بنام خدا

مبانی محاسبات نرم (۸۳۰۵۰۳۹۰۱)- نیم تمرین شماره ۳

توسط: هادی ویسی

بارم کلیه نمرات از ۲۰ است

ردیف	شماره دانشجویی		سوال ۱ 25%	سوال ۲ 75%	تاخیر ساعت	نمره کل از 20	علت کسر نمره سوال اول	علت کسر نمره سوال دوم
1	830503034		20.00	19.20	2.0	**19.01**	کامل	به طور کلی هیچ کدی برای محاسبه میانگین زمان اجرای هر epoch نوشته نشده است و طبعا عدد های گزارش شده به عنوان میانگین زمان اجرای هر دور قابل اعتماد نیستند ولی اعداد گزارش شده منطقی هستند - به طور دقیق تر (در بخش اول - مدل پایه غلط است ، لایه پیچش با 32 فیلتر در مدل قرار داده نشده است)(در بخش دوم - در کد، اشتباه مدل پایه همچنان وجود دارد و لایه پیچش دوم همچنان به جای 32 فیلتر، 128 فیتر دارد و در گزارش، اضافه کردن لایه dropout باعث اضافه شدن نویز به مدل نمیشود که اشتباها به آن اشاره شده است)(در بخش چهارم - در گزارش، اضافه کردن لایه های batch normalization باعث اضافه شدن نویز به مدل نمیشود که اشتباها به آن اشاره شده است ولی تفسیر ارائه شده درباره تاثیر این عملیات بر روی فعال ساز relu بسیار خوب بود)(در بخش هشتم - گزارش، به حد زیادی درست ولی باز اشتباها به اینکه لایه های dropout و batch normalization نویز به مدل اضافه میکنند اشاره شده است) (در بخش 12 - مدل vgg16 اشتباها به جای vgg19 استفاده شده است) - در کل عالی
2	830503033		0.00	0.00	0.0	**0.00**	عدم تحویل	عدم تحویل
3	830503035		19.00	13.00	0.0	**14.50**	قسمت هایی از نوشته ناخوانا هستند	به طور کلی، شرط توقف بر اساس loss روی validation در هیچ قسمتی از بخش های سوال انجام نشده است - به طور دقیق تر (در بخش اول - در کد، مدل پایه اشتباه است ، در لایه اول پیچش به جای 64 کانال ، 32 کانال قرار داده شده است و در گزارش، نتیجه گیری انجام شده اشتباه است. طبق نمودار مدل صراحتا overfit شده است و در گزارش حتی به بروز این اتفاق اشاره نشده است) (در بخش سوم - کد مدل 4 لایه زده نشده است و 2 بار مدل 3 لایه نوشته شده است. همچنین در گزارش، پدینگ انجام شده تاثیر به سزایی در افزایش حجم محاسبات مدل ندارد و اشتباها به آن به عنوان عامل طولانی تر شدن میانگین زمان اجرای هر epoch اشاره شده است – مقایسه معناداری بر اساس تاثیر افزایش لایه بین مدل 3 و 4 لایه انجام نشده است و اگر انجام می گرفت نیز غلط بود ، زیرا هر دو مدل پیاده سازی شده 3 لایه بودند) (در بخش چهارم - در گزارش، اضافه کردن لایه batch normalization توزیع داده ها را در هر لایه ثابت نگه نمیدارد بلکه داده های هر دسته را به توزیعی با میانگین صفر و انحراف از معیار قریب به 1 منتقل میکند. همچنین "به مدل کمک می کند" توضیحی کلی است و نحوه تاثیر گذاری لایه batch normalization بر مدل ذکر نشده است) (در بخش ششم - در گزارش، درمورد علت تاثیر فعال ساز swish بر عملکرد مدل توضیحی داده نشده است) (در بخش هفتم - در گزارش، به علت تاثیر بهینه سازهای SGD و ADAMW برعملکرد مدل هیچ اشاره نشده است و فقط به بهتر شدن یا نشدن عملکرد مدل اشاره شده است) (در بخش هشتم - در گزارش، مانند بخش قبلی به علت تاثیر سایز دسته بر عملکرد مدل هیچ اشاره ای نشده است و مقایسه ای معنادار بین دو سایز دسته تست شده انجام نگرفته است) (در بخش نهم - در کد، محتوای کرنل ها نیز باید در فرمت greyscale نمایش داده میشدند) (در بخش یازدهم - در ماتریس سردرگمی کشیده شده ترتیب کلاسی رعایت نشده . در ماتریس های سردرگمی درایه هایی که متعلق به یک سطر و ستون هستند ، نماینده یک کلاس هستند ولی در ماتریس کشیده شده این موضوع رعایت نشده است) (در بخش دوازدهم - در کد، در نمودار خطا ، منحنی loss برای داده های validation رسم نشده و فقط منحنی loss برای داده های train روی دو مدل pretrain رسم شده اند . همچنین برای مدل reznet50 دقت روی داده های آزمایش گزارش نشده است و
