

Основы практического использования нейронных сетей.

Лекция 5. Настройка гиперпараметров НС.

Дмитрий Буряк.
к.ф.-м.н.
dyb04@yandex.ru

Гиперпараметры НС

- ☐ Параметры НС, которые не настраиваются в процессе обучения.
- ☐ Влияют на емкость (способность к обобщению), скорость обучения и объем требуемой памяти НС.
- ☐ Параметры, определяющие архитектуру НС (число и размеры слоев)
- ☐ Параметры алгоритма обучения

Определение значений гиперпараметров

☐ «Ручной»

- требует понимания влияния гиперпараметров на свойства и характеристики сети.

☐ Автоматический

- требует большой объем дополнительных вычислений.

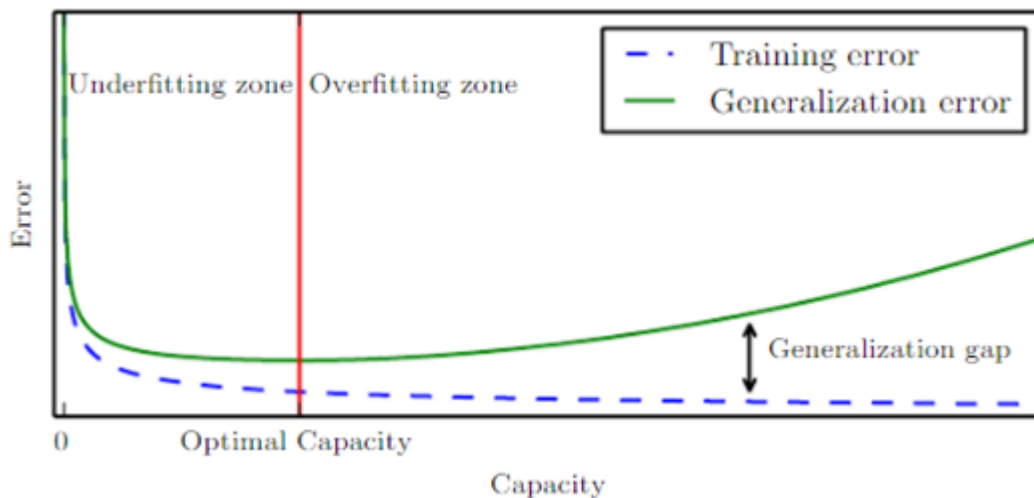
«Ручной» подбор значений гиперпараметров

- ❑ Цель – определить параметры НС, эффективная емкость которой соответствует сложности задачи.
- ❑ Емкость НС $\leftrightarrow$ Возможности алгоритма обучения $\leftrightarrow$ Переобучение.
- ❑ Эффективная емкость НС vs Емкость НС.

Гиперпараметр	Взаимосвязь с емкостью НС	Пояснение
Число нейронов во внутренних слоях	Емкость НС растет с ростом параметра	Увеличение числа внутренних нейронов ведет к росту числа параметров НС и, следовательно, емкости
Коэффициент обучения	Емкость НС растет, если параметр оптимален	Некорректный коэффициент обучения ведет к «ошибкам» алгоритма обучения и низкой емкости НС
Коэффициент регуляризации	Емкость НС растет с уменьшением параметра	Уменьшение коэффициента ведет к ослаблению дополнительных условий на веса сети и росту ее емкости.

Подтверждающая выборка

□ Анализ результатов на данных из подтверждающей выборки - один из основных способов настройки гиперпараметров



Интервалы сканирования гиперпараметров

- ❑ Масштаб поиска
 - оригинальный
 - логарифмический

- ❑ Типы гиперпараметров
 - непрерывный (коэффициент обучения, коэффициент регрессии и т.п.);
 - дискретный (количество нейронов);
 - бинарный (включение дополнительных этапов в процесс обработки)

Кросс-валидация

❑ Фиксированная подтверждающая выборка:

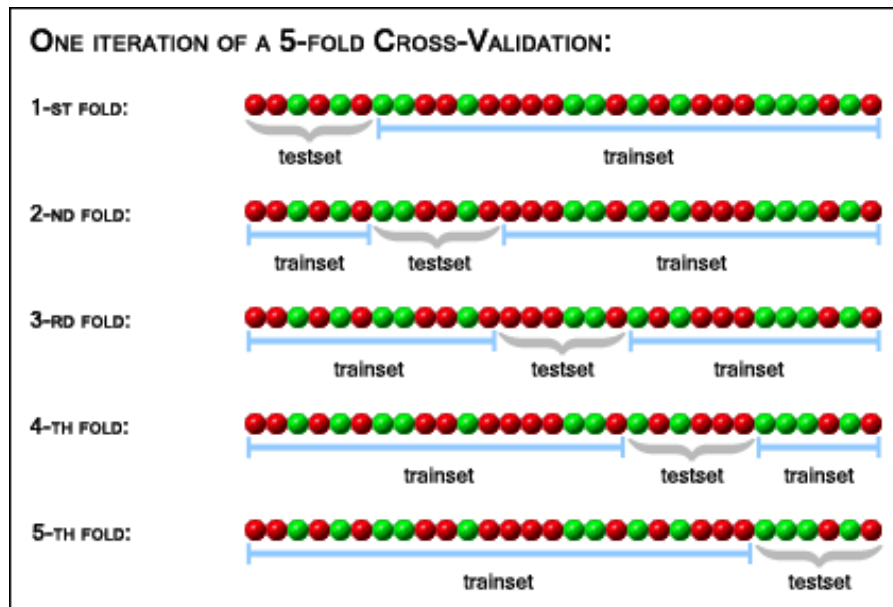
+ : меньший объем вычислений;

- : сильное влияние состава подтверждающей выборки.

❑ Кросс-валидация:

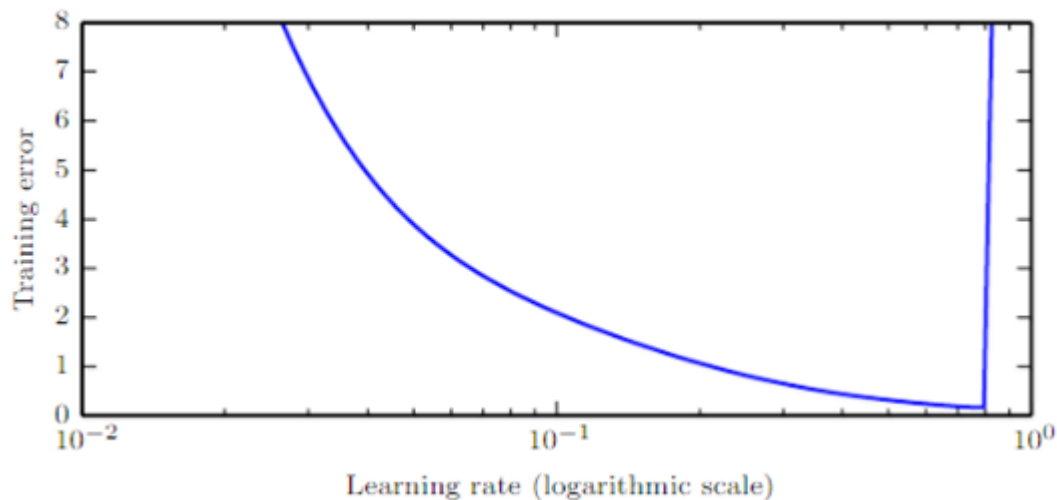
+ : эффективность, устойчивость;

- : большой объем вычислений.



Настройка коэффициента обучения

- ☐ Один из основных гиперпараметров.
- ☐ Цель – скорость уменьшения ошибки на обучающей выборке
- ☐ Требуется только обучающая выборка



Анализ результатов на обучающей и подтверждающей выборках

□ E_{train} – ошибка на обучающей выборке, E_{valid} – ошибка на подтверждающей выборке, E_{goal} – целевое значение ошибки.

□ $E_{train} > E_{goal} \rightarrow$ увеличение НС.

