

ПРАВИТЕЛЬСТВО РОССИЙСКОЙ ФЕДЕРАЦИИ  
ФГАОУ ВО НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ УНИВЕРСИТЕТ  
«ВЫСШАЯ ШКОЛА ЭКОНОМИКИ»

Факультет компьютерных наук  
Образовательная программа «Прикладная математика и информатика»

УДК 004.8

**Отчет об исследовательском проекте на тему:**

**Кластеризация аудио**

(промежуточный, этап 1)

**Выполнил студент:**

группы #БПМИ213, 3 курса

Бонич Дмитрий Сергеевич

**Принял руководитель проекта:**

Сендерович Александра Леонидовна

Научный сотрудник

Факультет компьютерных наук НИУ ВШЭ

Москва 2024

# Содержание

<b>Аннотация</b>	<b>3</b>
<b>1 Введение</b>	<b>4</b>
1.1 Постановка задачи . . . . .	4
1.2 Метрики качества . . . . .	4
1.3 Классические методы решения . . . . .	5
1.3.1 K-means . . . . .	5
1.3.2 DBSCAN . . . . .	5
1.3.3 MeanShift . . . . .	5
1.3.4 AgglomerativeClustering . . . . .	5
<b>2 Обзор литературы</b>	<b>5</b>
<b>3 План дальнейшей работы</b>	<b>5</b>
<b>4 Примеры</b>	<b>5</b>
4.1 Ссылки на статьи . . . . .	5
4.2 Рисунки . . . . .	6
4.3 Таблицы . . . . .	7
4.4 Формулы . . . . .	7
<b>Список литературы</b>	<b>8</b>

## **Аннотация**

Существующие методы глубинного обучения для кластеризации изображений работают довольно хорошо. Однако вопрос качественной кластеризации аудио остается открытым. В этой работе мы планируем адаптировать лучшие методы кластеризации изображений для задачи кластеризации аудио.

## **Ключевые слова**

Глубинное обучение, обучение без учителя, кластеризация

# 1 Введение

## 1.1 Постановка задачи

В задаче классификации мы имеем обучающую выборку, где для каждого объекта известен его класс и от нас требуется моделировать распределение классов на пространстве объектов. Задача кластеризации более сложная – необходимо разбить объекты на осмысленные группы не зная ни самих групп, ни их распределения.

## 1.2 Метрики качества

Чтобы замерить качество кластеризации как правило метод решает задачу на размеченном для кластеризации датасете, разумеется не используя метки классов. Полученные после кластеризации номера кластеров будем называть *псевдометками*. Псевдометки сравниваются с настоящими метками и качество определяется как некоторый вид корреляции между ними.

Одной из самых популярных метрик качества является NMI(Normalized Mutual Information). Пусть у нас есть пара случайных величин  $X$  и  $Y$ , тогда:

$$\text{NMI}(X, Y) = \frac{\text{KL}(p_{(X,Y)} | p_X \otimes p_Y)}{\sqrt{H(X)H(Y)}}$$

При подсчете NMI мы считаем распределения меток и псевдометок за  $X$  и  $Y$  и вычисляем *оценку максимального правдоподобия*<sup>1</sup>. NMI принимает значения от 0 до 1.

Еще одной используемой метрикой является ARI(Adjusted Rand Index). Это скорректированная версия метрики RI<sup>2</sup> (Rand Index). Определяется ARI следующим образом:

$$\text{ARI}(X, Y) = \frac{\text{RI}(X, Y) - \text{E}[\text{RI}(X, Y)]}{1 - \text{E}[\text{RI}(X, Y)]}$$

ARI принимает значения от 0 до 1.

Также в качестве метрики качества можно использовать долю правильных ответов как и в задаче классификации. Однако предварительно необходимо решить *задачу о назначениях*<sup>3</sup>. между псевдометками и метками. Мы переименуем псевдометки так, чтобы достичь максимального качества. Затем на переименованных псевдометках мы посчитаем долю правильных ответов, которую и используем в качестве метрики качества.

