

بِسْمِ اللَّهِ الرَّحْمَنِ الرَّحِيمِ



دانشگاه صنعتی اصفهان
دانشکده علوم ریاضی

تشخیص ناهنجاری در داده‌های چند متغیره به کمک یادگیری گراف مکانی-زمانی

پروژه کارشناسی ریاضیات و کاربردها

آیسان عزیزی

استاد راهنما

دکتر رضا مختاری

۱۴۰۵

کلیه حقوق مالکیت مادی و معنوی مربوط به این پروژه متعلق به دانشگاه صنعتی اصفهان و پدیدآورندگان است. این حقوق توسط دانشگاه صنعتی اصفهان و براساس خط مشی مالکیت فکری این دانشگاه، ارزش‌گذاری و سهم‌بندی خواهد شد. هر گونه بهره‌برداری از محتوا، نتایج یا اقدام برای تجاری‌سازی دستاوردهای این پروژه تنها با مجوز کتبی دانشگاه صنعتی اصفهان امکان‌پذیر است.

فهرست مطالب

چهار	فهرست مطالب
شش	چکیده
هفت	پیشگفتار
۱	۱ مقدمه
۴	۲ روش تشخیص ناهنجاری مبتنی بر یادگیری گراف مکانی-زمانی
۵	۱.۲ یادگیری هم‌بستگی میان متغیرها
۶	۱.۱.۲ انتخاب ارتباط‌های قوی‌تر
۷	۲.۲ گراف پیچش
۷	۳.۲ مدل‌سازی زمانی
۸	۴.۲ ترکیب اطلاعات زمانی و گرافی
۸	۵.۲ محاسبه خطای پیش‌بینی
۹	۱.۵.۲ نرمال‌سازی خطای پیش‌بینی
۹	۶.۲ امتیازدهی ناهنجاری با PCA
۱۰	۱.۶.۲ تعداد مؤلفه‌های اصلی
۱۰	۷.۲ محاسبه امتیاز نهایی ناهنجاری
۱۱	۱.۷.۲ آستانه‌گذاری
۱۲	۳ تعمیم چارچوب LTG به سری‌های زمانی مالی
۱۲	۱.۳ انگیزه تعمیم به حوزه مالی
۱۳	۱.۱.۳ بازتعریف مسئله مالی
۱۳	۲.۱.۳ تعریف گره‌ها و ویژگی‌ها

۱۴		تعریف ناهنجاری در بازار مالی	۳.۱.۳
۱۴		یادگیری ساختار وابستگی مالی	۲.۳
۱۵		مدل سازی زمانی در داده های مالی	۱.۲.۳
۱۵		نرمال سازی خطای پیش بینی در داده های مالی	۲.۲.۳
۱۶		امتیازدهی PCA برای ناهنجاری مالی	۳.۲.۳
۱۶		چالش تغییر رژیم و غیرایستایی	۳.۳
۱۶		جلوگیری از نشت اطلاعات	۱.۳.۳
۱۷		محدودیت های تعمیم مالی	۲.۳.۳
۱۷		محدودیت ها	۳.۳.۳
۱۹		نتیجه گیری	
۲۱		پیوست ۱	
۲۲		کتاب نامه	
۲۵		واژه نامه فارسی به انگلیسی	
۲۸		واژه نامه انگلیسی به فارسی	
۳۲		فهرست نمادها	
۳۴		فهرست کلمات اختصاری	

چکیده:

در این پروژه یک روش جدید مبتنی بر یادگیری گراف مکانی-زمانی برای تشخیص ناهنجاری در سری‌های زمانی چندمتغیره بررسی می‌شود. اساس این روش بر این ایده استوار است که رفتار هر متغیر علاوه بر الگوی زمانی خود، به روابط آن با سایر متغیرهای سامانه نیز وابسته است؛ از این رو، مدل ابتدا ارتباط میان متغیرها را مستقیماً از داده یاد می‌گیرد و آن‌ها را در قالب یک گراف جهت‌دار نمایش می‌دهد. سپس با ترکیب مدل‌سازی وابستگی‌های زمانی و انتقال اطلاعات میان متغیرهای مرتبط، مقدار گام زمانی بعد پیش‌بینی می‌شود. اختلاف میان مقدار واقعی و مقدار پیش‌بینی شده به عنوان خطای پیش‌بینی در نظر گرفته شده و پس از نرمال‌سازی، ساختار معمول این خطاها با استفاده از تحلیل مؤلفه‌های اصلی مدل می‌شود. در نهایت، میزان اختلاف الگوی خطای مشاهده‌شده با الگوی عادی به عنوان امتیاز ناهنجاری در نظر گرفته می‌شود. همچنین در این پروژه امکان تعمیم این چارچوب به سری‌های زمانی مالی بررسی می‌شود؛ به طوری که دارایی‌های مالی به عنوان گره‌های گراف و بازده یا سایر ویژگی‌های مالی به عنوان داده‌های زمانی در نظر گرفته می‌شوند. در این تعمیم، چالش‌هایی مانند غیرایستایی داده‌های مالی، تغییر رژیم، تغییر روابط میان دارایی‌ها و جلوگیری از نشت اطلاعات مورد توجه قرار می‌گیرند.

واژگان کلیدی:

سری زمانی چندمتغیره، تشخیص ناهنجاری، یادگیری گراف مکانی-زمانی، شبکه عصبی گراف، تحلیل مؤلفه‌های اصلی، سری زمانی مالی، تغییر رژیم.

پیشگفتار

تشخیص ناهنجاری در سری‌های زمانی چندمتغیره یکی از مسائل مهم در تحلیل داده‌های پیچیده است. در چنین داده‌هایی، رفتار هر متغیر تنها به مقادیر گذشته خود وابسته نیست، بلکه می‌تواند تحت تأثیر رفتار سایر متغیرهای سامانه نیز قرار گیرد. از این رو، بررسی هم‌زمان وابستگی‌های زمانی و روابط میان متغیرها می‌تواند نقش مهمی در شناسایی رفتارهای غیرعادی داشته باشد.

در این پروژه، یک روش جدید مبتنی بر یادگیری گراف مکانی-زمانی برای تشخیص ناهنجاری در سری‌های زمانی چندمتغیره بررسی می‌شود. اساس این روش بر یادگیری هم‌زمان ساختار ارتباطی میان متغیرها و الگوهای زمانی آن‌ها استوار است. در این چارچوب، هر متغیر به عنوان یک گره در گراف در نظر گرفته می‌شود و روابط میان گره‌ها مستقیماً از داده یاد گرفته می‌شوند.

برای مدل‌سازی وابستگی‌های زمانی، ابتدا از شبکه حافظه کوتاه‌مدت بلندمدت^۱ (LSTM) استفاده می‌شود. این شبکه با استفاده از سازوکار حافظه و دروازه‌های کنترلی، امکان حفظ و پردازش اطلاعات مربوط به وابستگی‌های زمانی در طول توالی را فراهم می‌کند.

در ادامه، شبکه پیچشی زمانی^۲ (TCN) برای استخراج الگوهای زمانی در مقیاس‌های مختلف به کار گرفته می‌شود. استفاده هم‌زمان از LSTM و TCN باعث می‌شود مدل بتواند هم وابستگی‌های ترتیبی و هم الگوهای زمانی محلی و چندمقیاسی را در نظر بگیرد.

برای مدل‌سازی روابط میان متغیرها، از شبکه پیچشی گرافی^۳ (GCN) استفاده می‌شود. در این بخش، اطلاعات میان گره‌هایی که بر اساس گراف آموخته شده با یکدیگر ارتباط دارند، منتقل و جمع می‌شود. به این ترتیب، GCN امکان استفاده از ساختار ارتباطی میان متغیرها را در کنار اطلاعات زمانی فراهم می‌کند.

پس از ترکیب ویژگی‌های استخراج شده توسط LSTM، TCN و GCN، مدل مقدار گام زمانی بعد را پیش‌بینی می‌کند. اختلاف میان مقدار واقعی و مقدار پیش‌بینی شده به عنوان خطای پیش‌بینی در نظر گرفته می‌شود. از آنجا که متغیرهای مختلف ممکن است دارای مقیاس و میزان نوسان متفاوتی باشند، خطاهای پیش‌بینی پیش از محاسبه امتیاز ناهنجاری نرمال‌سازی می‌شوند.

