**Тема 2.4. Нормализация – еще о регуляризации**

**План:**

1. Пакетная нормализация
2. Особенности батч-нормализации
3. Другие методы нормализации

**Пакетная нормализация**

В 2015 году была предложена техника пакетной нормализации – батчнормализации (batch normalization, batch- norm, технически будет правильней назвать межпакетная нормализация).

Идея метода батчнормализации заключалась в следующем. Различия статистических характеристик между такими мини пакетами замедляет обучение нейронной сети и повышает требования к точности подбора гиперпараметров и требования к инициализации весовых параметров – то есть повышает вероятность переобучения и снижает обобщающую способность. Под статистическими характеристиками авторы понимают в первую очередь среднее и дисперсию по каждому признаку для каждой подвыборки (мини-пакета) входных данных. Описанный эффект авторы назвали внутренний ковариационный сдвиг, internal covariate shift. Для решения проблемы ковариационного сдвига было предложено использование нормализации результатов работы каждого слоя нейронной сети. При этом статистические характеристики должны усредняться по всем мини-пакетам и использоваться при работе обученной нейронной сети. То есть батчнормализация ведет себя по-разному во время тренировки и во время работы.

**Особенности батч-нормализации**

Сам по себе принцип ковариационного сдвига не доказан до сих пор. Однако, известно, что батчнормалиазция работает. При использовании батч-нормализации сеть становится менее склонной к переобучению и менее чувствительна к выбору значений гиперпараметров. Однако, также сеть становится чувствительной к изменению размера батча или изменению входных изображений; также сеть имеет понижение регуляризационных свойств при небольшом размере пакета. Для корректной̆ работы батчнормализации рекомендуется использовать пакеты от 50−100 экземпляров данных.

**Другие методы нормализации**

Батчнормализация не всегда корректно работает для стратегии переноса обучения. Поэтому в ряде случаев предлагается заменять этот слой на альтернативы, например, нормализацию слоя или нормализацию весовых параметров сети.

**Резюме**

Использование батч-нормализации или ее аналогов остаётся сегодня наиболее популярных методом регуляризации в рамках архитектур сверточных нейронных сетей в задачах компьютерного зрения.