

МИНОБРНАУКИ РОССИИ
Федеральное государственное бюджетное образовательное учреждение
высшего образования
**«САРАТОВСКИЙ НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ
ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ
ИМЕНИ Н. Г. ЧЕРНЫШЕВСКОГО»**

Кафедра дискретной математики и информационных технологий

КЛАССИФИКАЦИЯ АУДИОФАЙЛОВ
АВТОРЕФЕРАТ БАКАЛАВРСКОЙ РАБОТЫ

студента 4 курса 421 группы
направления 09.03.01 — Информатика и вычислительная техника
факультета КНиИТ
Кравченко Матвея Сергеевича

Научный руководитель
профессор, д. ф.-м. н.

В. А. Молчанов

Заведующий кафедрой
доцент, к. ф.-м. н.

Л. Б. Тяпаев

ВВЕДЕНИЕ

В настоящее время большинство музыкальных композиций распространяется в цифровом виде через сеть Интернет с помощью многочисленных сервисов, количество которых стремительно растёт. Анализ музыкальных файлов может стать полезным инструментом, с помощью которого данные сервисы будут способны рекомендовать своим пользователям новую музыку, на основе жанровых и прочих предпочтений клиента. Кроме того, классификация аудиофайлов может быть полезна обладателям больших музыкальных коллекций для простой и быстрой сортировки.

Целью выпускной квалификационной работы является классификация аудиофайлов с помощью методов машинного обучения. Задачами работы являются:

- 1) изучение понятия классификации аудиофайлов;
- 2) постановка проблемы классификации аудиофайлов;
- 3) ознакомление с алгоритмом логистической регрессии;
- 4) реализация алгоритма логистической регрессии;
- 5) тестирование алгоритма логистической регрессии.

КРАТКОЕ СОДЕРЖАНИЕ

Первый раздел дипломной работы содержит в себе постановку проблемы классификации аудиофайлов.

Второй раздел дипломной работы посвящен рассмотрению основных понятий, связанных с классификацией аудиофайлов. Данный раздел содержит три подраздела, в первом из которых рассматривается понятие классификации. Во втором подразделе рассматриваются признаки классификации и описываются их основные типы. В третьем подразделе содержится описание трех основных классов алгоритмов классификации: экспертных систем, алгоритмов обучения без учителя, алгоритмов обучения с учителем.

Третий раздел дипломной работы посвящен признакам классификации. Данный раздел состоит из двух подразделов, в первом из которых описывается библиотека Python `scikits.talkbox`, используемая при создании программы. Во втором подразделе содержится формальное описание алгоритма MFCC (Mel-Frequency Cepstral Coefficients).

Четвертый раздел работы посвящен логистической регрессии. Данный раздел состоит из четырех подразделов, в первом из которых приводится математическое обоснование алгоритма. Во втором подразделе содержится формальное описание алгоритма логистической регрессии. В третьем подразделе приводится формальное описание алгоритма многоклассовой логистической регрессии. В четвертом подразделе описывается библиотека Python `scikit-learn`, реализующая данный алгоритм, и используемая при создании программы.

Пятый раздел содержит в себе описание реализованной в ходе выполнения дипломной работы программы. Данный раздел состоит из двух подразделов, в первом из которых содержится общее описание программы. Во втором подразделе описывается процесс взаимодействия с программой и описание интерфейса.

Шестой раздел дипломной работы содержит в себе результаты тестирования программы-классификатора. Данный раздел состоит из трех подразделов, в первом из которых описывается набор данных GTZAN, используемый в качестве проверочных данных, и приводится точность классификации, полученная в других работах, где был использован этот набор. Во втором подразделе приводятся примеры работы программы для шести музыкаль-

ных жанров (классика, кантри, джаз, металл, поп и рок) при разном размере тестовой выборки. В третьем подразделе приводится анализ результатов работы программы.

ЗАКЛЮЧЕНИЕ

В процессе написания данной работы изучено понятие классификации аудиофайлов, рассмотрен алгоритм логистической регрессии. Разработана программа с графическим интерфейсом, реализующая алгоритм логистической регрессии на языке Python. Было проведено тестирование данной программы для шести музыкальных жанров с различным размером обучающей выборки. Данный классификатор может найти свое применение в сервисах, рекомендующих музыку своим пользователям, а также в качестве сортировщика аудиофайлов для обладателей больших музыкальных коллекций.

СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ

- 1 Классификация [Электронный ресурс]. URL: <http://www.machinelearning.ru/wiki/index.php?title=%D0%9A%D0%BB%D0%B0%D1%81%D1%81%D0%B8%D1%84%D0%B8%D0%BA%D0%B0%D1%86%D0%B8%D1%8F> (дата обращения: 17.05.2020). Загл. с экрана. Яз. рус.
- 2 Ромацкий, Д. Б. Автоматическая классификация по музыкальным жанрам : дис. соис. ст. маг. : 010300 / Даниил Борисович Ромацкий ; науч. рук. А. И. Попков ; Минобрнауки РФ, Томский гос. ун-т. Томск, 2014.
- 3 McKay C. Issues in automatic musical genre classification. / C. McKay – Presented at the McGill Graduate Students Society Symposium, 2004.
- 4 Tzanetakis G., Cook P. Musical Genre Classification of Audio Signals. / G. Tzanetakis, P. Cook – IEEE Transactions on Speech and Audio Processing, vol. 10, no. 5, July 2002.
- 5 Automatic genre classification of music content: a survey / N. Scaringella, G. Zoia, D. Mlynek – Signal Processing Magazine, IEEE (Volume 23, Issue 2), March 2006.
- 6 Improving Music Genre Classification by Short-Time Feature Integration / A. Meng, P. Ahrendt, J. Larsen – IEEE International Conference on Acoustics, Speech, and Signal Processing, 2005.
- 7 Aucouturier J.J., Pachet F. Representing musical genre: a state of the art / J.J. Aucouturier, F. Pachet – Journal of New Music Research, vol. 32, no. 1, pp. 83-93, 2003.
- 8 scikits.talkbox 0.2.5 project description [Электронный ресурс]. URL: <https://pypi.org/project/scikits.talkbox/> (дата обращения: 08.05.2020). Загл. с экрана. Яз. англ.
- 9 Лекция 6. Признаки. Кепстральные коэффициенты MFCC [Электронный ресурс]. URL: <https://logic.pdmi.ras.ru/~sergey/oldsite/teaching/asr/notes-06-features.pdf> (дата обращения: 10.05.2020). Загл. с экрана. Яз. рус.
- 10 Лекции по линейным алгоритмам классификации [Электронный ресурс]. URL: <http://www.machinelearning.ru/wiki/images/6/68/voron-ML-Lin.pdf> (дата обращения: 14.05.2020). Загл. с экрана. Яз. рус.

- 11 Лекции по алгоритмам классификации. Семинар 6 [Электронный ресурс]. URL: <http://www.machinelearning.ru/wiki/images/4/48/Ml.mmp.s6.pdf> (дата обращения: 17.05.2020). Загл. с экрана. Яз. рус.
- 12 Введение в Scikit-learn [Электронный ресурс]. URL: <https://neurohive.io/ru/osnovy-data-science/vvedenie-v-scikit-learn/> (дата обращения: 08.05.2019). Загл. с экрана. Яз. рус.
- 13 wxPython. The GUI Toolkit for Python [Электронный ресурс]. URL: <https://www.wxpython.org/> (дата обращения: 02.06.2020). Загл. с экрана. Яз. англ.
- 14 GTZAN Genre Collection [Электронный ресурс]. URL: <http://marsyas.info/downloads/datasets.html> (дата обращения: 15.05.2020). Загл. с экрана. Яз. англ.
- 15 Music Genre Classification: A Multilinear Approach / I. Panagakis, E. Venetos, C. Kotropoulos – ISMIR 2008: Proceedings of the 9th International Conference of Music Retrieval, pp. 583-588, 2008.
- 16 Мелконян, А. А. Автоматическая классификация по музыкальным жанрам // Десятая юбилейная годовичная научная конференция : сб. научн. ст. / Российско-армянский университет. Ереван, 2016. С. 64-68.
- 17 Коэльо, Л. Построение систем машинного обучения на языке Python / В. Ричарт ; пер. А. А. Слинкина. 2-е издание. М. : ДМК Пресс, 2016. 302 с.
- 18 Python 2.7.18 documentation [Электронный ресурс]. URL: <https://www.wxpython.org/> (дата обращения: 16.05.2020). Загл. с экрана. Яз. англ.
- 19 wxPython. GUI Toolkit. Tutorial [Электронный ресурс]. URL: https://www.tutorialspoint.com/wxpython/wxpython_tutorial.pdf (дата обращения: 02.06.2020). Загл. с экрана. Яз. англ.
- 20 wxFormBuilder [Электронный ресурс]. URL: <https://github.com/wxFormBuilder/wxFormBuilder> (дата обращения: 02.06.2020). Загл. с экрана. Яз. англ.