4	830503031		16.00	17.13	0.0	**16.85**	توضیحات زیادی تحت الفظی ترجمه شده اند - روش جلوگیری از symmetry breaking در قسمت سوم توضیح داده نشده	در کل کد عالی - به طور دقیق تر (در بخش دوم - در گزارش، به تاثیر لایه dropout اشاره شده ولی به تاثیر کاهش تعداد کانال ها از 64 کانال به 32 کانال اشاره ای نشده است همچنین به اشتباه اشاره شده است که "تعداد تکرار های بیشتری نیاز بوده تا مدل کاملا واگرا شود")(در بخش سوم - در گزارش، به تفاوت های رخ داده در عملکرد مدل اشاره شده ولی به تفاوتی که اضافه کردن لایه های جدید در مدل و تعداد پارامترهای آن ایجاد میکند اشاره ای نشده است) (در بخش چهارم - در گزارش، مانند بخش قبلی به تغییر عملکرد مدل اشاره شده ولی به تفاوت رفتاری ای که اضافه کردن لایه batch normalization در عملکرد مدل ایجاد میکند، اشاره ای نشده است) (در بخش پنجم - در گزارش، تحلیل اشتباه است – استفاده از نرخ یادگیری 0.1 و 0.01 باعث تغییر نکردن وزن ها نمی شود بلکه باعث تغییر بیش از اندازه و مخرب وزن ها می شود) (در بخش ششم - در گزارش، به مزیت و یا عیب رفتاری دو تابع فعال ساز جدید نسبت به relu اشاره ای نشده است و فقط به تغییرات عملکرد مدل اشاره شده است) (در بخش هفتم - در گزارش، درباره تفاوت رفتاری بهینه ساز های SGD و ADAMW با فعال ساز مدل پایه ADAM هیچ اشاره ای نشده و فقط به تغیرات عملکرد مدل اشاره شده است) (در بخش هشتم - در گزارش، درباره علت تفاوت عملکردی مدل در استفاده از سایز های دسته مختلف صحبتی نشده و فقط به تغیرات عملکرد مدل اشاره شده است) در کل اگر تفاوتی در مدل و ساختار مدل ایجاد میکنید نه تنها باید به آن
5	830401036		19.50	16.72	2.0	**17.07**	استفاده از تابع فعال سازی relu مشکل exploding را مرتفع نمی کند	عدم گزارش و تحلیل میانگین زمان هر ایپاک در قسمت ب وپ وث وچ - عدم صحبت درمورد همگرایی در بخش ب و پ وت -نتیجه گیری ناقص قسمت ب - تصویر کرنل ها نمایش داده
6	830503032		15.00	14.00	0.0	**14.25**	توضیحات بیش از حد خلاصه نوشته شده اند - در بخش 4 ، تعداد روش های نام برده شده بسیار کم است و همچنین باید مزیت های هر روش به صورت مجزا و نسبت به هم نام برده میشد ، نه مزایای quantization به صورت کلی	کد تقریبا کامل به استثنای این مورد که میانگین زمان اجرای هر epoch در قسمت های مختلف سوال حساب نشده و همچنین در قسمت آخر به جای نمودار loss در حین یادگیری، نمودار خطای تشخیص رسم شده - به طور دقیق (در بخش اول - در گزارش، در صورتی که جواب شما به پدیده overfitting بله بود، طبق سوال باید راه حل پیشنهاد میدادید) (در بخش های 2 ، 3 ، 4 ، 5 ، 7 و 8 ، هیچ تحلیلی بر علت بهبود یا عدم بهبود مدل انجام نگرفته و فقط میزان آن گزارش شده) (در بخش نهم - در گزارش، از لفظ فیلتر به جای کرنل استفاده شده که اشتباه است و همچنین تحلیلی بر اطلاعاتی که کرنل ها به ما میدهند انجام نگرفته و فقط در ابتدا راجع به فیلر ها صحبت شده) (در بخش دوازدهم - در گزارش، اولین تحلیلی انجام شده غلط است، افزایش سایز تصویر هیچگاه باعث از بین رفتن جزئیات نمی شود و فقط در کاهش سایز تصویر این پدیده رخ میدهد. دو دلیل بعدی کمابیش قابل
7	830503036		0.00	0.00	0.0	**0.00**	عدم تحویل	عدم تحویل
