

بهبود صحت طبقه‌بندی گره‌ها در شبکه‌های عصبی گراف با استفاده از PageRank به عنوان یک ویژگی اضافی برای گره‌ها

حسین حسینی^۱، مهدی نقوی^{۲*}

^۱ دانشجوی دکتری هوش مصنوعی و رباتیک، دانشکده هوش مصنوعی و علوم شناختی، دانشگاه جامع امام حسین (ع)، تهران، ایران

^۲ استادیار، دانشکده هوش مصنوعی و علوم شناختی، دانشگاه جامع امام حسین (ع)، تهران، ایران

مقاله پژوهشی

چکیده

در این مقاله، راهکاری ساده و مؤثر برای بهبود عملکرد شبکه‌های عصبی گراف در مسئله طبقه‌بندی گره‌ها ارائه شده است. روش پیشنهادی شامل افزودن نمره PageRank به عنوان یک معیار جهانی مرکزیت به بردار ویژگی گره‌هاست تا اطلاعات زمینه‌ای فراتر از همسایگی محلی در مدل لحاظ شود. برای ارزیابی، سه معماری مشهور شبکه‌های عصبی گراف شامل شبکه‌های کانولوشنی گراف، شبکه‌های توجهی گراف و شبکه‌های نمونه‌گیری و تجمیع از روی گراف روی مجموعه داده‌های CiteSeer، Cora و PubMed آزمایش شده‌اند. عملکرد مدل‌ها با معیارهایی چون صحت، دقت، فراخوانی، نمره F1 و تحلیل‌های بصری مبتنی بر روش‌هایی همچون روش تحلیل مؤلفه‌های اصلی، روش تعبیه همسایه‌های تصادفی با توزیع t و روش تقریب و تصویرسازی یکنواخت چند فرم بررسی شده است. نتایج تجربی نشان می‌دهند که گنجاندن PageRank باعث بهبود معنی‌دار در دقت طبقه‌بندی می‌شود، به‌ویژه در معماری شبکه‌های توجهی گراف که از اطلاعات جهانی بهره بیشتری می‌برد. این روش در عین سادگی، هزینه محاسباتی کمی دارد و عملکرد پایدار و قابل اعتمادی را در داده‌های مختلف ارائه می‌دهد. در نهایت، مقاله امکان گسترش این رویکرد با استفاده از سایر معیارهای مرکزیت و کاربرد آن در گراف‌های بزرگ یا ناهمگن را نیز مطرح می‌سازد.

تاریخ دریافت:

۱۴۰۴/۳/۶

تاریخ پذیرش:

۱۴۰۴/۷/۱۶

کلیدواژه‌ها:

شبکه‌های عصبی گراف،
طبقه‌بندی گره‌ها، نمره
PageRank، شبکه‌های
کانولوشنی گراف، شبکه‌های
توجهی گراف، شبکه‌های
نمونه‌گیری و تجمیع از روی
گراف.

نویسنده مسئول:

hosseinhosseini@ihu.ac.ir

doi : 10.22034/ABMIR.2025.23200.1130

۱- مقدمه

کدگذاری شده‌اند. همچنین، پدیده هموارسازی بیش از حد^۶ که در آن نمایش گره‌ها با افزایش تعداد لایه‌های تجمیع غیرقابل تمایز می‌شوند، عملاً عمق و قابلیت بیان شبکه‌های عصبی گراف را محدود می‌کند [۵، ۶].

طبقه‌بندی گره‌ها که یکی از وظایف اصلی در یادگیری از گراف‌هاست، نمونه‌ای از این چالش‌هاست. طبقه‌بندی دقیق گره‌ها نیازمند آن است که مدل‌ها بتوانند اطلاعات محلی و جهانی را به‌خوبی تلفیق کنند و گره‌هایی را که در همسایگی نزدیک خود مشابه هستند ولی نقش‌های ساختاری متفاوتی در کل گراف دارند، از یکدیگر تمایز دهند. شبکه‌های عصبی گراف موجود اغلب در تعادل بین این دو نوع اطلاعات با مشکل مواجه می‌شوند که انگیزه‌ای برای تحقیقات در راستای روش‌هایی فراهم کرده است که بتوانند زمینه جهانی را به نمایش گره‌ها اضافه کنند.

یکی از راهکارهای امیدبخش، ادغام ویژگی‌های ساختاری حاصل از نظریه گراف است، از جمله درجه گره، ضریب خوشه‌بندی و انواع مختلف معیارهای مرکزیت. در میان این معیارها، PageRank به‌عنوان الگوریتمی گسترده‌ای برای سنجش اهمیت جهانی گره‌ها براساس ساختار لینکی گراف شناخته شده است. این الگوریتم در ابتدا برای رتبه‌بندی صفحات وب طراحی شد و نمره‌های بالاتری را به گره‌هایی اختصاص می‌دهد که اتصالات بیشتری دارند یا نقش پل بین بخش‌های مختلف شبکه را بازی می‌کنند. مطالعات اخیر به بررسی رابطه بین PageRank و شبکه‌های عصبی گراف پرداخته‌اند و معماری‌هایی را پیشنهاد داده‌اند که از PageRank شخصی‌سازی‌شده برای بهبود روش‌های انتشار استفاده می‌کنند [۵-۱۱].

با این حال، استفاده مستقیم از PageRank به عنوان یک ویژگی صریح برای گره‌ها به جای استفاده از آن به عنوان یک مکانیسم انتشار کمتر مورد توجه قرار گرفته است. به‌صورت شهودی،

داده‌های ساختاریافته به شکل گراف در بسیاری از کاربردهای دنیای واقعی رایج هستند، از جمله شبکه‌های اجتماعی و شبکه‌های نقل‌قول تا سیستم‌های زیستی و موتورهای پیشنهاددهنده. پیچیدگی ذاتی گراف‌ها که با اتصالات غیرمنظم و اطلاعات رابطه‌ای غنی مشخص می‌شوند، چالش‌های منحصربه‌فردی برای مدل‌های یادگیری ماشین ایجاد می‌کنند. رویکردهای سنتی مانند روش‌های کرنل و تکنیک‌های تعبیه^۱ سطحی اغلب ناتوانند الگوهای ساختاری محلی و جهانی را به‌طور هم‌زمان به دست آورند که این امر مؤثر بودن آن‌ها را در وظایفی چون طبقه‌بندی گره‌ها، پیش‌بینی لینک و تشخیص انجمن محدود می‌کند.

شبکه‌های عصبی گراف^۲ به‌عنوان یک رویکرد قدرتمند برای یادگیری از داده‌های ساختار گرافی درآمده‌اند. با تجمیع مداوم اطلاعات از همسایگی‌های محلی، شبکه‌های عصبی گراف قادرند نمایش‌های برجسته‌ای از گره‌ها یاد بگیرند که شامل اطلاعات ویژگی و توپولوژیکی می‌شوند. دهه گذشته شاهد پیشرفت‌های سریع در طراحی معماری شبکه‌های عصبی گراف بوده است؛ مدل‌های بنیادینی مانند شبکه‌های کانولوشنی گراف^۳، شبکه‌های توجهی گراف^۴ و شبکه‌های نمونه‌گیری و تجمیع از روی گراف^۵ معیارهای جدیدی برای صحت طبقه‌بندی گره‌ها تعیین کرده‌اند. این مدل‌ها کاربردهای گسترده‌ای در حوزه‌هایی چون کشف دارو، تشخیص کلاهبرداری، سیستم‌های پیشنهاددهنده و تحلیل شبکه‌های مغزی یافته‌اند [۱-۴].

با وجود موفقیت‌های آن‌ها، شبکه‌های عصبی گراف با چندین محدودیت مواجه هستند. بیشتر معماری‌ها به تبادل پیام درون یک تعداد ثابت از پرش‌ها^۷ متکی هستند که این امر میدان دید^۸ مدل را به همسایگی‌های محلی محدود می‌کند. این موضوع می‌تواند منجر به عملکرد زیربهبوده^۹ شود، به‌ویژه در گراف‌هایی که اطلاعات ساختاری مهم در وابستگی‌های بلندمدت یا الگوهای اتصال جهانی

^۶ Hops

^۷ Receptive field

^۸ Suboptimal

^۹ Over-smoothing

^۱ Embedding

^۲ Graph Neural Networks (GNNs)

^۳ Graph Convolutional Networks (GCNs)

^۴ Graph Attention Networks (GATs)

^۵ Graph SAmple and aggreGatE (GraphSAGE)

ج) بینش‌هایی درباره تعامل بین ویژگی‌های محلی و جهانی در شبکه‌های عصبی گراف ارائه می‌دهیم و اهمیت افزودن ویژگی‌های ساختاری را برجسته می‌کنیم.

د) محدودیت‌های روش پیشنهادی را برجسته می‌کنیم و مسیرهایی برای پژوهش‌های آتی را پیشنهاد می‌دهیم، از جمله ادغام سایر معیارهای مرکزیت و گسترش روش برای گراف‌های بزرگ‌تر یا ناهمگن.

ساختار مقاله به شرح زیر است:

بخش دوم: کارهای مرتبط را در مورد معماری‌های شبکه‌های عصبی گراف، ادغام ویژگی‌های ساختاری و استفاده از PageRank در یادگیری از گراف مرور می‌کند.

بخش سوم: روش پیشنهادی ما را به تفصیل توصیف می‌کند، شامل نحوه افزودن ویژگی و فرآیند آموزش مدل.

بخش چهارم: تنظیمات آزمایشی، مجموعه داده‌ها و معیارهای ارزیابی را توضیح می‌دهد.

