Національний технічний університет України

«Київський політехнічний інститут імені Ігоря Сікорського»

Радіотехнічний факультет

Кафедра радіоконструювання та виробництва радіоапаратури

**Дипломна робота**

**на тему: «Методи классифікації нот в акустичному сигналі»**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Керівник: | |  |  | Виконав | Красницький |
| Шпилька О.О. | |  |  | Група: | РТ-71 |
|  | |  |  |  | |
| Комісія: | |  |  | (ПІБ) | |
|  |  |  |  |  | |
|  |  |  |  | Дата | Підпис |

Київ — 2021 р.

1. Анотація

Дипломна робота викладена на 60 сторінках, містить 23 ілюстрацій, 2 таблиці та 4 додатки.

Метою цього проекту є огляд базових класифікаторів та створення алгоритму классифікації акустичних сигналів. А саме, такий класифікатор буде визначати яка нота грає на аудиозаписі.

1. Анотация

Дипломная работа изложена на 60 страницах, имеет 23 иллюстрации, 2 таблицы и 4 приложения.

Целью

Змі**с**т

[Анотація 2](#_Toc72355686)

[Анотация 3](#_Toc72355687)

[Зміст 4](#_Toc72355688)

[Вступ 5](#_Toc72355689)

[1 Теорія музики 6](#_Toc72355690)

[1.1 Теорія нот 6](#_Toc72355691)

[1.2 Фізика нот 8](#_Toc72355692)

[1.3 Фізика тембру 9](#_Toc72355693)

[2 Теорія класифікації 15](#_Toc72355694)

[2.1 Основи класифікації 15](#_Toc72355695)

[2.2 Байесівска класифікація. Гаусовський наївний баєсів класифікатор 16](#_Toc72355696)

[2.3 Методи регресивного аналізу. SVM 20](#_Toc72355697)

[3 Дослідження методів классифікації 23](#_Toc72355698)

[3.1 Емпіричні методи класифікації 23](#_Toc72355699)

[3.1.1 Выбор входного сигнала для классификации 23](#_Toc72355700)

[3.1.2 Длинна сигнала 23](#_Toc72355701)

[3.1.3 Домен сигнала (временной-частотный) 26](#_Toc72355702)

[3.2 Гаусівський наївний баєсів класифікатор (Gaussian naïve Bayes) 27](#_Toc72355703)

[3.3 SVM класифікатор 29](#_Toc72355704)

[3.4 Застосування машинного навчання з підкріпленням 29](#_Toc72355705)

1. Вступ

# Теорія музики

## Теорія нот

Нота – это графическая единица музикальной нотации [1] либо музыкальная элементарная единица композиции. Любая классическая композиция состоит из нот (звуков) и пауз. В графическом представлении ноты изображаются так, как показано на Рисунку 1.1.

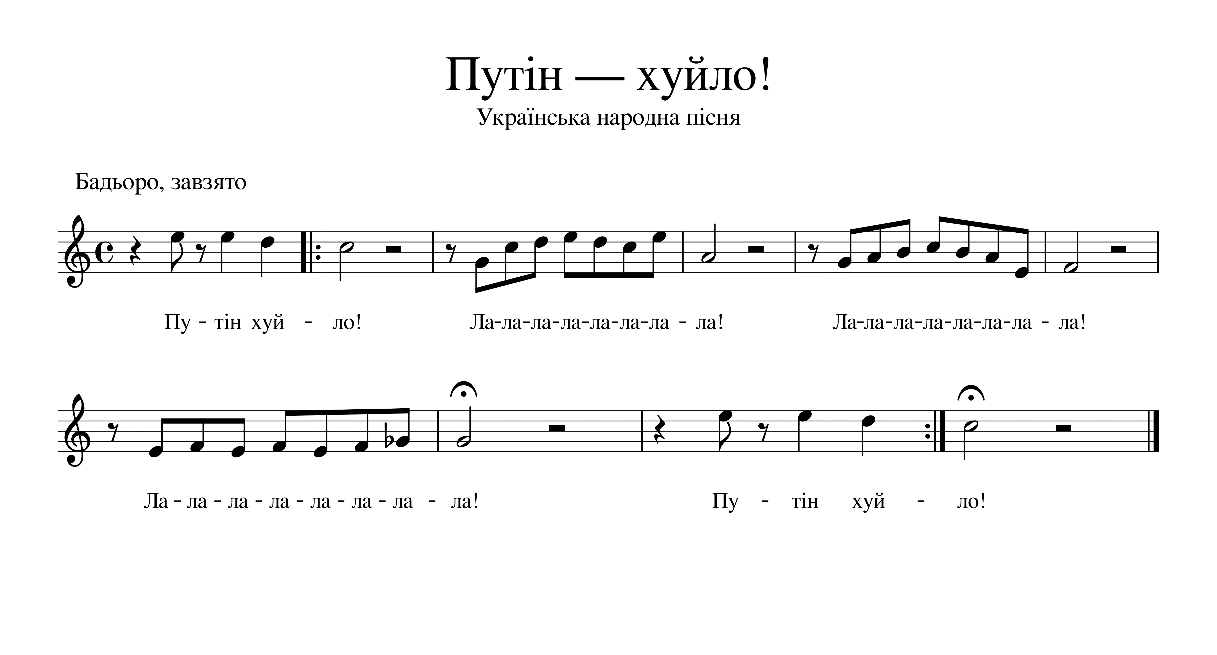


Рисунок 1.1 – Пример описания композиции нотной записью

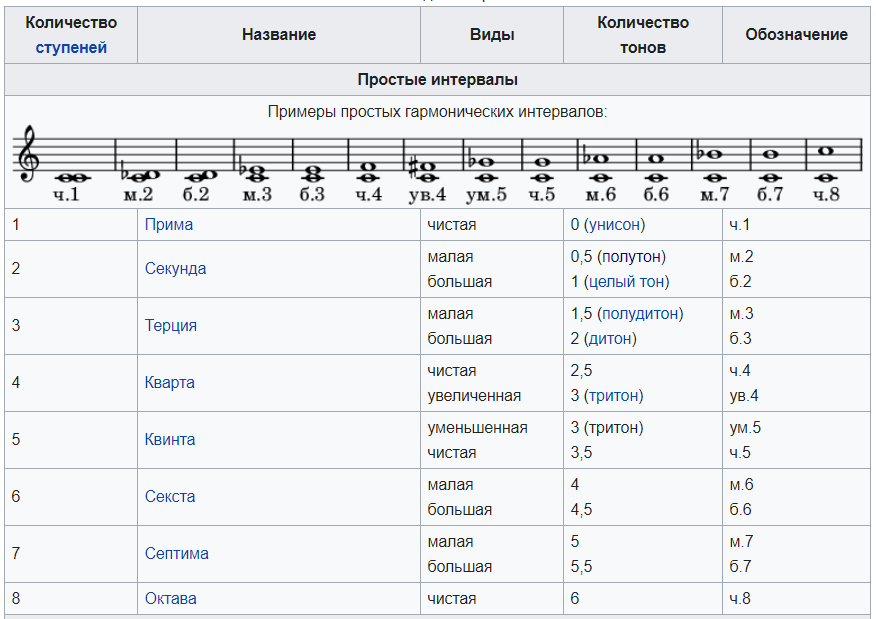
Физически ноты представляют собой вибрации в пределах диапазона звуковых частот человека. Всего нот 7: До, Ре, Ми, Фа, Соль, Ля, Си.

