کشف تقلب در بازار بورس اوراق بهادار با استفاده از کاربرد نامساوی چبیشف

سید جواد کاظمی‌تبار* و مجید شهباززاده

دانشکده برق و کامپیوتر، دانشگاه صنعتی نوشیروانی بابل، بابل، ایران

چکیده

یکی از راه‌های مشارکت افراد در توسعه اقتصادی کشور، سرمایه‌گذاری در بازار سرمایه و به‌خصوص، بورس اوراق بهادار است. به این منظور، بازارهای اوراق بهادار باید مورد اعتماد مردم و فعالان اقتصادی باشند. شفافیت و کارایی بازار می‌تواند حقوق و منافع سرمایه‌گذاران را حمایت کند و باعث رونق بازار شود. در این میان، بعضی از افراد با توجه به موقعیت خود از اطلاعات نهانی مربوط به بازار بورس اوراق بهادار سوء استفاده می‌کنند و باعث بی‌اعتمادی افراد به بازار سرمایه می‌شود؛ ازاین‌رو در این مقاله، با استفاده از کاربرد نامساوی چبیشف، روشی برای شناسایی افرادی که از اطلاعات نهانی، استفاده شخصی کرده و در مدت کوتاهی سود کلانی به‌دست آورده‌اند، ارائه شده است. به‌منظور استفاده از این روش دو فیلتر در نظر گرفته شده است، به‌طوری‌که فیلتر نخست تراکنش‌های بزرگ را شناسایی می‌کند و فیلتر دوم، افرادی که بیشترین سود حاصل از خرید و فروش سهام در مدت زمان اندک (سه روز)، به‌دست آورده‌اند؛ درحالی‌که دست‌کم یک تراکنش بزرگ در این حد فاصل زمانی رخ داده باشد، شناسایی می‌کند؛ سپس روش پیشنهادی، بر روی دو دسته از داده‌های واقعی بازار بورس اعمال شده است. با تغییر ضرایب فیلترها، می‌توان معیارهای مورد نظر را تغییر داد.

واژگان کلیدی: بازار بورس اوراق بهادار، دست‌کاری بازار، پیش‌روی، کشف تقلب، نامساوی چبیشف

Stock Market Fraud Detection, A Probabilistic Approach

Seyed Javad Kazemitabar* & Majid Shahbazzadeh

Department of Electrical Engineering, Babol Noshirvani University of Technology, Babol, Iran

Abstract

In order to have a fair market condition, it is crucial that regulators continuously monitor the stock market for possible fraud and market manipulation. There are many types of fraudulent activities defined in this context. In our paper we will be focusing on "front running". According to Association of Certified Fraud Examiners, front running is a form of insider information and thus is very difficult to detect. Front running is committed by brokerage firm employees when they are informed of a customer's large transaction request that could potentially change the price by a substantial amount. The fraudster then places his own order before that of the customer to enjoy the low price. Once the customer's order is placed and the prices are increased he will sell his shares and makes profit. Detecting front running requires not only statistical analysis, but also domain knowledge and filtering. For example, the authors learned from Tehran's Over The Counter (OTC) stock exchange officials that fraudsters may use cover-up accounts to hide their identity. Or they could delay selling their shares to avoid suspicion.

Before being able to present the case to a prosecutor, the analyst needs to determine whether predication exists. Only then, can he start testing and interpreting the collected data. Due to large volume of daily trades, the analyst needs to rely on computer algorithms to reduce the suspicious list. One way to do this is by assigning a risk score to each transaction. In our work we build two filters that determine the risk of each transaction based on the amount of statistical abnormality. We use the Chebyshev inequality to determine anomalous transactions. In the first phase we focus on detecting a large transaction that changes market

* Corresponding author

*نویسندهٔ عهده‌دار مکاتبات

• تاریخ ارسال مقاله: ۱۳۹۷/۳/۱۲ • تاریخ پذیرش: ۱۳۹۸/۶/۱۱ • تاریخ انتشار: ۱۳۹۹/۰۴/۰۱ • نوع مطالعه: کاربردی

price significantly. We then look at transactions around it to find people who made profit as a consequence of that large transaction. We tested our method on two different stocks the data for which was kindly provided to us by Tehran Exchange Market. The officials confirmed we were able to detect the fraudster.

Keywords: Stock Exchange, Market manipulation, Front running, Fraud detection, Chebyshev inequality

۱- مقدمه

بر اساس بند ۳۲ ماده ۱ قانون بازار اوراق بهادار جمهوری اسلامی ایران، هرگونه اطلاعات افشانشده برای عموم که به‌طور مستقیم و یا غیر مستقیم به اوراق بهادار، معاملات یا ناشر آن مربوط می‌شود و درصورت انتشار بر قیمت و یا تصمیم سرمایه‌گذاران برای معامله اوراق بهادار مربوطه تأثیر می‌گذارد، اطلاعات نهانی گفته می‌شود. در بازار اوراق بهادار به‌دلیل فاصله بین مدیریت و مالکیت در شرکت‌های بزرگ سهامی، اداره‌کنندگان این شرکت‌ها در موقعیت اطلاعاتی برتر نسبت به سهامداران عادی و دیگر اعضای بازار اوراق بهادار قرار گرفته‌اند و ازاین‌رو از یک سری اطلاعات نهانی در مورد شرکت، آگاهی می‌یابند که ممکن است به یک معامله ناعادلانه منجر شود [1]. تأمین برابری در معاملات بورسی، به‌دلیل عدم تقارن اطلاعات طرفین، ممکن نیست؛ بنابراین نیاز است، دولت در این زمینه با وضع مقررات خاص، حمایت و قدرت‌های ناموزون را به تعادل رساند [2].

هدف از تدوین یک نظام جامع برای اطلاعات نهانی شرکت‌های تجاری در بازار بورس اوراق بهادار، برقراری مساوات بین سرمایه‌گذاران و رعایت انصاف در این معاملات است که از طریق ممنوعیت معامله افرادی است که به‌دلیل دست‌یابی اطلاعات لازم از چشم‌انداز شرکت، بیش از سرمایه‌گذاران معمولی باخبرند [3,4].

اعتماد سرمایه‌گذاران به بازار سرمایه، به تضمین و امنیتی بستگی دارد که آن‌ها را در موقعیتی مساوی قرار دهد و در مقابل استفاده نامناسب از اطلاعات نهانی یا محرمانه از آن‌ها حمایت کند [5]؛ بنابراین عدالت و کارآمدی اقتصادی زیربنای ممنوعیت معامله با استفاده از اطلاعات نهانی است [6]. گزارش‌های مالی متقلبانه، می‌تواند برای سازمان، عواقب ویران‌گری داشته باشد و اعتماد عمومی در بازار سرمایه را از بین ببرد [7].

قیمت‌گذاری صحیح اوراق بهادار هنگامی اتفاق می‌افتد که بر اساس نظام عرضه و تقاضای بازار تنظیم شود. این امر زمانی محقق می‌شود که همه اطلاعات مرتبط با اوراق بهادار برای عموم منتشر شود [8]؛ ازاین‌رو به واسطه کاهش ریسک و امنیت سرمایه‌گذاران، موجبات جذب سرمایه‌گذاری فراهم آورده می‌شود. بنابراین هدف از تنظیم این مقررات، جلوگیری استفاده از آن دسته اطلاعاتی است که افراد، آن‌ها را به‌جای به‌کارگیری مهارت و تلاش، از طریق نامشروع مثل سوء استفاده از موقعیت خود در شرکت به‌دست می‌آورند و موجب بی‌اعتمادی افراد به بازار سرمایه می‌شود [1].

کارگزار به فرد یا سازمانی گفته می‌شود که بین خریدار و فروشنده ارتباط لازم را به‌وجود می‌آورد و در ازای آن حق‌العمل (حق کمیسیون) دریافت می‌کند. کارگزاری‌ها وکیل و نماینده تام‌الاختیار برای خرید و فروش سهام هستند و بدون واسطه آنها هیچ‌کس نمی‌تواند در تالار بورس خرید و فروش کند.

