



فصلنامه‌ی اقتصاد مقداری

صفحه ی اصلی وب سایت مجله:
www.iqe.scu.ac.ir
شاپا الکترونیکی: 2717-4271
شاپا چاپی: 2008-5850



**خوشه‌بندی فازی سری‌های زمانی مالی بر اساس سرریزهای نوسانات جهت‌دار:
شواهدی از سهام شرکت‌های پذیرفته‌شده در بورس تهران**
الهام فرزنگان*
* استادیار اقتصاد، گروه علوم اقتصادی و اجتماعی، دانشگاه بوعلی سینا- مجتمع آموزش عالی نهاوند (ویژه دختران)، همدان، ایران
(نویسنده‌ی مسئول).

اطلاعات مقاله	طبقه‌بندی JEL: C61, C34, C12
تاریخ دریافت: 7 شهریور 1395 تاریخ بازنگری: 8 اسفند 1395 تاریخ پذیرش: 13 فروردین 1396 انتشار آنلاین از تاریخ 22 مرداد 1400	واژگان کلیدی: تقسیم‌بندی فازی اطراف مدویدها، سرریزهای نوسانات جهت‌دار، معیار عدم مشابهت وزنی، فاصله نمایی، بورس تهران
ارتباط با نویسنده مسئول: ایمیل: e.farzanejan@basu.ac.ir 0000300237253189	آدرس پستی: همدان، چهارباغ شهید مصطفی احمدی روشن، دانشگاه بوعلی سینا- مجتمع آموزش عالی نهاوند (ویژه دختران)، کد پستی: 65178-38695

اطلاعات تکمیلی:

قدردانی: از داوران محترم بابت پیشنهادهای ارزشمند که موجب بهبود کیفیت این تحقیق شده است، قدردانی می‌نمایم.
تضاد منافع: نویسنده مقاله اعلام می‌کند که در انتشار مقاله ارائه شده تضاد منافی وجود ندارد.
منابع مالی: نویسنده هیچ‌گونه حمایت مالی برای تحقیق، تألیف و انتشار این مقاله دریافت نکرده‌اند.

چکیده

در این مقاله یک چارچوب برای خوشه‌بندی فازی سری‌های زمانی مالی معرفی می‌شود که بر مبنای سرریزهای نوسانات جهت‌دار قرار دارد. کشف خوشه‌ها در سرریزهای نوسانات بین سری‌های زمانی مالی بیش عمیقی در رابطه با ساختار بازار فراهم می‌کند که می‌تواند علاوه بر مدبران پرتفوی، برای سیاست‌گذاران نیز مفید باشد. بدین منظور، سرریزهای نوسانات جهت‌دار - "از سوی" و "بیسوی" سایر سهامها- بر اساس متدولوژی تجزیه واریانس خطای پیش‌بینی تعمیم‌یافته (GFEVD) اندازه گرفته می‌شوند. بعلاوه، برای اندازه‌گیری عدم مشابهت‌ها از فاصله اقلیدسی وزندار تعمیم‌یافته نسبت به سرریزهای جهت‌دار، استفاده می‌شود. سپس، با اتخاذ یک چارچوب فازی مبتنی بر الگوریتم تقسیم‌بندی اطراف مدویدها، تحت عنوان مدل خوشه‌بندی C-مدویدهای فازی مبتنی بر سرریز نوسانات (VS-FCMdd)، وزن‌های بهینه تعیین و سهامها نیز برحسب درجه عضویت گروه‌بندی می‌شوند. بعلاوه، به‌منظور خنثی کردن تأثیر سهام‌های پرت در فرآیند خوشه‌بندی، مدل C-مدویدهای فازی نمایی مبتنی بر سرریز نوسانات (VS-E-FCMdd)، نیز در نظر گرفته می‌شود که به دلیل در نظر گرفتن تبدیل نمایی از معیار عدم مشابهت وزنی، یک مدل استوار محسوب می‌شود. در پژوهش حاضر این دو رویکرد خوشه‌بندی نخستین بار برای بررسی اثرات سرریز در نمونه‌ای منتخب از 30 سهام شرکت از 28 صنعت فعال در بورس تهران، طی بازه زمانی 1387 لغایت 1402 در نظر گرفته می‌شود. بر اساس شاخص زی-بنی تعداد بهینه دو خوشه در سرریزهای نوسانات تعیین می‌شود. مقایسه نتایج نشان می‌دهد که

مدل $VS-E-FCMdd$ افزایش فازی را تولید و در خوشه‌بندی سرریزهای نوسانات سهام در بورس تهران کارآتر عمل می‌کند. یافته‌های تجربی حاصل از مدل $VS-E-FCMdd$ نشان می‌دهد که نهد سرریز "از سوی" برای خوشه‌بندی سهام‌ها از اهمیت بیشتری برخوردار است. به‌ویژه اینکه خوشه 1 حاوی سهام‌هایی است که سرریز بیشتری را "از سوی" سایرین دریافت ولی سرریز کمتری را "پاسوی" آن‌ها می‌فرستد.

ارجاع به مقاله:

فرزادگان، الهام. (1403). خوشه‌بندی فازی سری‌های زمانی مالی بر اساس سرریزهای نوسانات جهت‌دار: شواهدی از سهام شرکت‌های پذیرفته‌شده در بورس تهران. فصلنامه‌ی اقتصاد مقداری (بررسی‌های اقتصادی سابق)، دوره (شماره)، ص-ص.

 [10.22055/iqe.2019.26401.1899](https://doi.org/10.22055/iqe.2019.26401.1899)



© 2024 Nahavand Higher Education Complex, Bu-Ali Sina University, Hamedan, Iran. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution-NonCommercial 4.0 International (CC BY-NC 4.0 license) (<http://creativecommons.org/licenses/by-nc/4.0/>)

پذیرفتنی
فایل
انتشار

1- مقدمه

خوشه‌بندی یک تکنیک یادگیری غیر نظارتی است که برای پیدا کردن مشابهت‌ها یا تفاوت‌ها بین واحدهای آماری در یک مجموعه داده، مورد استفاده قرار می‌گیرد. آنچه در خوشه‌بندی مهم می‌باشد این است که چطور می‌توان عدم مشابهت‌ها را اندازه‌گیری نمود.

از دیدگاه روش‌شناسی سری‌های زمانی، عدم مشابهت‌ها را می‌توان با استفاده از معیارهای فاصله اندازه‌گیری نمود؛ فاصله اقلیدسی متعارف که فقط بر مبنای نزدیکی بودن مقادیر مشاهدات قرار دارد، نمی‌تواند معیار مناسبی برای سنجش عدم مشابهت بین سری‌های زمانی باشد؛ زیرا هیچ‌گونه فرضی راجع به فرایند تولید داده‌ها در نظر نمی‌گیرد.

در ادبیات، ارائه یک تعریف مناسب از عدم مشابهت هنگامی که سری‌های زمانی مالی واقعی مورد توجه باشد، پیچیده‌تر هم می‌شود. در واقع، برای دست یافتن به یک خوشه‌بندی معتبر و قابل قبول، حقایق تجربی شناخته‌شده در سری‌های زمانی مالی نیز می‌باید در تعریف عدم مشابهت‌ها قرار داده شود. برخی ویژگی‌های سری‌های زمانی مالی که در ادبیات برای خوشه‌بندی مورد استفاده قرار گرفته‌اند عبارتند از نوسانات بازده، حافظه بلندمدت، خودهمبستگی، واریانس ناهمسانی، گشتاورهای مراتب بالا، عدم تقارن و دنباله‌های پهن چگالی‌های تجربی، (Soltyk & Chan, 2023; Jondeau & Rockinger, 2012).

با وجود این، مرور ادبیات نشان می‌دهد که هنوز هم حقایق تجربی مهم و قابل توجهی در سری‌های زمانی مالی مشاهده می‌شود که تاکنون برای خوشه‌بندی مورد توجه قرار نگرفته است. یکی از این موارد، وجود اثرات سرریز در مقادیر مشاهده‌شده نوسانات است. کشف خوشه در سرریزهای نوسانات بین دارایی‌ها یا بین پرتفوی دارایی‌ها، و یا بین بازارهای دارایی مختلف، بینش عمیقی از ساختار بازارهای مالی فراهم می‌کند که می‌تواند برای مدیران پرتفوی و سیاست‌گذاران در شرایط وقوع بحران‌های مالی و نا اطمینانی گسترده در سطح اقتصاد، مفید واقع شود؛ بالأخص زمانی که سرریزهای نوسانات جهت‌دار در نظر گرفته می‌شود (Diebold & Yilmaz, 2009, 2012). داشتن اطلاعات راجع به انتقال شوک‌ها درون یک بازار از ارزش سیاسی بالایی برخوردار است زیرا مداخلات سیاسی به‌موقع می‌تواند گسترش احتمالی شوک‌ها فراتر از حدود قابل قبول را کنترل کند. علاوه بر این، مدیریت شوک درون یک سیستم نسبتاً ساده‌تر است زیرا می‌توان با بهره‌گیری از ویژگی‌های پویای سیستمی شوک‌ها و با نظارت بر نحوه حرکات آن‌ها، قدرت پیش‌بینی بهتری برای یک دارایی حاصل نمود.

نواوری این پژوهش از دیدگاه روش‌شناسی، در استفاده از رویکردی برای خوشه‌بندی فازی سری‌های زمانی است که بر مبنای سرریزهای نوسانات جهت‌دار قرار دارد. بدین منظور به پیروی از ادبیات موجود و بالأخص سرکتی و همکاران (2023)، دو شاخص سرریز نوسانات جهت‌دار – “از سوی” و “به سوی” سایرین – از طریق تجزیه واریانس خطای پیش‌بینی تعمیم‌یافته (GFEVD)، اندازه گرفته می‌شود (Cerqueti et al., 2023; Diebold & Yilmaz, 2012). علاوه، رویکرد تقسیم‌بندی اطراف مدویدها¹ (PAM) در یک چارچوب خوشه‌بندی فازی – تحت عنوان الگوریتم C-مدویدهای فازی مبتنی بر سرریز نوسانات² (VS-FCMdd) - اتخاذ می‌گردد؛ که با استفاده از آن می‌توان به گروه‌بندی سهم‌هایی که دارای درجه مشابهی از سرریز نوسانات هستند، پرداخت. از آنجایی که در این پژوهش دو بُعد سرریز – یعنی سهم‌های “دهنده” و “گیرنده” – مورد توجه قرار دارد، در تعریف معیار فاصله از یک سیستم وزن‌دهی خودکار استفاده می‌شود؛ به این صورت که اجازه داده می‌شود داده‌ها تصمیم بگیرند که کدامیک از دو بُعد سرریز برای خوشه‌بندی مجموعه داده‌ها، مناسب‌تر است. بعلاوه، برای کاهش تأثیر منفی مشاهدات پرت در روش خوشه‌بندی، یک رویکرد استوار بر مبنای تبدیل نمایی³ از مجذور فاصله اقلیدسی وزن‌دار (VS-E-FCMdd) نیز پیشنهاد می‌شود (D'Urso et al., 2016).

¹ Partitioning Around Medoids

² Volatility Spillover-based Fuzzy C-Medoids

³ Exponential Transformation

نوآوری پژوهش حاضر در مقایسه با مطالعات داخلی، بکار گرفتن رویکرد C- مدویدهای فازی است که برای نخستین بار از سرریزهای نوسانات جهت‌دار برای خوشه‌بندی سهام شرکت‌های فعال در صنایع مختلف بورس تهران استفاده می‌کند.