Полутон — наименьший интервал в традиционной и академической музыке Европы. Ноты Ми и Фа, а так же ноты Си и До находятся на расстоянии полутона, тогда как другие соседние ноты находятся на расстоянии тона.

В теории музыки используются различные интервалы. В Таблиці 1.1 показана классификация простых интервалов.

Знаками альтерации [3], такими как бемоль и диез, в нотной нотации обозначают ноту которая на пол тона ниже или выше соответственно. Для этого существуют отдельные ноты, которые называются через основные 7, к примеру «До диез» или «Ля бемоль». В силу того что 2 пары нот изначально были с расстоянием полутон – к изначальным 7ми нотам добавляется только 5. Итого получаем 12 нот на октаву.

Таблиця 1.1 – Класифікація простих інтервалів [2]



На пианино 7 основных нот показаны белыми клавишами, тогда как их диезы и бемоли показаны черными, що можна побачити на Рисунку 1.2.

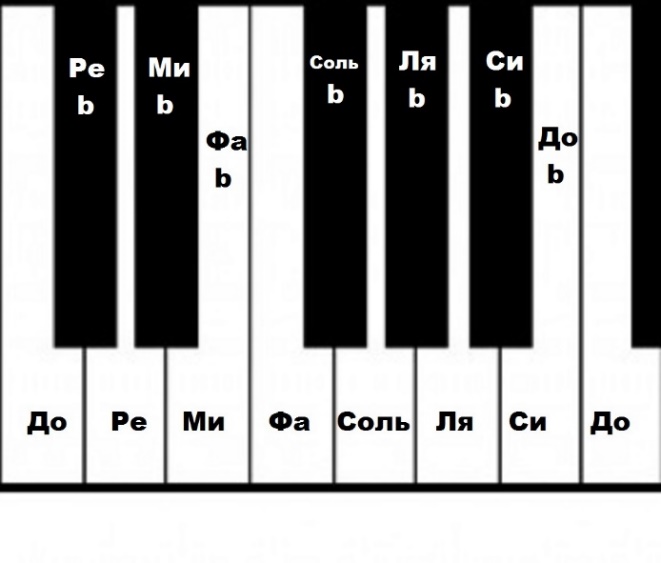


Рисунок 1.2 – Зображення октави на піаніно

Стоит заметить что диезы и бемоли взаимозаменяемы. Таким образом «До диез» и «Ре бемоль» одна и та же нота. Так как бемоли и диезы могут быть записаны в начале строки, у скрепичного ключа и распространяться на всю композицию ввели знак альтернации – бекар, который отменяет действие диеза и бемоля на выбранную ноту. Знаки дієза, бемоля та бекара можно побачити на Рисунку 1.3.

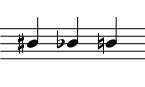


Рисунок 1.3 – Знаки альтерації з ліва направо: дієз, бемоль, бекар

## Фізика нот

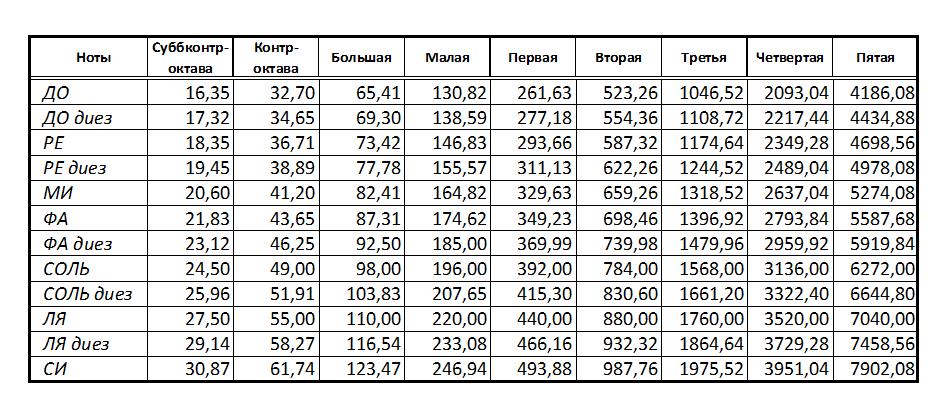
Фізично нота – це механічне коливання на визначеній частоті. Есть следующие системы распределения частот нотам:

1. Пифагоров строй
2. Чистый строй
3. Равномерно темперированный строй

ОПИСАНИЕ СТРОЕВ ДОБАВИТЬ!!!

Є декілька систем розподілу частот для кожної ноти. Найпоширеніший з них – равномерно темперирований строй [4]. Для такой системы верно следующее утверждение. Соотношение частот следующей и предыдущей ноты, расстояние между которыми – полутон, равно . За стандарт частоты принята нота Ля первой октавы с частотой 440Гц. От нее математически можно найти частоту любой ноты на любой октаве. На Таблиці 1.2 представлено соответствие нотам всех октав их частоты.

Таблиця 1.2 – Частоти нот кожної октави



Таким образом ноты с растоянием в октаву имеют кратные частоты. Нетрудно заметить, что с увеличением октавы частоты нот увеличиваются как показательная функция с основанием стемени равным 2. На Рисунку 1.4 каждым цветом показана функциональная зависимость частоты от октавы для каждой отдельной ноты. Синяя линия – До, зеленая – Си. Из рисунка видно, что с увеличением октавы увеличивается абсолютное значение частотного расстояния между нотами (относительное остается константным по принципу равномерно темперированного строя).

