

УДК 004.853

ОПРЕДЕЛЕНИЕ ЖАНРА МУЗЫКАЛЬНОГО ПРОИЗВЕДЕНИЯ С ПРИМЕНЕНИЕМ МНОГОСЛОЙНОГО ПЕРСЕПТРОНА

© Денисов М.Н., Лёзина И.В.

Самарский национальный исследовательский университет
имени академика С.П. Королева, г. Самара, Российская Федерация

e-mail: mdenisov2@mail.ru

В данной работе исследуется проблема определения жанра музыкального произведения. Жанр музыки является важным критерием для музыкальной классификации и может быть использован во многих задачах, таких как автоматическая рекомендация музыки, поиск похожих треков и т. д. Одним из наиболее распространенных типов нейронных сетей является многослойный перцептрон (Multilayer Perceptron, MLP), в данной работе будет использован именно он и обучается выбранная нейронная сеть методом обратного распространения ошибки [1]. На рисунке представлена модель многослойного перцептрона.

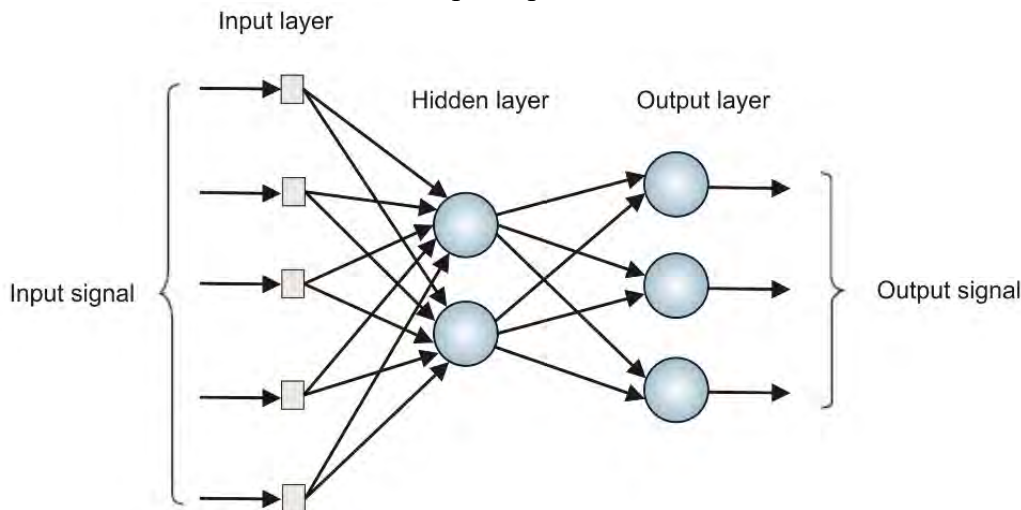


Рисунок – Модель многослойного перцептрона

Главной задачей является возможность классификации музыкальной композиции по десяти предоставленным жанрам: блюз, классика, кантри, диско, хип-хоп, джаз, метал, поп, рок, регги. Обучение происходит на наборе данных из открытого доступа.

Метод обратного распространения ошибки позволяет обучать многослойный перцептрон путем корректировки весовых коэффициентов в соответствии с минимизацией ошибки. Этот метод является одним из самых эффективных для обучения нейронных сетей и широко применяется в задачах классификации и анализа данных [2].

Для определения жанра музыкального произведения с помощью многослойного перцептрона в качестве входных параметров используются характеристики звукового сигнала. Среди таких характеристик можно выделить [3]:

- Темп (beats per minute, BPM) – количество ударов в минуту, используется для оценки скорости композиции;

- Среднеквадратичное значение (Root Mean Square, RMS) – мера амплитуды звука, позволяет оценить громкость композиции;
- Число переходов через ноль (Zero Crossing Rate, ZCR) – количество пересечений звукового сигнала через ноль за единицу времени, позволяет определить наличие сильных высокочастотных компонентов в сигнале;
- Коэффициенты мел-частотных кепстральных коэффициентов (Mel-frequency cepstral coefficients, MFCC) – мера спектральных характеристик звука, используется для анализа тембра и тональности музыки;
- Спектральный центроид (Spectral centroid) – средневзвешенное значение частот спектра, позволяет определить тональность музыки;
- Спектральная плотность (Spectral bandwidth) – функция, которая показывает распределение мощности сигнала по частотам;
- Спектральный контраст (Spectral Contrast) – мера различия между спектральными компонентами звука. Она используется для оценки спектральной изменчивости звукового сигнала и может помочь в определении стиля музыки;
- Спектральный откат (Spectral Rolloff) – частота, ниже которой расположена определенная доля спектральной энергии. Он используется для определения высоты тона и глубины звучания музыки;
- Хроматическая спектрограмма (Chroma STFT) – представление звука, основанное на частотах нот в музыке. Она используется для анализа тональности и гармонии музыки;
- Разложение на гармонические и перкуссионные компоненты (Harmonic-percussive separation, HPSS) – метод разделения звука на гармонические и перкуссионные составляющие. Он может использоваться для анализа гармонической и ритмической структуры музыки.

Использование этих параметров в качестве входных данных для многослойного перцептрона позволяет добиться высокой точности определения жанра музыкального произведения. Важно отметить, что для обучения нейронной сети необходимо большое количество данных, представляющих различные жанры музыки. Для выявления лучших показателей были проведены экспериментальные исследования эффективности разработанной системы и точности распознавания жанров при разном количестве нейронов в скрытом слое. Наилучшие результаты нейронная сеть показала при 75 нейронах в скрытом слое и коэффициенте скорости обучения, равном 0,15.

Таким образом, результаты исследования показывают, что использование многослойного перцептрона в определении жанра музыкальных произведений является эффективным методом. Однако дальнейшие исследования могут быть проведены для улучшения точности и эффективности алгоритма, включая использование более сложных нейронных сетей, улучшение качества и количества входных параметров, а также использование других методов машинного обучения.

Библиографический список

1. Осовский С., Нейронные сети для обработки информации / пер. с польского И.Д. Рудинского. М.: Финансы и статистика, 2002. 344с.
2. Academy yandex. Метод обратного распространения ошибки. URL: <https://academy.yandex.ru/handbook/ml/article/metod-obratnogo-rasprostraneniya-oshibki>.
3. Medium. Анализ аудиоданных с помощью глубокого обучения и Python. URL: <https://medium.com/nuances-of-programming/анализ-аудиоданных-с-помощью-глубокого-обучения-и-python-часть-1-2056fef8525e>.