

Школа: ИШИТР Инженерная школа информационных технологий и робототехники
 Направление подготовки 09.03.01 «Информатика и вычислительная техника»
 Отделение школы (НОЦ): Отделение информационных технологий

БАКАЛАВРСКАЯ РАБОТА

Тема работы
Кластерный анализ данных большой размерности

УДК 519.237:004.422.634:004.424.52

Студент

Группа	ФИО	Подпись	Дата
3-8В41	Гусева Виктория Сергеевна		

Руководитель ВКР

Должность	ФИО	Ученая степень, звание	Подпись	Дата
Доцент ОИТ	Буркатовская Ю.Б.	к.ф.-м.н		

КОНСУЛЬТАНТЫ ПО РАЗДЕЛАМ:

По разделу «Финансовый менеджмент, ресурсоэффективность и ресурсосбережение»

Должность	ФИО	Ученая степень, звание	Подпись	Дата
Доцент ОСГН	Креницына З.В.	к.т.н., Доцент		

По разделу «Социальная ответственность»

Должность	ФИО	Ученая степень, звание	Подпись	Дата
Ассистент ООД	Мезенцева И.Л.			

ДОПУСТИТЬ К ЗАЩИТЕ:

Руководитель ООП	ФИО	Ученая степень, звание	Подпись	Дата
09.03.01 Информатика и ВТ	Погребной А. В.	к.т.н., Доцент		

**ЗАПЛАНИРОВАННЫЕ РЕЗУЛЬТАТЫ ПО ОСНОВНОЙ ОБРАЗОВАТЕЛЬНОЙ
ПРОГРАММЕ ПОДГОТОВКИ БАКАЛАВРОВ 09.03.01 «ИНФОРМАТИКА И
ВЫЧИСЛИТЕЛЬНАЯ ТЕХНИКА», ИК ТПУ, ПРОФИЛЬ «СИСТЕМЫ
АВТОМАТИЗИРОВАННОГО ПРОЕКТИРОВАНИЯ»**

Код результата	Результат обучения (выпускник должен быть готов)
<i>Профессиональные компетенции</i>	
P1	Применять глубокие естественнонаучные и математические знания для решения научных и инженерных задач в области информатики и вычислительной техники.
P2	Применять глубокие специальные знания в области информатики и вычислительной техники для решения междисциплинарных инженерных задач.
P3	Ставить и решать инновационные задачи инженерного анализа, связанные с созданием аппаратных и программных средств информационных и автоматизированных систем, с использованием аналитических методов и сложных моделей.
P4	Выполнять инновационные инженерные проекты по разработке аппаратных и программных средств автоматизированных систем различного назначения с использованием современных методов проектирования, систем автоматизированного проектирования, передового опыта разработки конкурентно способных изделий.
P5	Планировать и проводить теоретические и экспериментальные исследования в области проектирования аппаратных и программных средств автоматизированных систем с использованием новейших достижений науки и техники, передового отечественного и зарубежного опыта. Критически оценивать полученные данные и делать выводы.
P6	Осуществлять авторское сопровождение процессов проектирования, внедрения и эксплуатации аппаратных и программных средств автоматизированных систем различного назначения.
<i>Универсальные компетенции</i>	
P7	Использовать глубокие знания по проектному менеджменту для ведения инновационной инженерной деятельности с учетом юридических аспектов защиты интеллектуальной собственности.
P8	Осуществлять коммуникации в профессиональной среде и в обществе в целом, активно владеть иностранным языком, разрабатывать документацию, презентовать и защищать результаты инновационной инженерной деятельности, в том числе на иностранном языке.

Код результата	Результат обучения (выпускник должен быть готов)
<i>Профессиональные компетенции</i>	
P9	Эффективно работать индивидуально и в качестве члена и руководителя группы, в том числе междисциплинарной и международной, при решении инновационных инженерных задач.
P10	Демонстрировать личную ответственность и ответственность за работу возглавляемого коллектива, приверженность и готовность следовать профессиональной этике и нормам ведения инновационной инженерной деятельности. Демонстрировать глубокие знания правовых, социальных, экологических и культурных аспектов инновационной инженерной деятельности.
P11	Демонстрировать способность к самостоятельному обучению, непрерывному самосовершенствованию в инженерной деятельности, способность к педагогической деятельности.

Министерство науки и высшего образования Российской Федерации
 федеральное государственное автономное
 образовательное учреждение высшего образования
 «Национальный исследовательский Томский политехнический университет» (ТПУ)

Школа ИШИТР Инженерная школа информационных технологий и робототехники
 Направление подготовки 09.03.01 «Информатика и вычислительная техника»
 Отделение школы (НОЦ): Отделение информационных технологий

УТВЕРЖДАЮ:
 Руководитель ООП
 _____ Погребной А. В.
 (Подпись) (Дата) (Ф.И.О.)

ЗАДАНИЕ
на выполнение выпускной квалификационной работы

В форме:

Бакалаврской работы

Студенту:

Группа	ФИО
3-8В41	Гусевой Виктории Сергеевне

Тема работы:

Кластерный анализ данных большой размерности.	
Утверждена приказом директора (дата, номер)	11.03.2019 г., № 1799/с

Срок сдачи студентом выполненной работы:	03.06.2019
--	------------

ТЕХНИЧЕСКОЕ ЗАДАНИЕ:

Исходные данные к работе	Реализация кластерного анализа аудиофайлов по спектральным характеристикам, собранных при помощи системы аудиомониторинга ульев BeePi.
---------------------------------	--

<p>Перечень подлежащих исследованию, проектированию и разработке вопросов <i>(аналитический обзор по литературным источникам с целью выяснения достижений мировой науки техники в рассматриваемой области; постановка задачи исследования, проектирования, конструирования; содержание процедуры исследования, проектирования, конструирования; обсуждение результатов выполненной работы; наименование дополнительных разделов, подлежащих разработке; заключение по работе).</i></p>	<p>Изучение существующих алгоритмов кластеризации данных. Исследование и реализация кластерного анализа аудиофайлов по спектральным характеристикам. Произведение расчетов оптимальной кластеризации для заданных данных.</p>
<p>Перечень графического материала <i>(с точным указанием обязательных чертежей)</i></p>	<p>Результаты работы программ, презентация в Microsoft PowerPoint 2010.</p>

<p>Консультанты по разделам выпускной квалификационной работы <i>(с указанием разделов)</i></p>	
<p>Раздел</p>	<p>Консультант</p>
<p>Финансовый менеджмент, ресурсоэффективность и ресурсосбережение</p>	<p>Креницына Зоя Васильевна</p>
<p>Социальная ответственность</p>	<p>Мезенцева Ирина Леонидовна</p>

<p>Дата выдачи задания на выполнение выпускной квалификационной работы по линейному графику</p>	
--	--

Задание выдал руководитель / консультант (при наличии):

<p>Должность</p>	<p>ФИО</p>	<p>Ученая степень, звание</p>	<p>Подпись</p>	<p>Дата</p>
<p>Доцент ОИТ</p>	<p>Буркатовская Ю.Б.</p>	<p>к.ф.-м.н</p>		

Задание принял к исполнению студент:

<p>Группа</p>	<p>ФИО</p>	<p>Подпись</p>	<p>Дата</p>
<p>3-8В41</p>	<p>Гусева Виктория Сергеевна</p>		

**ЗАДАНИЕ ДЛЯ РАЗДЕЛА
«ФИНАНСОВЫЙ МЕНЕДЖМЕНТ, РЕСУРСОЭФФЕКТИВНОСТЬ И
РЕСУРСОСБЕРЕЖЕНИЕ»**

Студенту:

Группа	ФИО
3-8В41	Гусева Виктория Сергеевна

Школа	ИШИТР	Отделение школы (НОЦ)	Информационных технологий
Уровень образования	Бакалавриат	Направление/специальность	09.03.01 «Информатика и ВТ»

Исходные данные к разделу «Финансовый менеджмент, ресурсоэффективность и ресурсосбережение»:

1. Стоимость ресурсов научного исследования (НИ): материально-технических, энергетических, финансовых, информационных и человеческих	Научно-техническое исследование проводится в компании МАУ «ЦСИ». В работе над проектом задействованы 2 человека: научный руководитель и студент-дипломник.
2. Нормы и нормативы расходования ресурсов	В соответствии с ГОСТ 14.322-83 «Нормирование расхода материалов» и ГОСТ Р 51541-99 «Энергосбережение. Энергетическая эффективность»; Минимальный размер оплаты труда в 2018 году составляет 11280 рублей.
3. Используемая система налогообложения, ставки налогов, отчислений, дисконтирования и кредитования	Отчисления по страховым взносам – 27,1% от ФОТ

Перечень вопросов, подлежащих исследованию, проектированию и разработке:

1. Оценка коммерческого потенциала, перспективности и альтернатив проведения НИ с позиции ресурсоэффективности и ресурсосбережения	- Потенциальные потребители результатов исследования;
2. Планирование и формирование бюджета научных исследований	- Структура работ в рамках научного исследования; - Определение трудоемкости выполнения работ; - Разработка графика проведения научного исследования; - Бюджет научно-технического исследования (НТИ)
3. Определение ресурсной (ресурсосберегающей), финансовой, бюджетной, социальной и экономической эффективности исследования	- Анализ и оценка научно-технического уровня проекта

Перечень графического материала (с точным указанием обязательных чертежей):

1. Оценка конкурентоспособности технических решений
2. Матрица SWOT
3. Альтернативы проведения НИ
4. График проведения и бюджет НИ
5. Оценка ресурсной, финансовой и экономической эффективности НИ

Дата выдачи задания для раздела по линейному графику	
---	--

Задание выдал консультант:

Должность	ФИО	Ученая степень, звание	Подпись	Дата
Доцент ОСГН ШБИП	Креницына Зоя Васильевна	К.т.н., доцент		

Задание принял к исполнению студент:

Группа	ФИО	Подпись	Дата
3-8В41	Гусева Виктория Сергеевна		

**ЗАДАНИЕ ДЛЯ РАЗДЕЛА
«СОЦИАЛЬНАЯ ОТВЕТСТВЕННОСТЬ»**

Студенту:

Группа	ФИО
3-8В41	Гусева Виктория Сергеевна

Школа	ИШИТР	Отделение (НОЦ)	Информационных технологий
Уровень образования	Бакалавриат	Направление/специальность	09.03.01 «Информатика и ВТ»

Тема ВКР: Кластерный анализ данных большой размерности

Исходные данные к разделу «Социальная ответственность»:	
1. Характеристика объекта исследования (вещество, материал, прибор, алгоритм, методика, рабочая зона) и области его применения	<ul style="list-style-type: none"> – Рабочее место сотрудника, исп. для получения результатов от поставленных целей организации; – Система, предназначенная для кластерного анализа аудиофайлов по спектральным характеристикам
Перечень вопросов, подлежащих исследованию, проектированию и разработке:	
1. Правовые и организационные вопросы обеспечения безопасности: <ul style="list-style-type: none"> – специальные (характерные при эксплуатации объекта исследования, проектируемой рабочей зоны) правовые нормы трудового законодательства; – организационные мероприятия при компоновке рабочей зоны. 	<ul style="list-style-type: none"> – ГОСТ Р 50923-96 «Дисплеи. Рабочее место оператора. Общие эргономические требования к производственной среде. Методы измерения» – СанПиН 2.2.2/2.4.1340-03 "Гигиенические требования к персональным электронно-вычислительным машинам и организации работы" (с изменениями на 21 июня 2016 года).
2. Производственная безопасность: 2.1. Анализ выявленных вредных и опасных факторов 2.2. Обоснование мероприятий по снижению воздействия	<ul style="list-style-type: none"> - отклонение показателей микроклимата; - превышение уровня шума; - отсутствие или недостаток естественного света; - недостаточная освещенность рабочей зоны; - повышенное значение напряжения в электрической цепи, замыкание которой может произойти через тело человека.
3. Экологическая безопасность:	<ul style="list-style-type: none"> – анализ воздействия объекта на атмосферу (выбросы); – анализ воздействия объекта на гидросферу (сбросы); – анализ воздействия объекта на

	литосферу (отходы).
4. Безопасность в чрезвычайных ситуациях:	Чрезвычайная ситуация техногенного характера – пожар. Установка общих правил поведения и эвакуации во время пожара

Дата выдачи задания для раздела по линейному графику	
---	--

Задание выдал консультант:

Должность	ФИО	Ученая степень, звание	Подпись	Дата
ассистент ООТД	Мезенцева И.Л.			

