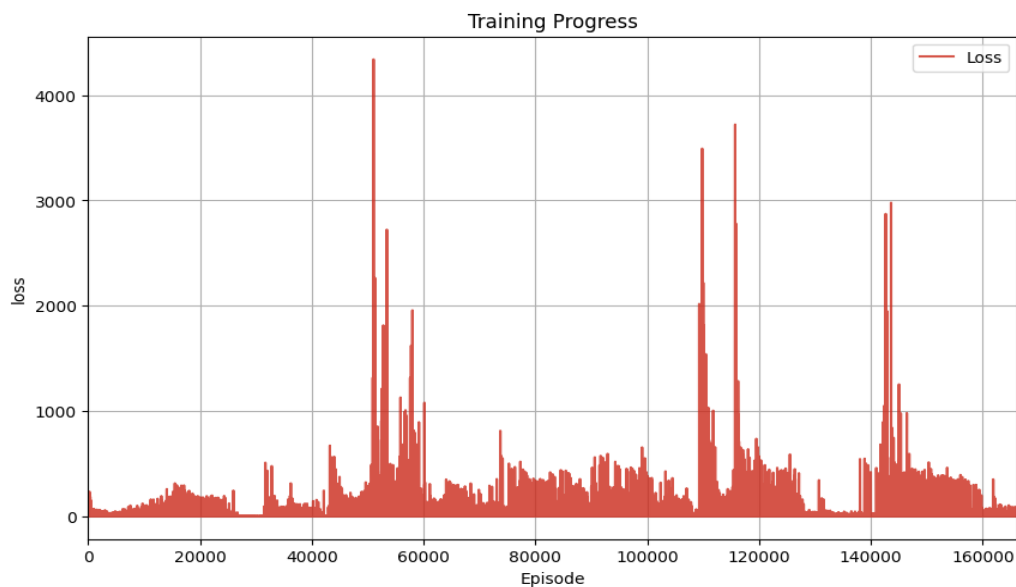
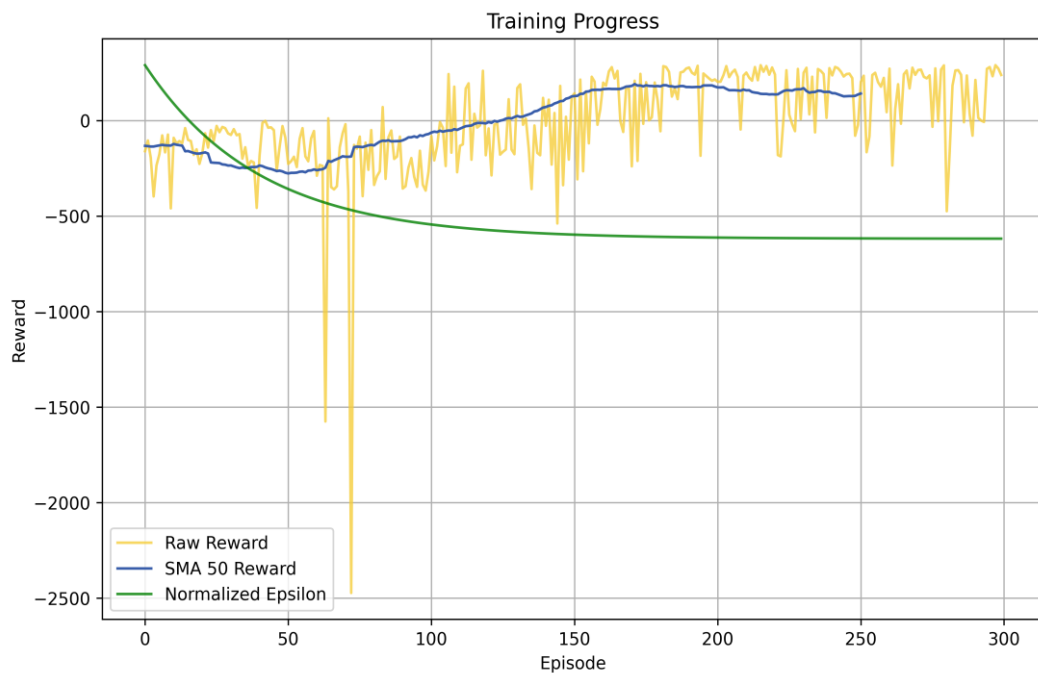


مقایسه الگوریتم ها DQN و D3QN

اول از همه این نکته رو بگم که برای این که هر دو این الگوریتم ها عمل کرد قابل قبولی داشته باشند باید سائیز بچی که برای آموزش استفاده میکنن بزرگ باشه در غیره این صورت عمل کرد خوبی نخواهند داشت و شبکه باید به حد خوبی بزرگ باشد هر دو این تخیرات به این علت است که فضای عمل و حالت بزرگ تری داریم.