8	830403064		16.00	14.80	2.0	**14.80**	مشکل تقارن به اشتباه توضیح داده شده است و روش های پیشنهادی برای حل این مشکل نیز متعاقبا اشتباه هستند	شرط early stopping در هیچکدام از بخش های سوال پیاده سازی نشده - به طور دقیق تر (بخش های 1، 2، 3، 6، 8، 10 و 11 از نظر کد، گزارش و تحلیل خوب هستند) (در بخش چهارم - در گزارش، درباره علت تاثیر لایه batch normalization بر عملکرد مدل صحبت نشده) (در بخش پنجم - در گزارش، تحلیل انجام شده برای نرخ یادگیری 0.01 اشتباه است، مدل در این حالت و طبق نموداری که در گزارش قرار داده شده، دچار بیش برازش نشده بلکه واگرا شده است که دقیقا نقطه مقابل آن است و در ادامه نیز اشتباها نوشته اید " با افزایش نرخ یادگیری ، احتمال بیش برازسازی افزایش میابد" که مجددا اشتباه است.) (دربخش هفتم - در کد، کد مربوط به adamw به درستی پیاده سازی نشده و طبعا اجرا نشده و در گزارش، به علت اینکه کد به درستی پیاده سازی نشده، نتیجه ای قرار نگرفته و تحلیل و مقایسه ای معنادار برای adamw انجام نگرفته است) (در بخش نهم - کد به درستی پیاده سازی نشده و گزارشی هم صورت نگرفته)
9	159403017		0.00	0.00	0.0	**0.00**	عدم تحویل	عدم تحویل
10	159403002		19.00	18.93	0.0	**18.95**	دلایل آورده شده برای مشکل exploding gradient اشتباه هستند به طور دقیق تر، توابع tanh و sigmoid باعث خطور این مشکل نمی شوند.	در کل کد و گزارش بسیار خوب - به طور دقیق تر (همه بخش ها به جز بخش های 7 و 12 عالی از نظر کد و گزارش) (در بخش هفتم - در گزارش، توضیحات بیشتری درباره نحوه کارکرد دو بهینه ساز در مقایسه با adam لازم بود) (در بخش دوازدهم - در کد، مدل vgg16 به جای vgg19 پیاده سازی شده – نمودار loss هنگام یادگیری برای داده های آموزش و ارزیابی رسم نشده – میانگین زمان اجرای هر epoch نیز محاسبه نشده و در گزارش، تحلیل انجام شده بر علت تفاوت عملکرد مدل های از پیش آموزش داده شده و مدل پیاده سازی شده در تمرین، کمابیش درست است – تحلیل دقیق تر در حضور نمودار loss در هنگام یادگیری بر روی داده های آموزش و ارزیابی روی مدل های از پیش آموزش داده ممکن و منطقی تر بود و همچنین بهتر بود تحلیلی بر آموزش کلیه لایه ها در مقایسه با آموزش لایه های تماما متصل در این
11	159403006		18.00	17.50	0.0	**17.63**	بهتر بود به زبان فارسی گزارش نوشته میشد - توضیحات همه بخش ها بیش از حد خلاصه هستند	هیچ کدی برای محاسبه میانگین زمان اجرای هر epoch نوشته نشده و اعداد گفته شده در گزارش قابل اعتماد نیستند - در کل کد و گزارش خوب بودند - به طور دقیق تر (به جز بخش های 6 ، 7 ، 11 و 12، همه بخش ها کامل از نظر کد و گزارش) (در بخش ششم - در گزارش، درباره علت تفاوت کارکرد swish نسبت به relu صحبت نشده و فقط نتایج گزارش شده) (در بخش هفتم - در گزارش، درباره علت تفاوت کارکرد adamw و sgd نسبت به adam صحبت نشده و فقط نتایج گزارش شده) (در بخش یازدهم - یا سایز تست ست بسیار کوچک است و یا ساب ستی بسیار کوچک از آن انتخاب شده زیرا ماتریس سر درگمی خیلی خلوت است – این ماتریس باید روی کل داده های تست ساخته میشد و نه بخش کوچکی از آن) (در بخش دوازدهم - در کد، نمودار loss هنگام یادگیری برای داده های آموزش و ارزیابی رسم نشده و در گزارش، دو مدل از پیش آموزش داده شده به جای مقایسه با مدل پیاده سازی شده در تمرین با یکدیگر مقایسه شده اند – همچنین بهتر بود تحلیلی بر آموزش کلیه لایه ها در مقایسه با آموزش لایه های تمام پیوسته به تنهایی، انجام میشد.)