Задание принял к исполнению студент:

Группа	ФИО	Подпись	Дата
3-8В41	Гусева Виктория Сергеевна		

РЕФЕРАТ

Выпускная квалификационная работа содержит 60 с., 10 рис., 10 источников.

Ключевые слова: кластерный анализ, спектральный анализ аудиофайлов, электронная система мониторинга улья BeePi, пакет Statistica, алгоритм k-средних, EM - масштабируемый алгоритм кластеризации.

Объектом исследования является набор аудиофайлов, записанных с помощью специального оборудования, установленного на ульях, а также различные спектральные характеристики этих файлов.

Цель работы – исследовать возможности кластерного анализа для выделения значимых признаков аудиофайлов, которые могут использоваться для классификации файлов. Рассматриваются 3 класса файлов: жужжание пчел, стрекот сверчков и различные шумы.

В процессе исследования были изучены и проанализированы различные типы спектральных характеристик и их применимость к кластеризации данных аудиофайлов.

В результате исследования был подобран набор спектральных характеристик достаточный для хорошей кластеризации, как на 3 кластера, так и для различных пар классов.

Область применения: автоматический аудиомониторинг пасек.

В будущем планируется провести достаточное количество экспериментов и использовать подобный набор параметров в системе автоматического аудиомониторинга ульев BeePi.

ОПРЕДЕЛЕНИЯ, ОБОЗНАЧЕНИЙ, СОКРАЩЕНИЯ И НОРМАТИВНЫЕ ССЫЛКИ

Кластерный анализ: это разбиение множества объектов на некоторые однородные подмножества, параметры которых изначально неизвестны.

Класс: группа элементов, известная изначально до применения алгоритма кластеризации.

Кластер: группа элементов, полученная с помощью алгоритма кластеризации.

Множество: определенная совокупность объектов. Объекты, из которых состоит множество, называются элементами множества.

K-means, EM: алгоритмы используемые для кластеризации.

Преобразование Фурье: операция, сопоставляющая вещественные переменные разных функций. Новая функция описывает гармонические колебания с разными частотами.

Оглавление

ВВЕДЕНИЕ	12
ОБЗОР ЛИТЕРАТУРЫ	13
1. ОБЪЕКТ И МЕТОДЫ ИССЛЕДОВАНИЯ	14
1.1 Общие сведения	14
1.2 История кластеризации	15
1.3 Примеры алгоритмов кластеризации	15
1.4 Пакет STATISTICA	16
2. РАСЧЕТЫ И АНАЛИТИКА	17
3. РЕЗУЛЬТАТЫ ПРОВЕДЕННОГО ИССЛЕДОВАНИЯ	19
4. ФИНАНСОВЫЙ МЕНЕДЖМЕНТ, РЕСУРСОЭФФЕКТИВНОСТЬ И РЕСУРСОСБЕРЕЖЕНИЕ	26
5. СОЦИАЛЬНАЯ ОТВЕТСТВЕННОСТЬ	42
ЗАКЛЮЧЕНИЕ	59
СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ	60

ВВЕДЕНИЕ

Кластерный анализ используется для разбиения совокупности объектов на однородные группы. Это задача многомерной классификации данных. Плюс кластерного анализа в том, что он позволяет производить разбиение объектов не по одному параметру, а по целому набору признаков. Так же кластерный анализ не накладывает никаких ограничений на вид рассматриваемых объектов, и позволяет рассматривать множество исходных данных практически произвольной природы.

В работе решается задача классификации аудиофайлов, которые были получены в результате аудиомониторинга пчелиных ульев. С помощью аудиомониторинга можно автоматизировать некоторые факторы пчелиных ульев. Необходимо использовать кластерный анализ, чтобы выполнить классификацию образцов звука, собранных электронным мониторингом ульев. Файлы относятся к одной из трех категорий: жужжание пчел, стрекот сверчков и шум.

ОБЗОР ЛИТЕРАТУРЫ

При написании данной работы были использованы научная и учебно-методическая литература, статьи в периодических изданиях.

Для изучения вопроса кластерного анализа аудиофайлов было проанализировано множество алгоритмов кластеризации, которые были описаны в работах[1, 2].

Основными источниками, которые смогли раскрыть необходимые теоретические основы, явились в большей части электронные ресурсы. В сфере статистического анализа первопроходцем являлся программный продукт Statistica, разработанный компанией StatSoft. Данный продукт был описан в источнике [3].

Необходимая информация по используемой мультисенсорной электронной системе мониторинга ульев BeePi была изучена в источнике[4].

1. ОБЪЕКТ И МЕТОДЫ ИССЛЕДОВАНИЯ

1.1 Общие сведения

Кластеризация - это разбиение множества объектов на некоторые однородные подмножества, параметры которых изначально неизвестны. Для некоторой задачи количество кластеров может быть произвольным или фиксированным. Пусть имеется набор данных $X_n = \{x_1, \dots, x_n\} \subset X$ ($n > 0$) и функция, определяющая степень сходства объектов, в большинстве случаев это функция расстояния между объектами $\rho(x_i, x_j)$.

Требуется разбить последовательность X_n на непересекающиеся подмножества так, чтобы каждый кластер состоял из объектов, близких по метрике ρ , а объекты разных кластеров существенно отличались. Алгоритм кластеризации – это функция $A: X \rightarrow Y$ которая любому объекту $x \in X$ ставит в соответствие метку кластера $y_i \in Y$. Чаще всего множество Y заранее не известно и дополнительной задачей является определение оптимального числа кластеров с точки зрения того или иного показателя качества кластеризации. Конечной целью процесса кластеризации является получение содержательных сведений о структуре исследуемых данных, что, как правило, является начальным этапом их более детального анализа.

В результате применения различных методов кластеризации могут быть получены неодинаковые результаты, это является следствием особенности работы того или иного алгоритма, которые следует учитывать при выборе метода кластеризации для конкретной задачи. [5]

Кластеризация может использоваться в таких областях как: информатика, экономика, маркетинг, лингвистика, астрономия.

Задачи кластеризации:

- понимание данных – разбиение на группы позволяет обработку данных упростить;
- сжатие данных – позволяет сократить исходные данные;

- обнаружение новизны – выделение объектов, не относящихся ни к одному из кластеров.

1.2 История кластеризации

Предшественником кластеризации являлась классификация – принцип распределения однородных объектов по группам по общему признаку. Одной из первых работ по классификации была – теория классификации и систематизации, предложенная Огюстеном Декандром для классификации растений в 1813 году. Вскоре классификация нашла применение в других науках, которые имели дело со множествами иерархически организованными объектами.

Как самостоятельная дисциплина кластерный анализ стал развиваться в XX веке. Одной из первых публикаций по данной теме является статья Яна Чекановского, написанная в 1911 году. В своей статье он предложил идею о структурной классификации, что содержала основную мысль кластерного анализа – выделение групп близких объектов.

Все же сам термин «кластерный анализ» был введен лишь в 1939 году английским ученым Трионом. Те методы, что используются сейчас, были предложены в 1960-е годы. В те года было написано огромное количество публикаций по данной теме.

Далее количество публикаций значительно сократилось. Объясняется это тем, что следующие года были потрачены на попытки осмысления существующих алгоритмов и способы их применения. Позже развитие продолжалось уже в контакте с теми областями, где используется кластерный анализ. [6]

1.3 Примеры алгоритмов кластеризации

В этом разделе будут приведены примеры кластеризации, которые будут использованы в работе.

K-means - целью данного алгоритма является разделение m наблюдений на k кластеров. Каждое наблюдение при этом относится к тому кластеру, к центру которого оно ближе всего. В качестве меры близости используется Евклидово расстояние; [7]

EM - масштабируемый алгоритм кластеризации. В основе алгоритма лежит предположение, что данные в каждом кластере подчиняются нормальному распределению. С учетом этого предположения можно определить математическое ожидание и дисперсию, которые соответствуют закону распределения элементов в кластере. [8]

1.4 Пакет STATISTICA

Для этого кластерного анализа был выбран пакет STATISTICA, так как он наиболее простой, широко востребован. Все расчеты программы отличаются высокой точностью, а подробное руководство по проведению статистических операций поможет сориентироваться в интерфейсе программы и выполнить расчеты безошибочно. Компьютерная программа анализа данных Statistica подходит для всех версий операционной системы Windows, как ранних версий (Windows 95, XP, Vista), так и для Windows 10. Существуют варианты как для 32-битных Windows, так и для операционных систем на 64 бита. Русскоязычный пакет прикладных программ Statistica доступен любому пользователю, независимо от технических характеристик его компьютера - даже на слабых моделях есть возможность выполнять все необходимые операции, поскольку программа имеет высокую степень оптимизации и потребляет небольшое количество ресурсов компьютера. Statistica является единственным пакетом, который способен максимально точно провести статистические расчёты, обладает быстродействием при работе с большим объёмом данных. [3]

2. РАСЧЕТЫ И АНАЛИТИКА

Пчеловоды определяют состояние своих пчелиных ульев, слушая их. Гудение пчел несет информацию о поведении колонии. Аудиомониторинг улья является ключевым компонентом ЭВМ, который привлекает значительные усилия в области исследований и разработок, потому что пчелы генерируют специфические звуки. Ранее заказчиком была разработана система для профилирования и анализа акустических сигнатур колоний пчел под названием BeePi. Зонд с микрофоном вставляется через отверстие, просверленное сзади улья. Микрофон подключен к компьютеру с аудиокартой и программным обеспечением для обработки звука, которое преобразовывает Фурье (FFT) в формат файла-контейнера для хранения записи оцифрованного аудиопотока (WAV). [4]

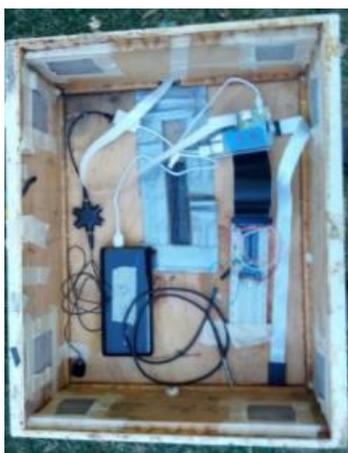


Рисунок 1 - Аппаратные компоненты системы BeePi



Рисунок 2 - Развернутый монитор системы VeePi

Заказчик получил классификацию путем ручной маркировки 9000 аудиофайлов, собранных упомянутой выше системой, прослушивая каждый файл и помещая его в одну из 3 непересекающихся категорий: жужжание пчел, щебетание сверчков и сторонний шум. Финальный набор данных включал в себя по 3000 образцов каждого типа категорий. Собранные аудиофайлы были поделены по переменным:

- 1-40 MFCC (Mel-frequency cepstral coefficients): мел - это единица высоты звука, основанная на восприятии этого звука нашими органами слуха. Высота звука не совсем линейно зависит от его частоты. Такая зависимость не претендует на большую точность и описывается простой формулой $m = 1125 \ln(1 + f/700)$. [9]
- 41-53 STFT (Short-time Fourier transform): представляет собой Фурье-преобразование связанных. Процедура вычислений - разделяет более длинный временной сигнал на более короткие сегменты равной длины. [10]
- 54-181 melspectrogram
- 182-188 STFT spectral contrast
- 189-193 tonnetz

3. РЕЗУЛЬТАТЫ ПРОВЕДЕННОГО ИССЛЕДОВАНИЯ

Нашей задачей являлась кластеризация аудиофайлов, собранных заказчиком, с помощью пакета Statistica. Необходимо, что бы программа при использовании различных алгоритмов отделила различные виды аудиофайлов друг от друга. Для этого возьмем по 200 переменных из каждой категории и поместим в общий файл.

Используя алгоритм k-средних, делая выборку по переменным 1-40 и выделяя 2 или 3 кластера нет точного деления по кластерам, возникает большая вероятность ошибки. Из графика видно, что средние значения данных переменных в различных кластерах сильно схожи. Объекты, у которых значение переменных между -10 и 10 уже могут принадлежать как 1, так 2 кластеру, так и 3. Для их кластеризации, возможно, подойдут другие переменные.

