Санкт-Петербургский государственный университет

***УСТИНОВА Василиса Алексеевна***

**Выпускная квалификационная работа**

***Разработка приложения для автоматического распознавания аккордов по музыкальным аудиосигналам***

Уровень образования: *бакалавриат*

Направление *50.03.01 «Искусства и гуманитарные науки»*

Основная образовательная программа *СВ.5045.2018 «Свободные искусства и науки»*

Научный руководитель:

Доцент кафедры проблем конвергенции естественных и гуманитарных наук, кандидат физико-математических наук

Черных Герман Анатольевич

Рецензент:

Затворницкий Александр Петрович

Санкт-Петербург

2022

СОДЕРЖАНИЕ

Введение………………………………..…………………………………........3

1. Проблематика работы………………………………………….…………..6

1. Разработанность темы………………………………….…………7

1. Цель и задачи исследования……………………………………………….7
2. Гармонизация мелодий……………………………..………...…………….8

3.1 Модель на основе сопоставления шаблонов…………..……….9

3.2 Модель на основе НММ………………………………………..10

3.3 Модель на основе генетического алгоритма……….……..…..12

* 1. Глубокая модель на основе BILSTM………………..………..14
  2. Глубокая многозадачная модель MTharmonizer…...…………15

1. Сравнительное исследование.………….…………………….......……....19
   1. Субъективная оценка ..…………………………….…............21
   2. Объективная оценка ..…...…………………………….….........24
   3. Оценка наилучшей модели...………………………………….26
2. Разработка программы для гармонизации мелодий………………….…27
   1. Функционал программы……………………………………..…27
3. Алгоритм работы программы…………………………………………….28
   1. Работа с wav файлами……………………………….…………28
   2. Преобразование Фурье…………………………………….…..29
   3. Pitch Class Profile (PCP) ……………………………………..…30
   4. Нормировка PCP……………………………………………….30
   5. Полносвязная нейронная сеть…………………………………30
      1. Архитектура полносвязной нейросети……………………..31
      2. Функция активации (Rectified Linear Unit) ………………..32
      3. Функция активации (сигмоидная функция)………………..33

6.6 Обучение нейронной сети…………………………………………34

6.6.1 Обучение с учителем…………………………………………34

6.6.2 Параметры нейронной сети…………………………………..36

6.6.3 Метрики качества…………………………………………….37

6.Предсказания модели…………………………………….………….38

7. Итоги……………………………………………………………….…….39

8. Будущая исследовательская деятельность……………………………..41

ВВЕДЕНИЕ

Термин гармония или гаромонизация используется для обозначения аккордового сопровождения. Задача автоматической гармонизации направлена на построение модели, которая генерирует последовательность аккордов исходя из заданной мелодии. То есть, под термином гармонизация мы подразумеваем разложение звука на аккорды. Все программы по гармонизации звуковых сигналов работают по принципу: программа слышит мелодию, определяет аккорды и выдает файл с последовательностью этих аккордов. Гармонизация музыки – сложная задача, существует много способов гармонизации одной и той же мелодии, а то, что делает ту или иную гармонизацию приятной часто зависит от музыкального жанра и других факторов.

Тональная музыка, которая охватывает большую часть западной музыки, определяет специфические мотивы отношений между аккордами на основе иерархической централизации - функции тональности связаны с «тоникой». То есть, тональная музыка подразумевает под собой зависимость от «тоники» - ноты, с которой произведение начинается и которым оно, как правило, заканчивается: все сводится к ней – так тоническая музыка и приобретает свой привычный и приятный человеческому уху гармоничный строй и приятное звучание (в отличие от атонической, где ноты не зависят друг от друга, не подчиняются законам тоники).

Создание приятной музыки часто зависит не только от технической стороны музыкальной грамоты, но и от тонкостей, долгосрочных зависимостей и культурных контекстов, которые могут быть легко доступны для композитора-человека, но будут очень трудно воспринимаемы и определяемы машиной.

В данной работе мы представляем исследование, оценивающее эффективность канонических подходов к решению поставленной задачи, а затем создаем программу по определению аккордов в музыкальной прогрессии. В сравнительном исследовании будут оцениваться модели, гармонизирующие мелодии с участием нескольких музыкальных инструментов, звучащих одновременно. Мы же на данном этапе своей научной деятельности создадим модель, которая будет способна распознавать гитарную музыку, виолончель и пианино. Гармонизация сложных оркестровых композиций будет рассмотрена нами, как поле для дальнейшего исследования. Сравнительное исследование включает в себя модель, основанную на согласовании шаблонов, скрытую модель Маркова, генетический алгоритм и два метода глубокого обучения. Сами мы создадим модель глубокого обучения с полносвязной архитектурой.

Оценка первых пяти моделей проводится на наборе данных Hookthe-ory Pianoroll triad Dataset (HTPD3 состоит всего из 48 трезвучий). В итоге мы получаем результаты объективной оценки с использованием шести различных метрик. К сожалению, данный датасет оказался закрытым, и мы не смогли провести на нем обучение нашей модели.

Мы же обучаем нашу модель на наборе данных, состоящем из 10 основных аккордов ("C", "D", "Dm", "E", "Em", "F", "G", "A", "Am", "Bm"). Мы берем такой набор аккордов, так как они являются основными и с помощью них можно будет добиться хороших результатов гармонизации мелодий с одним инструментом. Оценка качества получившейся проводится с помощью двух метрик: Categorical accuracy и Top-k Accuracy

1. ПРОБЛЕМАТИКА РАБОТЫ

В современном мире большую роль играет оперативность и мобильность: реакция на событие и ее результат сейчас должны быть мгновенны, а информация должна извлекаться без особого труда.

Многие музыканты часто сталкиваются с такой проблемой: на вечерах бардовских песен или на улицах им попадается замечательная мелодия, попадается и…пропадает бесследно - авторство неизвестно, слова не найдены. К сожалению, огромное количество песен и музыкальных произведений невозможно найти на просторах Интернета. Кропотливо подбирать нужную тональность – задача не из легких. Хотелось бы создать «Shazam», но только для музыкантов: программа слышит музыкальное произведение и выдает тебе аккорды.

Такие программы сейчас есть, однако, чтобы ей воспользоваться, нужно быть довольно опытным программистом: знать, как поместить файл с мелодией в программу, как запустить код, знать, по крайней мере, что такое «GitHub» - чего ты не сможешь сделать, имея только музыкальное образование. Данной работой нам хотелось открыть музыкантам возможность воспроизводить музыкальные произведения, аккордов которых нет в Интернете, самосовершенствоваться, уметь работать с кодом.

Наша работа делает доступным воспроизведение музыкантами произведения, автора и названия которого они не знают. Достаточно записать произведение на диктофон и подать этот файл в нашу программу. Результатом будет запись аккордов, по которым можно будет без труда сыграть интересующее произведение.

1. **Разработанность темы**

В прошлом было предпринято несколько попыток решения задачи гармонизации мелодий. До появления глубокого обучения наиболее активно использовался подход, основанный на скрытых марковских моделях. Например, Paiment (2006) и др. предложили древовидную HMM которая позволяет изучать нелокальные зависимости аккордов и кодировать вероятности субституции аккордов, взятые из психоакустики. [1] Они также представили новое представление для аккордов, которое кодирует относительные степени шкалы, а не абсолютные значения нот, и включили в свою модель подграф специально для его обработки

Temperly в 2009 году представил статистическую модель, которая генерирует и анализирует музыку по трём подструктурам: метрическая музыка, гармоническая музыка и структура потока. В генеративной части этой модели сначала генерируется метрическая структура, определяющая акценты ударов и биений, а затем, на основе этой метрический структуры генерируются гармоническая структура и прогрессия. [1]

Цусима (2018) и др. использовали обучение без надзора для обучения генеративных моделей НММ и PCFG для гармонизации, показав, что паттерны, изученные этими моделями, соответствуют категоризации, представленной функциональной гармонией. [1]

1. ЦЕЛЬ И ЗАДАЧИ ИССЛЕДОВАНИЯ

Мы разобрались в существующих подходах гармонизации мелодий (согласование шаблонов, скрытая Марковская модель (HMM), генетический алгоритм, и два варианта реализации модели на глубоких рекуррентных нейронных сетях (RMM).

