

پیش بینی شاخص بازار سهام به وسیله مدل مارکوف پنهان و روش خوشه بندی کا میانگین

سعید عسگری^۱

ناصر یزدانی^۲

محسن ناظم بکایی^۳

تاریخ پذیرش: ۹۵/۰۴/۲۰

تاریخ دریافت: ۹۴/۱۲/۰۵

چکیده

پیش بینی بازار سهام یک مسئله کلاسیک میباشد که تا کنون به طور گسترده ای به وسیله ابزارها و مدل های مختلف مورد تجزیه و تحلیل قرار گرفته است. تغییرات روند بازار سهام ناشی از تقابل نیرو های عرضه و تقاضا و دیگر عوامل اقتصادی است. تکنیک های اماری سنتی برای توصیف روند های فصلی و نامانایی داده های قیمت در بازار سهام ناتوان هستند. مدل مارکوف پنهان یکی از ابزار های بسیار قدرت مند در پردازش فرایندهای اتفاقی و دنباله های تصادفی است، که در پهنه وسیعی از کاربرد ها نظیر پردازش گفتار، تصویر و سیگنال استفاده شده است. این مدل یک مدل اماری بوده و کار برد اصلی آن در باز شناخت گفتار می باشد. اما از آنجا که یک مدل کلی برای فرایندهای تصادفی است می تواند در سری های زمانی مالی نیز به کار رود. در این تحقیق ما با استفاده از الگوریتم خوشه بندی K-mean به همراه مدل مارکوف پنهان بر روی متغیر های تکنیکال درصد تغییرات روزانه قیمت و حجم معاملات به پیش بینی شاخص بی ای ال ۲۰ (بلژیک) در یک روز آینده پرداخته ایم و در اخر دقت پیش بینی را با استفاده از معیار متوسط درصد خطاهای مطلق محاسبه کرده ایم که مقدار متوسط خطاهای مطلق برای پیش بینی مقدار روز آینده شاخص قیمتی بی ای ال ۲۰ در بازه زمانی ۱ دسامبر تا ۳۰ دسامبر ۲۰۱۵ عدد ۰,۳۴۶۳- درصد میباشد. سری های زمانی مقدار شاخص و حجم معاملات آن در بازه زمانی ۲۰۱۲/۱/۱ تا ۲۰۱۵/۱۲/۳۰ نمونه پژوهش را تشکیل میدهند. از داده های ۲۰۱۲/۱/۱ تا ۲۰۱۵/۱۱/۳۰ برای پرورش مدل و از داده های ۲۰۱۵/۱۱/۳۰ تا ۲۰۱۵/۱۲/۳۰ برای تست مدل و به دست آوردن دقت پیش بینی استفاده کرده ایم.

واژه های کلیدی: مدل مارکوف پنهان، شاخص قیمتی بی ای ال ۲۰، مدل های تکنیکال برای پیش بینی قیمت سهام، الگوریتم خوشه بندی کا میانگین.

۱- دانشجوی کارشناسی ارشد مدیریت بازرگانی (گرایش مالی) دانشگاه شاهد، تهران، ایران. saeed.mas23@yahoo.com

۲- استادیار دانشگاه شاهد، تهران، ایران، (نویسنده مسئول) n.yazdani@shahed.ac.ir

۳- استادیار دانشگاه شاهد، تهران، ایران. mnbokaei@yahoo.com

۱- مقدمه

از زمان دانشمندانی چون داو در قرن نوزدهم میلادی تاکنون مطالعات بسیاری برای پیش‌بینی قیمت سهام به روش‌های فاندamental، آماری و تکنیکال انجام شده است. ماهیت بازار سهام بسیار پیچیده و غیر خطی است. تغییرات در این بازار به عوامل مختلفی از جمله نرخ های سود(بهره)، اوراق خزانه، تغییرات و تحولات کلان اقتصادی و پیش‌بینی‌ها از ریسک بستگی دارد. سرمایه‌گذاران همیشه نمی‌توانند در چنین بازاری برنده باشد بنابراین نیاز به الگوریتم های غیر خطی برای کشف روابط پیچیده بین متغیر های این بازار بسیار محسوس می‌باشد. تحقیقات وسیعی در سال‌های اخیر برای دست یابی به یک مدل بهینه یا تقریباً بهینه برای پیش‌بینی تغییرات قیمت سهام انجام شده است. بسیاری از مطالعات انجام شده تکنیک‌های تجزیه و تحلیل آماری سری‌های زمانی مانند خود رگرسیون‌های میانگین متحرک یا به اصطلاح اریما و مدل های رگرسیون چند متغیره را به کار گرفته‌اند. بعلاوه سیستم هایی مبتنی بر هوش مصنوعی مانند شبکه های عصبی مصنوعی، شبکه‌های فازی و ترکیبی از شبکه‌های عصبی فازی، ماشین‌های بردار پشتیبان وجود دارند که هرکدام از این روش ها محدودیت های خاص خود را دارند. به عنوان نمونه تکنیک شبکه های عصبی مصنوعی به دلیل نحوه پیکربندی آن بسیار مسئله محور بوده (حسن و ناد^۱، ۲۰۰۵). مشکل اصلی شبکه‌های عصبی برآزش بیش از حد داده ها با درنظر گرفتن نویز می‌باشد و همچنین با تغییر دامنه داده ها نیاز به پیکر بندی مجدد دارد. در سال ۲۰۰۰ رامهی و شین^۲ از الگوریتم القا محور فازی برای پیش بینی سری‌های زمانی مالی استفاده کردند. اما بیشتر وقت‌ها دانش متخصص برای استخراج قوانین فازی نیاز است. همچنین مدل های رگرسیونی و اتو رگرسیونی تازمانی که فرض بر خطی بودن سری های زمانی باشد مناسب به نظر می‌رسند اما وقتی واریانس در سری‌های زمانی افزایش می‌یابد و

فرایند غیر خطی در سری‌های زمانی پیدا شود مشکل این روش ها بروز میکند(الغلیب^۳ و همکاران، ۲۰۱۴). مدل مارکوف پنهان در اواخر دهه ۱۹۶۰ میلادی برای شناسایی صدا معرفی گردید و در حال حاضر به سرعت در حال گسترش دامنه کاربردها می باشد. دو دلیل مهم برای این موفقیت وجود دارد. اول اینکه این مدل از لحاظ ساختار ریاضی بسیار قدرتمند است و به همین دلیل مبانی نظری بسیاری از کاربردها را شکل داده است. دوم اینکه مدل مارکوف پنهان اگر به صورت مناسبی ایجاد شود می تواند برای کاربردهای بسیاری مورد استفاده قرار گیرد. در سالهای اخیر محققان مدل مارکوف پنهان را به عنوان طبقه بندی کننده و پیش بینی کننده سیگنال های صوتی، توالی یاب DNA، شناسایی کننده ویژگی های دست خط ارائه داده اند(رابینر. ارال^۴، ۱۹۸۹).