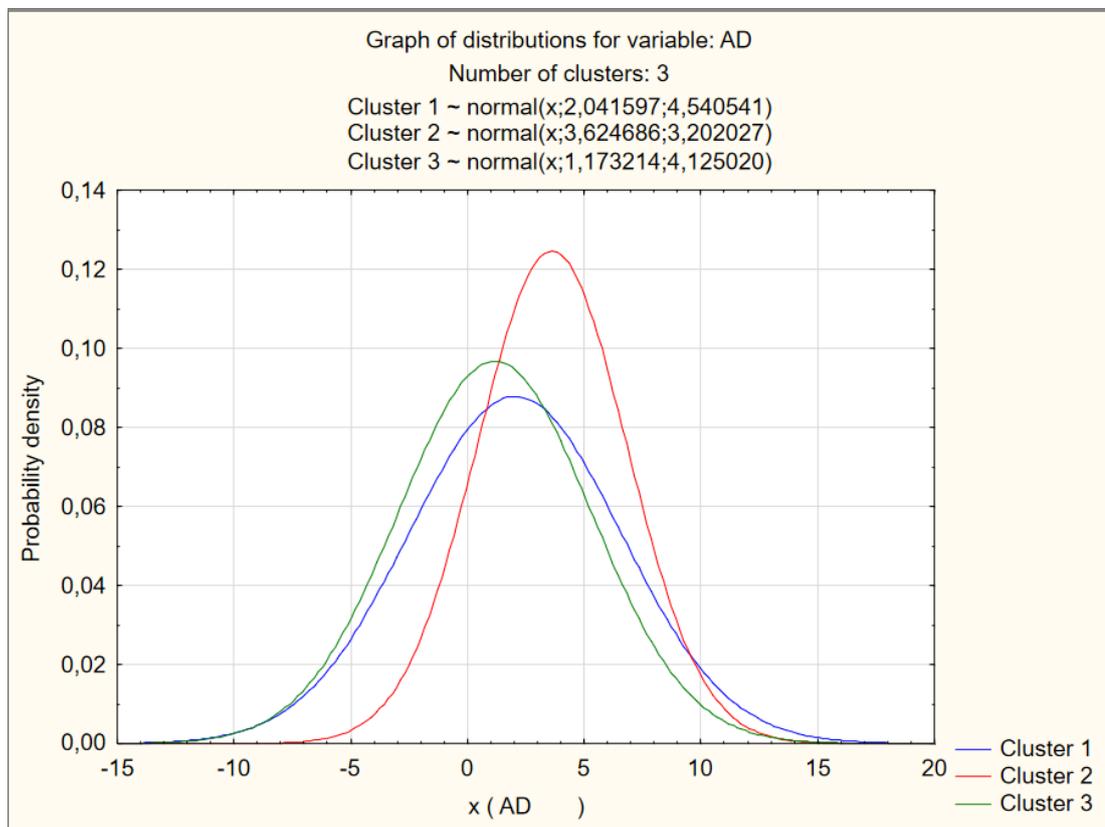


Рисунок 3 - Результат работы программы по алгоритму k-средних переменных 1-40

Такой же результат мы получим при использовании EM - масштабируемого алгоритма кластеризации. Следовательно, необходимо попробовать произвести кластеризацию по другим переменным. Далее возьмем переменные 41-53.

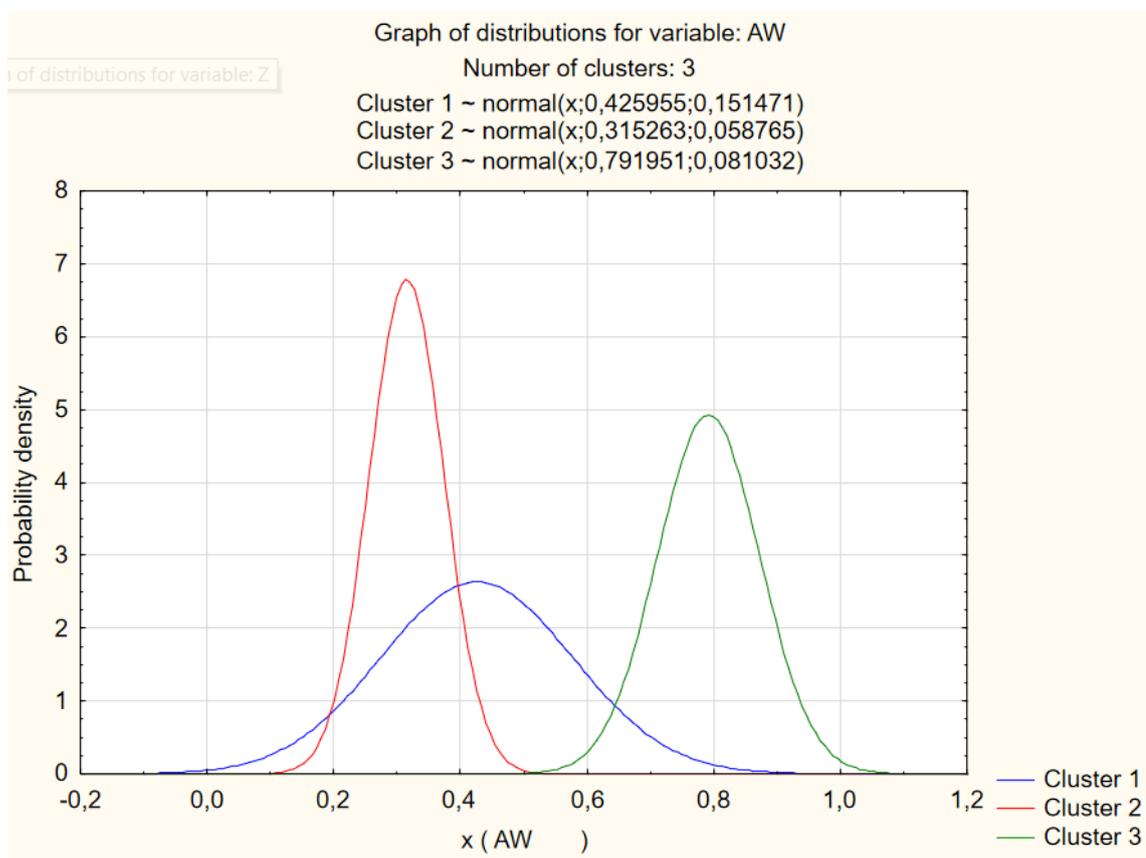


Рисунок 4 - Результат работы программы по алгоритму k-средних переменных 41-53

Из графика видно, что средние значения переменных отличаются лишь в кластерах 2 и 3. А кластер 1 имеет с обоими кластерами большую площадь, находящуюся под плотностями распределения. Из этого следует вывод, что данная кластеризация тоже нам не подходит.

Попробуем для кластеризации взять переменные 182-188 и использовать алгоритм k-средних.

ANOVA for continuous variables (Лист1 in все_звук)

Number of clusters: 3

Total number of training cases: 600

	Between SS	df	Within SS	df	F	p value
FZ	280,42	2	576,389	597	145,224	0,000000
GA	1533,21	2	542,734	597	843,255	0,000000
GB	540,98	2	236,093	597	683,973	0,000000
GC	14340,72	2	1745,702	597	2452,140	0,000000
GD	369,79	2	592,536	597	186,288	0,000000
GE	186,90	2	1133,325	597	49,227	0,000000
GF	0,00	2	0,129	597	5,962	0,002731

Рисунок 5 - p value для переменных 182-188

На данном этапе видно, что переменная GF обладает значением p value отличным от 0. Особенностью P-значений является их неустойчивость на эквивалентных выборках, что может стать препятствием для воспроизводимости результатов эксперимента. Следовательно произведем анализ не учитывая данную переменную.

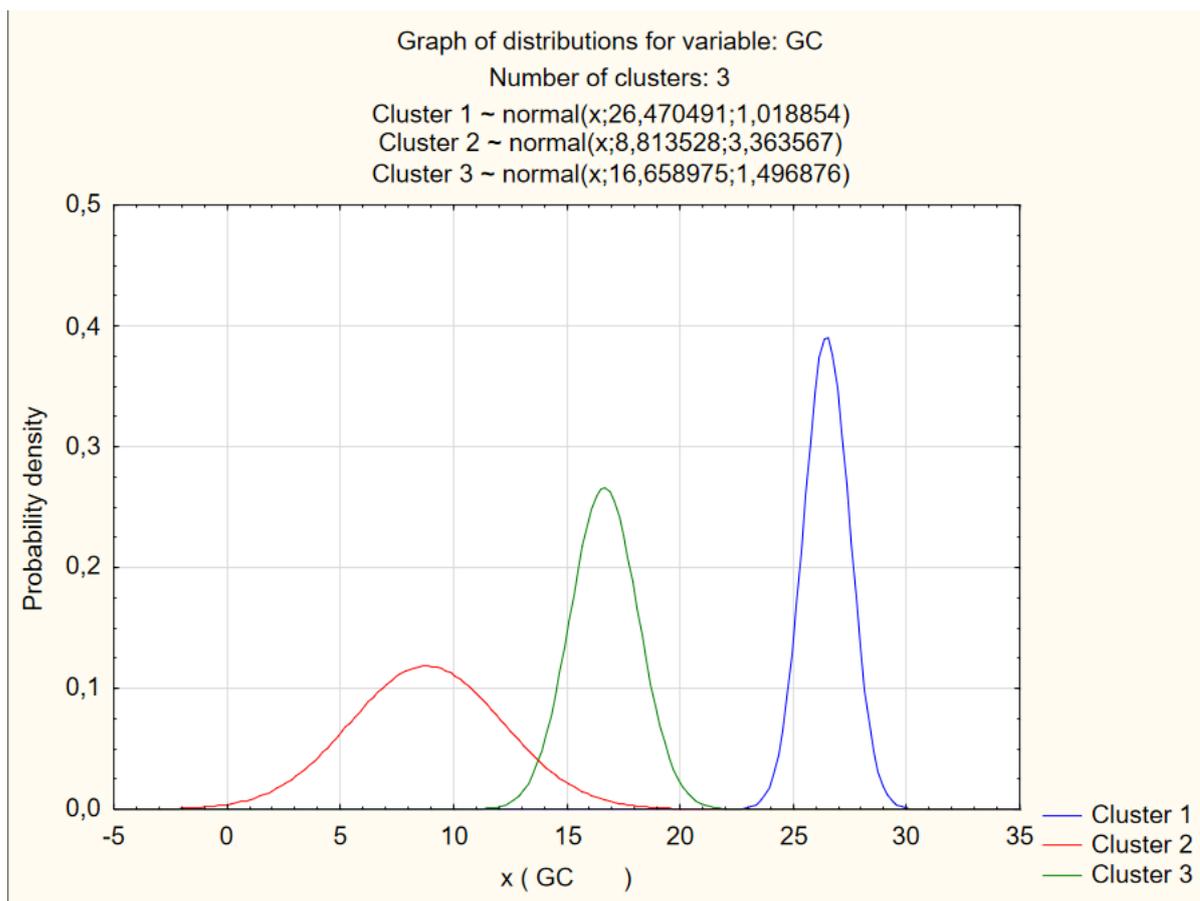


Рисунок 6 – Результат работы программы для переменных 182-187

Из графика видно, что средние значения переменных в различных кластерах сильно отличаются и при этом площадь, находящаяся под обеими плотностями распределения очень мала по сравнению с общей площадью под кривыми. Это говорит о том, что данные переменные могут быть использованы для кластеризации. Значит объекты, у которых значение данной переменной > 24 , можно отнести к 1 кластеру. Объекты, у которых значение переменной между 12 и 19 могут принадлежать как 2, так и 3 кластеру. Объекты, у которых значение переменной между 19 и 23, относятся к 3 кластеру, а объекты, у которых значение переменной < 12 , относятся ко 2 кластеру.

