مقاله پژوهشی

پیش‌بینی جهت حرکت قیمت سهام با استفاده از شبکه توجه گراف جهت‌دار

علیرضا جعفری و سامان هراتی‌زاده

چکیده: پیش‌بینی رفتار آینده بازار سهام به عنوان یک چالش حائز اهمیت در یادگیری ماشین توجه زیادی را به خود جلب کرده است و رویکردهای یادگیری عمیق، نتایج قابل قبولی را در این زمینه به دست آورده‌اند. مطالعات پیشین، اهمیت درنظرگرفتن داده‌های سهام مرتبط را در خلال عملکرد پیش‌بینی اثبات کرده‌اند. با وجود این، مدل‌سازی روابط بین سهام به عنوان یک گراف جهت‌دار و ساخت بازنمایی گره‌های این گراف به کمک مکانیزم توجه تا کنون مورد استفاده قرار نگرفته است. ما در این کار، چهارچوبی به نام DeepNet را معرفی می‌کنیم که یک شبکه جهت‌دار دودویی را از تأثیرات داده‌های سهام در بهبود دقت پیش‌بینی یکدیگر ایجاد می‌کند و با استفاده از شبکه توجه گراف، اهمیت گره‌های همسایه برای ساخت بازنمایی‌ها را در حین عملیات آموزش، کشف می‌نماید. ما مدل جدیدی از شبکه توجه گراف را برای استفاده در گراف‌های جهت‌دار توسعه دادیم که قادر است اهمیت بردار ویژگی گره‌ها برای ساخت بازنمایی را به صورت یک‌طرفه در نظر بگیرد. نهایتاً ارزیابی‌های ما بر روی داده‌های بازار سهام تهران نشان می‌دهد که مدل معرفی‌شده از دقت و MCC بالاتری نسبت به مدل‌های رقیب برخوردار است.

***کلیدواژه:* پیش‌بینی سهام، شبکه توجه گراف، شبکه عصبی گراف، گراف جهت‌دار، مدل مبتنی بر شبکه، یادگیری عمیق، یادگیری نیمه نظارت‌شده.**

۱- مقدمه

پیش‌بینی وضعیت آینده سهام به عنوان یک چالش حائز اهمیت در یادگیری ماشین همواره مورد توجه بسیاری از محققان بوده است؛ زیرا کمی بهبود دقت در پیش‌بینی می‌تواند به سودآوری عظیمی منجر گردد [۱]. تصمیم‌گیری برای سرمایه‌گذاری در یک سهام منوط به شناخت جهت حرکت آینده آن است و از این رو بسیاری از معامله‌گران بیش از آن که به ارزش دقیق یک سهام در آینده اهمیت بدهند، مشتاق هستند که درباره جهت تغییر قیمت آن بدانند [۲]. اخیراً استفاده از یادگیری عمیق برای پیش‌بینی بازار سهام کشورهای مختلف به طور گسترده مورد توجه بوده است. بررسی نتایج مطالعات گذشته نشان می‌دهد که پیش‌بینی سری‌های زمانی مالی با استفاده از الگوریتم‌های یادگیری ماشین و آموزش مدل با داده‌های قیمت گذشته سهام با دقت قابل قبولی انجام می‌شود [۳] تا [۶].

این مقاله در تاریخ ۲۷ اسفند ماه ۱۴۰۰ دریافت و در تاریخ ۱۸ شهریور ماه ۱۴۰۱ بازنگری شد.

علیرضا جعفری، دانشکده علوم و فنون نوین، دانشگاه تهران، تهران، ایران، (email: alireza.jafari7@ut.ac.ir).

سامان هراتی‌زاده (نویسنده مسئول)، عضو هیئت علمی دانشکده علوم و فنون نوین، دانشگاه تهران، تهران، ایران، (email: haratizadeh@ut.ac.ir).

با توسعه بازارهای مالی، شرکت‌ها به طور گسترده‌ای از طریق روابط مختلف با یکدیگر در ارتباط هستند [۴]. به طور مشخص، افزایش فروش و درآمد یک شرکت در کاهش قیمت سهام شرکت رقیب مؤثر است. یک رویکرد بدیع برای مدل‌سازی روابط میان مجموعه‌ای از موجودیت‌ها در یادگیری عمیق، استفاده از مدل‌سازی گرافی است که می‌توان آن را برای استخراج الگوهای سطح بالا در بین موجودیت‌ها تجزیه و تحلیل کرد.

در سال‌های اخیر با توسعه روش‌های مدل‌سازی شبکه‌ای داده‌های سهام، محققان از این رویکردها در پیش‌بینی بازارهای سهام استفاده کرده‌اند و رویکردهای مختلف شبکه‌ای در چندین آزمایش مختلف، توانایی خود را در پیش‌بینی رفتار آینده بازارهای مالی نشان داده‌اند [۷] تا [۱۰]. جعفری و هراتی‌زاده (۱۳۹۹) و کیا و همکاران (۲۰۱۸) شبکه‌هایی را ساختند که به نظر می‌رسید در مدل‌سازی روابط پیچیده بین حرکت سری‌های زمانی مالی مفید باشند [۸] و [۱۰]. اگرچه مدل‌های آنها نسبت به مدل‌هایی که ارتباط بین سهام‌ها را در نظر نمی‌گرفت بهتر عمل کرده است، با این حال محدودیت‌های قابل توجهی دارند. اکثر کارهای گذشته از روال ساده‌ای برای پیش‌بینی استفاده کرده‌اند و بیشتر آنها از روش‌های انتشار برچسب مبتنی بر گراف برای پیش‌بینی جهت حرکت سهام استفاده می‌کنند. همچنین شیوه آنها برای مدل‌سازی روابط بین سهام غالباً محدود به استفاده از همبستگی بین سهام‌ها در یک شبکه بدون جهت بوده و روابط مفید دیگر مغفول مانده است. در حالی که گراف‌های جهت‌دار می‌توانند روابط دقیق‌تری را مدل‌سازی کنند و این دسته از مدل‌ها به شدت وابسته به گراف زیربنایی هستند.

در چند سال اخیر با توسعه داده‌ها با ساختار شبکه‌ای و با تأکید بر توانایی‌های شبکه‌های عصبی مصنوعی، دسته دیگری از این مدل‌ها با نام شبکه عصبی گراف (GNN) معرفی شده است [۱۱]. این دسته از متدها سعی دارند بازنمایی‌هایی مفید را برای امر پیش‌بینی تولید کنند؛ در حالی که ساختار گراف زیربنایی را در نظر می‌گیرند و می‌توانند ارتباطات بین سهام را از ساختار شبکه بیاموزند. با این حال استفاده از این مدل‌ها برای پیش‌بینی سهام به علت نوظهوربودن این دسته از مدل‌ها مورد توجه نبوده و نیازمند مطالعه بیشتر است. اهمیت ساخت بازنمایی‌های سطح بالا از ویژگی‌های ابتدایی و سطح پایین برای پیش‌بینی سهام با استفاده از شبکه‌های عصبی مصنوعی در مطالعات گذشته اثبات شده است [۳]، [۱۲] و [۱۳]. با این حال، این مطالعات بر شبکه‌های عصبی با ورودی منظم و ترتیب‌دار مانند شبکه عصبی کانولوشن متمرکز هستند. همچنین اغلب این کارها متکی بر داده‌های خود سهام است و از درنظرگرفتن تأثیرات سهام بر یکدیگر در یک چارچوب قابل یادگیری غافل شده‌اند. درنظرگرفتن تأثیرات سهام‌ها بر یکدیگر به قدری حائز اهمیت است که حسین‌زاده و هراتی‌زاده (۲۰۱۹) و گندوز و همکاران (۲۰۱۸) در رویکردهایی ساده سعی

کردند تا مجموعه بزرگی از سری‌هایی زمانی مالی را که مرتبط می‌دانستند به عنوان ورودی برای مدل شبکه عصبی کانولوشن خود در نظر بگیرند [۵] و [۱۴].