Long Short-Term
¹Memory

Temporal Convolutional
²Network

Graph Convolutional
³Network

در مرحله بعد، تحلیل مؤلفه‌های اصلی^۴ (PCA) برای مدل‌سازی ساختار معمول خطاهای پیش‌بینی به کار گرفته می‌شود. بردار خطای نرمال‌شده با استفاده از PCA بازسازی می‌شود و اختلاف میان بردار اصلی و بردار بازسازی‌شده، مبنای محاسبه امتیاز ناهنجاری قرار می‌گیرد. هرچه این اختلاف بیشتر باشد، رفتار مشاهده‌شده فاصله بیشتری از الگوی عادی آموخته‌شده توسط مدل دارد.

در ادامه این پروژه، امکان تعمیم این چارچوب به سری‌های زمانی مالی نیز بررسی می‌شود. در این صورت‌بندی، دارایی‌های مالی به‌عنوان گره‌های گراف و بازده یا سایر ویژگی‌های مالی به‌عنوان داده‌های زمانی در نظر گرفته می‌شوند. هدف از این تعمیم، بررسی امکان استفاده از ساختار مکانی-زمانی مدل برای شناسایی رفتارهای غیرمعمول در یک سامانه مالی چندمتغیره است. در این راستا، مسائلی مانند غیرایستایی داده‌های مالی، تغییر رژیم، تغییر روابط میان دارایی‌ها و جلوگیری از نشت اطلاعات نیز مورد توجه قرار می‌گیرند.

فصل ۱

مقدمه

در این فصل مفاهیم پایه مورد نیاز شامل سری زمانی چندمتغیره، تشخیص ناهنجاری مبتنی بر پیش‌بینی، شبکه LSTM، شبکه پیچشی زمانی TCN، شبکه‌های عصبی گراف و روش PCA مرور می‌شوند. مطالب این بخش از منابع [۱، ۲، ۷، ۹، ۱۳] گردآوری شده‌اند.

تعریف ۱-۱ (سری زمانی چندمتغیره). فرض کنید

$$S = \{s_1, s_2, \dots, s_Z\}, \quad s_z \in \mathbb{R}^H,$$

که در آن Z تعداد نقاط زمانی و H تعداد متغیرها یا کانال‌هاست. هر بردار s_z وضعیت هم‌زمان سامانه را در زمان z نشان می‌دهد. در تشخیص ناهنجاری، هدف تعریف تابع امتیازدهی $\mathcal{X}(\cdot)$ است، به‌گونه‌ای که امتیاز مشاهده ناهنجار از مشاهده عادی بزرگ‌تر باشد. در نهایت با مقایسه امتیاز با یک آستانه، برچسب دودویی ناهنجاری تعیین می‌شود.

تعریف ۲-۱ (تشخیص ناهنجاری مبتنی بر پیش‌بینی). یکی از رویکردهای رایج این است که مدل فقط روی داده‌های عادی آموزش ببیند و از پنجره‌ای از مشاهدات گذشته برای پیش‌بینی مشاهده بعدی استفاده کند. اگر مدل ساختار عادی سامانه را به‌خوبی آموخته باشد، در حالت عادی خطای پیش‌بینی کوچک خواهد بود؛ در حالی‌که تغییر غیرمنتظره در الگوهای زمانی یا روابط میان متغیرها می‌تواند خطای پیش‌بینی را افزایش دهد. این خطا مبنای ساخت امتیاز ناهنجاری قرار می‌گیرد.

تعریف ۳-۱ (شبکه حافظه کوتاه مدت بلند مدت). شبکه LSTM نوعی شبکه بازگشتی است که با استفاده از حالت سلولی و دروازه‌های ورودی، فراموشی و خروجی، جریان اطلاعات در توالی را کنترل می‌کند. این ساختار برای حفظ وابستگی‌های زمانی دورتر نسبت به یک شبکه بازگشتی ساده مناسب‌تر است. در مدل مورد بررسی، خروجی LSTM ورودی بخش پیچش زمانی می‌شود تا دو سازوکار متفاوت برای استخراج ویژگی‌های زمانی با یکدیگر ترکیب شوند. برای یک گام زمانی z ، روابط دروازه‌ها را به صورت زیر بیان می‌کند:

$$f_z = \sigma(M_f[h_{z-1}, s_z] + b_f),$$

$$i_z = \sigma(M_i[h_{z-1}, s_z] + b_i),$$

$$o_z = \sigma(M_o[h_{z-1}, s_z] + b_o),$$

$$\tilde{C}_z = \sigma(M_c[h_{z-1}, s_z] + b_c),$$

و سپس

$$C_z = f_z \odot C_{z-1} + i_z \odot \tilde{C}_z,$$

$$h_z = o_z \odot \tanh(C_z).$$

تعریف ۴-۱. شبکه TCN با استفاده از پیچش‌های زمانی و پیچش‌های اتساع یافته می‌تواند اطلاعات مربوط به چندین گام زمانی را هم‌زمان بررسی کند. با افزایش ضریب اتساع در لایه‌های مختلف، شبکه می‌تواند به اطلاعات مربوط به زمان‌های دورتر نیز دسترسی داشته باشد، بدون اینکه به تعداد زیادی لایه نیاز داشته باشد. همچنین این ساختار نسبت به شبکه‌های بازگشتی امکان پردازش موازی بهتری دارد و می‌توان محدوده زمانی مورد بررسی مدل را راحت‌تر کنترل کرد. در چارچوب LTG از یک سازوکار دروازه‌ای در TCN استفاده می‌شود:

$$\text{TCN}(T^l, \Theta_l) = f_C(T^l, \Theta_l^c) \odot f_G(T^l, \Theta_l^g),$$

که f_C نقش فیلتر و f_G نقش دروازه را دارد.

تعریف ۵-۱ (شبکه عصبی و پیچشی گراف). گراف $G = (V, E)$ شامل مجموعه گره‌ها و یال‌هاست. در مسئله تشخیص ناهنجاری‌ها، هر گره یک متغیر سری زمانی را نمایندگی می‌کند و ماتریس مجاورت شدت رابطه میان متغیرها را مشخص می‌سازد. شبکه‌های عصبی گراف با انتشار و تجمیع پیام از همسایگان، نمایش هر گره را به‌روزرسانی می‌کنند. یکی از چالش‌ها در انتشار چندلایه اطلاعات، پدیده بیش‌هموارشدگی است؛ یعنی نمایش گره‌ها با افزایش عمق شبیه یکدیگر می‌شود.

تذکر ۱-۶. روش LTG برای کاهش این اثر، بخشی از حالت اولیه را در هر مرحله نگه می‌دارد و خروجی عمق‌های مختلف را وزن‌دهی می‌کند. ایده پایه آن به روش‌های ساده‌سازی پیچش گراف و انتشار چندگامی نزدیک است.

تعریف ۱-۷. PCA یک تبدیل خطی برای نمایش داده در زیرفضایی با ابعاد کمتر است. همچنین نقش آن در تشخیص ناهنجاری‌ها تنها کاهش بعد داده خام نیست؛ بلکه روی بردار خطای نرمال‌شده اعمال می‌شود.

تذکر ۱-۸. فاصله میان خطای اصلی و خطای بازسازی‌شده از مؤلفه‌های اصلی به‌عنوان امتیاز ناهنجاری استفاده می‌شود. این ایده باعث می‌شود تغییرات رایج یا نویزهای مشترک کمتر بر امتیاز نهایی غالب شوند.

تذکر ۱-۹. برای تشخیص کلی از معیارهای ارزیابی ROC-AUC و PRC-AUC استفاده می‌کند. معیار ROC-AUC توان تفکیک بین نمونه‌های عادی و ناهنجار را در آستانه‌های مختلف نشان می‌دهد. در داده‌های نامتوازن، PRC-AUC نیز اهمیت زیادی دارد، زیرا به‌طور مستقیم رابطه دقت مثبت و فراخوان را ارزیابی می‌کند. همچنین برای تشخیص زودهنگام، امتیاز F1 تحت محدودیت‌های تأخیر مختلف گزارش شده است.

