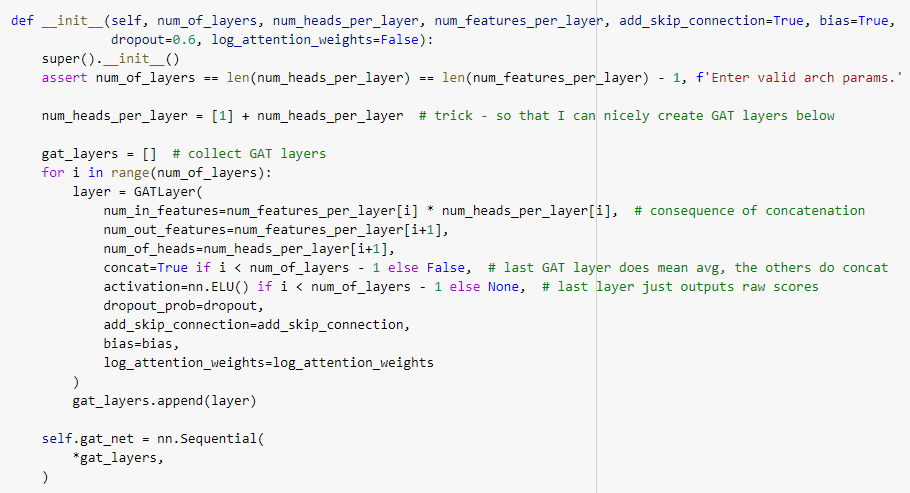
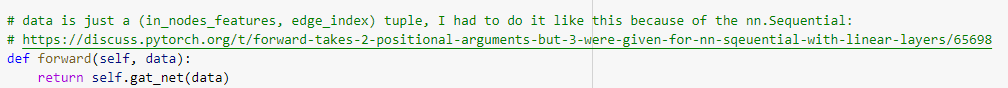


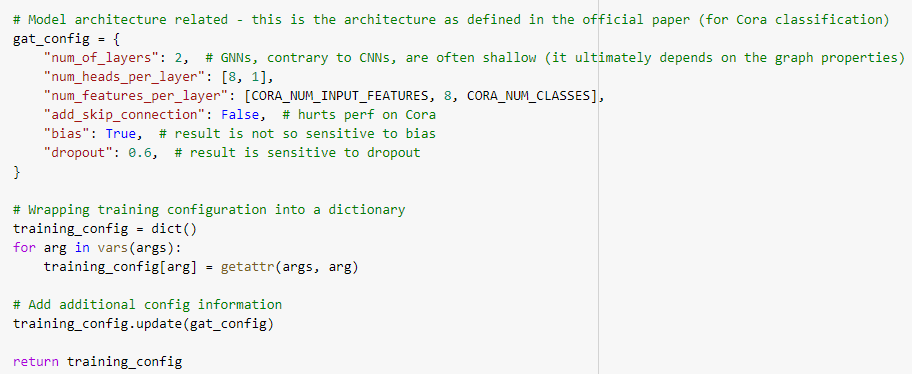
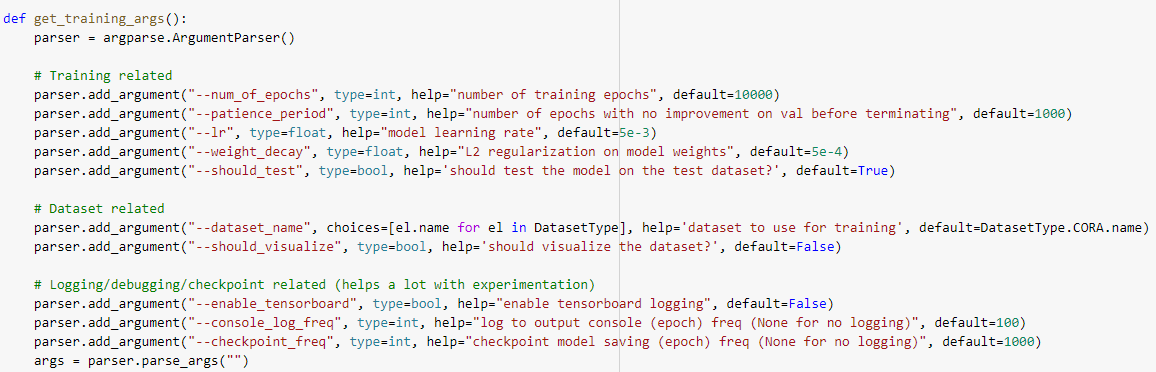
برای پیاده‌سازی مدل با استفاده از لایه‌های GAT کلاس GAT را تعریف کردیم.