Провели исследование производительности моделей, выявили из них наиболее точную модель для гармонизации мелодий, оценили ее сильные стороны и разобрали слабые.

Затем, исходя из всего вышеперечисленного, подобрали наиболее эффективный способ гармонизации и написали свою программу для распознавания мелодий.

Немаловажным шагом стало помещение получившейся программы в общий доступ и ее подробное комментирование и пояснение, для того чтобы музыканты могли ей пользоваться в личных целях: увеличивать свою производительность, пользоваться новыми инструментами написания музыкальных композиций, совершенствовать навыки работы с произведениями с помощью современных технологий.

1. ГАРМОНИЗАЦИЯ МЕЛОДИЙ

Для начала, разберемся с алгоритмом гармонизации мелодий. Модель гармонизации мелодии принимает на вход последовательность мелодии из тактов и генерирует соответствующую последовательность аккордов на выходе. [1] Последовательность аккордов определяется здесь как ряд аккордовых меток , где – длина последовательности. В данной работе каждая модель предсказывают метку аккорда для каждого полу бара, то есть . Каждая метка у выбирается из конечного словаря аккордов . Рассматриваются только аккорды триады для уменьшения сложности задачи - аккорды триады – это аккорды, состоящие из трех нот. Также мы рассматриваем состояние покоя, когда никаких аккордов не звучит, поэтому размер словаря аккордов составляет . [1]

Последовательность мелодий – это серия музыкальных нот. Для представления мелодий и использования их в качестве входных данных для моделей мы записываем признаки следующим образом . Модель гармонизации мелодии   
*f* (⋅) мoжет быть обучена путем минимизации потерь, вычисленных между истинным и выходом нашей модели , где  *-* входная мелодия.[1]

* 1. **Модель на основе сопоставления шаблонов**

Модель разделяет обучающие мелодии на полубары и строит профиль высоты тона для каждого сегмента, [15] чей профиль высоты тона наиболее точно совпадает. Если существует более одного возможного шаблона аккорда, который имеет наивысший балл соответствия, мы выбираем аккорд по принципу равномерного распределения среди возможных вариантов. [2] Такая модель называется моделью на основе сопоставления шаблонов поскольку лежащий в её основе метод сравнивает профиль данного сегмента мелодии с профилем аккордов шаблона. [2] В работе использовался профиль класса высоты тона Фуджишимы (PCP)[1]. PCP – это 12 мерный вектор признаков , где каждый элемент соответствует активности класса высоты тона. [1] PCP для каждой из аккордовых меток |С| строится путём установки элементов, соответствующих классам лада, которые являются частью аккорда, в единицу, а все остальные - в ноль.

Результаты этой модели более консервативный по дизайну и характеризуются интенсивным использованием аккордовых тонов. [1]Кроме того, эта модель устанавливает метку аккорда независимо для каждого полубара, не учитывая метки соседних аккордов или прогрессию аккордов во времени. [1]

**3.2 Модель на основе HMM**

НММ – это вероятностная схема для моделирования последовательностей с латентными или скрытыми переменными. [16] Представляет собой вероятностную модель временных рядов. Марковская модель описывает процесс, в котором скрытая марковская цепь генерирует случайную последовательность ненаблюдаемых состояний, а затем генерирует наблюдаемую случайную последовательность наблюдений для каждого состояния. [16] Каждое состояние генерирует наблюдение, которое приводит к последовательности наблюдения. [2]

Модели гармонизации мелодий на основе НММ рассматривают метки аккордов, как скрытые переменные и оценивают наиболее вероятную последовательность аккордов для данного набора нот мелодии. [1] В отличие от модели, основанной на сопоставлении шаблонов, эта модель учитывает взаимосвязь между соседними метками аккордов. [16] Такие модели (НММ) широко использовались в исследованиях по генерации аккордов и гармонизации мелодий до наступления эры глубокого обучения. [2] В этой модели сделаны следующие предположения:

1. Наблюдаемая последовательность мелодий Х = х1, х2, … хМ статистически смещена из-за скрытой последовательности аккордов Y = y1, Y2, … yM, которая должна быть оценена. [1]
2. х m зависит только от y m, ∀m [1, M]. [1]
3. y m зависит только от y m-1, ∀m [2, M]. [1]

Задача состоит в том, чтобы оценить наиболее вероятную скрытую последовательность =, …, . [1] Это равносильно максимизации апостериорной вероятности:

а членвероятностью перехода. Эта задача оптимизации может быть решена алгоритмом Витерби. [1]

Отходя от НММ, данная реализация использует РСР для подстановки нот мелодии, то есть для вычисления . Соответственно, здесь используются многомерные гауссовские распределения для моделирования вероятностей выбросов. Для каждой метки аккорда задается ковариационная матрица соответствующего гауссовского распределения в виде диагональной матрицы, затем вычисляется среднее значение и дисперсия для каждого измерения на основе РСР-изображений сегментов мелодии, связанных с данной меткой аккорда в обучающей выборке. [1]

Для расчёта вероятности перехода подсчитывается количество переходов между последними метками аккордов [16] (то есть биграммами), потом эти подсчеты нормализуются так, чтобы их сумма была равна единице для каждой предшествующей метки аккорда. [1] Если биграмма для предшествующей метки аккорда отсутствует - используется равномерное распределение. Чтобы избежать нулевой вероятности, мы сглаживаем распределение, интерполируя с предварительной вероятностью следующим образом:

что дает уточненную вероятность перехода P. Гиперпараметр эмпирически установлен на 0.08 путем проведения экспериментов на случайном 10% подмножестве обучающего множества. [1]

* 1. **Модель на основе генетического алгоритма**

Генетический алгоритм – это эвристический (практический) метод, не точный, но достаточный для решения задачи поиска, используемый для решения задач оптимизации и моделирования путём случайного подбора, комбинирования и вариации искомых параметров. [2] ГА примечателен использованием механизмов, аналогичных естественному отбору в природе. [13] Отличительной особенностью генетического алгоритма является акцент на использование оператора «скрещивания», который производит операцию рекомбинации решений-кандидатов, роль которой аналогична роли скрещивания в живой природе. [2]

Генетический алгоритм можно использовать, как в подходе, основанном на правилах, так и вероятностном подходе. [13] В первом случае необходимо разработать набор правил о том, какие условия должны быть соблюдены для музыкально приемлемых мелодий – фитнес функция формируется на основе этого набора правил. [2] Во втором случае фитнес функция формируется на основе статистического набора данных. Функция приспособленности (фитнес функция) - это вещественная или целочисленная функция одной или нескольких переменных, подлежащая оптимизации в результате работы генетического алгоритма, направляет эволюцию в сторону оптимального решения. [3]

В отличие от других реализованных моделей, модель на основе ГА принимает на вход вычислительный вектор признаков для каждой 16-й ноты (т.е. 1/4 такта). [1] Таким образом, представление мелодии имеет временное разрешение в 8 раз больше, чем у аккордовой прогрессии. Это значит, что и указывают на одну и ту же временную позицию. Исследуемая модель использует вероятностный подход, определяя фитнес-функцию на основе следующих элементов. Во-первых, условная вероятность каждого аккордовой прогрессии определяется как:

Где [] – функция потолка. Вероятность перехода хорды вычисляется как:

Условная вероятность каждого аккорда с учетом его позиции определяется как:

где, – временная позиция аккорда y m. Для простоты: , где – это функция модуляции. С помощью этого члена модель может узнать, что тонический аккорд обычно появляется в первой половине первого такта, в то время как доминантный аккорд появляется, как правило, во второй половине второго бара. [1] Наконец, в данной модели используется энтропия для оценки сложности последовательности аккордов, которая не должна быть слишком низкой, чтобы избежать монотонных аккордовых повторений. [1] Энтропия определяется как:

.