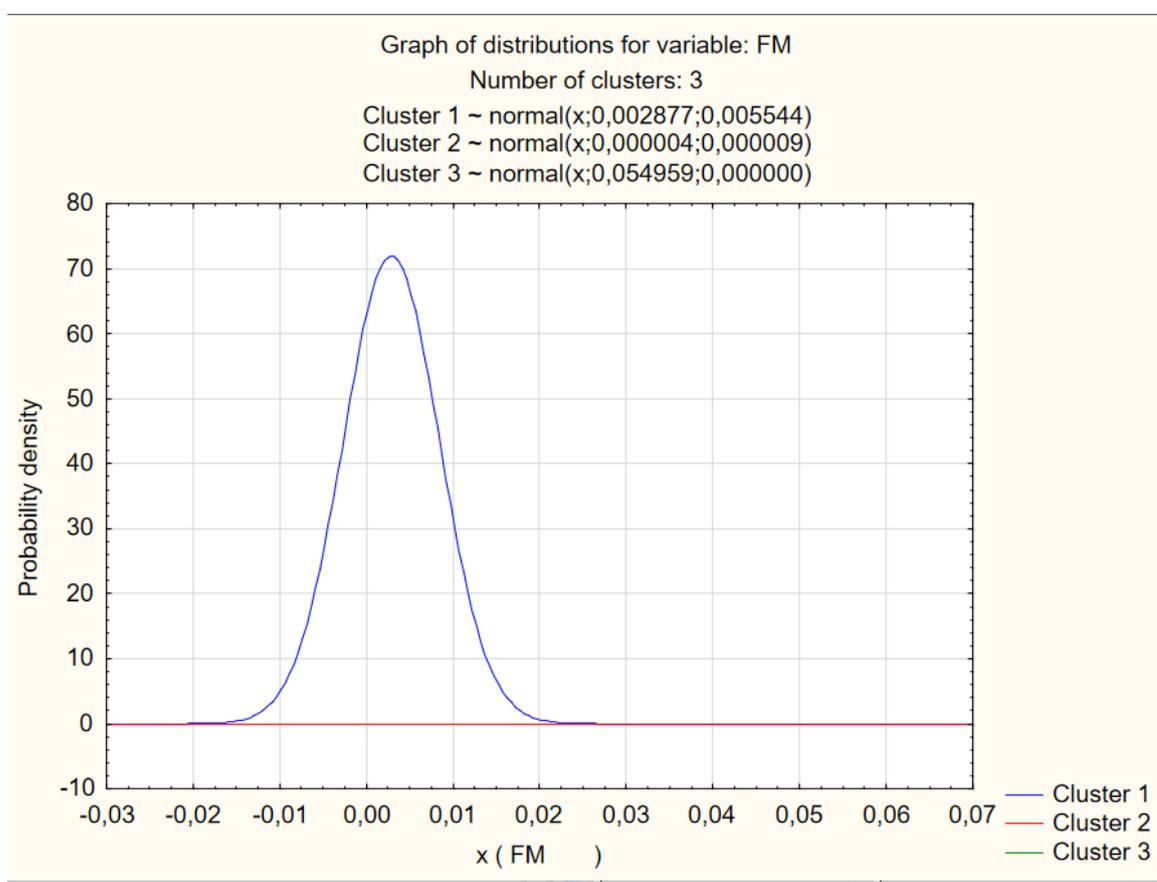


Рисунок 7 – результат работы программы с переменными 54-181

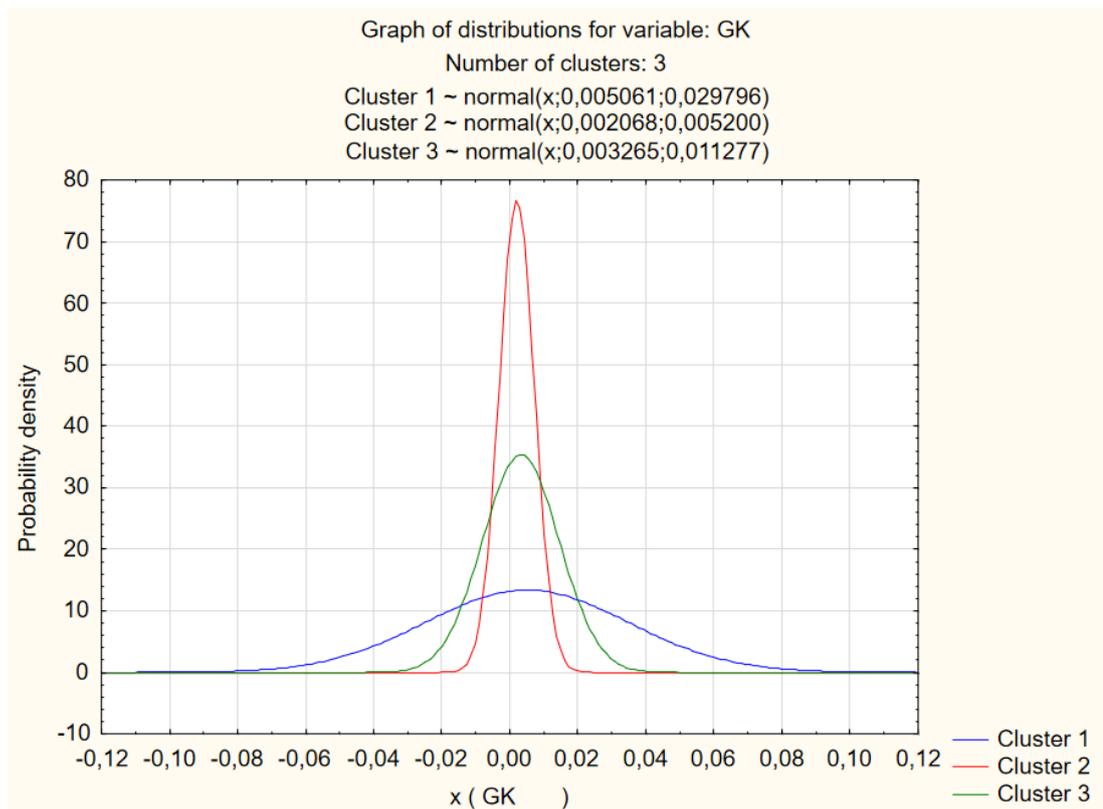


Рисунок 8 – Результат работы программы с переменными 189-193

На рисунках 7 и 8 на графиках видно, что при выборе переменных из диапазона 54-181 или 189-193 средние значения данных переменных в различных кластерах сильно схожи. Из этого следует, что для кластеризации следует подобрать другие переменные.

Далее возьмем по 200 переменных из 2 категорий: жужжание пчел и стрекот сверчков.

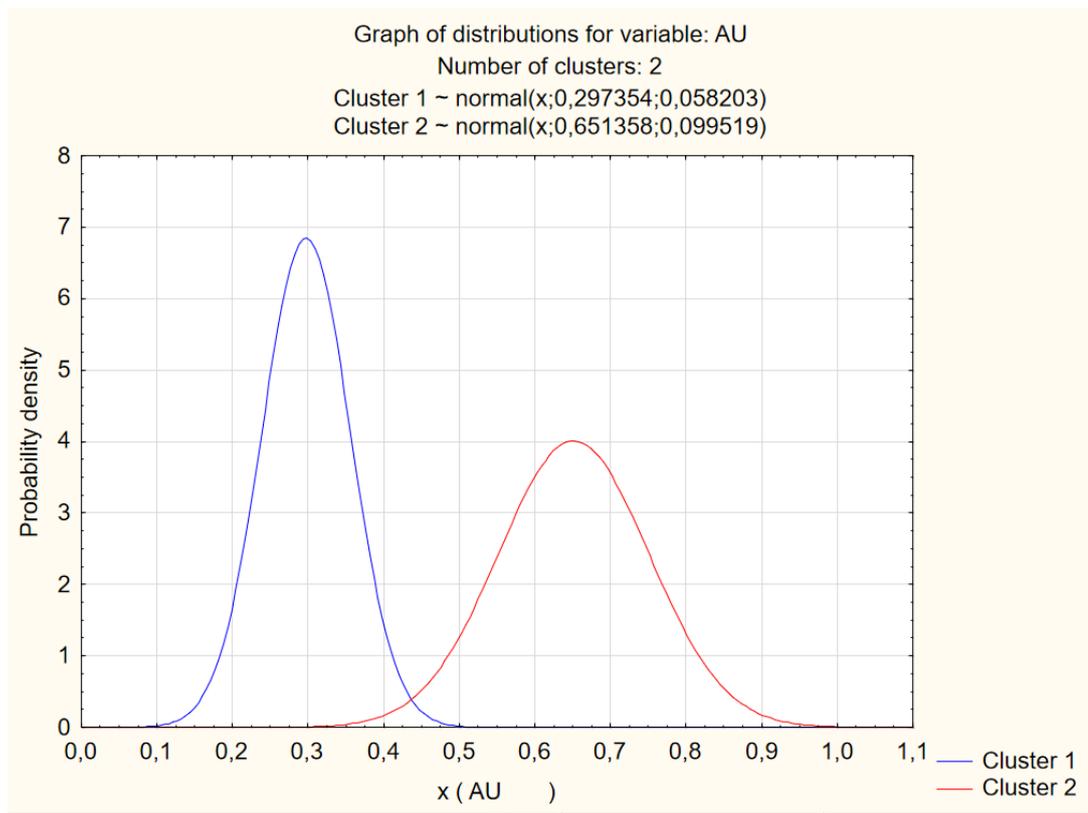


Рисунок 9 – Результат работы программы для переменных 41-53 алгоритмом k-средних

Из графика видно, что средние значения данных переменных в различных кластерах сильно отличаются и при этом площадь, находящаяся под обеими плотностями распределения очень мала по сравнению с общей площадью под кривыми. Это говорит о том, что данные переменные могут быть использованы для кластеризации. Значит объекты, у которых значение данной переменной $< 0,4$, с уверенностью могут быть отнесены к 1 кластеру. А объекты, у которых значение переменной $> 0,5$, можно отнести ко 2 кластеру. А объекты, у которых значение переменных между 0,4 и 0,5 уже могут принадлежать как 1, так и 2 кластеру. Для их кластеризации, возможно, подойдут другие переменные.

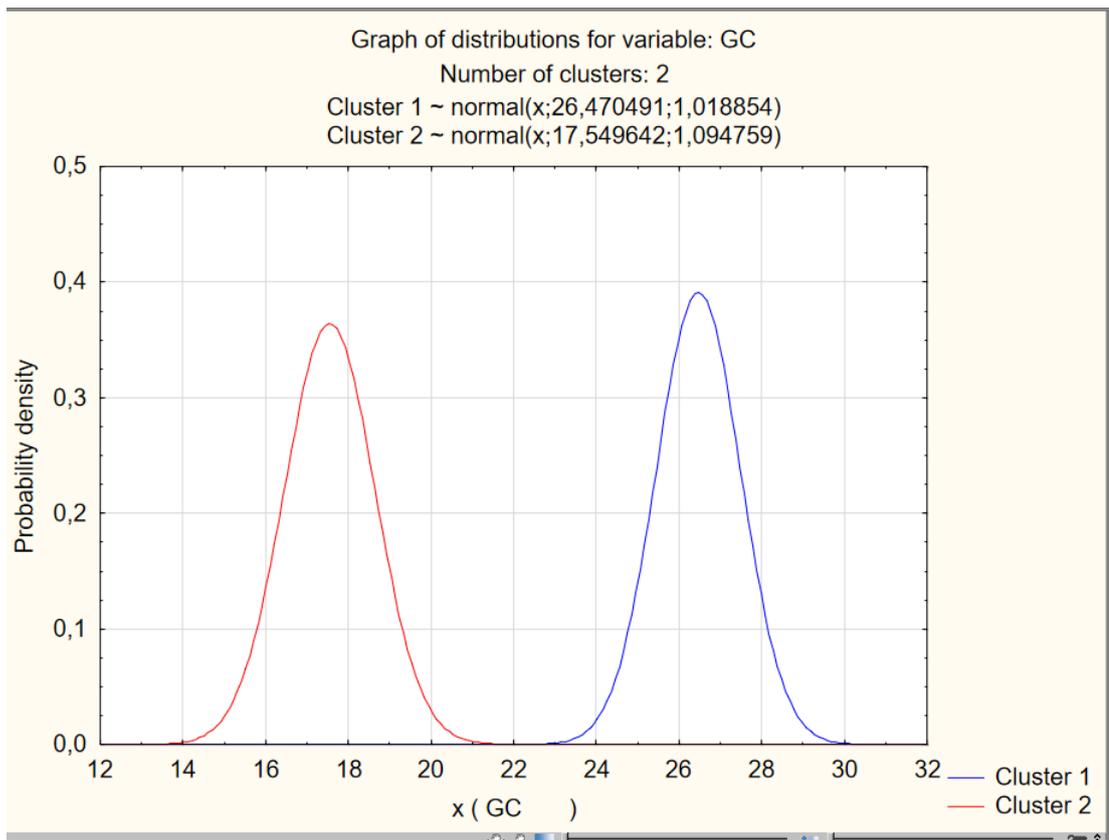


Рисунок 10 – Работа программы с переменными 182-188 по алгоритму k-средних

На данном графике видно, что переменные поделились на 2 кластера. Это говорит о том, что данные могут быть использованы для кластеризации. Объекты, у которых значения данной переменной <22 будут отнесены к кластеру 2, а у которых значения данной переменной >22 , будут отнесены к 1 кластеру.