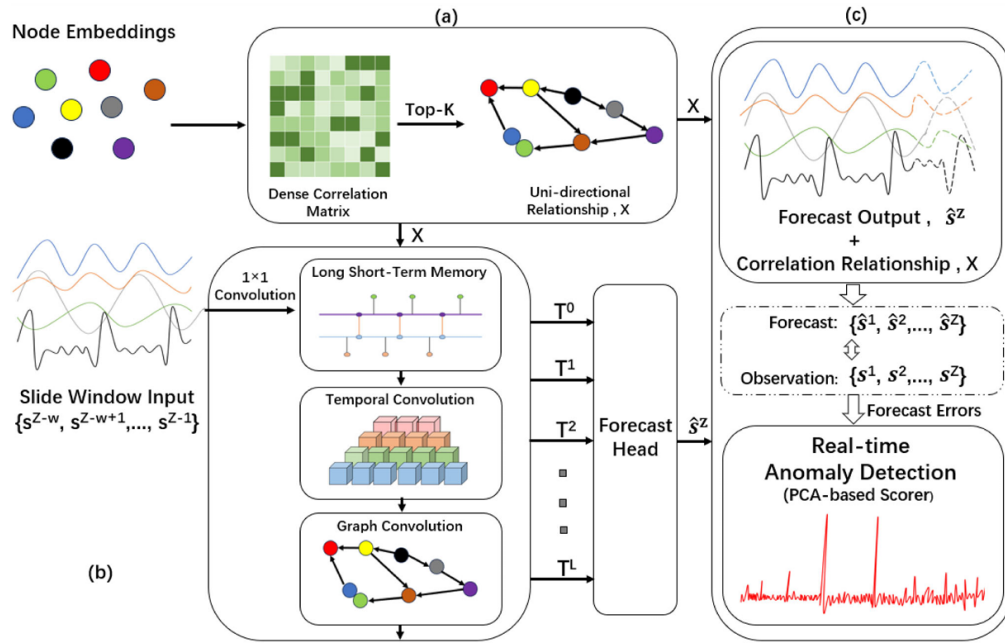
فصل ۲

روش تشخیص ناهنجاری مبتنی بر یادگیری گراف مکانی-زمانی

در این فصل، چارچوب LTG برای تشخیص ناهنجاری در سری‌های زمانی چندمتغیره بررسی می‌شود. ایده اصلی این روش آن است که برای تشخیص رفتار غیرعادی، تنها تاریخچه زمانی هر متغیر کافی نیست و روابط میان متغیرهای مختلف نیز باید در مدل در نظر گرفته شوند. چارچوب LTG ابتدا با استفاده از یک لایه یادگیری هم‌بستگی، روابط میان متغیرهای سری زمانی را یاد می‌گیرد و آن‌ها را در قالب یک گراف جهت‌دار نمایش می‌دهد. سپس شبکه‌های LSTM، TCN و GCN برای استخراج هم‌زمان وابستگی‌های زمانی و ساختاری استفاده می‌شوند. خروجی این بخش برای پیش‌بینی مقدار یک گام آینده به کار می‌رود. در مرحله تشخیص، اختلاف میان مقدار واقعی و مقدار پیش‌بینی‌شده محاسبه می‌شود. پس از نرمال‌سازی این خطا، روش تحلیل مؤلفه‌های اصلی PCA برای ساخت یک امتیاز ناهنجاری استفاده می‌شود. هرچه این امتیاز بزرگ‌تر باشد، احتمال وجود رفتار غیرعادی در داده بیشتر است. مدل LTG از سه بخش اصلی تشکیل شده است:

۱. یادگیری هم‌بستگی میان متغیرها و ساخت ماتریس مجاورت گراف؛
۲. استفاده از شبکه عصبی گراف مکانی-زمانی برای پیش‌بینی یک گام آینده؛
۳. تبدیل خطای پیش‌بینی به امتیاز ناهنجاری با استفاده از نرمال‌سازی و PCA.

ورودی مدل یک پنجره از مشاهدات گذشته سری زمانی است. هر متغیر یا کانال سری زمانی به‌عنوان یک گره در گراف در نظر گرفته می‌شود. گراف یادگرفته‌شده تعیین می‌کند که کدام متغیرها با یکدیگر ارتباط بیشتری دارند و اطلاعات میان کدام گره‌ها منتقل شود. شکل ۱.۲ ساختار کلی این چارچوب را نشان می‌دهد.



شکل ۱.۲: ساختار کلی چارچوب LTG.

۱.۲ یادگیری هم‌بستگی میان متغیرها

اولین بخش مدل LTG وظیفه یادگیری روابط میان متغیرهای مختلف سری زمانی را بر عهده دارد. در بسیاری از مسائل واقعی، ساختار ارتباطی واقعی بین متغیرها از قبل مشخص نیست. به همین دلیل، مدل باید این ساختار را به‌طور مستقیم از داده یاد بگیرد. برای H متغیر، دو ماتریس نمایش گره به‌صورت

$$H_1, H_2 \in \mathbb{R}^{H \times d}$$

در نظر گرفته می‌شوند که در آن d بعد نمایش هر گره است. سپس دو تبدیل قابل یادگیری روی این ماتریس‌ها اعمال می‌شود:

$$\tilde{H}_1 = \tanh(\alpha H_1 W_1),$$

$$\tilde{H}_2 = \tanh(\alpha H_2 W_2).$$

در ادامه، ماتریس ارتباط میان گره‌ها به صورت زیر ساخته می‌شود:

$$X = \text{ReLU} \left[\tanh \left(\alpha \left(\tilde{H}_1 \tilde{H}_2^T - \tilde{H}_2 \tilde{H}_1^T \right) \right) \right].$$

عبارت

$$\tilde{H}_1 \tilde{H}_2^T - \tilde{H}_2 \tilde{H}_1^T$$

باعث می‌شود روابط حاصل بتوانند جهت‌دار باشند. به بیان ساده، اگر متغیر i روی متغیر j ارتباط قوی داشته باشد، به طور لزوم رابطه معکوس با همان شدت وجود ندارد. پارامتر α نیز میزان اشباع تابع غیرخطی را کنترل می‌کند.

۱.۱.۲ انتخاب ارتباط‌های قوی‌تر

پس از تشکیل ماتریس ارتباط، نگه داشتن همه ارتباط‌های ممکن میان متغیرها هم از نظر محاسباتی هزینه‌بر است و هم می‌تواند باعث حفظ روابط ضعیف و کم‌اهمیت شود؛ به همین دلیل، برای هر گره تنها k ارتباط قوی‌تر نگه داشته می‌شود. اگر مجموعه همسایه‌های انتخاب‌شده برای گره i را با $\mathcal{I}_i$ نشان دهیم، داریم:

$$\mathcal{I}_i = \text{TopK}(X[i, :], k).$$

سپس

$$X[i, j] = 0 \quad \text{اگر} \quad j \notin \mathcal{I}_i.$$

به این ترتیب، برای هر متغیر فقط ارتباط‌های قوی‌تر باقی می‌مانند. این مرحله علاوه بر کاهش هزینه محاسباتی، باعث می‌شود مدل بر روابط مهم‌تر میان متغیرها تمرکز کند.

