

# 2024 ANN Project 6차 수행일지

- MNIST extended dataset을 이용한 CNN 모델 최적화 및 분석 -

■ 회의 정보

팀 명	7조
수행 제목	최종 모델 후보 하이퍼파라미터 변경
날짜	2024.06.03.월
시간	21:00 - 22:00
수행자 이름	정하연, 송준규
참여 인원	정하연, 박태현, 송준규, 양은주

■ 수행 내용 및 결과

수행 내용	1. DenseNet 개선하기								
	- learnin_rate, optimizer 조합 찾기								
	- activation function 실험								
	- dropout 실험								
	- augmentation 실험								
	- BatchNormalization								
	- 가중치 제한								
	- 노이즈 주입								
결과	1. DenseNet 개선하기								
	- learnin_rate, optimizer 조합 찾기 (0.01, Nesterov)								
	batch_size=2048								
	lr \ optimizer	Adam	AdaMax	Nadam	AdamW	RMSprop	Nesterov	Momentum	
	0.01	training_time	544.56	550.91	760.89	567.72	620.66	520.53	529.66
		val_loss	0.3642	0.3317	0.4253	0.4842	0.6007	0.5694	0.577
		val_accuracy	0.873	0.8842	0.8711	0.8707	0.8719	0.8932	0.8922
	0.001	training_time	537.02	569.95	774.14	561.67	614.44	536.29	519.2
		val_loss	0.6163	0.8789	0.9519	0.808	0.8885	1.1198	1.1106
		val_accuracy	0.8924	0.8893	0.8883	0.8861	0.8869	0.8867	0.8854
0.0001	training_time	545.23	565.66	751.28	557.15	613.29	542.48	519.65	
	val_loss	1.1086	1.1193	1.1327	0.7662	0.8024	0.9218	0.9276	
	val_accuracy	0.8867	0.8865	0.8877	0.8865	0.8895	0.8891	0.8889	

batch\_size=1024

lr \ optimizer		Adam	AdaMax	AdamW	Nesterov	Momentum
0.01	training_time	902.4	915.48	976.46	869.01	866.92
	val_loss	0.3468	0.3042	0.3891	0.311	0.3194
	val_accuracy	0.8793	0.8906	0.8766	0.9009	0.8993
	test_accuracy	0.866	0.869	0.868	0.904	0.895
0.001	training_time	895.94	913.5	929.1	906.08	868.29
	val_loss	0.3391	0.6115	0.578	0.5875	0.5953
	val_accuracy	0.8995	0.8925	0.8896	0.8892	0.8888
	test_accuracy	0.891	0.893	0.886	0.89	0.891

- activation function 실험

fixed values: lr=0.001, optimizer=Nesterov								
activation \ evaluate	training_time	val_loss	val_accuracy	test_accuracy	top-1 accuracy	top-5 accuracy	evaluation_time	inference_time
ReLU	892.24	0.2893	0.8976	0.8867	0.8867	0.9922	11.858	10.724
LeakyReLU	892.01	0.2805	0.8997	0.8936	0.8936	0.9932	12.534	11.473
ELU	900.86	0.2632	0.903	0.8945	0.8945	0.9961	11.979	11.272
Swish	1168.56	0.303	0.898	0.8984	0.8984	0.9922	13.849	13.366

- dropout 실험

dropout \ evaluate	training_time	val_loss	val_accuracy					
dropout_1	943.83	0.3013	0.8917	시간이 더 걸리고 accuracy가 조금 줄긴 했지만, training curve와 validation curve의 차이가 확연히 줄어들었기에 0.5->0.3으로 낮춰서 테스트 모델이 훈련 데이터에서의 성능을 저하시킬 정도로 규제가 강하게 작용 (과소적합) 모델이 훈련 데이터에서의 성능을 저하시킬 정도로 규제가 강하게 작용 (과소적합) loss가 가장 낮고, learning_curve가 가장 적게 진동함 -> 테스트 대상				
dropout_2	940.85	0.3642	0.8748					
dropout_3	973.92	0.4051	0.8705					
dropout_4	1037.98	0.4292	0.8553					
dropout_5	973.22	0.2948	0.8942					

### (1) Fully Connected Layer 뒤에 Dropout  
### (2) 각 Dense Block 뒤에 Dropout  
### (3) 각 Transition Block 뒤에 Dropout  
### (4) Dense Blocks 내부 Convolutional Layers 뒤에 Dropout  
### (5) 처음 Convolution Layer 뒤에 Dropout  
  
### (1-1) 0.5 -> 0.3  
### (1-1) 0.3 -> 0.2

dropout \ evaluate	training_time	val_loss	val_accuracy	test_accuracy	top-1 accuracy	top-5 accuracy	evaluation_time	inference_time
dropout_1_1	974.06	0.2847	0.8974	0.9072	0.9072	0.9912	12.504	11.357
dropout_5_1	973.52	0.2926	0.8939	0.8906	0.8906	0.9912	15.929	11.117

이제 learning\_curve가 진동하지 않는 조건을 확실히 찾음 -> early\_stopping(patience=10)으로 지정하고 개션 시작.

- augmentation 실험

augmentation \ evaluate	training_time	val_loss	val_accuracy	test_accuracy	top-1 accuracy	top-5 accuracy	evaluation_time	inference_time	
aug_1	1698.6555	0.3539	0.875	0.8818	0.8818	0.9902	12.207	10.174	validation curve > training_curve
aug_2	1920.1612	0.305	0.8863	0.8936	0.8936	0.9922	10.959	10.101	
aug_3	1986.3978	0.2885	0.8924	0.8936	0.8936	0.9932	10.659	10.114	
aug_4	1429.6499	0.2878	0.8935	0.8906	0.8906	0.9932	10.708	10.868	
aug_5	1994.2567	5.1104	0.024	0.0273	0.0273	0.1104	10.521	9.975	

aug\_3 (lr\_scheduler) validation\_curve가 아래 상승하지 않음. (과대적합)

## ■ 다음 수행 계획

다음 수행 계획	<p>- 다양한 모델 구조 설계, 최종보고서 작성 (송준규)</p> <p>- 설계된 모델 훈련 및 측정, 최종 PPT 제작 및 발표 준비 (정하연)</p> <p>- 통계적 검정 결과 정리 최종보고서 작성 (박태현)</p> <p>- 최종 PPT 제작 및 발표 준비 (양은주)</p>	
프로젝트	04.29(월)	프로젝트 개요 파악, 역할 분담, 수행계획서 작성
수행일정	05.06(월)	EMNIST 분석, LeNet 5 및 ResNet 50 스타다 및 학습

	05.13(월)	LeNet-5 및 ResNet-50 하이퍼파라미터 변경, 다양한 CNN 조사
	05.20(월)	다양한 CNN 모델 학습, 중간발표 PPT 제작, 대본 작성, 리허설
	05.22(수)	중간발표
	05.27(월)	CNN 모델 개발, 다양한 평가 metric 스터디, 실험 분석
	06.03(월)	최종발표 PPT 제작, 대본 작성, 리허설, 보고서 작성
	06.12(수)	최종발표