

Анализ методов очистки аудиозаписей при помощи нейронной сети с различной архитектурой

Авторы:

Удовин Илья Владимирович,
Щербаков Андрей Сергеевич
группа Б05-874, ФУПМ МФТИ



План

1. Введение
2. Распределение задачи и Структура программы
3. Физика шума и алгоритмы очистки
4. Анализ шума с помощью нейронных сетей
 - а. Классификация и распознавание
 - б. Немного о реализации
5. Демонстрация
6. Что дальше?



Почему, зачем и как?

- **Мотивация:** часто при записывании аудиозаписей большую проблему доставляют посторонние шумы, что очень влияет на качество и восприятие звука. Для повышения качества звука были придуманы определенные алгоритмы, очищающие запись, но их эффективность в “полевых” условиях неясна.
- **Цель:** научиться классифицировать шумы, определить, какие из наиболее популярных алгоритмов очистки, наиболее эффективны на разных типах шумов.
- **Средства:**
 - для классификации : нейронная сеть с архитектурой LSTM (+ небольшой анализ 2 других архитектур)
 - для очистки: различные алгоритмы, не использующие машинного обучения



Ход проекта и распределение задач

1. Щербаков Андрей:
 - а. Физика шума и алгоритмы очистки, язык: python
 - б. Архиватор проекта, презентация, отчет

2. Удовин Илья:
 - а. Разобраться с нейронными сетями с архитектурой LSTM
 - б. Написать классификатор, определить как оценивать “степень зашумленности” при помощи ML
 - с. Интерфейс и запуск программы



Структура программы

- Обучение классификатора:
 - Выбор базы данных для обучения (UrbanSound8k.csv в нашем случае), разбивка на классы
 - Первичная очистка и подготовка базы для обучения
 - Модель и Обучение, отображение степени зашумленности
- Выбор оптимального алгоритма очистки для каждого типа шума:
 - На обученной модели на той же базе данных смотрим степень зашумленности до и после очистки разными способами, выбираем наиболее оптимальный
- Анализ аудиозаписи:
 - Загрузка аудиозаписи через интерфейс
 - Классификация шумов в аудиозаписи
 - Очистка разными способами и их комбинациями
 - Анализ степени зашумленности в зависимости от различных методов очистки, выбор наиболее оптимального метода
 - Сохранение наиболее очищенной записи в файл



Физика шума и алгоритмы очистки

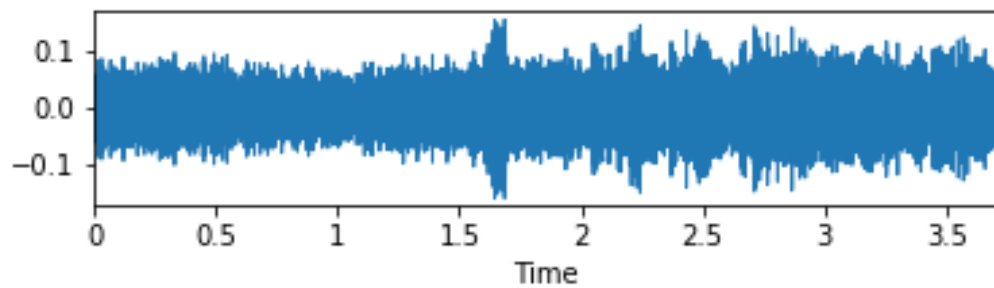
NOISE REDUCTION USING POWER:

$$\text{Centroid} = \frac{\sum_{n=0}^{N-1} f(n)x(n)}{\sum_{n=0}^{N-1} x(n)}$$

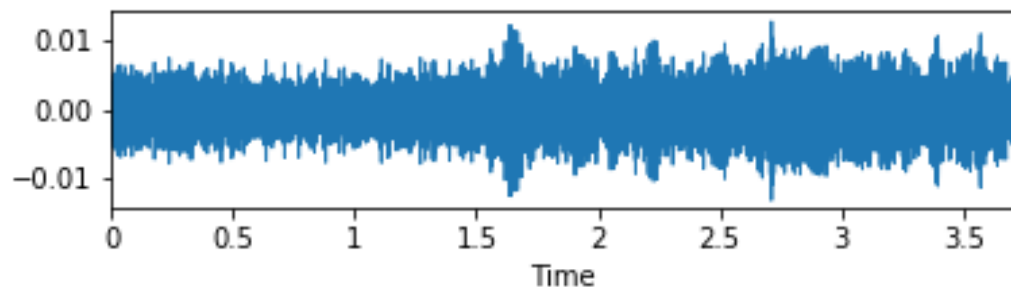
f — частота, x — спектральное значение

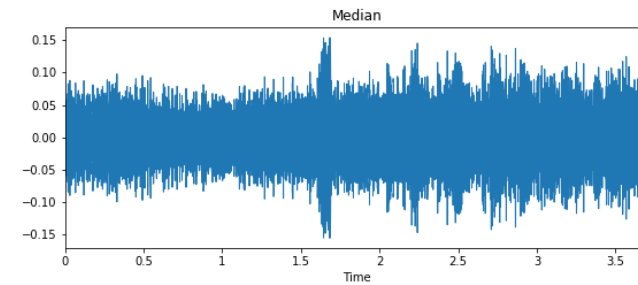
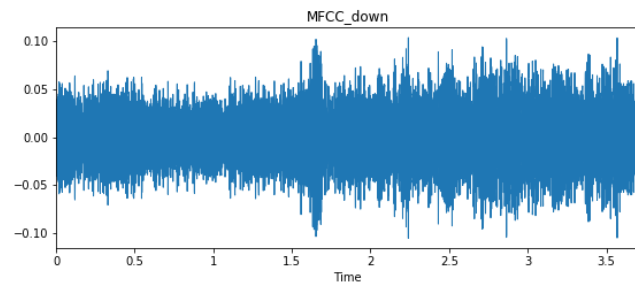
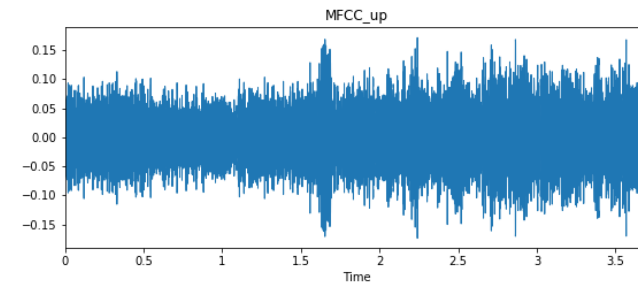
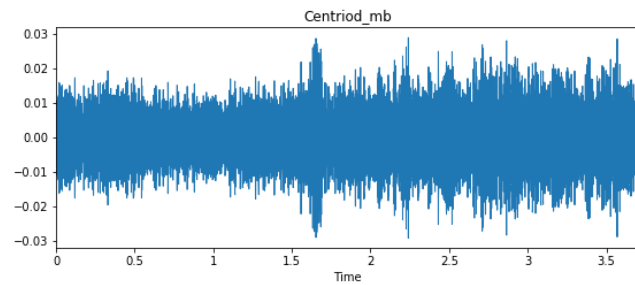
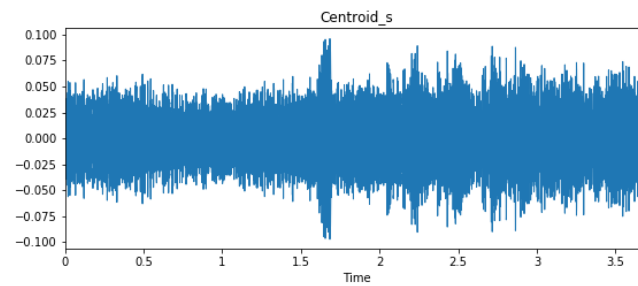
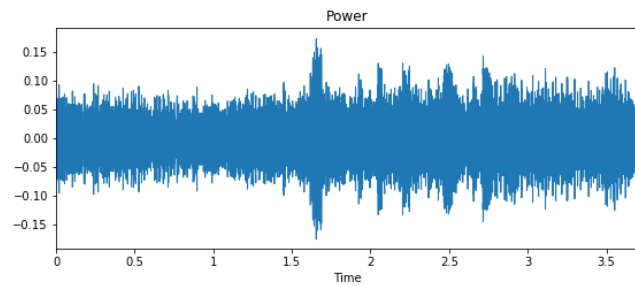


Before



After





Теперь о Классификации





Основные источники шумов на записи

кондиционер

двигатель

сигнал автомобиля

выстрел

крики детей

стройка

лай собаки

сирена

дрель

уличная музыка



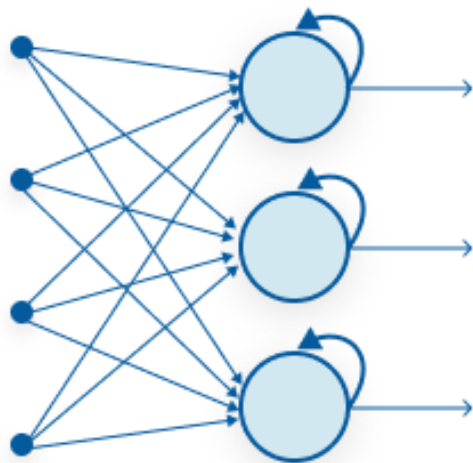
Задача распознавания

Каждому классу ставится в соответствие вещественный признак, принимающий значение от 0 до 1.