۲.۲ گراف پیچش

پس از ساخت گراف، باید اطلاعات بین گره‌های مرتبط منتقل شود. برای این منظور، چارچوب LTG از شبکه پیچشی گراف GCN استفاده می‌کند. فرض کنید $\tilde{X}$ ماتریس مجاورت نرمال‌شده گراف و G^{in} نمایش ورودی گره‌ها باشد. انتشار اطلاعات میان گره‌ها به صورت چندمرحله‌ای انجام می‌شود:

$$G^{k+1} = \beta G^{\text{in}} + (1 - \beta) \tilde{X} G^k, \quad k = 0, \dots, K.$$

در این رابطه، β مشخص می‌کند چه مقدار از اطلاعات اولیه هر گره در هر مرحله حفظ شود. اگر مقدار β صفر باشد، نمایش هر گره فقط بر اساس اطلاعات منتشرشده از گراف به‌روزرسانی خواهد شد. اما با قرار دادن یک مقدار مثبت برای β ، بخشی از اطلاعات اولیه گره در تمام مراحل باقی می‌ماند. این ویژگی برای کاهش پدیده بیش‌هموارشدگی مفید است. در بیش‌هموارشدگی، با افزایش تعداد مراحل انتشار، نمایش گره‌های مختلف بیش از حد به یکدیگر شبیه می‌شود. پس از انجام چند مرحله انتشار، خروجی مراحل مختلف با یکدیگر ترکیب می‌شوند:

$$G^{\text{out}} = \sum_{k=0}^K G^k \Theta_k.$$

در این رابطه، Θ_k ماتریس وزن قابل یادگیری مربوط به مرحله k است. از آنجا که گراف یادگرفته‌شده جهت‌دار است، مدل می‌تواند اطلاعات مربوط به جهت‌های مختلف ارتباط را با استفاده از X و X^T در نظر بگیرد.

۳.۲ مدل‌سازی زمانی

در کنار روابط میان متغیرها، رفتار هر متغیر در طول زمان نیز اهمیت زیادی دارد. چارچوب LTG برای مدل‌سازی این وابستگی‌ها از ترکیب LSTM و TCN استفاده می‌کند.

مدل‌سازی زمانی با LSTM

ابتدا توالی ورودی توسط شبکه LSTM پردازش می‌شود. هدف LSTM یادگیری اطلاعات ترتیبی و وابستگی‌هایی است که ممکن است در فاصله‌های زمانی به‌طور نسبی دور قرار داشته باشند. در هر گام زمانی، شبکه با استفاده از دروازه‌های ورودی، فراموشی و خروجی مشخص می‌کند چه مقدار از اطلاعات قبلی حفظ شود و چه مقدار اطلاعات جدید وارد حالت داخلی شبکه شود. خروجی این شبکه به‌عنوان ورودی بخش بعدی مدل استفاده می‌شود.

مدل سازی زمانی با TCN

پس از LSTM، خروجی زمانی وارد شبکه پیچشی زمانی TCN می شود. روش TCN با استفاده از پیچش های زمانی و اتساع یافته می تواند الگوهایی را در مقیاس های زمانی مختلف استخراج کند. برای حفظ اطلاعات ورودی در لایه های مختلف، مقاله از اتصال باقی مانده استفاده می کند:

$$T^{l+1} = \Phi(T^l, P_{l+1}) + \text{TCN}(T^l, \Theta_l).$$

در این رابطه، Φ ورودی مسیر باقی مانده را متناسب با طول خروجی لایه تنظیم می کند. در نتیجه، علاوه بر خروجی TCN، بخشی از اطلاعات ورودی نیز به طور مستقیم به مرحله بعد منتقل می شود.

۴.۲ ترکیب اطلاعات زمانی و گرافی

پس از استخراج ویژگی های زمانی، این اطلاعات وارد بخش گراف می شوند. ترکیب کلی شبکه مکانی-زمانی را می توان به صورت زیر نوشت:

$$T^{l+1} = \Phi(T^l, P_{l+1}) + \text{GCN}(\text{TCN}(T^l, \Theta_l), \Omega).$$

در این ساختار، TCN الگوهای زمانی را استخراج می کند و GCN اطلاعات مربوط به روابط میان متغیرهای مختلف را وارد مدل می کند. بنابراین خروجی هر گره فقط بر اساس تاریخچه زمانی خودش شکل نمی گیرد، بلکه اطلاعات گره های مرتبط نیز در آن تأثیر دارند.

مقدار اولیه این بخش از خروجی LSTM به دست می آید. در پایان، یک شبکه تمام متصل یا پرسپترون چندلایه خروجی شبکه را به مقدار پیش بینی شده یک گام آینده تبدیل می کند. اگر مقدار واقعی سری زمانی در زمان Z برابر با s_Z باشد، خروجی مدل را با $\hat{s}_Z$ نشان داده می شود.

۵.۲ محاسبه خطای پیش بینی

پس از پیش بینی مقدار آینده، اختلاف میان مقدار واقعی و مقدار پیش بینی شده برای هر متغیر محاسبه می شود. برای متغیر i در زمان Z داریم:

$$e_Z^i = s_Z^i - \hat{s}_Z^i.$$

اگر مدل رفتار عادی سامانه را به خوبی یاد گرفته باشد، در حالت عادی مقدار این خطا به طور نسبی کوچک خواهد بود اما اگر در داده رفتار غیر معمولی رخ دهد، پیش بینی مدل ممکن است با مقدار واقعی اختلاف بیشتری داشته باشد و در نتیجه خطای پیش بینی افزایش یابد. با این حال، استفاده مستقیم از خطای خام برای همه متغیرها مناسب نیست، زیرا متغیرهای مختلف ممکن است مقیاس ها و میزان تغییرات متفاوتی داشته باشند. به همین دلیل، مرحله بعدی نرمال سازی خطای پیش بینی است.

۱.۵.۲ نرمال سازی خطای پیش بینی

برای قابل مقایسه شدن خطاهای متغیرهای مختلف، می توان از یک روش نرمال سازی مقاوم استفاده می کند. خطای نرمال شده برای متغیر i به صورت زیر تعریف می شود:

$$\tilde{e}_Z^i = \frac{e_Z^i - \mu_Z^i}{\sigma_Z^i}.$$

در این رابطه، μ_Z^i میانه خطاهای اخیر متغیر i و σ_Z^i فاصله بین چارکی یا IQR همان خطاهاست. به بیان ساده، این مرحله بررسی می کند که خطای فعلی یک متغیر نسبت به خطاهایی که به تازگی برای همان متغیر مشاهده شده اند تا چه اندازه غیرعادی است. به این ترتیب، اگر یک متغیر به طور طبیعی نوسان بیشتری داشته باشد، صرفاً به دلیل بزرگ تر بودن مقیاس خطای آن به عنوان عامل اصلی ناهنجاری در نظر گرفته نمی شود. همچنین برای محاسبه میانه و IQR فقط از اطلاعات گذشته استفاده می شود. بنابراین روش برای تشخیص برخط به اطلاعات آینده وابسته نیست.

۶.۲ امتیازدهی ناهنجاری با PCA

پس از نرمال سازی خطای همه متغیرها، یک بردار خطای نرمال شده تشکیل می شود. این بردار را با $\tilde{A}_Z$ نشان می دهیم. روش PCA روی خطاهای مجموعه اعتبارسنجی اعمال می شود تا ساختار اصلی و رایج خطاهای مربوط به رفتار عادی سامانه شناسایی شود. اگر A_v میانگین خطاهای مجموعه اعتبارسنجی و J ماتریس بردارهای ویژه باشد، ابتدا بردار خطا به فضای مؤلفه های اصلی منتقل می شود:

$$R = (\tilde{A}_Z - A_v)J^T.$$

سپس فقط L مؤلفه اصلی نگه داشته می‌شوند:

$$\tilde{R} = R[:, 1:L],$$

و

$$\tilde{J} = J[:, 1:L].$$

حال بردار خطا با استفاده از این مؤلفه‌ها بازسازی می‌شود:

$$\tilde{A}_Z^{\text{PCA}} = \tilde{R}\tilde{J}^T + A_v.$$

اگر خطای مشاهده‌شده مشابه الگوی معمول خطاها باشد، بازسازی PCA به بردار اصلی نزدیک خواهد بود. اما اگر ساختار خطا غیرمعمول باشد، اختلاف بیشتری میان بردار اصلی و بازسازی‌شده ایجاد خواهد شد.

۱.۶.۲ تعداد مؤلفه‌های اصلی

تعداد مؤلفه‌های اصلی L نباید به صورت دلخواه انتخاب شود. تعداد مؤلفه‌ها به گونه‌ای تعیین می‌شود که خطای بازسازی روی مجموعه اعتبارسنجی مناسب باقی بماند. برای این منظور از معیار SMAPE استفاده می‌شود و تعداد مؤلفه‌های اصلی به گونه‌ای انتخاب می‌شود که مقدار این معیار روی مجموعه اعتبارسنجی کمتر از ۱۰ درصد باشد.