بخش پنجم: نتایج آزمایش‌ها را ارائه و تحلیل می‌کند و عملکرد مدل‌های مبنایی و مدل‌های اصلاح‌شده را با یکدیگر مقایسه می‌کند.

بخش ششم: مقاله را با جمع‌بندی یافته‌ها به انجام می‌رساند.

بخش هفتم: مسیرهای بالقوه برای پژوهش‌های آتی را مطرح می‌کند.

۲- کارهای مرتبط

شبکه‌های عصبی گرافی در سال‌های اخیر به سرعت پیشرفت کرده‌اند و انواع معماری‌ها و کاربردهای مختلفی را در بر گرفته‌اند. این بخش مروری است بر مدل‌های کلاسیک و جدید شبکه‌های عصبی گراف، رویکردهای بهبود طبقه‌بندی گره‌ها، تلفیق ویژگی‌های ساختاری و همچنین نقش PageRank و معیارهای مشابه در یادگیری ماشین مبتنی بر گراف.

PageRank اطلاعات ساختاری جهانی را در اختیار می‌گذارد که می‌تواند به عنوان مکملی برای تجمیع محلی عمل کند و قدرت تمایزدهندگی نمایش گره‌ها را افزایش دهد. با اضافه کردن نمره‌های PageRank به بردارهای ویژگی اصلی، فرض می‌کنیم که شبکه‌های عصبی گراف قادر خواهند بود گره‌ها را بهتر براساس اهمیت جهانی آن‌ها از یکدیگر تمایز دهند و این امر منجر به بهبود عملکرد طبقه‌بندی خواهد شد.

این مقاله به صورت سیستماتیک تأثیر ادغام PageRank به عنوان یک ویژگی اضافی برای طبقه‌بندی گره‌ها در چارچوب شبکه‌های عصبی گراف را بررسی می‌کند. ما روی سه معماری نمونه از شبکه‌های عصبی گراف شامل شبکه‌های کانولوشنی گراف، شبکه‌های توجیهی گراف و شبکه‌های نمونه‌گیری و تجمیع از روی گراف تمرکز کرده‌ایم و روش خود را روی مجموعه داده‌های استاندارد شبکه‌های نقل قول شامل Cora، CiteSeer و PubMed ارزیابی می‌کنیم. روش ما شامل محاسبه نمره‌های PageRank برای هر گره، افزون کردن ماتریس‌های ویژگی و آموزش مدل‌ها با پروتکل‌های استاندارد است. عملکرد مدل‌ها را با استفاده از معیارهای صحت^۱، دقت^۲، فراخوانی^۳، نمره F1^۴ و ماتریس‌های پراکندگی^۵ ارزیابی می‌کنیم، ضمن تصویرسازی نتایج با استفاده از روش‌های نمایش داده‌های چندبعدی در فضایی با بعد کمتر شامل روش تحلیل مؤلفه‌های اصلی^۶، روش تعبیه همسایه‌های تصادفی با توزیع t^۷ و روش تقریب و تصویرسازی یکنواخت چندفرم^۸.

یافته‌ها اصلی این تحقیق به صورت زیر خلاصه می‌شوند:

الف) ارزیابی تجربی جامعی از تأثیر افزودن PageRank بر صحت طبقه‌بندی گره‌ها در معماری‌ها و مجموعه داده‌های مختلف شبکه‌های عصبی گراف ارائه می‌دهیم.

ب) نشان می‌دهیم که افزودن PageRank به طور مداوم عملکرد را بهبود می‌دهد که بیشترین بهبود در مورد مدل شبکه‌های توجیهی گراف مشاهده می‌شود.

⁶ Principal Component Analysis (PCA)

⁷ T-Distributed Stochastic Neighbor Embedding (t-SNE)

⁸ Uniform Manifold Approximation and Projection (UMAP)

¹ Accuracy

² Precision

³ Recall

⁴ F1-score

⁵ Confusion matrix

۱-۲ مدل‌های کلاسیک شبکه‌های عصبی گراف

سه مدل بنیانی شبکه‌های کانولوشنی گراف، شبکه‌های توجهی گراف و شبکه‌های نمونه‌گیری و تجمع از روی گراف زمینه را برای یادگیری نمایشی گراف‌های مدرن فراهم کرده‌اند. شبکه کانولوشنی گرافی کیپف و ولینگ، با معرفی کانولوشن‌های طیفی گراف، قادر شد همسایگی‌ها را به صورت مؤثر تجمع کند و یادگیری نیمه-ناظری بر روی گراف‌های دارای برچسب گره‌ای را تسهیل کند [۱]. شبکه‌های کانولوشنی گراف عملکرد قوی خود را در شبکه‌های استنادی، گراف‌های اجتماعی و شبکه‌های بیولوژیکی نشان داده‌اند که از ویژگی‌های گره و اطلاعات مجاورت استفاده می‌کنند.

شبکه‌های توجهی گرافی، که توسط ولاچکوچ و همکاران مطرح شدند، شبکه‌های کانولوشنی گراف را با معرفی مکانیسم‌های توجه گسترش دادند که وزن‌های قابل یادگیری را به گره‌های مجاور اختصاص می‌دهند [۵]. این امر به مدل اجازه می‌دهد در حین تجمع، تمرکز خود را بر روی مهم‌ترین همسایه‌ها قرار دهد که منجر به افزایش استحکام و تفسیرپذیری می‌شود. شبکه‌های توجهی گراف در گراف‌های ناهمگن و پرتراکم ثابت کرده‌اند که تجمع یکنواخت ممکن است سیگنال‌های مهم را ضعیف کند. شبکه‌های نمونه‌گیری و تجمع از روی گراف که توسط همیلتون و همکاران توسعه یافت، با نمونه‌برداری از تعداد ثابتی از همسایه‌ها و تجمع ویژگی‌های آن‌ها از طریق عملیات میانگین، شبکه عصبی حافظه طولانی کوتاه مدت^۱ یا لایه ادغام^۲، مقیاس‌پذیری را بهبود بخشید [۷]. این رویکرد القایی^۳ امکان آموزش کارآمد بر روی گراف‌های بزرگ و تعمیم‌پذیری به گره‌های دیده نشده را فراهم کرده و شبکه‌های نمونه‌گیری و تجمع از روی گراف را برای شبکه‌های پویا یا در حال تحول مناسب می‌کند.

۲-۲ پیشرفت‌های اخیر در شبکه‌های عصبی گراف

در سال‌های اخیر، پژوهش‌های شبکه‌های عصبی گراف شاهد پیشرفت‌های قابل توجهی بوده است، به‌طوری که معماری‌ها و تکنیک‌های بهینه‌سازی جدیدی برای رفع محدودیت‌های مدل‌های کلاسیک ابداع شده‌اند. یکی از روندهای عمده توسعه، شبکه‌های عصبی گراف مقیاس‌پذیر است که قادرند با گراف‌های بزرگ شامل میلیون‌ها گره و یال کار کنند. به عنوان مثال، مدل PPRGo از میلیون‌ها گره و یال کار کنند. به عنوان مثال، مدل PPRGo از میلیون‌ها گره و یال کار کنند. به عنوان مثال، مدل PPRGo از میلیون‌ها گره و یال کار کنند.

جهت دیگر، افزایش قدرت بیانی^۵ شبکه‌های عصبی گراف از طریق تلفیق ساختارهای مرتبه بالاتر، مانند الگوهای شبکه‌ای^۶، زیرگراف‌ها یا یال‌های فرامنطقه‌ای است. شبکه‌های عصبی گرافی کوانتومی^۷ نیز در زمینه علم مواد مورد بررسی قرار گرفته‌اند و نشان داده‌اند که در مقایسه با مدل‌های کلاسیک، همگرایی سریع‌تر و خطای آزمون پایین‌تری دارند [۴]. مرورهایی از واک و همکاران و ژانگ و همکاران، دید جامعی از تکنیک‌های شتاب‌دهی شبکه‌های عصبی گراف و کاربردهای نوظهور در حوزه‌هایی مانند تصویربرداری مغزی و طبقه‌بندی نوار مغزی^۸ ارائه می‌دهند [۱-۳، ۱۲].

علاوه بر این، روش‌های انتشار جدید الهام‌گرفته از PageRank و نسخه‌های آن پیشنهاد شده‌اند. شبکه‌های عصبی گراف مبتنی بر PageRank تعمیم‌یافته^۹ به‌صورت تطبیقی وزن‌های انتشار را یاد می‌گیرد تا تعادلی بین اطلاعات محلی و جهانی ایجاد کند و در نتیجه قابلیت جهانی شدن را در گراف‌های هم‌خویش^{۱۰} و ناهم‌خویش^{۱۱} به دست آورد [۵]. PPNP^{۱۲} و تقریب سریع آن APPNP، از PageRank شخصی‌سازی شده برای گسترش دامنه

⁷ Quantum Graph Neural Networks (QGNN)

⁸ ElectroEncephaloGraphy (EEG)

⁹ Generalized PageRank Graph Neural Networks

¹⁰ Homophilic

¹¹ Heterophilic

¹² Personalized Propagation of Neural Predictions

¹ Long Short Term Memory (LSTM)

² Pooling

³ Inductive

⁴ State-of-the-art

⁵ Expressive power

⁶ Motifs

۲-۴ ویژگی‌های ساختاری در شبکه‌های عصبی

گراف

تلفیق ویژگی‌های ساختاری در شبکه‌های عصبی گراف به طور گسترده‌ای مطالعه شده است. مرکزیت درجه که تعداد اتصالات یک گره را کمی می‌کند، یک ویژگی ساده اما مؤثر برای تمایز دادن گره‌های مرکزی^۹ از گره‌های حاشیه‌ای^{۱۰} است. ضریب خوشه‌بندی میزان تمایل همسایه‌های یک گره به تشکیل مثلث^{۱۱} را اندازه‌گیری کرده و یکپارچگی محلی را منعکس می‌کند. بینیت^{۱۲} و مرکزیت بردار ویژه^{۱۳} دیدگاه‌های دیگری نسبت به اهمیت گره‌ها ارائه می‌دهند که نقش آن‌ها را به عنوان پل^{۱۳} یا تأثیرگذاران^{۱۴} در شبکه برجسته می‌کنند.

