

Назад

Далее

Упражнение

Добавить страницу в закладки

Этот элемент курса проверен как 'Методы машинного обучения'
вес: 1.0

В данном упражнении вам предстоит решить уже знакомую задачу классификации изображений – отделить изображения кошек от изображений собак, используя ансамбль моделей на основе стекинга.

В [предложенном архиве](#) находится выборка, включающая в себя изображения кошек и собак (по 500 изображений). Имя каждого изображения, для удобства, имеет следующий формат: cat/dog_номер_изображения.jpg в зависимости от того, какое животное присутствует на изображении. Данная выборка используется для обучения классификатора и его оценки.

[Следующая выборка](#) предназначена для классификации новых объектов после построения классификатора.

Важно! Используйте версию библиотеки scikit-learn==0.23.0, также Вам потребуется библиотека opencv

```
!pip install --upgrade pip
!pip install imutils
!pip install opencv-python
!pip install --upgrade scikit-learn==0.23.0
```

Для работы с изображениями и получения их гистограмм — характеристик распределения интенсивности изображения, следует воспользоваться следующей функцией и библиотекой **cv2**:

```
In [3]: def extract_histogram(image, bins=(8, 8, 8)):
        hist = cv2.calcHist([image], [0, 1, 2], None, bins, [0, 256, 0, 256, 0, 256])
        cv2.normalize(hist, hist)
        return hist.flatten()
```

Для чтения изображений также воспользуйтесь библиотекой cv2: **cv2.imread()**

При считывании набора данных (картинок) из папки для построения гистограммы отсортируйте названия по алфавиту, например, так (иначе тренировочные наборы данных могут отличаться даже при фиксированном random_seed):

```
imagePaths = sorted(list(paths.list_images('train')))
```

Котов закодируйте единицей, а собак — нулями.

В задачах используйте реализацию алгоритмов из библиотеки sklearn:

```
from sklearn.ensemble import BaggingClassifier
from sklearn.tree import DecisionTreeClassifier
from sklearn.svm import LinearSVC
from sklearn.ensemble import RandomForestClassifier
from sklearn.linear_model import LogisticRegression
from sklearn.ensemble import StackingClassifier
```

Если вы считаете, что написали верный код, но ответы не принимаются системой проверки, отправьте ссылку на свой блокнот в Google Colab с правами на редактирование или *.ipynb файл на нашу почту dc@itmo.ru. Не забывайте представиться и указать свой логин openedu.

Для выполнения работы **настоятельно** рекомендуем использовать сервис [Google Colab](#). Именно в этом случае мы можем гарантировать, что система зачтет ответы в случае правильного выполнения задания!

Задание

10.0/10.0 points (graded)

1. Обучите базовые алгоритмы на исходном наборе данных (предварительно закодируйте классы: Cat — 1, Dog — 0):

- Классификатор с мягким зазором и параметрами: `C = 1.74` , `random_state = 220` , остальные параметры по умолчанию;
- Бэггинг деревьев принятия решений. Параметры дерева: `criterion = 'entropy'` , `min_samples_leaf = 10` , `max_leaf_nodes = 20` , `random_state = 220` , остальные параметры по умолчанию. Параметры бэггинга: `n_estimators = 19` , `random_state = 220` ;
- Случайный лес с параметрами: `n_estimators = 19` , `criterion = 'entropy'` , `min_samples_leaf = 10` , `max_leaf_nodes = 20` , `random_state = 220` , остальные параметры по умолчанию.

2. Обучите метаалгоритм — логистическая регрессия: `solver='lbfgs'` , `random_state = 220` , остальные параметры по умолчанию. Оцените его точность при `cv = 2` .

3. Обучите модель стэкинга. Используйте 2-fold (`cv = 2`) кросс-валидацию для оценки.

Введите долю правильной классификации (Ассигасу):
Значение метрики округлите до сотых

0.851

0.851

Выполните предсказание для изображений, указанных ниже

Определите вероятность отнесения изображений к классу 1 с помощью обученного метаалгоритма. *Ответы округлите до тысячных.*

Файл cat.1009.jpg

0.474

0.474

Файл cat.1015.jpg

0.679

0.679

Файл dog.1014.jpg

0.676

0.676

Файл cat.1028.jpg

0.485

0.485

Отправить

Вы использовали 10 из 10 попыток

Назад

Далее

© Все права защищены

Каталог курсов
Каталог программ
Направления подготовки

О проекте
Вопрос-ответ
Задать вопрос
Системные требования

Пользовательское соглашение
Контактная информация
Контакты для СМИ
Политика в отношении перс. данных

POWERED BY
OPENedX®

Ru | En

Подписаться на новости
Открытого образования России

© 2022 Открытое образование

Введите ваш e-mail

Подписаться