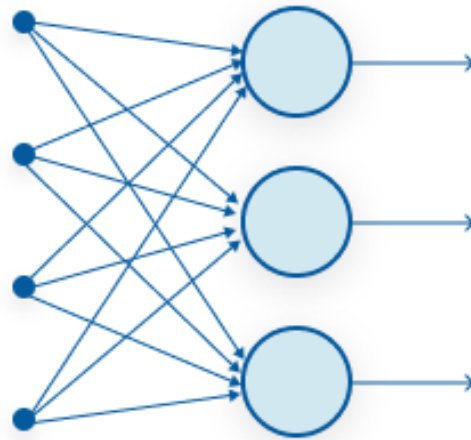
Для произвольной записи вычисляется вектор признаков. Значение каждого элемента — вероятность того, что шум на записи принадлежит данному классу.

Рекуррентные нейронные сети


Recurrent Neural Network structure



Recurrent Neural Network



Feed-Forward Neural Network



LSTM модель

(Long short-term memory)

Использует модули, запоминающие значения как на короткие, так и на длинные промежутки времени.

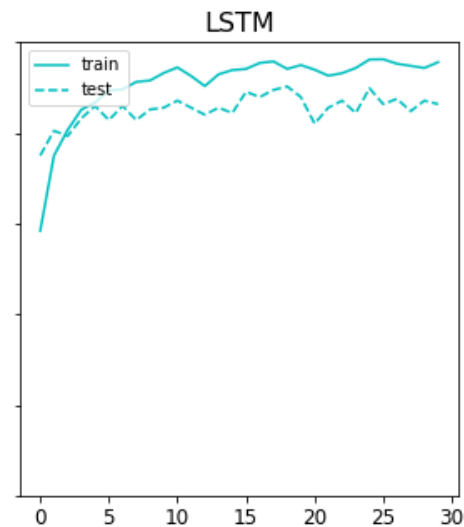
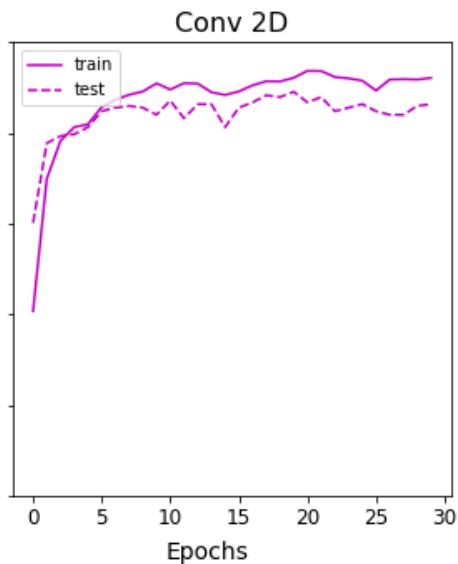
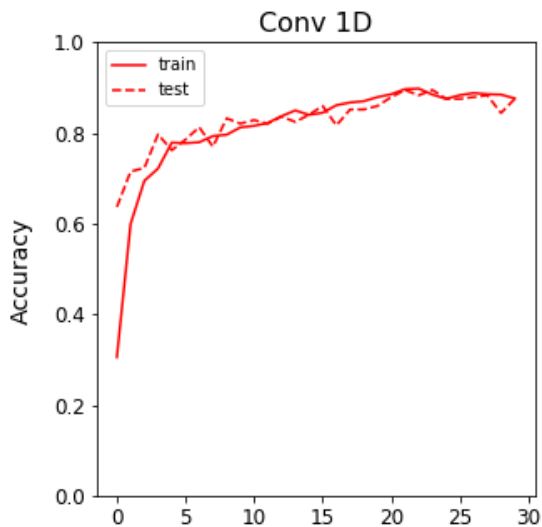
Модули группируются в блоки.

Подходит для задачи классификации. В нашем случае — выше точность и ниже ошибка по сравнению со сверточными нейронными сетями.

Немного о результатах

T = true, F = false, P = positive, N = negative

$$\text{accuracy} = \frac{TP + TN}{TP + TN + FP + FN}$$



Демонстрация



Что дальше?



Спасибо за внимание!

