

**ПРОЕКТИРОВАНИЕ СИСТЕМ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА.
РАЗРАБОТКА СИСТЕМЫ РАСПОЗНАВАНИЯ ОБРАЗОВ**

МОСКВА 2022

СОДЕРЖАНИЕ

ВВЕДЕНИЕ.....	4
1 РАСЧЕТЫ И ОБОСНОВАНИЕ ВЫБОРА ИНФОРМАТИВНЫХ ПЕРВИЧНЫХ ПРИЗНАКОВ, ОПИСЫВАЮЩИХ РАСПОЗНАВАЕМЫЕ ОБЪЕКТЫ.....	6
1.1 Информативность признаков по Фишеру.....	6
1.2 Метод корреляционных плеяд8	
1.3 Демонстрация оценок информативности каждого признака и графа взаимных связей признаков11	
2 ПОСТРОЕНИЕ ПРАВИЛ КЛАССИФИКАЦИИ.....	12
2.1 Метод дискриминантных функций12	
2.2 Метод К ближайших соседей14	
2.3 Результаты классификации объектов15	
3 РАЗРАБОТКА ИНТЕРАКТИВНОЙ ПРОГРАММЫ РАСПОЗНАВАНИЯ ОБРАЗОВ НА ЯЗЫКЕ C++.....	18
3.1 Обобщенная блок-схема программы19	
3.2 Детальная блок-схема процедур распознавания.....20	
3.3 Анализ результатов самостоятельного тестирования.....22	
3.4 Протокол самостоятельного тестирования.....22	
ЗАКЛЮЧЕНИЕ.....	22
СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ.....	24
ПРИЛОЖЕНИЕ А.....	25
ПРИЛОЖЕНИЕ Б.....	27
ПРИЛОЖЕНИЕ В.....	30
ПРИЛОЖЕНИЕ Г.....	36

ВВЕДЕНИЕ

Разработка эффективных систем распознавания объектов, явлений, ситуаций и процессов в настоящее время представляет собой задачу исключительной значимости. Наиболее важными направлениями применения данной системы являются:

- распознавание символов печатного и рукописного текстов;
- распознавание изображений, полученных в различных частотных диапазонах (оптическом, инфракрасном, радиочастотном, звуковом);
- распознавание речи;
- медицинская диагностика;
- системы безопасности;
- классификация и поиск в базах данных и знаний.

В настоящее время основное внимание уделяется разработке и построению больших систем распознавания, в которых описание объекта может достигать многих десятков и даже сотен признаков, при этом признаки могут иметь различную физическую природу. Например, при построении систем распознавания автомобилей могут быть использованы такие признаки, как марка машины, ее масса, тип двигателя, максимальная скорость, назначение (пассажирский, специальный, грузовой) и т. п.

В качестве предмета проводимых исследований выступает система распознавания объектов на основе их признакового описания в соответствии с заданными деревьями дихотомии.

Целью данной работы является создание алгоритма классификации образов и определения их принадлежности к одному из перечисленных классов. Для демонстрации полученных результатов была разработано интерактивное приложение.

В ходе работы над проектом были решены следующие задачи:

- проанализировать информативность первичных признаков, описывающих распознаваемые объекты;

- сгруппировать признаки с применением метода корреляционных плеяд;
- построить классификационные функции для узлов деревьев принятия решений;
- реализовать метод k-ближайших соседей;
- разработать интерактивную программу распознавания объектов на языке C++.

В основной части работы представлен сравнительный анализ выбора признакового пространства для каждого узла дерева дихотомии, реализованы основные методы классификации, особое внимание уделено их синтезу для достижения наилучшего эффекта.

1 РАСЧЕТЫ И ОБОСНОВАНИЕ ВЫБОРА ИНФОРМАТИВНЫХ ПЕРВИЧНЫХ ПРИЗНАКОВ, ОПИСЫВАЮЩИХ РАСПОЗНАВАЕМЫЕ ОБЪЕКТЫ

1.1 Информативность признаков по Фишеру

Одним из методов сокращения числа признаков является выделение наиболее информативных параметров. Для принятия того или иного решения оцениваются не все признаки объекта, а наиболее значимые с точки зрения их разделяющей способности. Тогда классификация строится для оценки только ограниченного числа информативных параметров, что упрощает его реализацию.

В такой постановке задачи возникает проблема выбора наиболее информативных параметров и, следовательно, исключения из рассмотрения параметров которые не существенны.

Информативность признака x по Фишеру для классов A и B вычисляется по формуле

$$I_x^{A/B} = \frac{(m_x^A - m_x^B)^2}{(\sigma_x^A)^2 + (\sigma_x^B)^2} \quad (1)$$

где m_x^A - среднее значение признака x у объектов класса A ;

m_x^B - среднее значение признака x у объектов класса B ;

σ_x^A - среднее квадратическое отклонение признака x у объектов класса A ;

σ_x^B - среднее квадратическое отклонение признака x у объектов класса B .

Основой для классификации является дерево дихотомии, поэтому все расчеты величины информативности ведутся согласно варианта дерева из задания.