در تعریف این مدل علاوه بر پارامترهای لایه‌های GAT پارامتری به نام num\_of\_layers برای مشخص کردن تعداد لایه‌های GAT تعریف شده است. تعداد لایه‌های GAT به مشخصات گراف بستگی دارد و هر مقدار دلخواهی نمی‌توان به آن اختصاص داد. تعداد ویژگی‌های ورودی هر لایه برابر حاصل‌ضرب تعداد خروجی لایه قبل و تعداد هدهای آن لایه است. در تمام لایه‌ها به جز لایه‌ی آخر خروجی هدها با هم concat می‌شوند و در لایه‌ی آخر میانگین آن‌ها حساب می‌شود. برای تمام لایه‌ها به جز لایه‌ی آخر از تابع فعالسازی ELU استفاده شده است و در لایه‌ی آخر هیچ نوع‌ تابع فعالسازی را اعمال نکردیم. در نهایت مدلی Sequential با کنار هم قرار گرفتن لایه‌ها ساخته می‌شود.



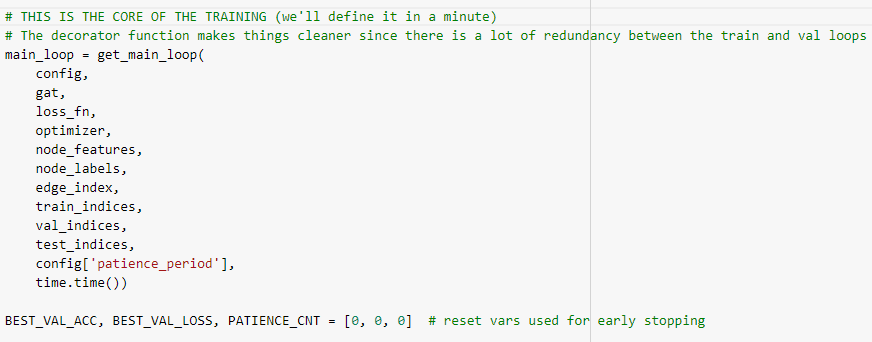
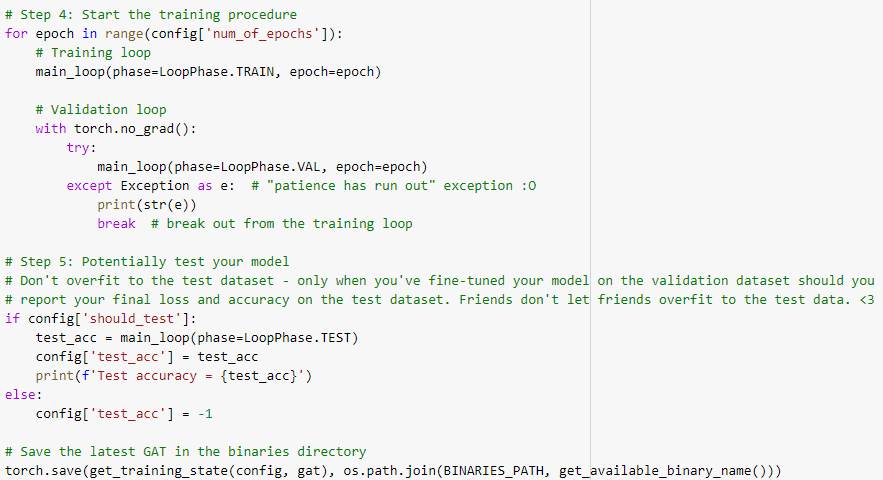
به وسیله‌ی تابع فوق مدل ساخته شده روی داده‌ی ورودی (یعنی گراف) اعمال می‌شود و خروجی لایه آخر به حل مسئله (مثلا طبقه‌بندی) کمک می‌کند.