4. ФИНАНСОВЫЙ МЕНЕДЖМЕНТ, РЕСУРСОЭФФЕКТИВНОСТЬ И РЕСУРСОСБЕРЕЖЕНИЕ

Потенциальные потребители результатов исследования

Разрабатываемый продукт собирает и выполняет кластерный анализ большого объема аудиофайлов по спектральным характеристикам при помощи математического пакета Statistica, а именно выполнен кластерный анализ аудиофайлов, собранных на пасеке и включающих все шумы происходящие там. В результате применения кластерного анализа строится карта, по которой моно будет определить соответствующие характеристики аудиофайлов для возможности попадания в нужный сегмент. Собранная информация будет помещаться в структуру в терминалах кластеров объектов, сходных между собой, будет отражена в иерархическом дереве различными ветвями. В результате анализа появится возможность обнаружить кластеры и интерпретировать их.

Основным потребителем является хозяин пасеки, по чьей просьбе был разработан этот проект. В дальнейшем данный проект может быть использован и другими людьми, имеющими пасеки при необходимости произвести анализ данных.

Сегментирование рынка — это универсальный способ разделения любой отрасли (в нашем случае будет считается собираемая статистика по сотруднику) на однородные группы, это позволит проанализировать ассортимент всех частей приложения, для этого необходимо составить карту сегментирования собираемых статистик для компаний.

Карта сегментирования рынка для статистики представлена в таблице 1.

Таблица 1 - Карта сегментирования рынка для проекта

		Статистика по сотрудникам компании			
		Активное приложение и заголовок его окна	Время активности пользователя за ПК	Действия за ПК во время активности	Продуктивность использования рабочего времени
Размер компании	Крупные		■		▨
	Средние	■	■		▨
	Мелкие	■		■	▨

где

■ - клиент часть

▨ - представление

SWOT-анализ

SWOT является аббревиатурой и подразделяется на следующие:

- strengths – сильные стороны проекта
- weaknesses – слабые стороны проекта
- opportunities – возможности проекта
- threats – угрозы для проекта

SWOT - анализ – это один из самых распространенных методов для оценки в комплексе внутренних и внешних факторов, влияющие на развитие компании. Анализ включает в себя свойства сильных и слабых сторон организации, а также возможности и угрозы со стороны внешней среды.

Матрица SWOT для проекта представлена в таблице 2.

Таблица 2 – Матрица SWOT проекта

		Сильные стороны	Слабые стороны
		С1. Система собирает данные без участия чел. фактора С2. Удобный способ представления данных С3. Упрощенный анализ собранных данных С4. Простота интерфейса	Сл1. Ограниченные функциональные возможности Сл2. Отсутствие прототипа разработки Сл3. Актуальность Сл4. Узконаправленное приложение

Возможности	В1. Появление дополнительного спроса на новый продукт В2. Контроль информации В3. Оптимизация системы		
Угрозы	У1. Узкая специализация и область применения разработки У2. Увеличение функционала конкурентами		

Второй этап состоит в выявлении соответствия сильных и слабых сторон научно-исследовательского проекта внешним условиям окружающей среды. Необходимо построить интерактивную матрицу проекта. Ее использование помогает разобраться с различными комбинациями взаимосвязей областей матрицы SWOT. Интерактивные матрицы представлены в таблицах 5-8.

Таблица – 5 интерактивная матрица для сильных сторон и возможностей

Направления развития		С1	С2	С3	С4
	В1	-	+	+	+
	В2	-	+	+	0
	В3	0	+	+	+

Таблица – 6 интерактивная матрица для слабых сторон и возможностей

Сдерживающие факторы		Сл1	Сл2	Сл3	Сл4
	В1	+	+	-	+
	В2	0	0	0	0
	В3	-	-	-	-

Таблица – 7 интерактивная матрица для сильных сторон и угроз

Угрозы развития		С1	С2	С3	С4
	У1	-	0	0	-
	У2	0	+	+	+

Таблица – 8 интерактивная матрица для слабых сторон и угроз

Уязвимости		Сл1	Сл2	Сл3	Сл4
	У1	-	-	+	+
	У2	0	-	0	0

В рамках третьего этапа должна быть составлена итоговая матрица SWOT-анализа.

Таблица – 9 итоговая матрица SWOT-анализа

	<p>Сильные стороны научно-исследовательского проекта: С1. Система собирает данные без участия чел. фактора С2. Удобный способ представления данных С3. Упрощенный анализ собранных данных С4. Простота интерфейса</p>	<p>Слабые стороны научно-исследовательского проекта: Сл1. Ограниченные функциональные возможности Сл2. Отсутствие прототипа разработки Сл3. Актуальность Сл4. Узконаправленное приложение</p>
<p>Возможности: В1. Появление дополнительного спроса на новый продукт В2. Контроль информации В3. Оптимизация системы</p>	<p>Направления развития: В1С2С3С4 – сильные стороны способствуют появлению дополнительного спроса на данную разработку В2С3С4 – удобный способ реализации способствует облегченному контролю информации В3С2С3С4 - оптимизировать систему с целью улучшения интерфейса и повышения производительности труда пользователя</p>	<p>Сдерживающие факторы: В1Сл1Сл2Сл4 – отсутствие и узконаправленность подобных решений приводит к низкому спросу</p>
<p>Угрозы: У1. Узкая специализация и область применения разработки У2. Увеличение функционала конкурентами</p>	<p>Угрозы развития: У2С2С3С4 – конкуренты реализуют более упрощенный и удобный способ анализа информации</p>	<p>Уязвимости: У1С3С4 – данная разработка актуальна малому кругу пользователей</p>

Положительные и слабые стороны, которые были обозначены в ходе проведенного анализа, дают возможность спланировать необходимые изменения, слабые стороны необходимо по возможности минимизировать, базируясь прежде всего на имеющихся сильных сторонах.

Структура работ в рамках научного исследования

Для организации правильного создания проектного решения необходимо:

1. определить структуру работ;
2. Установить процентную нагрузку на каждое лицо, участвующее в создании проекта;
3. рассчитать продолжительность работ;
4. построение графика проведения научных исследований
5. определить и спланировать занятость каждого из участников.

В данном проекте участвуют 2 непосредственных лица: Научный руководитель (далее Р) и инженер (далее И). Требуется создать линейный график для распределения результатов планируемых работ между Р и И.

Перечень этапов, работ и распределение исполнителей представлена в таблице 3.

Таблица 3 – Перечень этапов, работ и распределение исполнителей

Основные этапы	№ раб	Содержание работ	Должность исполнителя
Разработка технического задания	1	Составление и утверждение технического задания	Р и И
	2	Подбор и изучение материалов по теме проекта	Р и И
Выбор направления исследования	3	Выбор направления исследований	Р и И
	4	Календарное планирование работ по теме	И
	5	Проведение теоретических расчетов и обоснований	И
Проектирование	6	Создания базы данных	Р и И
	7	Выбор программы и алгоритмов кластеризации	Р и И
Разработка	8	Разработка базовой составляющей проекта	И
	9	Разработка схем и алгоритмов кластеризации	И
Тестирование	10	Наполнение промежуточных значений и данных для исследования и выявления нестабильностей	И
Теоретические и экспериментальные	11	Сопоставление результатов с эталонными результатами	Р и И

исследования			
Обобщения и оценка результатов	12	Оценка эффективности полученных результатов	И
	13	Определение целесообразности проведения ОКР	Р и И

Определение трудоемкости выполнения работ

Диаграмма Ганта (от англ. Gantt chart, также ленточная диаграмма, график Ганта) – горизонтальный, ленточный график, обусловленный столбчатыми элементами (гистограмма) в котором, данные представляются в виде протяженности временных отрезков, характеризующимися начала и конца выполняемых работ.

Для построения графика, длительность каждого из этапов следует перевести в календарные дни, для этого используется формула:

$$t_{ожi} = \frac{(3t_{мини} + 2t_{маxi})}{5}$$

где $t_{ожi}$ – ожидаемая трудоемкость выполнения i -ой работы, чел.-дн.;

$t_{мини}$ – минимально возможная трудоемкость выполнения i -ой работы (оптимистическая оценка: в предположении наиболее благоприятного стечения обстоятельств), чел.-дн.;

$t_{маxi}$ – максимально возможная трудоемкость i -ой работы (пессимистическая оценка: в предположении наиболее неблагоприятного стечения обстоятельств), чел.-дн.

Исходя из трудоемкости работ, определяется продолжительность каждой работы в рабочих днях T_p , учитывающая параллельность выполнения работ несколькими исполнителями. Вычисление необходимо для обоснованного расчета ЗП, так как удельный вес зарплаты в общей сметной стоимости научных исследований составляет приблизительно 65 %.

Продолжительность одной работы рассчитывается по формуле:

$$T_{pi} = \frac{T_{ож}}{\Psi_i}$$

где T_{pi} – продолжительность одной работы, раб. дней.;

$Ч_i$ – численность исполнителей, выполняющих одновременно одну и ту же работу на данном этапе, чел.

Разработка графика проведения научного исследования

Построение линейного графика зависит от длительности этапов на период рабочих дней, необходимо рассчитать эту длительность, а затем перевести в календарные дни.

Продолжительность выполнения этапов рассчитывается по формуле:

$$T_{ki} = T_{pi} * k_{\text{кал}}$$

где T_{ki} – продолжительность выполнения i -й работы в календарных днях;

T_{pi} – продолжительность выполнения i -й работы в рабочих днях;

$k_{\text{кал}}$ – коэффициент календарности.

Баланс рабочего времени – система показателей, характеризующих ресурсы рабочего времени работающих, их распределение по видам затрат

Баланс рабочего времени представлен в таблице 4.

Таблица 4 - Баланс рабочего времени

Показатели рабочего времени	Р	И
Календарное число дней	365	365
Количество нерабочих дней		
- выходные дни	48	44
- праздничные дни	14	14
Потери рабочего времени		
- отпуск	56	28
- невыходы по болезни	1	1

Действительный годовой фонд рабочего времени научно-технического исполнителя, раб.дн, рассчитывается по формуле:

$$F_{\text{д}} = \text{Календарные дни} - \text{выходные дни} - \text{праздничные дни} - \text{отпуск} \\ - \text{невыходные по болезни}$$

Коэффициент календарности рассчитывается по формуле:

$$k_{\text{кал}} = \frac{T_{\text{КАЛ}}}{T_{\text{кал}} - T_{\text{вых}} - T_{\text{ПД}}}$$

где $T_{\text{КАЛ}}$ – календарные дни;

$T_{ВД}$ – выходные дни;

$T_{ПД}$ – праздничные дни.

Расчет коэффициента календарности и действительного годового фонда, представлены в таблице 5.

Таблица 5 - Расчет коэффициента календарности и действительного годового фонда

Исполнители	$T_{КАЛ}$	$T_{ВД}$	$T_{ПД}$	$F_{Д}$	$k_{КАЛ}$
Р	365	52	14	250	1,22
И	365	52	14	274	1,22

Длительность и трудозатраты на выполнение этапов проекта приведены в таблице 6.

Календарный план-график проведения этапов проекта представлен на Рис. 1.

Таблица 6 - Длительность и трудозатраты на выполнение этапов проекта

Этапы	Исполнитель и	Продолжительность работ, дни						Длительность работ, чел/день			
		t_{mini}		t_{maxi}		$t_{ожи}$		T_{pi}		T_{ki}	
		Р	И	Р	И	Р	И	Р	И	Р	И
Составление и утверждение технического задания	Р и И	5	20	6	26	5.4	22.4	2.7	11.2	3	13
Подбор и изучение материалов по теме проекта	Р и И	1	9	2	10	1.4	9.4	0.7	4.7	1	6
Выбор направления исследований	Р и И	1	3	1	7	1	4.6	0.5	2.3	1	3
Календарное планирование работ по теме	И		1		3	0	1.8	0	1.8	0	2
Проведение теоретических расчетов и обоснований	И		5		10	0	7	0	7	0	8
Создания базы данных	Р и И	3	5	5	10	3.8	7	1.9	3.5	2	4

Выбор программы и алгоритмов кластеризации	Р и И	3	10	4	10	3.4	10	1.7	5	2	6
Разработка базовой составляющей проекта	И		5		7	0	5.8	0	5.8	0	7
Разработка схем и алгоритмов кластеризации	И		10		11	0	10.4	0	10.4	0	12
Наполнение промежуточных значений и данных для исследования и выявления нестабильностей	И		1		3	0	1.8	0	1.8	0	2
Сопоставление результатов экспериментов с теоретическими исследованиями	Р и И	1	1	1	2	1	1.4	0.5	0.7	1	1
Оценка эффективности полученных результатов	И		4		5	0	4.4	0	4.4	0	5
Определение целесообразности проведения ОКР	Р и И		3		3	0	3	0	1.5	0	2
Итоги		14	77	19	107	19.8	88	8	60.1	10	71

На основе полученных данных в таблице 3, строится календарный план-график. График строится для максимального по длительности исполнения работ в рамках научно-исследовательского проекта, с разделением по месяцам.