از آنجایی‌که کلیه سرمایه‌گذاران از طریق شرکت‌های کارگزاری اقدام به خرید و فروش می‌کنند، بنابراین کارگزاران اطلاعاتی را در مورد تراکنش‌هایی که در آینده نزدیک انجام می‌شود در اختیار دارند. در این میان، افراد سودجو با این حدس که بازار بورس در آینده‌ای نزدیک تکان خواهد خورد، زودتر دست به خرید سهام آن ورقه بهادار خواهند زد، تا در آینده‌ای نزدیک که مبلغ هر سهم افزایش پیدا خواهد کرد، دست به فروش سهام خریداری شده، بزنند. در این صورت بدون هیچ فعالیت و زحمتی به سودهای قابل توجهی دست پیدا خواهند کرد، به این رخداد "پیش‌روی[1]" گفته می‌شود. وجه تسمیه آن، به این صورت است که در زمان‌های قدیم، کارگزاران باید به‌صورت دستی تقاضای سفارش خرید و فروش مشتریان را به بازار بورس می‌بردند. کارگزارانی که قصد تقلب داشتند، باید زودتر می‌دویدند تا قبل از کارگزاران دیگر بتوانند سفارش خود را پیش از سفارش مشتریان به جریان بگذارند. برای توضیح بیشتر به مثال زیر توجه کنید.

فرض کنید کارگزاری، سفارش بزرگی در حدود ۴۰۰۰۰۰ سهم از یک مشتری دریافت می‌کند، کارگزار با علم به اینکه پس از این تراکنش بزرگ، بازار تکان خواهد خورد و مبلغ هر سهم افزایش پیدا خواهد کرد، قبل از اعمال سفارش مشتری، حدود ۳۰۰۰۰ سهم مشابه را به مبلغ هر

[1] Front running

سهم ۱۰۰۰۰ تومان، برای حساب خود خریداری می‌کند، سپس سفارش خرید ۴۰۰۰۰۰ سهم مشتری را اجرا می‌کند. در این موقع بازار تکان خورده است. اگر افزایش مبلغ هر سهم را ۲۰۰۰ تومان در نظر بگیریم، هر سهم ۱۲۰۰۰ تومان می‌شود. کارگزار سهم‌هایش را حتی اگر با مبلغ ۱۱۷۵۰ هم برای فروش بگذارد، در مدت زمان کوتاهی، سود قابل توجهی خواهد کرد، که در قسمت زیر محاسبه شده است.

(۱۱۷۵۰-۱۰۰۰۰) × ۳۰۰۰۰=۵۲۵۰۰۰۰۰

فرآیند پیش‌روی، برای افراد سودجویی که به این اطلاعات دسترسی دارند، اغوابرانگیز است و انجام این کار به‌شدت غیرقانونی است و برخورد قضایی با متخلفان را به دنبال دارد.

هدف از این مقاله شناسایی افرادی است که از موقعیت خود در شرکت سوء استفاده کرده و به اطلاعات نهانی دست یافته و از آن برای منافع شخصی خود استفاده می‌کنند تا به سودهای کلان در ظرف مدت کوتاهی دست یابند. طبق مطالعات نویسندگان تا کنون مورد مشابهی در این زمینه به‌صورت رسمی منتشر نشده است. در این مقاله، با استفاده از به‌کارگیری نامساوی چبیشف[1]، معیار قابل سنجشی براش شناسایی داده‌های مشکوک به تقلب در بازار بورس اوراق بهادار، ارائه شده است. درواقع دو فیلتر در نظر گرفته می‌شود که با توجه به نامساوی چبیشف، دو معیار به‌دست می‌آید. این دو معیار مقادیر بزرگی می‌توانند باشند؛ در هر مرحله داده‌هایی که از معیارهای به‌دست‌آمده بزرگ‌تر باشند، مد نظر ما هستند. این فیلترها بر روی داده‌های واقعی بورس اوراق بهادار پیاده‌سازی و نتایج آن بررسی شده است. داده‌های مورد استفاده در این پژوهش از طرف شرکت فرابورس ایران در اختیار نویسنده مسؤول این مقاله قرار گرفته است. نویسندگان بدین وسیله از عزیزان بخش بازرسی فرابورس سپاس‌گزاری می‌کنند. داده‌ها مربوط به معاملات دو نماد معاملاتی هستند. در نماد شماره یک همه معاملات از فروردین تا دی‌ماه سال ۹۳ و در نماد شماره دو همه معاملات از مرداد ماه تا اسفند ۹۴ به نویسندگان داده‌شد. به نویسندگان گفته شد که در هر یک از این دو نماد تقلب از نوع پیش‌روی صورت گرفته است. وظیفه نویسندگان پیدا کردن زمان دقیق رخداد این تقلب بوده است.

ساختار ادامه پژوهش به شرح زیر است: در بخش ۲، مطالعات انجام‌شده در زمینه دست‌کاری سهام در بازار بورس اوراق بهادار به‌طور مختصر بررسی و در بخش ۳، معیارهای قابل سنجش جهت شناسایی داده‌های مشکوک به تقلب با استفاده از به‌کارگیری نامساوی چبیشف توضیح داده شده و در بخش ۴، دو مجموعه از داده‌های واقعی بازار بورس اوراق بهادار، مورد بررسی قرار گرفته و معیارهای قابل سنجش بر روی این داده‌ها اعمال و در بخش ۵، بحث و نتیجه‌گیری حاصل از پیاده‌سازی این معیار بر روی داده‌های مورد نظر بیان شده است.

۲- مطالعات انجام‌شده

در این قسمت به بررسی چند مورد از مطالعات انجام‌شده به‌طور مختصر می‌پردازیم:

یوسنگ کیم و همکاران، بر اساس روش تحلیل گروهی همتا[2] که روشی بدون نظارت است، به شناسایی الگوهایی مشکوک به دست‌کاری قیمت سهام، پرداخته‌اند. این روش، رفتار غیرطبیعی هدف را با مقایسه‌اش با گروه همتا و اندازه‌گیری انحراف رفتار هدف را در مقایسه با همتایان خود مشخص می‌کند. همچنین با استفاده از ترکیب وزن‌های اعضای گروه یکتا، راه‌حلی برای بهبود روش پیشنهادی ارائه شده است. در این پژوهش، تغییر غیرطبیعی قیمت سهام در بازار بورس کشور کره جنوبی بررسی شده است [9].

دیاز و همکاران، رویکرد "جعبه باز[3]" را در استفاده از روش‌های داده‌کاوی برای تشخیص دست‌کاری روزانه قیمت با توجه به متغیرهای مالی را مطرح کردند. با استفاده از رویکرد تحلیلی که در این مقاله شرح داده شده است، الگوی دستکاری جدید تقلب شناسایی شده است. ویژگی‌های کدگذاری‌شده به‌عنوان درخت تصمیم‌گیری[4] می‌تواند به‌راحتی در سامانه‌های تشخیص تقلب استفاده شود [10].