بررسی‌های اخیر، رواج افزونگی ویژگی‌های ساختاری در طبقه‌بندی‌های مبتنی بر شبکه‌های عصبی گراف را به ویژه برای وظایفی مانند طبقه‌بندی نوار مغزی و تحلیل شبکه مغزی برجسته کرده‌اند [۳، ۱]. این ویژگی‌ها اغلب با ویژگی‌های اصلی گره ترکیب می‌شوند تا مدل‌ها بتوانند از هم محتوا و هم ساختار استفاده کنند. با این حال، انتخاب ویژگی‌ها و اهمیت نسبی آن‌ها همچنان سوالاتی باز در پژوهش است.

۲-۵ PageRank و معیارهای رتبه‌بندی مشابه در

یادگیری مبتنی بر گراف

PageRank که در ابتدا برای رتبه‌بندی صفحات وب توسعه یافت، توزیع ایستا یک حرکت تصادفی با بازنشانی^{۱۵} را محاسبه کرده و نمره بالاتری به گره‌هایی می‌دهد که از طریق مسیرهای زیادی قابل دسترسی باشند. اثربخشی آن در منعکس کردن اهمیت جهانی گره‌ها، منجر به استفاده گسترده آن در بازیابی اطلاعات، تحلیل شبکه‌های اجتماعی و سیستم‌های پیشنهادی شده است.

دریافتی بدون ایجاد پدیده هموارسازی بیش از حد استفاده کرده و طبقه‌بندی گره‌ها را بهبود می‌بخشند [۷].

۲-۳ رویکردهای بهبود صحت طبقه‌بندی گره

افزایش صحت طبقه‌بندی گره همچنان یکی از چالش‌های مرکزی در تحقیقات شبکه‌های عصبی گراف است. چندین استراتژی ارائه شده است، از جمله معماری‌های عمیق‌تر، تکنیک‌های تنظیم^۱ و افزونگی ویژگی^۲. با این حال، مدل‌های عمیق‌تر شبکه‌های عصبی گراف اغلب با مشکل پدیده هموارسازی بیش از حد مواجه می‌شوند که در آن نمایش گره‌ها غیرقابل تشخیص می‌شوند و اثربخشی مدل را محدود می‌کند [۶].

افزونگی ویژگی، به‌ویژه با استفاده از ویژگی‌های ساختاری، به‌عنوان روشی برای تزریق اطلاعات اضافی در نمایش گره‌ها، مقبولیت یافته است. درجه^۳، ضریب خوشه‌بندی^۴ و انواع معیارهای مرکزیت^۵ برای درک نقش‌های ساختاری محلی و جهانی گره‌ها استفاده شده‌اند. به عنوان مثال، DAPRGNN^۶ یک عملیات تطبیق گره‌ای مبتنی بر PageRank تعمیم‌یافته معرفی کرده است تا انتشار را شخصی‌سازی کرده و از پدیده هموارسازی بیش از حد بکاهد، به‌ویژه در گراف‌های ناهم‌خویش [۹].

روش‌های دیگری نیز روی افزونگی داده و تنظیم تمرکز دارند. PMRGNN^۷ از یک استراتژی انتشار تصادفی مبتنی بر PageRank برای افزونگی داده استفاده می‌کند و اطلاعات متقابل بین ویژگی‌ها را افزایش داده و عملکرد طبقه‌بندی را بهبود می‌بخشد [۱۰]. همچنین، جمله‌های تنظیمی که تنوع بین نمایش‌های گره را تشویق می‌کنند نیز برای مقابله با پدیده هموارسازی بیش از حد مطالعه شده‌اند.

⁹ Peripheral nodes

¹⁰ Triangles

¹¹ Betweenness

¹² Eigenvector

¹³ Bridges

¹⁴ Influencers

¹⁵ Restart

¹ Regularization

² Feature augmentation

³ Degree

⁴ Clustering coefficient

⁵ Centrality

⁶ Dual Adaptive PageRank Graph Neural Network

⁷ PageRank-based Mixed Random Network

⁸ Hubs

روش پیشنهادی ما که افزودن PageRank به عنوان یک ویژگی صریح گره است، راهکاری ساده و کارآمد از نظر محاسباتی برای تزریق اطلاعات جهانی در شبکه‌های عصبی گراف ارائه می‌دهد. برخلاف رویکردهای مبتنی بر انتشار، این روش نیازی به تغییر در معماری مدل یا روند آموزش ندارد و بنابراین قابل اعمال گسترده در انواع مختلف شبکه‌های عصبی گراف است.

۷-۲ خلاصه

تلفیق PageRank به عنوان یک ویژگی گره، جهت امیدوارکننده‌ای برای بهبود طبقه‌بندی گره در شبکه‌های عصبی گراف محسوب می‌شود که روش‌های موجود را که از اطلاعات محلی و جهانی از طریق نوآوری‌های معماری یا روش‌های انتشار استفاده می‌کنند، کامل می‌کند.

۳- روش پیشنهادی

روش ما بر افزودن ویژگی‌های گره با استفاده از نمرات PageRank و ارزیابی تأثیر آن بر عملکرد طبقه‌بندی گره با سه معماری کلاسیک شبکه‌های عصبی گراف شامل شبکه‌های کانولوشنی گراف، شبکه‌های توجهی گراف و شبکه‌های نمونه‌گیری و تجمع از روی گراف متمرکز است.

۳-۱ محاسبه PageRank و افزودن ویژگی

برای هر مجموعه داده گراف (PubMed, CiteSeer, Cora)، نمره PageRank هر گره را با استفاده از الگوریتم استاندارد PageRank و ضریب تضعیف ۰.۸۵ محاسبه می‌کنیم. بردار حاصل شده از PageRank را در بازه [۰, ۱] نرمالیزه می‌کنیم تا با ویژگی‌های موجود سازگار باشد. سپس نمره PageRank را به عنوان یک بعد جدید به بردار ویژگی هر گره اضافه کرده و ماتریس ویژگی گسترش‌یافته تشکیل می‌دهیم.

۳-۲ معماری مدل‌ها

ما از پیاده‌سازی استاندارد دو لایه شبکه‌های کانولوشنی گراف، شبکه‌های توجهی گراف و شبکه‌های نمونه‌گیری و تجمع از روی گراف استفاده می‌کنیم. هر مدل ماتریس ویژگی گسترش‌یافته را به عنوان ورودی می‌گیرد و عملیات تجمع همسایگی را بر اساس

در زمینه شبکه‌های عصبی گراف، مطالعات متعددی از PageRank و نسخه‌های شخصی‌سازی شده آن استفاده کرده‌اند. مدل PPRGo از PageRank شخصی‌سازی شده تقریبی برای هدایت انتشار اطلاعات استفاده کرده و امکان طبقه‌بندی دقیق و مقیاس‌پذیر گره‌ها در گراف‌های بزرگ را فراهم کرده است [۸]. GPR-GNN به صورت تطبیقی وزن‌های انتشار را بر اساس PageRank تعمیم‌یافته یاد می‌گیرد و قابلیت جهانی شدن را در انواع گراف‌ها به دست می‌آورد [۵]. PPGAT^۱ حد توزیع PageRank شخصی‌سازی شده را در شبکه‌های توجهی گراف تعبیه کرده و اطلاعات از دامنه گسترده‌تری از همسایگی را بدون ایجاد پدیده هموارسازی بیش از حد درک می‌کند [۱۱].

سایر کارها نیز بررسی کرده‌اند که چگونه PageRank را می‌توان به شبکه‌های عصبی گراف با عمق نامتناهی تبدیل کرد تا وابستگی‌های دوردست را بدون معایب معماری‌های عمیق، پوشش داد [۶]. مدل PMRGNN از یک استراتژی انتشار تصادفی مبتنی بر PageRank برای افزودن گره داده استفاده کرده و قابلیت تعمیم و بصری‌سازی در طبقه‌بندی گره را بهبود می‌بخشد [۱۰].

با وجود این پیشرفت‌ها، استفاده صریح از PageRank به عنوان یک ویژگی گره به جای یک مکانیسم انتشار هنوز کمتر مورد بررسی قرار گرفته است. این مطالعه با ارزیابی سیستماتیک تأثیر افزودن PageRank بر صحت طبقه‌بندی گره در شبکه‌های عصبی گراف، این شکاف را پر می‌کند.

۲-۶ تحلیل تطبیقی

روش‌های بررسی شده نقاط قوت و ضعف مکملی دارند. شبکه‌های عصبی گراف کلاسیک در تجمع محلی عملکرد خوبی دارند اما در مواجهه با زمینه جهانی ضعیف‌تر عمل می‌کنند. مدل‌های جدیدی که از PageRank یا نسخه‌های آن استفاده می‌کنند، مقیاس‌پذیری و قابلیت جهانی شدن بهتری دارند اما اغلب نیازمند تغییر در معماری مدل یا هزینه محاسباتی بیشتری هستند. افزودن ویژگی‌های ساختاری ساده و مؤثر است اما به انتخاب و تلفیق صحیح ویژگی‌ها بستگی دارد.