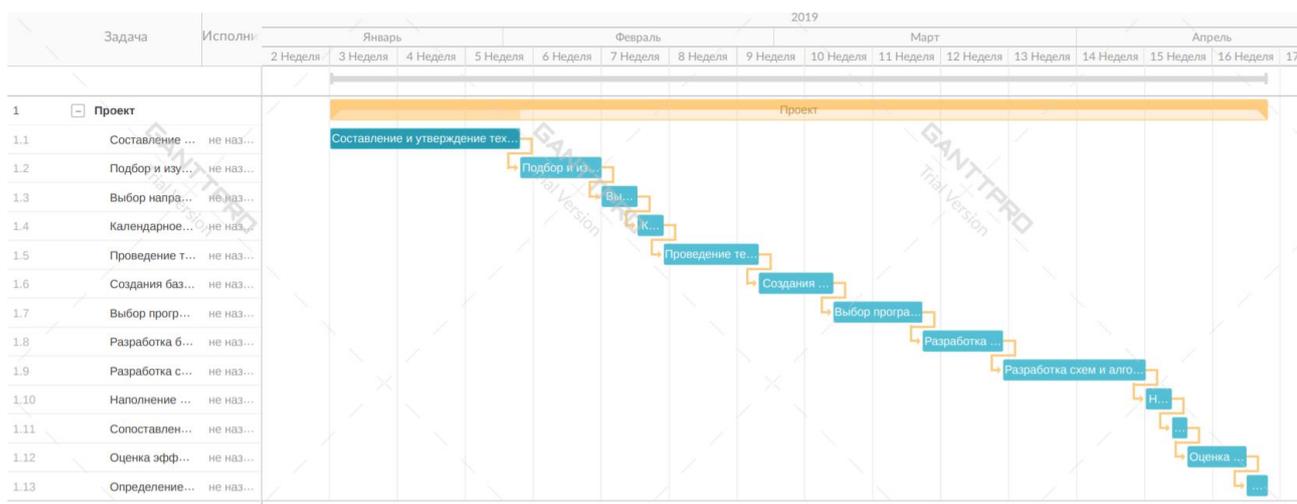


Рис.1. - Календарный план-график проведения этапов проекта, диаграмма Ганта

Общее число этапов становления проекта составило 13 шт. Ожидаемая трудоемкость работ для научного руководителя составила 8 дней, а для инженера 60 дней. Общая максимальная длительность выполнения работы составила $8 + 60 = 68$ (дней) календарных дней.

Бюджет научно-технического исследования (НТИ)

Необходимо рассчитать основную з/п для Р и И.

Средний оклад руководителя ТПУ в должности доцент, со степенью кандидата технических наук и без учета районного коэффициента, составляет 33664 (руб./мес.).

Оклад инженера составляет, 26300 (руб./мес.), без учета районного коэффициента.

Среднедневная заработная плата рассчитывается по формуле:

$$Z_{\text{дн}} = \frac{Z_{\text{осн}} * M}{F_{\text{д}}}$$

где $Z_{\text{м}}$ – месячный должностной оклад исполнителя, руб.;

M – количество месяцев работы без отпуска в течение года;

Рабочие месяцы с учетом праздничных и выходных, данные, заданные в таблице 4, и рассчитывается по формуле:

$$M = \frac{(T_{\text{кал}} - T_{\text{вых}} - T_{\text{ПД}})}{T_{\text{Мдн}}}$$

где $T_{\text{Мдн}}$ - среднемесячное число календарных дней (29.3)

Основная з/п исполнителя, рассчитывается по формуле:

$$Z_{\text{осн}} = Z_{\text{тс}} * k_p$$

где $Z_{\text{тс}}$ – заработная плата по тарифной ставке, руб.;

k_p – районного коэффициент, на текущий момент составляет 30%.

Расчёт основной заработной платы представлен в таблице 7.

Таблица 7 - Расчёт основной заработной платы

Исполнитель	$Z_{\text{тс}}$	k_p	$Z_{\text{осн}}$	F_d	M	$Z_{\text{дн}}$
Р	33664	1,3	43763,2	250	10,4	1820,55
И	26300	1,3	34190	274	10,4	1297,72
Итого						3118,27

Материальные затраты

В материальные затраты входят затраты на покупку программы, электроэнергию и канцелярские товары.

Расчет материальных затрат осуществляется по формуле:

$$Z_M = (1 + k_T) * \sum_{i=1}^m C_i * N_{\text{расх}i}$$

где m – количество видов материальных ресурсов, потребляемых при выполнении научного исследования;

$N_{\text{расх}i}$ – количество материальных ресурсов i -го вида, планируемых к использованию при выполнении научного исследования.

C_i – цена приобретения единицы i -го вида потребляемых материальных ресурсов (руб./шт.);

k_T – коэффициент, учитывающий транспортно-заготовительные расходы. Транспортные расходы равны 15% от стоимости материала.

Материальные затраты на создание проекта представлены в таблице 9.

Таблица 9 - Материальные затраты на создание проекта

Наименование	Цена за ед., руб.		Затраты на материалы, (З _м), руб.		Общие затраты на материалы, руб.	
	Р	И	Р	И	Р	И
Нотбук		35000		5250		40250
Персональный компьютер	30000		4500		34500	
Windows 2010	4000	4000	600	600	4600	4600
Microsoft Word	6699	6699	2300	2300	8999	8999
Statistica Basic Academic.	15000	15000	2250	2250	17250	17250
Канцелярия	500	3000	75	450	575	3450
Итого					65924	74549

Отчисления во внебюджетные фонды (страховые отчисления)

Отчисления во внебюджетные фонды рассчитывается по формуле:

$$З_{внеб} = k_{внеб} * З_{дн}$$

где $k_{внеб}$ - величина отчислений во внебюджетные фонды для образовательных учреждений равна 27,1%.

Общие затраты на основную з/п, рассчитывается по формуле:

$$З_{сум,осн} = (З_{исп1,дн} + З_{исп2,дн}) * \text{количество раб дней}$$

где $З_{исп1,дн}$ – основная з/п Р;

$З_{исп2,дн}$ – основная з/п И;

количество раб дней - соответствует таблице 8, затраченное время на Р - 8 раб/дней, а для И - 60 раб/дней.

где $З_{исп1,дн}$ – премиальная часть з/п Р;

$З_{исп2,дн}$ – премиальная часть з/п И;

Общее количество внебюджетные фондов на проект, рассчитывается по формуле:

$$Z_{\text{сум.внеб}} = Z_{\text{внеб}} * \text{количество раб дней}$$

Расчет внебюджетные фонды представлен в таблице 10.

Таблица 10- Расчет внебюджетные фонды представлен

Исполнитель	$Z_{\text{дн}}$	количество раб дней	$Z_{\text{внеб}}$
Р	1820,55	8	493,37
И	1297,72	60	351,68
Итого			845,05

$$Z_{\text{сум,осн}} = 1820,55 * 8 + 1297,72 * 60 = 94427,92$$

$$Z_{\text{сум.внеб}} = 493,37 * 8 + 351,05 * 60 = 25047,92$$

Накладные расходы

Накладные расходы рассчитаем, как 10 % от всей суммы, так как ТПУ берет процент за предоставление условий реализации проекта:

$$Z_{\text{накл}} = (\text{сумма статей } 1 \div 3) * k_{\text{нр}}$$

где $k_{\text{нр}}$ – коэффициент, учитывающий накладные расходы, в размере 16%.

Расчет величины накладных расходов приведен в таблице 11.

Таблица 11- Расчет величины накладных расходов

Исполнитель	(сумма статей 1÷3)	$Z_{\text{накл}}$ руб.
Исп.1	257948,7	41271,79
Исп. 2	299872,1	47979,53
Исп. 3	360323,8	57747,81

Формирование бюджета на ни проект

Бюджет затрат по каждому исполнению НТИ представлен в таблице 12.

Таблица 12 - Бюджет затрат по каждому исполнению НТИ

Наименование статьи	Сумма руб.		
	Исп. 1	Исп. 2	Исп. 3
1. Материальные затраты НТИ	140473	140473	140473

2. Затраты по основной з/п исполнителей темы	942427,75	125412,3	173446,8
3. Отчисления во внебюджетные фонды	25047,92	33986,74	47004,07
4. Накладные расходы	41271,79	47979,53	57747,81
5. Бюджет затрат НТИ	299220,5	347851,53	418671,6

Суммарный бюджет рассматриваемого НИ составил: 299220,5 руб.

Определение ресурсной, финансовой, бюджетной, социальной и экономической эффективности исследования

Интегральный, финансовый показатель разработки, рассчитывается по формуле:

$$I_{\text{исп}i,\text{финр}} = \frac{\Phi_{pi}}{\Phi_{\text{max}}}$$

Φ_{pi} – стоимость i -го варианта исполнения;

Φ_{max} – максимальная стоимость исполнения научно-исследовательского проекта.

$$I_{\text{исп}1,\text{финр}} = \frac{299220,5}{418671,6} = 0,71$$

$$I_{\text{исп}2,\text{финр}} = \frac{347851,53}{418671,6} = 0,83$$

$$I_{\text{исп}3,\text{финр}} = \frac{418671,6}{418671,6} = 1$$

Сравнение вариантов исполнения проекта представлено в таблице 10.

Таблица 10 – Сравнение вариантов исполнения проекта

Критерии \ Объект исследования	Весовой коэффициент параметра	Исп.1	Исп.2	Исп.3
1. Способствует росту производительности труда пользователя	0,3	5	3	4

2. Удобство в эксплуатации (соответствует требованиям потребителей)	0,2	5	2	3
3. Энергосбережение	0,1	4	3	3
4. Надежность	0,3	5	4	4
5. Материалоемкость	0,1	4	4	4
Итого	1			

Интегральный показатель ресурсоэффективности, рассчитывается по формуле:

$$I_{pi} = \sum a_i * b_i$$

где a_i – весовой коэффициент i -го варианта исполнения разработки;

b_i – бальная оценка i -го варианта исполнения разработки,

устанавливается экспертным путем по выбранной шкале оценивания;

n – число параметров сравнения.

$$I_{p-исп1} = 5 * 0,3 + 5 * 0,2 + 4 * 0,1 + 5 * 0,3 + 4 * 0,1 = 4,8;$$

$$I_{p-исп2} = 3 * 0,3 + 2 * 0,15 + 3 * 0,2 + 4 * 0,2 + 4 * 0,15 = 3,2;$$

$$I_{p-исп3} = 4 * 0,3 + 3 * 0,15 + 3 * 0,2 + 4 * 0,2 + 4 * 0,15 = 3,7.$$

Интегральный показатель эффективности, рассчитывается по формуле:

$$I_{испи} = \frac{I_{p-испи}}{I_{испи,финр}}$$

$$I_{исп1} = 4,8 / 0,71 = 6,76;$$

$$I_{исп2} = 3,2 / 0,83 = 3,85;$$

$$I_{исп3} = 3,7 / 1 = 3,7.$$

Сравнительная эффективность вариантов исполнения, рассчитывается по формуле:

$$\mathcal{E}_{ср} = \frac{I_{испи}}{I_{испи+1}}$$

$$\mathcal{E}_{\text{ср-исп1}} = 6,76 / 6,76 = 1;$$

$$\mathcal{E}_{\text{ср-исп2}} = 3,85 / 6,76 = 0,57;$$

$$\mathcal{E}_{\text{ср-исп3}} = 3,7 / 6,76 = 0,55.$$

Сравнительная эффективность разработки

Сравнительная эффективность разработки представлена в таблице 11.

Таблица 11 – Сравнительная эффективность разработки

№ п/п	Показатели	Исп.1	Исп.2	Исп.3
1	Интегральный финансовый показатель разработки	0,71	0,83	1
2	Интегральный показатель ресурсоэффективности разработки	4,8	3,2	3,7
3	Интегральный показатель эффективности	6,76	3,85	3,7
4	Сравнительная эффективность вариантов исполнения	1	0,57	0,55

Вывод

В ходе выполнения раздела «Финансовый менеджмент, ресурсоэффективность и ресурсосбережение» были, рассмотрены и решены задачи:

1. Оценка перспективности проекта и составление SWOT-анализа, что позволит конкурировать с уже существующими проектами.
2. Составлено расписание выполнения каждого этапа проекта и определено общее время, требуемое на создание ни.
3. Определено общее количество этапов для становления проекта, количество этапов составило 13 шт. Ожидаемая трудоемкость работ для научного руководителя составила 8 дней, а для инженера 60 дней. Общая максимальная длительность выполнения работы составила 68 (дней) календарных дней.
4. Суммарный бюджет, требуемый для создания НИ проекта, составил: 301631 руб.