2- مبانی نظری و پیشینه پژوهش

آن چیزی که در خوشه‌بندی از اهمیت زیادی برخوردار است چگونگی اندازه‌گیری عدم مشابهت‌های بین واحدها است. از نقطه‌نظر روش‌شناسی، روش‌های اندازه‌گیری عدم شباهت بین سری‌های زمانی را می‌توان در قالب سه رویکرد طبقه‌بندی نمود؛ روش‌های خوشه‌بندی مبتنی بر داده‌های خام سری‌های زمانی واقعی، روش‌های مبتنی بر مشخصه‌های سری‌های زمانی و روش‌های مبتنی بر تخمین‌های پارامترهای یک مدل آماری (Caiado et al., 2015). رویکردهای خوشه‌بندی مبتنی بر داده‌های خام، به‌طور معمول همه مشاهدات یا برخی موقعیت‌های سری‌های زمانی یا تبدیل‌های مناسبی از آن‌ها، مثل سرعت و شتاب را در نظر می‌گیرند. این رویکردها هیچ‌گونه فرضی راجع به فرآیند تولید داده‌ها لحاظ نمی‌کنند (Coppi et al., 2010; D'Urso et al., 2021). خوشه‌بندی بر مبنای مشاهدات صرف، به خاطر وجود نویز، گزینه مطلوبی نمی‌باشد. به همین خاطر، رویکردهای خوشه‌بندی مبتنی بر مشخصه، با مینا قرار دادن مشخصه‌های مناسبی از سری‌های زمانی به دنبال حل مشکل خوشه‌بندی داده‌های سری زمانی خام نویزی مورد توجه قرار گرفتند. آن دسته از مشخصه‌های سری‌های زمانی که به‌طور معمول برای تعیین معیار عدم مشابهت در نظر گرفته می‌شوند عبارت‌اند از در بُعد زمان: ضرایب همبستگی (Mantegna, 1999)، همبستگی متقابل (Alonso et al., 2021)، ساختار خودهمبستگی (D'Urso & Maharaj, 2009) و ساختار خودکواریانس کوانتیل⁴ (Lafuente-Rego et al., 2020)؛ در بُعد فرکانس: مشخصه‌های پریودوگرام (Caiado et al., 2006)، ضرایب کپسترال⁵ (Savvides et al., 2008) یا چگالی طیفی- متقابل کوانتیل (López-Oriona et al., 2022)؛ در بُعد موجک: تجزیه موجک (D'Urso et al., 2023). رویکردهای خوشه‌بندی مبتنی بر مدل، بر مبنای مشخصه‌هایی از مدل‌های برازش شده برای سری‌های زمانی قرار می‌گیرند. بر این اساس عدم مشابهت‌ها با استفاده از پارامترهای تخمین زده‌شده یا با استفاده از پسماندهای مدل‌های برازش شده، تعیین می‌شود. در این مورد روش‌های خوشه‌بندی عبارت‌اند از: مدل‌های ARMA (Piccolo, 1990)، مدل‌های GARCH (Otranto, 2008)، ضرایب اسپلاین⁶ (D'Urso et al., 2021)، پارامترهای توزیع‌ها (D'Urso, Maharaj, & Alonso, 2017).

در ادبیات خوشه‌بندی، محورهای پژوهش بر چگونگی استفاده از روش‌های خوشه‌بندی تحت حقایق تجربی مختلف مشاهده‌شده از سری‌های زمانی، قرار دارد. بدین منظور، رویکردهای خوشه‌بندی مختلفی معرفی شده است که می‌توانند این حقایق تجربی را در خود جای دهند. در رابطه با رویکردهای مبتنی بر مشخصه آماری سری‌های زمانی مالی، می‌توان به رویکردهایی که از آماره‌های مبتنی بر واریانس (Bastos & Caiado, 2014)، ضرایب کپسترال (D'Urso et al., 2020)، مولفه‌های هرست (Cerqueti & Mattera, 2023)، یا ترکیب‌هایی از مشخصه‌های مختلف داده‌های ایستا (Bastos & Caiado, 2021) استفاده می‌کنند، اشاره نمود. از سوی دیگر، در رابطه با رویکردهای مبتنی بر مدل‌سازی سری‌های زمانی مالی، مدل‌های خوشه‌بندی وجود دارد که بر پایه فرآیند ARMA (D'Urso et al., 2013)، فرآیند GARCH (D'Urso et al., 2013; D'Urso et al., 2016)، مدل خطای حاصل‌ضربی⁷ MEM (Otranto & Gargano, 2015) یا مدل‌های امتیاز-محور قرار داده‌شده‌اند به‌گونه‌ای که می‌توانند گشتاورهای مرتبه بالاتر را نیز در خود جای دهند (Cerqueti et al., 2022).

⁴ Quantile Autocovariance Structure

⁵ Cepstral Coefficients

⁶ Splines Coefficients

⁷ Multiplicative Error Model

با وجود گستردگی و تنوع مشخصه‌های در نظر گرفته‌شده، هنوز هم حقایق تجربی مهم و قابل‌ملاحظه‌ای مشاهده‌شده است که تاکنون برای خوشه‌بندی موردتوجه قرار نگرفته‌اند. یکی از این موارد، وجود اثرات سرریز در مقادیر مشاهده‌شده نوسانات است (Cerqueti et al., 2023). سرریزهای نوسانات می‌توانند بین سهام شرکت‌ها در یک بازار (Diebold & Yilmaz, 2014; Gillaizeau et al., 2019) یا بین صنایع فعال در یک بازار (Chen et al., 2022; Choi, 2022) به وجود آیند، اما می‌توانند به علت روابط میان سهام‌ها در سطح بین‌المللی نیز وجود داشته باشند (Buncic & Gisler, 2016). به‌علاوه، سرریزهای نوسانات می‌توانند حتی با لحاظ شوک‌های برونزا مثل قیمت نفت (Chang et al., 2013) یا دیگر نا اطمینانی‌ها (Cheuathonghua et al., 2019) هنوز هم وجود داشته باشند. سرکتی و همکاران (2023) بر کشف خوشه‌ها در سرریزهای نوسانات جهت‌دار (Diebold & Yilmaz, 2009, 2012) تأکید کرده‌اند؛ آن‌ها معتقدند که تجزیه و تحلیل سرریزهای نوسانات جهت‌دار امکان شناسایی دارایی‌های به‌اصطلاح "دهنده" و "گیرنده"⁸ را فراهم می‌کند. دارایی‌هایی دهنده محسوب می‌شوند که بیشتر از آنکه سرریز دریافت کنند، آن را برای سایر دارایی‌ها تولید می‌کنند؛ برعکس، دارایی‌های گیرنده سرریز را بیشتر جمع می‌کنند تا اینکه آن را برای سایرین تولید کنند. براین اساس، این محققین بین منابع سرریز "از سوی" و سرریز "به سوی"⁹ تمییز در نظر گرفته‌اند. وجود خوشه‌ها در سرریزهای نوسانات از نظر مفهومی دلالت بر وجود گروه‌هایی از سهام‌ها دارد که وقوع یک شوک در آن‌ها هردوی نوسانات غیرسیستماتیک و نوسانات مربوط به سایر دارایی‌ها را افزایش می‌دهد؛ درحالی‌که گروه‌هایی دیگر از سهام‌ها وجود دارد که بیشتر در معرض شوک‌های ناشی از سایر دارایی‌ها قرار دارند. به‌عبارت‌دیگر، وجود خوشه‌ها دلالت دارد بر اینکه میزان انتقال نوسانات میان گروه‌های سهام‌ها مختلف است.

نکته‌ای که در ادبیات خوشه‌بندی وجود دارد این است که شناسایی یک مرز مشخص و قطعی بین خوشه‌ها در بازار سهام واقعی، همیشه امکان‌پذیر نیست؛ در مقابل، خوشه‌بندی در یک چارچوب فازی، این امکان را فراهم می‌کند که هر سهم بتواند هم‌زمان متعلق به چند خوشه باشد با درجه عضویتی که نشان‌دهنده نا اطمینانی راجع به تخصیص آن سهم به هر خوشه است. براین اساس، می‌توان بر نا اطمینانی ذاتی قرارگرفته در خوشه‌بندی سری‌های زمانی مالی غلبه نمود (Cerqueti et al., 2023).

برخلاف اهمیت موضوع، شواهد تجربی در این زمینه برای بازارهای نوظهور از جمله بورس تهران با داشتن چشم‌انداز توسعه آتی، وجود ندارد. از این‌رو، پژوهش حاضر به دنبال پوشش این شکاف است و به پیروی از رویکرد پیشنهادی (Cerqueti et al., 2023)، یک مدل خوشه‌بندی فازی برای کشف گروه‌هایی از سهام‌ها که درجه مشابه سرریز دارند، استفاده می‌شود. برای این منظور، از الگوریتم PAM در چارچوب فازی استفاده می‌شود؛ مزیت رویکرد PAM در این است که از نقاط داده به‌عنوان مدویدها استفاده می‌کند (Van der Laan et al., 2003).

مسئله دیگری که ممکن است در فرآیند خوشه‌بندی با آن مواجه شویم زمانی است که داده‌های دنیای واقعی از بازارهایی جمع‌آوری شده باشند که متشکل از تعدادی اندک از سهام‌هایی باشند که از مقادیر بسیار زیاد یا مقادیر بسیار کم سرریز در مقایسه با دیگر سهام‌ها برخوردار هستند. این نوع الگوهای خلاف قاعده در سری‌های زمانی را برای مثال می‌توان در بازاری نسبتاً کوچک که تحت سلطه کامل چند شرکت بزرگ قرار دارد، مشاهده نمود؛ هر چند که این‌چنین الگوهایی می‌تواند در بازارهای بزرگ پیشرفته‌تر نیز وجود داشته باشد. در ادبیات، برای خنثی ساختن اثر منفی داده‌های پرت، روش‌های خوشه‌بندی فازی استوار پیشنهاد شده است: در رویکرد نویزی، سری‌های زمانی پرت به خوشه‌های نویزی اختصاص داده می‌شوند (D'Urso et al., 2013)؛ در رویکرد متریک، معیارهای فاصله که از ویژگی‌های استوار بودن برخوردار هستند تعریف می‌شود (D'Urso et al., 2015). در رویکرد پیراسته¹⁰، روش خوشه‌بندی بعد از حذف بخش معینی از داده‌های پرت، پیاده‌سازی می‌شود (D'Urso,

⁸ Givers and Receivers

⁹ "From" and "To" Sources of Spillovers

¹⁰ Trimmed Approach

(Massari, et al., 2017). در رویکرد اعمال وزن‌دهی، یک سیستم وزن‌دهی در فرآیند خوشه‌بندی وارد می‌شود تا به سری‌های زمانی پرت به‌طور عینی وزن پایین داده شود (D'Urso, 2005). براین اساس، در پژوهش حاضر به پیروی از (Cerqueti et al., 2023) روش خوشه‌بندی فازی بر پایه رویکرد وزن‌دهی بکار برده می‌شود؛ برای خنثی نمودن اثر داده‌های پرت، یک متریک استوار¹¹ مبتنی بر تبدیل نمایی از مجذور فاصله اقلیدسی وزن‌دار (D'Urso et al., 2016) اتخاذ می‌شود. بی و همکاران (2021) یک معیار فاصله وزنی متشکل از فاصله اقلیدسی، فاصله ساختار زمانی پویا¹² (DTW) و فاصله DTW مبتنی بر تفاضل دنباله داده‌ها را معرفی و از آن در تابع هدف الگوریتم خوشه‌بندی C-مدویدهای فازی و الگوریتم خوشه‌بندی ترکیبی C-میانگین فازی¹³ استفاده نمودند (Bei et al., 2021). این محققین تعداد خوشه‌های بهینه را با استفاده از الگوریتم بهینه‌سازی ازدحام ذرات به دست آورده‌اند. دی ارسو و همکاران (2016) مدل‌های خوشه‌بندی فازی استوار مختلفی را پیشنهاد نمودند که در تابع هدف آن‌ها، یک معیار فاصله اقلیدسی وزن‌دار با استفاده از پارامترهای تخمین زده‌شده مدل GARCH و یک سیستم وزن‌دهی مناسب، بکار برده می‌شود. این معیار فاصله بر مبنای مقایسه فاصله بین نوسان‌پذیری غیرشرطی سری‌های زمانی و فاصله بین نوسان‌پذیری زمان متغیر سری‌های زمانی، قرار دارد (D'Urso et al., 2016). ژو و همکاران (2016) یک الگوریتم "خوشه‌بندی C-مدویدهای شواهد"¹⁴ (EGMdd) معرفی کردند که بر مبنای مدویدهای وزن‌دار چندگانه قرار دارد از این نظر که برای نمایش کامل هر خوشه جنبه‌های مختلفی از ساختار گروه‌ها می‌باید در نظر گرفته شود (Zhu, 2016).

در داخل کشور، احسانی و آل‌محمد (2021) با استفاده از رویکرد خوشه‌بندی فازی اطراف مدوید مبتنی بر معیار فاصله اقلیدسی بهبودیافته بر اساس مدل اریما، به خوشه‌بندی 21 صنعت در بورس اوراق بهادار تهران پرداختند. سری‌های زمانی بر اساس شاخص سیلپوت سخت و سیلپوت فازی به 4 گروه خوشه‌بندی شده‌اند (Ehsani & Alemohammad, 2021). رعیتی شوازی و همکاران (2021) با استفاده از رویکرد خوشه‌بندی K-میانگین مبتنی بر سوابق مالی پرتفوی همپوشان، به خوشه‌بندی صنایع مختلف در بورس اوراق بهادار تهران پرداخته‌اند (Rayati Shavazi et al., 2021). اصغری و احمدی (2018) با استفاده از الگوریتم کی-مینز به خوشه‌بندی مشتریان دریافت‌کننده تسهیلات صندوق کارآفرینی امید پرداختند (Asghari & Ahmadi, 2018).

سینایی و همکاران (2014) با استفاده از رویکرد خوشه‌بندی K-میانگین مبتنی بر معیار فاصله مجذور اقلیدسی به رتبه‌بندی صندوق‌های سرمایه‌گذاری پرداختند. آن‌ها برای شناسایی متغیرهایی که بیشترین تأثیر را در خوشه‌بندی دارند از رویکرد ANOVA استفاده نمودند (Sinaei et al., 2014). فرید و پورحمیدی (2012) به طبقه‌بندی 338 سهام شرکت‌های منتخب در قالب 3 خوشه معین با استفاده از تحلیل خوشه‌بندی فازی C-میانگین (FCM) پرداختند (Farid & Pourhamidi, 2012). برخی پژوهشگران نیز از روش خوشه‌بندی برای تشکیل پرتفوی استفاده کرده‌اند (Rahmani Asl et al., 2021; Sinaei et al., 2016).