بازنمایی وضعیت گره با استفاده از شبکه عصبی گراف روشی است که به وسیله ادغام وضعیت گره‌ها با درنظرگرفتن ساختار گراف، نمایشی کارآمدتر را برای هر گره تولید می‌کند [۱۵]. ادغام و حفظ ارتباطات داده‌های سهام می‌تواند باعث ایجاد یادگیری بهتر در شبکه سهام شود [۳]. روش‌های شبکه عصبی گراف با یادگیری ارتباط بین سهام‌ها، یک چارچوب ذاتی برای تجزیه و تحلیل روابط بین شرکت‌ها در مجموعه داده را فراهم می‌کنند؛ در حالی که از توانایی‌های فوق‌العاده شبکه عصبی بهره می‌برند. گره‌ها در این چارچوب می‌توانند سری زمانی مالی (سهام) را همراه با مجموعه‌ای از ویژگی‌ها نشان دهند؛ در حالی که یال‌های گراف تأثیر داده‌های یک سهام برای بهبود دقت پیش‌بینی سهام دیگر را مدل می‌کنند. اطلاعات محلی در گراف باید در طول فرایند یادگیری در نظر گرفته شود و به همین منظور استفاده از شبکه توجه گراف برای تولید بازنمایی‌های گره‌ها توسط ما پیشنهاد شده است [۱۶]. شبکه توجه گراف با استفاده از مکانیزم توجه در خلال روند یادگیری، اهمیت ادغام هر گره همسایه را می‌آموزد.

ما در این مقاله، مدل DeepNet را معرفی می‌کنیم که برای پیش‌بینی جهت حرکت سهام برای یک روز آینده به عنوان یک مسئله کلاسیک دسته‌بندی استفاده شده است. مدل DeepNet سعی در پیشبرد یادگیری عمیق برای داده‌های سهام ساختاریافته به صورت گراف دارد که به طور کلی از شبکه توجه گراف برای یادگیری اهمیت وضعیت گره‌های همسایه برای ساخت بازنمایی مفیدتر برای پیش‌بینی استفاده می‌کند. الگوریتم شبکه توجه گراف برای گراف‌های بدون جهت طراحی شده و در محاسبه توجه برای یال‌های یک گره، ارتباط دوطرفه را در نظر می‌گیرد که برای پیش‌بینی سهام می‌تواند نویز شدیدی را به مدل وارد کند. اثرگذاری سهام الف برای پیش‌بینی سهام ب لزوماً به این معنا نیست که سهام ب نیز برای پیش‌بینی سهام الف مفید است. مدل DeepNet، شبکه توجه گراف جهت‌دار یا به اختصار DGAT را معرفی می‌کند که توسعه‌ای از الگوریتم شبکه توجه گراف است که می‌تواند روابط جهت‌دار را در خلال عملکرد آموزش بپذیرد و برخلاف شبکه توجه معمولی از مکانیزم توجه یک‌طرفه استفاده می‌کند.

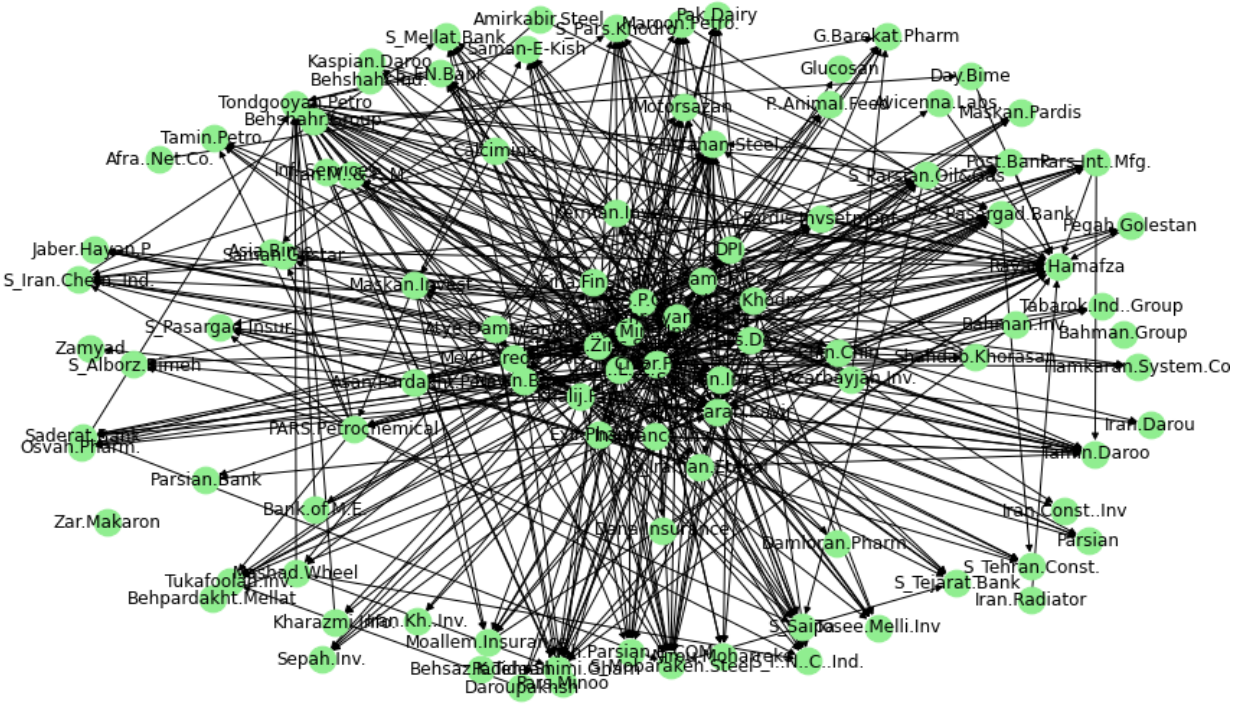
در بخش ۲ کارهای مرتبط و مدل‌های اصلی رقیب و در بخش ۳ ساختار شبکه و الگوریتم پیش‌بینی با جزئیات توضیح داده شده است. بخش ۴ مدل‌های ما را ارزیابی، نتایج پیش‌بینی را با سایر مدل‌های پایه، مقایسه و درباره نتایج بحث و آنها را تفسیر می‌کند. در پایان، بخش ۵ به نتیجه‌گیری و پیشنهاد برخی از مسیرها برای تحقیقات آینده در پیش‌بینی بازار سهام مبتنی بر مدل‌سازی گرافی می‌پردازد.

۲- ادبیات پایه و مطالعات پیشین

در پیش‌بینی بازار سهام با مدل‌های یادگیری عمیق گرافی از روش‌های مختلف یادگیری استفاده می‌شود که این روش‌ها را می‌توان در دامنه پیش‌بینی سهام به دو گروه تقسیم کرد: ۱) انتشار برچسب توسط شبکه و ۲) شبکه عصبی گراف.

در دسته اول از برخی برچسب‌ها و ساختار شبکه برای پیش‌بینی بازار سهام استفاده می‌شود. الگوریتم‌های انتشار برچسب و گسترش برچسب در آثار ژو (۲۰۰۵) توصیف شده‌اند [۱۷] و برای پیش‌بینی مالی در کارهای گذشته مورد استفاده بوده‌اند [۷] تا [۹]. این الگوریتم‌ها سعی دارند که برچسب‌های شناخته‌شده گره‌ها (جهت حرکت سهام) را در یک گراف به گره‌هایی با برچسب ناشناخته انتشار دهند. شین و همکاران (۲۰۱۳) شبکه‌ای بدون جهت از سهام‌ها را به عنوان گره‌ها که دارای برداری از برخی شاخص‌های فنی است در نظر گرفتند و یال‌ها را از فاصله اقلیدسی بین بردارهای گره‌ها ایجاد کردند. آنها از الگوریتم انتشار برچسب برای پیش‌بینی استفاده نمودند [۱۸]. کیا و همکاران (۲۰۱۸) سعی کردند که از سری زمانی بازارها یک شبکه همبستگی ایجاد کند و سپس از الگوریتم گسترش برچسب برای پیش‌بینی جهت بازارهای ناشناخته استفاده کردند [۱۰]. جعفری و هراتی‌زاده (۱۳۹۹) شبکه‌ای با استفاده از قانون اطمینان در مجموعه روابط انجمنی، ایجاد کردند و سعی داشتند به صورت دقیق‌تر از روابط بین سهام بهره‌برداری کنند. آنها برای هر سهام ۲ گره در شبکه در نظر گرفتند و سهام با وضعیت صعودی را از سهام با وضعیت نزولی مجزا کردند و سرانجام به وسیله انتشار برچسب به کمک الگوریتم پیچ رنگ با محاسبه اختلاف پیچ رنگ گره صعودی هر سهام از گره نزولی، اقدام به پیش‌بینی جهت حرکت قیمت سهام کردند [۸]. کیا و همکاران (۲۰۲۰) از الگوریتم گسترش برچسب و ساختار گراف جهت‌دار در یک شبکه مبتنی بر همبستگی استفاده کردند و عملکرد پیش‌بینی قابل توجهی را نسبت به مدل‌های رقیب به دست آوردند [۴]. با وجود این، کارهای گذشته همچنان از محدودیت‌های عمده‌ای رنج می‌برند. در این کارها یا برچسب‌ها با روش‌های غیر قابل اتکا انتخاب می‌شوند و یا از اختلاف زمان و فاصله جغرافیایی کشورها برای دسترسی به بخش کوچکی از برچسب‌ها استفاده می‌گردد و با کمک آن، سهام سایر نقاط جهان پیش‌بینی می‌شود و مدل‌های آنها توانایی پیش‌بینی سری‌های زمانی در داخل یک بازار را ندارند. همچنین تجزیه و تحلیل تأثیر سهام‌ها بر یکدیگر در داخل یک بازار سهام بسیار غنی‌تر و مفیدتر از تأثیر سهام سایر کشورها بر یکدیگر است [۱۹]. در روش معرفی‌شده DeepNet، علاوه بر توانایی پیش‌بینی روزانه تمام سهام یک بازار، روشی طراحی گردیده که بخش کوچکی از گره‌های شبکه را برچسب‌گذاری می‌کند که این امر مشکلات کارهای قبلی را برطرف کرده است.

