

The Lottery Ticket Hypothesis

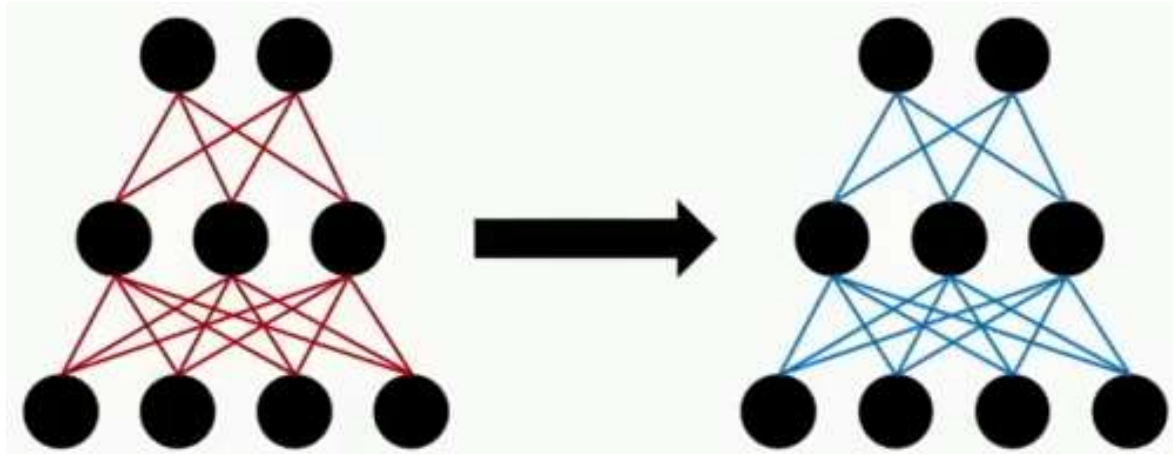
Шабалин Александр

16 марта 2020 г.

Network Pruning

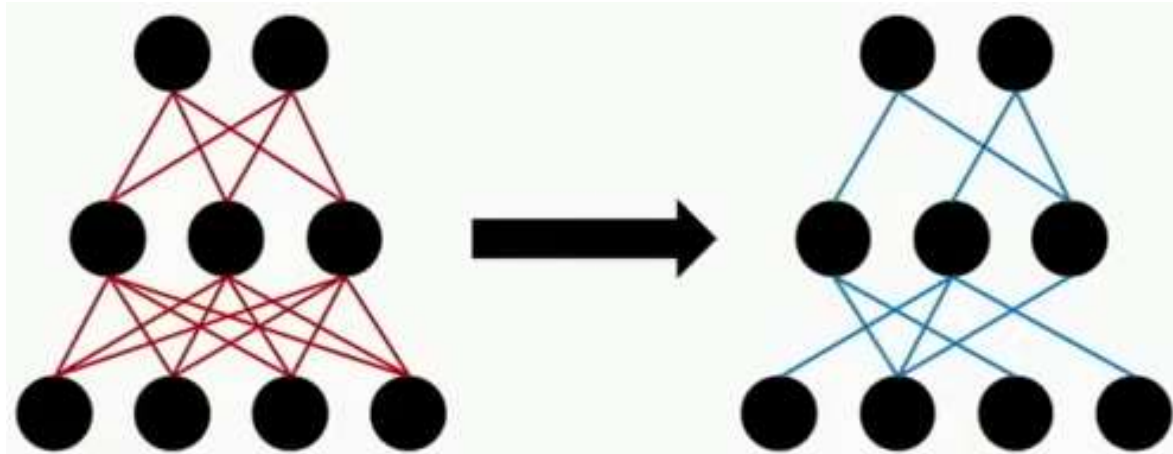
Network Pruning

1. Обучить нейронную сеть
2. Удалить лишние элементы
3. Дообучить полученную сеть (Fine-tune)
4. Повторить