در داخل کشور، رویکردهایی نیز برای بهبود الگوریتم خوشه‌بندی فازی معرفی شده که عملکرد آن‌ها عمدتاً برای داده‌های مصنوعی ارزیابی شده است. فرنام و درهمیرکی (2023) به معرفی یک الگوریتم FCM بهبودیافته مبتنی بر یک معیار فاصله پارامتری جدید برای داده‌های فازی پارامتری پرداختند (Farnam & Darehmiraiki, 2023). رفیعی و مرادی (2016) به معرفی یک الگوریتم C-میانگین فازی بهبودیافته با وزن‌دهی اتوماتیک و وزن‌دهی محلی ویژگی‌ها پرداختند. بعلاوه، این محققین از معیار شباهت غیراقلیدسی که اثر نویز در داده‌ها را کاهش می‌دهد استفاده کرده‌اند (Rafiei & Moradi, 2016).

¹¹ Robust Metric

¹² Dynamic Time Structuring

¹³ Fuzzy C-means Hybrid Clustering

¹⁴ Evidential c-Medoids Clustering

نوآوری پژوهش حاضر در مقایسه با مطالعات داخلی، بکار گرفتن الگوریتم C-مدوید فازی است که برای نخستین بار از سرریزهای نوسانات جهت‌دار بین سهام شرکت‌ها برای خوشه‌بندی سری‌های زمانی استفاده می‌کند. همچنین برخلاف مطالعه رفیعی و مرادی (2016) که تابع هدف را به صورت موزون در نظر گرفته‌اند (Rafiei & Moradi, 2016)، در پژوهش حاضر به خود فاصله اقلیدسی، وزن داده شده است. بعلاوه، برخلاف فرنام و درهمیرکی (1401) که با معرفی یک متر جدید به دنبال کاهش اثر منفی داده‌های پرت بوده‌اند در پژوهش حاضر از تبدیل نمایی فاصله اقلیدسی وزن‌دار استفاده شده است (Farnam & Darehmiraqi, 2023). در پایان، برخلاف همه مطالعات فوق، در پژوهش حاضر برای تعیین تعداد بهینه خوشه‌ها و ارزیابی مدل‌ها، از شاخص ژنی-جینی (1991) استفاده شده است (Xie & Beni, 1991).

3- روش‌شناسی پژوهش

3-1- اندازه‌گیری سرریزهای نوسانات جهت‌دار و عدم مشابهت

3-1-1- اندازه‌گیری سرریزها

دیبولد و بیلماز (2009) پیشنهاد کردند که در یک فرآیند خودرگرسیو برداری (VAR)، شاخص‌های سرریز را می‌توان بر حسب تجزیه واریانس خطای پیش‌بینی تعمیم‌یافته (GFEVD) محاسبه کرد (Pesaran & Shin, 1998) برای یک فرآیند N-متغیره VAR(P) کواریانس مانا ($i = 1, \dots, N$) به صورت زیر:

$$x_t = \sum_{p=1}^P \Phi_p x_{t-p} + \varepsilon_t, \quad t = 1, \dots, T \quad (1)$$

که Φ_p ماتریس پارامترها از بُعد $N \times N$ در وقفه p ام است و $\varepsilon_t \sim (0, \Sigma)$ بردار جملات خطا با توزیع i.i.d می‌باشد. از آنجاییکه برای اندازه‌گیری سرریزهای نوسانات، سری زمانی x_t می‌بایست شامل یک پراکسی از ریسک سهام نیز شود، به پیروی از پارکینسون (1980)، واریانس آمین سهام به صورت زیر محاسبه می‌شود (Parkinson, 1980):

$$\sigma_{it}^2 = \frac{1}{4 \ln 2} \left[\ln \left(\frac{p_{it}^{\max}}{p_{it}^{\min}} \right) \right]^2 \quad (2)$$

که p_{it}^{\min} و p_{it}^{\max} قیمت‌های ماکزیمم و مینیمم در روز معاملاتی t هستند. فرآیند VAR(1) کواریانس مانا را می‌توان بر حسب فرآیند میانگین متحرک برداری (VMA) نیز نوشت:

$$x_t = \sum_{p=0}^{\infty} A_p \varepsilon_{t-p} \quad (3)$$

که ماتریس ضرایب A_p از بُعد $N \times N$ بوده و از رابطه بازگشتی $A_p = \sum_{s=1}^p A_{p-s} \Phi_s$ پیروی می‌کند؛ ضرایب میانگین متحرک مهم‌ترین مؤلفه برای محاسبه شاخص‌های سرریز هستند.

برای محاسبه شاخص‌های سرریز، دیبولد و بیلماز (2012) رویکرد تجزیه واریانس خطای پیش‌بینی تعمیم‌یافته H-گام به جلو¹⁵ (GFEVD) را پیشنهاد کردند (Pesaran & Shin, 1998) که نشان‌دهنده کسری از واریانس خطای H-گام به جلو ($h = 1, \dots, H$) در پیش‌بینی سهام λ ام است که به علت بروز شوک به سهام λ ام می‌باشد ($i \neq j$). پس، GFEVD مربوط به آمین سهام- به علت بروز شوک به سهام λ ام- را می‌توان به صورت زیر محاسبه نمود:

¹⁵ H-step ahead Generalized Forecast Error Variance Decomposition

$$\theta_{ij}^g(H) = \frac{\sigma_{jj}^{-1} \sum_{h=0}^{H-1} (e_i' A_h \Sigma e_j)^2}{\sum_{h=0}^{H-1} (e_i' A_h \Sigma A_h' e_i)} \quad (4)$$

که θ_{ij}^g ، تجزیه واریانس خطای پیش‌بینی تعمیم‌یافته H-گام به جلو است؛ Σ ، ماتریس واریانس بردار خطای ε_t است، σ_{jj} ، انحراف معیار جمله خطای معادله زام - ناشی از سهام زام - در فرآیند VAR است؛ e_j برداری انتخابی است که برای آمین درایه مقدار یک می‌گیرد و مقدار صفر را برای مابقی درایه‌ها اختیار می‌کند. در این رویکرد نتایج تجزیه‌های واریانس خطای پیش‌بینی، دیگر متأثر از رتبه‌بندی متغیرها در سیستم نخواهد بود (Pesaran & Shin, 1998). به‌منظور بهره‌برداری از اطلاعات موجود در ماتریس تجزیه واریانس برای محاسبه شاخص سرریز نوسانات، هر درایه $\theta_{ij}^g(H)$ از ماتریس تجزیه واریانس خطای پیش‌بینی در طول هر سطر از این ماتریس، نرمال می‌شود:

$$\tilde{\theta}_{ij}^g(H) = \frac{\theta_{ij}^g(H)}{\sum_{j=1}^N \theta_{ij}^g(H)} \quad (5)$$

در اینصورت خواهیم داشت $\sum_{j=1}^N \tilde{\theta}_{ij}^g(H) = 1$ و $\sum_{i,j=1}^N \tilde{\theta}_{ij}^g(H) = N$ از روابط فوق می‌توان شاخص سرریز نوسانات کل را بر اساس سهم سرریزهای نوسانات از سوی همه سهامها (رابطه (5)) به کل واریانس خطای پیش‌بینی، تعریف نمود:

$$S^g(H) = \frac{\sum_{i,j=1}^N \tilde{\theta}_{ij}^g(H)}{N} \times 100 \quad (6)$$

اندازه‌گیری مقدار کل سرریزها درون یک بازار مالی (یا بین بازارهای مالی) نمی‌تواند برای اهداف خوشه‌بندی مفید واقع شود. از آنجاییکه، هر سهم خاص در یک بازار هم می‌تواند دهنده سرریز به‌سوی سایر سهامها باشد و هم دریافت‌کننده سرریز از سوی آنها؛ در رویکردهای خوشه‌بندی، علاوه بر اندازه سرریزها، اینکه سرریزها در چه جهتی سرایت می‌کنند نیز مورد توجه قرار دارد. شاخص‌های سرریز جهت‌دار این امکان را فراهم می‌کنند تا میزان سرریز نوسانات دریافت شده (داده‌شده) توسط سهام آم از سوی (به‌سوی) سایر سهامها در سیستم به‌صورت زیر اندازه‌گیری شود:

$$S_{i \rightarrow}^g(H) = \frac{\sum_{j=1}^N \tilde{\theta}_{ij}^g(H)}{N} \times 100 \quad (7)$$

برای اندازه‌گیری سرریزهای نوسانات که سهام آم از سوی سایرین دریافت می‌کند، و:

$$S_{\rightarrow i}^g(H) = \frac{\sum_{j=1}^N \tilde{\theta}_{ji}^g(H)}{N} \times 100 \quad (8)$$

برای اندازه‌گیری سرریزهای جهت‌دار تولیدشده توسط سهام آم به‌سوی سایر سهامها. براین اساس، رابطه (7) سرریز "از سوی" و رابطه (8) سرریز "به‌سوی"، تعریف می‌شوند.

3-1-2- معیار عدم مشابهت بر اساس سرریزهای وزنی

در این پژوهش، عدم مشابهت بین سهامها با توجه به توانایی آنها در تولید و دریافت سرریزهای ریسک، محاسبه می‌شود. بنابراین، یک رویکرد بدیهی برای اندازه‌گیری عدم مشابهت بین دو سهم i و j ، می‌تواند بر اساس فاصله‌های زیر باشد:

$$\text{From } d_{ii}^2 = \quad (9)$$

$$[S_{i \rightarrow}^g(H) - S_{i' \rightarrow}^g(H)]^2 = \frac{\left[\sum_{\substack{j=1 \\ j \neq i}}^N \tilde{\theta}_{ij}^g(H) - \sum_{\substack{j=1 \\ j \neq i'}}^N \tilde{\theta}_{ij'}^g(H) \right]^2}{N^2}$$

این فاصله برحسب سرریز "از سوی" در رابطه (7) محاسبه می‌شود و:

$$\text{to } d_{ii}^2 = \quad (10)$$

$$[S_{\rightarrow i}^g(H) - S_{\rightarrow i'}^g(H)]^2 = \frac{\left[\sum_{\substack{j=1 \\ j \neq i}}^N \tilde{\theta}_{ji}^g(H) - \sum_{\substack{j=1 \\ j \neq i'}}^N \tilde{\theta}_{ji'}^g(H) \right]^2}{N^2}$$

این فاصله برحسب سرریز "به سوی" در رابطه (8) محاسبه می‌شود. در واقع، در پژوهش حاضر از این واقعیت استفاده می‌شود که سهمها در یک بازار مفروض را می‌توان برحسب هر دوی اندازه و جهت سرریزها (روابط (7) و (8))، متمایز نمود. علت این است که سهمهایی کوچک در بازار می‌تواند وجود داشته باشد که مقدار بسیاری سرریز از سوی سایرین دریافت می‌کنند درحالی‌که مقدار اندکی سرریز به سوی آنها می‌دهند؛ به همین صورت، تعداد زیادی سهام در بازار می‌تواند وجود داشته باشد که مقدار زیادی سرریز تولید می‌کنند درحالی‌که مقدار اندکی سرریز از سایرین دریافت می‌کنند. به‌رحال در بازارهای واقعی، تمییز بین دهنده و گیرنده آسان نیست زیرا همه سهمها هم‌زمان تولیدکننده و دریافت‌کننده ولی با شدت‌های مختلف، هستند.

از این‌رو، برای محاسبه یک معیار عدم مشابهت مناسب می‌باید هر دو معیار فاصله (9) و (10) هم‌زمان باهم در نظر گرفته شوند. اما از آنجایی‌که بسته به بازار خاص مورد مطالعه، یکی از این دو بُعد سرریز می‌تواند از اهمیت بیشتری برای خوشه‌بندی برخوردار باشد اعمال مجموعه‌ای متفاوت از وزن‌ها بین سرریزهای "از سوی" و "به سوی"، مورد نیاز می‌باشد. بر این اساس، یک فاصله اقلیدسی مربعی وزن‌دار مشتمل بر یک سیستم وزن‌دهی مناسب برای هر دو مؤلفه سرریز، در نظر گرفته می‌شود (D'Urso et al., 2016):

$$d_{ii}^2(H) = \omega_1^2 \text{From } d_{ii}^2(H) + \omega_2^2 \text{To } d_{ii}^2(H) \quad (11)$$

که $d_{ii}^2(H)$ فاصله اقلیدسی مربعی وزن‌دار است؛ $\text{From } d_{ii}^2(H)$ ، فاصله اقلیدسی مربعی برحسب سرریز "از سوی" (9) و $\text{To } d_{ii}^2(H)$ ، برحسب سرریز "به سوی" (10) را مشخص می‌کنند؛ ω_1 و ω_2 عبارتند از وزن‌های مربوط به مؤلفه‌های سرریز "از سوی" و "به سوی". در اینجا فرض می‌شود $\omega_1 + \omega_2 = 1$ و $\omega_1, \omega_2 \geq 0$.