۲- مبانی نظری و مروری بر پیشینه پژوهش

سیستمی را درنظر بگیرید که ممکن است در هر مقطع زمانی در هر یک از N حالت مجزای S_1, S_2, \dots, S_n قرارگیرد. در زمان های مختلف و جدا از هم حالت سیستم بین N حالت تغییر میکند(یا احتمالاً در همان حالت قبلی میماند). زمان های جداگانه را t و حالت واقعی در زمان t را q_t مینامیم. در زنجیره مارکوف گسسته مرتبه اول، تعریف احتمالی حالت فعلی سیستم فقط به یک حالت قبلی آن بستگی دارد.

$$P [q_t = |q_{t-1} = S_i, q_{t-2} = S_k, \dots] = [q_t = |q_{t-1} = S_i]$$

احتمال انتقال از حالت i به حالت j به شکل زیر تعریف میشود.

$$a_{ij} = p[q_t = S_j | q_{t-1} = S_i], \quad 1 \leq i, j \leq N$$

مجموع ضرایب انتقال حالات برابر یک میباشند و هریک از ضرایب انتقال حالات بزرگتر از صفر هستند.

$$a_{ij} \geq 0, \quad \sum_{j=1}^n a_{ij} = 1.$$



نمودار ۱- مدل زنجیره مارکوف (رایبئر).

ار.ال، ۵، ۱۹۸۹)

افزایش بیش از ۵۰ درصد در حجم معاملات، کاهش بیش از ۱- در قیمت و افزایش زیر ۵۰ درصدی در قیمت سهام، کاهش بیشتر از ۱- درصدی در قیمت و کاهش زیر صفر حجم معاملات] داشته باشیم و این مشاهدات را با نماد هایی از O1 تا O9 نشان دهیم. به این فرایند گسسته سازی برداری می گویند(لاژوز، ۲۰۱۱). می خواهیم بدانیم که احتمال رخ دادن O_i در هر یک از سه حالت ممکن سیستم چقدر است؟ در واقع توزیع احتمال مشاهدات در حالت ها چگونه است؟

ماتریس مشاهدات یک ماتریس $M * N$ میباشد که در آن M تعداد مشاهدات ممکن و N حالت های ممکن سیستم میباشد. و بیانگر احتمال مشاهده شدن O در حالت S میباشد. اگر ماتریس مشاهدات را با B نشان دهیم (لاژوز، ۲۰۱۱):

$$B = \{b_{ij} | b_{ij} = p(o_t = v_j | q_t = s_i)\} = \{b_i(j)\} = p_{s*o}(s_i, o_j)$$

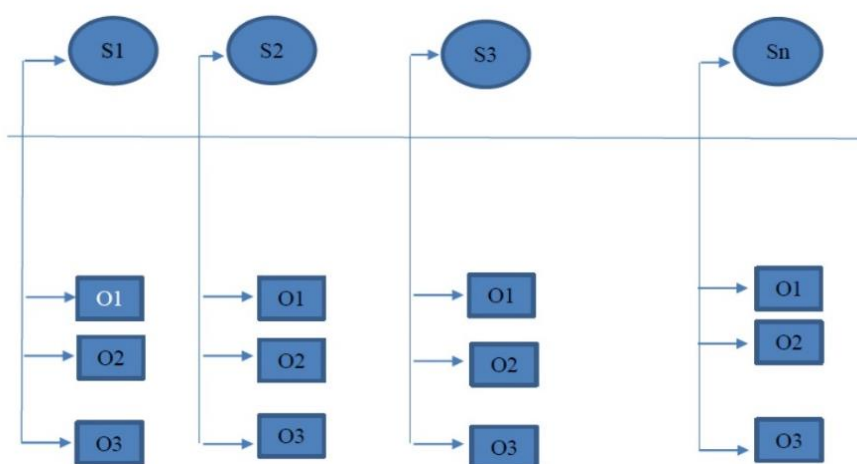
تا به اینجا ما مدل زنجیره مارکوف قابل مشاهده با مشاهدات گسسته را معرفی کردیم. اما در دنیای واقعی این مدل محدودیت های فراوانی دارد. در این قسمت مدل مارکوف را گسترش میدهیم. به صورتی که شامل حالت های غیر قابل مشاهده می باشد اما این حالت ها فقط از طریق تولید خروجی های خود (مشاهدات) قابل شناسایی هستند (رایبئر. ار.ال، ۱۹۸۹).

برای درک بهتر مدل مارکوف پنهان فرض کنید در بازار سهام مشاهدات $O_1, O_2, O_3, \dots, O_t$ از ترکیب تغییرات قیمت سهام و حجم معاملات موجود باشند. فرض کنید بازار در زمان های متفاوت در حالت های مختلفی به سر میبرد اما این حالت ها برای ما قابل مشاهده نمیباشند. در واقع ما نمیدانیم هر کدام از این مشاهدات از تغییرات قیمت سهام و حجم معاملات آن در چه حالتی از بازار در حال وقوع هستند. یک مدل مارکوف پنهان شامل اجزای زیر است (لاژوز، ۲۰۱۱).