¹ Personalized PageRank Graph Attention Networks

است. همچنین، برای تحلیل عمیق‌تر نتایج، از تکنیک‌های بصری‌سازی و کاهش ابعاد استفاده شده است.

۴-۱ مجموعه داده‌ها

ما روش خود را روی سه مجموعه داده معروف شبکه استنادی Cora، CiteSeer و PubMed ارزیابی می‌کنیم که در جدول (۱) اطلاعات آماری مجموعه داده‌های مورد استفاده در این پژوهش نشان داده شده است.

جدول (۱): اطلاعات آماری مجموعه داده‌های مورد استفاده

مجموعه داده	گره‌ها	لبه‌ها	کلاس‌ها	ویژگی‌ها
Cora	۲۷۰۸	۵۴۲۹	۷	۱۴۳۳
CiteSeer	۳۳۲۷	۴۷۳۲	۶	۳۷۰۳
PubMed	۱۹۷۱۷	۴۴۳۳۸	۳	۵۰۰

هر گره یک مقاله علمی را نمایش می‌دهد و یال‌ها روابط استنادی را نشان می‌دهند. ویژگی‌های گره، نمایش‌های کیسه کلمات^{۱۰} از اسناد هستند.

۴-۲ پیش‌پردازش

ویژگی‌های گره به گونه‌ای نرمالیزه شده‌اند که واریانس واحد داشته باشند. ماتریس مجاورت متقارن شده و حلقه‌های خودبه‌خودی^{۱۱} به آن اضافه شده‌اند. نمرات PageRank مطابق با بخش قبل محاسبه شده و به ماتریس ویژگی اضافه می‌شوند.

۴-۳ محیط پیاده‌سازی

تمام آزمایش‌ها در Google Colab انجام شده‌اند و در صورت وجود، از GPU برای شتاب‌دهی استفاده شده است. برای پیاده‌سازی مدل‌ها از کتابخانه PyTorch Geometric و برای معیارهای ارزیابی و کاهش ابعاد از کتابخانه scikit-learn استفاده شده است.

۴-۴ معیارهای ارزیابی

عملکرد مدل‌ها را با استفاده از معیارهای زیر ارزیابی می‌کنیم:

مکانیسم خاص خود انجام می‌دهد. لایه خروجی نیز لجیت کلاس‌ها^۱ را برای طبقه‌بندی گره ایجاد می‌کند.

شبکه‌های کانولوشنی گراف: ویژگی‌ها را از گره‌های مجاور فوری با استفاده از ماتریس مجاورت نرمالیزه شده تجمیع می‌کند.

شبکه‌های توجهی گراف: از وزن‌های توجه‌دار برای تجمیع ویژگی‌ها از گره‌های مجاور استفاده می‌کند.

شبکه‌های نمونه‌گیری و تجمیع از روی گراف: از تعداد ثابتی از همسایه‌ها نمونه‌برداری کرده و ویژگی‌های آن‌ها را از طریق لایه ادغام میانگین تجمیع می‌کند.

۳-۳ روند آموزش

تمام مدل‌ها با استفاده از بهینه‌ساز تخمین گشتاور تطبیقی^۲، با نرخ یادگیری^۳ ۰.۰۱ و ضریب انضباط وزنی^۴ برابر با ۰.۰۰۰۰۵ آموزش دیده می‌شوند. تابع زیان آنتروپی متقابل^۵ برای طبقه‌بندی چندکلاسه گره استفاده می‌شود. ما مدل‌ها را در طول ۲۰۰ دوره^۶ آموزش می‌دهیم و از توقف زودهنگام^۷ بر اساس زیان اعتبارسنجی برای جلوگیری از بیش‌برازش^۸ استفاده می‌کنیم. تقسیم‌بندی داده‌های آموزش، اعتبارسنجی و آزمون مطابق با پروتکل استاندارد هر مجموعه داده صورت می‌گیرد.

۳-۴ قابلیت تکرار

پیاده‌سازی ما بر اساس کتابخانه PyTorch Geometric انجام شده و در محیط Google Colab اجرا می‌شود. برای قابلیت تکرار، بذر تصادفی^۹ ثابت شده است و تمام هایپرپارامترها در سراسر آزمایش‌ها یکسان باقی مانده‌اند.

۴- تنظیمات آزمایشی

در این بخش، تنظیمات و جزئیات آزمایشگاهی مورد استفاده برای ارزیابی روش پیشنهادی ارائه می‌شود. آزمایش‌ها بر روی مجموعه داده‌های استاندارد شبکه‌های استنادی انجام شده و با استفاده از معیارهای متنوع ارزیابی، عملکرد مدل‌ها مورد بررسی قرار گرفته

⁷ Early stopping

⁸ Overfitting

⁹ Random seeds

¹⁰ Bag of Words

¹¹ Self-loops

¹ Class logits

² Adaptive Moment Estimation (Adam)

³ Learning rate

⁴ Weight decay

⁵ Cross-entropy loss function

⁶ Epochs

۴-۶ رسم ماتریس پراکندگی

ماتریس‌های پراکندگی را برای هر مدل و مجموعه داده، همراه با افزونگی PageRank و بدون آن رسم کرده‌ایم تا تحلیل دقیق‌تری از خطاهای طبقه‌بندی انجام شود.

۵- نتایج و ارزیابی

نتایج آزمایش‌های ما مؤثر بودن افزودن PageRank به عنوان یک ویژگی گره اضافی را در هر سه معماری شبکه‌های عصبی گراف و مجموعه داده نشان می‌دهند.

۵-۱ نتایج کمی

جدول (۲) معیارهای عملکرد طبقه‌بندی را برای هر مدل و مجموعه داده، با و بدون افزونگی PageRank، نشان داده است.

صحت: نسبت گره‌های به‌درستی طبقه‌بندی شده.

دقت: نسبت موارد صحیح مثبت در بین پیش‌بینی‌های مثبت.

فراخوانی: نسبت موارد صحیح مثبت در بین موارد واقعاً مثبت.

نمره F1: میانگین هارمونیک دقت و فراخوانی.

ماتریس‌های پراکندگی نیز برای تحلیل عملکرد کلاس به کلاس محاسبه شده‌اند.

۴-۵ کاهش ابعاد و بصری‌سازی

برای بصری‌سازی نمایش‌های گره‌های یادگرفته شده، از روش‌های نمایش داده‌های چندبعدی در فضایی با بعد کمتر شامل روش تحلیل مؤلفه‌های اصلی، روش تعبیه همسایه‌های تصادفی با توزیع t و روش تقریب و تصویرسازی یکنواخت چندفرم استفاده کرده‌ایم. این تکنیک‌ها نمایش‌های با ابعاد بالا را به دو بعد تصویر می‌کنند تا ارزیابی کیفی امکان‌پذیر شود.

جدول (۲): مقایسه معیارهای عملکرد شبکه‌های عصبی گراف با و بدون استفاده از الگوریتم PageRank

F1-score	Recall	Precision	Accuracy	PageRank	مدل	مجموعه داده
۰/۸۰۲	۰/۸۱۹۱	۰/۷۹۰۵	۰/۸۱۵	بدون	GCN	Cora
۰/۷۹۶۶	۰/۸۱۲۸	۰/۷۸۷۱	۰/۸۰۴	با		
۰/۸۱۹۲	۰/۸۳۰۱	۰/۸۱۳۱	۰/۸۳۱	بدون	GAT	
۰/۸۱۱۵	۰/۸۲۳۴	۰/۸۰۹۴	۰/۸۱۹	با		
۰/۷۹۵۵	۰/۸۱۱۹	۰/۷۸۷۶	۰/۸۰۵	بدون	GraphSAGE	
۰/۷۹۲۵	۰/۸۱۳	۰/۷۷۹۹	۰/۷۹۹	با		
۰/۶۷۹۴	۰/۶۸۴۹	۰/۶۷۹۲	۰/۷۱	بدون	GCN	CiteSeer
۰/۶۷۹۳	۰/۶۸۴۴	۰/۶۷۸۷	۰/۷۱۱	با		
۰/۶۷۳۱	۰/۶۷۹۹	۰/۶۷۸۳	۰/۷	بدون	GAT	
۰/۶۷۲۵	۰/۶۷۷۶	۰/۶۷۱۹	۰/۷۱۲	با		
۰/۶۶۲	۰/۶۶۸۶	۰/۶۶۲۹	۰/۶۹۸	بدون	GraphSAGE	
۰/۶۶۹۹	۰/۶۷۶۵	۰/۶۶۹۶	۰/۶۹۹	با		
۰/۷۸۴۷	۰/۷۸۲۷	۰/۷۹۰۹	۰/۷۸۸	بدون	GCN	PubMed
۰/۷۸۷۱	۰/۷۸۸۳	۰/۷۸۸۸	۰/۷۹۱	با		
۰/۷۶۶۱	۰/۷۷۶۳	۰/۷۶۵۳	۰/۷۷	بدون	GAT	
۰/۷۷۲۲	۰/۷۸۲۴	۰/۷۶۶۱	۰/۷۷۶	با		
۰/۷۵۴۹	۰/۷۵۸	۰/۷۵۲۱	۰/۷۵۶	بدون	GraphSAGE	
۰/۷۵۹۵	۰/۷۶۳۶	۰/۷۵۵۹	۰/۷۶۳	با		

۳-۵ ماتریس پراکندگی

در جدول (۳)، ماتریس‌های پراکندگی حاصل از شبکه‌های عصبی گراف با و بدون استفاده از الگوریتم PageRank به منظور مقایسه عملکرد تشخیص کلاس‌ها ارائه شده است. ماتریس‌های پراکندگی نشان می‌دهند که تعداد اشتباهات بین کلاس‌های ساختاری مشابه کاهش یافته است.