دسته دیگر از مدل‌های یادگیری عمیق در حوزه گراف، شبکه‌های عصبی گراف است که برای استفاده از داده‌ها با ساختار گراف طراحی شده است. GNN می‌تواند از ساختار شبکه تعاملات بازار سهام در طول عملیات پس‌انتشار استفاده کند و در مقایسه با مدل‌هایی که صرفاً از قیمت‌های تاریخی سهام استفاده می‌کنند برتری دارد [۱۱]. با این حال، استفاده از این مدل‌ها برای پیش‌بینی سهام به دلیل نوظهوربودن این مدل‌های عمیق، به اندازه کافی مورد مطالعه قرار نگرفته و به چند مطالعه محدود می‌شود [۳]، [۱۲]، [۱۳] و [۲۰]. در اولین آثار، ماتسوناگا و همکاران (۲۰۱۹) بازار سهام ژاپن را با استفاده از ساختارهای ساده شبکه عصبی گراف پیش‌بینی کردند و هدف آنها نشان‌دادن سودمندی ترکیب داده‌های واحدهای مالی مختلف برای پیش‌بینی سهام بر اساس گراف در این دسته از مدل‌ها بود [۲۰]. چن و همکاران (۲۰۱۸) شبکه‌ای از شرکت‌ها را بر اساس اطلاعات سرمایه‌گذاری مالی ایجاد کردند. آنها با استفاده از یک گراف ساده بدون جهت، یک مدل شبکه عصبی کانولوشن گراف را آموزش دادند و عملکرد پیش‌بینی آن را با مدل‌های تعبیه‌سازی گره سنتی مقایسه کردند [۱۲]. کیم و همکاران (۲۰۱۹) از یک شبکه توجه سلسله‌مراتبی برای پیش‌بینی سهام استفاده کردند و همچنین سعی نمودند تا شاخص کل سهام را به عنوان یک کار طبقه‌بندی گراف توسط GRU پیش‌بینی کنند. آنها برای شبکه زیربنایی از گراف کلاسیک بدون جهت ویکی داده که در کار قبلی استفاده شده بود، بهره بردند [۱۳].



شکل ۱: گراف جهت‌دار ساخته‌شده با ε برابر ۰/۱ از مجموعه داده مورد استفاده در این مقاله.

جعفری و هراتی‌زاده (۲۰۲۲) از یک ساختار شبکه کانولوشن گراف برای پیش‌بینی سهام‌های بازار بورس نیویورک استفاده کردند. گراف زیربنایی آنها با توجه به محدودیت مدل شبکه کانولوشن گراف برای اتخاذ گراف جهت‌دار، شبکه‌ای بدون جهت از تأثیرات دوطرفه در بازار سهام بود که باعث ایجاد نویز در برخی بازنمایی‌های ساخته‌شده می‌شد [۳]. اگرچه این مدل‌ها به عنوان اولین مدل‌های شبکه عصبی گراف برای پیش‌بینی بازار سهام به نتایج قابل قبولی دست یافته‌اند، اما همچنان دارای محدودیت‌های قابل توجهی هستند. این مدل‌ها به علت محدودیت شبکه کانولوشن و توجه گراف در اعمال به گراف‌های جهت‌دار، ارتباطات بین سهام را دوطرفه در نظر می‌گیرند و مدل خود را به یک چارچوب ساده محدود می‌کنند. همچنین نمی‌توانند از پتانسیل شبکه توجه گراف استفاده کنند زیرا شبکه توجه گراف ذاتاً به عنوان یک مدل نیمه نظارت‌شده طراحی گردیده که در اغلب مطالعات فوق نادیده گرفته شده است. علاوه بر این، کیفیت تولید بازنمایی‌ها در همه انواع شبکه‌های عصبی گراف به شدت به گراف زیربنایی بستگی دارد که در کارهای اولیه مورد توجه قرار نگرفته و از ساده‌ترین گراف‌ها استفاده شده است.

۳- روش پیشنهادی

در این بخش مدل معرفی‌شده توسط ما به نام DeepNet به طور کامل شرح داده خواهد شد. DeepNet در ابتدا با استفاده از الگوریتم تجزیه و تحلیل تفکیک خطی (LDA)، گراف جهت‌داری را از سهام بازار ایجاد می‌کند. هر یال در گراف، تأثیر سهام در گره مبدأ یال را در بهبود عملکرد پیش‌بینی سهام در مقصد یال نشان می‌دهد و هر گره در گراف دارای یک بردار وضعیت اولیه از سیگنال اندیکاتورهای فنی سهام است. ساخت گراف، بخش مهمی از الگوریتم آموزش ما است چون اساساً برای اعمال شبکه توجه گراف ایجاد می‌شود. در ادامه، بعد از هرس‌کردن یال‌ها با وزن کم، گراف با استفاده از مجموعه‌ای از الگوریتم‌های نظارت‌شده برچسب‌گذاری می‌شود. نهایتاً پیش‌بینی نهایی برای گره‌ها با استفاده از مدل معرفی‌شده جدیدی با نام شبکه توجه گراف جهت‌دار به دست می‌آید. شبکه توجه گراف جهت‌دار یک چارچوب نیمه نظارت‌شده است که با استفاده از گراف جهت‌دار ساخته‌شده از سهام‌ها، وضعیت اولیه گره‌ها و برچسب‌های اختصاص داده شده به برخی سهام‌ها آموزش داده می‌شود و نمایش جدید و سطح بالایی از وضعیت هر گره می‌سازد که با استفاده از آنها قادر است تا کل بازار سهام را برای روز بعد پیش‌بینی کند. شکل ۱ روش کلی مدل پیشنهادی را نشان می‌دهد. در ادامه، ابتدا روش ساخت شبکه و سپس شبکه توجه گراف جهت‌دار با جزئیات توضیح داده خواهد شد و در آخر، رویه آموزش و پیش‌بینی DeepNet شرح داده می‌شود.

۳-۱ ساخت شبکه

در این قسمت، جزئیات ساخت شبکه توضیح داده شده است. ما شبکه‌ای را به منظور پیش‌بینی جهت حرکت قیمت سهام ایجاد می‌کنیم. این شبکه با هدف مدل‌سازی تأثیر ادغام داده‌های سهام‌های مختلف بر یکدیگر ایجاد می‌شود تا از مفیدبودن بازنمایی‌های ساخته‌شده در مرحله بعد اطمینان حاصل کند. این گراف $G=(V,E)$ نامیده می‌شود که در آن V مجموعه گره‌ها و E مجموعه یال‌ها است. فرض می‌کنیم که n سهام در مجموعه داده‌ها وجود دارد: $Stocks=\{s_1,s_2,\ldots,s_n\}$. برای هر سهام s_i، یک گره در گراف G قرار خواهیم داد و بنابراین مجموعه گره‌ها برابر است با $V=\{v_1,v_2,\ldots,v_n\}$.