در تابع get\_training\_args پارامترهای (هایپرپارامترهای) آموزش مثل تعداد epoch‌ها، Learning Rate، Patience period و L2-regularization و پارامترهای مدل مقداردهی شده‌اند. این پارامترها به عنوان config به تابع train\_gat داده شدند.

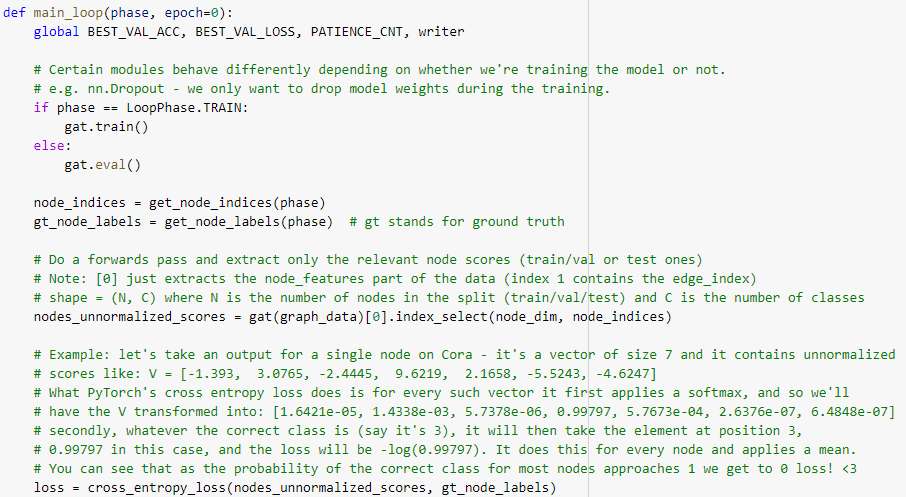
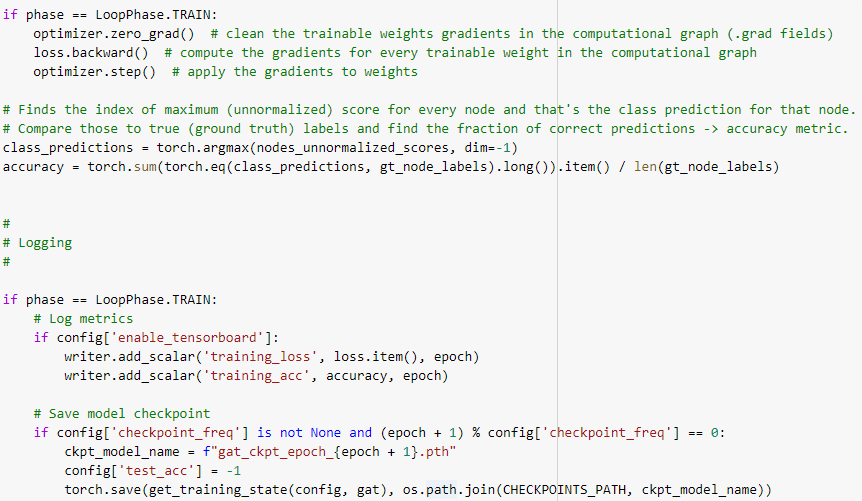
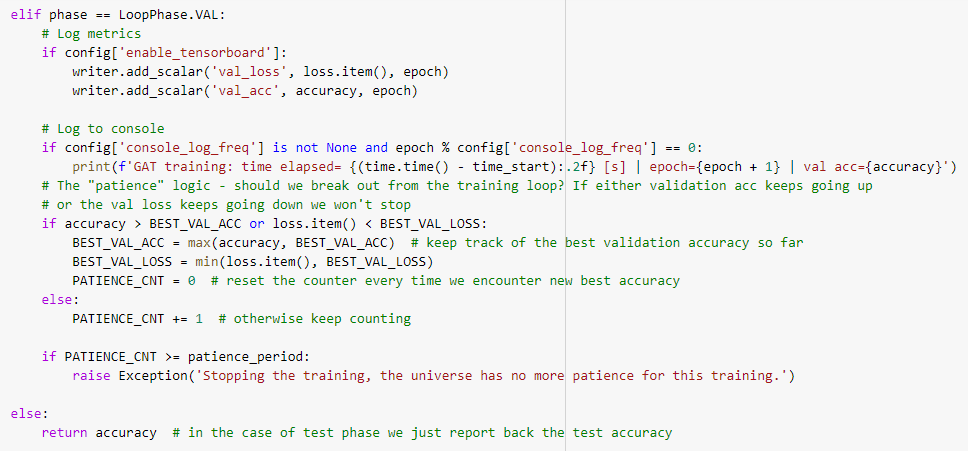