В фитнесс-функции оценивается, насколько вероятна энтропия *E*(*Y*) в данных. [1]

,

где *Е* – случайная величина энтропии хорды прогрессии и декретизируется на 0,25 (вероятность распределения получена из обучающих данных). [1]

Фитнес-функция *F*(*Y*)высчитывается следующим образом:

.

Здесь устанавливаются веса: w1, w2, w3, w4 = 1.0. [1]

**3.4 Глубокая модель на основе BILSTM**

Модель извлекает контекстную информацию из положительного и отрицательного временного направления. [4] Гармонический ритм в такой модели составляет пол бара, а словарь аккордов включает не только мажорные и минорные трезвучия, но, так же заниженные и дополнительные аккорды, а еще символ N.C. (т.е. |C) = 49). [1] Как показано на рисунке 1, эта модель имеет два скрытых BILSTM слоя, за которыми следует слой с полной связью. Показатель отсева и количество скрытых слоев и блокой выбираются эмпирически путем максимизации точности предсказания аккордов на случайном подмножестве обучающегно множества. [1] Обучение модели производилось с использованием минибатча градиентного спуска с катигориальной кросс энтропией в качестве функции стоимости. [1]

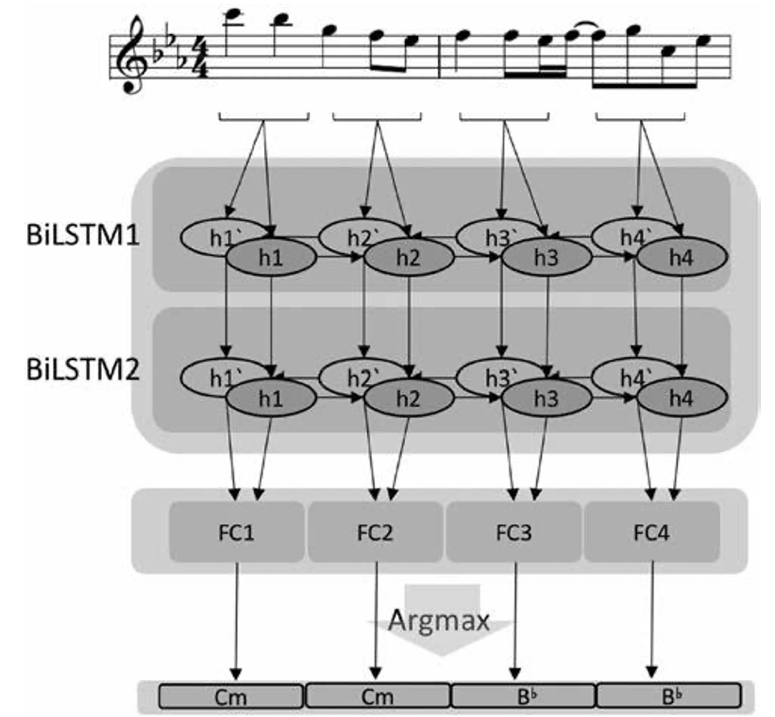


Рис. 1 архитектура нейронной сети на основе BILSTM

Из эмпирических наблюдений исследования выяснилось, что такая модель обучения имеет два недостатка для длинных фраз:

1. Чрезмерное использование общих аккордов таких как С, F и G - они повторяются слишком часто, что делает аккордовую прогрессию монотонной. [1]
2. Некорректная фразировка между аккордами и мелодией, в результате чего появляются частые повторения последовательности типа F C или G C. [1] Такая неточность может привести к каденции в начале или середине фразы, что потенциально вызывает ненужное ощущение окончания музыкальной фразы в середине аккорда.

**3.5 Глубокая многозадачная модель MTharmonizer**

Эти два недостатка обычного BILSTM алгоритма устраняются в расширенной модели BILSTM, которая называется MTharmonizer. Основная идея заключается в том, чтобы обучить модель предсказывать не только метки аккордов, но и функцию хорды. [1] Мульти задачной данная модель названа потому, что имеет дело сразу с двумя задачами одновременно: в данном методе функции аккордов и метки аккордов взаимосвязаны — это облегчает обучение модели аккордовым функциям, и информирует модель о том, какие метки аккордов имеют одну и ту же функцию, и, поэтому, могут быть взаимозаменяемыми. Такое многозадачное обучение помогает модели научиться правильной функциональной прогрессии, что, в свою очередь, приводит к улучшению гармонический фразировки относительно мелодии. Гармоническая фразировка – умение придавать мелодии необходимую долю выразительности, наделять ее должным характером и ритмикой.

[1]

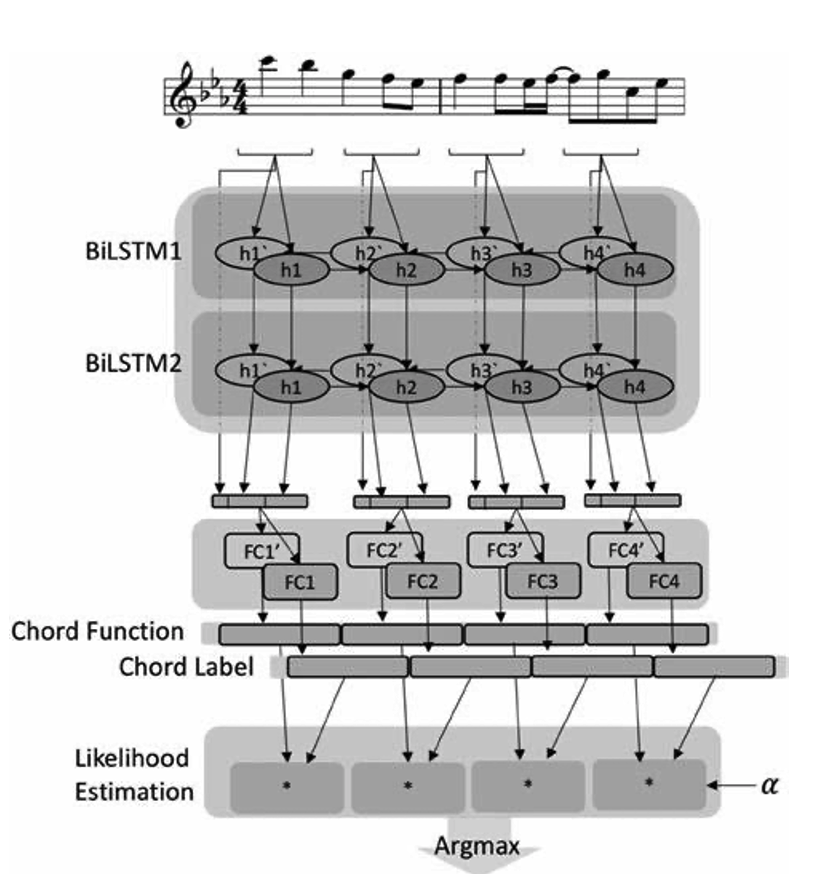
****

Рис. 2 архитектура нейронной сети на основе MTharmonizer

Функциональная гармония развивает отношения между аккордами и гаммами, а также, описывает как гармоническое движение направляет музыкальное восприятие и эмоции. В то время, как аккордовая прогрессия, состоящая из произвольно выбранных аккордов, обычно кажется бесцельной, аккорды прогрессии, которые следуют правилам функциональной характеристики, создают устойчивую тональность.