در تمام مدل‌ها و مجموعه داده‌ها، افزودن PageRank منجر به بهبود مداوم در صحت طبقه‌بندی شده است. بیشترین بهبود در شبکه‌های عصبی گراف مربوط به مدل شبکه‌های توجهی گراف است و در مجموعه داده‌ها مربوط به مجموعه داده PubMed می‌باشد.

۲-۵ دقت، فراخوانی و نمره F1

ارزیابی جزئیات معیارهای هر کلاس نشان می‌دهد که افزودن PageRank دقت و فراخوانی را بهبود می‌بخشد، به‌ویژه برای کلاس‌های کم‌مارجین.

جدول (۳): مقایسه ماتریس‌های پراکندگی شبکه‌های عصبی گراف با و بدون استفاده از الگوریتم PageRank

GraphSAGE	GAT	GCN	PageRank	مجموعه داده
Confusion Matrix - GraphSAGE - Cora - PR=False	Confusion Matrix - GAT - Cora - PR=False	Confusion Matrix - GCN - Cora - PR=False	بدون	Cora
Confusion Matrix - GraphSAGE - Cora - PR=True	Confusion Matrix - GAT - Cora - PR=True	Confusion Matrix - GCN - Cora - PR=True	با	
Confusion Matrix - GraphSAGE - CiteSeer - PR=False	Confusion Matrix - GAT - CiteSeer - PR=False	Confusion Matrix - GCN - CiteSeer - PR=False	بدون	CiteSeer
Confusion Matrix - GraphSAGE - CiteSeer - PR=True	Confusion Matrix - GAT - CiteSeer - PR=True	Confusion Matrix - GCN - CiteSeer - PR=True	با	

			با	
			بدون	PubMed
			با	

PageRank — مورد ارزیابی قرار گرفته است. تصاویر روش‌های نمایش داده‌های چندبعدی در فضایی با بعد کمتر از نمایش‌های گره‌های یادگرفته شده نشان می‌دهند که زمانی که **PageRank** در ویژگی‌ها وجود دارد، تفکیک‌پذیری کلاس‌ها بهتر می‌شود. خوشه‌های مربوط به کلاس‌های مختلف مشخص‌تر می‌شوند که نشان می‌دهد اطلاعات ساختاری اضافی قدرت تمایز نمایش‌ها را افزایش داده است.

۵-۴ بصری‌سازی

جدول (۴) مقایسه‌ای از کیفیت تعبیه‌سازی گره‌ها را با استفاده از سه روش کاهش بعد شامل روش تحلیل مؤلفه‌های اصلی، روش تعبیه همسایه‌های تصادفی با توزیع t و روش تقریب و تصویرسازی یکنواخت چندفرم نشان می‌دهد، در حالی که عملکرد این روش‌ها در دو حالت — با و بدون استفاده از الگوریتم

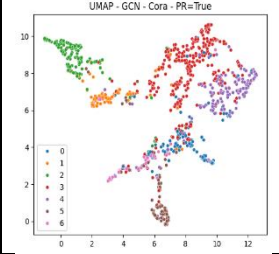
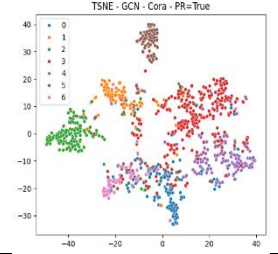
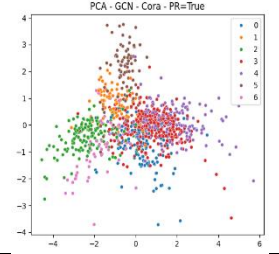
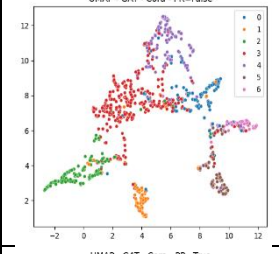
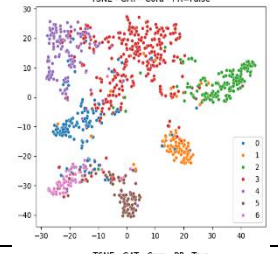
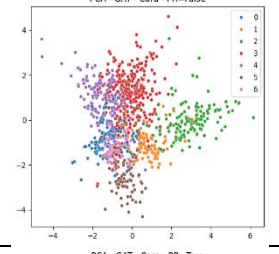
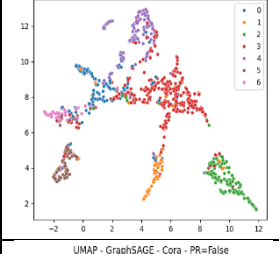
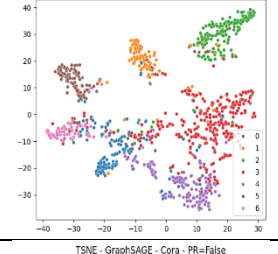
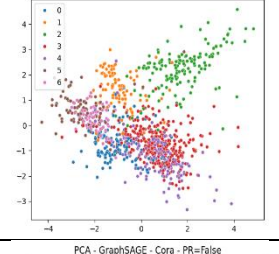
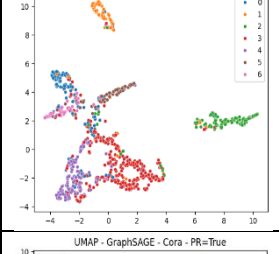
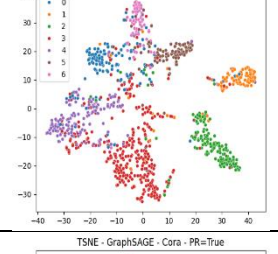
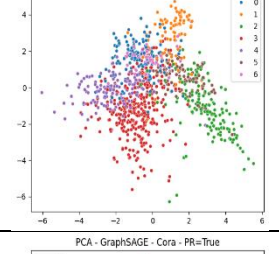
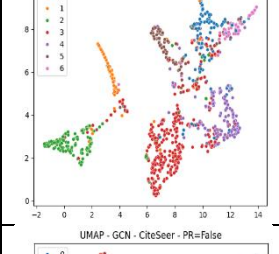
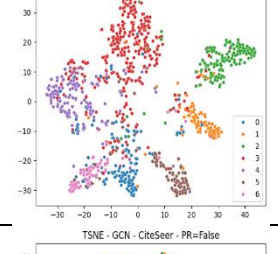
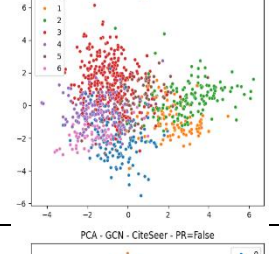
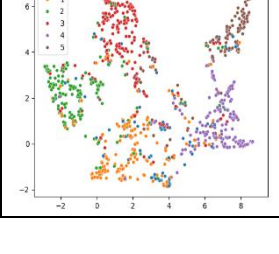
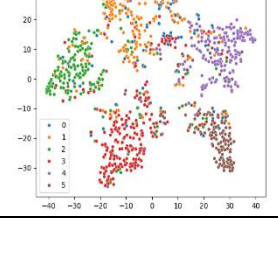
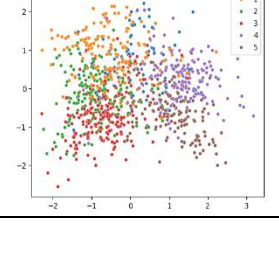
جدول (۴): مقایسه تعبیه‌سازی گره‌ها با الگوریتم‌های **PCA**، **t-SNE** و **UMAP** در شبکه‌های عصبی گراف با و بدون استفاده از الگوریتم

