

Дамашняе заданне 2

Аўласаў Уладзіслаў

Назва каманды: BSU_Aulasau_Uladzislau

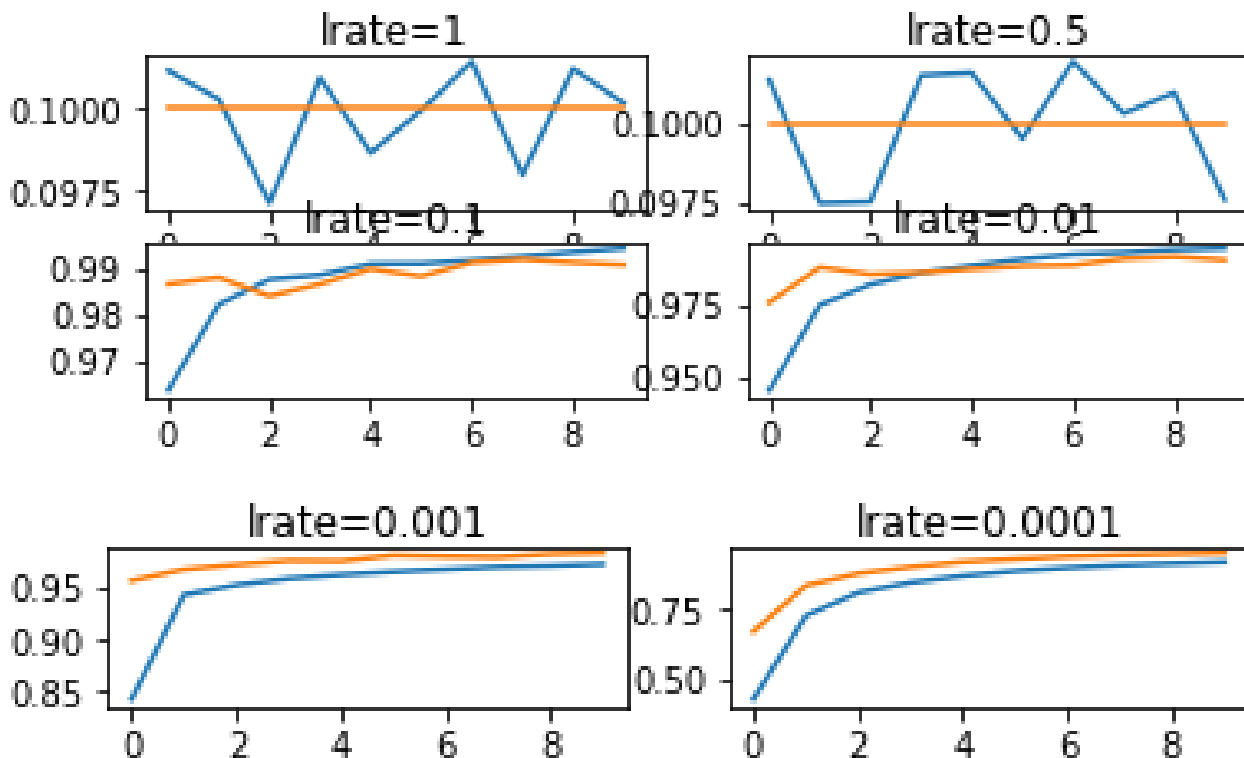
Заданне 1

Пабудаваў LeNet-5, прыкладаю файл LeNet-5.ipynb з сыходным кодам. Я разбіваў трэйн-сэт на 50 тысяч аб'ектаў для трэйна і 10 тысяч для валідацыі. У выніку на апошняй эпохе дакладнасць на гэтай валідацыі была 0.990999, але на некаторых эпохах раней яна была яшчэ лепей. У выніку дасылаў мадэль з лепшай дакладнасцю на валідацыі. Public score = 0.95860 дакладнасці.