Музыкальные теоретики аннотируют каждую степень шкалы в функции, такие как тоника и доминанта, основанные на ассоциации между аккордом и степенью определенной шкалы. [1] Эти функции объясняют, какую роль играет та или иная степень шкалы и связанный с ней, аккорд, который играет в музыкальной фразировке и композиции. [1] В частности, в при запуске MTharmonizer рассматриваются следующие функции:

* Тонизирующая функция служит для стабилизации и укрепления тонального центра
* Доминирующая функция обеспечивает сильное чувство движения назад к тональному центру. Например, прогрессия, которая движется от доминантной функциональной шкалы аккорда к тональной шкале аккорда сначала создает напряжение, а затем, разрешает его.
* Другие, которые охватывают все остальные аккорды, которые не являются ни тоническими, ни доминантными, а например, субдоминантными.

Было выявлено два потенциальных преимущества добавления функции аккордов к целевому выходу. Во-первых, в отличие от распределения аккордовых меток, распределение аккордовых функций относительно сбалансировано, что облегчает для модели обучение аккордовым функциям. [1] Во-вторых, поскольку метки аккордов и метки аккордовых функций взаимозависимы, добавление функций аккордов в качестве цели информирует модель о том, какие метки аккордов имеют одну и ту же функцию и поэтому могут быть взаимозаменяемыми. [1] Предполагается, что такое многозадачное обучение поможет модели научиться правильной функциональной прогрессии, что, в свою очередь, приведет к улучшению гармонической фразировки относительно мелодии. Функция потерь определяется как:

Где H – обозначает категориальную функцию перекрестной энтропии. [1]

f – ветвь предсказания метки хорды. [1]

g – ветвь предсказания функции хорды. [1]

= 0, модель сводится к однозадачной модели, можно просто записать как *Y.* В модели установлен равным 1,5, чтобы убедиться, что значение потерь из функций и одинаково масштабируются. [1] Две ветви f и g разделяют два BILSTM слоя, но не полностью связанный слой. [1] Эмпирически было обнаружено, что, если значение слишком мало, модель будет стремиться к гармонизации мелодии с аккордами с тоническими и доминантными функциями: итоговые последовательности аккордов, таким образом, будут лишены разнообразия. [1]

Выходами f и g являются значения вероятности для каждой метка аккорда и аккордовой функции, учитывая входную мелодию. Как показано на рисунке 2, при прогнозировании конечной последовательности аккордов следует полагаться на взвешенную комбинацию входов f и выходов g следующим образом:

\*

где h – это таблица поиска, которая отображает три функции аккордов для аккордовых меток |С|, а – это заранее определенный гиперпараметр, который позволяет исследователям увеличить важность правильного прогнозирования функции хорды по сравнению с вероятностью правильного предсказания метки для каждого аккорда. [1] В данной реализации установлено для тонических и доминантовых аккордов и для остальных аккордов, чтобы побудить модель выбирать аккордовые метки с меньшей вероятностью, то есть, для разнообразия прогрессии использовать «другие» аккорды. [1] Это, скорее всего, повлияет на среднюю часть последовательности аккордов, поскольку именно здесь вероятность наблюдения аккорда из трех функций, скорее всего, будет одинаковой. Именно поэтому, применение различных гиперпараметров имеет значение. Поскольку в исследовании, в основном, добавляется разнообразие в среднюю часть аккордовой последовательности, общая аккордовая прогрессия и фразировка не нарушаются. [1]

1. СРАВНИТЕЛЬНОЕ ИССЛЕДОВАНИЕ

Обучение всех моделей проводилось сначала на обучающем срезе HTPD3, а затем данные применялись к тестовому срезу HTPD3. В итоге, получается предсказание последовательностей аккордов.

1) Пример гармонизации (в мажорным ключе) из песни The Beatles: Hey Jude. Mы видим что в то время как модели без глубокого обучения меняют гармонизацию в разных фразах MTharmonizer генерирует прогрессию, которая красиво закрывает фразу. [1]



Рис. 3 пример гармонизации песни The Beatles: Hey Jude

2) Пример гармонизации (в минорном ключе) из песни ABBA Gimme Gimme Gimme Gimme. Как и в примере выше результат работы MTharmonizer кажется более разнообразным и функционально правильным. [1] Мы также видим, что результат работы ГА довольно интересен, например, с недиатоническим аккордом ре-бемоль мажор и закрытием музыкальной фразы пиккардийской терцией (то есть мажорным аккордом тоники в конце аккордовой последовательности, находящейся в минорном ключе). [1] Мы также видим, что методы без глубокого обучения похоже слабее справляются с тональность музыки. [1]



Рис. 4 Пример гармонизации песни ABBA Gimme, Gimme, Gimme, Gimme

**4.1 Субъективная оценка**

Рассмотрим результаты сравнительного исследования 2021 года, проведенного Yin-Cheng Yeh, Wen-Yi Hsiao и Benjamin Genchel. Исследователи из университета Джорджии (Атланта) и университета Нихон сравнили результаты производительности модели, основанной на согласовании шаблонов, скрытой модели Маркова, генетического алгоритма и двух методов глубокого обучения. [1]

На выборке из 200 людей был проведён опрос, в котором испытуемым предложили прослушать и оценить результаты гармонизации различных моделей. Сравнивались результаты, автоматически сгенерированной, программы с последовательностью, составленной человеком. Однако учитывая временные и когнитивные затраты было невозможно попросить каждого испытуемого оценить результаты каждой модели для каждого музыкального произведения в тестовом наборе (всего их 5538). [1] Испытуемые оценивали гармонизацию по следующим критериям:

* Гармоничность: степень, в которой аккордовая прогрессия успешно или приятно гармонизируют данную мелодию. Это соответствует тому, что измеряют метрики гармоничности мелодии/аккорда. [1]
* Интересность: степень, в которой аккордовое развитие звучит захватывающе, неожиданно, и вызывает положительную стимуляцию. Обращаем внимание на то, что здесь используется менее техничный термин «интересность», так как планируется получить отзывы и от людей, не имеющих музыкального образования. [1]
* Согласованность: согласованность между мелодией и аккордовой прогрессией с точки зрения гармоничности и фразировки. [1]
* Аккордовая прогрессия: насколько последовательна, приятна, или разумна аккордовая прогрессия сама по себе независимо от мелодии. [1]
* Общее качество данной гармонизации. [1]

Среди двух методов глубокого обучения, MTharmonizer превосходит BILSTM по всем критериям ранжирования, особенно по критерию интересности. Мы наблюдаем, что MTharmonizer действительно генерирует более разнообразные аккордовые прогрессе по сравнению с BILSTM, возможно, благодаря учёту функций. [1]

Среди четырёх метрик, оригинальные прогрессии, составленные человеком, имеют более высокие оценки по показателям «согласованность» и «общее», и самые низкие по показателю «интересность». Это говорит о том, что способы упрощения данных, например, использование только аккордов с корневым расположением трезвучий мог ограничить перцептивные качество музыки, в частности её разнообразие. [1]

В целом, результаты по аккордовой прогрессии (то есть связность аккордовой прогрессии самой по себе) лучше коррелирует с результатами по связности (то есть по связности между мелодией и аккордовыми последовательностями) чем интересность аккордовой прогрессии. [1] Это говорит о том, что аккордовая прогрессия, оцененная как интересная, может не звучать связано.