На рисунке 1 и рисунке 2 показаны два дерева последовательной дихотомии, где (A, B, C), (A, B), (A, C), (B, D) — синтетические классы-объединения.

Схемы расчета матрицы информативностей свойств по Фишеру для дерева слева и справа отличны, в частности:

- Дерево дихотомии 1:

1) $I_x^{(D)/(ABC)}$;

2) $I_x^{(C)/(AB)}$;

3) $I_x^{A/B}$.

- Дерево дихотомии 2:

1) $I_x^{(A,C)/(B,D)}$;

2) $I_x^{A/C}$;

3) $I_x^{B/D}$.

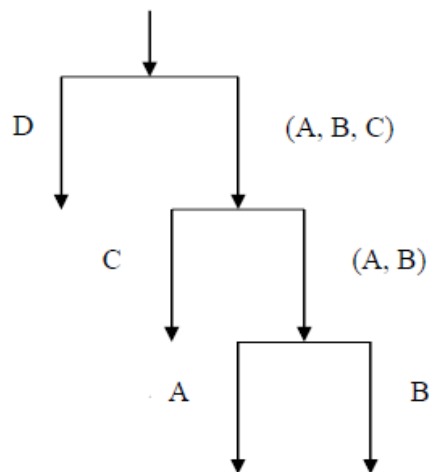


Рисунок 1 – Дерево дихотомии 1

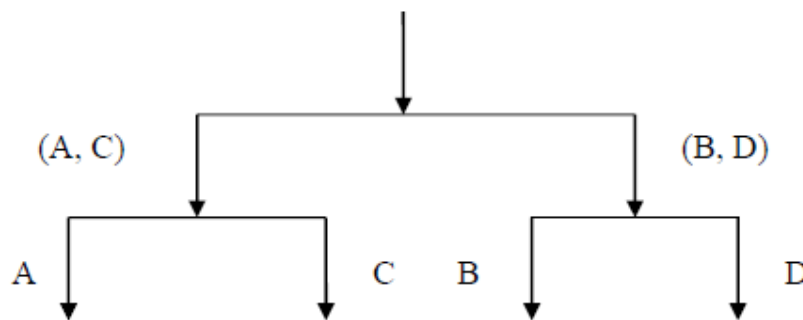


Рисунок 2 – Дерево дихотомии 2

1.2 Метод корреляционных плеяд

Метод корреляционных плеяд позволяет выделить связанные признаки по значению коэффициентов линейной корреляции между признаками классов объектов. В основе метода лежат вычисления матриц взаимной (парной) корреляции, т.е. меры линейной связи, признаков.

Метод состоит из двух этапов:

- Построение графа взаимных связей признаков (по заданной матрице корреляционных взаимосвязей признаков).

- Разбиение признаков на группы (по построенному графу взаимных связей признаков и заданному граничному значению для определения сильно связанных признаков).

Алгоритм построения графа взаимных связей признаков:

- 1) Находим в матрице корреляционных взаимосвязей признаков размерностью $n \times n$ максимальный недиагональный элемент.

- 2) Исходя из индексов найденного элемента, начинаем строить граф: индексы элемента определяют номера вершин графа, а значение найденного элемента значение, которое наносят над ребром графа, которое соединяет эти вершины.

3) Рассматриваем две строки матрицы, номера которых соответствуют номерам двух последних рассмотренных вершин графа. В этих строках находим максимальный недиагональный элемент, не рассмотренный ранее.

4) В матрице вычеркиваем строку и столбец с номерами, соответствующими номеру одной из двух последних рассмотренных строк графа, в которой не нашли максимальный недиагональный элемент не рассмотренный ранее.

5) Исходя из индексов найденного элемента строим очередную часть графа: индексы элемента определяют номера вершин графа, а значение найденного элемента значение, которое наносят над ребром графа, которое соединяет эти вершины.

6) Если количество вершин графа равно N, граф построен. Если нет - переходим к п.3.

Разбиение признаков на группы:

- после того как граф взаимных связей признаков построен, разбиваем признаки на группы: “разрываем” ребра графа, над которыми указаны значения, меньшие заданного граничного значения и смотрим, какие вершины графа остались связанными;

- для определения величины коэффициента взаимосвязи, были использованы эмпирическое значение величины коэффициента сильной связи $R = 0,75$.

В основе метода лежат вычисления матриц взаимной корреляции. Построение таких матриц также зависит от деревьев дихотомии (рисунок 1 и рисунок 2) и может определяться двумя соображениями:

- оценка взаимной связности всех объектов в узле **до** распределения;
- оценка взаимной связности всех объектов в узле **после** распределения.

Оценка взаимной связности всех объектов в узле **до** распределения для дерева дихотомии 1:

1) $R_{xy}(DABC)$;

2) $R_{xy}(ABC)$;

3) $R_{xy}(AB)$.

Оценка взаимной связности всех объектов в узле **после** распределения для дерева дихотомии 1:

1) $R_{xy}(D)$ или $R_{xy}(ABC)$;

2) $R_{xy}(C)$ или $R_{xy}(AB)$;

3) $R_{xy}(A)$ или $R_{xy}(B)$.