ماتریسی که در آن ضرایب a_{ij} نشان داده می شود ماتریس انتقال نام دارد. اگر مطابق مثال بالا یک سیستم با سه حالت داشته باشیم ماتریس انتقال آن شامل سه سطر و سه ستون است که در آن جمع عناصر هر سطر برابر یک می باشد. زنجیره بالا تا زمانی که حالت ها و مشاهدات مربوط به هر حالت آن آشکار و معلوم هستند، زنجیره مارکوف قابل مشاهده نامیده می شود. برای مثال اگر ما فرض کنیم که در بازار سهام سه حالت وجود داشته باشد. حالت گاو که بیانگر رشد بیشتر از یک درصدی شاخص بازار می باشد. حالت خرس که بیانگر کاهش بیشتر از ۱- درصدی شاخص باشد و در آخر حالت سوم که بیانگر تغییرات شاخص در بازه بین ۱ و ۱- می باشد. برای مثال اگر امروز در حالت بازار گاو باشیم چقدر احتمال وجود دارد تا روز بعد بازار در حالت خرس باشد؟ چقدر احتمال وجود دارد تا بازار در حالت گاو بماند؟ تمامی این سوالات به وسیله ماتریس انتقال قابل پاسخ گویی هستند. حال فرض کنید که ما مشاهداتی گسسته (جدا از هم) متشکل از ترکیب حجم معاملات و درصد تغییرات شاخص بازار به صورت [افزایش بیش از ۱ درصد در مقدار قیمت سهام به همراه افزایش بیش از ۵۰ درصد در حجم معاملات، افزایش بیش از یک درصد در قیمت و افزایش کمتر از ۵۰ درصد در حجم معاملات، افزایش بیش از یک درصد در قیمت سهام و کاهش زیر صفر در حجم معاملات، تغییر بین ۱ و ۱- درصدی در قیمت و افزایش کمتر از ۱ و ۱- درصدی در قیمت سهام و افزایش کمتر از ۵۰ درصد حجم معاملات، تغییر بین ۱ و ۱- درصدی در قیمت و کاهش زیر صفر حجم معاملات، کاهش بیش از ۱- درصدی در قیمت و

تعریف ریاضی رسمی مارکوف پنهان بر پایه یک ماتریس انتقال حالات و یک ماتریس احتمال مشاهدات که از احتمالات شرطی تشکیل شده میباشد (الیورسی، ۲۰۰۹).

- ۱) مجموعه ای از حالت های متناهی S
- ۲) مجموعه ای متناهی از نماد مشاهدات Q
- ۳) ماتریس انتقال حالات
- ۴) ماتریس احتمال مشاهدات



نمودار ۲- مدل مارکوف پنهان (الیورسی، ۲۰۰۹)

گیری در مورد حالت های پنهان در مدل مارکوف برای تجزیه و تحلیل سری های زمانی تغییرات تصادفی قیمت به دانش متخصص بستگی دارد و در بازار های مختلف با توجه به ویژگی های مختلف آن بازار متفاوت است. در اکثر تحقیقات خارجی که بازار مورد مطالعه بورس های نسبتاً کارا و باثبات بوده، حالت های مارکوف پنهان روند بازار با سه حالت (صعودی، نزولی، بدون تغییر) یا با دو حالت (صعودی، نزولی) و با حالت های بیشتر بوده است. اما در بازار های ناکارا و بی ثبات به نظر میرسد باید حالت های پنهان توسط شوک ها و حالات روحی رفتاری سرمایه گذاران تعریف شوند. مثلاً فرض کنید در یک مدل دو حالتی حالت اول حالتی است که سرمایه گذاران رفتار هیجانی ناشی از اخبار داشته اند و حالت دوم حالتی است که آنها رفتار منطقی در پیش گرفته اند و شوک خبری در بازار وجود ندارد. به هر حال آنچه در تحلیل مارکوف پنهان با حالات گسسته و قابل انتقال در پیش بینی قیمت سهام مهم است، این موضوع می باشد که در واقع

علاوه بر مدل مارکوف پنهان استاندارد تعدادی از انواع دیگر آن نیز ارائه و برای مدل سازی استفاده شده است (لاژوز، ۲۰۱۱).

- مارکوف پنهان داده-سناده^۹
- مارکوف پنهان مرتبه چندم^{۱۰}
- مارکوف پنهان با حالت های خاموش^{۱۱}
- مارکوف پنهان چپ به راست یا راست به چپ^{۱۲}
- مارکوف پنهان زوجی^{۱۳}
- مارکوف پنهان اتو رگرسیون^{۱۴}

میتوانیم فرض کنیم که رفتار بازار دو حالتی باشد، حالت اول شوک هیجانی بازار که باعث افزایش حساب گونه قیمت ها میشود و حالت دوم ثبات بازار که در این حالت قیمت ها به صورت معقول افزایش یا کاهش میابند. یا حتی میتوان حالت های پنهان را بسته به روند ده روزه یا سی روزه بازار در نظر گرفت. مانند مثال بازار گاو، خرس و بازار بدون تغییر. به عنوان مثال اگر میانگین قیمت های ده روز گذشته صعودی باشد روند بازار در ده روز گذشته بازار گاوی بوده و ...، تصمیم

۳) مسئله پرورش: با توجه به توالی مشاهدات $\lambda=(s, v, a, \pi)$ مدل مارکوف پنهانی $\{O_1, O_2, \dots, O_n\}$ را پیدا کنید که به بهترین نحو توالی مشاهدات را توضیح دهد. که به معنی پیدا کردن مقادیر $\lambda=(s, v, a, \pi)$ که $P(O|\lambda)$ را حداکثر میکند میباشد. به عبارات دیگر این مسئله تخمین پارامتر های مدل مارکوف پنهان با توجه به مجموعه مشاهدات میباشد که معمولاً به وسیله الگوریتم باثوم ویچ^{۲۰} پاسخ داده میشود. الگوریتم باثوم ویچ نوع خاصی از الگوریتم پیشینه سازی انتظار^{۲۱} (امید ریاضی) است.

