



دانشگاه صنعتی امیر کبیر
(پلی تکنیک تهران)

دانشکده مهندسی کامپیوتر

تحلیل شبکه‌های پیچیده

پروژه نهایی

استاد درس: دکتر مصطفی حقیر چهرقانی

بهمن ۱۴۰۱

مهلت تحویل تمرین :

۱۱:۵۹ ب.ظ

۱۴۰۱/۱۱/۲۳

سوال اول

الف) مجموعه دادگان DBLP را در نظر بگیرید و با استفاده از کتابخانه [PyTorch Geometric \(PyG\)](#) تعداد گراف، تعداد گره، تعداد یال، تعداد ویژگی هر گره، تعداد کلاس‌های موجود و دیگر ویژگی‌های این مجموعه را بررسی و گزارش کنید.

ب) در این بخش لازم است هر یک از شبکه‌های زیر را با شرایط خواسته شده پیاده سازی کرده و بر روی این مجموعه دادگان بیازمایید.^۱

▪ شبکه‌های پیچشی گرافی (GCN):

- ۱- با استفاده از یک یا چند لایه پیچشی ویژگی‌های مهم در این مجموعه دادگان را استخراج کنید.
- ۲- تعداد ویژگی‌های مربوط به هر لایه پنهان در مدل را با آزمون و خطا بدست آورید. (آزمون و گزارش حداقل سه مقدار الزامی است)
- ۳- با استفاده از یک لایه کاملاً متصل پیش‌بینی نهایی را انجام دهید.
- ۴- برای تعداد لایه‌های پیچشی و تعداد لایه‌های کاملاً متصل نهایی به ترتیب مقادیر یک تا سه لایه را آزمایش کرده و بهترین ساختار را گزارش دهید. برداشت خود را از میزان تاثیر هر کدام از ویژگی‌ها، تعداد لایه پیچشی، تعداد ویژگی‌های هر لایه پنهان و تعداد لایه کاملاً متصل بیان کنید.

▪ شبکه توجه گرافی (GAT):

- ۱- برای ارزیابی این مدل تعداد سرهای مکانیزم توجه را از ۱ تا ۱۰ تغییر دهید و نمودار دقت مدل بر حسب تعداد سرهای توجه را رسم کنید.

^۱ در پیاده سازی هر شبکه با تغییر ویژگی‌های شبکه سعی کنید بهترین مدل را بدست آورید.

۲- پس از یافتن تعداد سر بهینه در این مدل (با در نظر گرفتن نتیجه حاصل شده در قسمت قبل) با توجه به اهمیت متفاوت زیرفضاهای متناسب با هرسر برای هریک از آنها یک ضریب قابل آموزش را در نظر بگیرید.

▪ GAT-V2.

۱- در ارزیابی این مدل لازم است قسمت مشخص شده در فرمول این شبکه را با عملگرهای Min, Max, Hadamard جایگزین کرده و مقایسه کنید.

$$e(h_i, h_j) = a^T \text{LeakyReLU}(W \cdot \underline{[h_i || h_j]})$$

ج) در انتها از میان تمامی مدل‌های مورد آزمون قرار گرفته، بهترین مدل را گزارش کرده و دلایل برتری آن را توجیه کنید.

سوال دوم^۲

در این سوال لازم است پس از بررسی اجمالی مقالات زیر، یکی از آن‌ها را به همراه یک مجموعه دادگان **صرفاً متناظر با همان مقاله** (که در روبروی نام مقاله در جدول زیر آورده شده است)، انتخاب کرده و موارد خواسته شده را انجام دهید.

Paper Name	Appropriate Dataset
Self-Attention Graph Pooling	MUTAG, IMDB, ENZYMS, PTC
Simplifying Graph Convolutional Networks	Amazon Photos, Amazon Computers, DBLP
Spectral Clustering with Graph Neural Networks for Graph Pooling	MUTAG, ENZYMS, Amazon Computers, DBLP

الف) مجموعه دادگان و مقاله‌ای که انتخاب کرده‌اید را شرح دهید. بدین منظور لطفاً گزارشی در اندازه ۱ الی ۲ صفحه را تهیه کنید.

ب) مدل موجود در مقاله انتخاب شده را خودتان پیاده سازی کنید^۳. مدل را تنظیم کرده و عملکرد آن بر روی مجموعه دادگان انتخاب شده را با نتایج ارائه شده در مقاله مقایسه و بررسی کنید.

^۲ در حل این سوال می‌توانید از پیاده سازی مدل مقالات که در اینترنت موجود هستند کمک بگیرید.

^۳ بدین منظور می‌توانید از PyG, GraphGym, DeepSNAP استفاده کنید.

ج) در این بخش در صورت امکان یک ایده برای بهبود عملکرد مدل ارائه دهید. این ایده می‌تواند نوآوری خود شما باشد یا با خواندن مقالات دیگر ترکیب نقاط قوت چندین مدل باشد. ایده موردنظر را اجرا کنید و مقدار بهبود را توجیه و گزارش کنید. (این بخش دارای امتیاز اضافی می‌باشد)

نکاتی که در گزارش نهایی شما باید گنجانده شود:

- ۱- کدام مقاله و کدام مدل یا روش را اجرا کردید؟ مدل را توصیف کنید (ایده کلیدی، نوآوری، انگیزه، ارتباط مدل با کار پایه).
- ۲- کدام مجموعه داده را انتخاب کردید و چرا؟
- ۳- عملکرد روش موجود در مقایسه با عملکرد مدل پایه.
- ۴- ایده و انگیزه خود را برای بهبود بیشتر مدل‌ها یا روش‌های موجود در مقاله شرح دهید.

نکات مربوط به تحویل تمرین

- برای قسمت‌های مرتبط با پیاده‌سازی زبان مورد قبول، زبان پایتون است.
- نظم و خوانایی در نوشتن گزارش و کدها از اهمیت بالایی برخوردار است. کدهای خود را تا حد امکان کامنت‌گذاری کنید. خروجی کدها و نتایج سوالات را درون گزارش بنویسید و از توضیح اضافی کد و موارد دیگر خودداری فرمایید (کد بدون گزارش ارزشی ندارد).
- در پیاده‌سازی بخش‌های مختلف امکان استفاده از کتابخانه‌های آماده مربوط به کار با گراف را ندارید؛ مگر در مواردی که در صورت سوال ذکر شده باشد.
- برای هر سوال یک پوشه جداگانه در نظر گرفته، کد و مواردی از قبیل خروجی برنامه و نمودارها را در آن ذخیره نمایید. این پوشه‌ها را به همراه یک فایل report.pdf برای گزارش و توضیح سوالات، درون یک فایل فشرده شده با فرمت FinalProject_StdNumber.zip که StdNumber شماره دانشجویی شماست، در سامانه بارگذاری کنید.
- مطابق قوانین دانشگاه هر گونه کپی‌برداری ممنوع است و در صورت مشاهده، نمره هر دو طرف صفر داده می‌شود. اما استفاده از توضیحات موجود در اینترنت به منظور یادگیری الزاما با ذکر منبع بلامانع است.
- در صورت وجود هر گونه سوال یا ابهامی با ایمیل درس در تماس باشید:

cna.assistant@gmail.com