5. СОЦИАЛЬНАЯ ОТВЕТСТВЕННОСТЬ

Разработанная система позволит произвести кластерный анализ аудиофайлов по спектральным характеристикам. Пользователь будет иметь возможность произвести анализ аудиофайлов, записанных на пасеке.

Для удобства отделения различных аудиофайлов друг от друга будет использовано наша система для более гибкой и понятной демонстрации результатов в виде иерархического дерева различными ветвями.

Область применения: автоматический аудиомониторинг пасек.

Потенциальные пользователи владельцы и сотрудники пасек, которые будут пользоваться рабочим компьютером.

Правовые и организационные вопросы обеспечения безопасности

В современном мире программы для ЭВМ являются одними из самых распространенных объектов интеллектуальной собственности, а программисты стали представителями самой многочисленной из творческих профессий. Программы для ЭВМ охраняются авторским правом, но имеют собственную специфику, которую необходимо учитывать, при защите авторских прав на программное обеспечение.

Программный продукт в соответствии с требованиями ст. 1225 Гражданского кодекса РФ является результатом интеллектуальной деятельности и подлежит охране.

Согласно ст. 8 закона «О правовой охране программ для электронных вычислительных машин и баз данных» автором программы для ЭВМ или базы данных признается физическое лицо, в результате творческой деятельности которого они созданы.

Если программа для ЭВМ или база данных созданы совместной творческой деятельностью двух и более физических лиц, то независимо от того, состоит ли программа для ЭВМ или база данных из частей, каждая из которых имеет самостоятельное значение, или является неделимой, каждое из этих лиц признается автором такой программы.

Программой же признается «объективная форма представления совокупности данных и команд, предназначенных для функционирования ЭВМ и других компьютерных устройств с целью получения определенного результата». В том числе материалы, которые были задействованы в подготовке программы или полученные в ходе составления и разработки конечного программного продукта, а также создаваемые программой аудио-визуальные отображения.

Основным охраняемым элементом программы для ЭВМ является исходный текст. Он существует в языковой форме, т.е. доступен для понимания всем, кто знает соответствующий язык программирования. По этой причине среди объектов авторского права программы для ЭВМ приравнены к литературным произведениям. Этот факт объясняет многие вопросы, связанные с защитой авторских прав на программы для ЭВМ.

Под авторское право подпадают любые написанные программы (все виды программ на любом языке) оформленные и представленные в удобоваримой форме. Охрана прав не распространяется на использованные методы, принципы функционирования, способы написания, концепции, лежащие в основе программы или ее отдельных элементов. В мире программного обеспечения (ПО) не редка ситуация, когда программный продукт (частично) может быть повторен другим лицом, и для охраны программы требуется провести процедуру ее индивидуализации, простым языком запатентовать. Так как понятие плагиат в программном мире не существует, а схожесть показывает лишь оптимальность примененного метода или подхода в программировании.

Авторское право на программный продукт начинает действовать в момент создания самого продукта, осуществлять какую либо регистрацию для признания авторского права не требуется. Правообладатель написанного продукта имеет право, но не обязан использовать знак охраны авторского права, так называемый знак копирайта ©, имя и год выпуска программы. При отсутствии знаков охраны, автором считается тот, кто указан в качестве

последнего на оригинале программного продукта или его экземпляре. Так же следует помнить, что авторское право действует в течение всей жизни автора + 50 лет после его смерти.

Требования к эргономическим параметрам оборудования на рабочих местах с ПЭВМ представлены в следующих нормативных документах:

- ГОСТ Р 50923-96 Дисплей. Рабочее место оператора. Общие эргономические требования и требования к производственной среде.
- СанПиН 2.2.2/2.4.1340-03 Гигиенические требования к персональным электронно-вычислительным машинам и организации.

В соответствии с государственными стандартами и правовыми нормами обеспечения безопасности предусмотрена рациональная организация труда в течение смены, которая предусматривает:

- длительность рабочей смены не более 8 часов;
- установление двух регламентируемых перерывов (не менее 20 минут после 1-2 часов работы, не менее 30 минут после 2 часов работы);
- обеденный перерыв не менее 40 минут.

Обязательно предусмотрен предварительный медосмотр при приеме на работу и периодические медосмотры.

Каждый сотрудник должен пройти инструктаж по технике безопасности перед приемом на работу и в дальнейшем, должен быть пройден инструктаж по электробезопасности и охране труда.

Производственная безопасность

Согласно ГОСТ 12.0.003-2015, неблагоприятные производственные факторы по результирующему воздействию на организм работающего человека подразделяют на:

- на вредные производственные факторы, то есть факторы, приводящие к заболеванию, в том числе усугубляющие уже имеющиеся заболевания;
- опасные производственные факторы, то есть факторы, приводящие к травме, в том числе смертельной.

В соответствии с заданием заполнена информация по вредным и опасным факторам. Возможные опасные и вредные факторы приведены в таблице 1.

Таблица 1 - Возможные опасные и вредные факторы

Факторы (ГОСТ12.0.003-2015)	Этапы работ			Нормативные Документы
	Разработка	Изготовление	Эксплуатация	
1.Отклонение показателей микроклимата	+	+	+	- требования к значению напряженности устанавливаются СанПиН 2.2.2/2.4.1340-03 Гигиенические
2. Превышение уровня шума				требования к персональным электронно-вычислительным машинам и
3.Отсутствие или недостаток естественного света		+	+	организации работы; - требования к освещению устанавливаются СП 52.13330.2016
4.Недостаточная освещенность рабочей зоны	+	+	+	Естественное и искусственное освещение. Актуализированная редакция СНиП 23-05-95*. Дата
5.Повышенное значение напряжения в электрической цепи, замыкание которой	+		+	введения: 2017-05-08. Статус: действующий; - требования к значению отклонений

может произойти через тело человека				микроклимата устанавливаются СанПиН 2.2.4.548-96 Гигиенические требования к микроклимату производственных помещений
-------------------------------------	--	--	--	---

Анализ опасных и вредных производственных факторов

В соответствии с последовательностью в таблице 1. описываются выявленные вредные и опасные факторы. Каждый вредный фактор рассматривается по следующему плану:

- источник возникновения фактора;
- воздействие фактора на организм человека;
- приведение допустимых норм с необходимой размерностью, а также при возможности проводится анализ на соответствие нормам;
- предлагаемые средства защиты (коллективные и индивидуальные) для минимизации воздействия фактора.

Микроклимат в помещениях, оборудованных компьютерами

В помещении, в котором установлен компьютер, должен поддерживаться определенный микроклимат – влажность, запыленность, температура воздуха и т. д.

Допустимые нормы:

В соответствии с СанПиН 2.2.2/2.4.1340-03 в производственных помещениях, в которых работа с использованием ПЭВМ является основной и связана с нервно-эмоциональным напряжением, должны обеспечиваться оптимальные параметры микроклимата в соответствии с действующими санитарно–эпидемиологическими нормативами микроклимата производственных помещений.

Исходя из СанПин 2.2.4.548-96 значения температуры, влажности и скорости движения воздуха устанавливаются для рабочей зоны производственных помещений в зависимости от категории тяжести выполняемой работы, величины избытков явного тепла, выделяемого в помещении, и периода года.

В соответствии с СанПин 2.2.4.548-96, оптимальные микроклиматы для помещений с компьютерами представлены в таблице 2.

Таблица 2 - Оптимальные микроклиматы для помещения с компьютерами

Период года	Температура а воздуха, С ⁰	Температура поверхности й, С ⁰	Относительна я влажность воздуха, %	Скорость движения воздуха, м/с
Холодный	21 - 23	20 - 24	40 - 60	0,1
Теплый	23 - 25	22 - 26	40 - 60	0,1

Также для рабочей смены, количеством в 8 часов, используется нормы предельно допустимых значений микроклимата.

В соответствии с СанПин 2.2.4.548-96, предельно допустимые микроклиматы для помещения с компьютерами представлены в таблице 3.

Таблица 3 - Предельно допустимые значения микроклимата для помещения с компьютерами

Период года	Температура воздуха, °С		Температура поверхностей, °С	Относительная влажность воздуха, %	Скорость движения воздуха, м/с	
	диапазон ниже оптимальных величин	диапазон выше оптимальных величин			для диапазона температур воздуха ниже оптимальных величин	для диапазона температур воздуха выше оптимальных величин
Холодный	19,0- 20,9	23,1 - 24,0	18,0 - 25,0	15 - 75	0,1	0,2
Теплый	20,0- 21,9	24,1 - 28,0	19,0 - 29,0	15 - 75	0,1	0,3

Средства защиты:

Мероприятия по защите человека от неблагоприятного воздействия микроклимата.

Для регламентации времени работы в пределах рабочей смены в условиях микроклимата с температурой воздуха на рабочем месте выше или ниже допустимых величин используется защита временем.

Защита временем – уменьшение вредного действия неблагоприятных факторов рабочей среды и трудового процесса на работников за счет снижения времени их действия:

- введение внутрисменных перерывов;
- сокращение рабочего дня;
- увеличение продолжительности отпуска;
- ограничение стажа работы в данных условиях.

Уровень шума в помещениях, оборудованных компьютерами

В производственных помещениях при выполнении основных или вспомогательных работ с использованием ПЭВМ уровни шума на рабочих местах не должны превышать предельно допустимых значений, установленных для данных видов работ в соответствии с действующими санитарно-эпидемиологическими нормативами.

Источник возникновения фактора:

- машины вычислительные электронные цифровые;
- машины вычислительные электронные цифровые персональные (включая портативные ЭВМ);
- принтеры.

Воздействие фактора на организм человека:

- заболеванию нервной системы;
- нарушениям слуха.

В соответствии с СанПин 2.2.2/2.4.1340-03, допустимые значения уровней звукового давления для помещения с компьютерами представлены в таблице 4.

Таблица 4 - Допустимые значения уровней звукового давления для помещения с компьютерами

Уровни звукового давления в октавных полосах со среднегеометрическими частотами, Гц									Уровни звука в и эквивалентные уровни звука дБ А
Предприятия, учреждения и организации									
31,5	63	125	250	500	1000	2000	4000	8000	50
86 дБ	71 дБ	61 дБ	54 дБ	49 дБ	45 дБ	42 дБ	40 дБ	38 дБ	

Измерение уровня звука рассчитывалось на расстоянии равным 50 см. от поверхности оборудования, на высоте равной, расположению источника звука.

Рабочие помещения с ЭВМ не должны граничить с помещениями, в коих уровень шума выше нормируемых значений, представленные в ГОСТ 12.1.003-2014[20].

Средства защиты:

При организации рабочего места следует принимать все необходимые меры по снижению шума, воздействующего на человека на рабочих местах, до значений, не превышающих допустимые:

- применением средств и методов коллективной защиты по ГОСТ 12.1.029-80;
- применением средств индивидуальной защиты по ГОСТ 12.4.051-87.

Методы и средства коллективной защиты в зависимости от способа реализации подразделяются на строительно-акустические, архитектурно-планировочные и организационно-технические и включают в себя:

- изменение направленности излучения шума;

- применение звукоизоляции.

Недостаток освещенности в помещениях, оборудованных компьютерами

Освещения на рабочем месте - особый вид освещения, служащий, для создания оптимальных условий трудовой деятельности работников. Основная задача такого освещения: обеспечить такую освещенность рабочего помещения, которая будет наилучшим образом соответствовать характеру выполняемой работы и получения максимальной продуктивности от рабочей силы.

Искусственное освещение используется в помещениях, для которых испытывается недостаток естественного освещения, может так же использоваться для освещения помещений в те часы суток, когда естественное освещение отсутствует. Освещенность измеряется в Лк – люксах.

Источник возникновения фактора:

- лампы накаливания;
- люминесцентные лампы;
- газоразрядные лампы;

Допустимые нормы:

В соответствии с СанПиН 2.2.2/2.4.1340-03, требование к искусственной освещенности на рабочих местах с компьютерами 300-500 лк.

Естественное освещение в помещениях, оборудованных компьютерами

Естественное освещение создается путем солнечного воздействия света через световые проемы в помещении. В повседневной жизни используются два вида источников освещения, это лампы накаливания и газоразрядные лампы.

Источник возникновения фактора:

- Единственным источником естественного освещения является: Солнце, оно же является самым благоприятным для человеческого

глаза, посредством содержания ультрафиолетовых лучей, которые положительно влияют на здоровье человека.

