



Advanced Computer Vision

Практический курс

Савельева Юлия Олеговна, 3 семестр, 10.12.2020

Knowledge Distillation

Обучить нейронную сеть **Teacher** с двумя скрытыми слоями в 1200 нейронов на MNIST (ориентируйтесь на точность 96-97%)

Обучить нейронную сеть **Student** с двумя скрытыми слоями в 600 нейронов на MNIST (ориентируйтесь на точность 93%)


Данные точности получены с препроцессингом, который состоит только из нормализации со средним 0.5 и стандартным отклонением 0.5, а так при обучении обеих сетей **НЕ** использовалась регуляризация с помощью dropout и weight decay.

Knowledge Distillation

Реализовать целевую функцию из лекции по Knowledge Distillation:

$$L = L_{cls} + \lambda T^2 C = - \sum_{k=1}^N t^{(k)} \log \tilde{q}_k - \lambda T^2 \sum_{k=1}^N p_k \log q_k$$

$\tilde{\mathbf{q}} = \text{Softmax}(\mathbf{z})$ $\mathbf{q} = \text{Softmax}_T(\mathbf{z}, T)$



Свой градиент писать **НЕ** нужно!

Knowledge Distillation

Используйте уже обученную сеть **Teacher**, чтобы использовать ее предсказания в целевой функции.

Ориентируйтесь на описание обучения на MNIST в статье:

<https://arxiv.org/pdf/1503.02531.pdf>

В ходе экспериментов подберите параметры λ и T так, чтобы точность сети **Student2.0** стала выше и приблизилась к точности сети **Teacher** (ориентир 96%). Если точность при обучении не растет или даже падает, то параметры подобраны неверно.

Knowledge Distillation

Дополнительное задание

Посчитайте количество ошибок на классе «3» для сети *Student2.0*.

Опираясь на эксперименты статьи, исключите класс «3» из обучающей выборки, и обучите сеть *Student2.0(-3)*.

Посчитайте количество ошибок на классе «3» для сети *Student2.0(-3)*.

Увеличьте *bias* для последнего полносвязного слоя у выхода, соответствующего классу «3». Увеличьте его так, чтобы это не приводило к неправильной классификации других классов.

Посчитайте количество ошибок на классе «3» для сети *Student2.0(-3)*. Количество ошибок должно уменьшиться и приблизиться к количеству ошибок сети *Student2.0*.

Дедлайн 20.12.2020 03:00