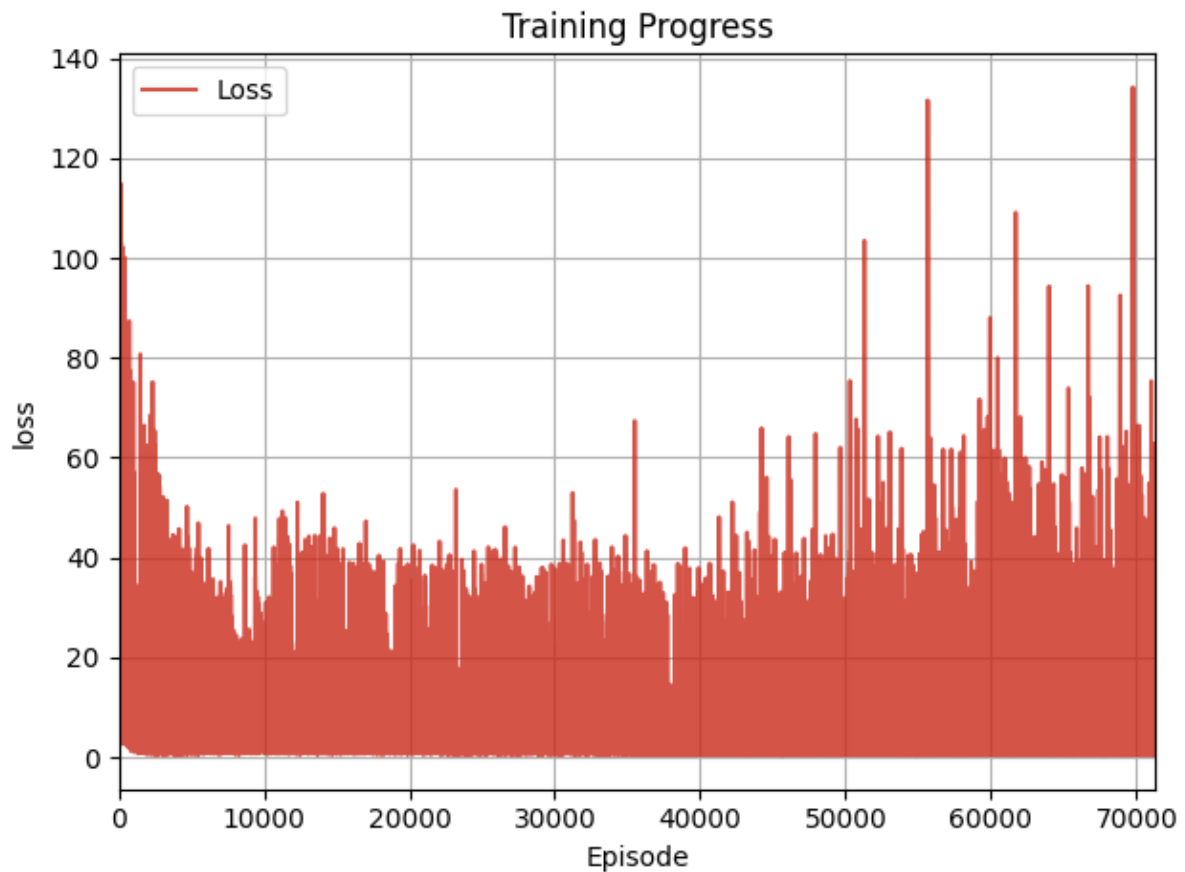
نمودار های زیر مربوط به الگوریتم D3QN هست



که اگر خوب دقت کنیم هر جا که لاس بالا رفته بعد از اون مدل شروع به یادگیری کرده و دقیقا از وقتی که لاس میاد پایین مدل عمل کرد بهتری داره

برای این که هر دو این دو الگوریتم رو تست کنید کافیه که مدل رو در حالت تست قرار دهید یعنی متخیر train رو برابر False قرار دهید همچنین از عمل کرد هر دو یک ویدیو قرار داده شده است در پوشه های هر کدام از الگوریتم ها به علاوه وزن های ذخیره شده

نمودار زیر مربوط به الگوریتم D3QN هست وقتی که بخش دو تمرین در آن اعمال شده است که تابع لاس برای داشتن کاوش های بیشتر بسیار نوسانی است من با ایده مقاله موافقم ولی من که تکرار 183 اون رو ران کردم دیدم تا پاداش های خیلی بد در حد -1500 دریافت میکنه برعکس حالت معمولی و دیگه آموزش رو ادامه ندادم





نمودار زیر مربوط به الگوریتم DQN هست اولین تفاوت در تعداد تکرار های مدل هست با مدل قبلی بعد از اون می بینیم که وقتی که در نمودار مدل قبلی بعد 150 تکرار به حداکثر پاداش میرسد ولی این مدل بعد از 500 تکرار به این هدف نمیرسد البته ذکر این نکته هم مهم این که هر تکرار در الگوریتم اول زمان بیشتری می برد و دلیل ان هم این است که محاسبات بیشتری دارد این که الگوریتم DQN میتواند این محیط را حل کند ولی نیاز به تکرار های بیشتر و تنظیم دقیق تر ابر متغیر ها دارد

مهم ترین کاری من تلاش کردم بکنم این بود که جلوی این نوسان ها در تابع لاس با بگیرم و موفق هم شدم ولی بعد از تکرار 305 مدل گیر میکرد وقت کافی برای حل این مشکل نبود اگر این مشکل رو فرصت داشتم که حل کنم قطعاً محیط حل میشد همچنین یکی از مهم ترین علت ها این مسعله بزرگ بودن فضای عمل و حالت است که دچار این نوسات ها می شویم

