

ИЗВЛЕЧЕНИЕ И КЛАССИФИКАЦИЯ ФОРМАНТ МЕТОДАМИ МАШИННОГО ОБУЧЕНИЯ

Саламатов К.А.¹, Смирнов А.А.¹, Целищева А.В.¹

¹) Уральский федеральный университет имени первого Президента России Б. Н. Ельцина, Екатеринбург, Россия
E-mail: kirillsalamatov20@gmail.com

FORMANT EXTRACTION AND CLASSIFICATION BY MACHINE LEARNING METHODS

Salamatov K.A.¹, Smirnov A.A.¹, Tselishcheva A.V.¹

¹) Ural Federal University named after the first President of Russia B.N. Yeltsin,
Ekaterinburg, Russia

In this paper, the possibilities of studying formants and their prediction using different machine learning models are considered.

В современном мире растёт количество информации вследствие чего появляется проблема при работе с таким большим объёмом. В том числе имеется множество аудиозаписей голоса собранных в наборы данных: «VOICED Database» [1], «Common Voice» [2] взятый с сайта Kaggle и прочие. Из аудиозаписей голоса можно извлечь форманты. Форманта (от лат. formans, род. падеж formantis - образующий) - область усиленных частичных тонов в спектре муз. звуков, звуков речи, а также сами эти призвуки, определяющие своеобразие тембра звуков. Тембр есть окраска звука, индивидуальность голоса человека.

Данные для исследования формант будут извлекаться из записи голоса по разработанному алгоритму, после данные обрабатываются по определённому порядку действий и алгоритмами машинного обучения. С помощью алгоритмов машинного обучения можно извлечь признаки, которые могут быть интерпретированы человеком и использованы в дальнейших различных целях. Целью данной работы является выявление модели машинного обучения, которая будет лучше всего справляться с прогнозом формант. И вместе с тем будут рассмотрены вопросы о классификации формант с помощью машинного обучения.

В перспективе данная работа поможет продолжить более глубокие исследования в областях: идентификации человека по голосу, определение проблем, болезней, связанных с голосом (связки, гортань, лёгкие), синтез голоса.

1. A new database of healthy and pathological voices / U. Cesari, G. De Pietro, E. Marciano, C. Niri, G. Sannino, and L. Verde, Naples, Computers & Electrical Engineering, 2018. – 11 p
2. <https://www.kaggle.com/datasets/mozillaorg/common-voice/data>