اوگات و همکارانش، برای شناسایی دست‌کاری بازار بورس اوراق بهادار کشور ترکیه، از شبکه‌های عصبی مصنوعی[5] (ANN) و ماشین بردار پشتیبان[6] (SVM) استفاده کرده‌اند. در این مقاله، میانگین بازده روزانه سهام دست‌کاری‌شده، میانگین بازده روزانه شاخص، میانگین تغییر روزانه در حجم معاملات و میانگین نوسانات روزانه، به‌عنوان متغیرهای مالی در نظر گرفته شده است. دقت، حساسیت و ویژگی‌های آماری طبقه‌بندی برای شبکه‌های عصبی مصنوعی و ماشین بردار پشتیبان با نتایج تجزیه و تحلیل

[1] Chebyshev

[2] Peer Group Analysis
[3] Open-Box
[4] Decision Tree
[5] Artificial Neural Network
[6] Support Vector Machine

تفکیک‌کننده[1] و رگرسیون لجستیکی[2] (logit)، مقایسه شده است. آن‌ها به این نتیجه رسیده‌اند که روش‌های داده‌کاوی (ANN و SVM) برای تشخیص دست‌کاری قیمت سهام از روش‌های آماری چندمتغیره (تجزیه و تحلیل تفکیک‌کننده و رگرسیون لجستیکی) مناسب‌تر است؛ به‌طوری‌که عملکرد روش‌های داده‌کاوی از نظر دقت طبقه‌بندی و حساسیت آماری بهبود یافته است [11].

پژوهش‌گران بازار اوراق بهادار، از نمودارها و تصاویر مختلف برای نظارت بر بازار استفاده می‌کنند. حساب‌رسان و تنظیم‌کنندگان بازار اوراق بهادار کشور کانادا، توجه زیادی به استفاده از روش‌های بصری در پیداکردن داده‌ها دارند؛ به‌طوری که با استفاده از این روش، الگوهایی را که در داده‌ها و یا سایر اطلاعات، قابل تشخیص نیست، می‌توان شناسایی کرد. روش‌های بصری به‌طرز قابل توجهی از جدول‌ها و نمودارها بهتر است [12].

تجسم سه‌بعدی، ابزاری برای نظارت بر جریان معاملات خام، قیمت و حجم است. در این روش، دادوستدهای بازار با مجموعه‌ای از پارامترها مقایسه و اگر آن‌ها خارج از محدوده باشند، هشدار داده می‌شود.

در تجسم سه‌بعدی، هر سلول یک ورقه بهادار را نشان می‌دهد؛ به‌طوری که اندازه یک سلول متناسب با سرمایه بازار است و رنگ یک سلول نشان‌دهنده تغییر قیمت آن است، به‌عنوان مثال رنگ سبز برای افزایش و رنگ قرمز برای کاهش قیمت به کار رفته است.

هوآنگ و همکاران، یک روش تجزیه و تحلیل بصری برای بازار سهام به نام استوکر[3] مطرح کردند که در آن از نقشه‌های درختی سه‌بعدی[4] برای نظارت بر بازار و از تحلیل شبکه‌های اجتماعی[5] برای نظارت بر فعالیت‌های کارگزاران استفاده شده است [13].

گل‌محمدی و همکاران در سال ۲۰۱۴، برای شناسایی دادوستدهای مشکوک در بازار سهام از روش‌های یادگیری به همراه ناظر استفاده کرده‌اند؛ ازاین‌رو بر روی داده‌های دست‌کاری‌شده سال ۲۰۰۳، الگوریتم‌های طبقه‌بندی و درخت‌های رگرسیونی[6]، درخت‌های استنباط شرطی[7]، سی۵،[8] جنگل تصادفی[9]، نایو بیز[10]، شبکه‌های عصبی، ماشین بردار پشتیبان و نزدیکترین همسایه[11] را پیاده‌سازی کرده‌اند. نتایج عملی نشان داده که الگوریتم "نایو بیز" با حساسیت ۰/۸۹، تشخیص ۰/۸۳ و دقت ۰/۸۳ با محاسبه معیار F ۰/۵۳، بهترین عملکرد را در میان الگوریتم‌های پیشنهادی داشته است [14].

راویسانکار و همکاران[12]، از روش‌های داده‌کاوی مانند شبکهٔ عصبی چندلایه پیش‌خور، ماشین بردار پشتیبان، برنامه‌ریزی ژنتیک، روش گروهی مدل‌سازی داده‌ها، رگرسیون لجستیک و شبکه عصبی احتمالی، برای شناسایی گزارش‌های مالی متقلبانه استفاده کردند. آنان هر یک از مدل‌های مزبور را برای ۲۰۲ شرکت چینی بررسی کردند. مدل‌های شبکه عصبی احتمالی و برنامه‌ریزی ژنتیک عملکرد بهتری نسبت به سایر مدل‌ها داشتند [15].

گلنسی و یاداف[13]، از رویکرد کمی‌سازی اطلاعات داده‌های متنی درخصوص کشف تقلب استفاده کردند. آنان با آزمون، بحث و تحلیل مدیریت، عملکرد مدل پیشنهادی خود را بررسی کردند. مدل پیشنهادی آنان عملکرد موفقی در خصوص تشخیص موارد متقلبانه داشت [16].

کاناپیکین و گروندین[14]، مدلی برای کشف تقلب در صورت‌های مالی به‌وسیله نسبت‌های مالی پیشنهاد کردند. آنان چهل صورت مالی متقلبانه و ۱۲۵ صورت مالی غیر متقلبانه را با استفاده از مدل رگرسیون لجستیک، مطالعه کردند. نتایج پژوهش آنان، کارایی مدل ارائه شده در پیش‌بینی گزارش مالی متقلبانه را تأیید کرده است [17].

لین[15] و همکاران، کارایی مدل‌های رگرسیون لجستیک، درخت تصمیم، شبکه عصبی مصنوعی را بررسی کردند. نتایج پژوهش آنان، کارایی مدل‌های شبکه عصبی و درخت تصمیم را در مقایسه با رگرسیون لجستیک تأیید می‌کند. نکته قابل توجه در پژوهش آنان، نرخ دسته‌بندی صحیح بیشتر از نود درصد شرکت‌های متقلب در کاربرد شبکه عصبی است [18].

اعتمادی و زلقی، به بررسی تقلب در صورت‌های مالی در شرکت‌های پذیرفته‌شده در بورس اوراق بهادار تهران با استفاده از برخی نسبت‌های مالی مرتبط پرداختند. نمونه آماری این پژوهش، شامل ۶۸ شرکت، در قالب ۳۴ شرکت

[1] Discriminant Analysis
[2] Logistics Regression
[3] Stockare
[4] 3D Treemap
[5] Social Network Analysis
[6] Classification And Regression Trees (CART)
[7] Conditional inference trees
[8] C5.0
[9] Random Forest
[10] Naïve Bayes
[11] K nearest neighbors (KNN)
[12] Ravisankar
[13] Glancy and Yadav
[14] Kanapickienė and Grundienė
[15] Lin

دارای نشانه‌های تقلب و ۳۴ شرکت فاقد نشانه‌های تقلب بوده است. آنان نُه نسبت مالی را به‌عنوان پیش‌بینی‌کننده های تقلب، انتخاب و از روش رگرسیون لجستیک برای تدوین مدل و شناسایی عوامل مرتبط با تقلب استفاده کردند. مدل مزبور، در طبقه‌بندی صحیح نمونه از نرخ دقت ۸۳/۸ درصد برخوردار بوده‌است [19].

فرقاندوست حقیقی، هاشمی و فروغی دهکردی، پژوهشی در زمینه ارتباط مدیریت سود و امکان تقلب در صورت‌های مالی انجام دادند. پژوهش آنان با استفاده از روش رگرسیون لجستیک، به این نتیجه دست یافت که در شرکت‌های با سابقه مدیریت سود، احتمال تقلب در صورت های مالی وجود دارد [20].

کاتسیس[1] و همکاران، عملکرد و کارایی روش‌های الگوریتم مورچگان، رگرسیون لجستیک، شبکه عصبی مصنوعی، درخت تصمیم‌گیری و شبکه گمانه‌زنی بیزی در کشف تقلب در صورت‌های مالی را بررسی کردند. الگوریتم مورچگان عملکرد موفق‌تری نسبت به دیگر روش‌ها داشت. رویکرد طبقه‌بندی و تابع برازش مورد استفاده در مدل مورچگان مبتنی بر مفهوم فاصله اقلیدسی بوده است [21].