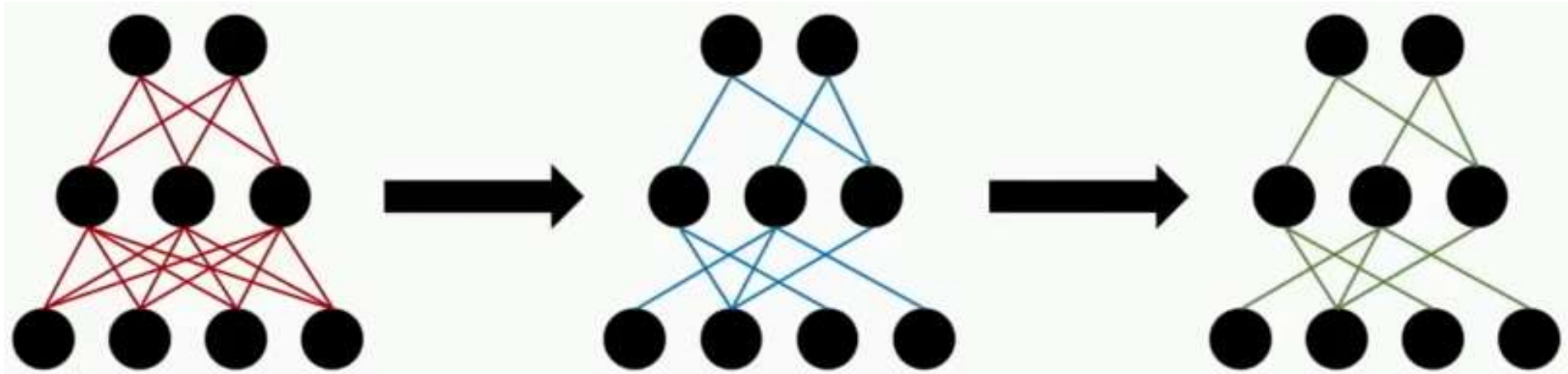
Network Pruning

1. Обучить нейронную сеть
2. Удалить лишние параметры
3. Дообучить полученную сеть (Fine-tune)
4. Повторить



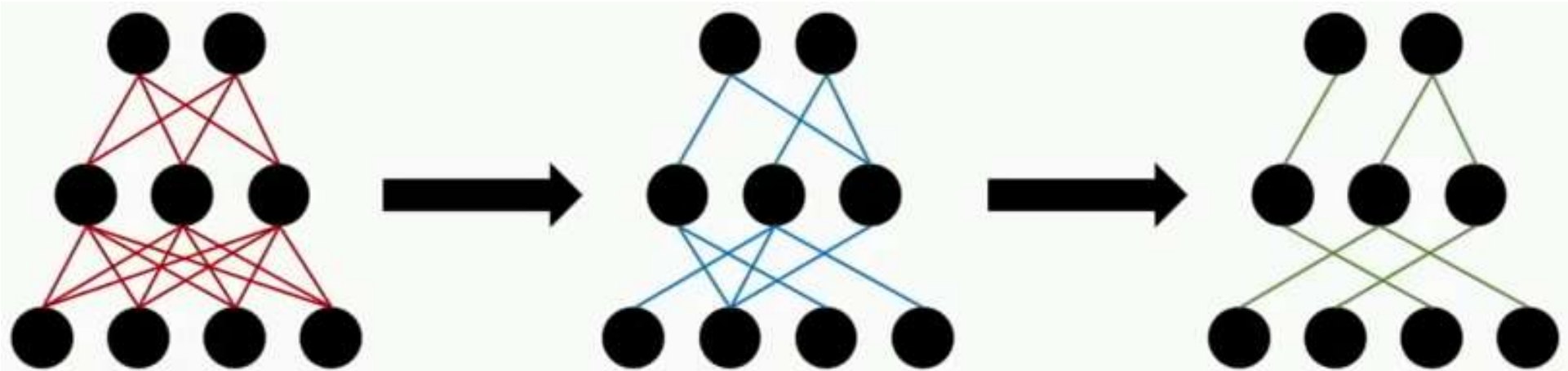
Network Pruning

1. Обучить нейронную сеть
2. Удалить лишние элементы
3. Дообучить полученную сеть (Fine-tune)
4. Повторить



Network Pruning

1. Обучить нейронную сеть
2. Удалить лишние элементы
3. Дообучить полученную сеть (Fine-tune)
4. Повторить



Network Pruning

Network	Top-1 Error	Top-5 Error	Parameters	Compression Rate
LeNet-300-100 Ref	1.64%	-	267K	
LeNet-300-100 Pruned	1.59%	-	22K	12×
LeNet-5 Ref	0.80%	-	431K	
LeNet-5 Pruned	0.77%	-	36K	12×
AlexNet Ref	42.78%	19.73%	61M	
AlexNet Pruned	42.77%	19.67%	6.7M	9×
VGG-16 Ref	31.50%	11.32%	138M	
VGG-16 Pruned	31.34%	10.88%	10.3M	13×

Если качество не падает, то почему бы не тренировать урезанные сети с нуля?

