتخاب ویژگی توضیح داده‌شد، k به‌عنوان مقدار آستانه ویژگی‌های انتخابی، مقادیر ۴،۲ و ۶ که حالت حدی‌ای است که در آن انتخاب ویژگی هیچ تاثیری بر مجموعه ویژگی‌های اولیه ندارد، اتخاذ می‌کند. در جدول ۲، متوسط دقت پیش‌بینی‌کننده، برای هر یک از مقادیر k آورده شده است.

جدول ۲. مقایسه دقت نتایج برای مقادیر مختلف K

تعداد ویژگی‌ها	دقت طبقه‌بندی‌کننده	
	آموزش (٪)	تست (%)
$k=\left\lfloor\frac{\lvert F\rvert}{4}\right\rfloor=۲$	٪ ۷۴/۵۰	٪ ۷۳/۵۲
$k=\left\lfloor\frac{2*\lvert F\rvert}{4}\right\rfloor=۴$	٪ ۷۳/۸۰	٪ ۷۱/۵۹
K=|F|=۶	٪ ۶۶/۲۰	٪ ۶۹/۹۱

این جدول نشان می‌دهد که با افزایش تعداد ویژگی‌ها، دقت طبقه‌بندی کننده کاهش می‌یابد. همانطور که در جدول ۳ خواهیم دید امتیاز فیشر برای شاخص‌های آماری بازده و انحراف استاندارد کمترین مقدار رادارد، که این مساله با این واقعیت که بازده و انحراف استاندارد توانایی چندانی در پیش‌بینی قیمت ندارند، هم‌خوانی دارد. همچنین یادآوری می‌شود که، مدل ماشین بردارپشتیبان به کاررفته، با استفاده از حجم‌های روزانه‌معاملات تعدیل‌یافته‌است، در نتیجه انتظار می‌رود که شاخص حجم جاری معاملات توانایی خوبی در پیش‌بینی قیمت داشته‌باشد، که این مورد هم با توجه به امتیاز فیشر به دست‌آمده برای این شاخص، تایید می‌شود.

جدول۳. نتایج امتیازات انتخاب ویژگی فیشر و دقت متوسط طبقه‌بندی با اضافه‌شدن هر ویژگی

ردیف	ویژگی	امتیاز فیشر	دقت متوسط (٪)
۱	OBV	۰/۰۸۲۷	٪۷۳/۵۰
۲	ATR	۰/۰۰۵۴	٪۷۳/۵۴
۳	%R	۰/۰۰۲۷	٪۷۰/۸۲
۴	RSI	۰/۰۰۰۴	٪۶۹/۵۰
۵	R	۰/۰۰۰۳	٪۶۸/۹۰
۶	σ	۰/۰۰۰۱	٪۶۹/۲۰

طبق نتایج جدول ۱، انتخاب K= ۲، منجر به بیشترین دقت می‌شود و علاوه بر این، پس از انجام فرآیند جست‌وجوی ترتیبی پیشرو پشتیبانی‌شده، دیده می‌شود که استفاده از دو ویژگی حجم جاری معاملات و دامنه متوسط واقعی منجر به بهترین نتایج می‌شود. نتایج امتیازات فیشر برای هر ویژگی و متوسط دقت پیش‌بینی با اضافه‌شدن هر ویژگی طبق فرآیند جست‌وجوی ترتیبی پیشروی پشتیبانی‌شده، به زیرمجموعه ویژگی‌ها در جدول ۲ آورده شده است.

مقایسه دقت پیش‌بینی با استفاده از مدل پیشنهادی و روش‌های رقیب. مقایسه دقت متوسط، بین طبقه‌بندی‌کننده‌های ماشین بردار پشتیبان وزن‌دهی‌شده با حجم معاملات و ماشین بردار پشتیبان ساده، همراه با ۴ حالت مختلف برای انتخاب ویژگی، در پیش‌بینی قیمت سهام، در جداول ۴ و ۵ آورده شده است. همان‌طور که مشاهده می‌شود، برای تمام روش‌های انتخاب ویژگی، ماشین بردارپشتیبان وزن‌دهی‌شده با حجم معاملات عملکرد بهتری نسبت به ماشین بردار پشتیبان ساده دارد.