برای ساخت گراف G از الگوریتم تجزیه و تحلیل تفکیک خطی (LDA) استفاده می‌شود. برای هر سهام s_i در مجموعه داده‌ها، ابتدا یک مدل LDA با استفاده از ویژگی‌های استخراج‌شده از داده‌های خود سهام، آموزش داده می‌شود و دقت پیش‌بینی آن در یک بازه ۱۰۰روزه قبل از روز آزمون محاسبه می‌گردد. سپس ویژگی‌های سهام s_i با هر یک از دیگر سهام‌های موجود در شبکه مانند s_j میانگین گرفته می‌شود تا بازنمایی ساده‌ای از داده‌های هر دو سهام ایجاد شود. مدل LDA دیگری با استفاده از ویژگی‌های تولیدشده جدید در بازه آموزش قبلی آموزش داده می‌شود و دقت پیش‌بینی آن در بازه ۱۰۰روزه محاسبه می‌گردد. اگر دقت پیش‌بینی مدلی که دو سهام را برای پیش‌بینی در نظر گرفته بود بیشتر از

حالت اول شد، این نشان‌دهنده تأثیر سهام s_j در بهبود عملکرد پیش‌بینی s_i است و یک یال از s_j به s_i در گراف قرار خواهد گرفت

$$Impact_{s_i,s_j} = Score_{LDA}(data^{s_i,s_j}) - Score_{LDA}(data^{s_i})$$
$$Impact_{s_i,s_j} \neq Impact_{s_j,s_i}$$
$$Set\ of\ vertices = \{v_1, v_2, \ldots, v_n\} \quad (1)$$
$$Set\ of\ edges = \begin{cases} W_{i,j} & \text{if } Impact_{s_i,s_j} > 0 \\ 0 & \text{otherwise} \end{cases}$$
$$i, j \in Set\ of\ vertices$$

بعد از محاسبه تأثیر تمامی سهام‌ها بر یکدیگر برای نادیده‌گرفتن یال‌ها با وزن کم و ازبین‌بردن نویز، از روش هرس ε نزدیک‌ترین همسایگی استفاده می‌کنیم که این روش با روش هرس k نزدیک‌ترین همسایه شباهت زیادی دارد. در هر دو الگوریتم با تعیین یک پارامتر، تعدادی از یال‌ها با وزن کم برای هر گره از گراف حذف می‌شوند، با این تفاوت که در روش ما یال‌های با وزن کمتر از ε در هر همسایگی از شبکه خارج می‌شوند؛ در حالی که k نزدیک‌ترین همسایه، k یال با بیشترین وزن را برای هر گره نگه می‌دارد که باعث ایجاد نویز شدید در گراف می‌شود؛ چرا که ممکن است یال‌ها با وزن پایین در شبکه باقی بمانند و یال‌هایی با وزن بالا برای گره‌ها هرس شوند. همچنین این کار، میانگین درجه تمامی گره‌ها را برابر می‌کند و هاب‌ها را از بین می‌برد و ساختار گراف اصلی را محدود می‌نماید. روش هرس ما باعث می‌شود که توانایی‌های شبکه توجه گراف جهت‌دار هرچه بیشتر نمایان شود. گراف ساخته‌شده دارای دو ویژگی مهم است. اولاً شبکه یک گراف جهت‌دار است و روابط برخلاف مطالعات گذشته بسیار دقیق‌تر مدل‌سازی شده و می‌توان با استفاده از n لایه توجه جهت‌دار، اطلاعات هر گره تا عمق n را در ساخت بازنمایی گره‌ها در نظر بگیرد. ثانیاً ساخت گراف از یال‌ها با بیشترین وزن به معنی اهمیت قابل توجه سهام در مبدأ یال در بهبود دقت پیش‌بینی سهام در مقصد یال است.

۳-۲ شبکه توجه گراف جهت‌دار (DGAT)

در این بخش قصد داریم تا مدل نوآورانه معرفی‌شده توسط خود را با نام شبکه توجه گراف جهت‌دار توضیح دهیم. شبکه توجه گراف جهت‌دار، توسعه‌ای از الگوریتم شبکه توجه گراف است. شبکه توجه گراف یا به اختصار GAT چارچوبی مؤثر برای یادگیری بازنمایی گره‌های گراف است و همچنین زیرمجموعه‌ای از شبکه‌های عصبی گراف به حساب می‌آید. GAT از مکانیزم توجه در حین عملکرد ساخت بازنمایی‌ها استفاده می‌کند. مکانیزم توجه، قابلیت آن را دارد که بر روی مهم‌ترین بخش‌های ورودی متمرکز شود و برخلاف شبکه کانولوشن گراف که در ساخت بازنمایی تنها به ساختار گراف متکی است، به گره‌های مختلف در گراف با مقادیر خاصی توجه می‌کند. با این حال، مدل GAT منحصراً برای گراف‌های بدون جهت طراحی شده و بنابراین GAT اهمیت ورودی‌ها (بردار وضعیت گره‌ها) را در یک رابطه دوطرفه بر روی یال‌ها محاسبه می‌کند که این می‌تواند توانایی آن را در گراف‌های جهت‌دار به طور چشم‌گیری محدود نماید.

گراف‌های جهت‌دار، روابط دقیق‌تری را بین گره‌های گراف مدل‌سازی می‌کنند و می‌توانند روابط بین سهام‌ها را با جزئیات با اهمیت بیشتری در نظر بگیرند. ما در این کار، شبکه توجه گراف جهت‌دار یا به اختصار DGAT را معرفی می‌کنیم که در محاسبه نمره توجه گره‌ها و ساخت بازنمایی‌ها متفاوت از GAT عمل می‌کند. مدل DGAT برخلاف GAT مکانیزم توجه را یک‌طرفه بر روی یال‌های گراف اعمال می‌کند و تأثیر بردار ویژگی گره در مبدأ یال را در ساخت بازنمایی برای گره در مقصد یال، محاسبه می‌کند.

ورودی مدل DGAT شامل دو جنبه است: ماتریس ویژگی‌های اولیه H و ماتریس مجاورت باینری گراف جهت‌دار A. یک لایه DGAT می‌تواند اطلاعات هر گره را به همسایگانی که به آنها یال خروجی دارد توزیع کند و اطلاعات سهامی را که به آن یال ورودی دارد در اختیار داشته باشد؛ در حالی که با استفاده از مکانیزم توجه بهترین تأثیر سهام همسایه را کشف می‌کند. با ادغام اطلاعات همسایگان، هر لایه DGAT نمایشی به‌روزشده برای هر گره ایجاد می‌کند.

لایه l از شبکه توجه گراف جهت‌دار به عنوان گام اولیه برای به دست آوردن قدرت کشف الگوهای پنهان برای تبدیل ویژگی‌های ورودی به ویژگی‌های سطح بالاتر، یک تبدیل خطی را با استفاده از ماتریس قابل یادگیری W برای تمامی گره‌ها انجام می‌دهد

$$z_i^{(l)} = W^{(l)} h_i \quad (2)$$

سپس یک نمره توجه غیر نرمال بین دو همسایه را محاسبه می‌کند. ابتدا بردارهای z دو گره به یکدیگر الحاق (||) می‌شوند (یا به صورت مشابه می‌توان میانگین دو بردار را در نظر گرفت). سپس حاصل‌ضرب نقطه‌ای آن با بردار وزن قابل یادگیری $a^{(l)}$ محاسبه می‌شود و در پایان، این لایه یک تبدیل غیر خطی LeakyReLU را اعمال می‌کند. معادله (۳) مهم‌ترین مرحله در الگوریتم است و این شکل از مکانیزم توجه به نام توجه افزایشی شناخته می‌شود

$$e_{ij}^{(l)} = \text{LeakyReLU}\left(a^{(l)^T}\left(z_i^{(l)} \middle\| z_j^{(l)}\right)\right) \quad (3)$$
$$\text{when is } j \to i$$

جایی که $e_{i,j}$ متناظر نمره توجه نرمال‌نشده یال بین گره i و j است. ساختار گراف با انجام توجه پوشانده‌شده[1] به مکانیزم تزریق می‌شود. به بیان دیگر، مقدار $e_{i,j}$ فقط برای گره‌ها با شرط $j \in N_i^{in}$ محاسبه می‌گردد، جایی که N_i^{in} گره‌هایی است که با یال ورودی به i متصل هستند. رابطه (۳) غیرنرمال است و نمی‌توان برای ساخت بازنمایی‌های جدید از آن استفاده کرد و به همین منظور، یک تابع $soft\max$ برای نرمال‌سازی وزن‌های ورودی به هر گره اعمال می‌شود

$$\alpha_{i,j} = soft\max(e_{i,j}) = \frac{\exp(e_{i,j})}{\sum_k \exp(e_{i,k})},\ k \in N_i^{in} \quad (4)$$