The Lottery Ticket Hypothesis

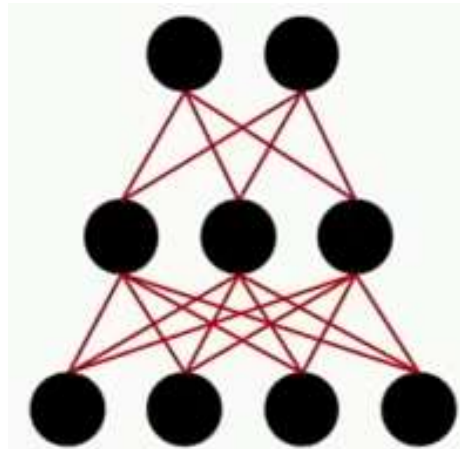
The Lottery Ticket Hypothesis

Гипотеза:

Случайно инициализированная нейронная сеть содержит в себе подсеть, которая при обучении на не большем числе итераций достигает не меньшей точности, чем исходная сеть.

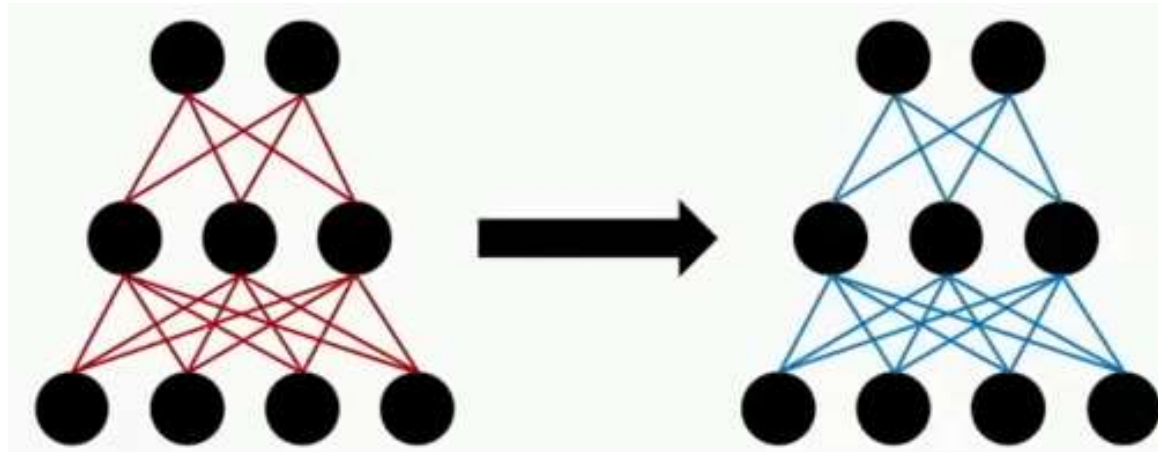
The Lottery Ticket Hypothesis

1. Создаем нейронную сеть со случайными весами
2. Обучаем ее и обрезаем лишние параметры
3. Возвращаем параметры к исходным значениям
4. Повторяем (2)-(3)