اصولاً، وزن‌های ω_1 و ω_2 را می‌توان به‌صورت ذهنی یا به‌طور عینی تعیین نمود. برخلاف رویکرد ذهنی که وزن‌ها با توجه به نظرات ذهنی یا نظرات متخصص تعیین می‌شوند؛ در رویکرد عینی، وزن‌ها در طی اجرای یک الگوریتم خوشه‌بندی مناسب، انتخاب می‌شوند. بر این اساس در این پژوهش، با استفاده از سیستم وزن‌دهی عینی اجازه داده می‌شود تا داده‌ها خود مناسب‌ترین مجموعه وزن‌ها را تعیین کنند.

متأسفانه ممکن است فاصله اقلیدسی مربعی (11) در شرایط وجود داده‌های نویزی یا مقادیر پرت، از قوت آماری برخوردار نباشد (D'Urso et al., 2021; D'Urso et al., 2023; Wu & Yang, 2002). وجود مقادیر پرت در شاخص‌های سرریز امکان‌پذیر است. به همین خاطر، وو و یانگ (2002) و دی ارسو و همکاران (2016)، یک روش خوشه‌بندی استوار بر مبنای تبدیل‌نمایی از عدم مشابهت، پیشنهاد کردند (D'Urso et al., 2016; Wu & Yang, 2002):

(12)
$$\exp d_{ii}^2(H) = 1 - \exp[-\beta d_{ii}^2(H)]$$
 به‌طوری‌که β یک مقدار ثابت مثبت است و $d_{ii}^2(H)$ معیار عدم مشابهت سرریز وزنی از رابطه (11) است. از نظر مفهومی، در تبدیل نمایی از معیار فاصله، وزن‌های کوچک‌تر به اشیائی اختصاص داده می‌شود که در مجموعه داده‌ها، نویزی هستند. مقدار β مطابق با رویکرد پیشنهادی وو و یانگ (2002) محاسبه می‌شود (Wu & Yang, 2002).

3-2- مدل‌های خوشه‌بندی فازی

مدل‌های خوشه‌بندی فازی مورد استفاده در این پژوهش بر مبنای معیار عدم مشابهت وزنی 3-1- (11) و (12) و تبدیل نمایی آن (12) قرار دارند. بدین منظور، در ابتدا معیارهای سرریز جهت‌دار (7) و (8) برای هر سهم λ محاسبه می‌شود که داده‌های ورودی روش خوشه‌بندی محسوب می‌شوند. این داده‌های ورودی را می‌توان به‌صورت یک ماتریس S یک ماتریس $S = \{S_i = [S_{i \rightarrow}^g(H), S_{\rightarrow i}^g(H)]\}$, $i = 1, \dots, N$ از بُعد $N \times 2$ نمایش داد:

$$S = \begin{bmatrix} S_{1 \rightarrow}^g(H) & S_{\rightarrow 1}^g(H) \\ \vdots & \vdots \\ S_{i \rightarrow}^g(H) & S_{\rightarrow i}^g(H) \\ \vdots & \vdots \\ S_{N \rightarrow}^g(H) & S_{\rightarrow N}^g(H) \end{bmatrix} \quad (13)$$

ستون‌ها، معیارهای سرریز "از سوی" (7) و "به سوی" (8) مربوط به N سهم هستند. در این پژوهش، دو رویکرد PAM در نظر گرفته می‌شود که در آن‌ها نمونه نماینده هر خوشه¹⁶، بانام مدوید، همان سهام واقعی است نه "اشیاء مجازی". این ویژگی قابلیت تفسیر خوشه‌ها و نماینده آن‌ها را امکان‌پذیر می‌سازد.

3-2-1- رویکرد C-مدویدهای فازی مبتنی بر سرریز نوسانات (VS-FCMdd)

با داشتن ماتریس S و معیار عدم مشابهت (11)، واحدها را می‌توان با استفاده از رویکرد VS-FCMdd که بر پایه مسئله کمینه‌سازی مربعی مقید قرار دارد، خوشه‌بندی نمود:

$$\min: \mathcal{F}_{VS-FCMdd}(S, U, \omega_1, \omega_2) = \sum_{l=1}^N \sum_{c=1}^C u_{lc}^m d_{lc}^2(H) \quad (14)$$

با قیود:

$$\sum_{c=1}^C u_{lc} = 1, \quad \omega_1, \omega_2 \geq 0, \quad \omega_1 + \omega_2 = 1 \quad (15)$$

که در آن U ماتریس درجه عضویت است؛ پارامتر $m \geq 1$ فازی بودن خوشه‌ها را کنترل می‌کند؛ $d_{lc}^2(H)$ عدم مشابهت وزنی بین λ امین واحد و مدوید خوشه c ام است. منظور از درجه عضویت u_{lc} ، درجه عضویت λ امین واحد در خوشه c ام می‌باشد. اگر $m = 1$ باشد، رابطه (14) به رویکرد PAM غیر فازی (کلاسیک) تبدیل می‌شود، که سهام i متعلق به خوشه c ام با عضویت باینری $\{0, 1\}$ $u_{lc} =$ است؛ جواب‌های بهینه مسئله برای درجه عضویت عبارت‌اند از:

¹⁶ Prototypes of each Cluster

$$u_{ic} = \frac{1}{\sum_{c'=1}^C \left[\frac{\omega_1^2 \text{From} d_{ic}^2(H) + \omega_2^2 \text{To} d_{ic}^2(H)}{\omega_1^2 \text{From} d_{ic}^2(H) + \omega_2^2 \text{To} d_{ic}^2(H)} \right]^{1/(m-1)}} \quad (16)$$

و برای وزن‌ها:

$$\omega_1 = \frac{\sum_{i=1}^N \sum_{c=1}^C u_{ic}^m \text{To} d_{ic}^2(H)}{\sum_{i=1}^N \sum_{c=1}^C u_{ic}^m [\text{From} d_{ic}^2(H) + \text{To} d_{ic}^2(H)]} \quad (17)$$

$$\omega_2 = \frac{\sum_{i=1}^N \sum_{c=1}^C u_{ic}^m \text{From} d_{ic}^2(H)}{\sum_{i=1}^N \sum_{c=1}^C u_{ic}^m [\text{From} d_{ic}^2(H) + \text{To} d_{ic}^2(H)]} \quad (18)$$

جواب‌های بهینه با استفاده از روش ضرایب لاگرانژ به دست آمده‌اند. واضح است که رابطه (16) به دو رابطه (17) و (18) بستگی دارد و برعکس؛ بنابراین، برای یافتن رامل‌ها، یک الگوریتم تکراری مورد نیاز است؛ بطوریکه در هر تکرار، میزان عضویت و نیز مدویدهای خوشه‌ها تغییر می‌کنند. اثبات این روابط در مقاله دی ارسو و همکاران (2016) شرح داده شده است (D'Urso et al., 2016). مراحل مورد نیاز برای اجرای الگوریتم VS-FCMdd به شرح جدول 1 است (Cerqueti et al., 2023).

جدول 1. الگوریتم خوشه‌بندی VS-FCMdd
مأخذ: (Cerqueti et al., 2023)

Table 1. Clustering Algorithm VS-FCMdd
Source: (Cerqueti et al., 2023)

مرحله 0: تخمین شاخص‌های سرریز جهت‌دار مطابق رابطه‌های (7) و (8)؛ مقداردهی اولیه برای تعداد خوشه‌ها C ، برای ماکزیمم تعداد تکرار max.iter و برای توان درجه عضویت (میزان فازی بودن) m . تولید تصادفی ماتریس درجه عضویت U مشروط به محدودیت (15) و متعاقباً انتخاب تعداد C مدوید از ماتریس داده‌ها S و تشکیل ماتریس H با درایه‌های s_c مربوط به مدوید خوشه c ام.
مرحله 1: محاسبه وزن‌های w مطابق روابط (17) و (18) از روی ماتریس‌های U و H که در تکرار قبل به دست آمده‌اند.
مرحله 2: به‌روزرسانی ماتریس مدوید H : برای $c = 1, \dots, C$
$$q = \underset{i' \leq N}{\operatorname{argmin}} \sum_{i'=1}^N u_{i'c}^m [\omega_1^2 \text{From} d_{i'c}^2(H) + \omega_2^2 \text{To} d_{i'c}^2(H)]$$

قرار می‌دهیم $s_c = q$. به‌روزرسانی ماتریس U مطابق رابطه (16) از روی ماتریس به‌روزرسانی شده H (مرحله 2) و وزن‌های w (مرحله 1).
مرحله 3: اگر تعداد تکرار برابر max.iter شود الگوریتم خاتمه می‌یابد، در غیر این صورت برو به مرحله 1.

3-2-2- رویکرد C-مدویدهای فازی نمایی مبتنی بر سرریز نوسانات (VS-E-FCMdd)

روش خوشه‌بندی استوار VS-E-FCMdd، از حل مسئله کمینه‌سازی زیر به دست می‌آید:

$$\min: \mathcal{F}_{VS-E-FCMdd}(S, U, \omega_1, \omega_2) = \sum_{l=1}^N \sum_{c=1}^C u_{lc}^m \exp d_{lc}^2(H) \quad (19)$$

با قیود:

$$\sum_{c=1}^C u_{ic} = 1, \quad \omega_1, \omega_2 \geq 0, \quad \omega_1 + \omega_2 = 1 \quad (20)$$

در این مورد $\exp d_{ic}^2(H)$ عدم مشابهت وزنی بین واحد λ_m و مدوید خوشه کام تعریف می‌شود. چنانچه داشته باشیم $\omega_1 = \omega_2 = 0/5$ و $m = 1$ ، رابطه (19) به PAM غیر فازی تحت عدم مشابهت نمایی، تبدیل می‌شود (Wu & Yang, 2002). راحل بهینه برای درجه عضویت مربوط به مسئله (19) با استفاده از روش ضرایب لاگرانژ به دست می‌آید (D'Urso et al., 2016):

$$u_{ic} = \frac{1}{\sum_{c=1}^C \left[\frac{1 - \exp\{-\beta[\omega_1^2 \text{From} d_{ic}^2(H) + \omega_2^2 \text{To} d_{ic}^2(H)]\}}{1 - \exp\{-\beta[\omega_1^2 \text{From} d_{ic}^2(H) + \omega_2^2 \text{To} d_{ic}^2(H)]\}} \right]^{1/(m-1)}} \quad (21)$$

و برای وزن‌ها:

$$\omega_1 = \frac{\sum_{i=1}^N \sum_{c=1}^C u_{ic}^m \text{To} d_{ic}^2(H) [1 - \exp d_{ic}^2(H)]}{\sum_{i=1}^N \sum_{c=1}^C u_{ic}^m [\text{From} d_{ic}^2(H) + \text{To} d_{ic}^2(H)] [1 - \exp d_{ic}^2(H)]} \quad (22)$$

$$\omega_2 = \frac{\sum_{i=1}^N \sum_{c=1}^C u_{ic}^m \text{From} d_{ic}^2(H) [1 - \exp d_{ic}^2(H)]}{\sum_{i=1}^N \sum_{c=1}^C u_{ic}^m [\text{From} d_{ic}^2(H) + \text{To} d_{ic}^2(H)] [1 - \exp d_{ic}^2(H)]} \quad (23)$$

همچنین راحل‌ها به‌طور تکراری به دست می‌آیند. الگوریتم مورد استفاده برای اجرای روش خوشه‌بندی استوار به شرح جدول 2 است.

β نیز بر طبق رویکرد پیشنهادی وو و یانگ (2002) به‌صورت زیر محاسبه می‌شود:

$$\beta = \left[\frac{\sum_{i=1}^N d^2(X_i, \bar{X}_q)}{N} \right]^{-1} \quad (24)$$

$$\left[\frac{\sum_{i=1}^N \omega^2 \text{From} d_{ii}^2(H) - (1 - \omega)^2 \text{To} d_{ii}^2(H)}{N} \right]^{-1} \\ = \left[\frac{\sum_{i=1}^N \omega^2 [S_{i \rightarrow}^g(H) - \tilde{S}_{q \rightarrow}^g(H)]^2 - (1 - \omega)^2 [S_{\rightarrow i}^g(H) - \tilde{S}_{\rightarrow q}^g(H)]^2}{N} \right]^{-1}$$

بطوریکه

$$\tilde{X}_q = \left(\tilde{S}_{q \rightarrow}^g(H), \tilde{S}_{\rightarrow q}^g(H) \right) : \quad (25)$$

$$q = \operatorname{argmin}_{1 \leq i \leq N} \sum_{i=1}^N d^2(X_i, \tilde{X}_i)$$

$\tilde{X}_q = \left(\tilde{S}_{q \rightarrow}^g(H), \tilde{S}_{\rightarrow q}^g(H) \right)$ مدوید واحد q است که نزدیکترین فاصله را با سایرین دارد (Wu & Yang, 2002). بر این اساس، معیار عدم مشابهت نمایی (12)-(13) وزن‌های بالاتر را به فاصله‌های پایین (بالا) که داده‌ها تغییرپذیری زیاد (کم) دارند، اختصاص می‌دهد.