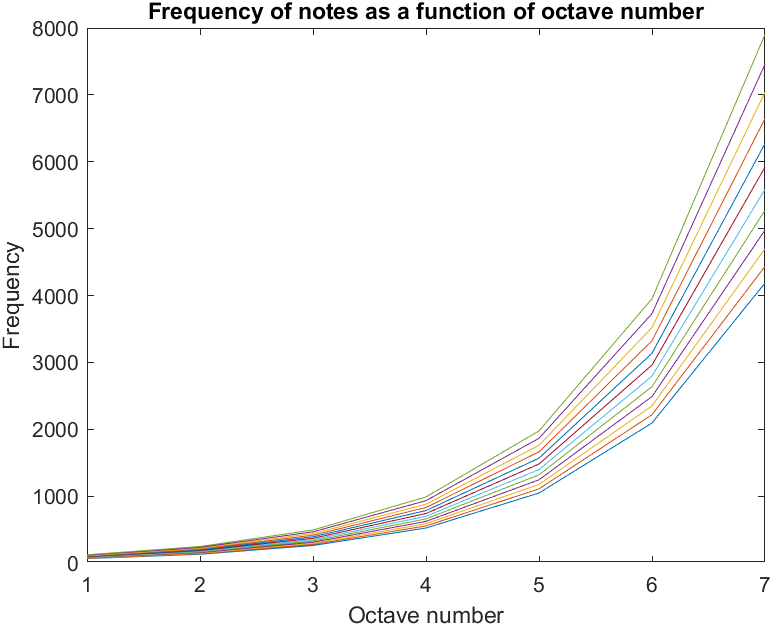


Рисунок 1.4 – Графік відповідністі нотам октав частот

## Фізика тембру

Ноты полностью определяются частотами, но звучание ноты одной и той же частоты на разных инструментах звуит с разной окраской. Это называется тембром.

Начнем с того, что что бы музыкальный инструмент извадал звук ноты, какая то его часть должна удовлетворять физическим условиям к механическим колебаниям на частоте ноты.

У струнных инструментов такими элементами есть струны. Их длинна определяет длинну половины длинны волны которую они могут возбудить в механическом колебинии. А натяжение, материал и прочие характеристики струны определяют скорость распространения этой волны. Таким образом создаются звучания на частотах нот.

У духовых музыкальных инструментов, таких как орган, используются трубы. Длинна трубы подбирается так, что бы в нее вложилась полна длинна волны что будет возбуждаться. Воздушный поток возбуждает в трубе механические колебания. (В зависимости от скорости воздушного потока могут возбуждаться разные кратные гармоники. Об этом лучше позже, но это важно!) В таких инструментах, как кларнет, частота задается закрытием/открытием клапанов, что определяет длинну стоячей волны внутри трубы, что показано на Рисунку 1.5.



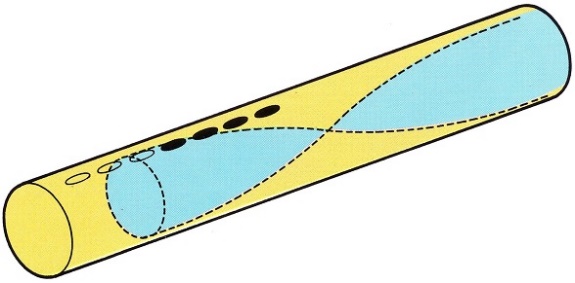
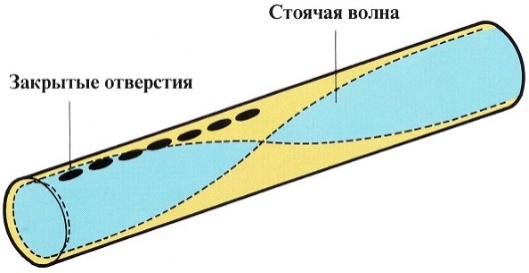


Рисунок 1.5 – Формування стоячої хвилі в трубі в залежності від стану клапонів[5]

Рассмотрим понятие тембра более подробно на примере струнных инструментов с точки зрения большей наглядности. Струна представляет собой длинный отрезок гибкого материала, находящийся в натянутом состоянии, благодаря чему может свободно колебаться. Изначально у струны есть 2 узла. На гитаре эти узлы расположены на нижнем и верхнем порожке соответственно. Єти участки гитары можно увидеть на Рисунку 1.6.



Рисунок 1.6 – Конструктивні частини класичної гітари

Таким образом при возбуждении струны она будет колебаться волнами, узлы которых ложатся в узлы на порожках. Докозательство этого факта можно найти из пособия математической физики в разделе «Свободные колебания струны с закрепленными концами»[6]. Так, самый низкочастотный возбужденный звук получается у волны с самой большой длинной волны. И половина длинны такой волны равна длинне всей струны. Такой тон называют основным, у него два узла и максимальная амплитуда колебания по середине струны. На Рисунку 1.7 ниже можно посмотреть как меняется состояние струны при колебаниив основном тоне со временем.

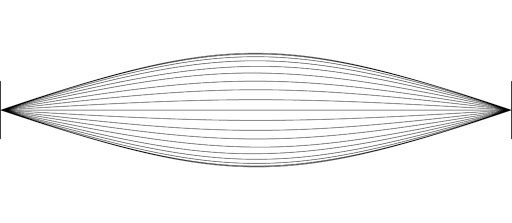


Рисунок 1.7 – Коливання струни на основному тоні в різні моменти часу

Кроме основного тона струна колеблется всеми другими волнами у которых отношение длинны струны и половины длинны дают целое число. Несложно прийти к тому, что эти волны будут кратны по частоте основному тону, что можно увидеть на Рисунку 1.8.

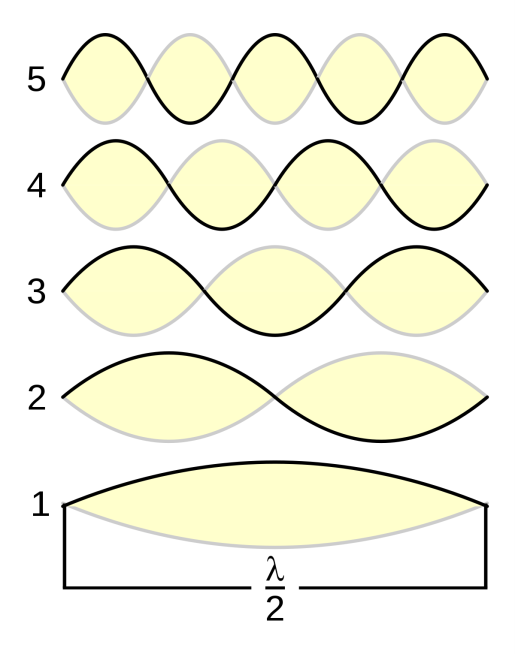


Рисунок 1.8 – Зображення довжин хвиль, що збуджуються в струні

Тони, що є кратними по частоті основному тону називають обертонами або гармоніками. Вони мають деякі властивості: частотна відстань між обертонами рівна частоті основного тону, зазвичай амплітуда основного тону є найвища.

Так как распределение амплитуд обертонов относительно основного тона есть индивидуальным для каждого инструмента, получаем тембры инструментов. Так, на Рисунку 1.9 можно увидеть распределение амплитуд обертонов некоторых инструментов.