12	159403005		20.00	20.00	3.0	**19.40**	-	-
13	159403010		19.00	15.50	1.0	**16.21**	بهتر بود به زبان فارسی گزارش نوشته میشد - توضیحات برخی قسمت ها بیش از حد خلاصه هستند	در کل تحلیل ها بسیار خلاصه بودند همچنین در کد محاسبه میانگین زمان اجرای هر epoch پیاده سازی نشده و اعداد آورده شده در گزارش قابل اعتماد نیستند - به طور دقیق تر (در بهش دوم - در گزارش، به تاثیر کاهش تعداد فیلتر های لایه اول اشاره نشده) (در بخش سوم - در گزارش، به تغییرات ساختاری مدل و تغییر در تعداد پارامترهای آن هیچ اشاره ای نشده) (در بخش چهارم - در گزارش، به نحوه کارکرد لایه های batch normalization و تاثیر آن ها در نحوه کارکرد مدل هیچ اشاره ای نشده) (در بخش ششم - در کد، به جای swish از selu استفاده شده و در گزارش، به علت پیاده سازی اشتباه تابع swish نتایج گزارش شده حائز ارزش نیستند و همچنین توضیحی درباره نحوه کارکرد این تابع داده نشده) (در بخش هفتم - در گزارش، توضیحات داده شده درباره نحوه کارکرد adamw بسیار مختصر و ناقص هستند و هیچ توضیحی درباره نحوه کارکرد sgd داده نشده) (در بخش نهم - در گزارش، درباره ارتباط وزن های کرنل و خروجی های لایه اول هیچ صحبتی نشده و به طور کلی درباره ویژگی های استخراج شده در لایه ها صحبت شده) (در بخش دوازدهم - در کد، نمودار loss هنگام یادگیری بر روی داده های آموزش و ارزیابی رسم نشده و در گزارش، بهتر بود تحلیلی بر آموزش کلیه لایه ها در مقایسه با آموزش لایه های تمام پیوسته به تنهایی، انجام میشد)
14	159403013		19.50	19.26	0.0	**19.32**	استفاده از تابع فعال سازی relu مشکل exploding را مرتفع نمی کند که در توضیح مشکل اشتباها به آن اشاره شده و ناسازگاری با توضیح خود تابع که در ادامه آورده شده دارد - به طور کلی عالی	در کل کد و گزارش بسیار خوب - به طور دقیق تر (همه بخش ها به جز بخش های 2، 7 و 12، عالی از نظر کد و گزارش) (در بخش دوم - در گزارش، تحلیل انجام شده درباره عملکرد لایه dropout اشتباه است . کند شدن همگرایی مدل، ناشی از پردازش های بیشتر مدل به دلیل حضور لایه dropout نیست، بلکه به علت کم تر فیت شدن مدل به داده های آموزش و افزایش تعمیم پذیری آن در اثر وجود این لایه هاست. درباره تاثیر کاهش تعداد فیلترهای لایه اول نیز صحبتی نشده) (در بخش هفتم - در گزارش، بهتر بود کمی بیشتر درباره نحوه کارکرد بهینه ساز SGD صحبت میشد) (در بخش دوازدهم - در کد، مدل vgg16 به جای vgg19 پیاده سازی شده ولی چون هر 3 مدل از پیش آموزش داده شده پیاده سازی شده است، نمره ای کسر نشد)