جدول ۴. عملکرد ماشین بردار پشتیبان وزن‌دهی‌شده با ۴حالت مختلف برای انتخاب ویژگی

طبقه‌بندی‌کننده+ روش انتخاب ویژگی	متوسط دقت پیش‌بینی
ماشین بردار پشتیبان وزن‌دهی‌شده	٪۶۹/۹
ماشین بردار پشتیبان وزن‌دهی‌شده+ بهره اطلاعات	٪۷۱/۴
ماشین بردار پشتیبان وزن‌دهی‌شده+عدم قطعیت متقارن	٪۷۰/۶۸
ماشین بردار پشتیبان وزن‌دهی‌شده+ انتخاب ویژگی برپایه همبستگی	٪۷۱/۳۱
ماشین بردار پشتیبان وزن‌دهی‌شده+ انتخاب ویژگی هیبرید	٪۷۳/۵

جدول ۵. عملکرد ماشین بردار پشتیبان وزن‌دهی‌شده با ۴ حالت مختلف برای انتخاب ویژگی

طبقه‌بندی‌کننده+ روش انتخاب ویژگی	متوسط دقت پیش‌بینی
ماشین بردار پشتیبان	٪۶۸/۳
ماشین بردار پشتیبان + بهره اطلاعات	٪۶۵/۱۲
ماشین بردار پشتیبان +عدم قطعیت متقارن	٪۶۵/۱۲
ماشین بردار پشتیبان + انتخاب ویژگی برپایه همبستگی	٪۶۳/۷
ماشین بردار پشتیبان + انتخاب ویژگی هیبرید	٪۷۰/۲۳

با توجه به اعداد موجود در جداول، می‌توان به برتری روش انتخاب ویژگی هیبرید نسبت به سایر روش‌ها پی‌برد. در ادامه به منظور تایید و اثبات برتری عملکرد مدل برپایه ماشین بردار پشتیبان وزن‌دهی‌شده نسبت به مدل ماشین بردار پشتیبان ساده، اقدام اجرای آزمون‌های فرض تی-زوجی[1] و همچنین آزمون فرض ناپارامتریک من-ویتنی[2] نموده‌ایم. علّت استفاده از آزمون من-ویتنی این است که آزمون تی-زوجی فرض می‌کند که توزیع داده‌ها نرمال است؛ درحالی‌که ممکن است داده‌های ما نرمال نباشند. آزمون من ویتنی دو آزمون نمونه‌ای رتبه‌ای را، برای تفاوت بین میانه‌های دو جمعیت انجام می‌دهد. طبق جدول ۶، نتایج مربوط به آزمون من-ویتنی، برای تمامی روش‌های انتخاب ویژگی، تفاوت قابل توجهی را بین دقت پیش‌بینی با ماشین بردار پشتیبان ساده و ماشین بردار پشتیبان تعدیل‌یافته، نشان می‌دهد. این به این معنی است که با وجود اینکه ماشین بردار پشتیبان ساده هم نتایج نسبتا خوبی را ارائه کرده است، ماشین بردار پشتیبان تعدیل‌یافته، در مورد تمامی روش‌های پیش‌بینی نتایج خوبی را ارائه کرده است. با توجه به نتایج ارائه‌شده، واضح است که برای نمونه داده‌های مورد استفاده ما، عملکرد ماشین بردار پشتیبان تعدیل یافته، برتری قابل توجهی نسبت به عملکرد ماشین بردار پشتیبان ساده دارد.

پژوهشگاه علوم انسانی و مطالعات فرهنگی
پرتال جامع علوم انسانی

1. Paired t-test
2. Mann-Whitney nonparametric test

جدول ۶. آزمون تی-زوجی و من-ویتنی، به منظور مقایسه بین دقت‌های ماشین بردار پشتیبان تعدیل یافته و ماشین بردار پشتیبان ساده

روش انتخاب ویژگی	طبقه‌بندی‌کننده	دقت (٪)	آزمون تی-زوجی		آزمون ناپارامتریک من-ویتنی	
			آماره t	p-value	آماره Z	p-value
روش هیبرید FSSFS	ماشین بردار پشتیبان تعدیل‌یافته	٪۷۳/۵	۱۳/۱۷۶	۰/۰۰۱	-۲/۳۶۹	۰/۰۰۱
	ماشین بردار پشتیبان ساده	٪۷۰/۲۳				
بهره اطلاعات IG	ماشین بردار پشتیبان تعدیل‌یافته	٪۷۱/۴	۴/۸۱۵	۰/۰۰۱	-۲/۰۴۶	۰/۰۰۲
	ماشین بردار پشتیبان ساده	٪۶۵/۱۲				
عدم قطعیت متقارن SU	ماشین بردار پشتیبان تعدیل‌یافته	٪۷۰/۶۸	۵/۵۵۶	۰/۰۰۱	-۱/۹۴۵	۰/۰۰۱
	ماشین بردار پشتیبان ساده	٪۶۵/۱۲				
انتخاب ویژگی برپایه همبستگی CFS	ماشین بردار پشتیبان تعدیل‌یافته	٪۷۱/۳۱	۱۳/۴۰۴	۰/۰۰۱	-۱/۶۰۲	۰/۰۰۳
	ماشین بردار پشتیبان ساده	٪۶۳/۷				