Хотя Генетический Алгоритм работает хуже, чем MTharmonizer по всем четырём метрикам, он довольно хорошо справляется с задачей интересности. Парный t - тест не выявил существенной разницы в производительности между генерируемыми генетическим алгоритмом прогрессиями и оригинальными прогрессиями, составленными человеком, по показателю интересности. [1]

В итоге стало понятно, что методы глубокого обучения имеют более высокие шансы на победу над методами без глубокого обучения по показателям гармоничности и общего качества.

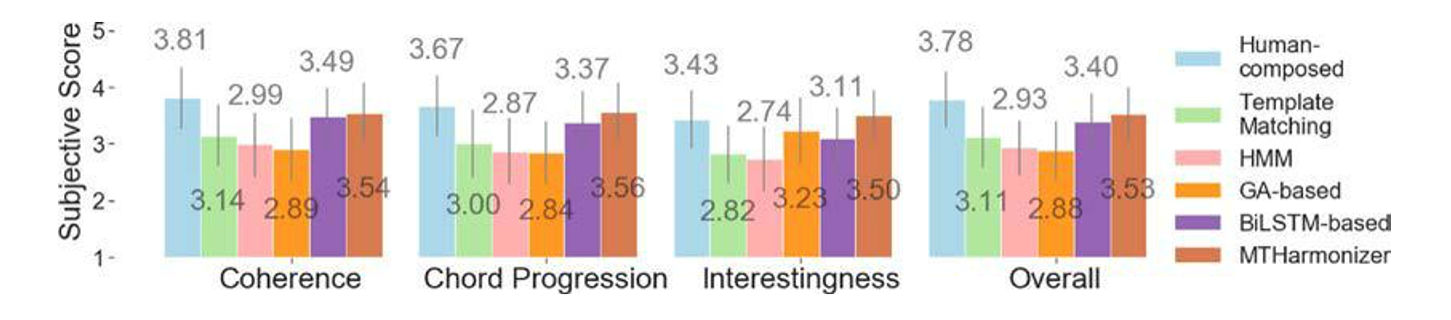
****

Рис. 5 график результатов субъективной оценки гармонизации мелодий

**4.2 Объективная оценка**



Рис. 6 результаты объективной оценки гармонизации мелодий

В данной таблице представлены баллы объективной оценки для различных методов гармонизации мелодий. Общая оценка складывается из «метрики гармоничности мелодии/аккорда» и «метрики аккордовых прогрессий», сопоставление которых дает результат гармонизации. Чем ближе к значениям образцов, составленных человеком, тем лучше модель моделирует обучающие данные. Жирным шифтом выделены значения, близкие к значениям образцов, составленных человеком по каждой метрике. Более высокие значения в CTnCTR и PCS и более низкие значения в MCTD могут говорить о высокой гармоничности мелодии. Более высокие значения в CHE и CC и более низкие значения CTD могут говорить о высоком разнообразии (которое может быть связано с интересностью) аккордовой прогрессии.

Методы на основе сопоставления шаблонов и HMM имеют высокие баллы в PCS и низкие в MCTD, что указывает на то, что гармонизация, создаваемые этими двумя методами, может быть слишком консервативной. В противоположность этому, GA имеет низкие баллы в PCS и высокие в MCTD, что указывает на слишком низкую гармоничность. [1] Эти результаты согласуются с субъективной оценкой, предполагая, что эти метрики могут отражать человеческое восприятие гармоничности между мелодией и аккордами.

Из результатов метрики аккордовых прогрессий, полученных с помощью СНЕ и СС, мы также видим, что прогрессии, созданные методами, основанными на сопоставлении шаблонов и НММ, кажутся недостаточно разнообразными. [1] В отличие от них, результаты работы Генетического Алгоритма отличаются большим разнообразием. [1]

Поскольку метод, основанный на ГА, получил более низкий рейтинг, чем методы согласования шаблонов и НММ по критерию «в общем» в субъективной оценке, похоже, что испытуемые больше заботятся о гармоничности, чем о разнообразии аккордовых прогрессий.

Сравнивая два метода глубокого обучения, видно, что MTHarmonizer использует больше неаккордовых тонов и большее количество уникальных аккордов, что говорит о его высоком критерии интересности: такая музыка будет звучать «не замылено».

**4.3 Оценка наилучшей модели**

В целом, результаты объективной оценки совпадают с результатами субъективной оценки. Трудно количественно определить, какие метрики лучше для каких целей и насколько полезны и точны эти метрике в целом. Поэтому, предлагается использовать их, в основном, для получения практического представления о результатах работы модели автоматической гармонизации мелодий, а не для оценки их качества. [1]

Объективный метрикой, например, можно пользоваться для отслеживания производительности модели во время разработки, прежде чем приступать к пользовательскому исследованию. Тем не менее, для оценки качества генерируемый музыки все ещё необходимы человеческие оценки.

Однако, все же, наиболее удачной по всем параметрам, можно назвать модель MTHarmonizer:

* MTHarmonizer очень близка к оригинальным прогрессиям, составленным человеком. Так же, данная модель достигает самой низкой ошибки предсказания аккордов. [1]
* На тестовом наборе точность предсказания аккордов для пяти моделей (сопоставления шаблонов, HMM, GA, BILSTM и MTHarmonizer) составляет **29%, 31%, 20%, 35%, 38%** соответственно. Точность предсказания функции аккорда составляет **62%, 61%, 55%, 65%, 69%.** [1]

Полученная оценка показывает, что модели глубокого обучения действительно работают лучше, чем модели без глубокого обучения по целому ряду аспектов включая гармоничность и интересность. [1] Более того, модель глубокого обучения, учитывающая функцию аккордов, достигает наилучшего результата, среди всех оценённых моделей. [1]

1. РАЗРАБОТКА ПРОГРАММЫ ДЛЯ ГАРМОНИЗАЦИИ МЕЛОДИЙ

Оценив производительность наиболее известных моделей гармонизации, мы разработали свою программу. Программа была написана в среде Google Collab на языке программирования Python. Простота этого языка помогает разработчикам создавать надежные системы, не отвлекаясь на синтаксис и технические нюансы.

* 1. **Функционал программы**

Наша программа принимает аудиофайл, обрабатывает его и выдает текстовые обозначения аккордов "C", "D", "Dm", "E", "Em", "F", "G", "A", "Am" или "Bm", по которым можно сыграть большинство существующих мелодий. В дальнейшей работе планируется пополнять словарь аккордов, чтобы увеличивать интересность аккордовых прогрессий и придавать мелодиям разнообразие звучаний.

В программе можно обозначить количество секунд, которое тот или иной аккорд проигрывается. С данной функцией каждому аккорду на выходе программы соответствует количество секунд, которое длится данный аккорд в музыкальном произведении – это может быть очень важно для начинающих музыкантов.

К слову, в нашей программе настроена обработка не только стереосигнала, но и моно-сигнала.

Одноканальный формат звука называется моно-сигналом. Он достигает ушей слушателя одновременно, а точка его поступления находится строго по центру. В моно-сигнале нет информации о том, что содержит левый правый канал, а нахождение сигнала относительно расположения слушателя не определено. Первый изобретенный формат звукового сигнала.

Стерео формат имеет двухканальный способ трансляции звука и используется во всех современных наушниках. Двухканальный способ трансляции означает, что информация о звуке в левом и правом ухе может различаться (бенуарный эффект).

Если в нашу нейронную сеть подается стерео-звук, то для предсказания я буру только первый его канал. Это было вычислено мной эмпирическим путем.