الگوریتم پیشینه سازی انتظار برای پیدا کردن بیشترین احتمال برآورد پارامتر های مدل های احتمالی وقتی که مدل به متغیر های غیر قابل مشاهده و پنهان بستگی دارد به کار میرود. الگوریتم پیشینه سازی انتظار بین اجرای مرحله انتظار (امید ریاضی) که انتظار احتمال را برآورد میکند، و حداکثر سازی که محتمل ترین برآورد^{۲۲} پارامتر هارا توسط حداکثر سازی احتمال مورد انتظار^{۲۳} که در مرحله امید ریاضی به دست آمده بود محاسبه میکند تکرار میشود. پارامتر هایی که در مرحله حداکثر سازی پیدا شده بودند برای شروع یک مرحله امید ریاضی دیگر به کار میروند و این فرایند تکرار میشود. دو مثال مشهور از الگوریتم پیشینه سازی انتظار، الگوریتم باثوم ویچ (یا پسر و پیشرو^{۲۴}) و الگوریتم درون - بیرون^{۲۵} هستند (ریچکا، ۲۰۱۱).

پیشینه پژوهش‌ها

در سال ۲۰۰۵ میلادی حسن و ناد^{۲۶} از مدل مارکوف پنهان برای پیش بینی قیمت سهام در بازار های مرتبط به هم استفاده کرده اند. آنها از سری های زمانی قیمت های آغازین، آخرین قیمت، بیشترین قیمت و کمترین قیمت در هر روز از سال ۲۰۰۲ تا ۲۰۰۴ برای آموزش مدل استفاده کردند و از داده های سه ماه بعد نیز برای تست مدل استفاده نمودند. سپس آنها داده‌هایی را که در گذشته بیشترین همسایگی

حالات برای ما قابل مشاهده هستند و ما خود با توجه به دانش تخصصی خود نسبت به بازار حالت ها را تعریف میکنیم و با دسته بندی و یا خوشه بندی داده ها به روش های مختلفی که در ادامه توضیح داده میشود حالات خروجی های مشاهده شده گذشته را تعریف میکنیم. در واقع تفاوت کاربرد مدل مارکوف پنهان در پیش بینی قیمت در بازار های سهام با دیگر کار برد های ان مانند شناسایی سیگنال های صوتی در این نکته میباشد.

در پیاده‌سازی و کاربرد های مارکوف پنهان معمولاً درگیر سه مسئله اساسی میشویم که در ادبیات موضوع از آنها با عنوانهای مسئله برآورد^{۱۵}، دی کده کردن^{۱۶} و پرورش^{۱۷} یاد میشود. این مسائل بنیادی و جنبه های دیگر مارکوف پنهان به طور گسترده ای در ادبیات موجود مطالعه شده اند. تئوری، کاربرد ها و پیاده سازی مارکوف پنهان مراتب بالاتر نسبتاً پیچیده تر هستند و بنابر این به صورت وسیع مورد تحقیق، توسعه و پیاده سازی قرار نگرفته اند. در ادامه، توضیحی بر مسائل موجود در ادبیات مارکوف پنهان آورده میشود (رابینر، ا.رال،^{۱۸} ۱۹۸۹):

۱) مسئله برآورد یا تخمین: باتوجه به مدل مارکوف پنهان $\lambda=(s, v, a, \pi)$ و توالی مشاهدات $O = O_1, O_2, \dots, O_t$

احتمال اینکه توالی مشاهده شده توسط مدل ایجاد شود چیست؟ یعنی تعیین $P(O|\lambda)$.

۲) مسئله دی کده کردن: باتوجه به مدل مارکوف پنهان $\lambda=(s, v, a, \pi)$ و توالی مشاهدات $O = O_1, O_2, \dots, O_t$ محتمل ترین توالی حالت پنهان $Q = q_1, q_2, \dots, q_t$ که میتوانند مشاهدات را تولید کنند چیست؟ یعنی پیدا کردن توالی حالات $Q = q_1, q_2, \dots, q_t$ که احتمال مشاهده ای $P(O, Q|\lambda)$ را حداکثر کند. که معمولاً با استفاده از الگوریتم وایتربی^{۱۹} پاسخ داده میشود.

در سال ۲۰۱۱ میلادی جارسلاو لائوز^{۲۷} در رساله دکتری خویش اقدام به ایجاد استراتژی‌های سرمایه گذاری بر پایه مدل مارکوف پنهان نمود. او همچنین یک الگوریتم با ۱۰ گام برای بدست آوردن پارامترهای مارکوف پنهان و پیش بینی قیمت سهام طراحی نمود. از مدل مارکوف پنهان برای پیش‌بینی رفتار شاخص S&P 500 و اتخاذ و ایجاد استراتژی‌های مناسب سرمایه گذاری استفاده نمود. در نتیجه گیری‌ها بیان شد که مدل مارکوف پنهان موفق بوده و در مقایسه با رگرسیون‌های میانگین متحرک بسیار بهتر ظاهر شده است.

در سال ۲۰۱۲ میلادی بهوان و گوپتا از روش HMM MAP^{۲۸} برای پیش بینی ارزش سهام با توجه به داده‌های تاریخی استفاده کردند. آنها برای آموزش HMM پیوسته خود از تغییرات کسری ارزش سهام بیشترین قیمت و کمترین قیمت درون یک روز استفاده کردند. سپس آنها از این مدل برای تصمیم‌گیری ماکزیمم ای پوستریوری^{۲۹} برای تمام مقدارهای ممکن برای روز آینده استفاده کردند. رهیافت خود را بر روی سهام‌های مختلف تست کردند و نتایج عملکرد آن را با تعدادی از روش‌های موجود که از HMM و ANN استفاده کرده بودند مقایسه کردند. مقایسه دقت پیش‌بینی AMP HMM با روش‌های دیگر نشان از بهبود دقت پیش‌بینی با استفاده از این روش بود.

در سال ۲۰۱۴ میلادی اسداله الغلیب^{۳۰} و همکاران تحقیقی مشابه تحقیق حسن و ناد که در سال ۲۰۰۵ انجام شده بود را در بازار هند انجام دادند و دقت پیش‌بینی مدل مارکوف پنهان را با دقت پیش‌بینی الگوریتم نزدیک‌ترین همسایه مقایسه کردند. آنها در نتیجه گیری خود بیان کردند که HMM نسبت به نزدیک‌ترین همسایه بسیار حرفه‌ای‌تر عمل میکند.

