

Основы практического использования нейронных сетей.

Лекция 4. Основные этапы разработки НС.

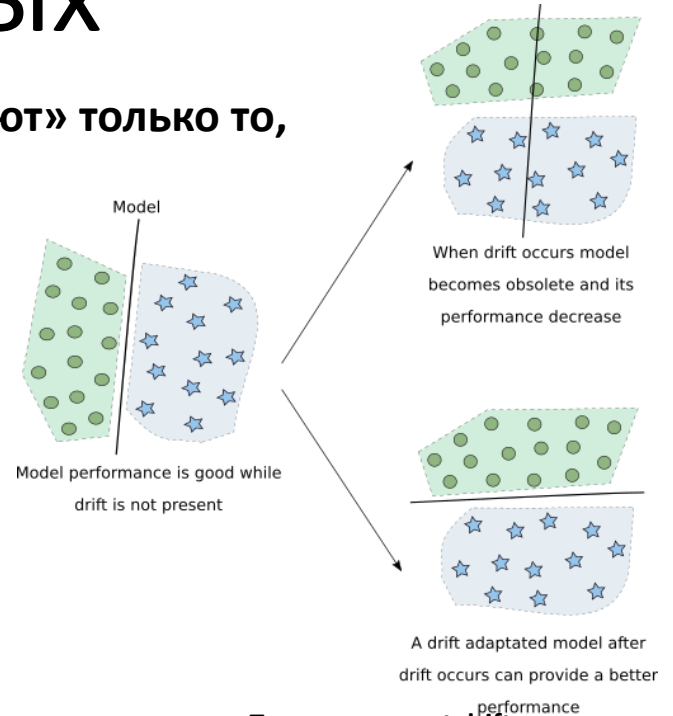
Дмитрий Буряк.
к.ф.-м.н.
dyb04@yandex.ru

Определение задачи

- ❑ Входные данные: 1D вектор, тензор, последовательность,...
- ❑ Тип задачи
 - Бинарная классификация: спам-детектор
 - Много-классовая классификация: распознавание команд
 - Классификация со множеством меток для каждого примера: поиск изображений
 - Регрессия: определение жизненных показателей человека (пульс, давление)
 - ...
- ❑ Выходные данные.

Сбор данных

- ❑ Алгоритмы машинного обучения «запоминают» только то, что было в обучающих данных.
- ❑ Основные свойства данных:
 - Объем
 - Репрезентативность: соответствие обучающих и продуктовых данных, concept drift, sampling bias, ...
 - Качество разметки
- ❑ Анализ данных: статистика меток, распределение элементов данных, ...



Пример concept drift

(<https://www.kdnuggets.com/2019/12/ravages-concept-drift-stream-learning-applications.html>)

Выбор целевых метрик

□ Показатели, которые определяют эффективность полученного решения.

Могут быть недифференцируемыми по весам НС.

□ Классификация

- Точность
- F1-мера
- Площадь под ROC

		POSITIVE	NEGATIVE
ACTUAL VALUES	POSITIVE	TP	FN
	NEGATIVE	FP	TN

$$Precision = \frac{TP}{TP + FP} \quad Recall = \frac{TP}{TP + FN}$$

$$Accuracy = \frac{TP + TN}{TP + FP + FN + TN}$$

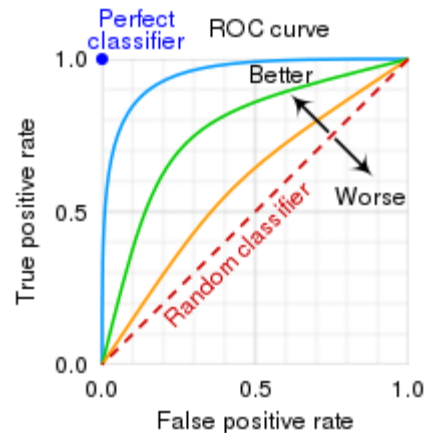
$$F1 \text{ Score} = 2 \times \frac{Precision \times Recall}{Precision + Recall}$$