در تابع train\_gat علاوه بر کنترل دسترسی به cpu یا gpu و بازگذاری دیتا، پارامترها و لایه GAT روی آن‌ها در تنظیمات transductive از تابع ضرر CrossEntropy و در تنظیمات inductive از تابع ضرر BCEWithLogitsLoss استفاده شده است. در ضمن در هر دو تنظیمات بهینه‌ساز Adam برای آموزش انتخاب شده‌است.

در همین تابع train\_gat و با استفاده از تابع get\_main\_loop مدل آموزش می‌بیند و عملکرد آن روی داده‌های آموزش، اعتبارسنجی و آزمون محاسبه می‌شود.

یک مورد را اینجا متوجه نشدم و ننوشتم. من بر حسب دیتاست CORA توضیحات مدل رو نوشتم. میخواستم اینجا درباره سنجیدن با accuracy در این دیتاست و سنجیدن با f1\_measure در دیتاست ppi بنویسم اما مطمئن نبودم دلیل تفاوت معیارها مربوط به inductive و transductive بودن است یا متوازن و نامتوازن بودن داده‌ها

یکی از توابع مهم که در فرایند آموزش مورد استفاده قرار گرفته تابع main\_loop است که بهینه‌سازی و مینیمم کردن تابع ضرر یا هزینه در آن صورت می‌گیرد. همچنین دسته‌بندی (یا classification) به این صورت است که اندیس بزرگترین score برگردانده می‌شود.

مورد دیگری که به آن اشاره نکردم این است که score دقیقا چطور از خروجی GAT محاسبه می‌شود. آیا هر لایه score دارد یا score خروجی لایه آخر است که به نظر میرسد این مورد صحیح باشد. در تابع get\_training\_args که در ابتدای صفحه‌ی 2 همین فایل آمده در قسمت gat\_config نوشته

"num\_of\_layers": 2,

"num\_heads\_per\_layer": [8, 1],

"num\_features\_per\_layer": [CORA\_NUM\_INPUT\_FEATURES, 8, CORA\_NUM\_CLASSES]

دو لایه داریم. تعداد هدهای توی این دو لایه هم مشخص شده اما انگار تعداد ویژگیها در هر لایه لیست سه عضو یعنی سه لایه دارد و در لایه سوم که قاعدتا لایه gat هست دسته‌بندی انجام شده اما مثل مدلهایی که تا الان می‌دیدیم جز لایه‌ها شمرده نشده.