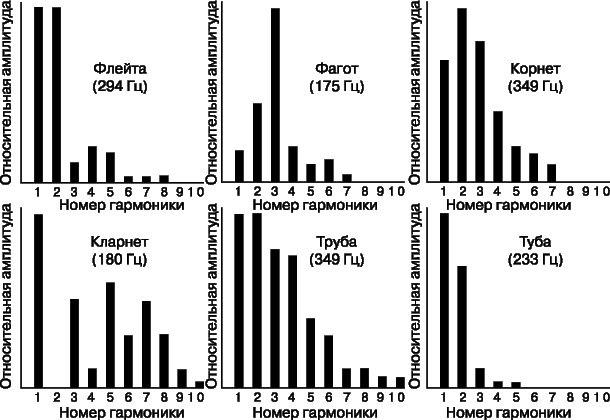


Рисунок 1.9 – Спектральні показники тембру флейти, фаготу, корнету, кларнету, труби та туби

Так же стоит упомянуть что и сам инструмент может усиливать некоторые обертоны. К примеру если дека скрипки резонирует с каким то тоном – он будет усилен в спектре. Картины колебаний верхней и нижней деки можно увидеть на

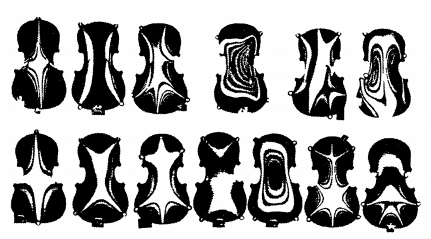


Рисунок 1.10 – Формы колебаний верхней и нижней деки [7]

Вырезы в деке скрипки, называемые эфы, созданы именно для усиления явления резонанса корпуса. В то же время полоски дерева, называемые обечайками, которые серепляют верхнюю и нижнюю плоскость деки, образуют особые своды. Геометрия сводов, их толщина и распределение так же влияет на силу звука и на тембр инструмента. Конструкцию скрипки можно увидеть на



Рисунок 1.11 – Конструктивні елементи скрипки

Распределение амплитуд не остается постоянным во времени, без вынужденных колебаний (внешней поддержки, смычок поток воздуха). Звучащие элементы музыкальных инструментов имеют разные коэффициенты затухания для разных частот, от чего отношение обертонов к основному тону может меняеться во времени. Так, на Рисунку 1.12 можно посмотреть на временные спектрограммы флейты и гобоя, а на Рисунку 1.13 на отношение обертонов гитарной струны по времени.

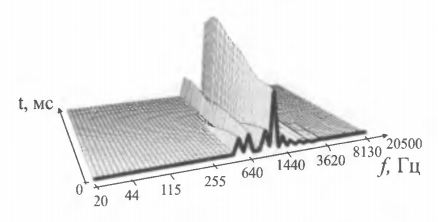
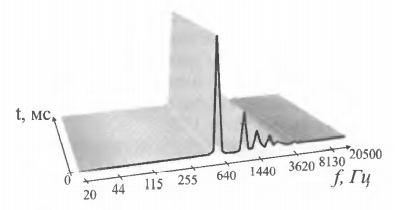


Рисунок 1.12 – Часові спектрограмми зліва направо: флейта, гобой[7]

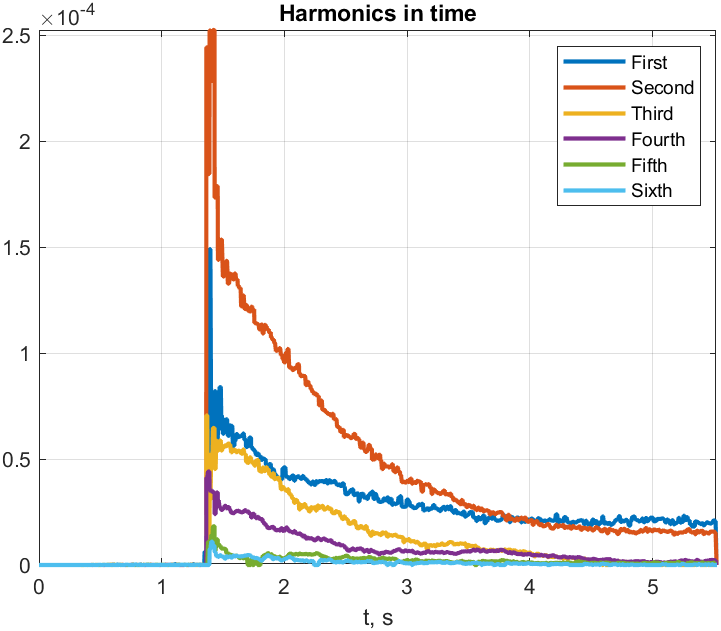


Рисунок 1.13 – Відношення обертонів до основнго тону в часі при збуджені струни гітари

# Теорія класифікації

## Основи класифікації

Основна задача класифікації є задача розділення об’єктів за певними характеристиками з множини об’єктів на певні класи. Класифікатором називається алгоритм, що навчаючись на скінченній группі об’єктів з заздалегіть відомими классами, може класифікувати довільний об’єкт з вхідної множини. Таку группу об’єктів для навчання називають вибіркою. Класифікація – процес присвоювання заданному об’єкту класу з множини класів.

В математичній статистиці задачі класифікації називаються також задачами дискретного аналізу. В машинному навчанні завдання класифікації вирішується, як правило, за допомогою методів штучної нейронної мережі при постановці експеримента у вигляді навчання з учителем.

Нехай - множина описів об'єктів, - множина номерів класів. Тоді існує певна невідома залежність , значення якої відомі тільки для скінченної вибірки для навчання . Задачею класифікації є побудова алгоритма що здатен класифікувати довільный об’єкт .

Характеристикою є відображення , де – множина допустимих значень характеристик. Об’єкт може бути представленим через вектор характеристик так, що . Характеристики можна ототожнювати із самими об'єктами. Нехай кожен об’єкт множини має ознак , а множина класів містить класів .

Так, характеристикою фотографії може бути матриця пікселей, що містять значення яскравості для кожного кольору. Слово може бути вредставлено як бінарний вектор характеристик, в якому одиниця стоїть на місці відповідному до слова. Такий метод незручний в використанні, бо потребує вектори довжиною в повну кількість слів словника для кожного слова. Альтернативним та більш бажаним є представлення слова в просторі , де значно менше кількості слів в словнику. Так слова розділяються на під-змісти, синоніми знаходяться ближче один до одного в мірному просторі, антоніми знаходяться далі.

Основна функціональність штучних нейронних мемрежей загалом поділяється на класифікацію та регресію. Нейронні мережі не використовують апріорних знань про об’єкти для класифікації, натомість мережа методом навчання з тренеровочної вибірки дізнається про закономірності між об’єктами. Такі класифікатори відрізняються від інших рекордною точністю та універсальністью однієї мережі до множини проблем класифікації. Так, система що класифікує фотографії автомобілів може бути змінена на систему, що класифікує фотографії тварин тільки зміною тренеровочної вибірки та зміною назв класів. Недоліками такого класифікатора може бути потреба в великій бази данних для навчання та тестування, великі часові затрати на навчання.

