

Report for assignment

담당조교 : 신우현 (wuhyun.shin@kasit.ac.kr)

이하연 (hayeon926@kaist.ac.kr)

보고자 : 오진호 (삼성 AI Expert 과정, starjinogl@gmail.com)

보고 내용 : 과제 A

1. SGVB에서 Reparametrization trick을 쓰는 이유는?

→ Stochastic Gradient Variational Bayes(SGVB)에서 Gradient variance를 낮추고 더 쉽고 빠르게 계산

2. Sparse VD에서 Local reparameterization Trick을 사용하는 이유는?

→ Droprate α 가 큰 영역에서 θ 에 대한 gradient variance가 매우 크기때문에, θ 와 무관한 σ 로 reparametrization 시켜줌

3. Sparse VD에서 weight의 sparsity가 유도되는 이유는?

→ Fixed rate Dropout에 비하여 weight별 독립적인 droprate 학습

Weight θ 는 Droprate $\alpha \rightarrow \infty$ 로 가면 무한대의 noise가 발생하므로, 이를 상쇄하기 위해 $\theta \rightarrow 0$ 으로 되어야 하고, 이는 weight을 dropout 하는 효과가 있음

4. Variational dropout의 prior가 log uniform인 이유는?

→ Zero 근처에서 높은 Probability density → weight에 적용할 경우 sparsity 유도 가능

보고 내용 : 과제 B

1. 타입 변경 한 뒤의 dropout_type 변수와 그때의 training log의 일부분을 캡처하여 보고서에 첨부해주세요. (1점)

① dropout_type = None

Training Log → train: epoch 60, (1.645 secs), cent 0.011133, acc 0.996267 test: epoch 60, (1.821 secs), cent 0.069962, acc 0.981400

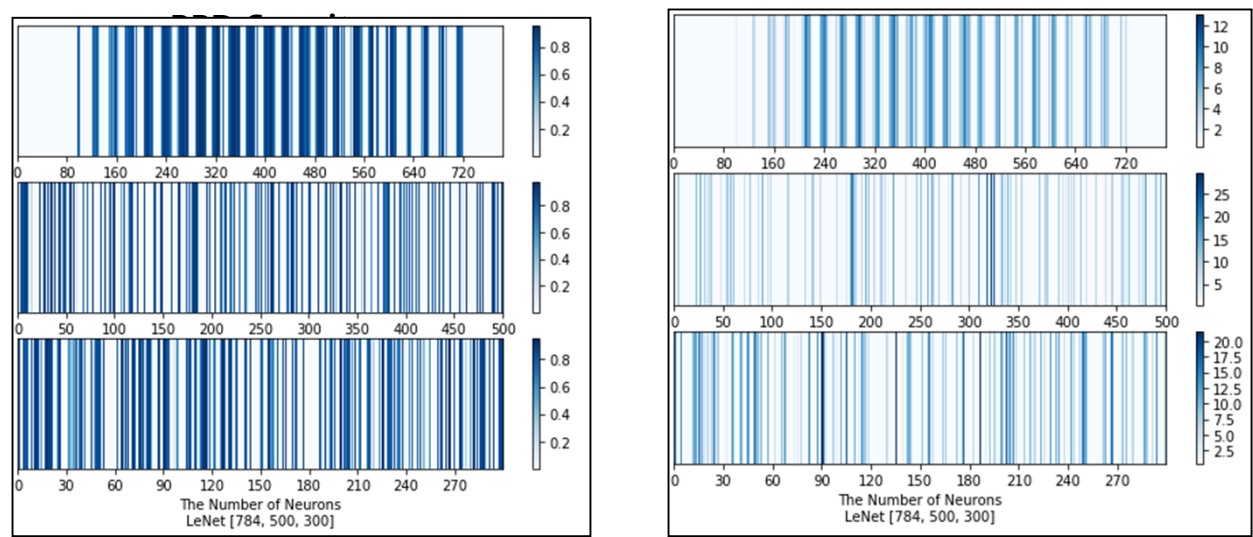
② dropout_type = 'bbdropout'

Training Log → train: epoch 60, (4.172 secs), cent 0.015482, acc 0.995583 test: epoch 60, (4.439 secs), cent 0.042999, acc 0.987800 Epoch 60 start, learning rate 0.000100 n_active: [366, 202, 159]

③ dropout_type = 'sbpd dropout'

Training Log → train: epoch 60, (9.375 secs), cent 0.004073, acc 0.999817 test: epoch 60, (11.244 secs), cent 0.048523, acc 0.984600 Epoch 60 start, learning rate 0.000100 n_active: [325, 147, 113]

2. BBD와 SBPD 레이어의 sparsity를 시각화한 그림을 보고서에 첨부해주세요. (2점)



3. 위 3가지 타입의 (Non) Dropout에 대해 아래 표를 완성하여 보고서에 첨부하세요. (2점)

Model	Test Acc(%)	Layer 1 Memory (%)	Layer 2 Memory (%)	Layer 3 Memory (%)
None	98.14	100	100	100
BBD	98.78	46.68	40.40	53.00
SBPD	98.46	41.45	29.40	37.67