Оценка взаимной связности всех объектов в узле **до** распределения для дерева дихотомии 2:

1) $R_{xy}(ACBD)$;

2) $R_{xy}(AC)$;

3) $R_{xy}(BD)$.

Оценка взаимной связности всех объектов в узле **после** распределения для дерева дихотомии 2:

1) $R_{xy}(AC)$ или $R_{xy}(BD)$;

2) $R_{xy}(A)$ или $R_{xy}(C)$;

3) $R_{xy}(B)$ или $R_{xy}(D)$.

По результатам применения метода корреляционных плеяд для каждого узла строится наборы информативных некоррелированных признаков. С учетом того, что классификация имеет линейный характер, в общем случае можно построить разные сочетания пар признаков.

В итоге может оказаться, что для классификации в разных узлах могут использоваться разные наборы признаков.

1.3 Демонстрация оценок информативности каждого признака и графа взаимных связей признаков

Таким образом, применив метод корреляционных плеяд, получаем разбиение признаков на группы, причем признаки, попадающие в одну группу, признаются сильно связанными, а признаки, попадающие в разные группы - слабо связанными. Далее, выделив наиболее информативные признаки, оставляем от каждой группы только один - самый информативный.

Набор отобранных таким образом наиболее информативных групповых признаков является основой для построения признакового пространства для решения задачи классификации.

Расчеты и обоснование выбора информативных первичных признаков:

- оценки информативности признаков для каждого узла двух деревьев дихотомии представлены в приложении А;
- графы взаимных связей признаков представлены в приложении Б;

2 ПОСТРОЕНИЕ ПРАВИЛ КЛАССИФИКАЦИИ

Основой процедуры является дерево последовательной дихотомии, где в каждом узле дерева выполняется классификация вида «объект класса А» - «не объект класса А». Одной из задач такого распознавания является измерение расстояния между объектом и классом. Причем в случае классификации с обучением нам известны только элементы обучающей выборки с метками принадлежности их тем или иным классам. Можно выделить несколько методов решения этой задачи:

- классификация с применением линейной дискриминантной функции;
- классификация с применением метода ближайшего соседа;
- классификация с применением метода 3-х ближайших соседей.

Для подсчета расстояния между объектами класса используется метрика Евклида:

$$d_2(x,y)=\|x-y\|_2=\sqrt{(x_1-y_1)^2 + \dots + (x_n-y_n)^2} \quad (2)$$

Конечным результатом классификации является результат мажоритарного голосования среди результатов, полученных от каждого метода, т.е. условие «объект класса Q» выполняется при результатах «2-из 3» и «3-из-3», поскольку логика дерева дихотомии подразумевает выбор двух альтернативных вариантов.

2.1 Метод дискриминантных функций

Данный метод заключается в нахождении функции, имеющей специальный вид:

$$F=(m_x^A-m_x^B)x^T-\frac{1}{2}(m_x^A-m_x^B)(m_x^A+m_x^B)^T \quad (3)$$

где F - дискриминантная функция;

m_x^A - вектор признаков, описывающий центр класса A ;

m_x^B - вектор признаков, описывающий центр класса B ;

x^T - вектор признаков, описывающий неопознанный объект.

Тогда метод распознавания объектов класса будет происходить следующим образом:

- Если $F(x) > 0$, то объект класса A ;

- Если $F(x) < 0$, то объект класса B .

Построение линейной дискриминантной функции:

Пусть

$$m_x^A=(m_1^A, m_2^A) \quad (4)$$

$$m_x^B=(m_1^B, m_2^B) \quad (5)$$

$$x^T=\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \quad (6)$$

Введем обозначения

$$c_1=\frac{(m_1^{A^2}-m_1^{B^2}+m_2^{A^2}-m_2^{B^2})}{2(m_1^A-m_1^B)} \quad (7)$$

$$c_2=\frac{m_2^A-m_2^B}{m_1^A-m_1^B} \quad (8)$$

$$c=m_1^A-m_1^B \quad (9)$$

Получаем

$$\frac{F}{c} = x_1 + c_2 x_2 - c_1 \quad (10)$$

Для распознавания объектов необходимо приравнять функцию к нулю. Таким образом условие классификации объектов показаны в формулах

$$x_1 + c_2 x_2 - c_1 > 0 \rightarrow \text{объект класса А} \quad (11)$$

$$x_1 + c_2 x_2 - c_1 < 0 \rightarrow \text{объект класса В} \quad (12)$$

На основании полученных средних значений признаков (приложение А) было построено 9 дискриминантных функций, из которых 6 относятся к первому дереву дихотомии, а 3 – ко второму. Это обусловлено тем, что наборы наиболее информативных признаков, посчитанные двумя способами, для второго дерева дихотомии не отличаются. Результаты вычисления дискриминантных функций представлены в приложении В.

2.2 Метод К ближайших соседей

Формально основой метода является гипотеза компактности: если метрика расстояния между примерами введена удачно, то схожие объекты вероятнее всего будут лежать в одном классе, чем в разных. Тогда для классификации объект присваивается тому классу, который является наиболее распространённым среди k соседей данного элемента, классы которых уже известны.

