

Основы практического использования нейронных сетей.

Практические задания.

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

Практическое задание (1)

- Задача: бинарная классификация изображений,
классы: есть солнцезащитные очки, нет солнцезащитных очков
- Формат изображений: 30x32, grayscale, jpg
- Примеры изображений:



Практическое задание (1)

Размеры выборок:

- Обучающая: 378 изображений
- Тестовая: 22140 изображений

Возможная проблема: переобучение

Возможные способы решения:

- Упрощение архитектуры нейронной сети
- Оптимальный выбор момента остановки процесса обучения
- Увеличение обучающей выборки за счёт преобразований исходных изображений
- Обучение нескольких сетей и некоторое усреднение предсказаний

Практическое задание (1)

train_tags.csv

%Имя_%Фамилия.csv

id,tag

id,sunglasses_probability

0,sunglasses

0,0.08292

1,open

1,0.45391

2,sunglasses

2,0.18530

...

...

377,sunglasses

22139,0.86927

Метрика качества:

$$Accuracy = \frac{1}{n} \sum_{i=1}^n (pred_i == y_i)$$

if $pred_i > 0.5$ $pred_i = 1$, *else* 0

Практическое задание (1)

Необходимо до зачёта прислать архив:

1. Файл с предсказаниями для изображений в тестовой выборке (формат на предыдущем слайде)
2. Исходный код решения: при запуске должны воспроизводиться предсказания на тестовой выборке
3. Небольшой отчёт, содержащий:
 - a. Описание архитектуры использованной нейронной сети (или несколько архитектур)
 - b. Описание действий, предпринятых для борьбы с переобучением

Практическое задание (2)

- Задача: аппроксимация функции (регрессия)
 $F: \mathbb{R}^{256} \rightarrow \mathbb{R}; y = F(x)$
- Размеры выборок:
 - обучающая: 1000 пар $(x, y)_i$
 - тестовая: 100 векторов x_i .
- Возможная проблема: локальный минимум, переобучение.
- Способы решения:
 - анализ входных данных;
 - выбор архитектуры НС;
 - подбор параметров обучения;
 - выбор момента остановки обучения.

Практическое задание (2)

- Формат файла с обучающей выборкой (train_regression.csv):
 - 1000 строк
 - строка i : $x_{i1}, x_{i2}, \dots, x_{i256}, y_i$
- Формат файла с тестовой выборкой (test_regression.csv):
 - 100 строк
 - строка i : $x_{i1}, x_{i2}, \dots, x_{i256}$
- Формат файла с результатами на тестовой выборке (%Имя_%Фамилия_regression.csv):
 - 100 строк
 - строка i : y_i
 - порядок строк должен быть таким же как в test_regression.csv

Практическое задание (2)

Необходимо прислать (dyb04@yandex.ru) архив со следующими файлами

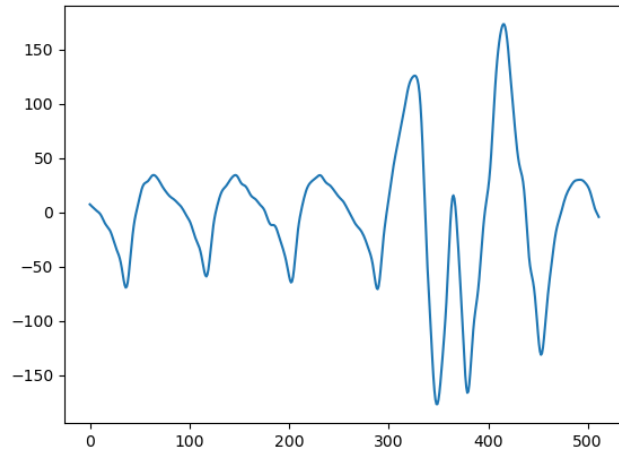
- Файл с результатами регрессии на тестовых данных (формат на предыдущем слайде);
- Исходный код решения;
- Отчет, содержащий:
 - описание архитектуры НС
 - что было сделано/исследовано для получения конечного результата.

Практическое задание (3)

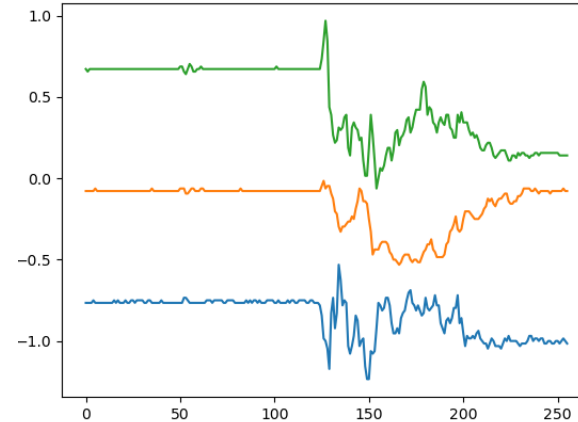
- Задача: оценка частоты сердечных сокращений (ЧСС) по сигналам оптического датчика и данным акселерометра.
- Сигналы получены с браслета, носимого на кисти.
В браслете установлены датчик PPG и 3-х осевой акселерометр.
- В ходе экспериментов пользователи выполняли различные активности.
Таким образом, сигналы PPG были искажены.
- Одновременно с записью сигналов с браслета выполняется регистрация ЭКГ, по которой вычислены истинные значения ЧСС.
Сигналы ЧСС усреднялись в окне длиной 8с.
- Необходимо разработать НС, которая вычисляет ЧСС по данным PPG и акселерометра.

Практическое задание (3)

- Входные данные:
 - сигнал PPG, длина 8с;
 - 3 сигнала акселерометра, длина 8с
- Выходные данные: значение ЧСС.



Пример сигнала PPG



Пример сигналов акселерометра

Практическое задание (3)

- Структура исходных данных:
 - Данные по каждому человеку находятся в отдельной папке: S1, S2, ...
 - Для каждого человека даны следующие файлы (в формате numpy):
 - bvr.npy – PPG сигналы, частота сэмплирования 64Гц
 - acc.npy – сигналы акселерометра, частота сэмплирования 32Гц
 - hr.npy – значения ЧСС (только для обучающих данных)
- Формат данных (N – число измерений для каждого человека)
 - bvr: numpy массив 512xN
 - acc: numpy массив 256x3xN (3 оси акселерометра)
 - hr: numpy массив N
- Формат результатов
 - numpy массив с вычисленными ЧСС, в том же порядке, как данные bvr и acc
 - для каждого человека из теста сформировать отдельный numpy файл.

Практическое задание (3)

Необходимо прислать (dyb04@yandex.ru) архив со следующими файлами

- Файл с результатами ЧСС на тестовых данных (формат на предыдущем слайде);
- Исходный код решения;
- Отчет, содержащий:
 - описание архитектуры НС
 - что было сделано/исследовано для получения конечного результата.