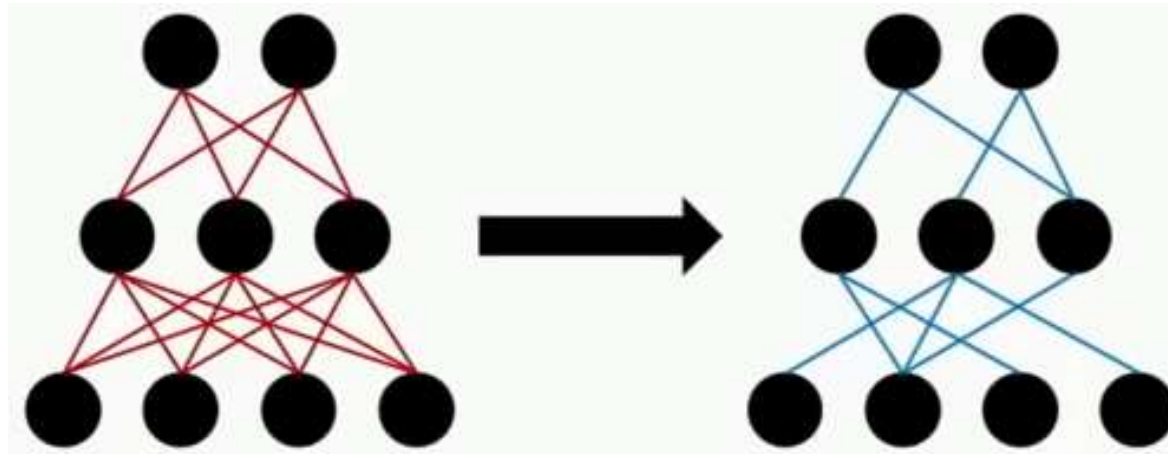
The Lottery Ticket Hypothesis

1. Создаем нейронную сеть со случайными весами
2. Обучаем ее и обрезаем лишние параметры
3. Возвращаем параметры к исходным значениям
4. Повторяем (2)-(3)



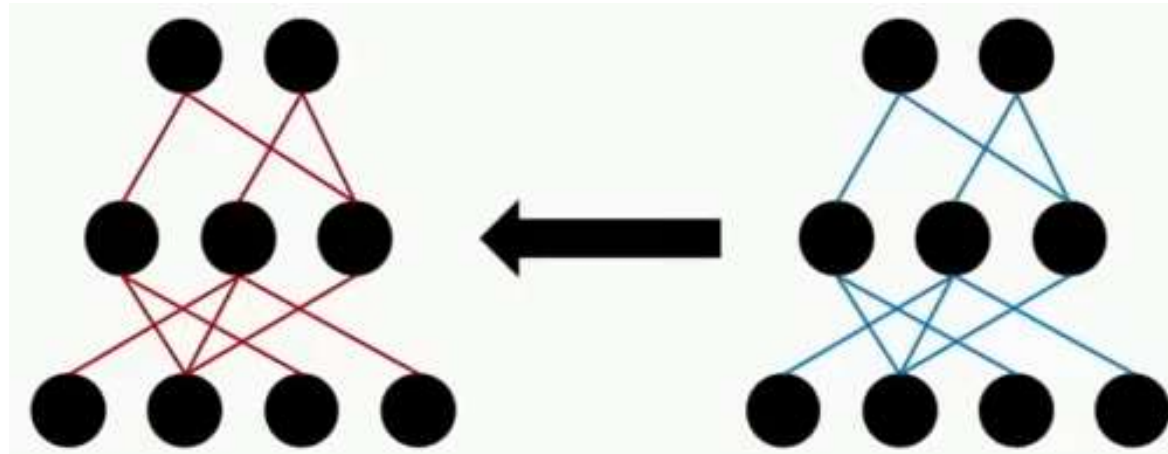
The Lottery Ticket Hypothesis

1. Создаем нейронную сеть со случайными весами
2. Обучаем ее и обрезаем лишние параметры
3. Возвращаем параметры к исходным значениям
4. Повторяем (2)-(3)



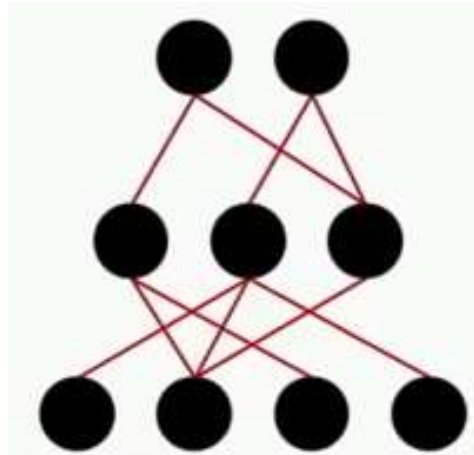
The Lottery Ticket Hypothesis

1. Создаем нейронную сеть со случайными весами
2. Обучаем ее и обрезаем лишние параметры
3. Возвращаем параметры к исходным значениям
4. Повторяем (2)-(3)