□ Регрессия

- Близость по мере

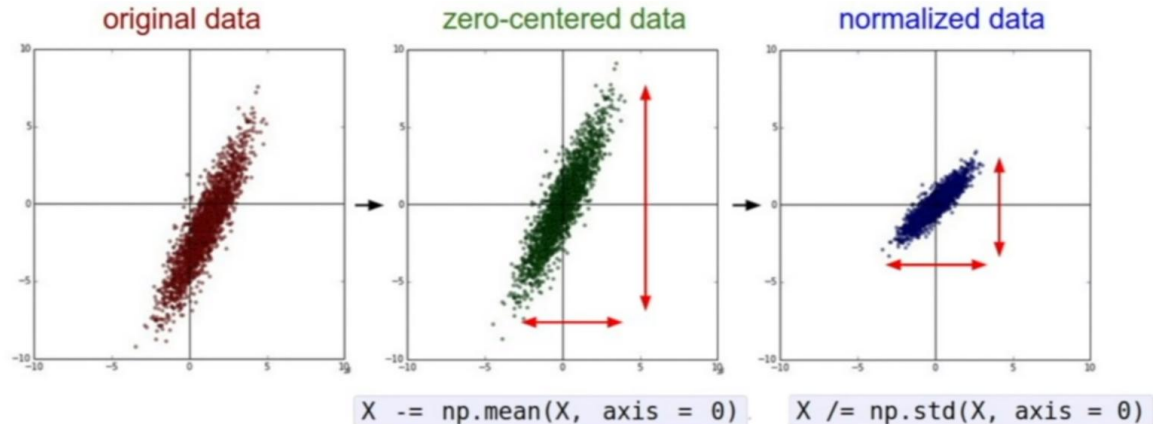
□ Выбор сценария оценки эффективности

- Использования валидационной выборки
- Кроссвалидация



Предобработка данных

- ❑ Векторизация данных
- ❑ Нормализация данных
 - Значения из небольшого интервала: $[0,1]$, $[-1,1]$, ...
 - Одинаковые диапазоны для каждого признака.
- ❑ Обработка пропущенных значений, выбросов



Подготовка к обучению НС

- ❑ Дизайн архитектуры НС
- ❑ Определение функции потерь

Тип задачи	Функция активации в последнем слое	Функция потерь
Бинарная классификация	Сигмоида	Binary crossentropy
Много-классовая классификация	Softmax	Categorical crossentropy
Множество меток на пример	Сигмоида	Binary crossentropy
Регрессия	Линейная	MSE

Основные причины низкой эффективности обучения НС

- Низкая эффективность = большая ошибка на тестовых данных.
- Проблемы с данными;
- Несоответствие архитектуры НС сложности задачи;
- Неоптимальные значения гиперпараметров;
- Переобучение;
- Ошибки в реализации.

Обозначения

- S_{train} – обучающая выборка; S_{test} – тестовая выборка;
- E_{train} – ошибка на обучающей выборке, E_{test} – ошибка на тестовой выборке, E_{goal} – целевое значение ошибки.
- $E_{test} > E_{goal}$

