

- Удобный просмотр объектов торговли (все систематизировано и в полной степени описано).
 - Оформление заказа.
- Достоинства данного сервиса:
- Удобный, быстрый, вариативный способ заказа.
 - Безопасность данных пользователя.
 - История покупок пользователя.
 - Удобный и информативный интерфейс.
 - Максимально полное и подробное описание торговых объектов.
 - Наличие специального Андроид приложения.
 - Гибкая и изменяемая серверная часть.
 - Возможность легкой модернизации за счет новейших технологий разработки.
- На рисунке 1 представлен скриншот интерфейса Web-сервиса.

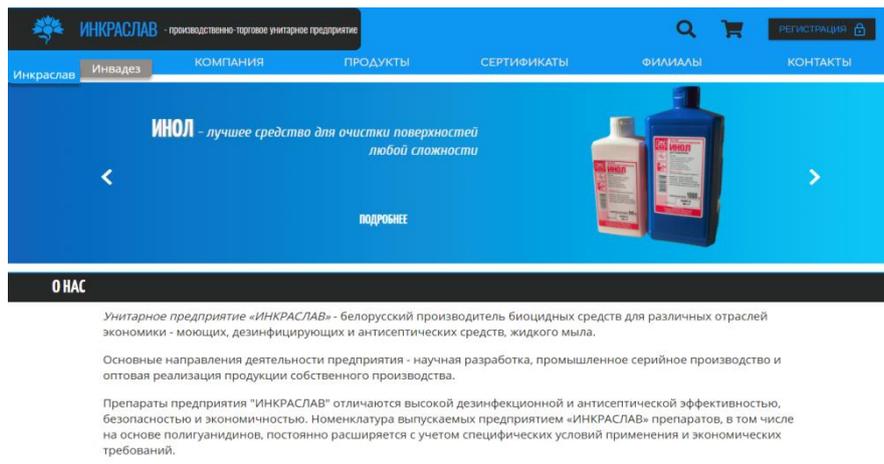


Рис. 1 – Интерфейс WEB-сервиса

Итогом работы является полностью рабочий Web-сервис и приложение, которые привлекут новых клиентов и обеспечат рост производства за счет качества товара и предоставления услуг на Web-сервисе.

Список использованных источников:

1. O'Reilly Media, Web Design in a Nutshell, 3rd Edition, A Desktop Quick Reference. O'Reilly Media. 2009, 832 P.
2. Thomas Connolly, Database Systems: A Practical Approach to Design, Implementation and Management with Learning SQL: A Step-By-Step Guide Using Oracle with Learning SQL. Addison Wesley. 2004.
3. O'Reilly Media. Learning React. Functional Web Development with React and Redux. O'ReillyMedia. 2017. 350 P.

РАСПОЗНАВАНИЕ ЖАНРА МУЗЫКАЛЬНЫХ КОМПОЗИЦИЙ С ПОМОЩЬЮ МАШИННОГО ОБУЧЕНИЯ

*Белорусский государственный университет информатики и радиоэлектроники
г. Минск, Республика Беларусь*

Дектярёв В.И.

Жвакина А. В. – к. т. наук, доцент

В настоящее время, в виду постоянно развивающихся веб-технологий и закономерно растущей численностью разнообразных сервисов, предоставляющих пользователям некоторую информацию, растёт конкуренция в различных сферах. В частности, появляется всё больше музыкальных сервисов. Для того, чтобы быть успешным на рынке, каждый из них должен выделиться некоторым функционалом, более особенным, чем возможность прослушивания композиций. Одним из вариантов такого функционала является анализ жанровых предпочтений пользователя по прослушанным композициям для последующего составления списка рекомендаций. Для анализа композиций на предмет различных характеристик зачастую используются нейронные сети. Они и были использованы в разработанном программном продукте для решения задачи.

Машинное обучение – класс методов искусственного интеллекта, характерной чертой которых является не прямое решение задачи, а обучение в процессе применения решений множества сходных задач. Для построения таких методов используются средства математической статистики, численных методов, методов оптимизации, теории вероятностей, теории графов, различные техники работы с данными в цифровой форме.

Искусственная нейронная сеть (ИНС) – математическая модель, а также её программное или

аппаратное воплощение, построенная по принципу организации и функционирования биологических нейронных сетей — сетей нервных клеток живого организма. Данное понятие сформировалось при попытке изучить и промоделировать процессы, являющиеся основой нашей мозговой активности. В контексте машинного обучения, нейронная сеть представляет собой частный случай методов распознавания образов. Данный механизм подходит для решения поставленной задачи: выделение характеристик в некотором наборе данных.