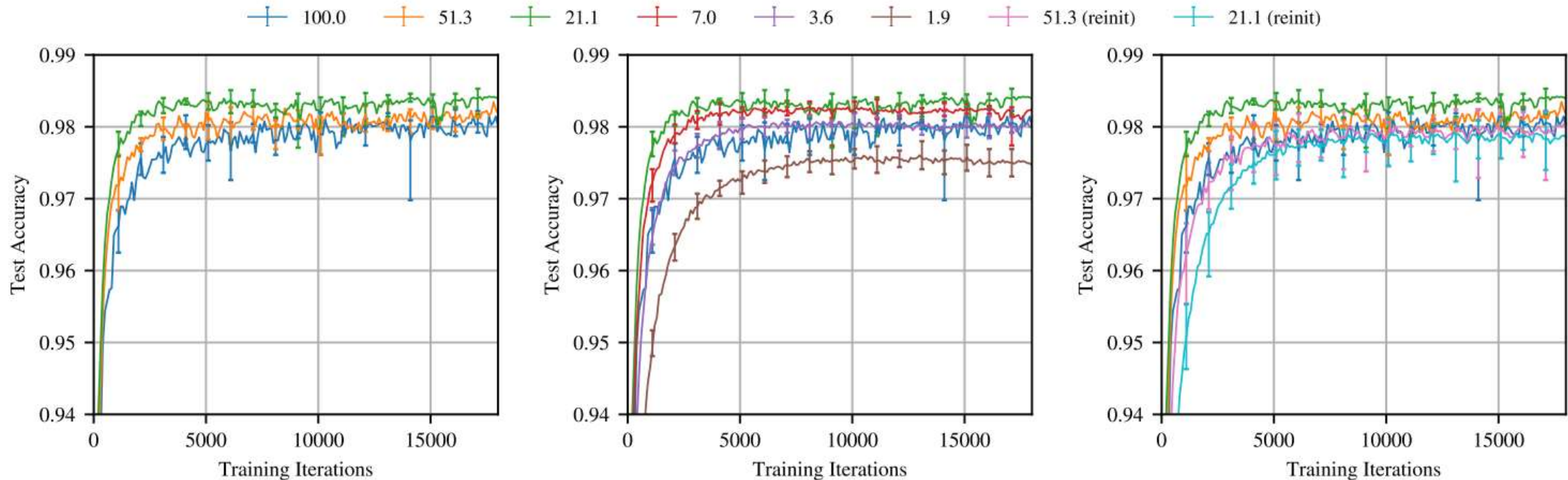
The Lottery Ticket Hypothesis

1. Создаем нейронную сеть со случайными весами
2. Обучаем ее и обрезаем лишние параметры
3. Возвращаем параметры к исходным значениям
4. Повторяем (2)-(3)



The Lottery Ticket Hypothesis

Получаем нейронную сеть малых размеров, 1-15% от исходного размера, которая быстрее обучается и достигает не меньшей точности.

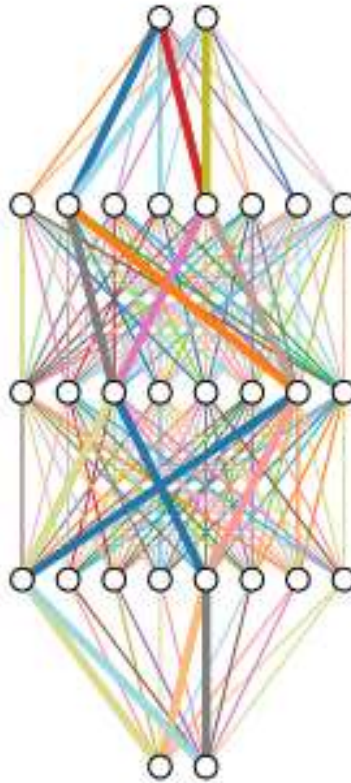


What's Hidden in a Randomly Weighted Neural Network?

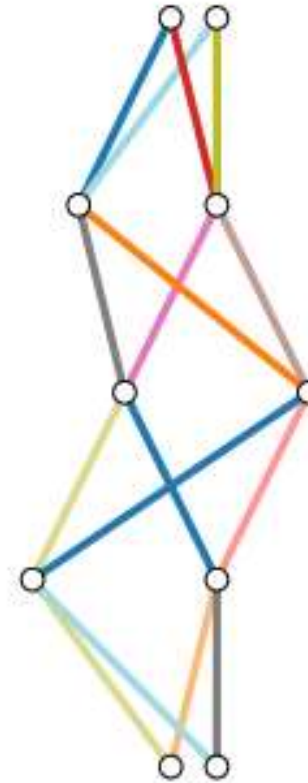
What's Hidden in a Randomly Weighted Neural Network?