Программа позволяет подбирать размер окна (его длительность в секундах) для разных инструментов. Напоминаем, что окно – это деление звукового сигнала: например, для распознавания гитарного звука, окно длится одну секунду – одна секунда = 1 звук/аккорд. То есть, мы можем сделать окно шире во временном промежутке для того, чтобы вместить в него более протяжные звуки виолончели или фортепиано. Таким образом, будеть улучшаться производительность модели.

1. АЛГОРИТМ РАБОТЫ ПРОГРАММЫ
   1. **Работа с wav файлами**

WAV - файлы можно прослушивать как и с мультимедийных плееров, так и с помощью многих цифровых приложений для Microsoft. [17] Так же, их можно просмотреть на платформах Linux и Mac. [17] По сравнению с файлами формата MP3, WAV файлы больше из-за того, что цифровой аудио-контент в них не сжимается. Звуковые данные, которые содержатся в этих файлах, называют сигналы, и данные сигналы можно реализововать с различной частотой дискретизации и битрейтом.

Для начала мы подготавливаем файл типа wav с нужной нам композицией и извлекаем оттуда частоту аудио и другую информацию. Данная информация показывает амплитуду для каждого фрагмента аудио, которую мы можем получить, благодаря PCM (Pulse-code modulation) – методу оцифровки фрагментов аналогового сигнала.

* 1. **Преобразование Фурье**

Затем преобразуем сигнал при помощи преобразования Фурье. Преобразование [Фурье](https://ru.wikipedia.org/wiki/%D0%A4%D1%83%D1%80%D1%8C%D0%B5,_%D0%96%D0%B0%D0%BD_%D0%91%D0%B0%D1%82%D0%B8%D1%81%D1%82_%D0%96%D0%BE%D0%B7%D0%B5%D1%84)— операция, сопоставляющая одной функции вещественной переменной другую функцию вещественной переменной. [10] В новой функции описываются амплитуды или коэффициенты при разложении первоначально взятой функции на простые составляющие — гармоничные звуковые колебания с разными значениями частот.

Преобразование Фурье функции f вещественной переменной является интегральным [10] и задаётся следующей формулой:

* 1. **Pitch Class Profile**

f^(ω)=12π∫−∞∞f(x)e−ixωdx.

Далее из этих данных мы формируем Pitch Class Profile. Это двенадцатимерный вектор, который показывает интенсивность двенадцати классов высот полутонов. Например, первый элемент PCP показывает, насколько сильный аккорд “C” в целом. [8]

, для

*–* это таблица, которая сопоставляет индекс части спектра с индексом PCP:

Где – это опорная частота, которая попадает в PCP(0).

– показывает частоту спектра в .

* 1. **Нормировка PCP**

После чего мы нормируем наш вектор интенсивности PCP для того, чтобы наша нейронная сеть лучше обучалась:

, для

Выполнив все преобразования в PCP, мы подаем двеннадцатимерный вектор в полносвязную нейронную сеть. [8]

* 1. **Полносвязная нейронная сеть**

В полносвязной нейронной сети прямого распространения каждый нейрон имеет связи с остальными нейронами, которые находятся в соседних слоях. В нейронной сети такого типа все связи имеют направление строго от входных нейронов к нейронам на выходе.

Такие нейронные сети хорошо подходят для задач классификации. Для этого, в последнем слое устанавливается такое количество нейронов, сколько классов мы хотим предсказать. В нашем случае количество аккордов равняется 10, соответственно последний слой состоит из 10 нейронов. Для задач такого типа, чаще всего функцией активации выбирается сигмоида – функция, показывающая вероятность принадлежности к тому или иному классу. Она возвращает число от 0 до 1, где 1 – точная принадлежность к определенному классу, а 0 – точная непринадлежность.

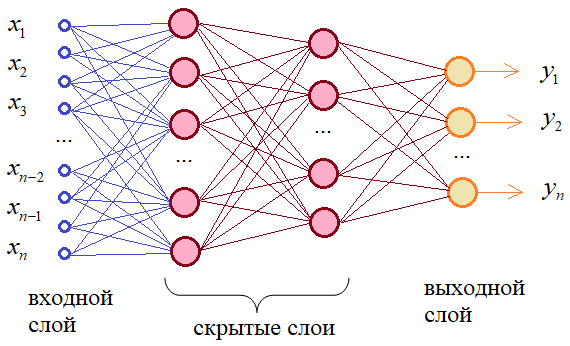


Рис. 7 архитектура полносвязой нейронной сети

На изображении вы можете увидеть входные слои нейронной сети слева (синий цвет), затем два скрытых слоя посередине (розовый цвет), и справа находятся выходные слои (желтый).

* + 1. Архитектура полносвязной нейронной сети

Наша нейронная сеть состоит из 3 слоев. Первый – он же входной – состоит из 30 нейронов и входной размерностью 12. Активационная функция ReLU.

Второй слой скрытый состоит из 20 нейронов с активационной функцией ReLU.

Последний слой выходной состоит из 10 нейронов. В итоге получается 10 выходов с активационной функцией сигмоида.

* + 1. Функция активации (Rectified Linear Unit)

Rectified Linear Unit — эта наиболее часто встречающаяся функция активации при глубоком обучении. Она возвращает 0, если аргумент отрицателен, однако, если аргумент положителен, функция возвращает тот же сигнал [18]:

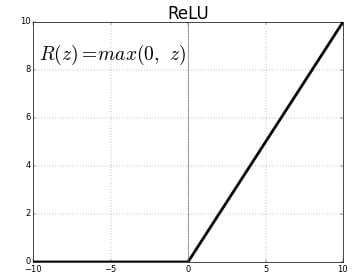


Рис. 8 график функции Rectified Linear Unit [11]

Вот несколько преимуществ функции ReLU перед сигмоидой и гиперболическим тангенсом:

1. Быстрый и легкий подсчет производной. Отрицательные аргументы принимают значение — 0, положительные — сами себя. [18]
2. Разреженность активации. При обучении сетей с большим количеством нейронов может возникнуть проблема производительности обучения при использовании гиперболического тангенса или сигмоидной функции – задействуются все нейроны сразу. Что же касается использования ReLU ­— данная функция приводит к нулю отрицательные значения входного сигнала, что облегчает математические вычисления. [18]
   * 1. Функция активации (сигмоидная функция)

Сигмоидную функцию также называют логистической. Она гладкая, нелинейная и монотонно возрастающая [11]:

Так как в особенности сигмоидной функции входит гладкий градиент, то наличие "прыжков" при подсчете выходного значения резко уменьшается. [11]

Кроме всего этого, значительным плюсом функции сигмоиды для значений и является способность "прижимается" [11] к одной из асимптот — это делает предсказания классов более четкими и интерпретируемыми.

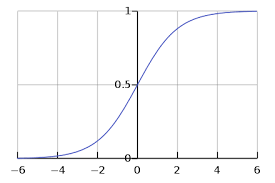


Рис. 9 график сигмоидной функции активации [11]

Значительным недостатком такой функции является то, что при определенных значениях появляется эффект исчезновения градиента. При нем производная близка к нулю, что может привести к медленному обучению или вовсе к недообучению сети.

* 1. **Обучение нейронной сети**

Для начала нами задается функция потерь, где в качестве функции потерь мы берем бинарную кросс энтропию:

Нашей следующей задачей является определение оптимизатора. Мы выбрали оптимизатор Аdam по нескольким причинам: в обучении нейрсетей Adam — один из самых эффективных алгоритмов для оптимизации. [5]

Его преимущества:

* Прост в реализации.
* Хорошая эффективность вычисления.
* Не требует много памяти.
* Идеален для решения задач с разреженными или очень шумными градиентами.
  + 1. Обучение с учителем

При обучении данной нейронной сети, мы использовали обучение с учителем. Обучение с учителем (**supervised learning**) предполагает, что обучать модель мы будем на заранее размеченных данных. [7] То, что мы имеем полностью размеченный датасет, означает, что примерам из обучающего набора уже соответствуют их правильные ответы на выходе.