15	830402001		18.50	17.20	0.0	**17.53**	اولین علت شکت تقارن غلط بیان شده است، به طور دقیق تر مقدار دهی اولیه یکسان وزن ها، باعث شکت تقارن نمیشود ، باعث بروز تقارن میشود - در کل عالی	تعداد لایه های dropout استفاده شده، از بخش 2 بعد، بساز زیاد است و این موضوع مدل های پیاده سازی شده را به شکلی منفی تحت تاثیر قرار داده که این تاثیر منفی از مقادیر loss بدست آمده روی داده های یادگیری و ارزیابی آشکار است و با عمیق تر شدن مدل ها این مشکل تشدید می شود، با صرف نظر از این مورد، بخش های 1، 5، 8، 9، 10 و 11، از نظر کد و گزارش کامل هستند - به طور دقیق تر (در بخش دوم - در گزارش، نتیجه گیری ها کمابیش درست هستند اما تعدادی نتیجه گیری غلط مانند "افزودن لایه‌های Dropout در این مدل به کنترل همگرایی کمک کرده است" در گزارش وجود دارد) (در بخش سوم - در گزارش، نتیجه گیری ها اشتباه هستند. توجه شود، افزایش پیچیدگی مدل عمیق باعث کاهش بیش برازش نمی شود بلکه آن را افزایش میدهد که در نتیجه گیری چندین بار به اشتباه به آن اشاره شده – علت دقت کم شما روی داده های یادگیری و پدیده ای که مشابه برازش کم تر نسبت به مدل 2 لایه است، استفاده بی رویه از لایه های dropout می باشد که خود، عملی اشتباه است) (در بخش چهارم - در گزارش، نتیجه گیری ها کمایش قابل قبول هستند اما تعدادی از آنها مانند " این کاهش دقت به احتمال زیاد ناشی از تغییرات در تنظیمات مدل و پیچیدگی محاسباتی لایه‌های Batch Normalization است." اشتباه هستند. توجه شود پیچیدگی محاسباتی یک لایه به هیچ شکلی نمی تواند باعث کاهش دقت یک مدل شود) (در بخش ششم - در گزارش، هیچ اشاره ای به نحوه کارکرد توابع leaky relu و swish، مقایسه رفتاری آنها با relu وعلت بهبود نتایج نشده است) (در بخش هفتم - در گزارش، هیچ اشاره ای به نحوه کارکرد بهینه ساز های SGD و AdamW و مقایسه رفتاری آنها با Adam وعلت بهبود نتایج نشده است) (در بخش دوازدهم - در کد، مدل efficientnetb0 و مدل reznet50 هر دو مشکل پیاده سازی دارند. در هنگام پیاده سازی مدل های از پیش آموزش دیده باید توجه شود که زمان دادن ورودی به آنها، داده های شما باید ابتدا به سایز درست ورودی این مدل ها تغییر سایز داده شوند که در کد شما رعایت نشده. در صورت رعایت نکردن این موضوع مدل از وزن های از پیش آموزش داده شده لایه ابتدایی استفاده نمی کند و سایز خروجی قسمت استخراج ویژگی نیز تحت تاثیر قرار میگیرد و شما از مزیت transfer learning بهره مند نخواهید شد و در گزارش، پیاده سازی اشتباه مدل های از پیش آموزش دیده شده باعث نتیجه گیری های اشتباه از اجرای آنها شده است همچنین از نوآوری شبکه به اشتباه برداشت شده است و به جای اشاره کردن به نوآوری هر کدام از 3
16	830402004		18.50	19.80	0.0	**19.48**	اشاره اشتباه به منفی بودن عملیات symmetry breaking ، به طور دقیق تر شکست تقارن مانع یادگیری موثر و کامل شبکه نمی شود بلکه باعث آن می شود و در ادامه نیز راه حل های ارائه شده در گزارش باعث جلوگیری از شکست تقارن نیستند بلکه باعث ایجاد شکست تقارن هستند و این لفظ اشتباه در نتیجه گیری این قسمت نیز استفاده شده است - به طور کل توضیحات همه بخش های	مدل ها از مصرف بی رویه لایه dropout اندکی رنج می برند که این موضوع با عمیق تر شدن مدل، شدیدتر نیز می شود و این موضوع در مدل 4 لایه در قسمت سوم خود را