

# РАСПОЗНАВАНИЕ ЭМОЦИЙ ПО ГОЛОСУ

Казаковцев В.Л.

Шабанов Д.А.

Глухов И.А.

Михайлов А.И.

# Выбор программных средств

- Почему Python?
- Список используемых библиотек:
- NumPy
- Pandas
- TensorFlow
- Skit-Learn
- Librosa

# Кодирование звуковой информации

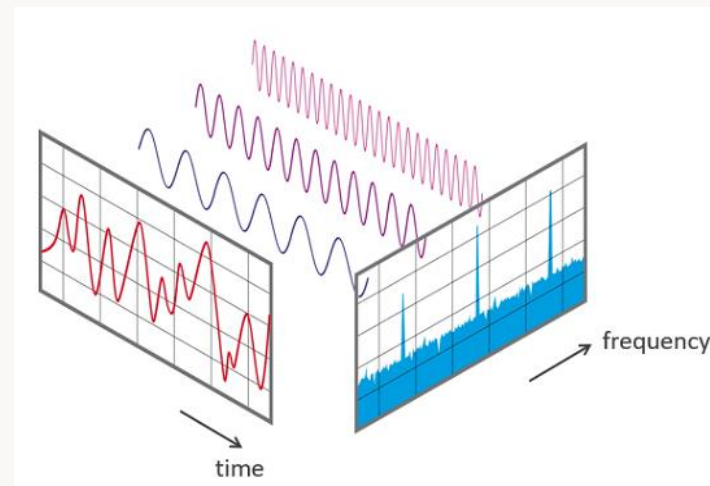
## МЕЛ

Человеческое ухо более чувствительно к изменениям звука на низких частотах, чем на высоких

Единица измерения высоты звука — мел:

$$mel = 1127.01048 \ln \left( 1 + \frac{freq}{700} \right)$$

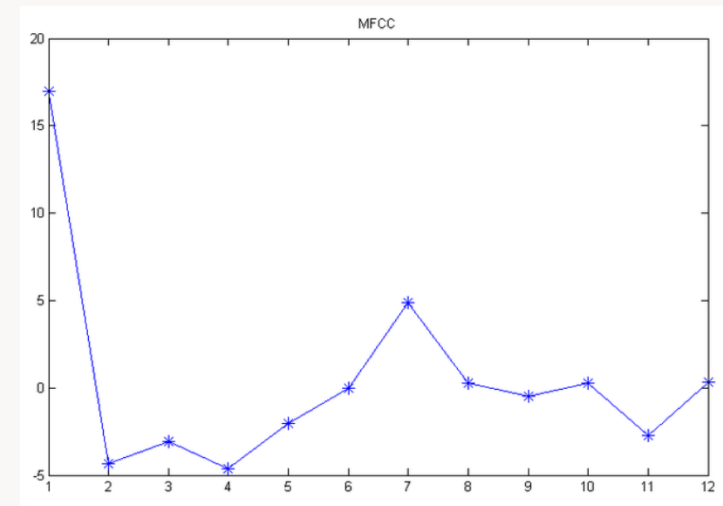
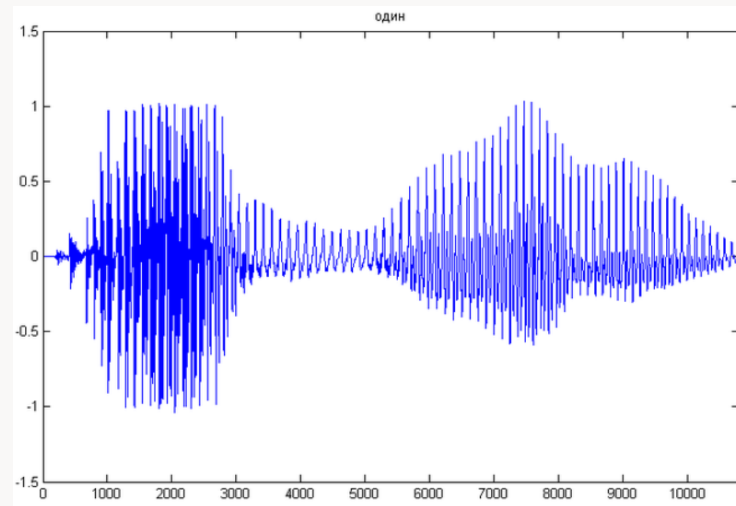
## АУДИОСИГНАЛ: ФУНКЦИЯ АМПЛИТУДЫ И ВРЕМЕНИ



# Мел- кепстральные коэффициенты

## ПРЕОБРАЗОВАНИЕ

- К записанному в дискретном виде сигналу применяем преобразование Фурье
- Составляем гребенку фильтров, используя оконную функцию
- Применяем дискретное косинусное преобразование

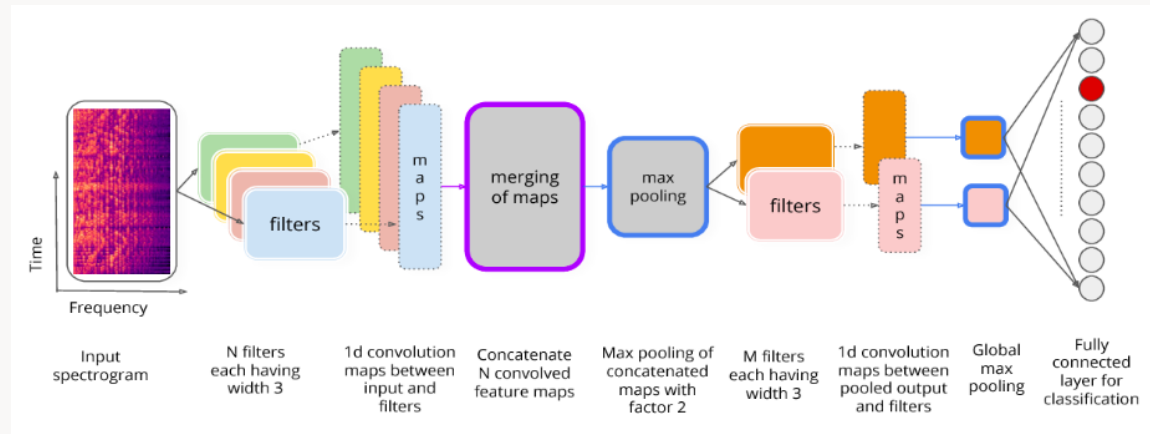


# Выбор модели

## СВЁРТОЧНАЯ НЕЙРОННАЯ СЕТЬ

- Слой свёртки
- Слой активации
- Слой субдискретизации

## СХЕМА



# Результаты

## МЕТРИКИ КАЧЕСТВА

Среднее значение *accuracy*: 91.4

Среднеквадратическое отклонение *accuracy*: 2.6

## ПРИМЕР ИСПОЛЬЗОВАНИЯ

```
In [1]: from voice_recognition import Recognizer
rec = Recognizer()
rec.recognize(['VoiceExamples/angry_man.wav',\
              'VoiceExamples/sad_vladimir.wav',\
              'VoiceExamples/calm_Daria.wav'], start_from=0.5)

1/1 [=====] - 0s 5ms/step
1/1 [=====] - 0s 2ms/step
1/1 [=====] - 0s 1ms/step

Out[1]: {'VoiceExamples/angry_man.wav': 'male_angry',
         'VoiceExamples/sad_vladimir.wav': 'male_sad',
         'VoiceExamples/calm_Daria.wav': 'female_calm'}
```

СПАСИБО ЗА  
ВНИМАНИЕ