۷.۲ محاسبه امتیاز نهایی ناهنجاری

پس از بازسازی بردار خطا توسط PCA، اختلاف میان خطای واقعی و خطای بازسازی‌شده محاسبه می‌شود. امتیاز ناهنجاری در زمان Z به صورت زیر تعریف می‌شود:

$$\mathcal{X}(Z) = \left\| \tilde{A}_Z^{\text{PCA}} - \tilde{A}_Z \right\|_1.$$

در این رابطه، $\|\cdot\|_1$ فاصله L_1 را نشان می‌دهد. هرچه مقدار $\mathcal{X}(Z)$ بزرگ‌تر باشد، اختلاف الگوی خطای فعلی با الگوی معمول خطاها بیشتر است. بنابراین مقدار بزرگ‌تر امتیاز نشان‌دهنده احتمال بیشتر وجود رفتار غیرعادی در زمان موردنظر است.

۱.۷.۲ آستانه‌گذاری

برای تبدیل امتیاز ناهنجاری به یک تصمیم نهایی، باید یک آستانه تعیین شود و می‌توان آستانه را برابر بیشترین امتیاز ناهنجاری مشاهده‌شده روی مجموعه اعتبارسنجی در نظر گرفت. اگر آستانه را با τ نمایش دهیم، نمونه Z در صورتی ناهنجار در نظر گرفته می‌شود که $\mathcal{X}(Z) > \tau$. مزیت این روش آن است که نیازی به تعیین دستی آستانه ندارد.

فصل ۳

تعمیم چارچوب LTG به سری‌های زمانی مالی

در این فصل، امکان تعمیم چارچوب LTG به سری‌های زمانی مالی بررسی می‌شود. ایده اصلی آن است که بازار مالی نیز می‌تواند به صورت یک سامانه چندمتغیره در نظر گرفته شود که در آن رفتار هر متغیر نه تنها به تاریخچه زمانی خودش، بلکه به رفتار سایر متغیرها نیز وابسته است. در صورت‌بندی پیشنهادی، دارایی‌های مالی به عنوان گره‌های یک گراف در نظر گرفته می‌شوند و بازده یا سایر ویژگی‌های مالی به عنوان سیگنال زمانی گره‌ها وارد مدل می‌شوند. لایه یادگیری هم‌بستگی وظیفه کشف ارتباط میان دارایی‌ها را بر عهده دارد و ترکیب LSTM، GCN و TCN برای مدل‌سازی هم‌زمان وابستگی‌های زمانی و روابط بین دارایی‌ها استفاده می‌شود. در انتهای فصل نیز محدودیت‌های این تعمیم، مسیرهای توسعه و نتیجه‌گیری نهایی پروژه بیان می‌شوند.

۱.۳ انگیزه تعمیم به حوزه مالی

چارچوب LTG برای تشخیص ناهنجاری در سری‌های زمانی چندمتغیره طراحی شده است. هسته اصلی این روش به یک کاربرد صنعتی مشخص محدود نیست. مدل ابتدا روابط میان متغیرها را از داده یاد می‌گیرد، سپس وابستگی‌های زمانی و گرافی را به صورت مشترک برای پیش‌بینی گام آینده مدل می‌کند و در نهایت انحراف خطای پیش‌بینی از الگوی عادی را به یک امتیاز ناهنجاری تبدیل می‌کند [۱۳]. در بازارهای مالی نیز چندین متغیر به طور هم‌زمان تغییر می‌کنند. رفتار یک دارایی می‌تواند با رفتار سایر دارایی‌ها، شاخص‌های بازار و متغیرهای ریسک مرتبط باشد. بنابراین به جای مدل‌سازی هر سری مالی به صورت مستقل، می‌توان بازار را به

شکل یک سامانه چندمتغیره در نظر گرفت و ساختار وابستگی میان دارایی‌ها را در قالب یک گراف مدل کرد.

۱.۱.۳ بازتعریف مسئله مالی

فرض کنید در هر لحظه زمانی t یک بردار مالی چندمتغیره در اختیار باشد:

$$\mathbf{x}_t = [x_t^{(1)}, x_t^{(2)}, \dots, x_t^{(H)}]^T.$$

در این رابطه، هر مؤلفه می‌تواند نماینده یک دارایی، شاخص یا یک ویژگی مالی مشخص باشد. در ساده‌ترین حالت، هر گره گراف یک دارایی است و مقدار گره در زمان t بازده همان دارایی را نشان می‌دهد. استفاده مستقیم از قیمت خام برای چنین مدلی همیشه مناسب نیست، زیرا سطح قیمت دارایی‌های مختلف می‌تواند بسیار متفاوت باشد. یک انتخاب متداول برای ورودی مدل، بازده لگاریتمی است

$$r_t^{(i)} = \log P_t^{(i)} - \log P_{t-1}^{(i)},$$

که در آن $P_t^{(i)}$ قیمت دارایی i در زمان t است. در یک نسخه کامل‌تر، ویژگی هر دارایی می‌تواند علاوه بر بازده شامل حجم معاملات، نوسان‌پذیری، دامنه تغییر قیمت و سایر متغیرهای مالی باشد.

۲.۱.۳ تعریف گره‌ها و ویژگی‌ها

برای تعمیم LTG به داده‌های مالی، چند شیوه برای تعریف گره‌های گراف وجود دارد.

۱. گره به‌عنوان دارایی: هر گره یک سهم، شاخص، ارز یا دارایی دیگر است. ویژگی زمانی هر گره می‌تواند بازده یا مجموعه‌ای از متغیرهای مربوط به همان دارایی باشد.

۲. گره به‌عنوان متغیر مالی: در یک دارایی مشخص، گره‌ها می‌توانند متغیرهایی مانند بازده، حجم معاملات و نوسان‌پذیری باشند.

۳. گراف چندبخشی: می‌توان گروه‌های متفاوتی از گره‌ها را برای دارایی‌ها و شاخص‌های کلان یا بازاری در نظر گرفت. این طراحی نیازمند تغییر بیشتری در معماری پایه است.

برای حفظ بیشترین سازگاری با معماری اصلی LTG، ساده‌ترین انتخاب این است که هر دارایی به‌عنوان یک گره در نظر گرفته شود و سری بازده آن ورودی زمانی گره باشد.

۳.۱.۳ تعریف ناهنجاری در بازار مالی

در این تعمیم، مفهوم ناهنجاری به طور لزوم به معنای تقلب یا یک رخداد منفی نیست. منظور از ناهنجاری، مشاهده‌ای است که با الگوی زمانی و روابط بین دارایی‌هایی که مدل به عنوان رفتار معمول آموخته است سازگار نباشد. چنین رخدادی می‌تواند شامل موارد زیر باشد:

- شوک ناگهانی و غیرمعمول در یک دارایی؛
 - حرکت هم‌زمان غیرمعمول چند دارایی؛
 - شکسته شدن رابطه‌ای که مدل میان دو یا چند دارایی آموخته است؛
 - تغییر رژیم بازار که باعث افزایش پایدار خطای پیش‌بینی شود؛
 - خطای داده یا مشاهده‌ای که با ساختار معمول بازار سازگار نباشد.
- بنابراین امتیاز ناهنجاری در این چارچوب باید به عنوان یک سیگنال آماری تفسیر شود و نه به طور مستقیم به عنوان یک سیگنال خرید یا فروش.