به خوبی نشان داده است. توجه شود که لایه های dropout معمولا بعد از لایه های تماما متصل استفاده می شوند. با صرف نظر کردن از این مورد، تمامی بخش ها به جز بخش 4 و 12 از نظر کد و گزارش کامل هستند - به طور دقیق تر (در بخش چهارم - در گزارش، توضیح داده شده درباره نحوه کارکرد لایه batch normalization بسیار کلی است و نیاز به توضیحی دقیق تر درباره عملیات صورت گرفته در این لایه وجود دارد) (در بخش دوازدهم - در کد، گذاشتن لایه dropout بعد از قسمت فریز شده شبکه های از پیش آموزش دیده شده، کار مناسبی نیست و به شکلی منفی آموزش این شبکه ها را تحت تاثیر قرار میدهد. لایه های dropout باید فقط بعد از لایه های تماما متصل استفاده شوند. این مورد در پیاده سازی شبکه efficientnetb0 رعایت نشده بود ولی در reznet50 به درستی رعایت شده بود و گزارش
17	830402005		20.00	17.15	4.0	**17.15**		تحلیل ناقص در قسمت ب - عدم تحلیل میانگین زمان هر ایپاک در قسمت ب و پ و ت - تحلیل نادرست برای شبکه با 3 لایه در قسمت پ -تحلیل ناقص قسمت پ- عدم نمایش قسمت خ - همچنین کد قسمت خ مشکل دارد- عدم بیان نوع اوری شبکه ef - عدم
18	159403003		17.25	18.40	0.0	**18.11**	مشکل تقارن به اشتباه توضیح داده شده است -بخش بهینه ساز بسیار	عدم تحلیل همگرایی در قسمت ب و پ- تحلیل ناقص برای میانگین زمان هر ایپاک در قسمت پ -عدم توضیح batch normalization - عدم تحلیل بخش ت - عدم بحث در مورد
19	159403015		18.00	19.75	0.0	**19.31**	مشکل تقارن به اشتباه توضیح داده	عدم بیان نوع اوری شبکه EfficientNet - گزارش بسیار کامل و خوب نوشته شده است
20	830402095		0.00	0.00	0.0	**0.00**	عدم تحویل	عدم تحویل
21	830402058		18.00	14.20	23.0	**11.67**	مشکل تقارن به اشتباه توضیح داده شده است	عدم تحلیل همگرایی شبکه در قسمت الف -عدم انجام بخش ر - عدم تحلیل بخش ب و پ و ت وث و ج وچ و ح - عدم توضیح batch normalization - نمودار برای خطا sgd رسم نشده
22	159403018		20.00	17.00	1.0	**17.57**		همگرایی در قسمت الف و ب و پ و ث و ح به درستی بیان نشده است - عدم بیان نوع اوری شبکه های پری ترین در قسمت ر - نتیجه گیری ناقص در قسمت ر- بخش ر کد نبود اما
23	159403004		20.00	15.80	0.0	**16.85**		توضیح نادرست برای همگرایی در بخش الف و ب وپ و ث و ج و چ و ح- عدم توضیح علت بهبود در قسمت پ - عدم توضیح batch normalization -تعداد ایپاک کم در قسمت ث - تحلیل نادرست در قسمت ث و عدم مقایسه - نتیجه گیری نادرست در قسمت ر - عدم
24	830402023		18.00	17.35	7.0	**16.29**	مشکل تقارن به اشتباه توضیح داده شده است	عدم صحبت درباره همگرایی در ب و پ و ت و ج و چ وح- نتیجه ناقص در قسمت پ -عدم توضیح batch normalization - عدم توضیح برای قسمت خ - تحلیل ناقص برای قسمت ر و
25	830403011		18.00	18.70	0.0	**18.53**	مشکل تقارن به اشتباه توضیح داده	عدم بیان همگرایی در بخش ب و ج - دقت بسیار پایین بر روی شبکه های پری ترین و