Є багато методів класифікацій, роздивимось декілька найбільш популярних.

## Байесівска класифікація. Гаусовський наївний баєсів класифікатор

Почнемо з ймовірнісного формулювання задачі класифікації. Припускається що множина пар «об'єкт, клас» є ймовірнісним простором з невідомою ймовірнісною мірою . Є скінченна навчальна вибірка спостережень , згенерована згідно з ймовірнісною мірою . Необхідно побудувати алгоритм здатний класифікувати довільный об’єкт .

Байесівський класифікатор – йомвірнісний класифікатор, що використовує теорему Байеса для визначення ймовірності приналежності об’єкта до одного з класів. Наївний байесівський класифікатор (надалі БК) працює при припущенні незалежності ознак об’єкта . Наївний БК іноді працює краще ніж нейронні мережі, що спостерігається в випадках з обеженою вибіркою.

Тобто, якщо аналізуючи ознаки об'єкта можна однозначно визначити, до якого класу він належить, байєсів класифікатор повідомить ймовірність приналежності до цього класу. У проміжних же випадках, коли об’єкт може з різною ймовірністю належати до різних класів, результатом роботи класифікатора буде вектор, компоненти якого є ймовірностями приналежності до того чи іншого класу.

Можна бачити, що ідеальний байєсів класифікатор в якомусь сенсі є оптимальним. Його результат не може бути поліпшений, тому що у всіх випадках, коли можлива однозначна відповідь, він його дасть — а в тих випадках, коли відповідь неоднозначна, результат кількісно характеризує міру цієї неоднозначності.

Разом з тим, в оптимальності криється і основний недолік ідеального байєсового класифікатора: для його побудови потрібна вибірка, що містить всі можливі комбінації ознак — а розмір такої вибірки експоненціально зростає із зростанням числа ознак. Для подолання описаної вище проблеми на практиці використовують наївний БК — класифікатор, побудований на основі припущення про незалежність змінних, обмежившись лише впливом кожної змінної окремо на приналежність образу до одного з класів. [11]

Байєсівський класифікатор заснований на принципі максимуму апостеріорної ймовірності. Для об'єкта класифікації обчислюються функції правдоподібності кожного з класів, по ним обчислюються апостеріорні ймовірності класів. Об'єкт відноситься до того класу, для якого апостериорная ймовірність максимальна.

Приймаючи за множину класів, де один клас, X за множину об’єктів, за об’єкт, а за алгоритм класифікації, рішення про належність об’єкту до класу може бути визначено за формулою (1.1)

Іншими словами алгоритм зводиться до максимуму апостеріорної ймовірності. Клас, ймовірність пренадлежності до якого буде максимальним зі всієї множини класів для конкретного об’єкту буде дорівнювати .

Згідно з теоремою Байеса вираз апостеріорної ймовірності вище може бути переписано в більш явному вигляді (1.2).

|  |  |
| --- | --- |
|  | () |

де – ймовірність що у об’єкта класу будуть ознаки ;

– апріорна ймовірність зустрічі классу серед множини об’єктів;

– апріорна ймовірність з якою зустрічається такий набір ознак.

При пошуку максимуму з формули константами з виразу апостеріорної ймовірності (1.2) можна знехтувати. Так, не залежить від класу, отже може бути відкинутим. Якщо розглядати на практиці, то частіше всього , так як фотографії пікселі яких повністю співпадають не зустрічаються, якщо тільки одна фотографія не зроблена з першої. (?)

Застосовуючи наївну БК припустимо, що ознаки , не залежать одне від одного. Тоді чисельник формули (1.2) можна спростити.

де – вектор ознак, що є взаємозамінним поняттям до об’єкту.

А що б уникнути надмірно малих чисел при множенні великого числа ймовірностей – праву частину можна прологарифмувати, так як максимум від цього не зміниться.

Неперервні залежності, такі як , як правило, оцінюються через нормальний розподіл. Тобто на практиці БК використовує підходи Гаусовського класифікатора[12] (надалі ГК). Принцип ГК полягає в тому, що залежності нам відомі заздалегідь і ми знаємо що їх форма розподілу – гаусівська (1.5). Такий підход є популярним, так як багато природніх процесів, що можуть класифікуватись мають нормальний розподіл параметрів.

де – математичне очікування класу для ознаки ;

– середнеквадратичне відхилення тої ж ознаки.

В якості математичного очікування і дисперсії обчислюються середнє арифметичне (1.6) і середнє квадратичне відхилення (1.7) відповідно.

де M – кількість об’єктів з навчальної вибірки класу , відповідно – i-та ознака p-го об’єкту.

Відповідно для класифікації Наївним БК з використанням принципів ГК потрібно буде зберігти математичних очікувань та середньоквадратічніх відхілень. Тобто на кожен клас з множени потрібно зберігти параметри нормального розподілу для кожної ознаки .

Такий метод широко застосовується для класифікації об’єктів, ознаки яких слабо зв’язані один з одним. Тобто, для тих, ознаки яких можна прийняти незалежними без особливих втрат точнсті.

Розглянемо приклад класифікації. Для навчання класифікатора можна взяти базу данних ознак вин, таких як процент алкоголю, вміст яблучної кислоти, вміст і лужність золи та іншні[13]. Згідно статті[14] видно, що в порівнянні з деревом ухвалення рішення та методом k-найближчих сусідів наївний БК показав, що він значно класифікує з вищою точністью та в значно меншій мірі залежить від параметрів оптимізації, хоча займає більше часу на обрахунки.

## Методи регресивного аналізу. SVM

Задачу классификации можно решить используя регрессионный анализ. Регрессионный анализ – это набор статистических методов исследования влияния независимых переменных или признаков , на независимую переменную [8].

Регрессионный анализ используют, когда для набра значений определено условное математическое ожидание. Уравнение регрессии в общем виде выглядит следующим образом [8]

Применительно к классификации аккустических сигналов это значит, что регрессионный анализ применим только в случае, когда между отчетами спектра и классом из множества классов есть определенная функциональная зависимость и при конкретно заданых значениях спектра есть соответствующий им класс. Функция называется регрессией велечины Y по величинам .

При использовании линейной регрессии и метода метода наименьших квадратов для поиска коэфициентов, записывают следующее уравнение

В нем матрица опорных коэффициентов имеет размерность NxK, где N – количество признаков на объект, а K – колличество классов для класификации.

Такая функция даст вектор-столбец размерностью 1xK, в котором числа в каждой ячейке будут обозначать вероятность принадлежности объекта к соответствующему классу.

https://habrastorage.org/getpro/habr/post_images/c66/15f/a9e/c6615fa9ee7da03c3d6e0b506ff03c48.png

Рассматрим метод опорных векторов (англ. support vector machine, SVM), как один из методов классификации и регрессии.