در سال ۱۳۹۱ شمسی مقصود امیری به همراه مهدی بیگلری کامی یک مدل خطی برای پیش‌بینی قیمت‌اتی سهام با استفاده از زنجیره مارکوف طراحی نمودند. آنها با ترکیب متغیرهای حجم معاملات و در

(مطابقت) را با داده‌های مورد نظر داشتند را پیدا کرده و با ضرب نمودن ارزش‌های آنها تغییرات قیمت را محاسبه نمودند تا قیمت پایانی سهم را در روز بعد پیش‌بینی نمایند. سپس دقت پیش‌بینی خود را با استفاده از MAPE به دست آورده و نتایج را با دقت پیش‌بینی به روش شبکه‌های عصبی مصنوعی مقایسه کردند که نتایج پیش‌بینی بهبود دقت پیش‌بینی به وسیله HMM را نشان می‌دهد. آنها استدلال داشتند که ضعف اولیه شبکه‌های عصبی عدم توانایی آن در توضیح مدل‌هاست.

در سال ۲۰۰۶ میلادی حسن و ناد مدلی ترکیبی متشکل از مارکوف پنهان، شبکه‌های عصبی مصنوعی و الگوریتم ژنتیک برای پیش‌بینی رفتار بازارهای مالی ارائه دادند. آنها از روش شبکه‌های عصبی برای مرتب‌سازی و تبدیل قیمت‌های روزانه سهام به ورودی‌های مدل مارکوف پنهان استفاده نمودند. سپس الگوریتم ژنتیک را به منظور بهینه‌سازی پارامترهای اولیه مدل مارکوف پنهان به کار بردند. مدل مارکوف حل شده و نهایی برای شناسایی و کنار هم گذاشتن الگوهای مشابه تاریخی به کار برده شد. اختلاف قیمت روزهای مشابه در الگو و قیمت روزاتی محاسبه شد. در نهایت میانگین وزنی این اختلاف قیمت برای پیش‌بینی قیمت‌های آتی به دست آورده شد. آنها با این مدل قیمت‌تعدادی سهم از بخش‌ای تی را پیش‌بینی نمودند و نتایج را با مدل‌های سنتی مقایسه کردند.

در سال ۲۰۰۹ میلادی حسن برای پیش‌بینی قیمت سهام از ترکیبی از مدل مارکوف پنهان و مدل‌های فازی استفاده کرد. او با استفاده از مدل مارکوف پنهان الگوی سری‌های زمانی را شناسایی کرد و سپس با استفاده از قوانین فازی برای به دست آوردن مقدار پیش‌بینی تلاش کرد. نتایج حاصل از کار او بهبود دقت پیش‌بینی با استفاده از این مدل را در مقایسه با سایر مدل‌ها از قبیل ANN, ARIMA تایید کرد.

$\text{dotproduct}(A_i P_t E(\text{state} - \text{returns}))$

که A بیانگر ماتریس انتقال و P^3 احتمالات شرطی قرار داشتن مشاهده در هر حالت و E تغییرات مورد انتظار قیمت در هر حالت هستند.

گسسته سازی ورودی ها میتواند بر اساس سری های زمانی یک متغیر واحد مانند قیمت باشد. البته گسسته سازی حرفه ای که برای مسائل واقعی کار برد داشته باشد میتواند شامل دو یا چند متغیر مربوط به هم مانند قیمت، حجم معاملات، مقدار شاخص صنعت مربوطه، قیمت سهام مرتبط با سهام مورد بررسی و مقدار شاخص کل بورس باشد (ریجکا، ۲۰۱۱). ترکیب درصد روزانه تغییرات شاخص و حجم معاملات مشاهدات ما را تشکیل می دهند و برای گسسته سازی این مشاهدات از جدول زیر استفاده می کنیم.

ما برای مدل مارکوف پنهان ۳ حالت در نظر می گیریم. هرچه حالت های مدل بیشتر باشند دقت پیش بینی افزایش میابد اما فرایند محاسبات پیچیده تر میشود، تا جایی که دیگر با افزایش حالت ها بهبود معنی داری در مدل ایجاد نمی شود (لاژوز، ۲۰۱۱). گام بعدی برای اجرای مدل مارکوف پنهان انتخاب یک راه موثر برای تخمین اولیه ماتریس انتقال و مشاهدات است. این انتخاب در فرآیند پرورش مدل مارکوف پنهان بسیار حیاتی می باشد. معمولاً مساله پرورش مدل مارکوف پنهان به کمک الگوریتم باثوم ویج حل می شود که برآورد بیشینه احتمال^{۳۲} را محاسبه می کند. باثوم ویج ماتریس های انتقال حالات و مشاهدات اولیه را به صورتی که احتمال تطبیق آنها با توالی مشاهدات حداکثر شود تعدیل می کند. یعنی حداکثر کردن $P(O|\lambda)$ و این حقیقت که تخمین باثوم ویج فقط در یک نقطه حداکثر محلی^{۳۳} تلاقی^{۳۴} می شود می شود شناخته شده است (الیور، سی، ۲۰۰۹). در بیشتر مسائل دنیای واقعی سطح بهینه سازی $P(O|\lambda)$ بسیار پیچیده است و تعداد زیادی نقطه حداکثر محلی وجود دارد. بعلاوه راه حلی برای پیدا کردن نقطه حداکثر کلی^{۳۵} وجود ندارد.

صد تغییرات قیمت ۹ وضعیت را تعریف کردند و برای هر یک از این متغیر ها سه سطح مثبت و منفی و خنثی تعریف نمودند. ایشان مدل پیش نهادی خود را با استفاده از شاخص داوجونز مورد آزمایش قرار دادند.