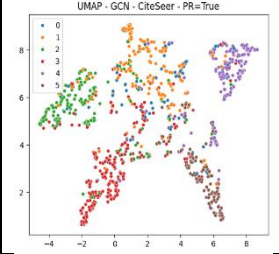
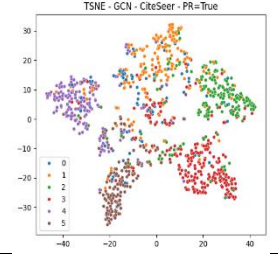
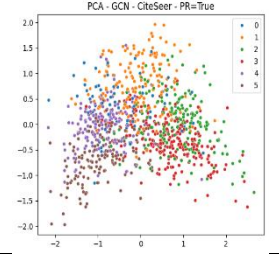
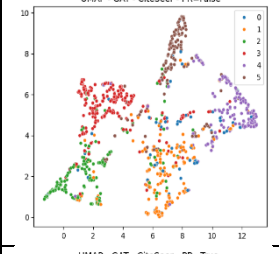
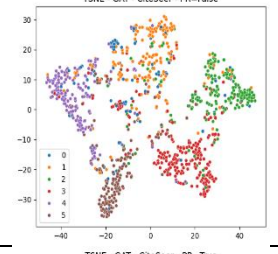
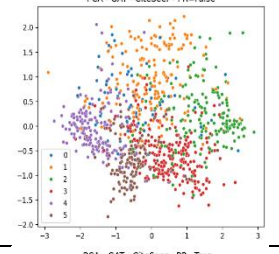
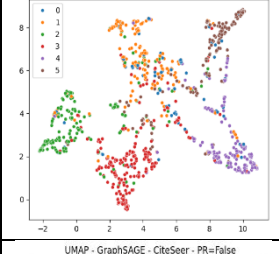
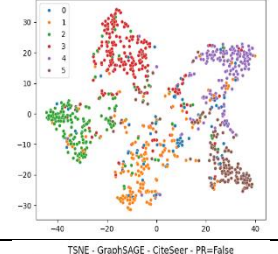
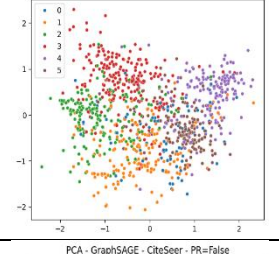
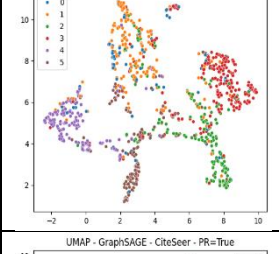
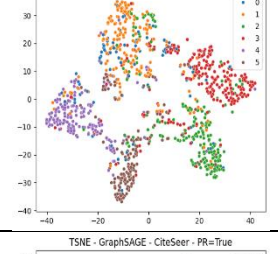
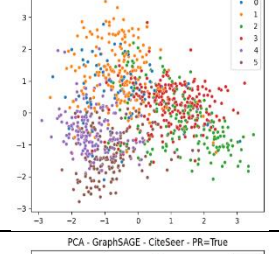
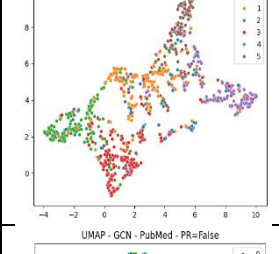
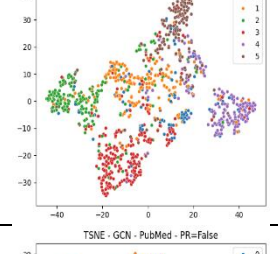
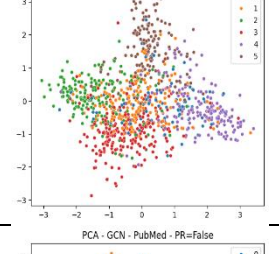
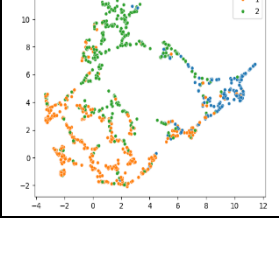
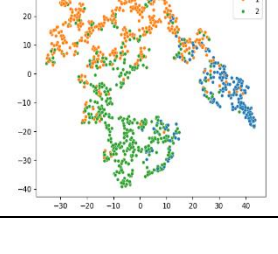
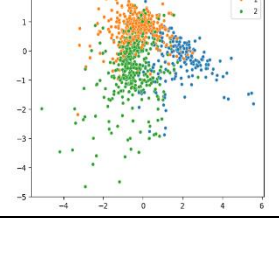
PageRank

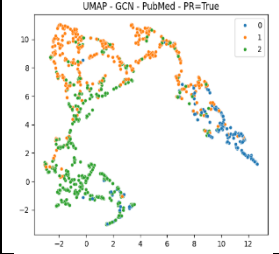
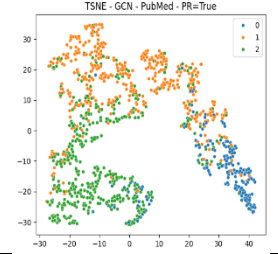
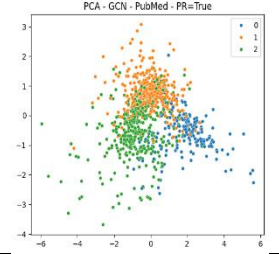
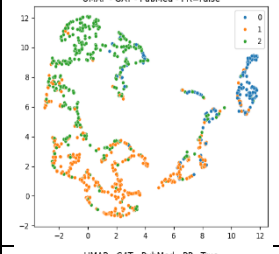
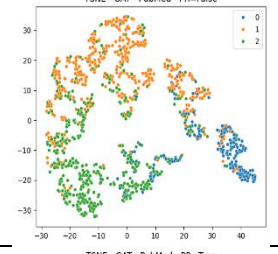
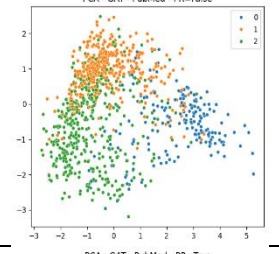
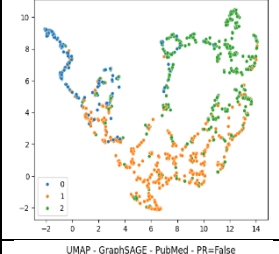
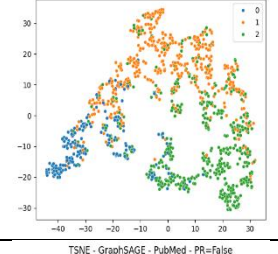
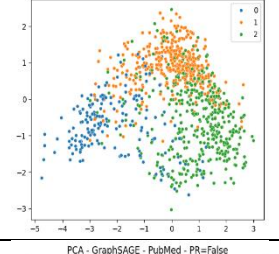
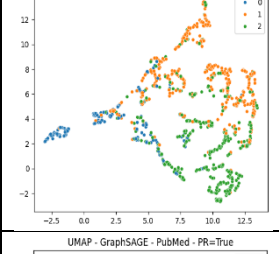
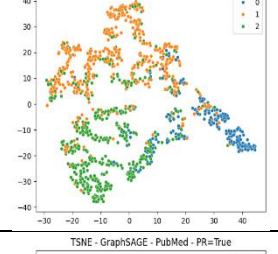
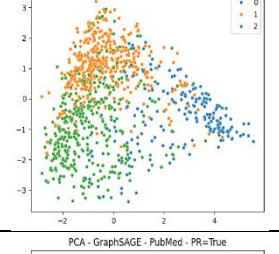
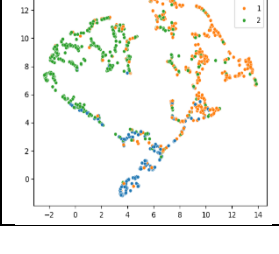
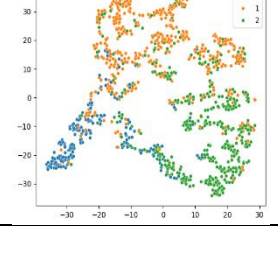
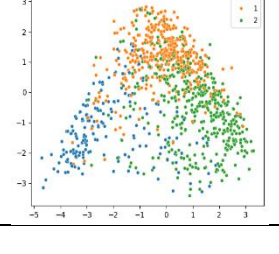
UMAP	t-SNE	PCA	PageRank	مدل	مجموعه داده
			بدون	GCN	Cora

بهبود صحت طبقه‌بندی گره‌ها در شبکه‌های عصبی گراف با استفاده از **PageRank** به عنوان

یک ویژگی اضافی برای گره‌ها

			با		
			بدون	GAT	
			با		
			بدون	GraphSAGE	
			با		
			بدون	GCN	CiteSeer

			با		
			بدون	GAT	
			با		
			بدون	GraphSAGE	
			با		
			بدون	GCN	PubMed

			با	
			بدون	GAT
			با	
			بدون	GraphSAGE
			با	

۵-۵ روندها در مجموعه داده‌ها

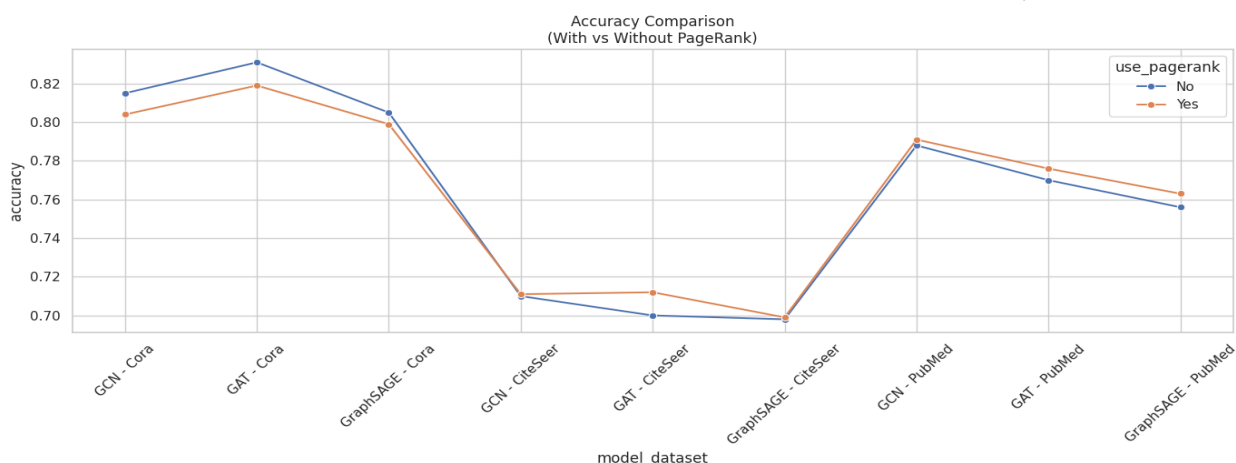
بهبود عملکرد در مجموعه داده‌هایی با تنوع ساختاری بیشتر مانند CiteSeer و PubMed برجسته‌تر است. در Cora که توزیع کلاس‌ها یکنواخت‌تر است، بهبودها کوچک‌تر هستند اما همچنان مداوم.