Идея классификации с помощью метода опорных векторов заключается в поиске и построении функции гиперплоскости (гиперповерхности?) которая разделит объекти выборки оптимальным способом. Оптимальной называется та поверхность, у которой наибольшее расстояние до объектов разделяющих классов. Пример функции разделяющей классы в двумерном пространстве можно посмотреть на Рисунку 2.1.

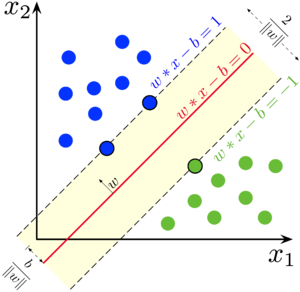


Рисунок 2.3 – Оптимальная разделяющая гиперплоскость в

Очевидно, что как и большенство природных физических явлений, классификация аккустических сигналов не сможет быть разделена линейной функцией. Линейно разделяемые выборки встречаются только математически. Для данного применения воспользуемся SVM с магким отступом (англ. soft-margin SVM): [9]

Для нелинейного разделения над признаками сперва проводят нелинейные операции, таким образом поверхность которая будет разделять кластеры принимает вид

<https://www.machinelearningmastery.ru/prior-over-functions-gaussian-process-1c58e8c40272/>

[https://habr.com/ru/post/101338/#:~:text=%D0%9A%D0%BB%D0%B0%D1%81%D1%82%D0%B5%D1%80%D0%B8%D0%B7%D0%B0%D1%86%D0%B8%D1%8F%20(%D0%B8%D0%BB%D0%B8%20%D0%BA%D0%BB%D0%B0%D1%81%D1%82%D0%B5%D1%80%D0%BD%D1%8B%D0%B9%20%D0%B0%D0%BD%D0%B0%D0%BB%D0%B8%D0%B7)%20%E2%80%94,%D0%B1%D1%8B%D1%82%D1%8C%20%D0%BA%D0%B0%D0%BA%20%D0%BC%D0%BE%D0%B6%D0%BD%D0%BE%20%D0%B1%D0%BE%D0%BB%D0%B5%D0%B5%20%D0%BE%D1%82%D0%BB%D0%B8%D1%87%D0%BD%D1%8B](https://habr.com/ru/post/101338/#:~:text=%D0%9A%D0%BB%D0%B0%D1%81%D1%82%D0%B5%D1%80%D0%B8%D0%B7%D0%B0%D1%86%D0%B8%D1%8F%20(%D0%B8%D0%BB%D0%B8%20%D0%BA%D0%BB%D0%B0%D1%81%D1%82%D0%B5%D1%80%D0%BD%D1%8B%D0%B9%20%D0%B0%D0%BD%D0%B0%D0%BB%D0%B8%D0%B7)%20%E2%80%94,%D0%B1%D1%8B%D1%82%D1%8C%20%D0%).

# Дослідження методів классифікації

## Емпіричні методи класифікації

В данной работе в качестве объектов для классификации представлены акустические сигналы, а признаками выступают спектральные состовляющие сигнала. Судя по частотам на которых найдены локальные максимумы спектра можно судить о звучащей ноте.

В слышимом диапазоне для человека существуют 84 ноты. Соответственно, с учетом возможной музыкальной паузы, аккустический сигнал может должен отоситься как минимум к одному из 85-ти классов.

/\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*/

/\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\* В РАЗДЕЛ 3 \*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*/

### Выбор входного сигнала для классификации

### Длинна сигнала

Задача классификации определить каждую ноту композиции. Исходя их этого можно определить необходимую длинную временной вырезки как такую, в которую ляжет самая короткая нота композиции.

Ноты могут быть сколь-угодно короткими, так что ограничимся теми, которые чаще всего используются, а именно 1/32. Ноты короче 1/32 встречаются очень редко. Это объясняется тем, что что бы задать временную длительность ноты автор сперва записывает чему равен bmp (bit per minute), если не указан – используется значение по умолчанию120, а потом нотами задает количество тактов на частоту. Так самая длинная нота – целая, состоит из 4 тактов. Соответственно, половинная нота длится 2 такта, четвертная 1 такт и тд. На Рисунок 2.1 можна побачити як дві половинні ноти складают цілу, дві четвертинні складають половинну і так далі.

Для наглядности опишу четкий принцип определения длительности ноты. Значение bpm задает сколько целых тактов будет сыграно за минуту. Поделив это число на 60 и после подношения к минус первой степени получаем время в секундах, затрачиваемое на один такт.

В таком случае длительность самой длинной – а именно целой ноті, будет длиться секунд. Так как мы условились самой короткой нотой как 1/32, то длительность такой ноты будет . Принимая во внимение случай по умолчанию как наиболее популярный приведу расчеты с данными числами.

Имея время наименьшей ноты и учитывая стандартную частоту дискретизации можем найти как количество отчетов такого вырезанного сигнала, так и частотный шаг в его спектре.

Первое, что можно заметить – это частотный шаг, который чем шаг между некоторыми нотами. С Таблиці 1.2 можно просчитать, что шаг между нот варьируется с 4Гц до 444Гц начиная с Большой и заканчивая Пятой октавой. Прежде всего для ускорения некоторых дальнейших расчетов, таких как Быстрое Преобразование Фурье, доцільно округлить число отчетов до степени двойки, тогда

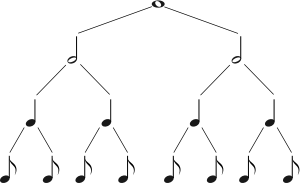


Рисунок 2.1 – Ділення нот по їх тривалості

Глядя на спектр нот, между которіми расстояние менее частотного шага можно наблюдать неблагоприятную картину, в которой разные ноты повышают амплитуду или аккамулируют в одних и тех же спектральных состовляющих. Что бы избежать такого эффекта можно применить интерполяцию спектра.

Интерполяцию спектра можно получить несложным образом – добавив нулей в сигнал, который будет преобразован в фурье. Таким образом мы искусственно увеличиваем колличество отчетов N. Однако у такогометода есть свои недостатки. Главный из них – это пульсации в спектре. Они объясняются тем, что по сути мы применили оконное преобразование фурье к большому сигналу, в котором прямоугольной функцией хемминга, путем перемножения, был выделен наш анализируемый сигнал. Соответственно, если во временной области сигналы перемножились, то в спектральной области спектры этих сигналов свернулись. Выходит что от каждой синусоидальной гармоники будет расходиться синк, что может перекрыть низкоуровневые гармоники пульсациями более высоких гармоник. Данный процесс можно увидеть на Рисунок 2.2.