То есть, к примеру, при обучении с учителем на данных из фотографий разных пород кошек мы заранее подготавливаем связку фотография – порода. [8] Метка класса породы выступает в роле учителя, как бы, говоря нейронной сети, какой ответ мы желаем увидеть на ту или иную фотографию.

Обучение с учителем прекрасно подходит для решения задач классификации и регрессии. Как мы помним, перед нами стоит именно задача классификации.

В задачах на классификацию алгоритм делает предсказание дискретного значения. Каждому значению соответствует номер класса. В нашем случае, класс содержит возможные звучания конкретного аккорда. Так как аккордов у нас всего 10, то мы имеем 10 классов соответственно: "C", "D", "Dm", "E", "Em", "F", "G", "A", "Am", "Bm". В обучающем датасете с фотографиями кошек будет некоторое количество классов, где каждый класс будет содержать в себе конкретную породу кошки. [8]

Качество такого алгоритма можно оценить тем, насколько точно он предскажет метки аккорда для такого трезвучия из датасета, которое он еще не слышал [7]. Например, сильно выражена нота «До» и слабо нота «Соль».

Таким образом, обучение с учителем подходит для нашей цели, так как мы имеем достаточный набор уже размеченных данных, на которые будет ориентироваться сеть при обучении. На самом деле, набор данных у нас не большой – но на нем мы можем обучить нейросеть распознавать основные аккорды, из которых состоит любая мелодия. То есть, используя такой набор аккордов, вполне можно достичь хорошего и полезного результата гармонизации мелодий.

Наш исходный датасет был разделен на тестовые данные «трейн» и тестовые данные «тест». Это сделано, чтобы проверить качество модели на тех примерах, которые она еще не видела: тестовые данные модель получает впервые, они отличаются от «трейна», поэтому важно, чтобы нейронная сеть адекватно реагировала на подачу новых данных из нашего датасета. [12]

Возвращаясь назад, стоит заметить, что в нашем случае набор данных представляет из себя 200 примеров всевозможного звучания одних и тех же нот для каждого конкретного аккорда. Каждый пример обозначается как 12-мерный вектор, который получается после преобразования PCP.

* + 1. Параметры нейронной сети

Продолжая разговор об обучении модели, стоит отметить функцию fit. [12] Мы передаем на нее ответы (у) – правильные ответы (метка аккорда) и (х) – численно закодированный аккорд.

Так же мы задаем количество эпох. Количество эпох определяет сколько раз тренировочные данные будут прогоняться через нейронную сеть в момент обучения. Наша нейронная сеть содержит 100 эпох.

В свою очередь, каждая эпоха делится на «батчи» (от англ. batch). Это маленькие партии обучающей выборки, которые передаются в нейронную сеть за раз. После прогона такого батча, как раз и идет корректирование весов связей между нейронами, что и называется обучением нейронной сети. В нашем случае, размером батча является 10.

* + 1. Метрики качества.

Метрики качества вычисляются в ходе тестового прогона нейронной сети, чтобы понять, как она предсказывает результаты гармонизации.

В работе была проведена гармонизация мелодий следующих музыкальных инструментов: гитара, пианино и виолончель. Для гитары точность предсказания составила 99 процентов. Для пианино точность составила 88 процентов. Аккорды виолончели нейронная сеть предсказала чуть хуже - с точностью до 81 процента. Необходимо отметить, что для предсказания в десять классов это хорошая точность.

Как мы видим, для гитарной музыки предсказания получились наиболее точные. Все это потому, что тягучие звуки виолончели искажают начало и конец ноты, и их становится довольно сложно правильно разграничить. Что касается пианино, то тут четкому восприятию мешает педаль, которая так же делает звуки менее ясными и отчетливыми. У гитары же, из-за сравнительно недолгого колебания струны, переходы от аккорда к аккорду ярко выражены, что не вызывает проблем при гармонизации мелодии.

Что касается Top-k Accuracy, то она у нас равняется 100 процентов для всех музыкальных инструментов. Эта метрика оценки точности классификации вычисляет находится ли правильный ответ в числе k предсказаний наиболее вероятных меток класса. То есть, насколько часто ответ появляется в топе предсказаний.

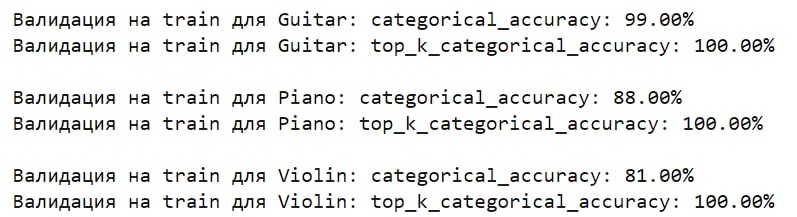


Рис. 10 оценки качества гармонизации нашей модели

**6.7 Предсказания модели**

Для предсказания результатов мы должны проделать те же операции по предобработке звукового сигнала, а затем прогнать получившийся 12-мерный вектор через нейронную сеть. гармонизации мелодии не слишком отличается от описанного выше алгоритма.

Для начала мы выполним все те же преобразования для входного сигнала, а именно:

1. Поделим песню на «окна» в некоторых промежутках времени.
2. Для каждого такого «окна» выполним преобразование Фурье.
3. Далее из этих данных мы формируем Pitch Class Profile - двенадцатимерный вектор интенсивности высот полутонов.
4. Затем, мы подаем этот вектор в нейронную сеть.
5. На выходе мы получим текстовый файл с буквенными описаниями аккордов. Мы видим его на рисунке 11. Эта последовательность аккордов песни «About a girl» группы «Nirvana». Ниже на рисунке 12 представлены оригинальные аккорды, взятые с сайта Ultimate guitar. Хоть песня и не сложная, но хороший результат вдохновляет на последующие достижения в области музыкальной гармонизации.



Рис. 11 текстовый файл с аккордами – результат работы модели

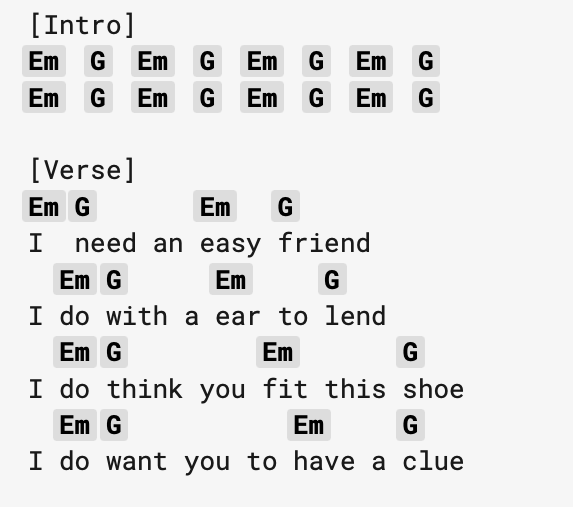


Рис. 12 отрывок песни с аккордами, взятый с сайта Ultimate guitar

1. ИТОГИ

Проанализировав различные модели гармонизации (согласование шаблонов, скрытая Марковская модель (HMM), генетический алгоритм, глубокая нейронная сеть на основе BILSTM и глубокая нейронная сеть MTharmonizer) и выяснив, что наилучший результат дают модели, основанные на глубоком обучении, мы создали свою модель с использованием глубокой полносвязной нейронной сети. Программа для гармонизации мелодий показала очень хорошие результаты для распознавания гитарной музыки. Предполагается, что данная модель нейронной сети будет хорошо распознавать бардовские песни, лирические композиции, рок музыку, так как в музыке такого типа, как правило ведущий инструмент – гитара. Для гитары точность предсказания составила 99 процентов.