۳- روش شناسی پژوهش

در بخش های پیشین جنبه های تئوری مدل مارکوف پنهان بیان شد. هدف اصلی این پژوهش پیاده سازی تئوری های موجود برای رسیدن به پیش بینی درصد تغییرات مقدار شاخص بی ای ال ۲۰ در یک روز آینده است. بی ای ال ۲۰ شاخص قیمتی میباشد که متشکل از قیمت های سهام ۲۰ شرکت است که بیشترین سرمایه را در بورس بروکسل دارند. جامعه آماری این تحقیق تمامی مقادیر شاخص بی ای ال ۲۰ به همراه حجم معاملات آن به صورت روزانه میباشد. سری های زمانی مقدار شاخص و حجم معاملات آن در بازه زمانی ۲۰۱۲/۱/۱ تا ۲۰۱۵/۱۲/۳۰ نمونه پژوهش را تشکیل می دهند. از داده های ۲۰۱۲/۱/۱ تا ۲۰۱۵/۱۱/۳۰ برای پرورش مدل و از داده های ۲۰۱۵/۱۱/۳۰ تا ۲۰۱۵/۱۲/۳۰ برای تست مدل و به دست آوردن دقت پیش بینی استفاده کرده ایم. داده ها از پایگاه اینترنتی www.investing.com استخراج شده اند. برای پیاده سازی مدل مارکوف پنهان گام های زیر برداشته میشود:

- ۱) تبدیل مشاهدات پیوسته قیمت و حجم معاملات روزانه به مشاهدات گسسته با فرایند کمی سازی برداری.
- ۲) انتخاب تعداد حالت های پنهان مدل .
- ۳) تخمین اولیه ماتریس انتقال حالات مدل.
- ۴) تخمین اولیه ماتریس مشاهدات مدل.
- ۵) محاسبه برآورد بیشترین احتمال ماتریس انتقال و ماتریس مشاهدات.
- ۶) محاسبه توالی حالات احتمالی مشاهدات (P).
- ۷) پیش بینی مقدار شاخص در آینده با استفاده از .

$HMM \text{ forecast} =$



جدول ۱- نقشه تبدیل مشاهدات و ایجاد ورودی های مدل مارکوف پنهان (لاژوز، ۲۰۱۱)

O	درصد تغییرات شاخص	درصد تغییرات حجم معاملات	O	درصد تغییرات شاخص	درصد تغییرات حجم معاملات
۱	$۲ >$	> ۵۰	۱۷	$۰ = < < -۰,۵$	> ۵۰
۲	$۲ >$	$۵۰ = < < ۰$	۱۸	$۰ = < < -۰,۵$	$۵۰ = < < ۰$
۳	$۲ >$	$۰ = < < -۵۰$	۱۹	$۰ = < < -۰,۵$	$۰ = < < -۵۰$
۴	$۲ >$	$> = -۵۰$	۲۰	$۰ = < < -۰,۵$	$> = -۵۰$
۵	$۲ = < < ۱$	> ۵۰	۲۱	$-۰,۵ = < < -۱$	> ۵۰
۶	$۲ = < < ۱$	$۵۰ = < < ۰$	۲۲	$-۰,۵ = < < -۱$	$۵۰ = < < ۰$
۷	$۲ = < < ۱$	$۰ = < < -۵۰$	۲۳	$-۰,۵ = < < -۱$	$۰ = < < -۵۰$
۸	$۲ = < < ۱$	$> = -۵۰$	۲۴	$-۰,۵ = < < -۱$	$> = -۵۰$
۹	$۱ = < < ۰,۵$	> ۵۰	۲۵	$-۱ = < < -۲$	> ۵۰
۱۰	$۱ = < < ۰,۵$	$۵۰ = < < ۰$	۲۶	$-۱ = < < -۲$	$۵۰ = < < ۰$
۱۱	$۱ = < < ۰,۵$	$۰ = < < -۵۰$	۲۷	$-۱ = < < -۲$	$۰ = < < -۵۰$
۱۲	$۱ = < < ۰,۵$	$> = -۵۰$	۲۸	$-۱ = < < -۲$	$> = -۵۰$
۱۳	$۰,۵ = < < ۰$	> ۵۰	۲۹	$-۲ = <$	> ۵۰
۱۴	$۰,۵ = < < ۰$	$۵۰ = < < ۰$	۳۰	$-۲ = <$	$۵۰ = < < ۰$
۱۵	$۰,۵ = < < ۰$	$۰ = < < -۵۰$	۳۱	$-۲ = <$	$۰ = < < -۵۰$
۱۶	$۰,۵ = < < ۰$	$> = -۵۰$	۳۲	$-۲ = <$	$> = -۵۰$

احتمال انتقال به حالت دو بسیار کم است و احتمال باقی ماندن در حالت یک زیاد است (بازار سهام باثبات) (لاژوز، ۲۰۱۱). شاید بهترین راه حل ترکیب نظر کارشناسان با روش های هوش مصنوعی باشد. به هر حال ما در این پژوهش از روش خوشه بندی کامیابگین برای رسیدن به حدس اولیه ماتریس های انتقال حالات و مشاهدات کمک میگیریم و بعد از پرورش دادن مدل توسط الگوریتم بائوم ویچ و برآورد بیشینه احتمال پارامتر های مدل ادامه مراحل چهار گانه را که در بالا ذکر شد طی میکنیم تا به پیش بینی شاخص بی ای ال ۲۰ در روز آینده برسیم. سپس با استفاده از داده های ۲۰۱۵/۱۱/۳۰ تا ۲۰۱۵/۱۲/۳۰ دقت پیش بینی را توسط معیار متوسط مطلق درصد خطاها^{۳۷} محاسبه میکنیم. متوسط مطلق درصد خطاها در امار یکی از روش های به دست آوردن درصد خطای پیش بینی است. و شیوه محاسبه آن به صورت فرمول مقابل است.