جدول 2. الگوریتم خوشه‌بندی VS-E-FCMdd
ماخذ: (Cerqueti et al., 2023)

Table 2. Clustering Algorithm VS-E-FCMdd
Source: (Cerqueti et al., 2023)

مرحله 0: تخمین شاخص‌های سرریز جهت‌دار در رابطه‌های (7) و (8)(8)(8)؛ مقداردهی اولیه برای تعداد خوشه‌ها C ، ماکزیمم تعداد تکرار $\max.iter$ ، توان درجه عضویت m و محاسبه مقدار β . تولید تصادفی ماتریس درجه عضویت U مشروط به محدودیت (20) و متعاقباً انتخاب تعداد C مدوید از ماتریس داده‌ها S و تشکیل ماتریس H با درایه‌های S_c مربوط به مدوید خوشه C .
مرحله 1: محاسبه وزن‌های w مطابق روابط (22) و (23) از روی ماتریس‌های U و H که در تکرار قبل به دست آمدند.
مرحله 2: به‌روزرسانی ماتریس مدوید H : برای $c = 1, \dots, C$

$$q = \operatorname{argmin}_{1 \leq i \leq N} \sum_{i=1}^N u_{i^c}^m \left\{ 1 - \exp \left[-\beta \left(\omega_{1From}^2 d_{i^c}^2(H) + \omega_{2To}^2 d_{i^c}^2(H) \right) \right] \right\}$$

قرار می‌دهیم $s_q = S_c$. به‌روزرسانی ماتریس U مطابق رابطه (21) از روی ماتریس به‌روزرسانی شده H (مرحله 2) و وزن‌های w (مرحله 1).
مرحله 3: اگر تعداد تکرار برابر $\max.iter$ شود الگوریتم خاتمه می‌یابد، در غیر اینصورت برو به مرحله 1.

4- یافته‌های پژوهش

4-1- خوشه‌بندی سری‌های زمانی بورس تهران

چارچوب نظری معرفی شده، برای خوشه‌بندی سری‌های زمانی بورس تهران، بکار برده می‌شود. جامعه آماری در نظر گرفته شده شامل کلیه شرکت‌های پذیرفته شده در بورس تهران است. مجموعه نمونه، شامل سهام شرکت‌هایی می‌شود که سری زمانی روزانه کاملی را طی بازه 1387/6/5 تا 1402/6/5، فراهم می‌کنند. انتخاب شرکت‌ها به گونه‌ای است که تا حد امکان اکثر صنایع فعال در بورس تهران را پوشش دهد. اما سری زمانی هیچ‌کدام از شرکت‌ها از اندازه یکسانی برخوردار نیست زیرا مشاهدات از تاریخ یکسانی شروع نمی‌شود و بنابراین اطلاعات برخی سری‌های زمانی تنها برای دوره زمانی کوتاهتری موجود است. همان‌طور که دی ارسو و همکاران (2016) نشان داده‌اند طبقه‌بندی سری‌های زمانی با اندازه‌های متفاوت هم امکان‌پذیر است چراکه مدل‌سازی سری‌های زمانی در یک چارچوب پارامتریک (فرآیند VAR) انجام می‌شود (D'Urso et al., 2016). به این ترتیب، از بین شرکت‌هایی که بیشترین تعداد مشاهدات را طی بازه زمانی مذکور تشکیل می‌دهند تعداد $N = 30$ سهام که متعلق به 27 صنعت است، به عنوان نمونه آماری انتخاب گردید (Error! Reference source not found.0).
داده‌های مورد نیاز از نرم‌افزار ره‌آوردنویس 3 استخراج و کلیه تجزیه و تحلیل‌ها با استفاده از نرم‌افزارهای RStudio و MATLAB انجام شده‌اند.

17 به دلیل بالا رفتن تعداد صفحات، امکان بررسی همه شرکت‌ها و ارائه گزارش برای همه آن‌ها وجود نداشت. ر.ک. به (Cerqueti et al., 2023; D'Urso et al., 2015).

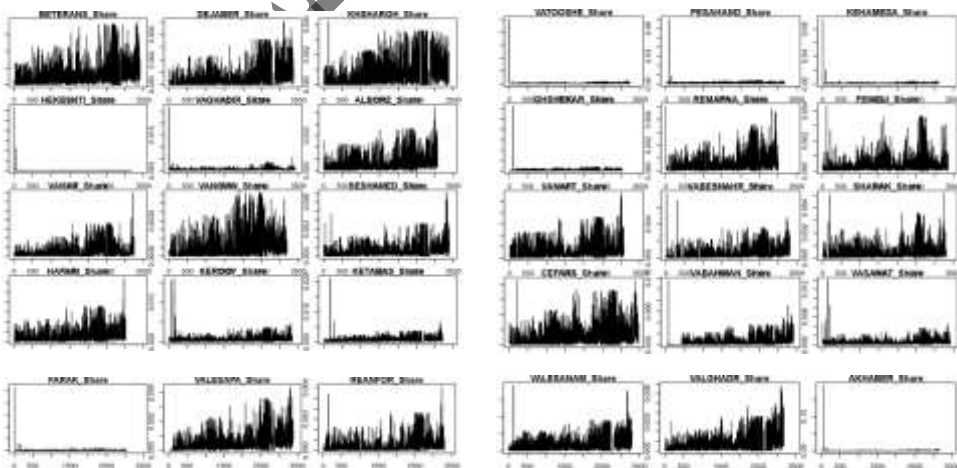
جدول 3. اسامی صنایع و نماد شرکت‌های منتخب
مأخذ: نتایج پژوهش

Table 3. Name of selected industries and symbols of selected companies

Source: Research results

نماد شرکت	نام صنعت	نماد شرکت	نام صنعت
و صنعت	سرمایه‌گذاری‌ها	ولصنم	واسطه‌گری‌های مالی و پولی
ولسایا	واسطه‌گری‌های مالی و پولی	ولغدر	واسطه‌گری‌های مالی و پولی
رانفور	رایانه	اخابر	مخابرات
بترانس	دستگاه‌های برقی	فاراگ	محصولات فلزی
دجابر	دارویی	وتوشه	ماشین‌آلات و تجهیزات
خشرق	خودرو و قطعات	پسهند	لاستیک و پلاستیک
حکشتی	حمل و نقل انبارداری و ارتباطات	کهمدی	کالی غیر فلزی
وغذیر	چندرشته‌ای صنعتی	قشکر	قند و شکر
البرز	بیمه و بازنشتگی	رمینا	فنی و مهندسی
وکار	بانک‌ها و مؤسسات اعتباری	فملی	فلزات اساسی
ونوین	بانک‌ها و مؤسسات اعتباری	ولفت	فراورده‌های نفتی
ثشاهد	انبوهسازی املاک و مستغلات	ویشیر	غذایی بجز قند و شکر
حفاری	استخراج نفت گاز جز اکتشاف	شاراک	شیمیایی
کروی	استخراج کانه‌های فلزی	سفارین	سیمان آهک گچ
کطیس	استخراج زغال‌سنگ	وبهنم	سرمایه‌گذاری‌ها

(2) تخمین‌های به‌دست‌آمده برای نوسانات سهام شرکت‌ها در شکل 1 به تصویر کشیده شده است. آماره‌های توصیفی شامل میانگین و واریانس لگاریتم قیمت‌های روزانه ماکزیمم و مینیمم و همچنین متوسط مقدار نوسانات که از رابطه (2) برحسب درصد به‌دست‌آمده‌اند، در جدول 4 گزارش شده است. ملاحظه می‌شود که سهام "پسهند" و "حکشتی" بیشترین مقدار نوسانات و سهام "فملی" و "وغذیر" پایین‌ترین متوسط مقدار نوسانات را طی دوره زمانی پژوهش دارا هستند. هرچند که اکثر سهام‌ها مقادیر نوسانات متوسط مشابهی را نشان می‌دهند، اما هدف اصلی در این پژوهش خوشه‌بندی مجموعه داده‌ها بر اساس مقادیر مختلف سرریزهای نوسانات "از سوی" و "به سوی" سهام‌ها است.



شکل 1. سری‌های زمانی نوسانات تخمین زده شده سهام شرکت‌های منتخب
مأخذ: رابطه (2) و نتایج پژوهش

Figure 1. Estimated volatility time series of stocks of selected companies
Source: Research results from (2)

همان‌گونه که دیبولد و بیلماز (2012) نشان داده‌اند شاخص‌های سرریز چندان تحت تأثیر انتخاب مرتبه P در مدل VAR و انتخاب افق H -گام به جلو برای محاسبه $GFEVD$ قرار نمی‌گیرند (Diebold & Yilmaz, 2012). این محققین، افق $H = 10$ را در پژوهش خود پیشنهاد کرده‌اند. بر این اساس و در این پژوهش بر طبق معیار اطلاعات آکاییک (AIC)، وقفه بهینه $P = 5$ انتخاب می‌شود. سپس، تخمین‌های دو شاخص سرریز نوسانات "از سوی" و "به سوی" بر حسب درصد، در جدول 5 نشان داده می‌شود.

از جدول 5، ملاحظه می‌شود که سهام "حکشتی" و "قشکر" به دلیل داشتن به ترتیب مقدار بسیار بزرگ و مقدار کوچک از هر دو معیار سرریز، پرت محسوب می‌شوند.

جدول 4. آماره‌های توصیفی قیمت سهام شرکت‌های منتخب
مأخذ: نتایج پژوهش

Table 4. Descriptive Statistics of log prices of stocks of selected companies
Source: Research results

میانگین نوسانات (%)	مینیم قیمت		ماکزیم قیمت		تعداد مشاهدات	
	انحراف معیار	میانگین	انحراف معیار	میانگین		
0/0719	0/8589	6/6794	0/8651	6/717	3458	ولسنم
0/0783	1/1438	6/5437	1/1505	6/5836	3415	ولغدر
0/0503	1/1526	7/2821	1/1598	7/3113	3413	اخابر
0/0739	0/6143	6/7039	0/6169	6/7421	3405	فارااک
0/0661	1/8153	6/0013	1/8247	6/0363	3392	وتوشه
0/0812	1/5388	7/9793	1/5442	8/0195	3383	پسپهند
0/0657	1/3170	6/3557	1/3241	6/3902	3378	کههدا
0/0769	1/3221	5/9539	1/3254	5/9919	3369	قشکر
0/0550	1/4599	7/2058	1/4686	7/2392	3358	رمینا
0/0402	1/0161	6/5504	1/0185	6/5787	3334	فملی
0/0668	1/0731	6/6238	1/0806	6/6602	3316	ونفت
0/0589	1/0365	7/0711	1/0452	7/1043	3309	وبشهر
0/0440	1/5600	7/6496	1/5665	7/6791	3304	شاراک
0/0589	1/4385	7/2273	1/4456	7/2616	3244	سفارس
0/0608	1/3293	5/7508	1/3398	5/7826	3226	وبهمین
0/0537	0/9628	5/8402	0/9702	5/8731	3211	وصنعت
0/0582	0/4886	6/7414	0/4962	6/7759	3191	ولسایا
0/0468	1/6218	6/5623	1/6298	6/5911	3179	رانفور
0/0574	0/7207	6/5736	0/7258	6/6079	3170	بترانس
0/0541	1/4687	7/2531	1/4791	7/2829	3163	دجابر
0/0712	1/0178	7/3611	1/0250	7/3972	3063	خشرق
0/0797	1/1842	7/1661	1/1921	7/1967	3181	حکشتی
0/0431	1/3868	7/0908	1/3909	7/1208	3026	وغدیر
0/0635	1/3677	5/2877	1/3742	5/3237	3418	البرز
0/0489	0/7962	6/3990	0/7989	6/4289	3003	وکار
0/0544	0/8141	6/9515	0/8224	6/9831	3296	ونوین
0/0608	1/3638	7/1919	1/3681	7/2267	3120	نشاهد
0/0653	0/9026	7/2120	0/9097	7/2491	3040	حقاری
0/0656	1/2282	6/4613	1/2323	6/4982	3019	کروی
0/0740	1/5377	7/1775	1/5472	7/2151	3073	کطیس

شکل 2، نتایج جدول 5 را برحسب مقادیر استانداردشده، به تصویر می‌کشد: سهامی که مقادیر (استاندارد شده) سرریز بالای صفر دارد، سرریز را بیشتر از متوسط، به دیگران می‌دهد (دریافت می‌کند)؛ درحالی‌که سهامی با مقادیر (استاندارد شده) سرریز کمتر از صفر، سرریز را کمتر از مقدار متوسط، به دیگر سهامها می‌دهد (دریافت می‌کند). با نگاه به شکل 2 از سمت محور X (سرریز "از سوی")، تفاوت بین سهام شرکت‌ها بیشتر نمایان می‌شود؛ سهام "حکشتی"، "اخبر" و "کهمد" دورتر از سایر سهامها و در گوشه سمت راست و بالای نمودار تنها قرار گرفته‌اند. واضح است که این سهامها واحدهای پرت محسوب می‌شوند. به‌رحال دیده می‌شود که سهام شرکت "قشکر" با داشتن مقادیر خیلی کوچک برای هر دو معیار سرریز "از سوی" و سرریز "به سوی"، در گوشه سمت چپ و پایین نمودار تنها قرار گرفته است. در واقع، این شرکت هم می‌تواند یک کاندید برای واحد پرت محسوب شود اما پرت بودن سه سهام متکثر مشهودتر است.