---

<sup>1</sup>далее ОМП

<sup>2</sup>[https://en.wikipedia.org/wiki/Rand\\_index](https://en.wikipedia.org/wiki/Rand_index)

<sup>3</sup>[https://en.wikipedia.org/wiki/Assignment\\_problem](https://en.wikipedia.org/wiki/Assignment_problem)

## 1.3 Классические методы решения

С точки зрения классических методов кластеризуемые объекты это точки в многомерном пространстве. Такие методы обычно имеют итерационную природу и решают задачу выделения кластеров точек находя их скопления, области с повышенной плотностью, некоторые структуры. Посмотрим на несколько примеров.

### 1.3.1 K-means

Одним из самых простых и известных методов является K-means. Он работает на базе EM-алгоритма<sup>4</sup>. K-means итерационно пытается найти два неизвестных набора переменных — номера кластеров для объектов и центры этих кластеров. Количество кластеров фиксировано и задается до начала работы алгоритма в качестве гиперпараметра  $k$ .

В начале своей работы классический K-means инициализирует все центры кластеров случайно. Затем чередуются E и M шаги до сходимости. На E-шаге мы назначаем каждому объекту номер кластера к центру которого объект ближе всего в качестве псевдометки. На M-шаге мы вычисляем ОМП для каждого центра кластера, то есть берем в качестве центра кластера усредненную точку из всех объектов с псевдометкой этого кластера.

### 1.3.2 DBSCAN

### 1.3.3 MeanShift

### 1.3.4 AgglomerativeClustering

## 2 Обзор литературы

## 3 План дальнейшей работы

## 4 Примеры

### 4.1 Ссылки на статьи

Ссылки на статьи оформляются с помощью пакета `biblatex`, например [1]. В описании статье в `bib` файле нужно обязательно указывать место публикации работы (журнал или конференцию) и год. Обратите внимание, что для описания статей из разных источников в списке литературы используются разные команды в `bib` файле: статья из журнала [3],

---

<sup>4</sup>[https://en.wikipedia.org/wiki/Expectation%E2%80%93maximization\\_algorithm](https://en.wikipedia.org/wiki/Expectation%E2%80%93maximization_algorithm)

статья с конференции [1], книга [6], глава книги [5]. Если статья еще не опубликована нигде, а только выложена на arXiv, то на нее тоже можно сослаться [2], но предпочтительно ссылаться на опубликованную версию, если она уже существует. Если вы хотите сослаться на сайт, то можно либо так же внести его в список литературы [4] (рекомендуется, если таких ссылок у вас много из-за особенностей темы вашего проекта), либо использовать ссылку внизу страницы<sup>5</sup>. При работе с онлайн ресурсами не забывайте указывать дату обращения к этому ресурсу, так как в отличие от опубликованных статей, эти ресурсы могут измениться в любой момент.