تاپر[2]، طی پژوهشی با عنوان مسئولیت حسابرسان برای کشف خطاها و تقلب در صورت‌های مالی، اقدام به بررسی عوامل ایجاد ریسک مرتبط با وقوع تقلب در فعالیت های حسابرسی نمود. این عوامل در پنج طبقه اصلی شامل بدهی‌ها، مشتریان، صاحبان سهام، حسابهای سود و زیانی و موجودی کالا مورد تحلیل قرار گرفتند. پژوهش‌گر در این پژوهش به وظایف گوناگون حسابرسان در پیش‌گیری و کشف تقلب در قالب این پنج طبقه اشاره کرده است [22].

۳- روش پیشنهادی

همانطور که می‌دانیم روزانه تراکنش‌های زیادی در بازار بورس در حال انجام‌شدن است؛ ازاین‌رو اگر بخواهیم در طول چند هفته و یا چند ماه، داده‌های بورس را بررسی کنیم، با حجم بسیار بالایی از اطلاعات روبه‌رو هستیم.

با توجه به شکل (۱) و جدول (۱) مشاهده می‌شود، در توزیع نرمال، احتمال اینکه قدر مطلق داده‌های مورد بررسی بزرگ‌تر از $u+\sigma$ باشد، ۳۲٪ است، این احتمال برای بیشتر بودن از $u+2\sigma$، به ۵٪ و همچنین برای بزرگتربودن از $u+3\sigma$، فقط به ۰/۳٪ می‌رسد.

در نظریه احتمالات، نابرابری چبیشف این قضیه را بیان می‌کند که در هر توزیع، احتمال آن‌که برآمد یک متغیر تصادفی بیش از k برابر انحراف معیار با میانگین فاصله داشته باشد، حداکثر برابر با $\frac{1}{k^2}$ است.

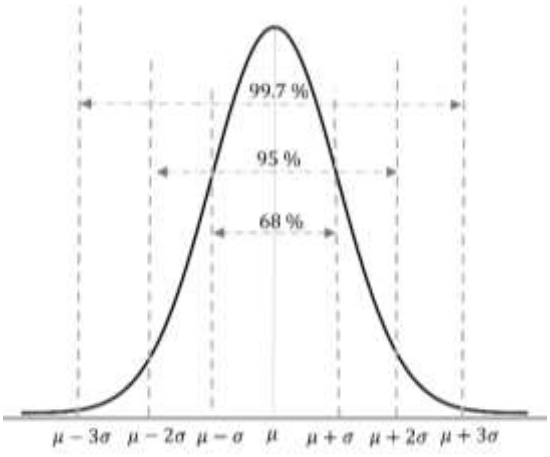
قضیه نامساوی چبیشف: اگر X متغیری تصادفی با مقدارمتوسط متناهی μ و واریانس σ^2 باشد، آن‌گاه برای هر $k>0$:

$$\Pr\{|X-\mu|\geq k\}\leq\frac{\sigma^2}{k^2} \qquad (۱)$$

این قضیه در مورد هر توزیع دلخواه صادق است. بنابراین فارغ از آن‌که توزیعی را برای داده‌های بورس در نظر بگیریم، به‌دنبال تراکنش‌های بزرگ می‌گردیم. برای به‌دست آوردن معیارهای قابل سنجش نیاز به استفاده از دو فیلتر داریم.

میانگین و انحراف معیار رابطه (۱)، در بخش ۳-۱ (فیلتر اول) و در بخش ۳-۲ (فیلتر دوم) به‌صورت جداگانه توضیح داده شده است. ابتدا به بحث در مورد فیلتر نخست می‌پردازیم.

الگوریتم روش پیشنهادی در شکل (۲) آورده شده است.



(شکل-۱): توزیع احتمال نرمال [23]

(Figure-1): Normal probability distribution [23]

(جدول-۱): جدول توزیع نرمال [23]

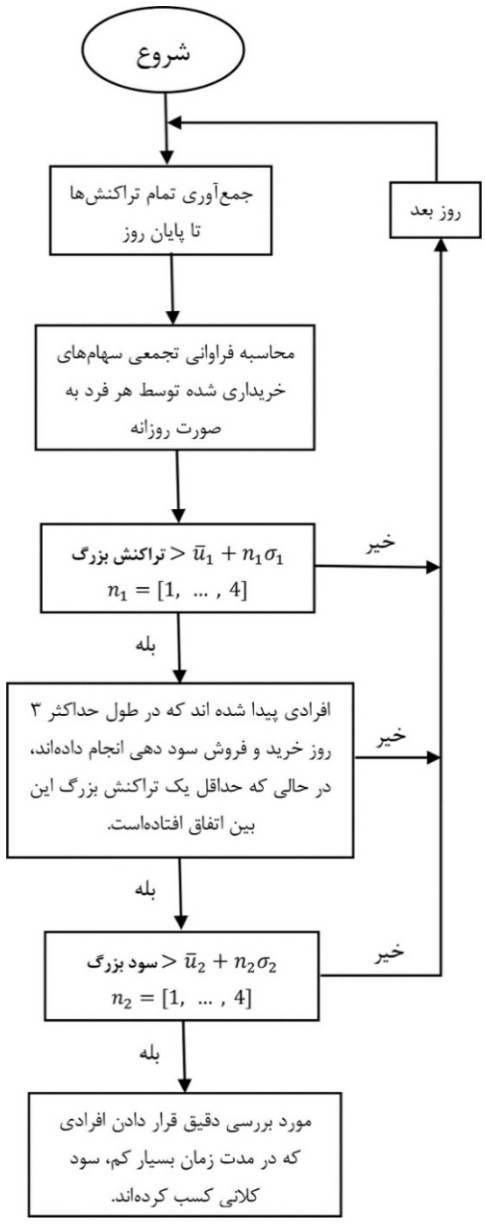
(Table-1): Normal distribution table [23]

n	$p=F(u+n\sigma)-F(u-n\sigma)$	$1-p$
1	0.68	0.32
2	0.95	0.05
3	0.997	0.003

[1] Katsis
[2] Topor

۳-۱- فیلتر نخست

برای ایجاد فیلتر نخست، تعداد سهام خریداری‌شده توسط هر فرد مورد بررسی قرار گرفته است. به این صورت که، تمام سهام خریده‌شده توسط هر فرد، به‌صورت روزانه جمع زده می‌شود؛ سپس اطلاعات مربوط به فراوانی‌های تجمعی سهام به‌دست‌آمده برای تمام افراد به‌دست آمده است. با استفاده از این فراوانی‌ها، معیار تعریف‌شده‌ای در نظر می‌گیریم که آن را معیار نخست می‌نامیم. فراوانی تجمعی سهام خریده‌شده توسط هر فرد، در روز جاری را با این معیار مقایسه می‌کنیم. اگر از معیار تعریف شده نخست، بزرگ‌تر باشد، آن را تراکنش بزرگ تعریف می‌کنیم. برای توضیح بیشتر، درواقع به شیوهٔ شکل (۳) عمل می‌کنیم.



(شکل-۲): نمودار جعبه‌ای روش پیشنهادی

(Figure-2): Block diagram of proposed approach

معیار تعریف‌شده نخست، برای شناسایی تراکنش‌های بزرگ به‌صورت $u_1 + n_1\sigma_1$ در نظر گرفته می‌شود. به‌طوری که u_1، میانگین فراوانی تجمعی‌های به‌دست‌آمده در روزهای قبل و σ_1، انحراف از معیار آن‌ها است.