در این پژوهش به علت وجود سهام کاندید به‌عنوان واحد پرت، عملکرد دو مدل VS-FCMdd و VS-E-FCMdd در فرآیند خوشه‌بندی فازی مقایسه می‌گردد. برای انتخاب تعداد بهینه خوشه‌ها نیز از شاخص ژنی-بنی استفاده می‌شود (Xie & Beni, 1991)؛ به این صورت که دو الگوریتم خوشه‌بندی به ازای مقادیر $C = \{2, 3, 4, 5\}$ و $m = 1/5$ اجرا و شاخص ژنی-بنی محاسبه می‌شود. بر طبق نتایج در شکل 3، مقدار شاخص ژنی-بنی تحت هر دوی این مدل‌ها در $C = 2$ مینیمم می‌شود. لازم به ذکر است که مقدار این شاخص در $C = 5$ و برای مدل VS-FCMdd، بی‌نهایت بدست آمده است. همچنین، مدل VS-FCMdd وزن‌های $\omega_1 = 0/30$ و $\omega_2 = 0/70$ و مدل استوار VS-E-FCMdd وزن‌های $\omega_1 = 0/82$ و $\omega_2 = 0/18$ را به معیار عدم مشابهت (11)، اختصاص می‌دهند.

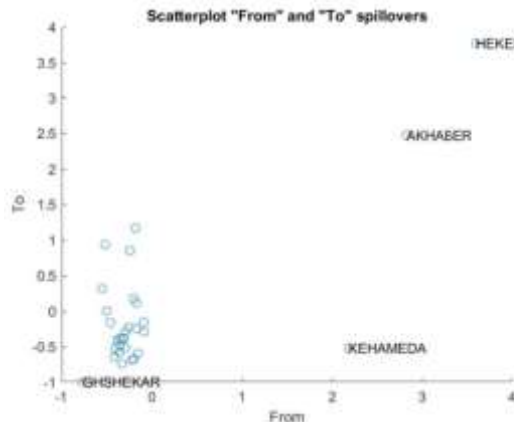
در جدول 6 نتایج حاصل برای منویدهای هر خوشه و درجه عضویت سهامها نشان داده شده است. یک تفاوت قابل توجه در نتایج به دست آمده از دو مدل خوشه‌بندی برای واحدهای پرت مشاهده می‌شود؛ درجه عضویت حاصل شده از مدل VS-E-FCMdd در مقایسه با مدل VS-FCMdd برای همه سهامها و در هر دو خوشه کوچکتر است. بالأخص، مدل VS-FCMdd، دو سهام "حکشتی" و "اخبر" را به خوشه 2، به ترتیب با درجه عضویت 0/5842 و 0/6121 و سهام "کهمد" را به خوشه 1 با درجه عضویت 0/6453

جدول 5. تخمین سرریزهای نوسانات جهت‌دار: "از سوی" و "به سوی" سایر سهام
مأخذ: نتایج پژوهش

Table 5. Estimations of Directional spillover volatilities: "From" and "To" others
Source: Research results

"به سوی"	"از سوی"	سهام
0/3476	0/3211	ولصنم
0/3093	0/3017	ولغدر
1/1195	1/4828	اخبر
0/6809	0/3614	فراک
0/7669	0/3838	وتوشه
0/5359	0/2483	پسهند
0/3106	1/2488	کهمد
0/1817	0/1691	قشکر
0/4994	0/3760	رینیا
0/3480	0/3367	قملی
0/3736	0/3436	ونفت
0/2657	0/3700	وبشهر
0/3211	0/3219	شاراک
0/2499	0/3297	سفارس
0/3407	0/3087	وبهم
0/3889	0/3559	وصنعت
0/4828	0/3874	ولسایا
0/4071	0/2817	رانفور
0/2946	0/3184	بترانس

0/3721	0/4198	دجابر
0/3458	0/3297	خشرق
1/4698	1/7681	حکشتی
0/3849	0/3882	وغدیر
0/2689	0/3777	البرز
0/3124	0/3426	وکار
0/2774	0/2987	ونوین
0/2899	0/3935	نشاهد
0/4091	0/4174	حفاری
0/7039	0/2604	کروی
0/4508	0/2657	کطیس



شکل 2. نمودار پراکنندگی شاخص‌های سرریز نوسان‌پذیری "از سوی" و "به سوی" (استاندارد شده). سهام شرکت‌های "حکشتی"، "اخیر"، "کهمد" و "قشکر" بر اساس هر دوی این شاخص‌ها پرت محسوب می‌شوند.

مأخذ: یافته‌های پژوهش

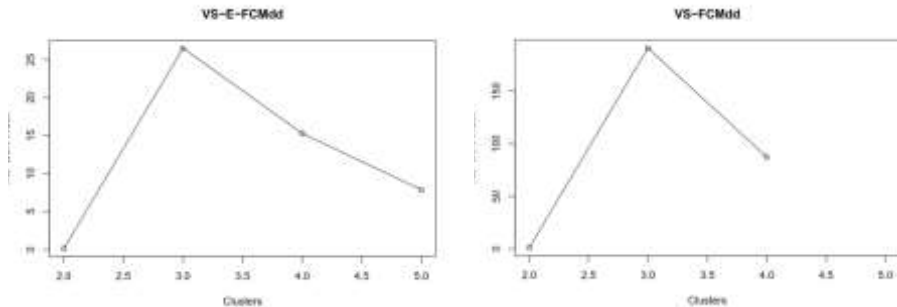
Figure 2. Scatterplot of "From" and "To" volatility spillover indicators (in standardized value). Both indicators suggest that the HEKESHTI, AKHABER, KEHAMEDA, and GHGHSHEKAR stocks are outliers.

Source: Research results

اختصاص می‌دهد؛ اما مدل استوار VS-E-FCMdd این سهام را به هر دو خوشه 1 و 2 با درجه عضویت 0/5 اختصاص می‌دهد؛ علت آن به اثرات هموارکنندگی تبدیل نمایی از عدم مشابهت (رابطه 11))، مربوط می‌شود. به بیان دیگر، مدل VS-E-FCMdd واحدهای پرت را با اختصاص درجه عضویتی نزدیک به مقدار $1/C$ شناسایی می‌کند که در یک چارچوب مبتنی بر رویکرد استوار، دور از انتظار نخواهد بود؛ برعکس، مدل VS-FCMdd به اشتباه واحد پرت را به خوشه با درجه عضویت بالاتر، اختصاص می‌دهد.

بعلاوه، مدل VS-E-FCMdd یک درجه عضویت کوچک حدود 0/5 را برای سهام "قشکر" نیز اختصاص داده است. به دلیل اینکه مقادیر سرریز جهت‌دار "از سوی" و "به سوی" برای این سهام در مقایسه با سایرین پایین‌تر بوده است، در سمت پایین و چپ 0 قرار گرفته است.

اما مدل VS-FCMdd سهام "قشکر" را به خوشه 1 با درجه عضویت بالاتر 0/8378 اختصاص می‌دهد. در کنار نتایج خوشه‌بندی فوق، با مراجعه به 0 می‌توان اینگونه



شکل 3. مقادیر شاخص زی-بنی (1991) (محور y) حاصل از تعداد خوشه‌های مختلف (محور x) به تصویر کشیده شده است. شکل سمت راست مربوط به مدل VS-FCMdd و شکل سمت چپ مربوط به مدل VS-E-FCMdd است. تعداد بهینه خوشه عددی است که شاخص زی-بنی را مینیمم می‌سازد. برای مدل VS-FCMdd، مقدار این شاخص در $C = 5$ بی‌نهایت شده است. مأخذ: یافته‌های پژوهش

Figure 3. the Xie-Beni index obtained for different values of clusters is depicted. VS-FCMdd is in the right and VS-E-FCMdd is in the left. The optimal number of clusters minimizes the Xie-Beni index. For the VS-FCMdd, the value of this index at $C=5$ is infinite.

Source: Research results

استدلال نمود که به هر حال این سهام از لحاظ سرریزهای نوسانات واحد پرت محسوب می‌شود. سهام "تشکر" جزو دسته شرکت‌هایی در بورس تهران است که از ارزش بازاری کوچکی برخوردار می‌باشد؛ از این رو هر دو مقدار سرریز جهت‌دار برای این سهام کمترین مقدار را دارد؛ علی‌رغم اینکه این سهام بیشتر از مقداری که از دیگران دریافت می‌کند، به آن‌ها سرریز می‌دهد اما این مقدار هم از سایر سهامها بسیار کوچکتر است. بنابراین، به نظر می‌رسد افزایش حاصل شده از VS-FCMdd از قدرت توضیح دهندگی ضعیف‌تری نسبت به افزایش حاصل از مدل استوار، برخوردار است.

همچنین، از 0 ملاحظه می‌شود که در مدل خوشه‌بندی VS-FCMdd، واحدهای فازی دیگری، بجز آن‌هایی که می‌توان به عنوان پرت در نظر گرفت، وجود ندارند. در مقابل، در مدل VS-E-FCMdd سهام‌های "وتوشه"، "فاراک"، "گروی" و "پسهند" فازی هستند زیرا درجه عضویت بسیار پایین دارند و می‌توانند به خوشه 1 و خوشه 2 متعلق باشند (D'Urso et al., 2013).

جدول 6. نتایج حاصل از خوشه‌بندی بر اساس مدل‌های FCMdd و Exp-FCMdd: مدویدها و درجات عضویت مأخذ: نتایج پژوهش

Table 6. Clustering results from FCMdd AND Exp-FCMdd: Medoids and membership degrees

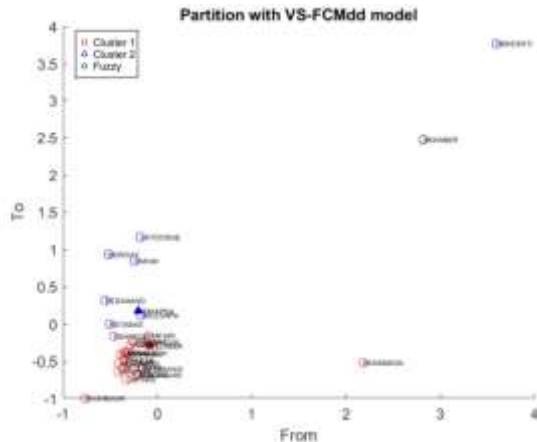
VS-E-FCMdd		VS-FCMdd		سهام
ونفت (خوشه 2)	وسنعت (خوشه 1)	رمینا (خوشه 2)	داجار (خوشه 1)	
0/8511	0/1489	0/0102	0/9897	ولصنم
0/7304	0/2696	0/0299	0/9701	ولغدر
0/5	0/5	0/6121	0/3879	اخابر
0/4227	0/5773	0/8943	0/1057	فاراک
0/4218	0/5782	0/8263	0/1737	وتوشه
0/5905	0/4095	0/9822	0/0178	پسهند
0/4999	0/5000	0/3547	0/6453	کههدا
0/5535	0/4465	0/1622	0/8378	تشکر
0/2333	0/7667	1	0	رمینا
0/9699	0/0301	0/0064	0/9936	فملی
1	0	0/0045	0/9955	ونفت

0/3565	0/6435	0/0445	0/9555	وبشهر
0/8374	0/1626	0/0180	0/9819	شاراک
0/7486	0/2514	0/0644	0/9356	سفارس
0/7706	0/2294	0/0155	0/9845	وبهمن
0	1	0/0070	0/9930	وصنعت
0/2467	0/7533	0/9994	0/0006	ولسایا
0/6644	0/3356	0/1790	0/8210	رانفور
0/7915	0/2085	0/0334	0/9666	بترانس
0/3368	0/6632	0	1	دجابر
0/9190	0/0810	0/0083	0/9917	خشرق
0/5	0/5	0/5842	0/4158	حکشتی
0/2156	0/7844	0/0007	0/9993	وغدیر
0/3261	0/6739	0/0409	0/9591	الیرز
0/8632	0/1368	0/0173	0/9827	وکار
0/7083	0/2917	0/0510	0/9490	ونوین
0/3106	0/6894	0/0240	0/9760	نثشاهد
0/3270	0/6730	0/0255	0/9745	حفاری
0/5597	0/4403	0/8704	0/1296	کروی
0/6256	0/3744	0/8407	0/1593	کطیس

از آنجایی که این سهم‌ها در بین سایر سهم‌ها و در مرکز 0 قرار گرفته‌اند، واحدهای فازی واقعی به نظر می‌رسند.