$$M=1/n \sum (At-Ft/At)$$

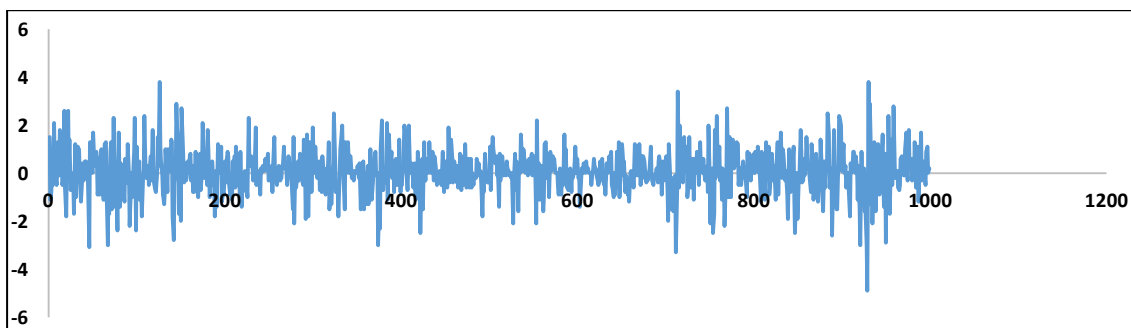
بنابراین ما نیاز به روشی داریم که به وسیله آن ماتریس انتقال و ماتریس مشاهدات را به صورت اولیه تعیین کنیم تا الگوریتم بائوم ویچ بتواند پاسخی را بیابد که ارائه دهنده نقطه ماکزیمم کلی باشد و یا به صورت زیادی به آن نزدیک باشد (رابینر، ارال، ۱۹۸۹). این مسئله میتواند مسئله کاوش از نوع هوش مصنوعی باشد (الین ریچ، کوین کینگ ۱۹۹۱). تکنیک های زیادی از جمله خوشه بندی کا میانگین^{۳۶}، شبیه سازی احتمالی مونت کارلو یا دیگر روش های خوشه بندی می توانند برای پیدا کردن پارامترهای اولیه مدل مارکوف پنهان به کار گرفته شوند. روش دیگری که معمولاً بر تکنیک های هوش مصنوعی ترجیح داده می شود استفاده از نظر کارشناسان مالی برای پیدا کردن ماتریس انتقال حالات و مشاهدات اولیه میباشد. مثلاً از نظر کارشناسان، بازار های باثبات با حالت هایی شناخته می شوند که در طی زمان احتمال انتقال بین آنها کم می باشد. یعنی اگر اکنون در حالت یک باشیم

برای پرورش مدل استفاده شده اند شامل درصد تغییرات روزانه شاخص و درصد تغییرات حجم معاملات روزانه شاخص از تاریخ ۲۰۱۲/۱/۱ تا ۲۰۱۵/۱۱/۳۰ و قسمت دوم داده هایی که برای به دست آوردن دقت پیش بینی مدل استفاده شده اند شامل مقادیر روزانه شاخص از تاریخ ۲۰۱۵/۱۱/۳۰ تا ۲۰۱۵/۱۲/۳۰. داده های قسمت اول که شامل هزار مشاهده می شد را به روش کمی سازی برداری گسسته سازی کردیم.

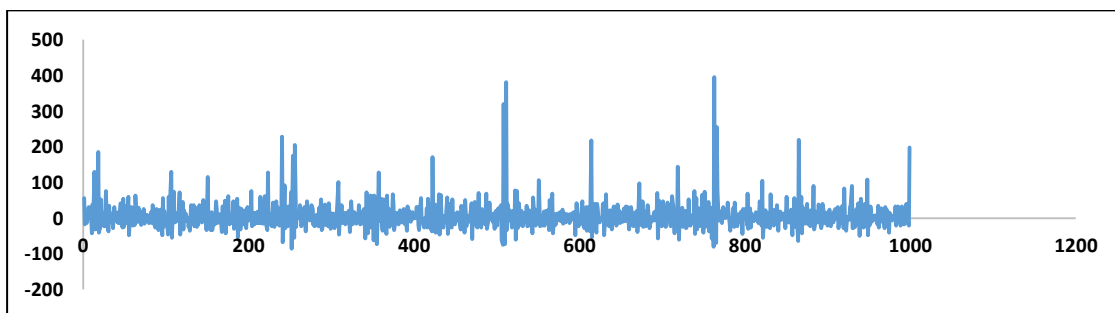
که n تعداد پیش بینی ها میباشد و At مقدار واقعی متغیر و Ft مقدار پیش بینی شده متغیر میباشد. اگر فرمول بالا را در صد ضرب کنیم در صد خطا را به دست می آوریم (تافلایس^{۳۸}، ۲۰۱۵).

۴- نتایج پژوهش

ما مدل مارکوف پنهان را به کمک توابع نرم افزار متلب پیاده سازی کردیم، داده های پژوهش به دو قسمت تقسیم میشوند. قسمت اول داده هایی که



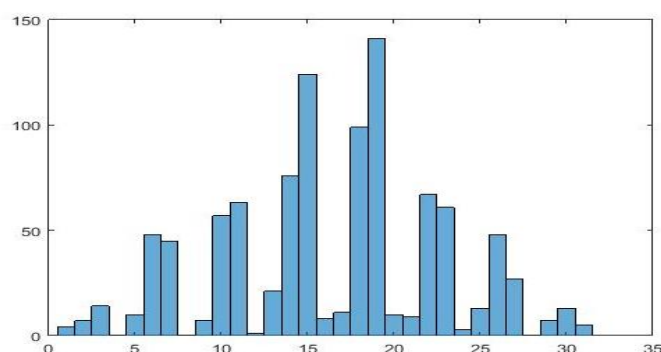
نمودار ۳- درصد تغییرات روزانه مقدار شاخص قیمتی بی ای ال ۲۰ از تاریخ ۲۰۱۲/۱/۱ تا ۲۰۱۵/۱۱/۳۰ شامل هزار مشاهده



نمودار ۴- درصد تغییرات روزانه حجم معاملات شاخص قیمتی بی ای ال ۲۰ از تاریخ ۲۰۱۲/۱/۱ تا ۲۰۱۵/۱۱/۳۰ شامل هزار مشاهده

جدول ۲- ماتریس انتقال حالات

1	0.26361	0.186082	0.550308
2	0.230301	0.151584	0.618114
3	0.285386	0.106596	0.608018



نمودار ۵- فراوانی مشاهدات گسسته

جدول ۳- مقایسه نتایج مقدار پیش بینی شده شاخص با مقدار واقعی شاخص به روش مارکوف پنهان

تاریخ	مقدار واقعی	مقدار پیش بینی
30.15.DEC	3734.02	3750.122
29.15.DEC	3753.05	3690.671
28.15.DEC	3688.4	3693.984
24.15.DEC	3691.71	3686.632
23.15.DEC	3689.51	3616.23
22.15.DEC	3614.72	3615.585
21.15.DEC	3614.23	3651.017
18.15.DEC	3648.77	3673.222
17.15.DEC	3676.09	3612.63
16.15.DEC	3615.45	3579.386
15.15.DEC	3582.18	3490.452
14.15.DEC	3492.9	3536.06
11.15.DEC	3538.54	3580.183
10.15.DEC	3577.98	3586.546
09.15.DEC	3585.39	3614.445
08.15.DEC	3616.98	3664.259
07.15.DEC	3667.12	3643.522
04.15.DEC	3641.28	3645.153
03.15.DEC	3647.71	3735.239
02.15.DEC	3732.94	3746.406
01.15.DEC	3744.1	3763.206

متوسط در صد خطای مطلق پیش بینی عدد ۰,۰۳۴۶۳- میباشد.