در رابطه فوق $\alpha_{i,j}$ برابر با توجه نرمال‌شده از بردار ویژگی گره j برای ساخت بازنمایی گره i است و نهایتاً معادله کلی به صورت رابطه زیر بازنویسی می‌شود

$$\alpha_{i,j} = \frac{\exp\left(\text{LeakyReLU}\left(a^{(l)^T}\left(z_i^{(l)} \middle\| z_j^{(l)}\right)\right)\right)}{\sum_k \exp\left(\text{LeakyReLU}\left(a^{(l)^T}\left(z_i^{(l)} \middle\| z_k^{(l)}\right)\right)\right)},\ k \in N_i^{in} \quad (5)$$

همچنین DGAT از توجه چندسر[2] برای غنی‌سازی ظرفیت مدل و تثبیت فرایند یادگیری استفاده می‌کند. به طور مشخص، K مکانیزم توجه مستقل، اجرا و با استفاده از (۶) خروجی لایه تولید می‌شود. باید توجه کرد که خروجی مکانیزم‌ها را می‌توان با هم الحاق کرد یا میانگین گرفت؛ ولی

1. Masked Attention
2. Multi-Head Attention

برای لایه آخر که پیش‌بینی نهایی مدل را تولید می‌کند ضروری است تا از میانگین استفاده شود

$$Average: h_i^{l+1} = \sigma(\sum_{k=1}^{K}\sum_{N_i^{in}} \alpha_{ij}^k W^K h_j^l)$$
$$Concatenation: h_i^{l+1} = \Big\|_{k=1}^{K} \sigma(\sum_{N_i^{in}} \alpha_{ij}^k W^K h_j^l) \quad (۶)$$

در شبکه توجه، گراف جهت‌دار مانند تمامی شبکه‌های عصبی گراف ماتریس همانی با ماتریس مجاورت گراف جمع می‌شود تا از مشارکت بردار ویژگی هر گره برای ساخت بازنمایی آن اطمینان حاصل گردد.

معماری یک مدل DGAT از مجموعه‌ای از لایه‌های توجه جهت‌دار تشکیل می‌شود که هر یک از آنها توابع فعال‌سازی Swich را برای افزایش توانایی غیر خطی استفاده می‌کنند. ما در DeepNet از ۲ لایه توجه جهت‌دار برای پیش‌بینی جهت حرکت قیمت سهام‌ها استفاده می‌کنیم. مدل DGAT مانند اکثر شبکه‌های عصبی گراف به صورت نیمه نظارت‌شده عمل می‌کند و خطای محاسبه‌شده بر اساس گره‌های برچسب‌دار گراف را برای به‌روزرسانی وزن‌های قابل یادگیری منتشر می‌کند. ما از تمام سهام برچسب‌دار موجود در گراف و تلفات آنتروپی متقابل به عنوان تابع خطا برای آموزش مدل استفاده می‌کنیم

$$Loss = -\sum_{i \in Z_L}\sum_{f=1}^{F} Z_{i,f} \ln Y_{i,f} \quad (۷)$$

جایی که Z_L مجموعه سهام برچسب‌دار، F کلاس‌های جهت حرکت قیمت سهام و Y مقدار پیش‌بینی‌شده برای هر کلاس توسط مدل DGAT است.

۳-۳ رویه آموزش و پیش‌بینی

ابتدا در این بخش، روش برچسب‌گذاری گراف و آموزش مدل DGAT را توضیح می‌دهیم و در ادامه، رویه پیش‌بینی نهایی سهام‌ها را بیان خواهیم کرد.

۳-۳-۱ روش تخصیص برچسب‌ها

همان طور که ذکر گردید، DeepNet یک چارچوب نیمه نظارت‌شده است و برای آموزش DGAT به تعدادی برچسب قابل اعتماد از آینده برخی سهام‌ها نیاز دارد. مشابه مدل معرفی‌شده با نام PLD در مدل جعفری و هراتی‌زاده (۲۰۲۲)، ما مدلی را برای شناسایی برچسب‌های قابل اعتماد استفاده می‌کنیم که از الگوریتم‌های یادگیری نظارت‌شده برای پیش‌بینی برچسب‌های روز آینده استفاده می‌کند [۳]. الگوریتم‌های یادگیری تحت نظارت دارای قابلیت‌های منحصربه‌فردی در پیش‌بینی سری‌های زمانی مالی هستند. همچنین سری‌های زمانی مالی در دوره‌های زمانی مختلف، رفتارهای متفاوتی در جهت حرکت دارند و این باعث می‌شود که بهترین مدل برای پیش‌بینی سری‌های زمانی مالی در دوره‌های زمانی مختلف متفاوت باشد. در الگوریتم DeepNet با استفاده از مجموعه‌ای از الگوریتم‌های یادگیری نظارت‌شده مانند جنگل تصادفی، تجزیه و تحلیل تفکیک خطی، ماشین بردار پشتیبان، درخت تصمیم، پرسپترون چندلایه و ...، حرکت قیمت سهام در روز آینده پیش‌بینی می‌شود. این مدل‌ها آموزش داده می‌شوند تا بر اساس تاریخچه قیمت هر سهام و تعدادی از اندیکاتورهای تحلیل فنی، جهت حرکت قیمت سهام در روز بعد را پیش‌بینی کنند. با استفاده از امتیاز پیش‌بینی بر روی داده‌های اعتبارسنجی، دقیق‌ترین مدل برای هر سهام مشخص می‌شود و سپس فقط در صورتی که امتیاز بهترین مدل در پیش‌بینی داده‌های اعتبارسنجی بالاتر از آستانه‌ای مشخص‌شده باشد، جهت پیش‌بینی‌شده سهام مورد نظر برای آموزش DGAT استفاده می‌گردد. با روش فوق هر سری زمانی به ازای هر روز با مدل متفاوتی پیش‌بینی می‌شود که آن مدل، بهترین مدل برای پیش‌بینی آن سری زمانی در روز مشخص‌شده است. مقدار آستانه از داده‌های اعتبارسنجی به دست می‌آید.

۳-۳-۲ رویه پیش‌بینی

الگوریتم DeepNet ابتدا شبکه‌ای جهت‌دار از تأثیرات سهام بازار را ایجاد و سپس با استفاده از روش PLD برخی گره‌ها را به عنوان کاندیدا برای برچسب‌گذاری انتخاب می‌نماید. این گره‌ها دقت قابل قبولی را در داده‌های اعتبارسنجی با استفاده از مجموعه‌ای از مدل‌های نظارت‌شده کسب کرده‌اند. در مرحله بعد، جهت حرکت پیش‌بینی‌شده برای گره‌های کاندیدا با استفاده از بهترین مدل نظارت‌شده برای هر سهام در شبکه قرار می‌گیرد و برچسب نهایی برای این گره‌ها در نظر گرفته می‌شود. همچنین DeepNet از این گره‌های برچسب‌دار برای آموزش شبکه توجه گراف جهت‌دار، استفاده و گره‌های باقیمانده در گراف را برچسب‌گذاری می‌کند. روش PLD با درنظرگرفتن آستانه امتیاز سخت‌گیرانه، درصد کمی از گره‌های گراف را برچسب‌گذاری می‌کند که این باعث می‌شود دقیق‌ترین برچسب‌ها در شبکه قرار گیرند و گره‌هایی که در مجموعه کاندیداها قرار نگرفته‌اند با استفاده از روش پیچیده‌تر و منابع اطلاعاتی بیشتر پیش‌بینی شود. آستانه امتیاز با تعداد گره‌های برچسب‌گذاری‌شده، رابطه معکوس و با دقت برچسب‌های این گره‌ها رابطه مستقیم دارد.