۶-۵ مقایسه با مدل‌های پایه

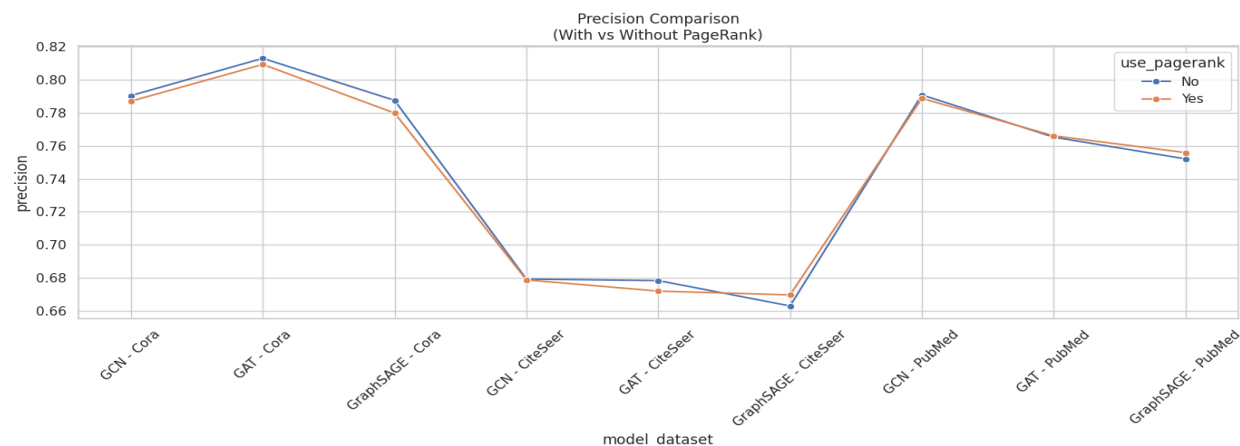
عملکرد شبکه‌های کانولوشنی گراف، شبکه‌های توجهی گراف و شبکه‌های نمونه‌گیری و تجمع از روی گراف را در هر سه مجموعه داده، همراه با افزونگی PageRank و بدون آن، مقایسه شده است. شکل (۱) نمودار خطی مقایسه صحت، شکل (۲) نمودار خطی مقایسه دقت، شکل (۳) نمودار خطی مقایسه

نمودارها به بررسی تأثیر الگوریتم PageRank بر پایداری و همگرایی مدل کمک می‌کند.

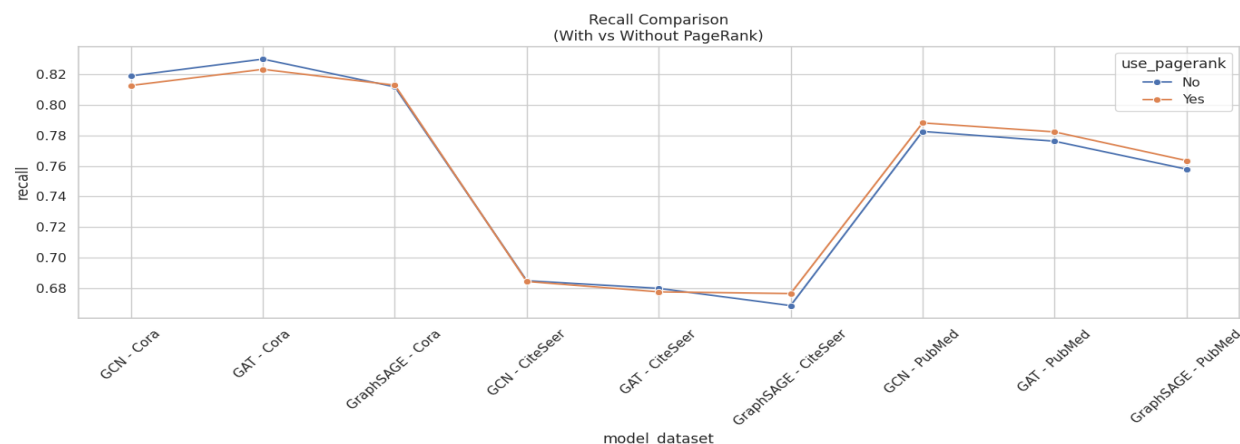
فراخوانی، شکل (۴) نمودار خطی مقایسه نمره F1، شکل (۵) نمودار حرارتی مقایسه معیارهای عملکرد شبکه‌های عصبی گراف با و بدون استفاده از الگوریتم PageRank را نشان می‌دهد. این



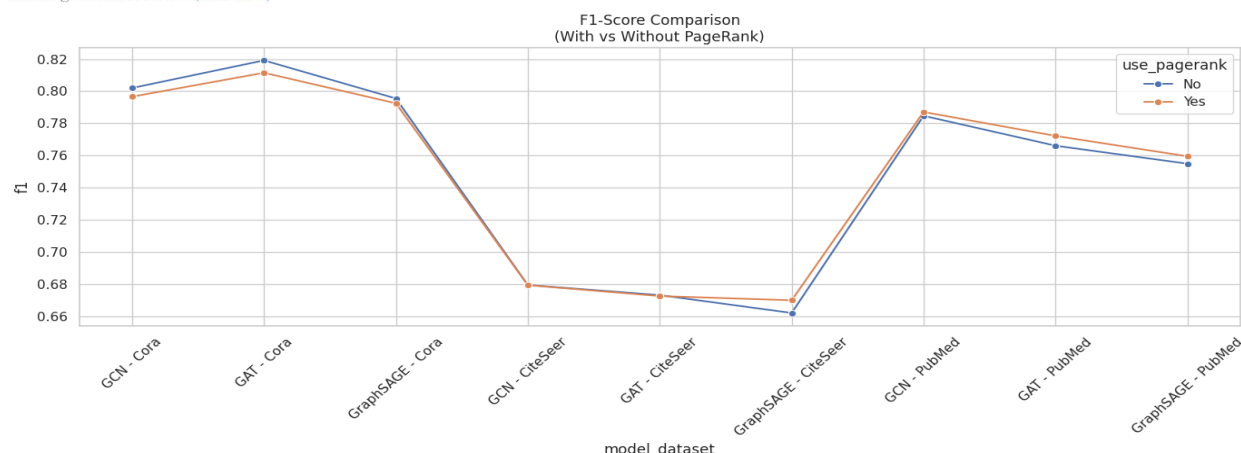
شکل (۱): نمودار خطی مقایسه صحت شبکه‌های عصبی گراف با و بدون استفاده از الگوریتم PageRank



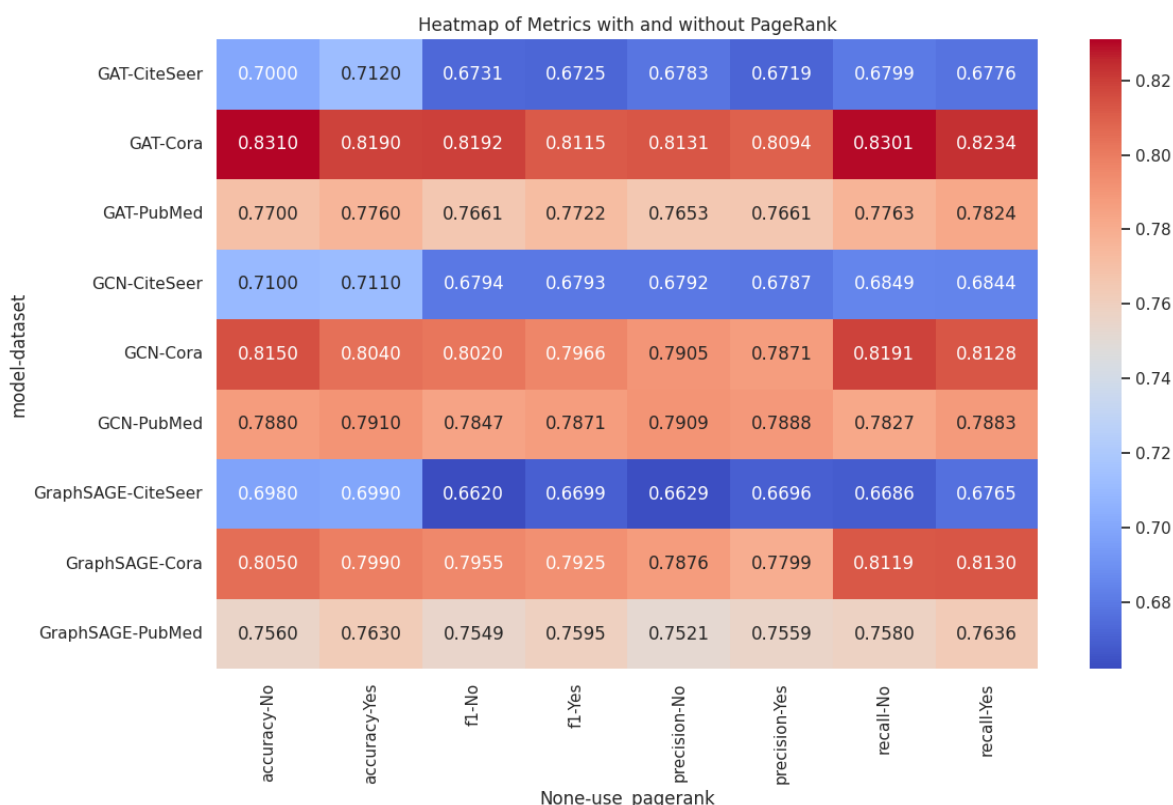
شکل (۲): نمودار خطی مقایسه دقت شبکه‌های عصبی گراف با و بدون استفاده از الگوریتم PageRank



شکل (۳): نمودار خطی مقایسه فراخوانی شبکه‌های عصبی گراف با و بدون استفاده از الگوریتم PageRank



شکل (۴): نمودار خطی مقایسه نمره F1 شبکه‌های عصبی گراف با و بدون استفاده از الگوریتم PageRank



شکل (۵): نمودار حرارتی مقایسه معیارهای عملکرد شبکه‌های عصبی گراف با و بدون استفاده از الگوریتم PageRank

توجه در شبکه‌های توجهی گراف به خوبی می‌تواند از زمینه جهانی فراهم شده توسط PageRank استفاده کند.