□ $E_{valid} > E_{goal}$

- $E_{valid} = E_{train} + \text{Gap}$;

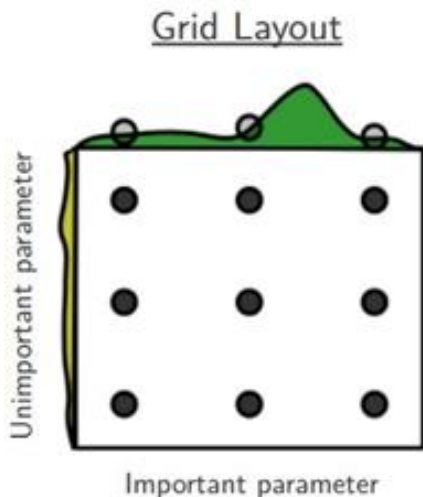
- уменьшение значения Gap (регуляризация, увеличение обучающей выборки).

Автоматический поиск значений гиперпараметров

- ❑ Создание «обертки» для тестирования различных наборов значений гиперпараметров и выбора оптимального.
- ❑ Функция оптимизации – значение ошибки на подтверждающей выборке.
- ❑ «Обертка» имеет собственные гиперпараметры.
- ❑ Подходы к формированию наборов значений гиперпараметров:
 - регулярный
 - случайный

Регулярное формирование наборов значений гиперпараметров

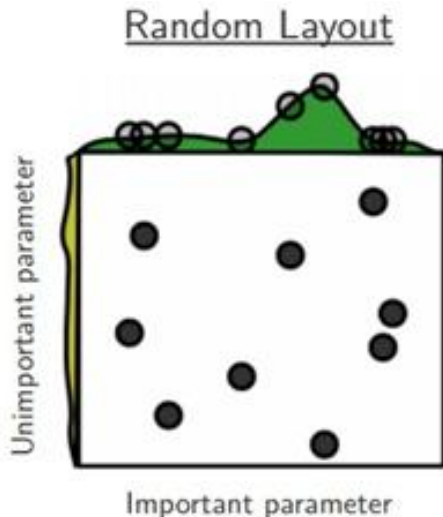
- ☐ Выбор фиксированного числа значений каждого гиперпараметра.
- ☐ Формирование комбинаций выбранных значений.



- ☐ Повторный поиск на границах интервала и на более частой сетке в выбранных областях.
- ☐ Низкая эффективная размерность пространства гиперпараметров.
- ☐ Различная значимость гиперпараметров на емкость НС.
- ☐ Экспоненциальный рост числа вычислений.
- ☐ Эффективен для малого числа гиперпараметров.

Случайное формирование наборов значений гиперпараметров

- ❑ Случайный выбор вектора значений гиперпараметров в соответствии с заданным распределением.

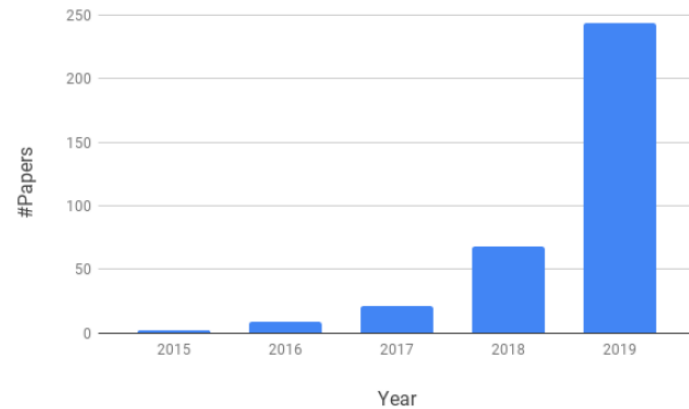


- ❑ Повторный поиск в выбранных областях.
- ❑ Устойчив к пространствам низкой эффективной размерности.
- ❑ Более эффективный метод относительно числа требуемых вычислений.

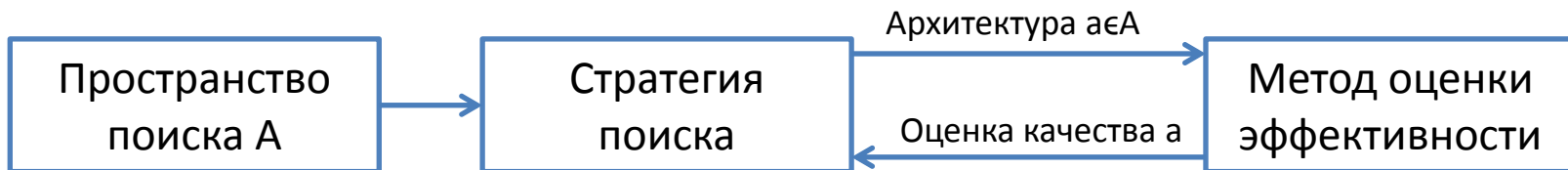
Автоматизированный поиск архитектуры НС (1)

