

Определение видов птиц по записи их пения

Выполнили: Тин Павел, Папа-Дмитриева Юлия

17 мая 2022 г.

1. Аннотация

Для решения поставленной задачи методами машинного обучения, кратко охарактеризованной в названии, использовали Случайный лес и Градиентный бустинг. Дополнительно представили визуализации признаков и проанализировали их на предмет экстракции большего количества признаков для обучения моделей. Звуковые дорожки искусственно разбивались на короткие двухсекундные записи, вырезали тишину. Полученных данных получилось много, но из-за особенностей предобработки звука имели характерные повторяющиеся шумы. Для сложения задачи моделям были внесены дополнительные шумы. Дополнительно провели проверку работы модели Случайного леса с разными параметрами, установили наилучшие. Градиентный бустинг показал большую точность, но занял больше времени для обучения.

2. Введение

Основной целью проекта является сборка датафрейма с записями пения птиц России и построение определителя животных до вида. Решаемая задача может быть востребована у орнитологов, локально изучающих ареалы обитания разных родов и видов птиц.

Далее мы рассмотрим поэтапное решение поставленной задачи. Определимся с необходимыми библиотеками, выделим несколько видов, найдём интересующие соответствующие видам аудиодорожки, получим признаки и внесём их в общую базу и обучим модели.

3. Понятия и обозначения

1. Термины

Спектрограмма – изображение, показывающее зависимость спектральной плотности мощности сигнала от времени. двумерная диаграмма: на горизонтальной оси представлено время, по вертикальной оси — частота; третье измерение с указанием амплитуды на определенной частоте в конкретный момент времени представлено интенсивностью или цветом каждой точки изображения.

Фурье-преобразование – преобразование, позволяющее представить непрерывную функцию (сигнал) в виде бесконечного ряда тригонометрических функций и дающее частотный спектр.

Мел-кепстральные коэффициенты – небольшой набор признаков (обычно около 10–20, в нашем случае 20), которые кратко описывают общую форму спектральной огибающей.

2. Пакеты и библиотеки

Librosa – пакет Python для анализа музыки и видео. С его помощью мы получаем необходимые признаки для рандомного дерева.

Sklearn – пакет Python для машинного обучения.

Pydub – библиотека Python для обработки звуковых дорожек. С её помощью мы переконвертировали mp3 файлы в ogg для дальнейшей работы в *librosa*.

Audiosegment – модуль Pydub.

Pandas – библиотека для обработки и анализа данных.

Numpy – библиотека для быстрой работы с массивами.

Seaborn – библиотека для визуализации данных.

3. Названия птиц

Ardea cinerea – серая цапля.

Emberiza citrinella – обыкновенная овсянка.

Erithacus rubecula – зарянка.

Fringilla coelebs – зяблик.

Nycticorax nycticorax – обыкновенная кваква, или ночная цапля.

Oenanthe oenanthe – каменка.

Forpus passerinus – домовый воробей.

Pyrrhula pyrrhula – обыкновенный снегирь.

4. Ход работы

В первую очередь нам потребовалось создать базу данных с признаками. Для этого брали аудиодорожки длительностью ≤ 2 секунд удаления абсолютной тишины и путём разбиения получившихся сегментов.

1. Получение аудиодорожек

В дополнение к каждому mp3 файлу с записью голоса шёл файл в формате json, в котором были размечены начало и конец участков не с полной тишиной. Это позволило нам работать уже с имеющими смысл звуковыми участками и разрезать их на мелкие части по 2 секунды, при этом оставшиеся участки длиной ≤ 2 сек были также оставлены. Здесь мы сразу сохраняли наши дорожки в формате mp3 в ogg. Аудиодорожки этого формата весят примерно в 2 раза меньше, чем mp3, и поддерживаются пакетом *librosa*.

2. Получение признаков и создание DataFrame

После того, как база аудиодорожек была готова, мы приступили к выделению признаков с помощью функций пакета *librosa*. В качестве последних взяли спектрограмму, преобразование Фурье и мел-кепстральные коэффициенты. Первые две функции на выходе давали распределение по 1025 частотам для каждой порезанной дорожки, т.е. по 1025 признаков, а мел-кепстральные

коэффициенты - 20 признаков. В сумме получили 2070 признаков, которые соединили в один DataFrame.

3. Обучение моделей

Обучение модели на признаках проводили с использованием 2-х алгоритмов: случайный лес и метод градиентного спуска. Первый метод дал точность 0.84 при глубине, равной 8, в то время как второй, при вдвое меньшей глубине, показал точность выше -0.90.

4. Оценка результатов

Модели работают, но в связи с тем, что мы меняли их параметры вручную, результат кажется искусственным. Попробуем узнать наилучшие параметры. Используя GridSearchCV, выяснили, что наилучший результат может быть получен при максимальной глубине деревьев 9, признаке расщепления sqrt и количестве деревьев 93. Подставив наилучшие параметры, добились точности 0.86, которая всё равно меньше, чем у градиентного бустинга. Визуализируем полученный результат с помощью матрицы корреляции. Из неё видно, что достоверные результаты получены для всех птиц, кроме 9-й, которую модель путает с 3-й.

5. Выводы

Градиентный бустинг работает сильно дольше, но даже при меньшей глубине деревьев даёт высокую точность. Чтобы добиться примерно похожей точности от случайного леса потребовалось увеличить глубину деревьев в два раза.