## 4.2 Рисунки

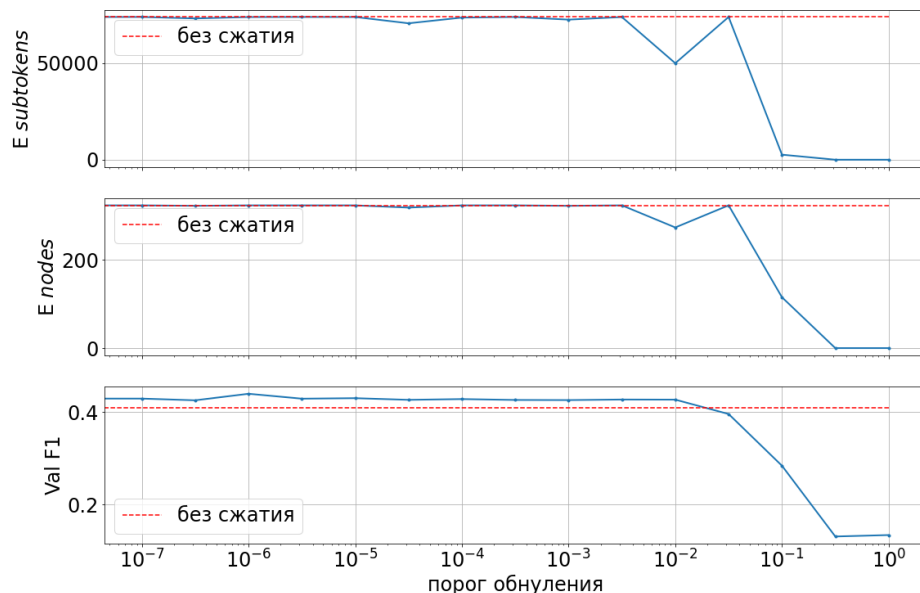


Рис. 4.1: Пример графика. Тут должна быть подпись, поясняющая что происходит на рисунке (краткая, но достаточная для понимания основной идеи графика).

Все рисунки в тексте должны иметь подписи и вы на них должны ссылаться в тексте. Например, на Рисунке 4.1 изображен пример графика. Не забывайте подписывать все оси на графиках, добавлять легенду и пояснять все обозначения, а также используйте адекватного размера шрифты и толщину линий на графиках (все должно быть видно и понятно без многократного увеличения). На рисунке из примера явно не хватает обозначения синей линии в легенде.

<sup>5</sup>Книги доступны по ссылке: <http://www-cs-faculty.stanford.edu/~uno/abcde.html>, дата обр. 16.05.2013

## 4.3 Таблицы

Все таблицы в тексте тоже должны иметь подписи и вы на них должны ссылаться в тексте. Например, в Таблице 4.1 показаны результаты примерного эксперимента.

Таблица 4.1: Пример таблички. Тут должна быть подпись, поясняющая что происходит в таблице (краткая, но по делу).

	Val			Test			nodes	subtokens
	Prec	Rec	F1	Prec	Rec	F1		
запуск 1	0.4894	0.3775	0.4263	0.4824	0.3683	0.4177	10029	179
запуск 2	0.4887	0.3739	0.4237	0.4891	0.3724	0.4228	10039	177
запуск 3	0.4820	0.3751	0.4219	0.4838	0.3677	0.4178	10037	180
<b>среднее</b>	<b>0.4867</b>	<b>0.3755</b>	<b>0.4239</b>	<b>0.4851</b>	<b>0.3695</b>	<b>0.4195</b>		
<b>дисперсия</b>	0.0041	0.0019	0.0022	0.0036	0.0025	0.0029		

## 4.4 Формулы

Формулы стоит центрировать, а также нумеровать, если вы ссылаете на них в тексте. Также не забывайте пояснять все обозначения в формулах. Например, запишем следующую задачу оптимизации:

$$\theta^* = \min_{\theta} F(\theta), \quad (1)$$

где  $F$  – квадратичная функция от параметра  $\theta$ . При необходимости, далее в тексте можно сослаться на формулу (1). При этом, в зависимости от конкретных формул, можно использовать разные слова: формула, уравнение, задача оптимизации и т.п.

## Список литературы

- [1] Nadezhda Chirkova, Ekaterina Lobacheva и Dmitry Vetrov. “Bayesian Compression for Natural Language Processing”. в: *Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP)*. 2018.
- [2] Nadezhda Chirkova, Ekaterina Lobacheva и Dmitry Vetrov. “Bayesian Compression for Natural Language Processing”. в: *arXiv preprint, arXiv:1810.10927, version 2* (2018).
- [3] George D. Greenwade. “The Comprehensive Tex Archive Network (CTAN)”. в: *TUGBoat* 14.3 (1993), с. 342—351.
- [4] Donald Knuth. *Knuth: Computers and Typesetting*. URL: <http://www-cs-faculty.stanford.edu/~uno/abcde.html> (дата обр. 16.05.2013).
- [5] Donald E. Knuth. “Fundamental Algorithms”. в: Addison-Wesley, 1973. тл. 1.2.
- [6] Donald E. Knuth. *The Art of Computer Programming*. Four volumes. Seven volumes planned. Addison-Wesley, 1968.