Рассмотри метод более подробно (рисунки 2.1 и 2.2), требуется провести классификацию объекта (красный квадрат) на основе дерева последовательной дихотомии. Объекты разделяются на классы А и В (зеленые и синие круги соответственно). Рассмотрим метод К ближайших соседей:

- в случае $k=1$ отыскивается ближайший объект и классифицируется относительно него. На рисунке 2.2 синий круг лежит ближе к красному квадрату, чем остальные, поэтому объекту присваивается класс В;

- в более общем случае находятся k ближайших объектов и классификация выполняется мажоритарным голосованием. На рисунке 2.1 красный квадрат будет распознан как класс А, так как внутри круга, состоящем из трех ближайших элементов, два зеленых круга в противовес одному синему.

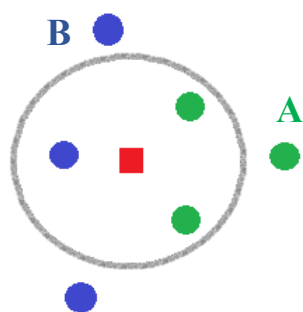


Рисунок 2.1 – Классификация с $k=3$ соседями

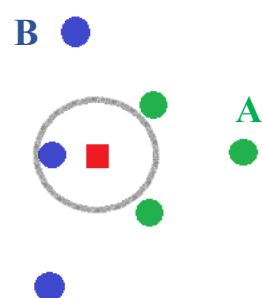


Рисунок 2.2 – Классификация с $k=1$ соседом

2.3 Результаты классификации объектов

Результаты классификации объектов из обучающей и тестовой выборках а также графики классифицированных объектов в выбранном двумерном пространстве признаков согласно построенным правилам классификации для каждого узла дерева дихотомии представлены в приложении В. На основании полученных данных можно сделать вывод, что второе дерево дихотомии показывает лучший результат, поэтому именно оно было взято за основу при реализации интерактивного приложения. Также экспериментальным путем были выявлены наиболее эффективные наборы признаков, при которых классификация имеет лучшее качество:

- признаки 1 и 4 для первого узла дерева;
- признаки 4 и 5 для второго узла дерева;

- признаки 5 и 9 для третьего узла дерева.

Графики классифицированных объектов, полученные с помощью наиболее эффективных параметров распознавания представлены на рисунках 2.3-2.5.