روز	اسم خریدار	سهم
1	A	x_1
1	A	x_2
1	B	x_3
2	C	y_1
3	B	z_1
3	A	z_2
3	A	z_3

روز	اسم خریدار	سهم
1	A	$x_1 + x_2$
1	B	x_3
2	C	y_1
3	B	z_1
3	A	$z_2 + z_3$

(شکل-۳): توزیع خرید روزانه سهام

(Figure-3): Daily stock purchase distribution

طبیعتاً هرچه ضریب n_1 را بیشتر در نظر بگیریم، معیار مقایسه، بزرگ‌تر می‌شود و تعداد تراکنش‌های کمتری در روز جاری به‌دست می‌آیند که از این معیار بزرگ‌تر باشد. درواقع این تراکنش بزرگ، همان سفارش بزرگی است که توسط یک خریدار به کارگزار ارائه شده است و پس از اعمال آن بازار تکان خواهد خورد.

تا این مرحله، سفارش‌های بزرگ شناسایی شده است؛ سپس باید بررسی کنیم، که آیا فرد یا افرادی وجود دارند که در مدت سه روز، خرید و فروش سوددهی انجام داده‌اند، درحالی‌که دست‌کم یک سفارش بزرگ در این حد فاصله زمانی رخ داده باشد یا خیر؟

پس از بررسی، یک سری افراد با مشخصات یادشده به‌دست می‌آیند؛ اما باید، فیلتر دیگری را در نظر گرفت تا افرادی که از معیار مورد نظر دوم، درصد سود بیشتری کسب کرده‌اند، مشخص کند.

۳-۲- فیلتر دوم

برای ایجاد فیلتر دوم، درصدهای سود به‌دست‌آمده مورد بررسی قرار گرفته است، به طوری که این درصدها از طریق رابطه (۲) محاسبه می‌شوند:

$$\text{درصد سود} = \frac{(o_1 \times t_1) - (o_2 \times t_2)}{(o_2 \times t_2)} \times 100 \quad (۲)$$

o_1= تعداد فروش سهام

t_1= قیمت فروش سهام

o_2= تعداد خرید سهام

t_2= قیمت خرید سهام

بنابراین، برای هر خریدار، درصد سود مختص به خود به‌دست آورده می‌شود.

در این مرحله، معیار دوم را بر روی درصدهای سود به‌دست آمده در نظر می‌گیریم که به‌صورت $u_2 + n_2\sigma_2$، تعریف می‌شود. به طوری که u_2 میانگین درصد سودهای به دست آمده و σ_2 انحراف معیار آن‌ها است.

با اعمال معیار دوم، قصد داریم افرادی که بیشترین سود حاصل از خرید و فروش سهام را در مدت سه روز به دست آورده‌اند، درحالی‌که در این مدت دست‌کم یک سفارش بزرگ وجود داشته باشد، را شناسایی کنیم. اگر درصد سود فردی از معیار دوم در نظر گرفته‌شده بیشتر باشد، فرد مورد نظر باید بررسی و سوابق فرد باید مطالعه شود.

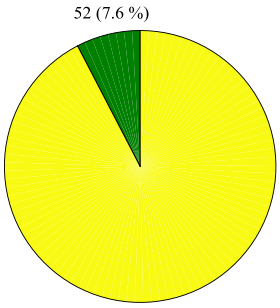
۴- پیاده‌سازی و ارزیابی نتایج

در این مقاله داده‌های مربوط به تراکنش‌های بورس در مدت ده ماه، برای یک سهام ورقه بهادار خاص مورد بررسی قرار گرفته است. این داده‌ها شامل ۲۳۲۲ سطر اطلاعات هستند که زمان تراکنش، نام خریدار، نام فروشنده، تعداد سهم و قیمت تراکنش را شامل می‌شوند.

طبق روش توضیح داده‌شده در بخش (۳-۱)، ابتدا مجموع تمام سهم‌های هر خریدار، به‌صورت روزانه محاسبه می‌شود؛ بعد از اعمال این قسمت، تعداد سطرهای داده‌های بورس به ۶۸۱ سطر کاهش می‌یابد، به این معناست که سطرهایی که زمان تراکنش و نام خریدار یکسانی دارند، در هم ادغام شده‌اند و سهام خریداری‌شده آن‌ها با هم جمع شده‌است.

دو پارامتر n_1 و n_2 تأثیر به‌سزایی در عملکرد فیلترهای تعریف شده دارند، به‌طوری که n_1 بر فیلتر سنجش تراکنش بزرگ تأثیر می‌گذارد و n_2 در شناسایی افرادی که از خرید و فروش سهام، بیش‌ترین سود را برده‌اند، تأثیرگزار است.

در آزمایش نخست، $n_1 = 1$در نظر گرفته می‌شود. به این معناست که تراکنش‌هایی که از $u_1 + \sigma_1$ بزرگتر باشد، به‌عنوان تراکنش بزرگ در نظر گرفته می‌شود.

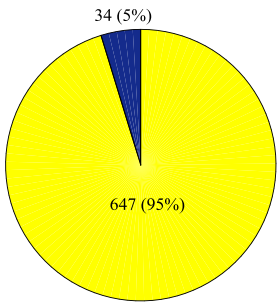


(شکل-۴): تراکنش‌های بزرگ به‌ازای $n_1 = 1$

(Figure-4): Large transaction for $n_1 = 1$

همانطور که در شکل (۴) دیده می‌شود، به‌ازای $n_1 = 1$، ۵۲ تراکنش بزرگ مشاهده می‌شود، درحالی که کل دایره شامل ۶۸۱ تراکنش است.

تعداد افرادی که در بازه سه روزه خرید و فروش انجام داده و همچنین سود کرده‌اند، درحالی‌که تراکنش یا تراکنش‌های بزرگی با معیار تعریف‌شده در طول این مدت، رخ داده باشد، در شکل (۵) مشاهده می‌شود. همان‌طور که مشاهده می‌شود در شکل (۵)، ۳۴ فرد با ویژگی‌های یادشده مشاهده می‌شود؛ اما تمام این افراد، سود هنگفتی نکرده‌اند. بنابراین باید فیلتر دوم را به‌کار گرفت، به این صورت افرادی را که بیش از معیار $u_2 + n_2\sigma_2$ سود کرده‌اند، شناسایی کنیم.



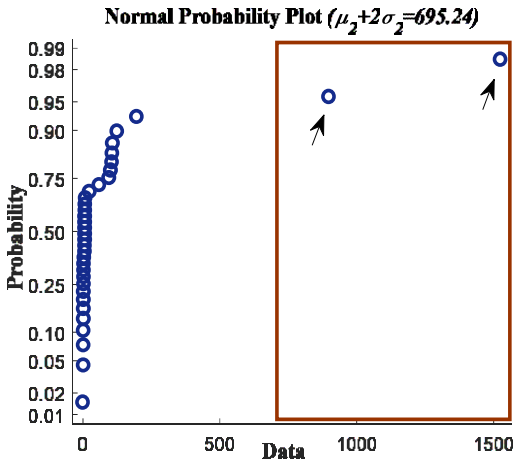
(شکل-۵): افرادی که تراکنش بزرگ بین خرید و فروش سود ده آن‌ها در طول سه روز رخ داده‌است، به‌ازای $n_1 = 1$

(Figure-5): People with a large and profitable transaction for whom buying and selling occurs within 3 days for $n_1 = 1$

ازاین‌رو با درنظرگرفتن $n_2 = 2$، معیار دوم به‌صورت زیر به‌دست می‌آید:

$$u_2 + 2\sigma_2 = 695.24$$

بنابراین نتایج به‌صورت زیر حاصل خواهد شد:



(شکل-۶): شناسایی نهایی افراد پس از فیلتر دوم به‌ازای $n_2 = 2$

(Figure-6): Final detection after second filter for $n_2 = 2$

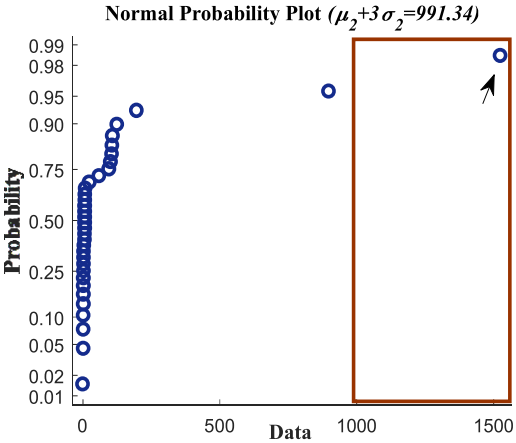
همان‌طور که مشاهده می‌شود، دو فرد شناسایی شد که سودی بیشتر از معیار تعریف‌شده کسب کرده‌اند.