Рекуррентные нейронные сети – один из видов нейронных сетей, в котором связи между элементами образуют направленную последовательность. Благодаря такой архитектуре появляется возможность просчитать последовательные пространственные цепочки или серии событий во времени. В отличие от многослойных перцептронов, рекуррентные сети могут использовать свою внутреннюю память для обработки последовательностей произвольной длины. Долгая краткосрочная память (англ. *Longshort-term memory*; *LSTM*) – одна из разновидностей архитектуры рекуррентных нейронных сетей, предложенная в 1997 году Сеппом Хохрайтером и Юргеном Шмидхубером. Как и большинство рекуррентных нейронных сетей, LSTM-сеть является универсальной в том смысле, что при достаточном числе элементов сети она может выполнить любое вычисление, на которое способен стандартный компьютер. Для вычислений необходима соответствующая матрица весов: она может рассматриваться как программа. LSTM-сети выделяются среди прочих рекуррентных сетей благодаря приспособленности к обучению на задачах классификации, обработки и прогнозирования временных рядов в случаях, когда важные события разделены временными интервалами с неопределённой продолжительностью. В контексте поставленной задачи данное свойство особенно эффективно, так как жанр музыкальных композиций может быть опознан не только по всей композиции целиком, но также и по её ключевым фрагментам.

Прямой анализ звуковых сигналов во временной области может потенциально занять много времени в зависимости от длительности и качества записи, и он сам по себе не является наиболее эффективным методом, так как данный объём информации будет избыточным, если анализировать его на какие-либо общие характеристики. Составление спектрограмм и их дальнейший анализ будет быстрее, но всё же не достаточно производительным. На текущий момент одним из наиболее рациональных представлений записи для дальнейшего анализа является метод мел-частотных кепстральных коэффициентов, который широко применяется для составления характеристик речевых сигналов. Данные коэффициенты представляют лог мощности спектра в мел частотной области. Описывают мощность огибающей спектра, которая характеризует модель речевого тракта. Получаются путём преобразования Фурье исходного сигнала, отображения значений спектра на мел-шкалу и последующего дискретного косинусного преобразования значений на мел-шкале. Полученные значения амплитуд спектра и будут являться целевыми коэффициентами.

Для реализации системы распознавания жанра используется нейронная сеть, на вход которой подаётся преобразованный аудиофайл, а на выходе – один из 10 жанров музыки: блюз, классическая, кантри, диско, хип-хоп, джаз, металл, поп, регги, рок. То есть выходом нейронной сети является 10-мерный вектор со значениями вероятности принадлежности композиции к тому или иному жанру.

В качестве основной модели нейронной сети используется рекуррентная модель LSTM, так как она позволяет анализировать данные переменной длины и работает гораздо эффективнее в задачах распознавания длинных фрагментов, чем стандартный многослойный перцептрон: она способна запоминать контекстное состояние и состояние системы на предыдущем этапе, что позволит оценивать композицию в целом, а не только её части.

Перед передачей аудиозаписи в нейронную сеть, планируется её начальное преобразование. Механизм преобразования заключается в выделении мел-частотных кепстральных коэффициентов, применимых в распознавании речи: они способны обозначить ключевые особенности аудиозаписи.

Для обучения нейронной сети была найдена база из миллиона экземпляров метаданных аудиозаписей, находящихся в открытом доступе: <https://labrosa.ee.columbia.edu/millionsong/>, а также открытая база из 1000 реальных аудиозаписей GZTAN: <http://marsyas.info/downloads/datasets.html>. Сеть может обучаться на одной из данных баз в зависимости от формата целевых входных данных.

На данный момент различные крупные музыкальные сервисы уже имеют систему подбора композиций на основе прослушанных, однако их алгоритмы являются закрытыми и базируются на тегах аудиофайлов. Разработанный продукт базируется на анализе самой аудиозаписи и не требует наличия тегов у записи.

Список использованных источников:

1. И. Заенцев. Нейронные сети: основные модели – Воронеж, 1999. – 74 с.
2. Nielsen M. Neural Networks and deep learning. [Электронный ресурс]. – Режим доступа: <http://neuralnetworksanddeeplearning.com>. – Дата доступа: 15.02.2018.
3. das, Asutosh, Manas Ranjan Jena, Kalyan Kumar Barik. Mel-Frequency Cepstral Coefficient (MFCC) - a Novel Method for Speaker Recognition. – Digital Technologies, 2014 – 195 с.
4. Долгая краткосрочная память [Электронный ресурс]. – Режим доступа: https://ru.wikipedia.org/wiki/Долгая_краткосрочная_память – Дата доступа: 25.02.2018