۲.۳ یادگیری ساختار وابستگی مالی

در نسخه اصلی LTG، لایه یادگیری هم‌بستگی^۱ ماتریس مجاورت را به طور مستقیم از نمایش‌های قابل یادگیری گره‌ها ایجاد می‌کند. همین ایده را می‌توان در حوزه مالی برای یادگیری ارتباط میان دارایی‌ها استفاده کرد. در این حالت، مدل لازم نیست یک ماتریس هم‌بستگی ثابت از پیش دریافت کند. در عوض، گرافی متناسب با مسئله پیش‌بینی یاد می‌گیرد. با این حال، یک تفاوت مهم میان داده‌های مالی و بسیاری از داده‌های صنعتی وجود دارد؛ روابط مالی ممکن است در طول زمان تغییر کنند. به عنوان مثال، دو دارایی ممکن است در یک دوره ارتباط قوی داشته باشند، اما در یک دوره دیگر رفتار متفاوتی نشان دهند. بنابراین یکی از توسعه‌های طبیعی مدل آن است که گراف از حالت ثابت به یک گراف وابسته به زمان تبدیل شود. اگر $\mathcal{W}_t$ پنجره اطلاعات موجود تا زمان t باشد، می‌توان به صورت مفهومی نوشت:

$$X_t = \mathcal{G}_\theta(\mathcal{W}_t),$$

که در آن $\mathcal{G}_\theta$ یک تابع قابل یادگیری برای ساخت ماتریس مجاورت در زمان t است. در اولین مرحله پیاده‌سازی می‌توان همان گراف ثابت LTG را حفظ کرد و سپس در توسعه‌های بعدی آن را با یک گراف پویا مقایسه کرد.

^۱Correlation Learning

۱۰۲.۳ مدل سازی زمانی در داده های مالی

بخش زمانی LTG نیز می تواند برای داده های مالی به کار رود. شبکه LSTM برای استخراج وابستگی های ترتیبی و اطلاعاتی که در بازه های زمانی طولانی تر وجود دارند استفاده می شود. در مقابل، TCN با پیچش های زمانی و اتساع یافته می تواند الگوهای محلی و چندمقیاسی را استخراج کند.

پس از استخراج ویژگی های زمانی، این ویژگی ها با ساختار گرافی ترکیب می شوند. در نتیجه، پیش بینی یک دارایی فقط بر اساس تاریخچه همان دارایی انجام نمی شود، بلکه اطلاعات دارایی های مرتبط نیز در پیش بینی آن دخالت می کنند.

برای یک پنجره به طول T ، هدف پیش بینی را می توان به صورت زیر نوشت:

$$\hat{\mathbf{r}}_{t+1} = f_{\theta}(\mathbf{r}_{t-T+1}, \dots, \mathbf{r}_t; X_t),$$

که در آن $\mathbf{r}_t$ بردار بازده تمام دارایی ها و X_t گراف مالی در زمان t است.

۲۰۲.۳ نرمال سازی خطای پیش بینی در داده های مالی

برای هر دارایی i خطای پیش بینی در زمان t به صورت زیر تعریف می شود:

$$e_t^{(i)} = \left| r_t^{(i)} - \hat{r}_t^{(i)} \right|.$$

از آنجا که دارایی های مختلف ممکن است سطح نوسان متفاوتی داشته باشند، مقایسه مستقیم این خطاها می تواند باعث شود دارایی های پرنوسان بر امتیاز کلی غالب شوند. بنابراین می توان از همان نرمال سازی مقاوم مورد استفاده در LTG استفاده کرد:

$$\tilde{e}_t^{(i)} = \frac{e_t^{(i)} - \text{median}\left(e_{t-W:t-1}^{(i)}\right)}{\text{IQR}\left(e_{t-W:t-1}^{(i)}\right) + \varepsilon}.$$

در این رابطه، W طول پنجره مرجع و ε یک مقدار کوچک برای جلوگیری از تقسیم بر صفر است. این نرمال سازی فقط باید با استفاده از اطلاعات گذشته انجام شود تا اطلاعات آینده وارد فرایند تشخیص نشوند.

۳.۲.۳ امتیازدهی PCA برای ناهنجاری مالی

پس از نرمال سازی خطای تمام دارایی ها، بردار زیر تشکیل می شود:

$$\tilde{\mathbf{e}}_t = [\tilde{e}_t^{(1)}, \dots, \tilde{e}_t^{(H)}]^T.$$

سپس می توان همان منطق PCA در مدل اصلی را حفظ کرد. روش PCA روی خطاهای یک دوره مرجع برازش داده می شود و بردار خطا با استفاده از مؤلفه های اصلی بازسازی می شود. امتیاز ناهنجاری مالی پیشنهادی برابر است با:

$$S_t = \|\tilde{\mathbf{e}}_t - \tilde{\mathbf{e}}_t^{\text{PCA}}\|_1.$$

افزایش S_t نشان می دهد که الگوی خطای فعلی با ساختار متداول خطاهای بین دارایی ها تفاوت بیشتری دارد. در یک نسخه توسعه یافته می توان علاوه بر امتیاز کل، سهم هر دارایی را نیز در امتیاز ناهنجاری گزارش کرد.

۳.۳ چالش تغییر رژیم و غیرایستایی

یکی از مهم ترین چالش های داده های مالی، غیرایستایی آن هاست. اگر مدل تنها روی یک دوره تاریخی آموزش ببیند، ممکن است تغییر عادی در شرایط بازار را به عنوان ناهنجاری تشخیص دهد. برای کاهش این مشکل، چند راهکار قابل بررسی است:

۱. آموزش یا به روز رسانی مدل با استفاده از پنجره غلتان؛
۲. بازبرازش دوره ای نرمال ساز خطا و PCA؛
۳. اضافه کردن یک سازوکار تشخیص تغییر رژیم؛
۴. مقایسه گراف ثابت با گراف پویا؛
۵. استفاده از چند مقیاس زمانی برای تفکیک شوک کوتاه مدت از تغییرات پایدار.

۱.۳.۳ جلوگیری از نشت اطلاعات

در مسائل مالی، ترتیب زمانی داده اهمیت بسیار زیادی دارد. مجموعه آموزش باید قبل از مجموعه اعتبارسنجی و مجموعه اعتبارسنجی نیز قبل از مجموعه آزمون قرار گیرد. همچنین تمام مراحل زیر باید فقط بر اساس اطلاعات گذشته انجام شوند:

- استانداردسازی یا نرمال‌سازی ویژگی‌ها؛
- برازش PCA؛
- انتخاب آستانه ناهنجاری؛
- تنظیم ابرپارامترهای مدل؛
- ساخت گراف در نسخه‌های پویا.

استفاده از اطلاعات آینده می‌تواند باعث شود عملکرد مدل به‌صورت غیرواقعی بهتر از مقدار واقعی دیده شود.

۲.۳.۳ محدودیت‌های تعمیم مالی

این تعمیم با چند محدودیت مهم روبه‌رو است.

- نخست، مفهوم ناهنجاری مالی به کاربرد بستگی دارد و هر تغییر بزرگ قیمت به‌طور لزوم به معنای وجود ناهنجاری نیست.
 - دوم، تغییر رژیم می‌تواند هم توزیع داده و هم ساختار گراف را تغییر دهد.
 - سوم، داده‌های مالی ممکن است دارای زمان‌های معامله متفاوت، مقادیر مفقود و ناهم‌زمانی باشند.
 - چهارم، امتیاز ناهنجاری بالا به‌تنهایی وجود رابطه علت و معلولی یا یک فرصت معاملاتی را اثبات نمی‌کند.
- بنابراین این چارچوب باید در درجه اول به‌عنوان یک ابزار پایش و تشخیص تغییرات غیرمعمول در ساختار چندمتغیره بازار در نظر گرفته شود.