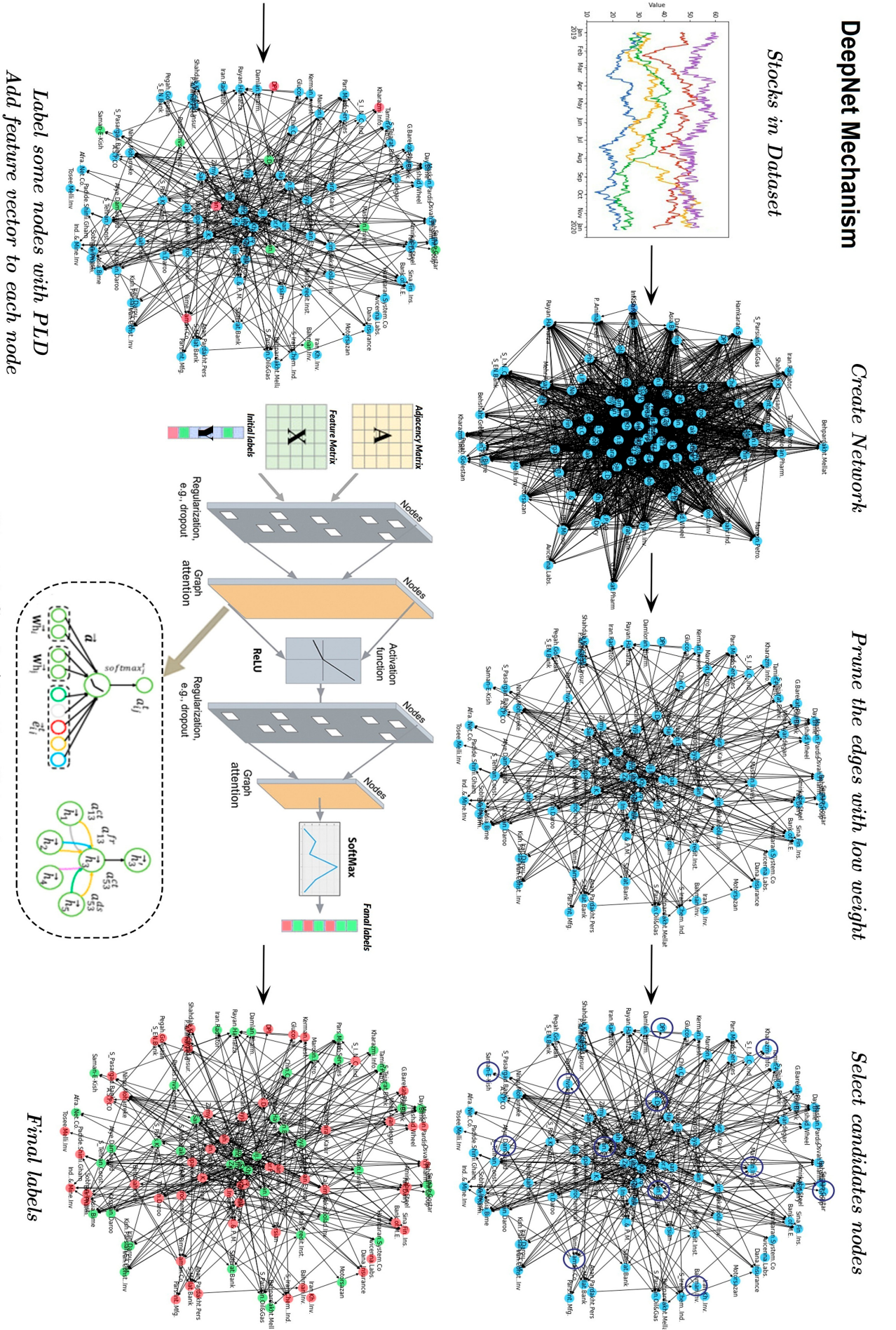
مدل DGAT برای پیش‌بینی جهت حرکت قیمت سهام، علاوه بر درنظرگرفتن وضعیت اولیه سهام‌ها، از ارتباط بین سهام در یک ساختار گرافی استفاده می‌کند و با داشتن دید مناسبی از آینده بازار سهام با استفاده از گره‌های برچسب‌دار، برچسب نهایی گره‌های باقیمانده را پیش‌بینی می‌نماید. معماری اجزای تشکیل دهنده‌ی مدل پیش‌بینی در شکل (۲) آمده است.

٤- ارزیابی

در این بخش، نتایج پیش‌بینی مدل پیشنهادی ارائه و ارزیابی شده است. ابتدا مجموعه داده را توصیف می‌کنیم و پس از آن، پیاده‌سازی و سپس روش ارزیابی آمده و نهایتاً نتایج مدل ما و مدل‌های پایه و بهترین مدل رقیب ارائه شده است.

تمام سری‌های زمانی سهام مورد استفاده در این تحقیق از سایت رسمی بورس تهران[1] جمع‌آوری گردیده است. داده‌های ما شامل قیمت بسته‌شدن، بازشدن، قیمت اولین و آخرین معامله و حجم معاملات روزانه سهام است. مجموعه داده ما متشکل از ۱۰ سهام برتر بر اساس اندازه سرمایه از ۱۰ صنعت بزرگ بازار سهام تهران یعنی رایانه و فعالیت‌های وابسته به آن، فلزات اساسی، انبوه‌سازی املاک و مستغلات، خودرو و ساخت قطعات، بیمه و صندوق بازنشستگی به جز تأمین اجتماعی، بانک‌ها و مؤسسات اعتباری، سرمایه‌گذاری‌ها، محصولات شیمیایی، مواد و محصولات دارویی و محصولات غذایی و آشامیدنی به‌جز قند و شکر جمع‌آوری شده که مجموعاً شامل ۱۰۰ سری زمانی می‌گردد. بازه زمانی مورد استفاده برای ساخت شبکه از ۲۰۱۹/۰۱/۰۱ تا ۲۰۲۱/۰۲/۲۰ است. ما برای آموزش مدل‌های نظارت‌شده در PLD از تمامی سابقه سهام برای آموزش مدل‌ها استفاده می‌کنیم و بازه اعتبارسنجی برای کشف سهام کاندیدا یک پنجره ۲۰روزه قبل از روز آزمون است. نهایتاً بازه آزمون ما

1. http://tsetmc.com



شکل ۲: معماری اجزای تشکیل‌دهنده مدل پیش‌بینی DeepNet.

جدول ۱: ارزیابی و نتایج.

Model	Accuracy (%)	MCC
Support Vector Machine	۵۳٫۸۴	۰٫۰۴۱
Random Forest	۵۸٫۲۴	۰٫۰۹۱
Linear Discriminant Analysis	۶۲٫۲۸	۰٫۱۲۱
[۲۱] ANN + PCA	۶۵٫۱۹	۰٫۱۶۵
[۴] DiMexRank	۶۵٫۹۸	۰٫۱۷۸
[۸] LabelNet	۶۶٫۳۲	۰٫۱۹۴
[۳] GCNET	۶۸٫۲۲	۰٫۲۰۲
DeepNet (this work)	۶۹٫۹۳	۰٫۲۶۷

شامل ۶۵ روز از بازار سهام از تاریخ ۲۰۲۱/۰۲/۲۰ تا ۲۰۲۱/۰۶/۲۱ است که شامل ۶۰۶۳ نمونه برای پیش‌بینی می‌شود. لینک دسترسی به مجموعه داده‌های[1] مورد استفاده در این تحقیق در پانویس آمده است.

٤-١ پیاده‌سازی و معماری DeepNet

امروزه، تعداد زیادی کتابخانه و پکیج یادگیری عمیق تولید شده است. ما در این کار از مجموعه Keras و TensorFlow برای پیاده‌سازی و اجرای مدل DeepNet و DGAT استفاده کرده‌ایم [۲۲]. برای معماری شبکه توجه گراف جهت‌دار با درنظرگرفتن تعداد ویژگی‌ها و نمونه‌های آموزشی، ما از ۸ کانال برای لایه اول و ۲ کانال برای پیش‌بینی ۲ کلاس در لایه دوم (آخر) استفاده کرده‌ایم. همچنین لایه اول از مکانیزم توجه چند سر، متشکل از ۳ مکانیزم توجه مجزا بهره می‌برد و در طول فرایند آموزش، بردارهای ویژگی با هم میانگین گرفته می‌شوند. تابع فعال‌سازی برای لایه اول Swish و لایه آخر Sigmoid است. از Adam با نرخ یادگیری ۰٫۰۰۱ برای آموزش DGAT استفاده شده [۲۳] و برای جلوگیری از بیش‌برازشی از حد مناسب L2 با نرخ ۰٫۰۰۱ و میزان چشم‌پوشی ۰٫۳ استفاده می‌گردد. همچنین برای دستیابی به کدهای این مقاله[2] آدرس اینترنتی در پانویس آورده شده است.

