

Основы практического использования нейронных сетей.

Практические задания.

Задание 4

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

Задание 4 (1)

- Задача: Классификация на несколько классов
 - Распознавание голосовых ключевых слов: 10 слов + незарегистрированные слова (other).
 - Слова (в скобках номер класса): bed (0), cat (1), dog (2), eight (3), five (4), go (5), happy (6), left (7), nine (8), on (9), other (10).
 - Длина аудио записи для каждого слова 1 секунда
- Исходные данные:
 - Мел спектрограмма для каждой записи (40 коэффициентов, для 51 отсчета).
 - Исходные аудио файлы из обучающей выборки.
 - Обучающая выборка: от 500 до 4750 спектрограмм для слова.
 - Тестовая выборка: 300 спектрограмм для классов 0-9, 750 – для класса 10
- Цель:
 - построить НС для классификации на несколько классов.
 - провести классификацию данных из тестовой выборки.

Задание 4 (2)

- Предлагаемая схема решения:
 - 2D сверточная НС: вход - мел спектрограмма (2D), выход – 11 нейронов (softmax)
 - Функция потерь: функция правдоподобия (categorical cross entropy, sparse categorical cross entropy)
- Возможная проблема: несбалансированная выборка, переобучение, локальный минимум.
- Способы решения:
 - введение весов классов;
 - схема сэмплирования пакетов;
 - применение регуляризации весов, dropout;
 - выбор архитектуры НС;
 - подбор параметров обучения;
 - выбор момента остановки обучения.

- регуляризация
- dropout
- кросс-валидация

$$L = \sum_c w_c \sum_{i=1}^{N_c} l(x_i^c, y_i^c)$$

Задание 4 (3)

- Обучающая выборка:

- Мел спектрограммы записаны в файлах pickle (*.pkl)
- Файлы распределены по папкам, соответствующие классам
- Код чтение из файла:

```
import pickle  
with open(filename, 'rb') as fid:  
    data=pickle.load(fid)
```

- Доступны исходные аудио файлы. По желанию можно начинать работать с ними, например, увеличивать выборку добавлением шумов. Код вычисления мел спектрограммы:

```
import numpy as np  
import librosa  
d=(d-np.mean(d))/np.std(d) # d – аудио сигнал, считанный из wav файла  
data = librosa.feature.melspectrogram(y=d, sr=16000, n_fft=512, hop_length=320, window=640,  
    n_mels=40 , fmax=8000, dtype=np.float32)
```

Задание 4 (4)

- Тестовое множество собрано в одной папке
- Формат файла с результатами на тестовом множестве (%Имя_%Фамилия_task3.csv):
 - <имя_файла_pkl>, <номер_класса>
- Метрика точности: среднее значение F1-меры по всем классам – вычислить F1 для каждого класса и усреднить полученные значения.
(<https://habr.com/ru/company/ods/blog/328372/>,
https://scikit-learn.org/stable/modules/generated/sklearn.metrics.f1_score.html)

Задание 4 (5)

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

- Файл с результатами классификации на тестовых данных (формат на предыдущем слайде);
- Исходный код решения (keras/tensorflow, pytorch), который можно запустить;
- Отчет (doc, pdf), содержащий:
 - описание архитектуры НС
 - что было сделано/исследовано для получения конечного результата.
 - пояснения, почему было принято решение остановиться на этом варианте.