به منظور بررسی افزایشها با جزئیات بیشتر، نتایج خوشه‌بندی‌های فوق در شکل 4 و شکل 50 به تصویر کشیده شده است. سهم‌هایی که متعلق به هر کدام از دو خوشه هستند با رنگ قرمز و آبی مشخص شده‌اند، اما سهم‌های با درجه عضویت بسیار پایین، با رنگ سیاه مشخص شده‌اند. به نظر می‌رسد افزایش به دست آمده از مدل VS-E-FCMdd، یک الگوی سرریز "کم" را در مقابل سرریز "زیاد" نشان می‌دهد؛ به این صورت که سهم‌های قرار گرفته در سمت راست و بالای نمودار دارای سطوح سرریز "زیاد" و سهم‌های قرار گرفته در سمت راست و پایین نمودار سهم‌هایی هستند که از سطوح "کم" سرریز برخوردار می‌باشند. اما در مقابل، افزایش به دست آمده از مدل VS-FCMdd، مشاهدات را به دو خوشه، یک خوشه با سرریز "از سوی" زیاد و خوشه دیگر با سرریز "از سوی" کم گروه‌بندی نموده است؛ سهم‌های قرار گرفته در سمت راست و پایین نمودار دارای سطوح سرریز "از سوی" زیاد ("به سوی" کم) و سهم‌های قرار گرفته در سمت چپ و بالای نمودار سهم‌هایی هستند که از سطوح "از سوی" کم ("به سوی" زیاد) برخوردار هستند.

شکل 4 افزایشها به دست آمده از مدل VS-FCMdd که یک مدل استوار نیست را نشان می‌دهد. ملاحظه می‌شود که سهام "حکشتی" به خوشه 2 با رنگ آبی و سهام "کهمدا" و "قشکر" به خوشه 1 با رنگ قرمز اختصاص داده شده‌اند؛ هر چند که این سهم‌ها از مدویدهای خوشه مربوطه که به ترتیب با مربع قرمز رنگ و مثلث آبی رنگ مشخص شده‌اند، خیلی فاصله دارند. افزایش فوق الگوی سرریز "از سوی" زیاد را در مقابل الگوی سرریز "از سوی" کم تأیید می‌کند اما مشاهدات حدی هم وجود دارد که به اشتباه به خوشه‌ها اختصاص داده شده‌اند؛ سهام "اخبر" علی‌رغم داشتن مقادیر زیاد سرریزهای "از سوی" و "به سوی"، به دلیل اینکه درجه عضویتی کمتر دارد، فازی محسوب می‌شود که با رنگ سیاه نشان داده شده است؛ در واقع، سهام "اخبر" که در نقطه میانه دو افزایش مشخص شده در شکل 4 قرار گرفته به عنوان واحد فازی تشخیص داده شده است.



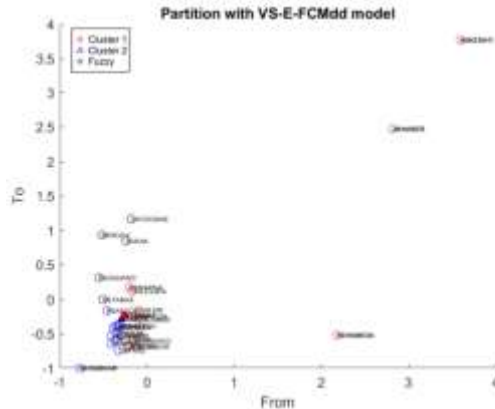
شکل 4. افرازهای به دست آمده از VS-FCMdd. سهامهای فرار گرفته در خوشه 1 و خوشه 2 به ترتیب با رنگ قرمز و رنگ آبی مشخص شده‌اند. سهامهای با رنگ سیاه واحدهایی را مشخص می‌کنند که درجه عضویت بسیار کوچک دارند که براین اساس می‌توانند به عنوان واحدهای فازی در نظر گرفته شوند.

مأخذ: یافته‌های پژوهش و (Cerqueti et al., 2023)

Figure 4. Partitions estimated from VS-FCMdd. The stocks in clusters 1 and 2 are indicated in red and blue, respectively. The black stocks are units with low degrees of membership, which could be considered fuzzy units.

Source: Research results and (Cerqueti et al., 2023)

برعکس، شکل 5 نشان می‌دهد که مدل VS-E-FCMdd یک افراز به دو خوشه با سرریزهای جهت‌دار "کم" در مقابل سرریزهای جهت‌دار "زیاد" را نتیجه می‌دهد؛ اما مشاهدات حدی را از این خوشه‌بندی جدا کرده است. در واقع، سهامهای فرار گرفته در قسمت بالا و چپ شکل 5، با رنگ سیاه مشخص شده‌اند و باقیمانده سهامها به خوشه‌ها اختصاص داده شده‌اند. از این رو، می‌توان به افراز قابل قبول‌تری برحسب خوشه‌های "کم" و "زیاد" (از نظر سرریز) دست یافت. بر طبق افراز حاصل از مدل VS-E-FCMdd، سهامهای "وتوشه"، "فاراک"، "کروی" و "پسهند"، فازی به دست آمده‌اند. این سهامها مقدار سرریز جهت‌دار "از سوی" زیر صفر دارند، اما مقدار سرریز "به سوی" بالای صفر دارند. به علاوه، همه سهامهای مشخص شده با رنگ قرمز و آبی در شکل 5، حول مدویدهای خوشه متعلق به آن بطور فشرده‌تر قرار گرفته‌اند. از این رو، به نظر می‌رسد مدل خوشه‌بندی استوار عملکرد برتری در خوشه‌بندی مجموعه داده‌ها دارد.



شکل 5. افزایش‌های به دست آمده از VS-E-FCMdd. سهم‌های قرار گرفته در خوشه 1 و خوشه 2 به ترتیب با رنگ قرمز و رنگ آبی مشخص شده‌اند. سهم‌های با رنگ سیاه واحدهایی را مشخص می‌کنند که درجه عضویت بسیار پایین دارند که بر این اساس می‌توانند به عنوان واحدهای فازی در نظر گرفته شوند.
مأخذ: یافته‌های پژوهش و (Cerqueti et al., 2023)

Figure 5. Partitions estimated from VS-E-FCMdd. The stocks in clusters 1 and 2 are indicated in red and blue, respectively. The black stocks are units with low degrees of membership, which could be considered fuzzy units.

Source: Research results and (Cerqueti et al., 2023)

5- نتیجه‌گیری

ارائه یک تعریف مناسب از معیار عدم مشابهت میان سری‌های زمانی در مدل‌های خوشه‌بندی مبتنی بر عدم مشابهت از اهمیت زیادی برخوردار است. هنگامی که سری‌های زمانی مالی مورد توجه باشد، برای رسیدن به یک خوشه‌بندی قابل قبول، قواعد تجربی سری‌های زمانی مالی نیز می‌باید مورد توجه قرار بگیرد. در بخش ادبیات نظری بسیاری از رویکردهای خوشه‌بندی که این قوانین تجربی را در نظر گرفته‌اند، معرفی گردید. با این حال، هنوز هم قوانین تجربی مهمی وجود دارد که تاکنون برای خوشه‌بندی در نظر گرفته نشده‌اند، اما از اهمیت خاصی برای متخصصین در صنعت مالی و سیاست‌گذاران برخوردار هستند. یکی از این موارد سرریز نوسانات است.

سهم اصلی پژوهش حاضر در ادبیات مربوطه، معرفی دو مدل خوشه‌بندی فازی مبتنی بر درجه و جهت سرریز نوسانات است. از نقطه نظر روش‌شناسی، در پژوهش حاضر سرریزهای نوسانات جهت‌دار با استفاده از رویکرد تجزیه واریانس خطای پیش‌بینی (GFEVD) اندازه گرفته می‌شوند. یک معیار عدم مشابهت بین سهم‌ها نیز بر اساس سرریزهای نوسانات جهت‌دار معرفی می‌شود. سپس در یک چارچوب فازی، دو مدل خوشه‌بندی PAM پیشنهاد می‌شود که مجموعه وزن‌های بهینه را به طور تکراری محاسبه می‌کند بطوریکه بر اساس آن درجه عضویت که گروه‌بندی نهایی مشاهدات را تعیین می‌کند، محاسبه می‌شود. اولین الگوریتم در نظر گرفته شده مدل VS-FCMdd است. بعلاوه، مدل VS-E-FCMdd نیز پیشنهاد شده است که با استفاده از یک تبدیل نمایی از عدم مشابهت وزنی، در مقابل داده‌های پرت استوار خواهد بود.

این دو مدل خوشه‌بندی نخستین بار برای بررسی ساختار بورس تهران از نظر سرریزهای نوسانات بکار برده می‌شود. بر طبق نتایج تجربی، هر دوی مدل‌های خوشه‌بندی یک ساختار دو گروهی را در مجموعه داده‌ها تأیید می‌کنند، یعنی سطوح "زیاد" و "کم" سرریزهای جهت‌دار. به‌رحال به دلیل وجود برخی سهام پرت، مدل استوار VS-E-FCMdd افزایش قابل قبول‌تری حاصل می‌کند. یافته‌ها

همچنین نشان می‌دهد که سهام شرکت‌های فعال در بورس تهران عمدتاً برحسب سرریز "از سوی" تمییز داده می‌شوند، یعنی مقداری از سرریز که هر سهام از دیگر سهام دریافت می‌کند. این شواهد تجربی از بورس تهران با بهره گرفتن از سیستم وزن‌دهی به دست آمده است که در خلال الگوریتم خوشه‌بندی محاسبه می‌شود بطوریکه مدل VS-E-FCMdd وزن بیشتر را به بُعد سرریز "از سوی" اختصاص می‌دهد.

یکی از عمده محدودیت‌های پیش روی پژوهش حاضر، استفاده از یک معیار ایستا برای اندازه‌گیری سرریزهای جهت‌دار است که پیشنهاد می‌شود در پژوهش‌های آتی با استفاده از تئوری پنجره غلتان، شاخص‌های سرریز پویا محاسبه شوند. در واقع، آنالیز مبتنی بر معیارهای پویا بینش ارزشمندی از وابستگی متقابل و سرایت ریسک درون سیستم‌های مالی فراهم می‌کند که در شرایط وقوع شوک‌های برون‌زا و دوره‌های اشفتگی بازار می‌تواند بوجود آید. براین اساس، می‌توان الگوریتم‌های خوشه‌بندی دیگری را با هدف جایگزینی یک معیار سرریز پویا در تابع هدف، تعمیم داد. بعلاوه، پیشنهاد می‌شود از معیارهای سرریز دیگری علاوه بر معیارهای سرریز زمان متغیر برای خوشه‌بندی استفاده شود. پیشنهاد دیگر، تعمیم مدل‌های خوشه‌بندی استوار جدید به‌عنوان جایگزین برای مدل VS-E-FCMdd، است. تکنیک‌های خوشه‌بندی استوار برای پرداختن به سری‌های زمانی مالی بسیار کاربردی و مفید می‌باشند که نتایج پژوهش حاضر نیز بر صحت این ادعا تاکید می‌کند. در پایان، پیشنهاد می‌شود مدل‌های خوشه‌بندی دیگری در چارچوب فازی، برای گروه‌بندی سری‌های زمانی نامانا و با ساختار غیرخطی، تعمیم داده شود.

Acknowledgments: we appreciate Reviewers' insightful suggestions on improving the quality of the research.

Conflict of Interest: The author declares no conflict of interest.

Funding: The author received no financial support for the research, authorship, and publication of this article.