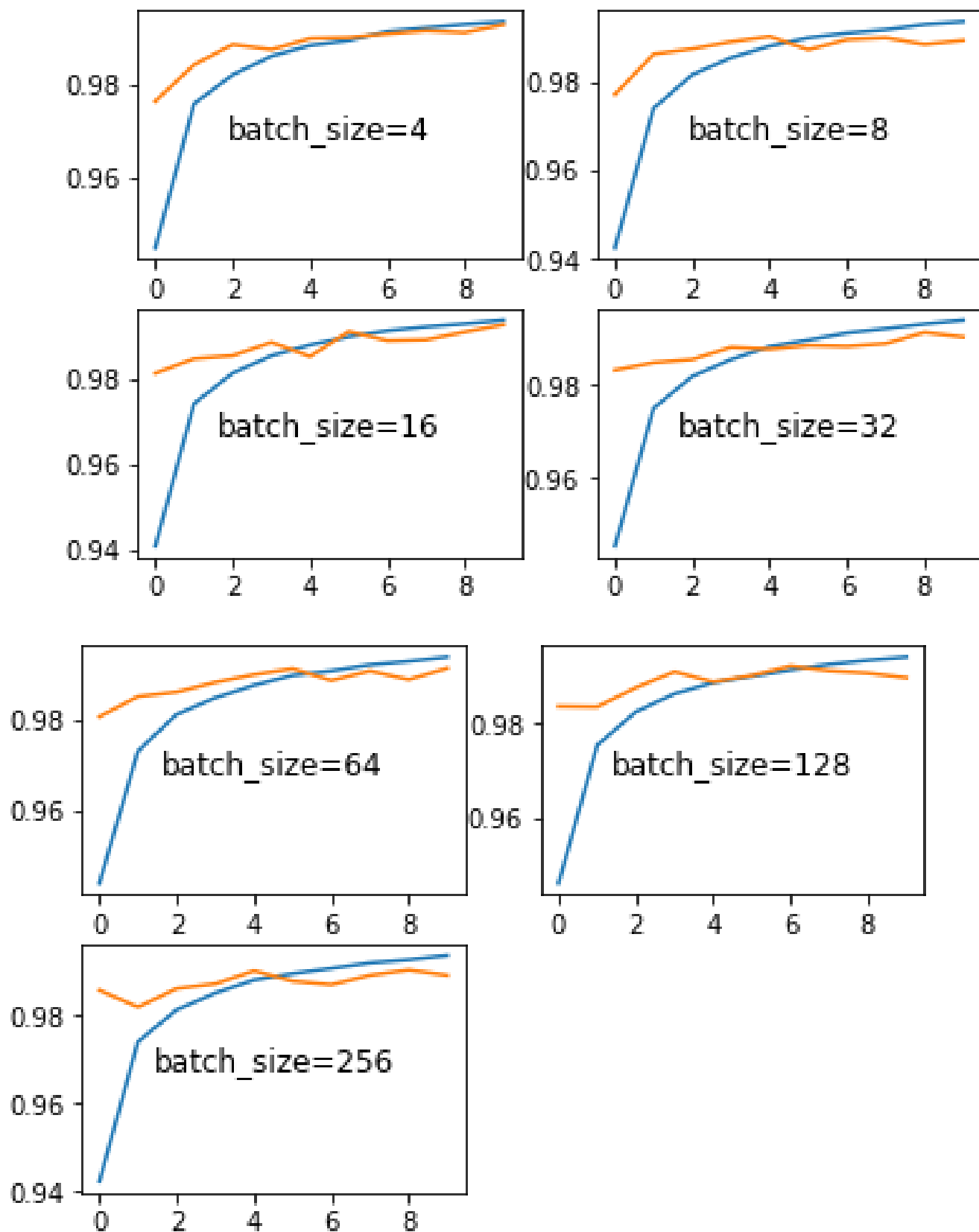
Заданне 2 і Заданне 3

Для гэтых заданняў я знізіў колькасць эпох да 10, каб сеткі не вучыліся вечна.

Для learning rate узяў значэнні [1, 0.5, 0.1, 0.01, 0.001, 0.0001], для batch size – значэнні [4, 8, 16, 32, 64, 128, 256] пры дэфолтным learning rate.



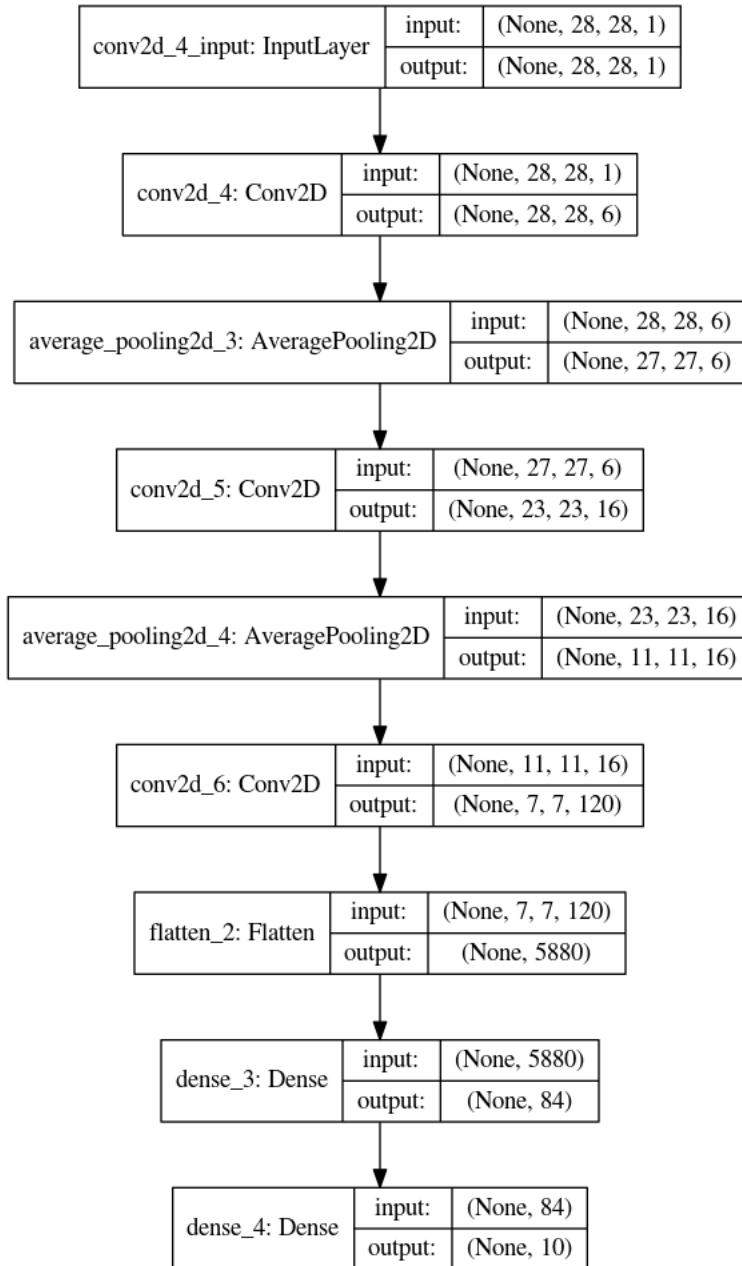
Як можна ўбачыць, добрыя рэзультаты lrate у нас на [0.1, 0.01, 0.001]. Пры вялікім значэнні ў нас зусім няма сатбільнасці, градыенты спуск не можа трапіць у свае экстрэмы. Пры нізкіх значэннях збежнасць алгарытма нізкая, навучанне не эфектыўнае і нам патрабуецца значна больш эпох для дасягнення хаця б нейкіх рэзультатаў.