A neural network τ which achieves good performance



Randomly initialized neural network N

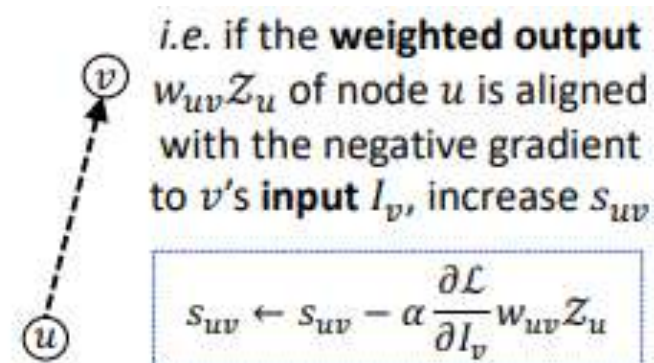
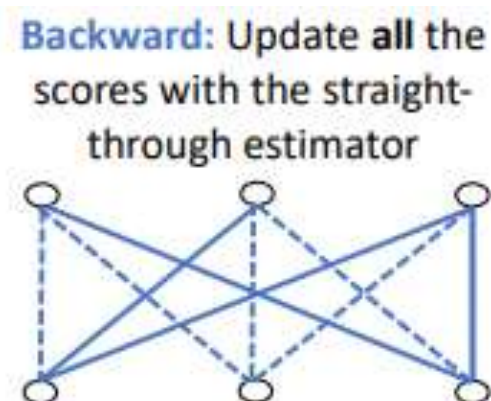
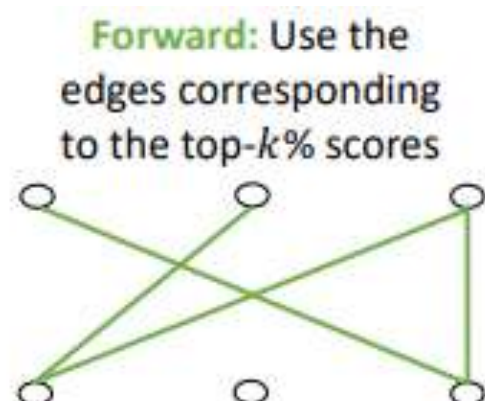
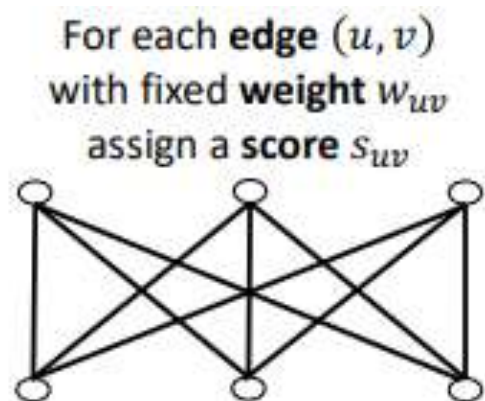