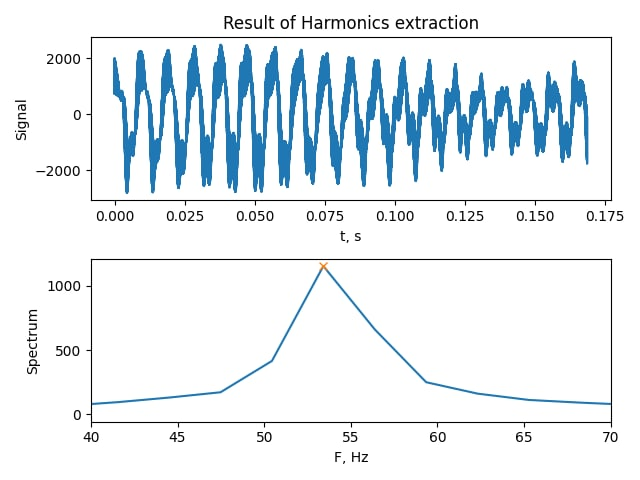
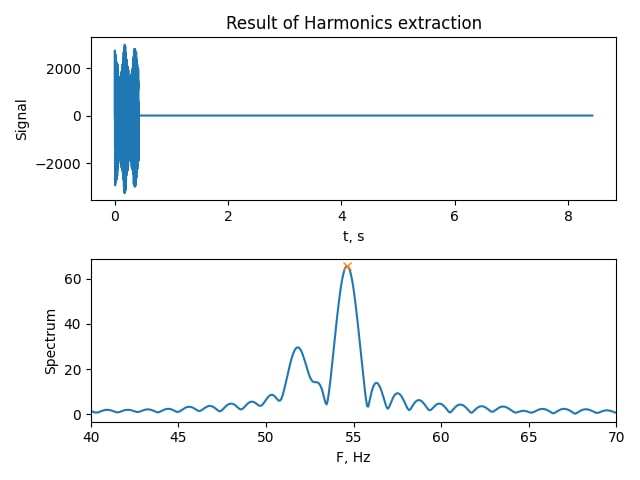
 

Рисунок 2.2 – Сигналы во временной обалсти и их спектры. Первый случай – акустический сигнал в котором заложен звук двух нот взят с такой длинной, что бы частотный шаг превышал расстояние между нотами в спектре. Второй случай – тот же сигнал был дополнен нулями в временной области, что позволило различить спектральные состовляющие нот.

Лучше всего подобрать такое количество нулей для заполнения во временной области, что бы частотный шаг был значительно меньше минимального шага между нот. Таким шагом был выбран шаг в 1Гц

В таком случае общая длинна акустического сигнала с нулевыми дополнениями должна быть равна частоте дискретизации, согласно предыдущим формулам, а значит нулей должно быть добавлено

### Домен сигнала (временной-частотный)

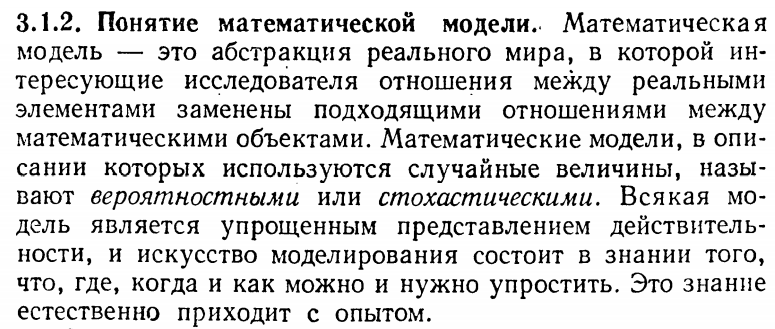
Рассматривая детальнее задачу классификации аккустических сигналов и принимая во внимание физические особенности звучания можно заметить, что именно частота звука определит его ноту. В таком случае уместнее (доцільно) перейти из временного домена в частотный и анализировать участок сигнала по его спектру.

Есть несколько методов перехода во временной домен, к примеру используя Быстрое Преобразование Фурье (fft) или вейвлет-преобразования. Вейвлет преобразования дают временно-частотную картину на которой можно четко увидеть в какой момент какие частоты появились, что решило бы проблему захвата участка сигнала с двумя нотами\*\*\*. Однако, в большинстве случаем оконное преобразование фурье может заменить более сложное для анализа вейвлет преобразование. Исходя из вышесказанного начнем с применения более простого фурье преобразования.

Быстрое преобразование фурье[10] – это алгоритм, который позволяет сократить время расчета дискретного преобразования фурье. Тогда как обычное дискретное преобразование фурье выполняет расчеты за время , быстрое преобразование Фурье справляется за .

/\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\* В РАЗДЕЛ 3 \*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*/

/\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*/



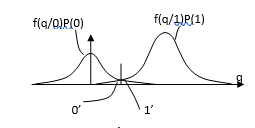
## Гаусівський наївний баєсів класифікатор (Gaussian naïve Bayes)

**КАК РАБОТАЕТ 1**

Применительно к классификации акустического сигнала с определением звучащей ноты выбор класса определится максимумом вероятности с которой сигнал относится к классу.

На практике инструменты могут быть неточно настроены или расстраиваться во время игры. В таком случае частоты нот могут иметь погрешности. Предположив, что отклонениче частот от номинальных значений происходит по нормальному закону, экспериментально можно найти параметры распределения, а именно – дисперсию.

Таким образом получим Гауссовские распределения на каждую ноту. Таким образом можно найти оптимальные пороговые частоты для разделения нот.



Другой подход с использованием Байесовской классификации заключается в том, что для каждой ноты можно построить свое гауссовское распределение, аргументы которого будут равны всем спектральным состовляющим акустического сигнала. Таким образом, для конкретной ноты наибольшая дисперсия будет на частотах спектра которые кратны основному тону, так как нота может звучать с разной громкостью. В то же время участки спектра отдаленные от обертонов должны быть на относительно низком уровне, что будет свидетельствовать о том, что другие ноты не играются.

Так, в программе берутся спектральные участки длинной в 1024 отчета. Для каждой из 84 нот будет создано Гауссовское распределение 1024 переменных и отдельное распределение для отсутствия ноты.

Применив такой алгоритм к всему множеству сигналов получим последовательность нот и пауз композиции.

/\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*/

/\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\* В РАЗДЕЛ 3 \*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*/

Раскрывая понятие сигнала можем расписать какие именно признаки (features) будут приниматься в расчеты.

Где – i-тый отчет спектра сигнала.

/\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\* В РАЗДЕЛ 3 \*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*/

/\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*/

Применимо к классификации аккустических сигналов – для обучения классификатора, а именно для подбора параметров для подсчета вероятности, могут быть использованы базы данных готовых музыкальных произведений с заведомо известными последовательностями нот. Таким образом заранее классификатор будет знать какие ноты вероятнее будут звучать. К примеру ноты суб-контр октав и контр октав практически не встречаются в музыке.

Нахождение априорной вероятности можно осуществить следующим методом. Собираем базу данных из множества популярных композиций. Находим все ноты которые там были сыграны. Находим часть ноты С среди всех остальных. Так и получаем априорную вероятность появления данной ноты.