مدل DeepNet دارای دو ابرپارامتر است که باید با استفاده از داده‌های اعتبارسنجی محاسبه شوند: ۱) مقدار ε برای هرس‌کردن یال‌ها در ساخت شبکه و ۲) مقدار آستانه امتیاز در برچسب‌گذاری گره‌ها. آستانه امتیاز محاسبه‌شده بر روی داده‌های مورد استفاده کار ما در این مقاله برابر ۰٫۸۵ می‌باشد و مقدار ε برابر ۰٫۱ مقداردهی شده است.

٤-٢ معیارهای ارزیابی

به دنبال کارهای قبلی برای پیش‌بینی سهام [۳]، [۴] و [۸]، ما معیار استاندارد دقت و ضریب همبستگی متیوز (MCC) را به عنوان معیارهای ارزیابی اتخاذ می‌کنیم. معیار MCC به پیش‌بینی همه کلاس‌های موجود برای همه سهام موجود توجه می‌کند. با توجه به ماتریس سردرگمی[3] $\begin{pmatrix} tp & fn \\ fp & tn \end{pmatrix}$ شامل تعداد نمونه‌هایی که به عنوان مثبت واقعی، مثبت کاذب، منفی واقعی و منفی کاذب طبقه‌بندی شده‌اند، دقت و MCC با استفاده از معادلات زیر محاسبه می‌شوند

$$ACC = \frac{tp + tn}{tp + tn + fp + fn} \qquad (۸)$$

1. https://github.com/ut-kdd/DeepNet-Dataset
2. https://github.com/ut-kdd/DeepNet
3. Confusion Matrix

جدول ۲: تحلیل ساختار شبکه.

Method	Accuracy (%)
DeepNet with Correlation graph	۶۴٫۴۸
DeepNet with introduced graph	۶۹٫۹۳

$$MCC = \frac{tp \times tn - fp \times fn}{\sqrt{(tp + fp)(tp + fn)(tn + fp)(tn + fn)}} \qquad (۹)$$

٤-٣ نتایج

ما مدل خود را با مجموعه کاملی از مدل‌های رقیب مقایسه کرده‌ایم که شامل الگوریتم‌های نظارت‌شده معروف، مدل‌های مبتنی بر شبکه، مدل‌های شبکه عصبی عمیق مصنوعی و مدل‌های شبکه عصبی عمیق گراف است. نتایج پیش‌بینی بر روی داده‌های آزمون مورد استفاده در این مقاله در جدول ۱ آورده شده و همان طور که مشخص است، مدل DeepNet با اختلاف قابل توجهی نسبت به مدل‌های رقیب و مدل‌های نظارت‌شده، نتیجه بهتری را کسب کرده است. بعد از آن، مدل GCNET با دقت ۶۸٫۲۲ بهترین مدل در بین مدل‌های رقیب بوده که از ساختار گرافی و نوعی از شبکه عصبی گراف استفاده می‌کند. سپس به ترتیب مدل‌های LabelNet و DiMexRank بهترین نتایج را کسب کرده‌اند که هر دو مدلی مبتنی بر شبکه هستند. نتایج حاکی از آن است که مدل‌های شبکه ای که از اطلاعات سهام‌های مرتبط برای پیش‌بینی استفاده کرده‌اند، نسبت به مدل‌های شبکه عصبی مصنوعی که فقط داده‌های یک سهام را در نظر می‌گیرند، برتر بوده و دقت پیش بینی بالاتری کسب کرده‌اند و این امر نشان دهنده توانایی مدل‌های شبکه‌ای در پیش بینی بازار سهام است.

استفاده‌کردن از ساختار شبکه در مدل‌سازی بازار سهام همواره مورد توجه محققان بوده است؛ اما استفاده مستقیم از این شبکه برای پیش‌بینی سهام کمتر مورد توجه قرار گرفته است. شبکه‌ها دارای خصوصیات منحصربه‌فردی هستند که باید در پیش‌بینی سهام مورد توجه قرار گیرند. در این کار، استفاده از شبکه‌های عصبی گراف و ترکیب اطلاعات پنهان بین سهام با توجه به روابط و توپولوژی گراف ساخته‌شده، ایده اصلی ما بوده است. درنظرگرفتن اطلاعات سهام و وضعیت سهام مرتبط، باعث افزایش دقت پیش‌بینی گردیده است. شبکه توجه گراف علاوه بر قدرت شبکه‌های عصبی مصنوعی در کشف روابط پنهان بین ویژگی‌ها می‌تواند با استفاده از ساختار گراف و مکانیزم توجه روابط بین گره‌ها (سهام) را نیز کشف کند که این کار برای پیش‌بینی بازار سهام مفید است.

٤-٣-١ تحلیل ساختار شبکه

برای مشاهده نحوه عملکرد شبکه معرفی‌شده در مدل DeepNet، رویکرد دیگری برای مدل‌سازی روابط بین سهام را جایگزین شبکه معرفی‌شده در کار خود کردیم. همان طور که در بخش بررسی ادبیات توضیح دادیم، رایج‌ترین ساختار شبکه برای پیش‌بینی سهام شبکه همبستگی است. ما شبکه‌ای از سهام ایجاد کردیم که در آن یال‌ها نشان‌دهنده همبستگی بین قیمت‌های گذشته سهام است. سپس از همان تکنیک هرس اعمال‌شده برای شبکه معرفی‌شده استفاده کردیم تا یال‌ها با همبستگی قوی در شبکه باقی بمانند و شبکه هرس‌شده همچنان متصل باشد. ما دیگر جزئیات الگوریتم DeepNet را بدون تغییر نگه داشتیم و عملکرد آن را در پیش‌بینی حرکات قیمت سهام ارزیابی کردیم. نتایج در جدول ۲ خلاصه شده است.

جدول ۳: تحلیل شبکه توجه گراف جهت‌دار.

Method	Accuracy (%)
GCN	۶۷٫۵۸
GAT Undirected graph	۶۸٫۳۱
DeepNet (this work)	۶۹٫۹۳

٤-۳-۲ تحلیل شبکه توجه گراف جهت‌دار

برای مقایسه مدل DGAT معرفی‌شده با دیگر انواع شبکه عصبی گراف مانند شبکه کانولوشن گراف و همچنین شبکه توجه گراف با ورودی بدون جهت، ما GCN و GAT را در مدل خود قرار دادیم و یال‌های گراف اولیه خود را بدون جهت در نظر گرفتیم. همان طور که در جدول نتایج ۳ آمده است، مدل DeepNet با شبکه توجه گراف جهت‌دار توانسته با دقت پیش‌بینی در حدود ۲ درصد بالاتر از دو مدل دیگر در پیش‌بینی جهت حرکت قیمت سهام بهتر عمل کند که توانمندی مدل معرفی‌شده را نشان می‌دهد.

۵- نتیجه‌گیری و تحقیقات بیشتر

ما در این مقاله، یک چارچوب پیش‌بینی سهام با استفاده از ساختار شبکه نوآورانه و شبکه عصبی گراف را ارائه داده‌ایم و همچنین مدل جدیدی را به نام شبکه توجه گراف جهت‌دار معرفی کرده‌ایم. مدل ما به وسیله تعداد زیادی الگوریتم نظارت‌شده تصمیم می‌گیرد چه موقع از داده‌های تاریخی سهام برای استخراج برچسب‌های مورد نیاز استفاده کند؛ در حالی که بهترین مدل نظارت‌شده برای هر سهام را در هر دوره زمانی پیدا می‌کند. با این کار، مشکل نداشتن برچسب‌های قابل اعتماد برای روز بعد در بازار سهام را حل کرده‌ایم. ما توانایی مدل‌های مبتنی بر گراف نیمه نظارت‌شده را در پیش‌بینی نشان دادیم که می‌تواند مدل‌های نظارت‌شده و همچنین مدل‌های رقیب قدرتمند در ادبیات یادگیری عمیق را شکست دهد. نشان داده شد که در بیشتر دوره‌های زمانی بررسی‌گردیده، اطلاعات اخیر در مورد سهام‌های مرتبط برای پیش‌بینی یک سهام، بهتر از داده‌های گذشته خود سهام است و مدل پیشنهادی ما در این تحقیق، عملکرد بهتری در دقت پیش‌بینی و MCC نسبت به مدل‌های دیگر دارد.