Reference

- Alonso, A. M., D'Urso, P., Gamboa, C., & Guerrero, V. (2021). Cophenetic-based fuzzy clustering of time series by linear dependency. *International Journal of Approximate Reasoning*, 137, 136-114. doi:<https://doi.org/10.1016/j.ijar.2021.07.006>
- Asghari, F & ,Ahmadi, F. (2018). A hybrid model for appraising and forecasting loan repayments (case study: karafarini omid fund). *Economics Research*, 18(71), 223-185. doi:10.22054/joer.2018.9833 [in persian]
- Bastos, J. A., & Caiado, J. (2014). Clustering financial time series with variance ratio statistics. *Quantitative Finance*, 14(12), 2133-2121. doi: 14697688.2012.726736/10.1080
- Bastos, J. A., & Caiado, J. (2021). On the classification of financial data with domain agnostic features. *International Journal of Approximate Reasoning*, 138, 11-11. doi:<https://doi.org/10.1016/j.ijar.2021.07.008>
- Bei, H., Mao, Y., Wang, W., & Zhang, X. (2021). Fuzzy clustering method based on improved weighted distance. *Mathematical Problems in Engineering*, 2021, 6687202. doi: 6687202/2021/10.1155

- Buncic, D., & Gisler, K. I. M. (2016). Global equity market volatility spillovers: A broader role for the United States. *International Journal of Forecasting*, 32(4), 1339-1317. doi:<https://doi.org/10.1016/j.ijforecast.2016.05.001>
- Caiado, J., Crato, N., & Peña, D. (2006). A periodogram-based metric for time series classification. *Computational Statistics & Data Analysis*, 50(10), 2684-2668. doi:<https://doi.org/10.1016/j.csda.2005.04.012>
- Caiado, J., Maharaj, E. A., & D'Urso, P. (2015). Time-series clustering. In *Handbook of Cluster Analysis* (pp. 285-262). Chapman and Hall/CRC.
- Cerqueti, R., D'Urso, P., De Giovanni, L., Giacalone, M., & Mattera, R. (2022). Weighted score-driven fuzzy clustering of time series with a financial application. *Expert Systems with Applications*, 198, 116752. doi:<https://doi.org/10.1016/j.eswa.2022.116752>
- Cerqueti, R., D'Urso, P., De Giovanni, L., Mattera, R., & Vitale, V. (2023). Fuzzy clustering of financial time series based on volatility spillovers. *Annals of Operations Research*. doi:10.1007/s 10479-023-05560-7
- Cerqueti, R., & Mattera, R. (2023). Fuzzy clustering of time series with time-varying memory. *International Journal of Approximate Reasoning*, 153, 218-193. doi:<https://doi.org/10.1016/j.ijar.2022.11.021>
- Chang, C.-L., McAleer, M., & Tansuchat, R. (2013). Conditional correlations and volatility spillovers between crude oil and stock index returns. *The North American Journal of Economics and Finance*, 25, 138-116. doi:<https://doi.org/10.1016/j.najef.2012.06.002>
- Chen, Y., Chiu, J., & Chung, H. (2022). Givers or receivers? return and volatility spillovers between Fintech and the Traditional Financial Industry. *Finance Research Letters*, 46, 102458. doi:<https://doi.org/10.1016/j.frl.2021.102458>
- Cheuathonghua, M., Padungsaksawasdi, C., Boonchoo, P., & Tongurai, J. (2019). Extreme spillovers of VIX fear index to international equity markets. *Financial Markets and Portfolio Management*, 33(1), 38-31. doi:10.1007/s 11408-018-0323-6
- Choi, S.-Y. (2022). Dynamic volatility spillovers between industries in the US stock market: Evidence from the COVID- 19pandemic and Black Monday. *The North American Journal of Economics and Finance*, 59, 101614. doi:<https://doi.org/10.1016/j.najef.2021.101614>

- Coppi, R., D'Urso, P., & Giordani, P. (2010). A fuzzy clustering model for multivariate spatial time series. *Journal of Classification*, 27, 88-54. doi:<https://doi.org/10.1007/s00357-010-9043-y>
- D'Urso, P. (2005). Fuzzy clustering for data time arrays with inlier and outlier time trajectories. *IEEE Transactions on Fuzzy Systems*, 13(5), .604-583doi:10.1109/TFUZZ.2005.856565
- D'Urso, P., De Giovanni, L., Maharaj, E. A., Brito, P., & Teles, P. (2023). Wavelet-based fuzzy clustering of interval time series. *International Journal of Approximate Reasoning*, 152, 159-136. doi:<https://doi.org/10.1016/j.ijar.2022.09.014>
- D'Urso, P., De Giovanni, L., & Massari, R. (2015). Time series clustering by a robust autoregressive metric with application to air pollution. *Chemometrics and Intelligent Laboratory Systems*, 141, 124-107. doi:<https://doi.org/10.1016/j.chemolab.2014.11.003>
- D'Urso, P., De Giovanni, L., & Massari, R. (2016). GARCH-based robust clustering of time series. *Fuzzy Sets and Systems*, 305, 28-21. doi:<https://doi.org/10.1016/j.fss.2016.01.010>
- D'Urso, P., Maharaj, E. A., & Alonso, A. M. (2017). Fuzzy clustering of time series using extremes. *Fuzzy Sets and Systems*, 318, 79-56. doi:<https://doi.org/10.1016/j.fss.2016.10.006>
- D'Urso, P., Massari, R., Cappelli, C., & De Giovanni, L. (2017). Autoregressive metric-based trimmed fuzzy clustering with an application to PM 10time series. *Chemometrics and Intelligent Laboratory Systems*, 161, 26-15. doi:<https://doi.org/10.1016/j.chemolab.2016.11.016>
- D'Urso, P., De Giovanni, L., & Massari, R. (2021). Trimmed fuzzy clustering of financial time series based on dynamic time warping. *Annals of Operations Research*, 299, 1395-1379. doi:<https://doi.org/10.1007/s10479-019-03284-1>
- D'Urso, P., De Giovanni, L., Massari, R., D'Ecclesia, R. L., & Maharaj, E. A. (2020). Cepstral-based clustering of financial time series. *Expert Systems with Applications*, 161, 113705. doi:<https://doi.org/10.1016/j.eswa.2020.113705>
- D'Urso, P., De Giovanni, L., Massari, R., & Di Lallo, D. (2013). Noise fuzzy clustering of time series by autoregressive metric. *METRON*, 71(3), .243-217doi:10.1007/s40300-013-0024-x
- D'Urso, P., De Giovanni, L., & Vitale, V. (2023). A robust method for clustering football players with mixed attributes. *Annals of Operations Research*, 325(1), 36-39. doi:<https://doi.org/10.1007/s10479-022-04558-x>

- D'Urso, P & Maharaj, E. A. (2009). Autocorrelation-based fuzzy clustering of time series. *Fuzzy Sets and Systems*, 160(24), 3589-3565. doi:<https://doi.org/10.1016/j.fss.2009.04.013>
- Diebold, F. X., & Yilmaz, K. (2009). Measuring financial asset return and volatility spillovers, with application to global equity markets. *The Economic Journal*, 119(534), 171-158. doi:<https://doi.org/10.1111/j.1468-0297.2008.02208.x>
- Diebold, F. X., & Yilmaz, K. (2012). Better to give than to receive: Predictive directional measurement of volatility spillovers. *International Journal of Forecasting*, 28(1), 66-57. doi:<https://doi.org/10.1016/j.ijforecast.2011.02.006>
- Diebold, F. X., & Yilmaz, K. (2014). On the network topology of variance decompositions: Measuring the connectedness of financial firms. *Journal of Econometrics*, 182(1), 134-119. doi:<https://doi.org/10.1016/j.jeconom.2014.04.012>
- Ehsani, A., & Alemohammad, S. (2021). Fuzzy clustering of Tehran Stock Exchange Industries Index time series based on Autoregressive model. *4th International Conference on Soft Computing*. [in persian]
- Farid, D., & Pourhamidi, M. (2012). Classifying stocks of listed companies on tehran stock exchange using fuzzy cluster analysis. *Financial Accounting Research*, 4(3), 128-105. doi:[20.1001.1.2312/23405.1391.4.3.8.8](https://doi.org/10.2312/23405.1391.4.3.8.8) [in persian]
- Farnam, M., & Darehmiraki, M. (2023). A parametric method for fuzzy data clustering. *Fuzzy Systems and its Applications*, 5(2), 119-193. doi:[10.22034/jfsa.2023.331678.1138](https://doi.org/10.22034/jfsa.2023.331678.1138)[in persian]
- Gillaizeau, M., Jayasekera, R., Maaitah, A., Mishra, T., Parhi, M., & Volokitina, E. (2019). Giver and the receiver: Understanding spillover effects and predictive power in cross-market Bitcoin prices. *International Review of Financial Analysis*, 63, 104-186. doi:<https://doi.org/10.1016/j.irfa.2019.03.005>
- Jondeau, E., & Rockinger, M. (2012). On the importance of time variability in higher moments for asset allocation. *Journal of Financial Econometrics*, 10(1), 123-184. doi:[10.1093/jfinec/nbr006](https://doi.org/10.1093/jfinec/nbr006)
- Lafuente-Rego ,B., D'Urso, P., & Vilar, J. A. (2020). Robust fuzzy clustering based on quantile autocovariances. *Statistical Papers*, 61(6), 2448-2393. doi:[10.1007/s00362-018-1053-6](https://doi.org/10.1007/s00362-018-1053-6)
- López-Oriona, Á., Vilar, J. A., & D'Urso, P. (2022). Quantile-based fuzzy clustering of multivariate time series in the frequency domain. *Fuzzy Sets and Systems*, 443, 154-115. doi:[10.1016/j.fss.2022.02.015](https://doi.org/10.1016/j.fss.2022.02.015)

- Mantegna, R. N. (1999). Hierarchical structure in financial markets. *The European Physical Journal B - Condensed Matter and Complex Systems*, 11(1), 197-193. doi:10.1007/s 100510050929
- Otranto, E. (2008). Clustering heteroskedastic time series by model-based procedures. *Computational Statistics & Data Analysis*, 4698- , (10)52 .4685doi:[10.1016/j.csda.2008.03.020](https://doi.org/10.1016/j.csda.2008.03.020)
- Otranto, E., & Gargano, R. (2015). Financial clustering in presence of dominant markets. *Advances in Data Analysis and Classification*, 9(3), 339-315. doi:10.1007/s11634-014-0189-z
- Parkinson ,M. (1980). The extreme value method for estimating the variance of the rate of return. *The Journal of Business*, 53(1), 65-61. doi:[http://www.jstor.org/stable/ 2352357](http://www.jstor.org/stable/2352357)
- Pesaran, H. H., & Shin, Y. (1998). Generalized impulse response analysis in linear multivariate models. *Economics Letters*, 58(1), 29-17. doi:[https://doi.org/10.1016/S0165-1765\(97\)00214-0](https://doi.org/10.1016/S0165-1765(97)00214-0)
- Piccolo, D. (1990). A distance measure for classifying ARIMA models. *Journal of Time Series Analysis*, 11(2), 164-153. doi:[10.1111/j.1467-9892.1990.tb00048.x](https://doi.org/10.1111/j.1467-9892.1990.tb00048.x)
- Rafiei, S., & Moradi, P. (2016). Improving performance of fuzzy C-means clustering algorithm using automatic local feature weighting. *Tabriz Journal of Electrical Engineering*, 46(2), 86-75. [in persian]
- Rahmani Asl, F., Sinaei, H. a., & Neis, A. (2021). Forming an portfolio of Exchange Traded Funds with the Clustering and UTADIS Models. *Quarterly Journal of Quantitative Economics*, 18(1), 79-96. doi:10.22055/jqe.2020.32330.2204
- Rayati Shavazi, A., Blue, G., Ebrahimi Sarv-e-Olia, M. H., & Amiri, M. (2021). Financial contagion based on overlapping portfolio in TEPIX industrial groups. *Financial Management Strategy*, 9(3), 102-179. doi:10.22051/jfm.2019.27654.2175[in persian]
- Savvides, A., Promponas, V. J., & Fokianos, K. (2008). Clustering of biological time series by cepstral coefficients based distances. *Pattern Recognition*, 41(7), 2412-2398. doi:[10.1016/j.patcog.2008.01.002](https://doi.org/10.1016/j.patcog.2008.01.002)
- Sinaei, H. a., Mehrabi, A., Basiir zadeh, H., & Samandar, M. (2016). Appling an expected regret minimization in constructing portfolio of mutual funds. *Quarterly Journal of Quantitative Economics*, 13(1), 119-140. doi:10.22055/jqe.2016.12329
- Sinaei, H., Basirzad, H., & Samandar, M. (2014). Application of clustering analysis in assessing the performance of mutual funds. *Financial Management Strategy*, 2(1), 20-21. doi:[10.22051/jfm. 2015.974](https://doi.org/10.22051/jfm.2015.974)[in persian]

- Soltyk, S. J., & Chan, F. (2023). Modeling time-varying higher-order conditional moments: A survey. *Journal of Economic Surveys*, 37(1), 57-33. doi:[10.1111/joes.12481](https://doi.org/10.1111/joes.12481)
- Van der Laan, M., Pollard, K., & Bryan, J. (2003). A new partitioning around medoids algorithm. *Journal of Statistical Computation and Simulation*, 73(8), 584-575. doi: 0094965031000136012/10.1080
- Wu, K.-L., & Yang, M.-S. (2002). Alternative c-means clustering algorithms. *Pattern Recognition*, 35(10), 2278-2267. doi:[10.1016/S0031-3203\(01\)00197-2](https://doi.org/10.1016/S0031-3203(01)00197-2)
- Xie, X. L., & Beni, G. (1991). A new fuzzy clustering validity criterion and its application to color image segmentation. Paper presented at the Proceedings of the 1991 IEEE International Symposium on Intelligent Control.
- Zhu, L. (2016). Improvement of weighted Euclidean distance clustering algorithm. *Computer & Digital Engineering*, 3(44), 424-421