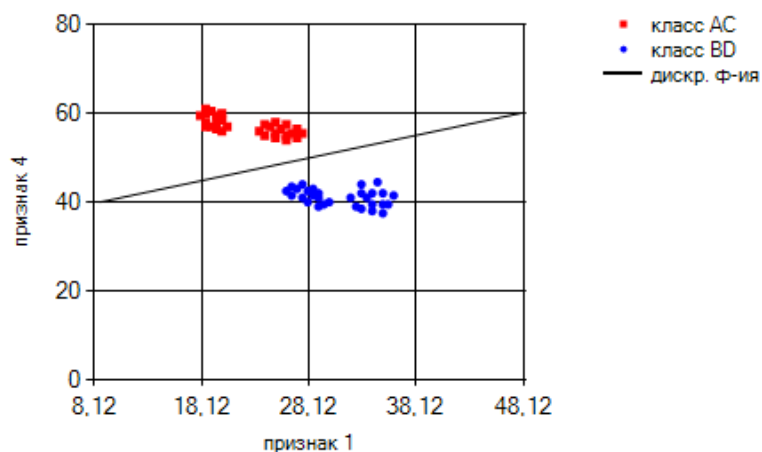


Рисунок 2.3 – Классификация объектов в 1 узле

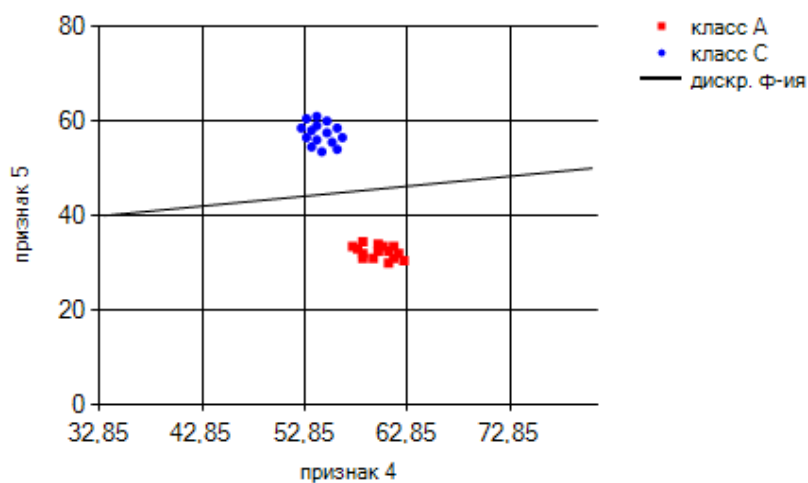


Рисунок 2.4 – Классификация объектов в 2 узле

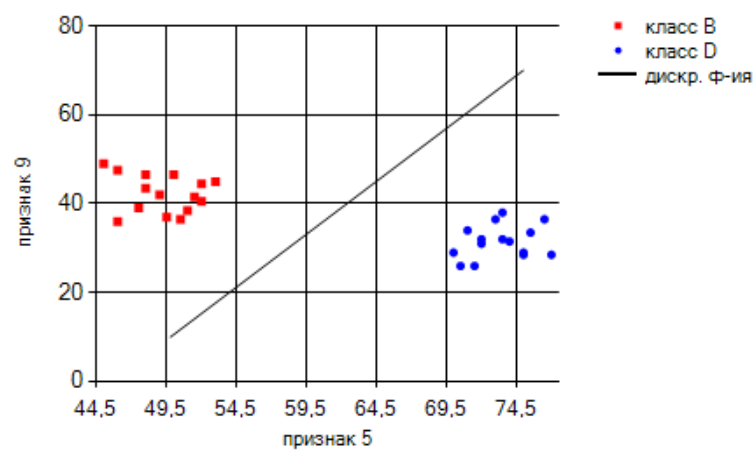


Рисунок 2.5 – Классификация объектов в 3 узле

3 РАЗРАБОТКА ИНТЕРАКТИВНОЙ ПРОГРАММЫ РАСПОЗНАВАНИЯ ОБРАЗОВ НА ЯЗЫКЕ C++

На основе результатов, описанных выше, было разработано интерактивное приложение на языке C++, реализующее алгоритм распознавания объектов класса A, B, C и D. Решение о принадлежности распознаваемого объекта классу в соответствии с заданным деревом принятия решений принимается голосованием в каждом узле дерева результатов классификации тремя методами: методом линейных дискриминантных функций, а также методом одного и трех ближайших соседей.

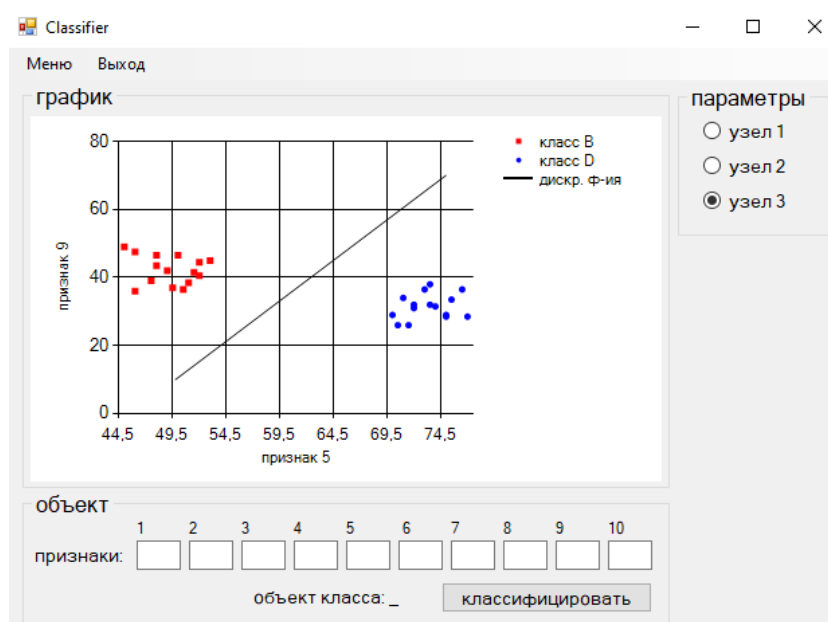


Рисунок 3.1 – Окно приложения

Окно приложения условно можно разделить на 3 части: область построения графиков, раздел параметров и область ввода данных. К основным функциям приложения можно отнести выбор узла дерева, т.е. рассмотрение классификации на разных этапах, само распределение объекта и реализация меню инструментов.

3.1 Обобщенная блок-схема программы

Интерактивное приложение предназначено для демонстрации результатов выполнения классификации, каждое событие обрабатывается своей процедурой, которая, в свою очередь, обновляет область построения графика.

Главную роль в приложении занимает класс Classifier, который содержит в себе основные наборы данных для классификации. Именно он отвечает за распознавание объектов на основании 3 методов, рассмотренных ранее. Блок-схема реализации такого класса представлена на рисунке 3.2.