26	159403008		0.00	0.00	0.0	**0.00**	عدم تحویل	عدم تحویل
27	159403020		0.00	0.00	0.0	**0.00**	عدم تحویل	عدم تحویل
28	159403019		17.00	17.28	3.0	**16.69**	راه حل برای مشکلات و علت Exploding و Vanishing به صورت تفکیک شده بیان نشده -مشکل تقارن به اشتباه توضیح داده شده	تعداد ایپاک کم در شبکه با 3 لایه در قسمت پ و ت وث وج و چ و ح- عدم توضیح batch normalization - در اکثر شبکه های به دلیل استفاده از تعداد ایپاک کم شبکه به خوبی ترین نشده است - عدم توضیح نو اوری شبکه های پری ترین - عدم گزارش قسمت خ کد موجود بود اما اجرا نشد) - دقت بسیار پایین در شبکه پری ترین
29	830402009		18.00	15.50	0.0	**16.13**	مشکل تقارن به اشتباه توضیح داده شده است	قسمت الف به خوبی ترین نشده است و خطا و دقت مناسب نیست - نتیجه گرفته شده برای قسمت الف با توجه به نتیجه نادرست است - عدم بیان همگرایی در قسمت الف و ب و پ و ج و چ و ح- نتیجه گیری ناقص قسمت ج - دقت بسیار پایین در قسمت چ - دقت بسیار پایین در قسمت ح - خروجی نمایش داده شده درقسمت خ ناقص است (تعداد کرنل ها نمایش داده شده کم است - خروجی از لایه اول نمایش ندادند) - عدم توضیح و نواروی
30	830402077		20.00	16.75	1.0	**17.39**		عدم گزارش زمان اجرا در هیچ قسمت - عدم تحلیل در قسمت ب - عدم تحلیل در قسمت پ- عدم تحلیل در قسمت ت - تحلیل ناقص ث - عدم تحلیل در قسمت ج- تحیل ناقص قسمت ح- عدم نمایش قسمت خ - عدم قرار دادن بخش ذ در گزارش - دو مدل در قسمت ر
31	830402079		20.00	19.30	0.0	**19.48**		
32	830503030		20.00	17.00	0.0	**17.75**		نتیجه قسمت پ ناقص است (توضیحی در مورد علت بهبود نشده است - توضیح ناقص همگرایی) - عدم بررسی زمان همگرایی در قسمت ت-عدم بیان دلیل بهبود شبکه در قسمت ت -توضیح نادرست همگرایی در قسمت ج - عدم بیان علت بهبود شبکه در قسمت چ -عدم تحلیل قسمت چ -نتیجه نادرست قسمت ح برای بچ 128 -عدم توضیح شبکه های
33	150901039		20.00	17.66	0.0	**18.25**	توضیحات کامل و جامع	کد بسیار خوب - بخش های 1، 3، 5، 7، 8، 10 و 11، از نظر کد و گزارش کامل هستند - به طور دقیق تر (در بخش دوم - در گزارش، به نحوه کارکرد لایه dropout آن و تاثیر آن بر روی مدل هیچ اشاره ای نشده) (در بخش چهارم - در گزارش، به نحوه کارکرد لایه dropout آن و تاثیر آن بر روی مدل هیچ اشاره ای نشده) (در بخش ششم - در گزارش، به نحوه کارکرد دو تابع leaky relu و swish اشاره ای نشده است) (در بخش نهم - در گزارش، درباره ارتباط تصاویر نقشه های ویژگی و وزن های کرنل و تصاویر ورودی، هیچ صحبتی نشده) (در بخش دوازدهم - در کد، نمودار loss روی داده های یادگیری و ارزیابی در حین آموزش رسم نشده و
34	159403021		0.00	0.00	0.0	**0.00**	عدم تحویل -	عدم تحویل -
35	159302001					**0.00**	عدم تحویل -	عدم تحویل -
	کمینه		**0.0**	**0.0**	**0.0**	**0.0**		
	بیشینه		**20.0**	**20.0**	**23.0**	**19.5**		
	میانگین		**14.8**	**13.7**	**1.4**	**13.3**		