Допустимые нормы:

В соответствии с СанПиН 2.2.1/2.1.1.1278-03, коэффициент естественной освещённости (КЕО) - это параметр, характеризующий количество естественного света, поступающего в помещение. требования к естественному освещению общественных зданий представлены в таблице 7.

Таблица 6 - Требования к естественному освещению общественных зданий

Рабочая поверхность и плоскость нормирования КЕО и освещенности (Г - горизонтальная, В - вертикальная) и высота плоскости над полом, м	КЕО E_n , %	
	при верхнем или комбинированном освещении	при боковом освещении
Г-0,8	3,0	1,0

Средства защиты:

При естественном освещении следует применять средства солнцезащиты, снижающие перепады яркостей между естественным светом и свечением экрана ПК:

- пленки с металлизированным покрытием;
- регулируемые жалюзи с вертикальными панелями;
- светорассеивающие шторы;

Электромагнитные поля в помещениях, оборудованных компьютерами

Ионизирующее излучение – поток микрочастиц, способных ионизировать вещество. Компьютер является источником электростатического и электромагнитного поля, в котором электромагнитные поля контролируются в двух диапазонах: от 5 Гц до 2 кГц, от 2 до 400 кГц.

Источник возникновения фактора:

Компьютер имеет сразу два источника электромагнитного излучения – это монитор и системный блок.

Основным источником электромагнитных излучений от мониторов ПЭВМ (ПК) является трансформатор высокой частоты строчной развертки.

Влияние фактора на организм человека:

- нарушения центральной нервной системы;
- отрицательное влияние на иммунологическую реактивность организма;
- нарушаются процессы иммуногенеза;
- возникновение аутоиммунитета;
- может способствовать неспецифическому угнетению иммуногенеза.

Допустимые нормы:

В соответствии с СанПиН 2.2.4.3359-16, ПДУ постоянного магнитного поля, приведена в таблице 7.

Таблица 7 - ПДУ постоянного магнитного поля

Время воздействия за рабочий день, минуты	Общее		Локальное	
	ПДУ напряженности, кА/м	ПДУ магнитной индукции, мТл	ПДУ напряженности, кА/м	ПДУ магнитной индукции, мТл
61 - 480	8	10	12	15

Предельно допустимые уровни электромагнитного поля частотой 50

Гц.

Средства защиты:

- мероприятия по сертификации, для операторов ЭВМ;
- организационно-технические мероприятия;
- увеличение относительной влажности воздуха до 65-75 %;
- уменьшение генерации электростатических зарядов или их отвод с наэлектризованного материала, что достигается путем заземления металлических электропроводных элементов оборудования,

Обоснование мероприятий по снижению уровней воздействия опасных и вредных факторов на исследователя (работающего)

Минимальным требованиям к рабочим местам офисного персонала является соответствие их ряду основных норм:

1. на каждого сотрудника отводится минимум 4 кв.м площади без учета используемой техники, дополнительной мебели, проходов между рабочими местами.
2. рабочие места сотрудников, чья работа связана с повышенной концентрацией и высокими нагрузками на нервную систему, отделяются перегородками высотой 1,5-2 м.
3. соблюдение температурного режима при стандартном 8-часовом рабочем дне в диапазоне от 20 до 28°C в зависимости от сезона и интенсивности труда.

Обеспечение комфортного уровня освещенности рабочего места, включающей как естественное освещение (обязательное требование при работе с компьютерной техникой), так и искусственное.

В соответствии с ГОСТ 12.2.032-78, оптимальное рабочее место оператора ЭВМ представлено далее.

Работа с применением персональных ПЭВМ (ПК) сопряжена со значительными зрительными и нервно-психологическими нагрузками, что повышает требования к организации труда пользователей ПК.

Конструкция рабочей мебели должна обеспечивать возможность индивидуальной регулировки, соответственно росту работающего, и создавать удобную позу. Часто используемые предметы труда и органы управления должны находиться в оптимальной рабочей зоне.

Конструкция рабочего стола должна обеспечивать оптимальное размещение на рабочей поверхности используемого оборудования с учетом его количественных и конструктивных особенностей, а также характера выполняемой работы.

Высота рабочей поверхности стола должна регулироваться в пределах 680-800 мм, при отсутствии такой возможности его высота должна быть не менее 725 мм.

На поверхности рабочего стола для документов необходимо предусматривать размещение специальной подставки, расстояние которой от глаз должно быть аналогичным расстоянию от глаз до клавиатуры.

Модульными размерами рабочей поверхности стола, на основании которых должны рассчитываться конструктивные размеры, следует считать: ширину 800, 1000, 1200 и 1400 мм, глубину 800 и 1000 мм при нерегулируемой его высоте, равной 725 мм.

Под столешницей рабочего стола должно быть свободное пространство для ног с размерами по высоте не менее 600 мм, по ширине 500 мм, по глубине 650 мм.

Конструкция рабочего стула должна обеспечивать поддержание рациональной рабочей позы при работе, что позволит изменять позу для снижения статического напряжения мышц шейно-плечевой области и спины для предупреждения развития утомления.

Тип рабочего стула должен выбираться в зависимости от характера выполняемой работы.

Рабочий стул должен быть подъемно-поворотным и с регулируемым углом наклона сиденья и спинки, а также расстоянием спинки от переднего края сиденья. При этом регулировка каждого параметра должна быть независимой, легко осуществляемой и иметь надежную фиксацию.

Рабочее кресло должно иметь подлокотники. Ширина и глубина поверхности сиденья должна составлять не менее 400 мм. Высота опорной поверхности спинки должна быть не менее 300 мм, ширина - не менее 380 мм. Радиус ее кривизны в горизонтальной плоскости 400 мм. Угол наклона спинки должен изменяться в пределах 90-110° к плоскости сидения.

Материал покрытия рабочего кресла должен обеспечивать возможность легкой очистки от загрязнений. Поверхности сидения и спинки

должны быть полумягкими, с нескользящим, неэлектризующим и воздухопроницаемым покрытием.

На рабочем месте необходимо оборудовать подставку для ног. Ее длина должна составлять 400 мм, ширина 350 мм. Необходимо предусматривать регулировку высоты подставки в пределах до 150 мм и угла ее наклона до 20 градусов. Поверхность подставки должна быть рифленой и иметь по переднему краю бортик высотой 10 мм.

При организации рабочих мест для работы на ПЭВМ необходимо предусматривать:

- пространство по глубине не менее 850 мм с учетом выступающих частей оборудования для нахождения человека-оператора;
- пространство для ног глубиной и высотой не менее 150 мм и шириной не менее 530 мм;
- расположение устройств ввода-вывода информации, обеспечивающее оптимальную видимость экрана;
- легкую досягаемость органов ручного управления в зоне моторного поля: по высоте - 900-1300 мм, по глубине - 400-500 мм;
- расположение экрана ПЭВМ в месте рабочей зоны, обеспечивающее устройство зрительного наблюдения в вертикальной плоскости под углом +30 градусов от нормальной линии взгляда оператора, а также устройство использования ПЭВМ (ввод-вывод информации при корректировке основных параметров технологического процесса, отладка программ и др.), одновременно с выполнением основных производственных операций (наблюдение за зоной обработки на станке с программным управлением и др.);
- возможность поворота экрана вокруг горизонтальной и вертикальной осей.

Клавиатуру следует располагать на поверхности стола на расстоянии 100-300 мм от края, обращенного к пользователю, или на специальной

регулируемой по высоте рабочей поверхности, отделенной от основной столешницы.

Экологическая безопасность

В общем случае под охраной окружающей среды характеризуется различного рода мероприятиями влияющие на следующие природные зоны:

- атмосфера;
- гидросфера;
- литосфера.

Разработанный проект особо сильно не влияет на окружение, размеры санитарно-защитных зон используется для предприятия пятого класса - 50 м.

Источники загрязнений для гидросферы не выявлены и поэтому методы защиты не разработаны.

Источники загрязнения литосферы: основным источником загрязнения является бытовой мусор.

Методы защиты от выбросов в литосферу: применять мероприятия по устранению мусора в производственных помещениях.

Источники загрязнения атмосферы: при изготовлении ЖК-дисплеев, происходит выброс (CO₂) NF₃.

Методы защиты от выбросов в атмосферу: использовать замену только расходных материалов у компьютеров, а не их целиком.

Безопасность в чрезвычайных ситуациях

Чрезвычайная ситуация - состояние, при котором возникает угроза жизни и здоровья человека также наносящее ущерб имуществу и природной среде.

Наиболее типичной ЧС для помещений с оборудованием компьютеров является пожар. Пожар может возникнуть вследствие причин: замыкание, искрение, неосторожное обращение с огнём, курение.

Разработка превентивных мер по предупреждению ЧС

Пожарная безопасность – организационные и/или технические мероприятия направленные на обеспечение безопасности людей.

Для работы за устройствами нужно знать и выполнять определённый список правил поведения:

- Ни в коем случае не курить, выходить только на улицу;
- Не бросайте непотушенные сигареты на землю, либо сразу затапывайте их ногой;
- Все неисправные электроприборы должны быть сданы в ремонт, исправные периодически подвергаться профилактическому обслуживанию;
- При перепадах напряжения в сети, обязательно обратитесь в жилищно-эксплуатационную службу, чтобы выяснить и устранить причину их возникновения;
- При уходе из квартиры, не оставляйте электроприборы, компьютеры и другую технику в режиме ожидания: стоит выключить их из сети.

Разработка действий в случае возникновения ЧС

В случае возникновения пожара, необходимо, предпринять меры по эвакуации персонала из помещения в соответствии с планом эвакуации здания, расположенный на каждом этаже здания. При отсутствии прямых угроз здоровью и жизни, необходимо произвести тушение огнетушителем. При потере контроля и/или стремительного разрастания пожара, необходимо эвакуироваться и ждать пожарную службу. Как только происходит возникновение пожара происходит запуск системы пожаротушения, издав и передает сигнал о пожаре на пункт пожарной станции сигнал, в случае если система не была установлена или еще по каким-либо причинам, система не сработала, необходимо самостоятельно вызывать пожарную службу по телефону 01, либо с мобильного 101.

Вывод

В ходе изучения и составления раздела «социальная ответственность» были получены навыки по оптимизации и улучшению рабочей среды операторов ЭВМ, выработаны и применены допустимые нормы различных факторов, что являются неотъемлемой частью разрабатываемого проекта.

Также рассмотрены ЧС, которые могут возникать при внедрении проекта. Выработаны действия при возникновении ЧС на территории, использующей проект, что позволит быстро реагировать в случае ЧС.

ЗАКЛЮЧЕНИЕ

В ходе выполнения данной выпускной квалификационной работы для кластеризации аудиофайлов по спектральным характеристикам, были рассмотрены и проанализированы методы кластеризации, применимые к алгоритмам кластеризации различных типов. В качестве исходных данных были взяты аудиофайлы, собранные при помощи мультисенсорной электронной системы мониторинга ульев BeePi.

В результате проведенного исследования было выяснено, что, делая выборку по переменным 1-40, программа не выдает точного деления по кластерам, так же как переменные 54-181 и 189-193. График показал, что по переменным 182-187 хорошо делятся данные на 3 кластера. Так же стало понятно, что при взятии по 200 переменных всего из 2 категорий: жужжание пчел и стрекот сверчков, программа выдает четкое деление данных на 2 кластера. Это и позволит отделить жужжание пчел от стрекота сверчков.

СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ

1. ELKAN, C. Using the triangle inequality to accelerate kmeans // ICML. — 2003. — Vol. 3. — P. 147–153.
2. НЕЙСКИЙ, И. М. Классификация и сравнение методов кластеризации // Интеллектуальные технологии и системы. Сборник учебно-методических работ и статей аспирантов и студентов. — М.: НОК «CLAIM», 2006. — Выпуск 8. — С. 130–142.
3. <http://statsoft.ru/products/overview/>
4. http://journals.tsu.ru/informatics/&journal_page=archive&id=1773&article_id=39530
5. https://ami.nstu.ru/~vms/lecture/data_mining/kurs_klaster.htm
6. Г.Е. Козлов. Реферат по истории и философии науки на тему: история развития кластерного анализа. [Электронный ресурс]. Систем. требования: Adobe Acrobat Reader. – URL:
7. <http://statistica.ru/theory/klasterizatsiya-metod-k-srednikh/>
8. <https://basegroup.ru/community/articles/em>
9. <https://habr.com/ru/post/140828/>
10. https://en.wikipedia.org/wiki/Short-time_Fourier_transfor