۳.۳.۳ محدودیت‌ها

محدودیت‌های اصلی روش عبارت‌اند از:

- آستانه خودکار ممکن است با آستانه بهینه فاصله داشته باشد؛
- تفسیر علت ناهنجاری و سهم هر متغیر هنوز نیازمند توسعه است؛
- ساختار گراف در شرایط پویا ممکن است تغییر کند؛

- پیچیدگی محاسباتی برای کاربرد برخط باید بررسی شود؛
- کیفیت نتیجه به انتخاب پنجره زمانی، تعداد همسایه‌ها و تنظیم ابرپارامترها وابسته است؛

الگوریتم ۱ پروتکل ارزیابی روش مبتنی بر LTG و PCA

Require: داده‌های قیمت و ویژگی‌های مالی، طول پنجره w ، و مجموعه‌های آموزش، اعتبارسنجی و آزمون

Ensure: خطای پیش‌بینی، امتیاز ناهنجاری، و رخدادهای شناسایی شده

۱: تبدیل قیمت‌ها به بازده و استخراج ویژگی‌های مالی

۲: تقسیم زمانی داده‌ها به مجموعه‌های آموزش، اعتبارسنجی و آزمون

۳: تشکیل پنجره‌های ورودی با طول w

۴: آموزش مدل LTG و یادگیری روابط میان دارایی‌ها

۵: پیش‌بینی مقدار گام بعد و محاسبه خطای هر دارایی

۶: نرمال‌سازی مقاوم خطاها با استفاده از میانه و IQR

۷: برازش PCA بر بردارهای خطای مجموعه اعتبارسنجی

۸: محاسبه امتیاز ناهنجاری برای داده‌های آزمون

۹: تعیین آستانه بر اساس مجموعه اعتبارسنجی

۱۰: if امتیاز ناهنجاری از آستانه بیشتر باشد then

۱۱: ثبت زمان به‌عنوان رخداد ناهنجار

۱۲: محاسبه سهم دارایی‌ها در رخداد

نتیجه‌گیری

در این پروژه، ساختار اصلی چارچوب LTG بررسی شد. این چارچوب ابتدا روابط میان متغیرهای سری زمانی را با استفاده از لایه یادگیری هم‌بستگی یاد می‌گیرد و یک گراف جهت‌دار ایجاد می‌کند. سپس، LSTM و TCN برای استخراج اطلاعات زمانی و GCN برای انتقال اطلاعات میان متغیرهای مرتبط به کار می‌روند. خروجی شبکه برای پیش‌بینی یک گام آینده استفاده می‌شود. پس از آن، خطای پیش‌بینی هر متغیر محاسبه و با استفاده از میان و IQR نرمال‌سازی می‌شود.

در مرحله نهایی، PCA الگوی اصلی خطاهای عادی را مدل می‌کند و فاصله میان خطای نرمال‌شده و بازسازی آن به‌عنوان امتیاز ناهنجاری در نظر گرفته می‌شود. نتایج بررسی آستانه‌گذاری نشان می‌دهد که میان آستانه خودکار و آستانه بهینه حاصل از بیشینه‌سازی معیار F1 اختلاف وجود دارد. بنابراین، تعیین آستانه مناسب همچنان یکی از مسائل قابل بررسی در ادامه این کار است.

در ادامه، چارچوب LTG در قالب یک سناریوی مالی مورد بررسی قرار گرفت. در این صورت‌بندی، دارایی‌ها به‌عنوان گره‌های گراف و بازده‌ها و سایر ویژگی‌های مالی به‌عنوان ورودی‌های زمانی مدل در نظر گرفته می‌شوند. همچنین، در کاربرد مالی عواملی مانند غیرایستایی، تغییر رژیم، هم‌ترازی زمانی و جلوگیری از نشت اطلاعات اهمیت ویژه‌ای دارند. این صورت‌بندی در این پروژه در سطح طراحی روش بررسی شده است و ارزیابی تجربی آن انجام نشده است.

ادامه این کار می‌تواند بر پیاده‌سازی و ارزیابی تجربی نسخه مالی متمرکز باشد. در این راستا، انتخاب مجموعه مشخصی از دارایی‌ها، تعیین تناوب زمانی داده‌ها، مقایسه گراف ثابت LTG با گراف پویا، بررسی اثر بازده، حجم معاملات و نوسان‌پذیری به‌عنوان ورودی، طراحی ارزیابی زمانی بدون نشت اطلاعات، تحلیل سهم هر دارایی در امتیاز ناهنجاری، بررسی رفتار مدل در دوره‌های تغییر رژیم و بازبرازش دوره‌ای PCA و آستانه ناهنجاری از جمله موارد قابل بررسی هستند. هر نتیجه عددی درباره این صورت‌بندی باید پس از اجرای آزمایش و ثبت دقیق تنظیمات گزارش شود.

به‌طور خلاصه، چارچوب LTG نمونه‌ای از مدل‌هایی است که در کنار وابستگی زمانی، ساختار ارتباطی میان متغیرها را نیز به‌صورت صریح در نظر می‌گیرد. در این روش، اطلاعات زمانی و ساختار گرافی به‌صورت هم‌زمان برای پیش‌بینی و تشخیص ناهنجاری به کار گرفته می‌شوند. صورت‌بندی مالی بررسی‌شده در این پروژه نشان می‌دهد که همین ساختار می‌تواند برای مدل‌سازی روابط میان دارایی‌های مالی نیز مورد استفاده قرار گیرد. با این حال، بررسی اعتبار عملی آن در حوزه مالی نیازمند اجرای آزمایش‌های واقعی، کنترل دقیق نشت اطلاعات و بررسی عملکرد مدل در شرایط مختلف بازار است.

پیوست ۱

چک‌لیست ارزیابی مالی

در این پیوست، مهم‌ترین موارد مورد نیاز برای پیاده‌سازی و ارزیابی تجربی چارچوب LTG در داده‌های مالی بیان می‌شود. ثبت این موارد برای جلوگیری از نشت اطلاعات و امکان بازتولید نتایج ضروری است.

۱. فهرست دارایی‌ها، منبع داده‌ها و تناوب نمونه‌برداری؛
۲. تعریف قیمت، فرمول بازده و سایر ویژگی‌های ورودی؛
۳. روش مدیریت داده‌های مفقود و هم‌ترازی زمانی دارایی‌ها؛
۴. مرز زمانی مجموعه‌های آموزش، اعتبارسنجی و آزمون؛
۵. روش نرمال‌سازی و اطمینان از عدم استفاده از اطلاعات آینده؛
۶. طول پنجره ورودی و افق پیش‌بینی؛
۷. تنظیمات آموزش شامل seed، optimizer، نرخ یادگیری، اندازه batch و تعداد epoch؛
۸. تنظیمات LSTM، TCN و GCN و نوع گراف مورد استفاده؛
۹. روش برازش PCA و تعیین آستانه تشخیص ناهنجاری؛
۱۰. تعریف ناهنجاری، معیارهای ارزیابی و روش‌های مبنای؛
۱۱. نسخه کد و کتابخانه‌های نرم‌افزاری مورد استفاده.

پیش از گزارش نتایج عددی، تمام مراحل پیش‌پردازش، نرمال‌سازی، آموزش، برازش PCA و تعیین آستانه باید ثبت شوند. همچنین باید اطمینان حاصل شود که اطلاعات آینده در هیچ‌یک از مراحل آموزش و ارزیابی استفاده نشده است.

کتاب نامه

- [1] Abu-El-Haija, S., Perozzi, B., Kapoor, A., Alipourfard, N., Lerman, K., Harutyunyan, H., & Ver Steeg, G. (2019). MixHop: Higher-order graph convolutional architectures via sparsified neighborhood mixing. *Proceedings of the International Conference on Machine Learning*. 21–29.
- [2] Bai, S., Kolter, J. Z., & Koltun, V. (2018). An empirical evaluation of generic convolutional and recurrent networks for sequence modeling. *arXiv preprint arXiv:1803.01271*.
- [3] Deng, A., & Hooi, B. (2021). Graph neural network-based anomaly detection in multivariate time series. *Proceedings of the AAAI Conference on Artificial Intelligence*, 35. 4027–4035.
- [4] Park, D., Hoshi, Y., & Kemp, C. C. (2018). A multimodal anomaly detector for robot-assisted feeding using an LSTM-based variational autoencoder. *IEEE Robotics and Automation Letters*, 3. 1544–1551.
- [5] Ren, H., Xu, B., Wang, Y., Yi, C., Huang, C., Kou, X., Xing, T., Yang, M., Tong, J., & Zhang, Q. (2019). Time-series anomaly detection service at Microsoft. *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. 3009–3017.
- [6] Su, Y., Zhao, Y., Niu, C., Liu, R., Sun, W., & Pei, D. (2019). Robust anomaly detection for multivariate time series through stochastic recurrent neural network. *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. 2828–2837.