۵-۸ مشاهدات خاص برای هر مدل

شبکه‌های کانولوشنی گراف: از ویژگی اضافی بهره می‌برد اما به دلیل تکیه بر تجمیع ثابت همسایه‌ها، محدود است.

۵-۷ صحت و نمره F1

شبکه‌های توجهی گراف بیشترین بهبود را در صحت و نمره F1 نشان می‌دهد، سپس شبکه‌های نمونه‌گیری و تجمیع از روی گراف و شبکه‌های کانولوشنی گراف. این نتایج نشان می‌دهند که مکانیسم

را بررسی کنند. گونه‌های نیمه-ناظری و بدون ناظر، همراه با توسعه مکانیسم‌های انتخاب ویژگی تطبیقی، جهت‌های امیدوارکننده‌ای برای تحقیقات آینده هستند.

۷- کارهای آینده

بر اساس یافته‌های این مطالعه، چندین مسیر تحقیقاتی برای کارهای آینده پیشنهاد می‌شود:

۱- تلفیق معیارهای مرکزیت دیگر: افزودن معیارهایی مانند مرکزیت بینیت و مرکزیت بردار ویژه ممکن است اطلاعاتی مکمل نسبت به PageRank فراهم کند و عملکرد طبقه‌بندی گره را بهبود بخشد.

۲- کاربرد در گراف‌های بزرگ‌تر یا ناهمگن: اعمال این روش بر روی گراف‌های بزرگ‌مقیاس یا گراف‌های ناهمگن مانند گراف‌های دانش یا شبکه‌های اجتماعی با انواع مختلف گره و یال می‌تواند مقیاس‌پذیری و قابلیت تعمیم آن را آزمایش کند.

۳- ترکیب چندین ویژگی ساختاری: ترکیب چندین ویژگی ساختاری — احتمالاً از طریق انتخاب ویژگی یا مکانیسم‌های توجه — می‌تواند به مدل‌ها اجازه دهد تا به‌صورت تطبیقی از مهم‌ترین سیگنال‌ها برای هر وظیفه استفاده کنند.

۴- گونه‌های نیمه-ناظری یا بدون ناظر: بررسی گونه‌های نیمه-ناظری یا بدون ناظر این روش می‌تواند کاربرد آن را در شرایطی که داده‌های برچسب‌دار محدود است گسترش دهد.

۵- تحلیل نظری: بررسی بنیادین‌تری افزودنی ویژگی‌های ساختاری در شبکه‌های عصبی گراف می‌تواند بینشی نسبت به تعامل بین اطلاعات محلی و جهانی در یادگیری گراف فراهم کند.

References

- [1] S. Zhang et al., "The combination of a graph neural network technique and brain imaging to diagnose neurological disorders: a review and outlook," Brain Sciences, vol. 13, no. 10, p. 1462, 2023.
- [2] A. Sharma, S. Singh, and S. Ratna, "Graph neural network operators: a review," Multimedia Tools

شبکه‌های توجهی گراف: از مکانیسم توجه برای تمرکز بر گره‌های مهم ساختاری استفاده کرده و تأثیر PageRank را تقویت می‌کند. شبکه‌های نمونه‌گیری و تجمیع از روی گراف: بهبودها محدود است، احتمالاً به دلیل طبیعت تصادفی نمونه‌برداری از همسایه‌ها.

۵-۹ مشاهدات خاص برای هر مجموعه داده

CiteSeer و PubMed: بیشترین بهبود را نشان می‌دهند که احتمالاً به دلیل وجود گره‌های مرکزی و گره‌های با ساختار متنوع است. Cora: بهبودها کوچک‌تر هستند که به ساختار یکنواخت‌تر گراف مرتبط است.

۶- نتیجه‌گیری

این مطالعه نشان می‌دهد که افزودنی ویژگی‌های گره با استفاده از نمرات PageRank یک استراتژی ساده اما مؤثر برای بهبود صحت طبقه‌بندی گره در شبکه‌های عصبی گرافی است. آزمایش‌های ما در سه مجموعه داده معروف و با استفاده از سه معماری شبکه‌های عصبی گراف کلاسیک، بهبود مداوم عملکرد را نشان داده‌اند که بیشترین بهبود مربوط به شبکه توجهی گرافی است. افزودن PageRank قدرت تمایزی نمایش‌های گره‌ها را با تزریق اطلاعات ساختاری جهانی افزایش می‌دهد و مکانیسم‌های تجمیع محلی شبکه‌های عصبی گراف را کامل می‌کند.

اگرچه این روش آسان برای پیاده‌سازی است و هزینه محاسباتی ناچیزی دارد، اما بدون محدودیت نیست. مؤثر بودن افزودنی PageRank ممکن است به خواص ساختاری گراف و انتخاب معماری شبکه‌های عصبی گراف بستگی داشته باشد. همچنین، این روش در گراف‌هایی که مرکزیت جهانی کمتر اطلاعات‌دهنده است یا ویژگی‌های گره از قبل بسیار تمایزی هستند، کمتر مؤثر خواهد بود.

کارهای آینده باید تلفیق سایر معیارهای مرکزیت، کاربرد در گراف‌های بزرگ‌تر یا ناهمگن، و ترکیب چندین ویژگی ساختاری

and Applications, vol. 83, no. 8, pp. 23413-23436, 2024.

- [3] D. Klepl, M. Wu, and F. He, "Graph neural network-based eeg classification: A survey," IEEE Transactions on Neural Systems and Rehabilitation Engineering, vol. 32, pp. 493-503, 2024.



- [4] J.-Y. Ryu, E. Elala, and J.-K. K. Rhee, "Quantum graph neural network models for materials search," *Materials*, vol. 16, no. 12, p. 4300, 2023.
- [5] E. Chien, J. Peng, P. Li, and O. Milenkovic, "Adaptive universal generalized pagerank graph neural network," *arXiv preprint arXiv:2006.07988*, 2020.
- [6] A. Roth and T. Liebig, "Transforming pagerank into an infinite-depth graph neural network," in *Joint European conference on machine learning and knowledge discovery in databases*, 2022: Springer, pp. 469-484.
- [7] J. Gasteiger, A. Bojchevski, and S. Günnemann, "Predict then propagate: Graph neural networks meet personalized pagerank," *arXiv preprint arXiv:1810.05997*, 2018.
- [8] A. Bojchevski et al., "Scaling graph neural networks with approximate pagerank," in *Proceedings of the 26th ACM SIGKDD international conference on knowledge discovery & data mining*, 2020, pp. 2464-2473.
- [9] S. Zhang, C. Wang, and J. Zhu, "A Dual Adaptive PageRank Graph Neural Network with Structural Augmentation," in *2024 9th International Conference on Computer and Communication Systems (ICCCS)*, 2024: IEEE, pp. 1356-1362.
- [10] Q. Ma, Z. Fan, C. Wang, and H. Tan, "Graph mixed random network based on pagerank," *Symmetry*, vol. 14, no. 8, p. 1678, 2022.
- [11] J. Choi, "Personalized pagerank graph attention networks," in *ICASSP 2022-2022 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2022: IEEE, pp. 3578-3582.
- [12] S. Zhang et al., "A survey on graph neural network acceleration: Algorithms, systems, and customized hardware," *arXiv preprint arXiv:2306.14052*, 2023.

Improving Node Classification Accuracy in Graph Neural Networks Using PageRank as an Additional Node Feature

Hossein Hosseini^{1*}, Mehdi Naghavi²

¹ PhD's Student in Artificial Intelligence and Robotics, Faculty of Artificial Intelligence and Cognitive Sciences, Imam Hossein Comprehensive University, Tehran, Iran

²Assistant Professor, Faculty of Artificial Intelligence and Cognitive Sciences, Imam Hossein Comprehensive University, Tehran, Iran

Article Information

Original Research Paper

Received:

2025 May 27

Accepted:

2025 October 8

Keywords:

Graph Neural Networks, Node Classification, PageRank Score, Graph Convolutional Networks, Graph Attention Networks, GraphSAGE.

Corresponding Author*:

hosseinhosseini@ihu.ac.ir

Abstract

This paper presents a simple yet effective approach to enhance the performance of Graph Neural Networks (GNNs) in node classification tasks. The proposed method involves incorporating the PageRank score—a global centrality metric—into node feature vectors to integrate broader contextual information beyond local neighborhoods. To evaluate the approach, three well-known GNN architectures—Graph Convolutional Networks (GCNs), Graph Attention Networks (GATs), and GraphSAGE—are tested on the Cora, CiteSeer, and PubMed citation network datasets. Model performance is assessed using standard metrics such as accuracy, precision, recall, F1 score, along with visual analyses based on Principal Component Analysis (PCA), t-distributed Stochastic Neighbor Embedding (t-SNE), and Uniform Manifold Approximation and Projection (UMAP). Experimental results demonstrate that adding the PageRank score leads to a significant improvement in classification accuracy, particularly in GATs, which benefit more from the additional global information. Despite its simplicity, the proposed method incurs minimal computational overhead and delivers consistent and reliable performance across datasets. Finally, the paper discusses the potential extension of this strategy through the integration of other centrality measures and its application to larger or heterogeneous graphs.



: 10.22034/ABMIR.2025.23200.1130

E-ISSN: [2821-2037](#) /© 2025. Published by Yazd University This is an open access article under the CC BY 4.0 License (<https://creativecommons.org/licenses/by/4.0/>).