۵. بحث و نتیجه‌گیری

در این پژوهش مدل ترکیبی جدید از ماشین بردار پشتیبان وزن‌دهی‌شده با حجم معاملات واقعی و انتخاب ویژگی هیبرید، ارائه شد. همچنین به منظور اثبات توانایی و قابلیت مدل پیشنهادی در پیش‌بینی درست و قابل اطمینان روندهای کوتاه‌مدت، نتایج این مدل با انتخاب ویژگی‌های مختلف و همین‌طور با نتایج طبقه‌بندی‌کننده ماشین بردار پشتیبان ساده مقایسه شد. همان‌طور که دیدیم برای هر دو طبقه‌بندی‌کننده، دقت پیش‌بینی روش انتخاب ویژگی هیبرید پیشنهادی از سایر تکنیک‌های انتخاب ویژگی بیشتر است. همچنین برای اثبات برتری ماشین بردار پشتیبان وزن‌دهی‌شده نسبت به ماشین بردار پشتیبان ساده، اقدام به ارائه استراتژی معاملاتی بر پایه هر دو روش نمودیم. همان‌طور که با توجه به جداول و شکل‌ها دیده می‌شود، نتایج روش پیشنهادی در این مورد هم به میزان قابل توجهی بهتر بود. همچنین لازم به یادآوری است که در این پژوهش از معتبرسازی مقطعی استفاده شده است و در هربار اجرای الگوریتم از ۸۰ درصد از داده‌ها برای آموزش و ۲۰ درصد برای آزمون استفاده شده و استراتژی معاملاتی بر روی داده‌های ۲ سال آخر پیاده‌سازی شده است؛ بنابراین می‌توان ادعا کرد که دقت‌های ارائه شده برای پیش‌بینی روش‌های مختلف که در جداول ارائه شده‌اند و همچنین نتایج استراتژی،

قابل اطمینان می‌باشد.

علاوه‌براین، نتایج نشان می‌دهد که شاخص‌های حجم جاری معاملات و دامنه متوسط واقعی با دقت قابل قبولی توانایی پیش‌بینی قیمت سهم را دارند. که این مورد با وزن‌دهی‌شده بودن ماشین بردار پشتیبان با حجم معاملات هم‌خوانی دارد.

٦. پیشنهادها

از جمله مواردی که با توجه به نتایج این پژوهش، می‌تواند محل مطالعه و بررسی واقع شود، تابع وزن‌دهی به تابع جریمه در مدل طبقه‌بندی‌کننده ماشین بردار پشتیبان می‌باشد. باید بررسی شود که آیا روش بهتری برای این وزن‌دهی وجود دارد یا خیر. خصوصا وارد کردن نوسانات به این وزن‌دهی، به‌عنوان عامل پیش‌بینی‌کننده، می‌تواند موجب بهبود مدل ارائه‌شده در این پژوهش شود. همچنین بررسی سایر روش‌های انتخاب ویژگی، به منظور افزایش دقت پیش‌بینی، می‌تواند محل مطالعات بیشتر واقع‌شود. علاوه‌براین، گسترش مجموعه داده‌های آموزش، به‌طوری‌که علاوه‌بر شاخص‌های تحلیل تکنیکال، شامل عواملی از اقتصاد کلان محیط نیز بشود، می‌تواند حوزه جالبی برای بهبود مدل و عملکرد آن باشد.