Что же касается фортепиано и виолончели, здесь модель обучилась чуть хуже – на это повлияли длинные и тягучие особенности звуков данных инструментов. Но это означает только пространство для дальнейших исследований.

Работа выложена в общий доступ, к ней написаны подробные комментарии с целью максимально понятно донести до людей цель и устройство программы. Ссылку прикрепляю: <https://github.com/Ustinova1234567/chrod_recognition>

1. БУДУЩАЯ ИССЛЕДОВАТЕЛЬСКАЯ ДЕЯТЕЛЬНОСТЬ

В ходе проведения исследования выяснилось, что наша программа плохо распознает аккорд “G”. Это связано с тем, что по амплитуде “G” похож на аккорд “F”. В дальнейшей работе данные будут размечены руками, а «сэмплов» будет не 200, как сейчас, а намного больше, что позволит избежать путаницы в аккордах.

Так же в наших интересах усовершенствовать данную программу до уровня распознавания музыки с несколькими инструментами, звучащими одновременно.

Более того, было бы полезно провести свое исследование оценки качества распознавания мелодий на большой выборке. Для этого нам нужно прописать критерии хорошей, правильной и приятной гармонии. Все-таки, мы выложили код в общий доступ и хотим, чтобы он был полезен людям.

СПИСОК ИСПОЛЬЗОВАННОЙ ЛИТЕРАТУРЫ

[1] “Автоматическая гармонизация мелодий с помощью аккордов триады”. Yin-Cheng Yeh, Wen-Yi Hsiao, Benjamin Genchel. (2021) [Электронный ресурс]. URL: <https://www.tandfonline.com/doi/full/10.1080/09298215.2021.1873392> (дата обращения 21.08.2021)

[2] “Deep learning techniques for music generation: A survey”. Briot, J. P., Hadjeres, G., & Pachet, F. (2017). Preprint arXiv:1709.01620. [[Google Scholar]](http://scholar.google.com/scholar?hl=en&q=Briot%2C+J.+P.%2C+Hadjeres%2C+G.%2C+%26+Pachet%2C+F.+%282017%29.+Deep+learning+techniques+for+music+generation%3A+A+survey.+Preprint+arXiv%3A1709.01620.) (дата обращения 08.04.2022)

[3] “Генетические алгоритмы: фитнесс – функция”. [Электронный ресурс]. URL: <https://coderlessons.com/tutorials/akademicheskii/izuchite-geneticheskie-algoritmy/geneticheskie-algoritmy-fitnes-funktsiia> (дата обращения 05.04.2022)

[4] “Chord generation from symbolic melody using BLSTM networks. In Proceedings of the international society for music information retrieval conference”. Lim, H., Rhyu, S., & Lee, K. (2017). (pp. 621–627). [[Google Scholar]](http://scholar.google.com/scholar?hl=en&q=Lim%2C+H.%2C+Rhyu%2C+S.%2C+%26+Lee%2C+K.+%282017%29.+Chord+generation+from+symbolic+melody+using+BLSTM+networks.+In+Proceedings+of+the+international+society+for+music+information+retrieval+conference+%28pp.+621%E2%80%93627%29.) (дата обращения 08.03.2022)

[5] “Automatic functional harmonic analysis”. De Haas, W. B., Magalhães, J. P., Wiering, F., & Veltkamp, R. C. (2014). Computer Music Journal, 37(4), 37–53. https://doi.org/10.1162/COMJ\_a\_00209 [[Crossref]](https://www.tandfonline.com/servlet/linkout?suffix=CIT0007&dbid=16&doi=10.1080%2F09298215.2021.1873392&key=10.1162%2FCOMJ_a_00209), [[Web of Science ®]](https://www.tandfonline.com/servlet/linkout?suffix=CIT0007&dbid=128&doi=10.1080%2F09298215.2021.1873392&key=000329510700004), [[Google Scholar]](http://scholar.google.com/scholar_lookup?hl=en&volume=37&publication_year=2014&pages=37-53&issue=4&author=W.+B.+De+Haas&author=J.+P.+Magalh%C3%A3es&author=F.+Wiering&author=R.+C.+Veltkamp&title=Automatic+functional+harmonic+analysis&doi=https%3A%2F%2Fdoi.org%2F10.1162%2FCOMJ_a_00209) (дата обращения 08.03.2022 )

[6] “Realtime Chord Recognition of Musical Sound: a System Using Common Lisp Music”, Takuya Fujishima, 1999 (дата обращения 05.04.2022)

[7] “Neural Networks For Musical Chords Recognition”, J. Osmalsyj, J-J. Emrechts, S. Pierard, M. Van Droogenbroeck, 2012 (дата обращения 15.04.2022)

[8] “Neural networks for musical chords recognition” J. Osmalskyj, J-J. Embrechts, S. Piérard, M. Van Droogenbroeck. NTELSIG Laboratory, University of Liège, Departement EECS (дата обращения 15.04.2022)

# [9] “Структура и принцип работы полносвязных нейронных сетей”. [Электронный ресурс]. URL: <https://proproprogs.ru/neural_network/struktura-i-princip-rabotypolnosvyaznyh-neyronnyh-setey> (дата обращения 20.04.2022)

## [10] “Fourier Transform at Action”. [Электронный ресурс]. URL: <http://makeloft.xyz/lessons/signal-processing/fourier-transform> (дата обращения 20.04.2022)

# [11] “Функции активации нейросети: сигмоида, линейная, ступенчатая, ReLu, than” [Электронный ресурс]. URL: <https://neurohive.io/ru/osnovy-data-science/activation-functions/> (дата обращения 16.05.2022)

# [12] “Deep Learning Techniques for Music Generation -- A Survey”. [Jean-Pierre Briot](https://arxiv.org/search/cs?searchtype=author&query=Briot%2C+J), [Gaëtan Hadjeres](https://arxiv.org/search/cs?searchtype=author&query=Hadjeres%2C+G), [François-David Pachet](https://arxiv.org/search/cs?searchtype=author&query=Pachet%2C+F). [Электронный ресурс]. URL: <https://arxiv.org/abs/1709.01620> (дата обращения 16.05.2022)

[13] “Метод генетического программирования на основе моделирования противоборствующих популяций автоматных программ”. Кулев Владимир Анатольевич. [Электронный ресурс]. URL: <http://is.ifmo.ru/diploma-theses/_kulev_report.pdf> (дата обращения 21.05.2022)

# [14] “Сопоставление с образцом - Pattern matching”. [Электронный ресурс]. URL: <https://wiki5.ru/wiki/Pattern_matching> (дата обращения 23.05.2022)

# [15] “Bidirectional LSTM”. [Электронный ресурс]. URL: <https://paperswithcode.com/method/bilstm> (дата обращения 21.05.2022)

# [16] “Скрытая марковская модель (HMM)”. [Электронный ресурс]. URL: <https://russianblogs.com/article/2372866572/> (дата обращения 29.04.2022)

# [17] “WAV.  Расширение файла”. [Электронный ресурс]. URL: <https://www.reviversoft.com/ru/file-extensions/wav> (дата обращения 22.05.2022)

[18] “Практики реализации нейронных сетей”. [Электронный ресурс]. URL: <https://neerc.ifmo.ru/wiki/index.php?title=Практики_реализации_нейронных_сетей>