A subnetwork τ' of N

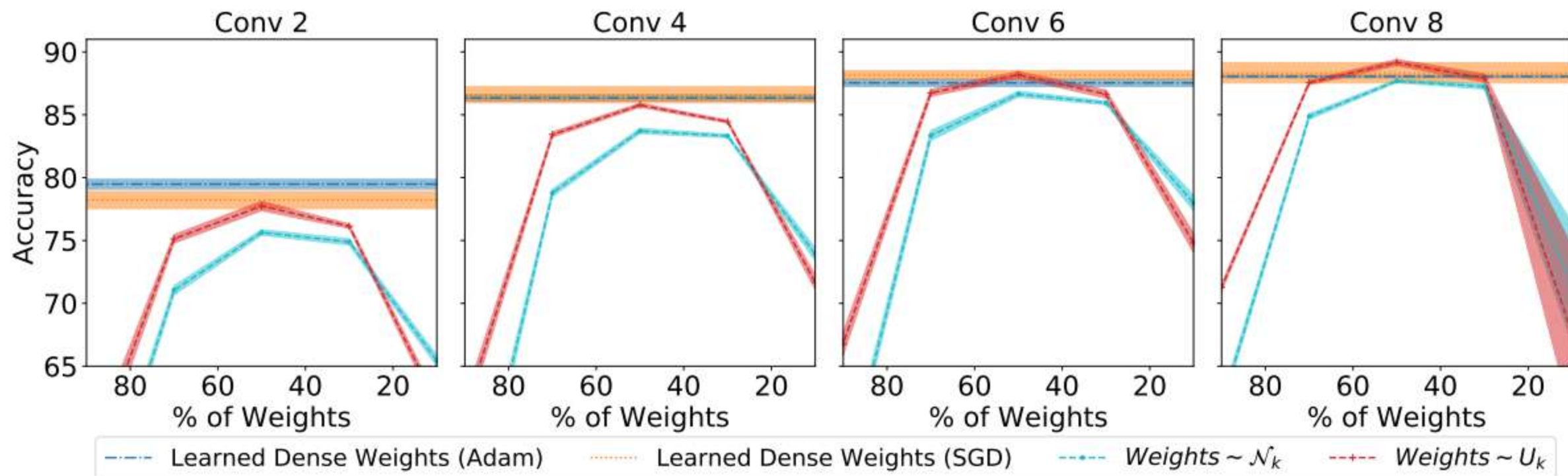
What's Hidden in a Randomly Weighted Neural Network?

- Xavier initialization: $std_i = \sqrt{\frac{2}{n_{i-1} + n_i}}$
- Kaiming initialization: $std_i = \sqrt{\frac{2}{n_{i-1}}}$