□ Автоматизированный поиск архитектуры (NAS):

- Пространство поиска – среди каких архитектур искать.
- Стратегия поиска – метод поиска (оптимизации).
- Оценка эффективности – как оценивать качество архитектуры.



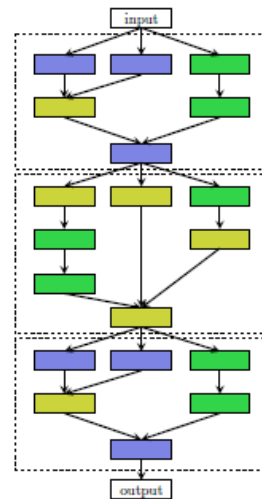
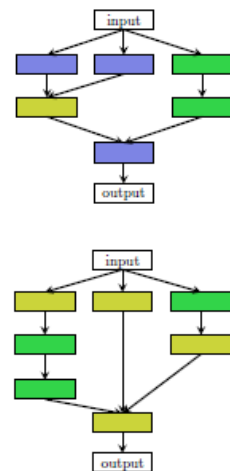
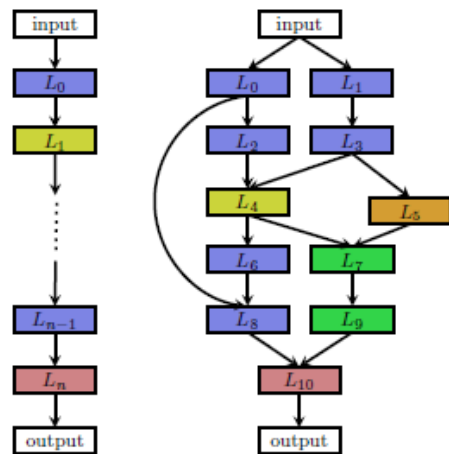
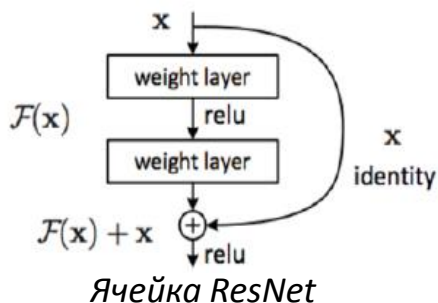
Число статей, посвященных NAS в год
(M. Lindauer et al. Best Practices for Scientific Research on Neural Architecture Search, 2020)



Автоматизированный поиск архитектуры НС (2)

□ Пространство поиска

- Пространство сетей со структурой цепочки.
- Оптимизация структуры ячеек (сеть строится из одинаковых ячеек): с сохранением и уменьшением размерности (например, ResNet).



Примеры ячеек и
результующей сети

Автоматизированный поиск архитектуры НС (3)

❑ Стратегия поиска

- Случайный поиск, Байесовская оптимизация, эволюционные методы, обучение с подкреплением, градиентные методы.

❑ Оценка эффективности

- Обучить – протестировать.
- Уменьшить число вычислений: ограничить время обучения, размерность и т.п.
- Экстраполяция кривой обучения.
- Инициализация весов на основе протестированных архитектур.
- One-Short Architecture Search – все архитектуры – подграфы суперграфа.

Примеры алгоритмов

❑ NEAT – Neuroevolution of Augmenting Topologies (2002)

- Для полносвязанных НС.
- Применение генетических алгоритмов.
- Кодирования архитектуры НС, реализация операций скрещивания и мутации.
- Расширения для глубоких НС: DeepNeat, CoDeepNeat.

❑ DARTS - Differentiable Neural Architecture Search (2019)

- Поиск «оптимальной» ячейки, НС строится из ячеек.
- Использование весов уже обученных сетей.
- Параметризация архитектуры → применение метода градиентного спуска.
- Вычислительная эффективность ~100х.