Нельга сказаць, што batch size моцна ўплывае на дакладнасць, прынамсі пры абранай намі архітэктурцы і параметрах. Але па тэорыі, гэты параметр уплывае на тое, як часта нашыя вагі ў нейроннай сетцы будуць аднаўляцца, то бок на тое, наколькі плаўна будзе змяняцца градыент. Пры маленькіх значэннях узрасце ўплыў ўсялякіх шумоў, але хуткасць збежнасці павінна ўзрасці, калі ў сярэднем усё добра. Пры большых значэннях збежнасць не такая хуткая, але ацэнкі акуратней і плаўней.

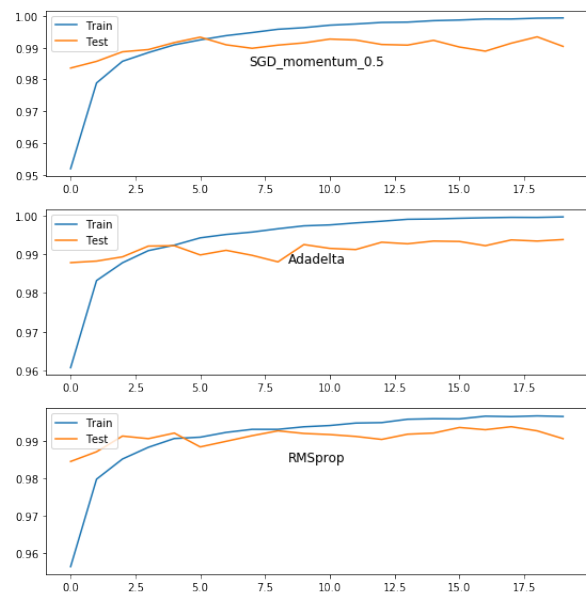
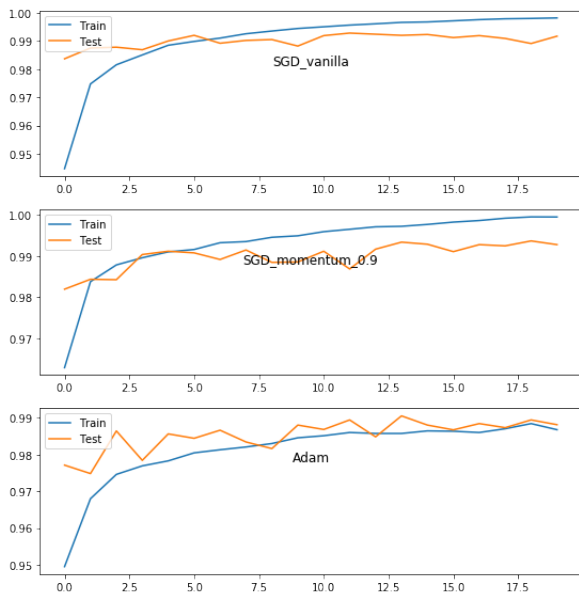
Заданне 4

Пакуль я недастаткова разбіраюся ў нейронных сетках, каб пабудаваць сваю добрую паспяховую, таму эксперыментавал з LeNet-5. Вось яе граф:



Заданне 5-6

Запусціў з рознымі аптымайзерамі.



З большага, паўсюль добрыя рэзультаты, нельга прам вызначыць нешта паасобку. Але тое, як Адам на тэсце выдае лепшы рэзультат за трэйн, варта адзначыць.