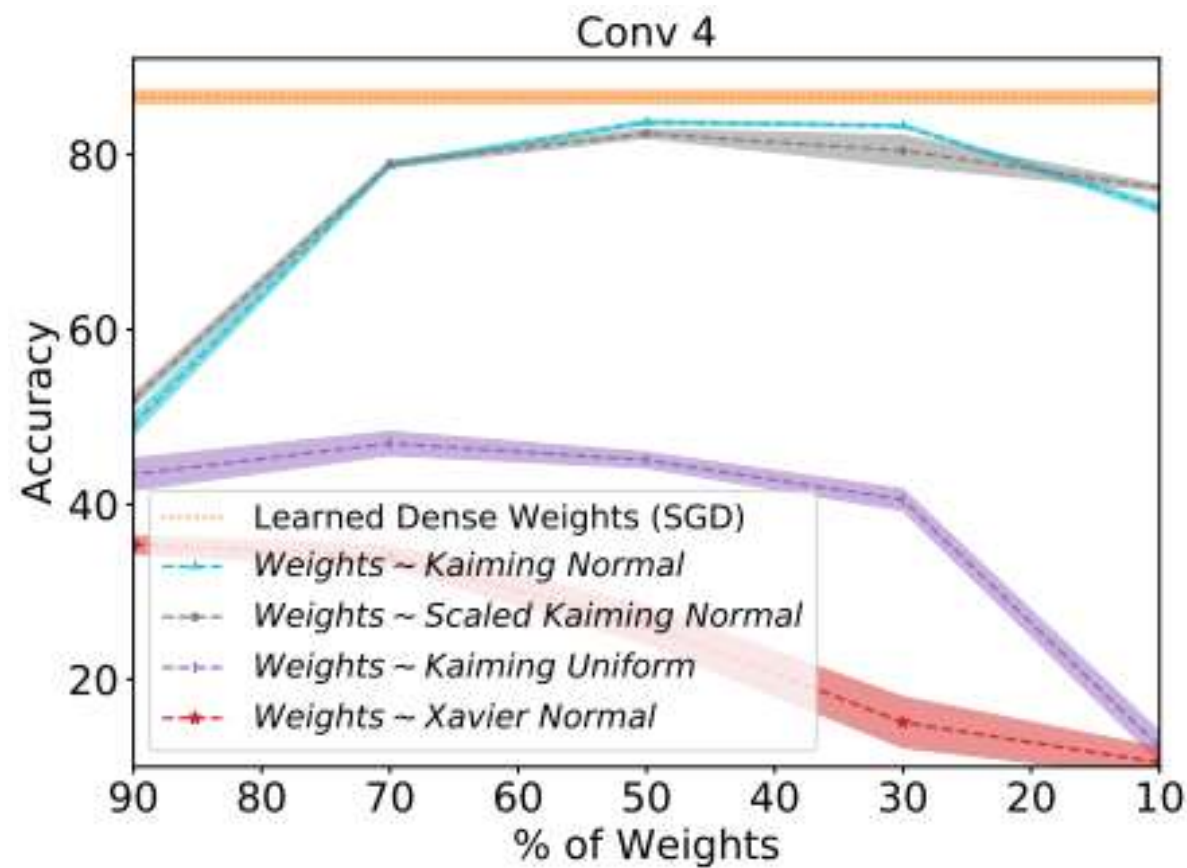
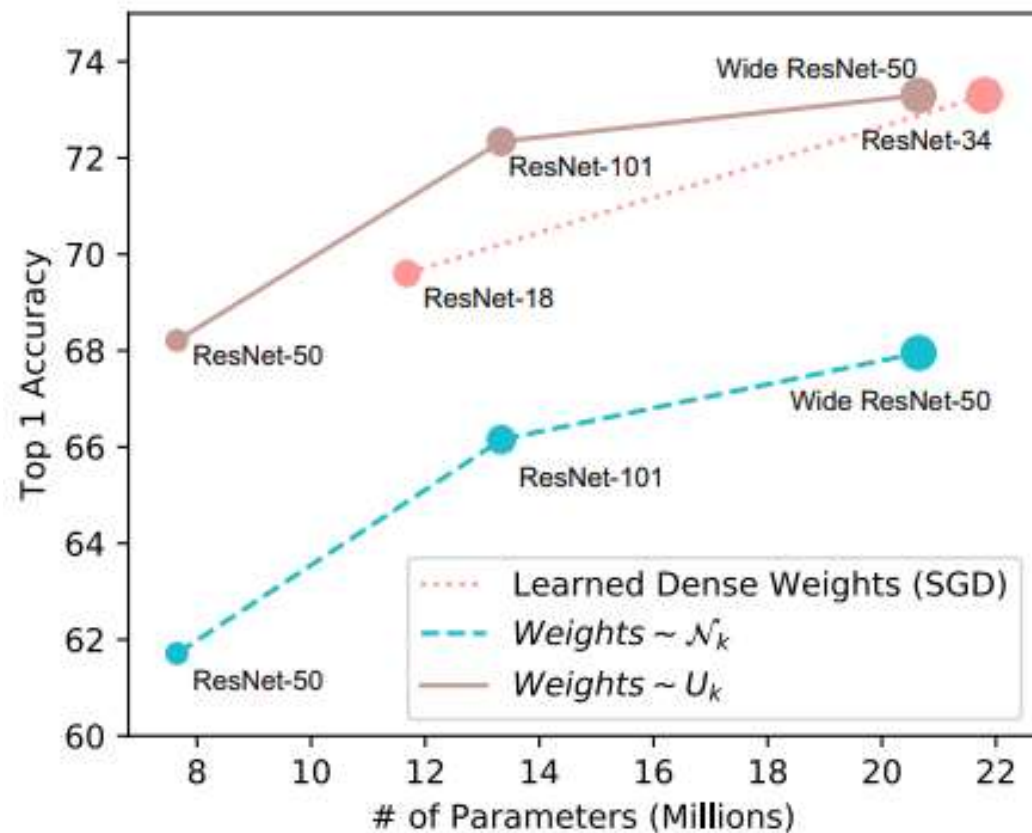
n_i - количество нейронов на i-ом слое



What's Hidden in a Randomly Weighted Neural Network?



What's Hidden in a Randomly Weighted Neural Network?



Вопросы

1. Сформулировать гипотезу The Lottery Tickets.
2. Алгоритм нахождения подсети в The Lottery Tickets Hypothesis
3. Алгоритм нахождения подсети показывающей хорошие результаты без обучения.

ИСТОЧНИКИ

- The Lottery Ticket Hypothesis: Finding Sparse, Trainable Neural Networks (2019)
- One ticket to win them all: generalizing lottery ticket initializations across datasets and optimizers (2019)
- What's Hidden in a Randomly Weighted Neural Network? (2019)