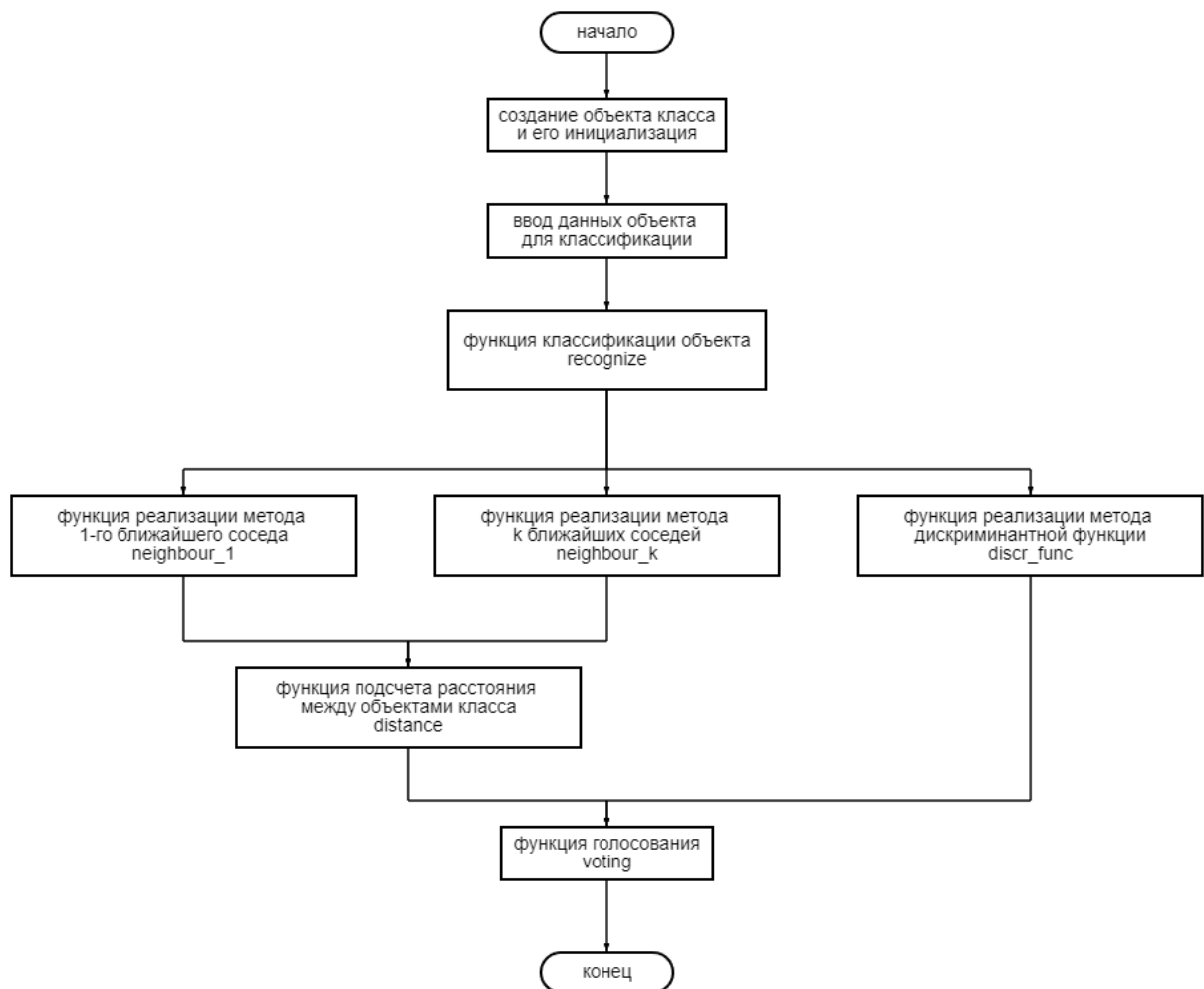


Рисунок 3.2 – Общая схема алгоритма классификации

3.2 Детальная блок-схема процедур распознавания

После ввода данных объекта, его признаки передаются на вход функции распознавания recognize. Она, в свою очередь, осуществляет обход дерева дихотомии и вызов функции мажоритарного голосования voting от каждого узла этого дерева. Функция голосования основана на применении трех методов классификации: метод дискриминантной функции discr_func, метод одного ближайшего соседа neighbor_1, метод трех ближайших соседей neighbor_k.

Рассмотрим более подробно реализацию алгоритма каждого метода классификации:

- Метод дискриминантной функции (рисунок 3.3). Данный алгоритм основан на применении посчитанной ранее дискриминантной функции. Сама классификация происходит посредством выяснения взаимного положения точки, характеризующей объект, и прямой в пространстве выбранных признаков.