ناحیه قهوه‌ای‌رنگ، محدوده‌ای است که افراد مورد نظر، از معیار تعریف‌شده سود بیشتری کرده‌اند و دایره‌های آبی‌رنگ، همان ۳۴ نفری هستند که پس از اعمال فیلتر نخست مشخص شدند.

اکنون اگر $n_2 = 3$ در نظر گرفته شود، طبیعتاً باید افراد کمتری را شناسایی کنیم که از معیار زیر سود بیشتری کرده باشند:

$$u_2 + 3\sigma_2 = 991.34$$

پس از اعمال فیلتر، نتایج به‌صورت زیر حاصل می‌شود:



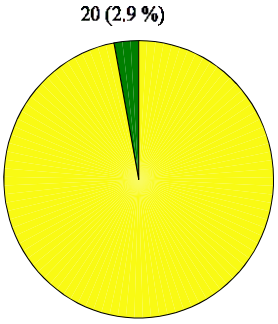
(شکل-۷): شناسایی نهایی افراد پس از فیلتر دوم به ازای $n_2 = 3$

(Figure-7): Final detection after second filter for $n_2 = 3$

همانطور که دیده می‌شود، تنها یک نفر شناسایی می‌شود که از معیار تعریف‌شده بیشتر سود کرده باشد.

در آزمایش دوم، درصورتی‌که $n_1 = 2$ در نظر گرفته شود، طبیعتاً تراکنش‌های بزرگ کمتری نسبت به حالت قبل به‌دست می‌آید.

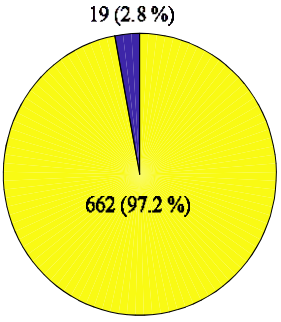
همان‌طورکه در شکل (۸) مشاهده می‌شود، فقط بیست تراکنش بزرگ در ۶۸۱ تراکنش مشاهده می‌شود که از معیار در نظر گرفته‌شده، به‌ازای $n_1 = 2$ بزرگ‌تر است، این درحالی است که به‌ازای $n_1 = 1$، ۵۲ تراکنش بزرگ مشاهده می‌شد.



(شکل-۸): تراکنش‌های بزرگ به‌ازای $n_1 = 2$

(Figure-8): Large transaction for $n_1 = 2$

با توجه به شکل (۹)، ۱۹ فرد پیدا می‌شود، که تراکنش یا تراکنش‌های بزرگی بین خرید و فروش آن‌ها در طول سه روز اتفاق افتاده و این دادوستد برایشان سودده بوده است.

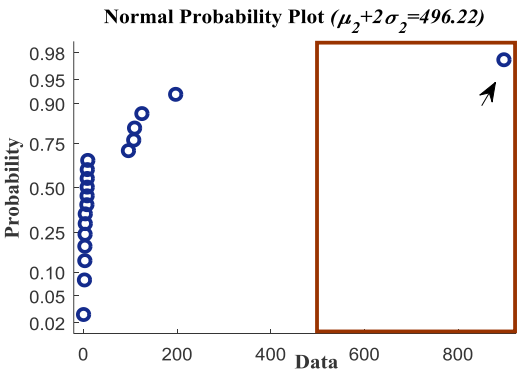


(شکل-۹): افرادی که تراکنش بزرگ بین خرید و فروش سود ده آن‌ها در طول سه روز رخ داده‌است، به‌ازای $n_1 = 2$

(Figure-9): People with a large and profitable transaction for whom buying and selling occurs within 3 days for $n_1 = 2$

با اعمال فیلتر دوم، به‌ازای $n_2 = 2$، فقط یک فرد مشاهده می‌شود که سود حاصل از خرید و فروش سهامش با اختلاف از معیار زیر بیشتر بوده است.

$$u_2 + 2\sigma_2 = 496.22$$

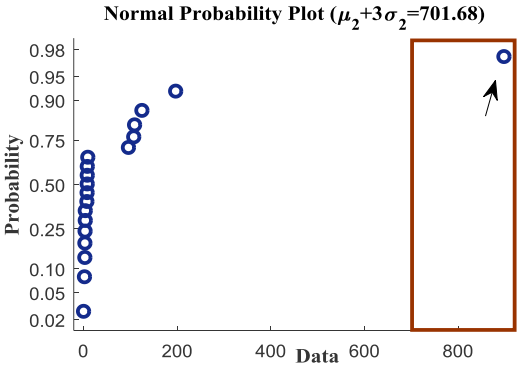


(شکل-۱۰): شناسایی نهایی افراد پس از فیلتر دوم به ازای $n_2 = 2$

(Figure-10): Final detection after second filter for $n_2 = 2$

اکنون n_2 را افزایش می‌دهیم و برای $n_2 = 3$ قصد داریم افراد نهایی را شناسایی کنیم. معیار تعریف‌شده به صورت زیر است:

$$u_2 + 3\sigma_2 = 701.68$$



(شکل-۱۱): شناسایی نهایی افراد پس از فیلتر دوم به‌ازای $n_2 = 3$

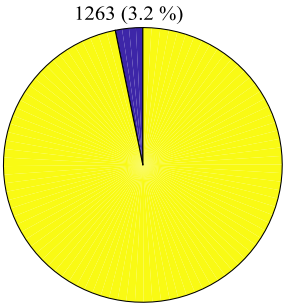
(Figure-11): Final detection after second filter for $n_2 = 3$

همان‌طور که دیده می‌شود، برای $n_2 = 3$ در شکل (۱۱)، همانند $n_2 = 2$ در شکل (۱۰)، فقط یک نفر شناسایی شده است. به این دلیل است، که درصد سود فرد مورد نظر از بقیه داده‌ها فاصله بسیار زیادی دارد.

برای بررسی یک مطالعه موردی دیگر، تراکنش‌های مربوط به سهام یک ورقه بهادار خاص دیگری در مدت هشت ماه مورد بررسی قرار داده شده است.

این داده‌ها شامل ۱۵۱۳۷۵ سطر اطلاعات است که پس از محاسبه فراوانی‌های تجمعی سهام هر خریدار، به صورت روزانه، داده‌های مورد نظر به ۳۹۷۴۸ سطر می‌یابد.

چون با تعداد اطلاعات بالایی سر و کار داریم، معیار تراکنش بزرگ را این دفعه به‌صورت $u_1 + 4\sigma_1$ در نظر می‌گیریم. افرادی که بین خرید و فروش سهامشان، تراکنش یا تراکنش‌های بزرگی با معیار تعریف‌شده، رخ داده است و سود حاصل از این داد و ستد بیشتر از ۰/۵ برابر قیمت سهام خریده شده باشد، ۱۲۴۳ تراکنش است که در شکل (۱۲) دیده می‌شود. به این دلیل ۰/۵ برابر قیمت سهام خریده، در نظر گرفته شده است؛ چون با تعداد بالایی داده روبه‌رو هستیم و خیلی از افراد درصد سود کمی کسب کرده‌اند، بنابراین می‌توان آن داده‌ها را نادیده گرفت.