پژوهشگاه علوم انسانی و مطالعات فرهنگی
پرتال جامع علوم انسانی

منابع

1. Abu-Mostafa, Y. S., & Atiya, A. F. (1996). Introduction to financial forecasting. *Applied Intelligence, 6(3),* 205˚213.
2. Badri, Dawaloo & Dori Nokarani (1395). Studying macroeconomic factors on stock market performance, *Journal of financial management perspective. 13*: (9-35).
3. Chavarnakul, Thira, and David Enke (2008). Intelligent technical analysis based equivolume charting for stock trading using neural networks. *Expert Systems with Applications 34.2*: 1004-1017.
4. Chen, Y.-W., & Lin, C.-J., (2005). Combining SVMs with various feature selection strategies. Available from: <http://www.csie.ntu.edu.tw/~cjlin/papers/features.pdf>.
5. Choudry, R. & Grag, K. (2008). A Hybrid Machine Learning System for Stock Market Forecasting. *World Academy of Science, Engineering and Technology, 39.*
6. Cortes, C., & Vapnik, V. (1995). Support-vector networks. *Machine Learning, 20,* 273˚297.
7. Cristianini, N., & Shawe-Taylor, J. (2000). An introduction to support vector machines.Cambridge: Cambridge University Press.
8. Dash, M. et al., (2002). Feature selection for clustering ˚ a filter solution. In Proceedings of the second international conference on data mining :115˚122.
9. Ebadi, Forcasting Stock index price in Tehran stock exchange using artificial network, Thesis for master degree, Bu Ali Sina University, Faculty of economy and social science.
10. Eslami Bidgoli (1393). Shabanpoorfard, Relevency of trading volume and return case study in Tehran stock exchange considering intraday volume and excluding the U effect, *Journal of Financial Management Perspective, 6:* 45-63.
11. Fallahpour, Golarzi, Fatoorechian, Stock Price Movement Prediction Using Support Vector Machine Based on Genetic Algorithm in Tehran Stock Exchange Market. *Journal of financial research. 15(2),* 269-288.
12. Huang, C.-L., Chen, M.-C., & Wang, C.-J. (2007). Credit scoring with a data mining approach based on support vector machines. *Expert Systems with Applications, 33(4),* 847˚856.
13. Huang, C.-J., Dian-Xiu & Chuang, Y.-T. (2007). Application of wrapper approach and composite classifier to the stock trend prediction. *Expert Systems with Applications, 34(4),* 2870˚2878.
14. Hashemi, Analysis of behavioral factors effect on forcasting stock price using regression feedforward artificial neural network, Thesis for master degree, University of Science and Culture.
15. Kohavi, Ron, and George H. John (1997). Wrappers for feature subset selection. Artificial intelligence 97.1: 273-324.
16. Lawrence, S., Giles, C. L., & Tsoi, A.-C. (1997). Lessons in neural network training: Over fitting may be harder than expected. In Proceedings of the fourteenth national conference on artificial intelligence, AAAl-97: 540˚545.
17. Lee, Ming-Chi (2009). Using support vector machine with a hybrid feature selection method to the stock trend prediction. *Expert Systems with Applications 36.8*: 10896-10904.
18. Leung, M. T., Daouk, H., & Chen, A. S. (2000). Forecasting stock indices: A

comparison of classification and level estimation models. *International Journal of Forecasting, 16,* 173˚190.
19. Min, J. H., & Lee, Y.-C. (2005). Bankruptcy prediction using support vector machine with optimal choice of kernel function parameters. *Expert Systems with Applications, 28(4),* 603˚614.
20. Monajemi, Ebrazi, Rayati, (1388). Stock price prediction in Tehran stock exchange using artificial neural network. *Journal of financial economy, 6(3),* 1-26.
21. Tan, T. Z., Quek, C., & See, Ng. G. (2007). Biological brain-inspired genetic complementary learning for stock market and bank failure prediction. *Computational Intelligence, 23(2),* 236˚261.
22. Theodoridis, S., & Koutroumbas, K. (2008). Pattern Recognition. Academic Press.
23. Yu, L., & Liu, H., (2003). Feature selection for high-dimensional data: A fast correlation-based filter solution. In Proceedings of the 20th international conference on machine learning: 856˚863.
24. bikowski, Kamil. (2014). Using volume weighted support vector machines with walk forward testing and feature selection for the purpose of creating stock trading strategy. *Expert Systems with Applications.* 10.001.
25. Zhang, Xiangzhou, et al (2014). A causal feature selection algorithm for stock prediction modeling. *Neurocomputing, 142*: 48-59.

پژوهشگاه علوم انسانی و مطالعات فرهنگی
پرتال جامع علوم انسانی