۵- نتیجه گیری و بحث

بیواینفورماتیک، بیو فیزیک و غیره کاربرد دارد، می توان موفقیت این مدل را به تئوری های پربار و غنی در عین حال پیچیده ریاضیات که پایه و اساس این روش را تشکیل می دهند نسبت داد.

در این پژوهش از مدل مارکوف پنهان به همراه الگوریتم کا میانگین برای پیش بینی یکی از مهم ترین مجهولات بازار سهام یعنی قیمت روز بعد استفاده کردیم. هدف ما از ارائه این مدل در دسترس قرار دادن

مدل مارکوف پنهان یک روش انعطاف پذیر است. دارای تعمیم پذیری بسیار بالا برای مدل نمودن سیستم های پویا که میتوانند در سری های زمانی یک متغیره یا چند متغیره مشاهده شوند، حتی اگر حالت سیستم مورد بررسی خود قابل مشاهده نباشد. در حالی که این مدل در زمینه های علمی زیادی از جمله شناسایی الگو ها توسط رایانه، رمز یابی، مالی،

- Doctoral dissertation, Oklahoma State University, United States of America
- * Md. Rafiul Hassan and Baikunth Nath, (2005), "stock market forecasting using hidden markov model: a new approach", Proceedings of the 2005 5th International Conference on Intelligent Systems Design and Applications (ISDA'05)
 - * Rabiner R L, (1989), A Tutorial on Hidden Markov Models and Selected Applications in Speech Recognition, Proceedings of the IEEE, Vol. 77(2), pp 257-286.
 - * Oliver C. Ibe, (2009) "Markov Processes for Stochastic Modeling", Elsevier Academic Press Inc, 2009.
 - * Rijeka, Przemyslaw Dymarski, (2011), "Hidden Markov Models, Theory and Applications" Publisher: InTech, Chapters published April 19, 2011 under CC BY-NC-SA 3.0 license
 - * Tofallis (2015). "A Better Measure of Relative Prediction Accuracy for Model Selection and Model Estimation", Journal of the Operational Research Society, 66(8), 1352-1362.
 - * Zhang, Y. (2004), "Prediction of Financial Time Series with Hidden Markov Models", Doctoral dissertation, Simon Fraser University, British Columbia, Canada

این روش برای موفقیت سرمایه گذاران در بازار سهام است. در بازار سهام مدل مارکوف پنهان نواقص و ضعف های مدل های دیگر پیش بینی سری های زمانی را دارا نبوده و از پایه های قوی آماری برخوردار است. همان طور که مشاهده می شود رقم متوسط درصد خطاهای مطلق مقدار بسیار ناچیزی است اما می توان با به کار بردن تکنیک های خاصی دقت پیش بینی مدل مارکوف پنهان را از این نیز بیشتر بهبود داد. مثلاً می توان در مرحله پیش بینی متوسط تغییرات در هر حالت را از مطالعه رفتار پیشین و تطبیق دادن دوره نزدیک به پیش بینی با دوره ای که رفتار سری های زمانی بیشترین شباهت را با آن دارد به دست آورد. کاری که حسن و ناد در پژوهش خود برای افزایش دقت پیش بینی قیمت سهام به وسیله مدل مارکوف پنهان انجام دادند. ورودی های مارکوف پنهان همچنین می توانند به جای داده های گسسته داده های پیوسته باشند و میتوان از متغیرهای دیگر از جمله تغییرات شاخص کل و یا تغییرات شاخص صنعت و سهام های مرتبط به سهام مورد پیش بینی به جای صرف حجم معاملات و درصد تغییرات قیمت استفاده نمود.

یادداشت ها

1. Hassan and Nath
2. Romahi and Shen
3. Al Galib
4. Rabiner R L
5. Rabiner R L
6. LAJOS
7. Rabiner R L
8. LAJOS
9. Input-output HMM
10. Nth order HMM
11. Silent HMM
12. Left to right or right to left HMM
13. Coupled HMM
14. Autoregressive HMM
15. Evaluation
16. Decoding
17. Learning
18. Rabiner R L
19. viterbi
20. Baum-welch
21. Expectation maximization(EM)
22. Maximum likelihood estimates
23. Expected likelihood

فهرست منابع

- * تهرانی، رضا، نوربخش، عسگر، (۱۳۸۸) "مدیریت سرمایه گذاری" نشر نگاه دانش، ص ۴۸۰، ۴۷۹، ۶۹
- * راعی، رضا، تلنگی، احمد، (۱۳۸۳) "مدیریت سرمایه گذاری پیشرفته" انتشارات سمت، ص ۶۷، ۶۹، ۷۰.
- * Al Galib, A., Alam, M. and Rahman, R.M. (2014) 'Prediction of stock price based on hidden Markov model and nearest neighbour algorithm', Int. J. Information and Decision Sciences, Vol. 6, No. 3, pp.262-292.
- * Elaine Rich and Kevin Knight, (1991) "Artificial Intelligence", 2nd edition, McGraw-Hill, Inc., 1991.
- * JAROSLAV LAJOS, (2015) "COMPUTER MODELING USING HIDDEN MARKOV MODEL APPROACH APPLIED TO THE FINANCIAL MARKETS"



24. Forward-backward
25. Inside-outside
26. Hassan and Nath
27. JAROSLAV LAJOS
28. Maximum a Posteriori Hidden Markov Model
29. Maximum a Posteriori
30. Asadullah Al Galib
31. Posterior state probability
32. Maximum likelihood estimate
33. Local maximum
34. convergens
35. Global maximum
36. K-means clustering
37. Mean absolute percentage error(MAPE)
38. Tofallis