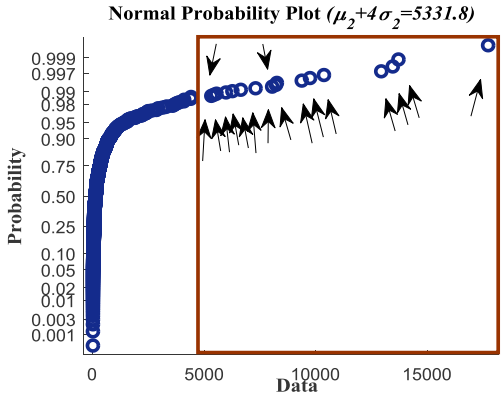


(شکل-۱۲): افرادی که تراکنش بزرگ بین خرید و فروش سود ده آن‌ها در طول سه روز رخ داده است، به‌ازای $n_1 = 4$

(Figure-12): People with a large and profitable transaction for whom buying and selling occurs within 3 days for $n_1 = 4$

با توجه به حجم بالای داده‌ها، فیلتر دوم به‌ازای $n_2 = 4$، در نظر گرفته شده است. به‌طوری‌که معیار سنجش فیلتر دوم، به‌صورت زیر تعریف می‌شود:

$$u_2 + 4\sigma_2 = 5331.8$$



(شکل-۱۳): شناسایی نهایی افراد پس از فیلتر دوم به‌ازای $n_2 = 4$

(Figure-13): Final detection after second filter for $n_2 = 4$

با توجه به شکل (۱۳)، در بین ۳۹۷۴۸ تراکنش، فقط هفده فرد پیدا می‌شوند که در عرض سه روز، خرید و فروش انجام داده‌اند و در عین حال سود قابل توجهی هم کسب کرده‌اند و همچنین تراکنش بزرگی بین این داد و ستد اتفاق افتاده باشد. افرادی که پس از اعمال این فیلترها مشخص می‌شوند، در صورت لزوم می‌توان مورد بررسی قرار داد.

در این مقاله، به‌منظور کشف تقلب در بازار بورس و شناسایی رخداد پیش‌روی، معیار قابل سنجشی با استفاده از به‌کارگیری نامساوی چبیشف ارائه شده است. با استفاده از قضیه چبیشف، فارغ از آن که توزیعی برای داده‌های بازار بورس در نظر گفته شود، به‌دنبال پیداکردن داده‌های‌های بزرگ هستیم. روش پیشنهادی، با استفاده از دو فیلتر معرفی شده است، به‌طوری که فیلتر نخست، معیار قابل سنجش برای شناسایی تراکنش بزرگ را محاسبه می‌کند و فیلتر دوم معیار قابل سنجش برای شناسایی افرادی است که بیشترین سود حاصل از خرید و فروش سهام در مدت زمان اندک، به‌دست آورده‌اند، درحالی‌که دست‌کم یک تراکنش بزرگ در این مدت رخ داده باشد. مدت زمان اندک در این مقاله، سه روز در نظر گرفته شده و دلیل تعریفش، آن است که احتمال این‌که فردی در این مدت زمان کم، سود کلانی را کسب کند، بسیار کم است، مگر این‌که از تراکنش‌های آینده نزدیک بازار، باخبر باشد.

۵- بحث و نتیجه‌گیری

در این مقاله روش پیشنهادی بر روی دو دسته از داده‌های بازار بورس اعمال شده است. برای دسته نخست که شامل ۲۳۲۲ سطر اطلاعات می‌باشد. ابتدا معیار تراکنش بزرگ به‌صورت $u_1 + \sigma_1$، در نظر گرفته شده است. مشاهده می‌شود ۵۲ تراکنش بزرگ رخ داده است، به‌طوری‌که ۳۴ نفر این تراکنش بزرگ بین خرید و فروش سودده آن‌ها که در مدت سه روز اتفاق افتاده است. معیار دوم به‌صورت $u_2 + 2\sigma_2$ بر روی درصدهای سود ۳۴ نفر اعمال شده است، که پس از اعمال فیلتر دوم، همان‌طورکه در شکل (۶) مشاهده می‌شود، دو فرد شناسایی شده است که از معیار دوم درنظر گرفته‌شده، سود بیشتری کرده‌اند؛ سپس معیار دوم را افزایش می‌دهیم و $u_2 + 3\sigma_2$ را بر روی درصدهای سود همان ۳۴ نفر اعمال می‌کنیم؛ همان‌طورکه در شکل (۷) مشاهده می‌شود، تعداد افراد شناسایی‌شده به یک نفر کاهش یافته است. در آزمایشی دیگر، معیار تراکنش بزرگ را برابر با $u_1 + 2\sigma_1$ قرار می‌دهیم. همان‌طورکه در شکل‌های (۸) و (۹) مشاهده می‌شود، بیست تراکنش بزرگ وجود دارد که نوزده نفر این تراکنش بزرگ بین خرید و فروش سودده آن‌ها که در مدت سه روز انجام شده، رخ داده است. با اعمال فیلتر دوم برای معیارهای $u_2 + 2\sigma_2$ و $u_2 + 3\sigma_2$، درنهایت فقط یک نفر شناسایی می‌شود که سود کلانی کسب کرده است.

به‌ازای $n_2 = 2$، وقتی n_1 را از ۱ به ۲ افزایش می‌دهیم، همان‌طورکه به‌ترتیب در شکل‌های (۶) و (۱۰) مشاهده می‌شود، دایره‌های آبی‌رنگ کاهش یافته است. به این معناست که معیار سنجش تراکنش بزرگ، بزرگتر شده است؛ ازاین‌رو تعداد افراد کم‌تری پیدا می‌شوند که در مدت سه روز خرید و فروش کرده‌اند و در عین حال سود هم کرده باشند. همان‌طورکه دیده می‌شود، به‌ازای $n_1 = 1$، ۳۴ نفر شناسایی شد، ولی به‌ازای $n_1 = 2$، این تعداد افراد به ۱۹ نفر کاهش یافت، درحقیقت دایره‌های آبی‌رنگ شکل (۱۰)، زیرمجموعه دایره‌های آبی‌رنگ شکل (۶) است، فقط به‌دلیل افزایش معیار سنجش تراکنش بزرگ، تعداد افراد مشخص شده کمتر شده است، ازاین‌رو، حتی فردی که در شکل (۶)، بیشترین سود (حدود ۱۵۰۰٪) را کسب کرده است، در شکل (۱۰) مشاهده نمی‌شود، به این دلیل است، که تراکنش بزرگ انجام‌شده بین خرید و فروش این فرد، از $u_1 + \sigma_1$ بزرگ‌تر بوده و از $u_1 + 2\sigma_1$ کوچک‌تر بوده است، بنابراین به‌ازای $n_1 = 2$، فرد مورد نظر شناسایی نشده است.

برای دسته دوم از داده‌های واقعی بورس که شامل ۱۵۱۳۷۵ سطر اطلاعات است. معیار نخست و دوم به‌صورت $u_1 + 4\sigma_1$ و $u_2 + 4\sigma_2$، در نظر گرفته شده است، که در شکل (۱۳) مشاهده می‌شود، فقط هفده نفر شناسایی شده‌اند که از معیار تعریف‌شده سود بیشتری کسب کرده‌اند.

نتایج این پژوهش با واقعیت انطباق دارد. اما آنچه که به چشم می‌خورد این است که مانند هر روش بدون نظارت وجود هشدارهای غلط در روش یادشده مشهود است. از آنجایی که روش پیشنهادی در این مقاله فقط به کل جامعه آماری نگاه می‌کند، یک راه‌حل برای کاهش تعداد هشدارهای غلط این است که از تحلیل رفتاری فردی کاربران استفاده شود. برای مثال می‌توان برای هر کاربر از روی رفتار گذشته وی فراوانی انجام معاملات و یا الگوی اندازه ریالی معاملات را به‌دست آورد؛ سپس می‌توان با استفاده از معیارهای تعیین فاصله رفتاری، میزان تغییر رفتار را در تراکنش‌های آینده به‌دست آورد و با تعیین آستانه صحیح نرخ هشدارهای غلط را کاهش داد.