хороший туториал по байесу - <https://blog.uiam.sk/the-naive-probability-is-an-expert/>

<http://www.machinelearning.ru/wiki/index.php?title=%D0%91%D0%B0%D0%B9%D0%B5%D1%81%D0%BE%D0%B2%D1%81%D0%BA%D0%B8%D0%B9_%D0%BA%D0%BB%D0%B0%D1%81%D1%81%D0%B8%D1%84%D0%B8%D0%BA%D0%B0%D1%82%D0%BE%D1%80>

## SVM класифікатор

фівф

## Застосування машинного навчання з підкріпленням

References

[1]Музыкальная нотация

<https://ru.wikipedia.org/wiki/%D0%9C%D1%83%D0%B7%D1%8B%D0%BA%D0%B0%D0%BB%D1%8C%D0%BD%D0%B0%D1%8F_%D0%BD%D0%BE%D1%82%D0%B0%D1%86%D0%B8%D1%8F>

[2] Интервал (музыка)

<https://ru.wikipedia.org/wiki/%D0%98%D0%BD%D1%82%D0%B5%D1%80%D0%B2%D0%B0%D0%BB_(%D0%BC%D1%83%D0%B7%D1%8B%D0%BA%D0%B0)>

[3] Знаки альтерации

<https://ru.wikipedia.org/wiki/%D0%97%D0%BD%D0%B0%D0%BA%D0%B8_%D0%B0%D0%BB%D1%8C%D1%82%D0%B5%D1%80%D0%B0%D1%86%D0%B8%D0%B8>

[4] Равномерно темперированный строй

<https://ru.wikipedia.org/wiki/%D0%A0%D0%B0%D0%B2%D0%BD%D0%BE%D0%BC%D0%B5%D1%80%D0%BD%D0%BE_%D1%82%D0%B5%D0%BC%D0%BF%D0%B5%D1%80%D0%B8%D1%80%D0%BE%D0%B2%D0%B0%D0%BD%D0%BD%D1%8B%D0%B9_%D1%81%D1%82%D1%80%D0%BE%D0%B9>

[5] Статья про тона инструментов

<http://information-technology.ru/sci-pop-articles/23-physics/268-kak-poluchayut-raznye-tona>

[6] Пособие – «Колебание струны». Математическая физика

<http://gukitkafmi.narod.ru/files/INShitov/string.pdf>

[7] Учебник – Музыкальная акустика, Ирина Алдошина и Рой Приттс

Прочитать начиная с 280 стр!!!

[8] Регрессионный анализ –

<https://ru.wikipedia.org/wiki/%D0%A0%D0%B5%D0%B3%D1%80%D0%B5%D1%81%D1%81%D0%B8%D0%BE%D0%BD%D0%BD%D1%8B%D0%B9_%D0%B0%D0%BD%D0%B0%D0%BB%D0%B8%D0%B7>

[9] Метод опорных векторов –

[https://neerc.ifmo.ru/wiki/index.php?title=%D0%9C%D0%B5%D1%82%D0%BE%D0%B4\_%D0%BE%D0%BF%D0%BE%D1%80%D0%BD%D1%8B%D1%85\_%D0%B2%D0%B5%D0%BA%D1%82%D0%BE%D1%80%D0%BE%D0%B2\_(SVM)#:~:text=%D0%9C%D0%B5%D1%82%D0%BE%D0%B4%20%D0%BE%D0%BF%D0%BE%D1%80%D0%BD%D1%8B%D1%85%20%D0%B2%D0%B5%D0%BA%D1%82%D0%BE%D1%80%D0%BE%D0%B2%20(%D0%B0%D0%BD%D0%B3%D0%BB.,%D1%80%D0%B0%D0%B7%D0%B4%D0%B5%D0%BB%D1%8F%D1%8E%D1%89%D0%B5%D0%B9%20%D0%BE%D0%B1%D1%8A%D0%B5%D0%BA%D1%82%D1%8B%20%D0%B2%D1%8B%D0%B1%D0%BE%D1%80%D0%BA%D0%B8%20%D0%BE%D0%BF%D1%82%D0%B8%D0%BC%D0%B0%D0%BB%D1%8C%D0%BD%D1%8B%D0%BC%20%D1%81%D0%BF%D0%BE%D1%81%D0%BE%D0%B1%D0%BE%D0%BC](https://neerc.ifmo.ru/wiki/index.php?title=%D0%9C%D0%B5%D1%82%D0%BE%D0%B4_%D0%BE%D0%BF%D0%BE%D1%80%D0%BD%D1%8B%D1%85_%D0%B2%D0%B5%D0%BA%D1%82%D0%BE%D1%80%D0%BE%D0%B2_(SVM)#:~:text=%D0%9C%D0%B5%D1%82%D0%BE%D0%B4%20%D0%BE%D0%BF%D0%BE%D1%80%D0%BD%D1%8B%D1%85%20%D0%B2%D0%B5%D0%BA%D1%82%D0%BE%D1%80%D0%BE%D0%B2%20(%D0%B0%D0%BD%D0%B3%D0%BB.,%D1%80%D0%B0%D0%B7%D0%B4%D0%B5%D0%BB%D1%8F%D1%8E%D1%89%D0%B5%D0%B9%20%D0%BE%D0%B1%D1%8A%D0%).

[10] Быстрое Преобразование Фурье –

<https://ru.wikipedia.org/wiki/%D0%91%D1%8B%D1%81%D1%82%D1%80%D0%BE%D0%B5_%D0%BF%D1%80%D0%B5%D0%BE%D0%B1%D1%80%D0%B0%D0%B7%D0%BE%D0%B2%D0%B0%D0%BD%D0%B8%D0%B5_%D0%A4%D1%83%D1%80%D1%8C%D0%B5>

[11] Наївний баєсів класифікатор –

<https://uk.wikipedia.org/wiki/%D0%9D%D0%B0%D1%97%D0%B2%D0%BD%D0%B8%D0%B9_%D0%B1%D0%B0%D1%94%D1%81%D1%96%D0%B2_%D0%BA%D0%BB%D0%B0%D1%81%D0%B8%D1%84%D1%96%D0%BA%D0%B0%D1%82%D0%BE%D1%80>

[12] Гаусовський класифікатор –

<https://learnmachinelearning.wikia.org/ru/wiki/%D0%93%D0%B0%D1%83%D1%81%D1%81%D0%BE%D0%B2_%D0%BA%D0%BB%D0%B0%D1%81%D1%81%D0%B8%D1%84%D0%B8%D0%BA%D0%B0%D1%82%D0%BE%D1%80>

[13] База данних для касифікування вина

<https://archive.ics.uci.edu/ml/datasets/Wine>

[14] Стаття про застосування Байесівського класифікатора

<https://blog.uiam.sk/the-naive-probability-is-an-expert/>