Анализ ошибки на обучающей выборке

❑ $E_{train} > E_{goal}$

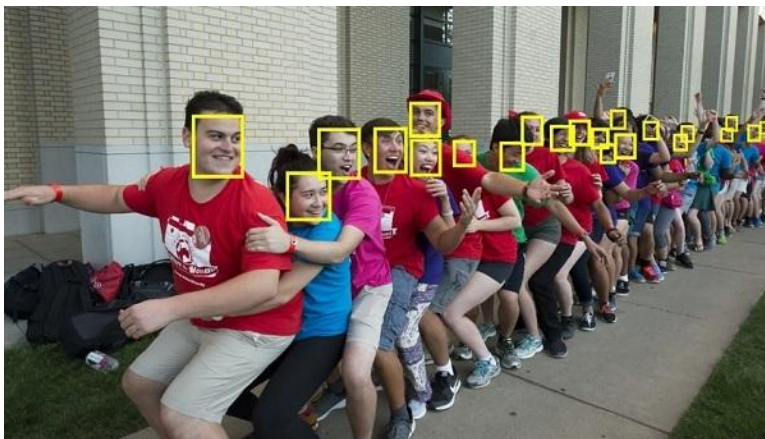
- ❑ Увеличить размер НС;
- ❑ Алгоритм обучения
- ❑ Оптимизировать значения гиперпараметров алгоритма обучения
- ❑ «Затухание» градиента: batch normalization, residual connection,...
- ❑ Анализ качества исходных данных
 - низкое значение сигнал-шум;
 - ошибки алгоритма предобработки;
 - недостоверные референсные значения;
 - несбалансированная выборка.

Анализ ошибки на тестовой выборке

- $E_{train} < E_{goal}$ и $E_{test} > E_{goal}$
- Увеличить размер S_{train} ;
- Уменьшить размер НС;
- Регуляризация, dropout, аугментация, ...;
- Подбор алгоритма обучения;
- Несоответствие S_{train} и S_{test} .

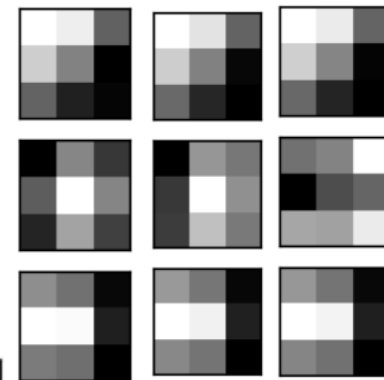
Анализ вычислений в НС

- ❑ Визуализация результатов:
 - согласованность статистических показателей и практики применения
- ❑ Визуализация, анализ результатов с наибольшими ошибками;



Анализ вычислений в НС (2)

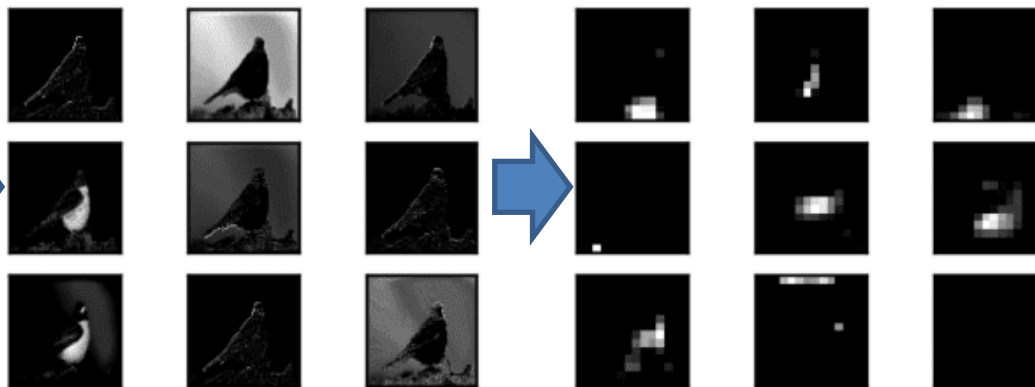
- Обучение на подвыборке меньшего размера;
- Анализ внутренних состояний сети:
 - матрица весов/фильтры сверточного слоя;
 - карты признаков;
 - гистограмма выходов нейронов;
- Deconvnet



Фильтры
первого слоя VGG16



Входное изображение
VGG16



Примеры карт признаков
1-го слоя VGG16

Примеры карт признаков
внутренних VGG16

Вопросы

- ❑ Основные свойства данных при их использовании для обучения и тестирования НС?
- ❑ В чем отличие функции потерь от целевой метрики, используемой для оценки эффективности НС?
- ❑ Что может быть причиной, когда ошибка на обучающей выборке превосходит целевое значение ошибки?
- ❑ Какие действия следует предпринимать, чтобы добиться уменьшения ошибки на тестовой выборке?