۶- مراجع 6- References

[۱] جنیدی لعیا ، نوروزی محمد ، "شناخت ماهیت اطلاعات نهانی در بورس اوراق بهادار"، *فصلنامه حقوق*، دوره ۳۹، شماره ۳ ، ۱۳۸۸.

[1] L. Joneidi, M. Norouzi, “Recognizing hidden information in stock market,” *The Law Quarterly,* vol. 39, no. 3, 2009.

[۲] کاتوزیان ناصر. "نظریه عمومی تعهدات"، تهران، میزان. ۱۳۹۱.

[2] N. Katouzian, “General Theory of Commitment”, Mizan Publications, Tehran 2012.

[3] M. E. Roszkowski, "Business law: Principles, cases and policy", Newyork. Addison Wesley. 1997.

[4] S. M. Bainbridge, "Insider treading", Newyork. Foundation. 1999.

[۵] شریفی سید الهام الدین. "بررسی تطبیقی جنبه‌های حقوقی معاملات در بازار بورس با استفاده از اطلاعات محرمانه". پژوهش‌های حقوقی، شماره ۶. ۱۳۸۳.

[5] S. E. Sharifi, “Investigating legal aspects of stock market transactions via private information”, *The Journal of Legal Research*, no. 6, 2004.

[6] H. E. Leland, "Insider treading: should it be prohibited?” *The journal of political economy*, 1992.

[۷] داداشی ایمان، کردمنجیری سجاد، مریم برادران، "تأثیر ساختار حسابرسی داخلی بر احتمال تقلب در صورت های مالی شرکت‌های پذیرفته شده در بورس اوراق بهادار تهران". دانش حسابرسی. دوره ۱۸. شماره ۷۰. ۱۳۹۷.

[7] I. Dadashi, S. Kord-Monjiri, M. Baradaran, “The impact of internal auditing on Financial Statement Fraud in companies registered in Tehran stock market”, The Auditing Science, vol. 18, no. 70, 2018.

[۸] انصاری علی، حیدری سورشجانی مریم، "تحلیل حقوقی تأیید بورس در معاملات اوراق بهادار". فصلنامه بورس اوراق بهادار. شماره ۲۷ سال هفتم. ۱۳۹۳.

[8] A. Ansari, M. H. Sourashjani, “Legal analysis of stock market transactions”, Stock Market Quarterly, no. 27, 2014.

[9] Y. Kim, and S. Y. Sohn, "Stock fraud detection using peer group analysis," *Expert Systems with Applications.* vol. 39, no. 10, pp. 8986-92, 2012.

[10] D. Diaz, B. Theodoulidis, and P. Sampaio, "Analysis of stock market manipulations using knowledge discovery techniques applied to intraday trade prices," *Expert Systems with Applications,* vol. 38, no. 10, pp. 12757-71, 2011.

[11] H. Öğüt, M. M. Doğanay, and R. Aktaş, "Detecting stock-price manipulation in an emerging market: The case of Turkey," *Expert Systems with Applications.* Vol. 36, no. 9, pp. 11944-49, 2009.

[12] K. Golmohammadi, and O. R. Zaiane, "Data mining applications for fraud detection in securities market", *In Intelligence and Security Informatics Conference (EISIC),* pp. 107-114, 2012.

[13] M. L. Huang, J. Liang, and Q. V. Nguyen, "A visualization approach for frauds detection in financial market. In Information Visualisation", *2009 13th International Conference. IEEE*, pp. 197-202,

[14] K. Golmohammadi, O. R. Zaiane, D. Díaz, "Detecting stock market manipulation using supervised learning algorithms," *In2014 International Conference on Data Science and Advanced Analytics (DSAA),* pp. 435-441, 2014.

[15] P. Ravisankar, V. Ravi, G. R. Rao, and I. Bose, "Detection of financial statement fraud and feature selection using data mining techniques," *Decision Support Systems*, *50*(2), pp. 491-500, 2011.

[16] F. H. Glancy, and S. B. Yadav, "A computational model for financial reporting fraud detection," *Decision Support Systems*, no.50(3), pp. 595-601, 2011.

[17] R. Kanapickienė, and Ž. Grundienė, "The model of fraud detection in financial statements by means of financial ratios", *Procedia: Social and Behavioral Sciences*, *213*, pp. 321-327, 2015.

[18] C. C. Lin, A. A. Chiu, S. Y. Huang, and D. C. Yen, "Detecting the financial statement fraud: The analysis of the differences between data mining techniques and experts’ judgments".- *Knowledge-Based Systems*, no. *89*, pp.459-470, 2015.

[۱۹] اعتمادی حسین، زلقی حسن، "کاربرد رگرسیون لجستیک درشناسایی گزارشگری مالی متقلبانه". دانش حسابرسی. دوره ۱۳. شماره ۵۱. ۱۳۹۲.

[19] H. Etemadi, H. Zolfi, “Application of logistic regression in financial statement fraud detection”, The Auditing Science, vol. 13, no. 51, 2013.

[20] فرقاندوست حقیقی، کامبیز، هاشمی، عباس و فروغی دهکردی، امین، "مطالعه رابطه مدیریت سود و امکان تقلب در صورت‌های مالی شرکت‌های پذیرفته شده در بورس اوراق بهادار تهران"، مجله دانش حسابرسی، دوره ۴، شماره ۲۱، ص ۶۸-۴۷، ۱۳۹۳.

[20] K. F. Haghighi, A. Hashemi, A. F. Dehkordi, “A study of correlation between profit management and financial statement fraud in the companies registered in Tehran stock market”, The Auditing Science, vol. 4, no. 21, pp. 47-68, 2014.

[21] C. D. Katsis, Y. Goletsis, P. V. Boufounou, G. Stylios, and E. Koumanakos, "Using ants to detect fraudulent financial statements", *Journal of Applied Finance and Banking*, no.2(6), pp. 73, 2012.

[22] D. I. Topor, "The Auditor's Responsibility for Finding Errors and Fraud from Financial Situations: Case Study", *International Journal of Academic Research in Accounting, Finance and Management Sciences*, no.7(1), pp. 342-352, 2017.

[23] B. W. Silverman, "*Density estimation for statistics and data analysis*," Routledge, 2018.

سیدجواد کاظمی تبار دوره کارشناسی خود را در سال ۱۳۸۲ در دانشگاه صنعتی شریف به پایان رسانده است. در سال ۱۳۸۷ وی مدرک دکترای خود را در رشته مخابرات در دانشگاه کالیفرنیا در شهر ارواین کسب کرد و تا سال ۱۳۹۳ در شرکت‌های مختلف مهندسی در آمریکا به فعالیت پرداخت. از جمله از سال ۱۳۹۱ تا ۱۳۹۳ در سیلیوکون ولی به‌عنوان متخصص داده‌کاوی در شرکت گاردین آنالیتیکس به کشف تقلب‌های بانکی کمک می‌کرد. از سال ۱۳۹۴ وی هم‌اکنون عضو هیأت علمی دانشگاه صنعتی نوشیروانی بابل است.

نشانی رایانامه ایشان عبارت است از:

j.kazemitabar@nit.ac.ir

مجید شهباززاده تحصیلات خود را در مقاطع کارشناسی برق-مخابرات و کارشناسی ارشد برق-کنترل به‌ترتیب در سال‌های ۱۳۹۴ و ۱۳۹۶ در دانشگاه صنعتی نوشیروانی بابل به پایان رسانده و درحال حاضر دانشجوی مقطع دکترای برق-کنترل در دانشگاه صنعتی نوشیروانی بابل است. زمینه‌های پژوهشی مورد علاقه وی کنترل مقاوم، کنترل پیش‌بین و رباتیک است.

نشانی رایانامه ایشان عبارت است از:

m.shahbazzadeh@stu.nit.ac.ir