پیشنهاد ما برای تحقیقات آینده، استفاده از شبکه‌هایی است که می‌توانند داده‌های گذشته سهام را در بیش از یک روز مدل‌سازی کنند. در تحقیقات اولیه ما متوجه شدیم که استفاده از ویژگی‌های ترازنشده به عنوان ورودی شبکه توجه گراف، نویز شدیدی را به مدل، وارد و عملکرد پیش‌بینی را تضعیف می‌کند. همچنین استفاده از بیش از دو یا سه لایه شبکه توجه گراف باعث می‌شود تا اطلاعات همه گره‌ها در بازنمایی وضعیت همه سهام منتشر شود و ویژگی محلی‌سازی و استفاده از اطلاعات سهام‌های مربوط به سهام مورد پیش‌بینی از بین رود. ما در این کار از مجموعه‌ای از الگوریتم‌های نظارت‌شده ساده برای استخراج برچسب‌های قابل اطمینان استفاده کردیم. بهبود عملکرد روش استخراج برچسب می‌تواند دقت کلی پیش‌بینی را افزایش دهد. در کارهای آینده می‌توان از مدل‌هایی پیچیده و دقیق‌تر برای استخراج برچسب استفاده کرد. همچنین گره‌های دارای برچسب در الگوریتم‌های مبتنی بر گراف مهم هستند و انتخاب گره‌ها برای برچسب‌گذاری می‌تواند یک مسیر تحقیقاتی جذاب باشد. افزودن سری‌های زمانی دیگر از منابع مختلف داده مانند قیمت نفت، طلا، شاخص صنایع و غیره نیز می‌تواند به پیش‌بینی مدل کمک کند. ما به عنوان ورودی برای شبکه توجه گراف از سیگنال‌های شاخص فنی استفاده کردیم که وضعیت سری‌های زمانی را فقط از نظر مالی نشان می‌دهد. استفاده از ویژگی‌های استخراج‌شده از داده‌های متنی یا شبکه‌های اجتماعی می‌تواند جالب توجه باشد. یادگیری نقاط چرخش به جای جهت حرکت و تغییر ساختار مدل به طبقه‌بندی ۳ کلاس (به عنوان مثال: بالا، پایین، ناشناخته) برای سری‌های زمانی مالی می‌تواند موضوع جالبی برای تحقیقات آینده باشد.

مراجع

[1] A. Timmermann and C. Granger, "Efficient market hypothesis and forecasting," *Int. J. Forecast*, vol. 20, no., pp. 15-27, 2004.

[2] J. Yao and C. L. Tan, "A case study on using neural networks to perform technical forecasting of forex," *Neurocomputing*, vol. 34, no. 1-4, pp. 79-98, Sept. 2000.

[3] A. Jafari and S. Haratizadeh, *GCNET: Graph-Based Prediction of Stock Price Movement Using Graph Convolutional Network*, arXiv preprint arXiv:2203.11091, 2022.

[4] A. N. Kia, S. Haratizadeh, and S. B. Shouraki, "Network-based direction of movement prediction in financial markets," *Engineering Applications of Artificial Intelligence*, vol. 88, Article ID: 103340, Feb. 2020.

[5] E. Hoseinzade and S. Haratizadeh, "CNNnpred: CNN-based stock market prediction using a diverse set of variables," *Expert Systems with Applications*, vol. 129, pp. 273-285, Sept. 2019.

[6] F. Jovanovic and C. Schinckus, *Econophysics and Financial Economics: An Emerging Dialogue*, Oxford Uni Press, 2017.

[7] K. Park and H. Shin, "Stock price prediction based on a complex interrelation network of economic factors," *Engineering Applications of Artificial Intelligence*, vol. 26, no. 5-6, pp. 1550-1561, May/Jun. 2013.

[۸] ع. ر. جعفری و س. هراتیزاده، "پیشبینی جهت حرکت قیمت سهام با استفاده از یک مدل انتشار برچسب مبتنی بر شبکه و یادگیری نظارتشده،" *مجموعه مقالات بیست و ششمین کنفرانس بین‌المللی کامپیوتر*، انجمن کامپیوتر ایران، ۶ صص.، تهران، ایران، ۱۳-۱۴ اسفند ۱۳۹۹.

[9] D. Shah, H. Isah, and F. Zulkernine, "Stock market analysis: a review and taxonomy of prediction techniques," *International J. of Financial Studies*, vol. 7, no. 2, Article ID: 70200026, 22 pp., 2019.

[10] A. N. Kia, S. Haratizadeh, and S. B. Shouraki, "A hybrid supervised semi-supervised graph-based model to predict one-day ahead movement of global stock markets and commodity prices," *Expert Systems with Applications*, vol. 105, pp. 159-173, Sept. 2018.

[11] F. Scarselli, M. Gori, A. C. Tsoi, M. Hagenbuchner, and G. Monfardini, "The graph neural network model," *IEEE Trans. on Neural Networks*, vol. 20, no. 1, pp. 61-80, Jan. 2009.

[12] J. Chen, T. Ma, and C. Xiao, *FastGCN: Fast Learning with Graph Convolutional Networks via Importance Sampling*, arXiv preprint arXiv: 1801.10247, 2018.

[13] R. Kim, C. H. So, M. Jeong, S. Lee, J. Kim, and J. Kang, *HATS: A Hierarchical Graph Attention Network for Stock Movement Prediction*, arXiv preprint arXiv: 1908.07999, 2019.

[14] H. Gunduz, Y. Yaslan, and Z. Cataltepe, "Intraday prediction of Borsa Istanbul using convolutional neural networks and feature correlations," *Knowledge-Based Systems*, vol. 137, pp. 138-148, Dec. 2017.

[15] T. N. Kipf and M. Welling, *Semi-Supervised Classification with Graph Convolutional Networks*, arXiv preprint arXiv: 1609.02907, 2016.

[16] P. Veličković, *et al*., *Graph Attention Networks*, arXiv preprint arXiv: 1710.10903, 2017.

[17] X. J. Zhu, *Semi-Supervised Learning Literature Survey*, Tech. Rep., University of Wisconsin-Madison Department of Computer Sciences, 2005.

[18] H. Shin, T. Hou, K. Park, C. K. Park, and S. Choi, "Prediction of movement direction in crude oil prices based on semi-supervised learning," *Decision Support Systems*, vol. 55, no. 1, pp. 348-358, Apr. 2013.

[19] R. N. Mantegna, "Hierarchical structure in financial markets," *The European Physical Journal B - Condensed Matter and Complex Systems*, vol. 11, pp. 193-197, 1999.

[20] D. Matsunaga, T. Suzumura, and T. Takahashi, *Exploring Graph Neural Networks for Stock Market Predictions with Rolling Window Analysis*, arXiv preprint arXiv: 1909.10660, 2019.

[21] X. Zhong and D. Enke, "Forecasting daily stock market return using dimensionality reduction," *Expert Systems with Applications*, vol. 67, pp. 126-139, Jan. 2017.

[22] F. Chollet, *et al.*, *Keras: The python Deep Learning Library, Astrophysics Source Code Library*, 2018.
[23] K. Da, *A Method for Stochastic Optimization*, arXiv preprint arXiv: 1412.6980, 2014.

علیرضا جعفری در سال ۱۳۹۸ مدرک کارشناسی مهندسی کامپیوتر خود را از دانشگاه آزاد اسلامی و در سال ۱۴۰۱ مدرک کارشناسی ارشد مهندسی فناوری اطلاعات خود را از دانشگاه تهران دریافت نمود و هم‌اکنون دانشجوی دکتری علوم کامپیوتر در دانشگاه ایالتی ویرجینیا در ایالات متحده آمریکا است. زمینه‌های تحقیقاتی مورد علاقه ایشان عبارتند از: یادگیری ماشین، علوم شبکه و تئوری گراف، شبکه‌های عصبی مبتنی بر گراف و پیش‌بینی بازارهای مالی.

سامان هراتی‌زاده تحصیلات خود را در مقاطع کارشناسی، کارشناسی ارشد و دکتری مهندسی کامپیوتر در دانشگاه صنعتی شریف به‌ترتیب در سال‌های ۱۳۷۸، ۱۳۸۰ و ۱۳۸۷ به پایان رسانده است. دکتر هراتی‌زاده از سال ۱۳۹۰ عضو هیأت علمی دانشکده علوم و فنون نوین دانشگاه تهران است و زمینه‌های پژوهشی مورد علاقه وی کاربردهای متنوعی از یادگیری ماشین و مدل‌سازی و تحلیل شبکه را شامل می‌شود.