- [7] Wu, F., Souza, A., Zhang, T., Fifty, C., Yu, T., & Weinberger, K. (2019). Simplifying graph convolutional networks. *Proceedings of the International Conference on Machine Learning*. 6861–6871.
- [8] Wu, Z., Pan, S., Long, G., Jiang, J., Chang, X., & Zhang, C. (2020). Connecting the dots: Multivariate time series forecasting with graph neural networks. *Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. 753–763.
- [9] Xia, F., Sun, K., Yu, S., Aziz, A., Wan, L., Pan, S., & Liu, H. (2021). Graph learning: A survey. *IEEE Transactions on Artificial Intelligence*, 2. 109–127.
- [10] Zamanzadeh Darban, Z., Webb, G. I., Pan, S., Aggarwal, C., & Salehi, M. (2024). Deep learning for time series anomaly detection: A survey. *ACM Computing Surveys*, 57. 1–42.
- [11] Zhang, C., Song, D., Chen, Y., Feng, X., Lumezanu, C., Cheng, W., Ni, J., Zong, B., Chen, H., & Chawla, N. V. (2019). A deep neural network for unsupervised anomaly detection and diagnosis in multivariate time series data. *Proceedings of the AAAI Conference on Artificial Intelligence*, 33. 1409–1416.
- [12] Zhang, H., Wu, B., Yuan, X., Pan, S., Tong, H., & Pei, J. (2024). Trustworthy graph neural networks: Aspects, methods, and trends. *Proceedings of the IEEE*.
- [13] Zhang, J., Wang, X., Yang, Y., Miao, H., & Yang, S. (2025). A novel anomaly detection method for multivariate time series based on spatial-temporal graph learning. *Journal of King Saud University Computer and Information Sciences*, 37. DOI: 10.1007/s44443-025-00024-3.
- [14] Zheng, Y., Koh, H. Y., Jin, M., Chi, L., Phan, K. T., Pan, S., Chen, Y.-P. P., & Xiang, W. (2023). Correlation-aware spatial-temporal graph learning for multivariate time-series anomaly detection. *IEEE Transactions on Neural Networks and Learning Systems*.

- [15] Zhu, H., Yi, C., Rho, S., Liu, S., & Jiang, F. (2023). An interpretable multivariate time-series anomaly detection method in cyber-physical systems based on adaptive mask. *IEEE Internet of Things Journal*.

واژه‌نامه فارسی به انگلیسی

آ

Threshold آستانه

ا

Residual connection اتصال باقیمانده

Anomaly score امتیاز ناهنجاری

Top-K انتخاب k مقدار برتر

ب

Logarithmic return بازده لگاریتمی

Embedding بازنمایی نهفته

Oversmoothing بیش‌هموارشدگی

پ

Prediction پیش‌بینی

One-step-ahead prediction پیش‌بینی یک گام آینده

ت

Principal component analysis تحلیل مؤلفه‌های اصلی

Anomaly detection تشخیص ناهنجاری

Regime change تغییر رژیم

خ

Prediction error خطای پیش‌بینی

س

Time series سری زمانی

Multivariate time series سری زمانی چندمتغیره

ش

Graph neural network شبکه عصبی گراف

Temporal convolutional network شبکه پیچشی زمانی

Graph convolutional network شبکه پیچشی گراف

ف

Interquartile range فاصله بین چارکی

غ

Non-stationarity غیرایستایی

ک

Channel کانال

Convolution پیچش

Graph convolution پیچش گراف

گ

Graph گراف

Dynamic graph گراف پویا

Directed graph گراف جهت‌دار
Node گره

م

Adjacency matrix ماتریس مجاورت
Training set مجموعه آموزش
Test set مجموعه آزمون
Validation set مجموعه اعتبارسنجی
Median میانه

ن

Anomaly ناهنجاری
Normalization نرمال‌سازی
Data leakage نشت اطلاعات

و

Feature ویژگی

ی

Spatial-temporal graph learning یادگیری گراف مکانی-زمانی
Correlation learning یادگیری هم‌بستگی
Edge یال

واژه‌نامه انگلیسی به فارسی

A

Adjacency matrix		ماتریس مجاورت
Anomaly		ناهنجاری
Anomaly detection		تشخیص ناهنجاری
Anomaly score		امتیاز ناهنجاری

C

Channel		کانال
Convolution		پیچش
Correlation learning		یادگیری هم‌بستگی

D

Data leakage		نشت اطلاعات
Directed graph		گراف جهت‌دار
Dynamic graph		گراف پویا

E

Edge		یال
Embedding		بازنمایی نهفته

F

Feature ویژگی

Forecasting پیش‌بینی

G

Graph گراف

Graph convolution پیچش گراف

Graph convolutional network شبکه پیچشی گراف

Graph neural network شبکه عصبی گراف

I

Interquartile range فاصله بین چارکی

L

Logarithmic return بازده لگاریتمی

Long short-term memory حافظه کوتاه‌مدت بلندمدت

M

Median میانه

Multivariate time series سری زمانی چندمتغیره

N

Node..... گره
Non-stationarity..... غیرایستایی
Normalization..... نرمال‌سازی

O

One-step-ahead prediction..... پیش‌بینی یک گام آینده
Oversmoothing..... بیش‌هموارشدگی

P

Prediction..... پیش‌بینی
Prediction error..... خطای پیش‌بینی
Principal component analysis..... تحلیل مؤلفه‌های اصلی

R

Regime change..... تغییر رژیم
Residual connection..... اتصال باقیمانده

S

Spatial-temporal graph learning..... یادگیری گراف مکانی-زمانی

T

Temporal convolutional network	شبکه پیچشی زمانی
Test set	مجموعه آزمون
Threshold	آستانه
Time series	سری زمانی
Top-K	انتخاب k مقدار برتر
Training set	مجموعه آموزش

V

Validation set	مجموعه اعتبارسنجی
----------------	-------------------

فهرست نمادها

نماد	مفهوم
S	سری زمانی چندمتغیره
s_z	بردار مشاهدات در زمان z
Z	تعداد نقاط زمانی
H	تعداد متغیرها یا کانال‌های سری زمانی
d	بعد بازنمایی گره
X	ماتریس مجاورت آموخته‌شده گراف
k	تعداد همسایه‌های حفظ‌شده در مرحله انتخاب ارتباط‌های قوی‌تر
α	پارامتر کنترل اشباع در لایه یادگیری هم‌بستگی
β	ضریب حفظ اطلاعات اولیه در انتشار گراف
G^k	نمایش گره‌ها در مرحله k انتشار گراف
Θ_k	ماتریس وزن قابل یادگیری در مرحله k انتشار گراف

خطای پیش‌بینی متغیر i در زمان Z	e_Z^i
خطای پیش‌بینی نرمال‌شده متغیر i در زمان Z	$\tilde{e}_Z^i$
بردار خطاهای نرمال‌شده در زمان Z	$\tilde{A}_Z$
امتیاز ناهنجاری در زمان Z	$\mathcal{X}(Z)$
آستانه تشخیص ناهنجاری	τ
تعداد مؤلفه‌های اصلی انتخاب‌شده	L
قیمت دارایی i در زمان t	$P_t^{(i)}$
بازده لگاریتمی دارایی i در زمان t	$r_t^{(i)}$
ماتریس وابستگی یا گراف مالی در زمان t	X_t
امتیاز ناهنجاری مالی در زمان t	S_t
طول پنجره مرجع برای نرمال‌سازی خطا	W
مقدار کوچک برای جلوگیری از تقسیم بر صفر	ε

فهرست کلمات اختصاری

کلمه کامل	معنی	اختصار
LSTM-TCN-GCN	چارچوب پیشنهادی مقاله	LTG
Long Short-Term Memory	حافظه کوتاه مدت بلند مدت	LSTM
Temporal Convolutional Network	شبکه کانولوشنی زمانی	TCN
Graph Convolutional Network	شبکه کانولوشنی گراف	GCN
Graph Neural Network	شبکه عصبی گراف	GNN
Principal Component Analysis	تحلیل مؤلفه های اصلی	PCA
Interquartile Range	فاصله بین چارکی	IQR
Symmetric MAPE	میانگین درصد خطای مطلق متقارن	sMAPE