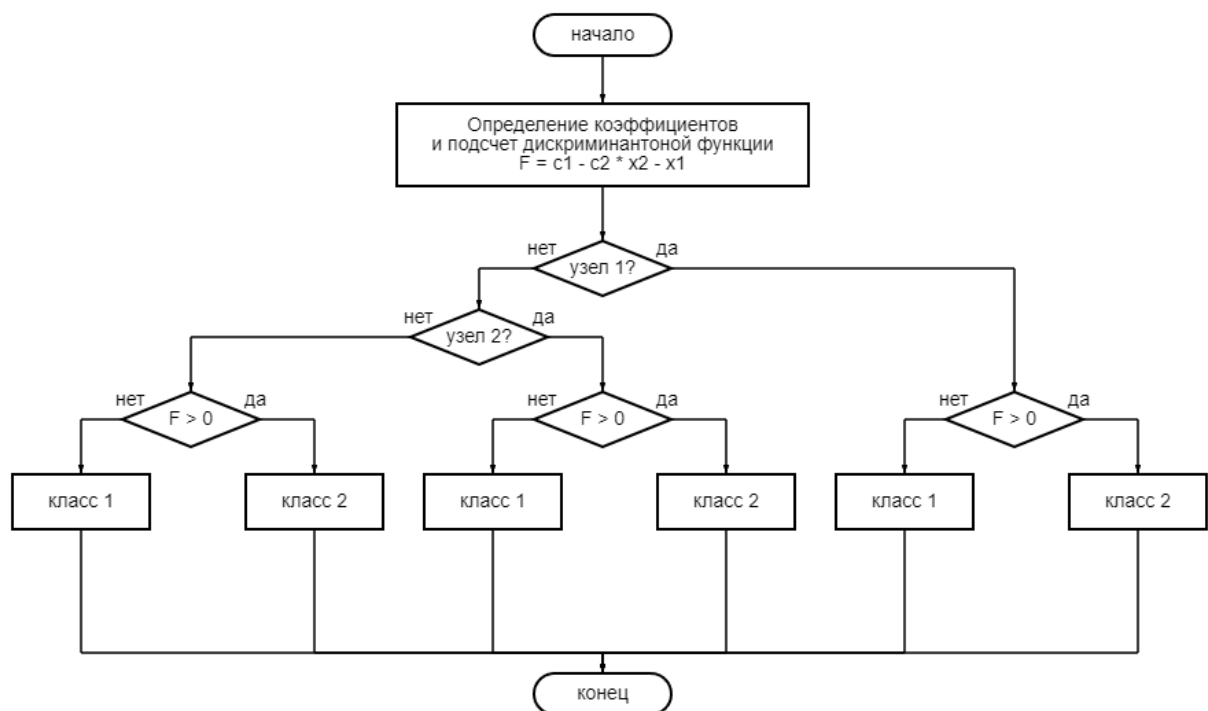


Рисунок 3.3 – Алгоритм, реализующий метод дискриминантной функции

- Алгоритм одного ближайшего соседа (приложение Г):

- 1) создаем переменную $\min1$ для 1-го класса и помещаем в нее расстояние от объекта до класса;
- 2) пробегаем по объектам всего класса и ищем минимальный элемент;
- 3) повторяем п. 1-2 для 2-го класса с переменной $\min2$;
- 4) на основании сравнения полученных минимумов принимается решение о принадлежности тому или иному классу.

- Алгоритм K ($K>1$) ближайших соседей (приложение Г):

- 1) создаем массив минимумов из k элементов для 1-го класса и осуществляем начальную инициализацию, т.е. помещаем в него расстояние от объекта до класса;
- 2) сортируем полученный массив;
- 3) пробегаем по объектам всего класса и изменяем массив, в случае нахождения меньшего значения расстояния, сдвигаем остальные элементы. В итоге получаем массив минимальных расстояний до объекта класса 1;
- 4) повторяем п. 1-3 для 2-го класса;
- 5) решение о принадлежности определенному классу принимается посредством мажоритарного голосования.

Данный подход, применяемый к классификации объектов, показывает хорошее качество распределения и является устойчивым к ошибкам одного из методов, что осуществляется функцией мажоритарного голосования.

3.3 Анализ результатов самостоятельного тестирования объектов из заданной обучающей выборки

Для самостоятельной проверки правильности созданной обучающей выборки и выведенных формул я провел расчет, взяв три точки разных классов и соотнесли их, это показано в приложении В. Эти точки были взяты вблизи средних значений признаков соответствующих классов. В ходе тестирования ошибок не было выявлено.

3.4 Протокол самостоятельного тестирования разработанной программы распознавания образов

Протокол самостоятельного тестирования разработанной программы распознавания образов показан в Приложении В. В ходе проверки векторов ошибок не было обнаружено.

ЗАКЛЮЧЕНИЕ

Системы распознавания в настоящее время получают все большее распространение, и трудно назвать такую сферу деятельности, где бы не использовалась данная технология. Проектирование такой систем является достаточно сложным процессом, на первом этапе которого происходит первоначальное разбиение объектов на классы в заданном признаковом пространстве. Такое описание позволяет найти в некотором смысле наилучшие границы разделения классов, а после выбора алгоритма классификации, решить задачу поиска рабочего, т.е. наиболее эффективного, пространства признаков и множества классов. Попутно необходимо решать задачи оптимизации процесса распознавания, оценки эффективности

системы в различных режимах ее функционирования и т. д. Именно такой комплексный подход позволяет построить реальную систему классификации.

В соответствии с поставленными целями и задачами по созданию системы распознавания объектов на основании их признакового описания, в ходе проделанной работы был разработан алгоритм, наиболее эффективно классифицирующий элементы для заданных деревьев дихотомии, а также разработано интерактивное приложение для демонстрации полученных результатов.

При оценке разработанной системы стоит отметить детально проработанный подход к классификации объектов, учитывающий самые разные нюансы, связанные с выбором признакового пространства. Также особого внимания заслуживает устойчивость к “плохим” исходным данным, что положительно сказывается на работе системы.

СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ

- 1 Горелик А.Л., Скрипкин В.А. Методы распознавания. Учебное пособие. М.: Высшая школа, 1984. – 208 с.
- 2 Чабан Л.Н. Теория и алгоритмы распознавания образов. Учебное пособие. М.: МИИГАиК. – 2004. – 70 с.
- 3 Колесникова С.И. Методы анализа информативности разнотипных признаков. Томск. – 2009. – 12 с.
- 4 Студми. Учебные материалы для студентов в 2013 г.: математические методы в психологии [Электронный ресурс]. – 2013. –
URL: https://studme.org/238029/matematika_himiya_fizik/kriteriy_fishera (дата обращения 27.11.2022)
- 5 Хабр в 2011 г.: Приближенное решение алгебраических и трансцендентных уравнений [Электронный ресурс]. – 2011. –
URL: <https://habr.com/ru/post/132366/> (дата обращения 27.11.2022)

ПРИЛОЖЕНИЕ

ПРИЛОЖЕНИЕ А

Таблица А.1 - Информативность признаков для узла 1 дерева дихотомии 1

D	ср. значение	37.00	14.73	30.00	40.27	73.67	15.33	15.00	16.07	31.93	30.67
	ср. кв. отклонение	2.25	2.79	5.27	4.02	4.19	2.57	2.63	3.66	7.20	5.69
ABC	ср. значение	23.53	15.16	29.60	52.36	46.58	15.22	14.80	17.58	35.29	29.64
	ср. кв. отклонение	5.60	2.76	5.60	8.00	11.19	2.76	2.80	4.96	9.87	5.53
информативность		4.98	0.01	0.00	1.82	5.14	0.00	0.00	0.06	0.08	0.02

Таблица А.2 - Информативность признаков для узла 2 дерева дихотомии 1

C	ср. значение	29.00	15.27	29.87	54.80	57.60	15.47	14.93	16.60	33.47	29.20
	ср. кв. отклонение	2.34	2.67	5.49	2.37	4.53	2.70	2.74	5.02	9.99	5.50
AB	ср. значение	20.80	15.10	29.47	51.13	41.07	15.10	14.73	18.07	36.20	29.87
	ср. кв. отклонение	4.68	2.80	5.66	9.42	9.30	2.77	2.83	4.86	9.68	5.54
информативность		2.46	0.00	0.00	0.14	2.55	0.01	0.00	0.04	0.04	0.01

Таблица А.3 - Информативность признаков для узла 3 дерева дихотомии 1

A	ср. значение	16.53	14.67	29.33	60.07	32.60	15.47	14.67	14.93	29.93	29.73
	ср. кв. отклонение	1.45	2.87	5.73	3.02	2.65	2.70	2.87	3.23	6.48	5.56
B	ср. значение	25.07	15.53	29.60	42.20	49.53	14.73	14.80	21.20	42.47	30.00
	ср. кв. отклонение	2.29	2.65	5.57	2.93	4.76	2.79	2.79	4.15	8.19	5.51
информативность		9.93	0.05	0.00	18.04	9.66	0.04	0.00	1.42	1.44	0.00

Таблица А.4 - Информативность признаков для узла 1 дерева дихотомии 2

AC	ср. значение	22.77	14.97	29.60	57.43	45.10	15.47	14.80	15.77	31.70	29.47
	ср. кв. отклонение	6.53	2.79	5.62	3.78	13.04	2.70	2.81	4.30	8.60	5.54
BD	ср. значение	31.03	15.13	29.80	41.23	61.60	15.03	14.90	18.63	37.20	30.33
	ср. кв. отклонение	6.38	2.75	5.42	3.65	12.87	2.70	2.71	4.68	9.34	5.61
информативность		0.82	0.00	0.00	9.50	0.81	0.01	0.00	0.20	0.19	0.01

Таблица А.5 - Информативность признаков для узла 2 дерева дихотомии 2

А	ср. значение	16.53	14.67	29.33	60.07	32.60	15.47	14.67	14.93	29.93	29.73
	ср. кв. отклонение	1.45	2.87	5.73	3.02	2.65	2.70	2.87	3.23	6.48	5.56
С	ср. значение	29.00	15.27	29.87	54.80	57.60	15.47	14.93	16.60	33.47	29.20
	ср. кв. отклонение	2.34	2.67	5.49	2.37	4.53	2.70	2.74	5.02	9.99	5.50
информативность		20.52	0.02	0.00	1.88	22.69	0.00	0.00	0.08	0.09	0.00

Таблица А.6 - Информативность признаков для узла 3 дерева дихотомии 2

В	ср. значение	25.07	15.53	29.60	42.20	49.53	14.73	14.80	21.20	42.47	30.00
	ср. кв. отклонение	2.29	2.65	5.57	2.93	4.76	2.79	2.79	4.15	8.19	5.51
D	ср. значение	37.00	14.73	30.00	40.27	73.67	15.33	15.00	16.07	31.93	30.67
	ср. кв. отклонение	2.25	2.79	5.27	4.02	4.19	2.57	2.63	3.66	7.20	5.69
информативность		13.81	0.04	0.00	0.15	14.49	0.02	0.00	0.86	0.93	0.01

ПРИЛОЖЕНИЕ Б

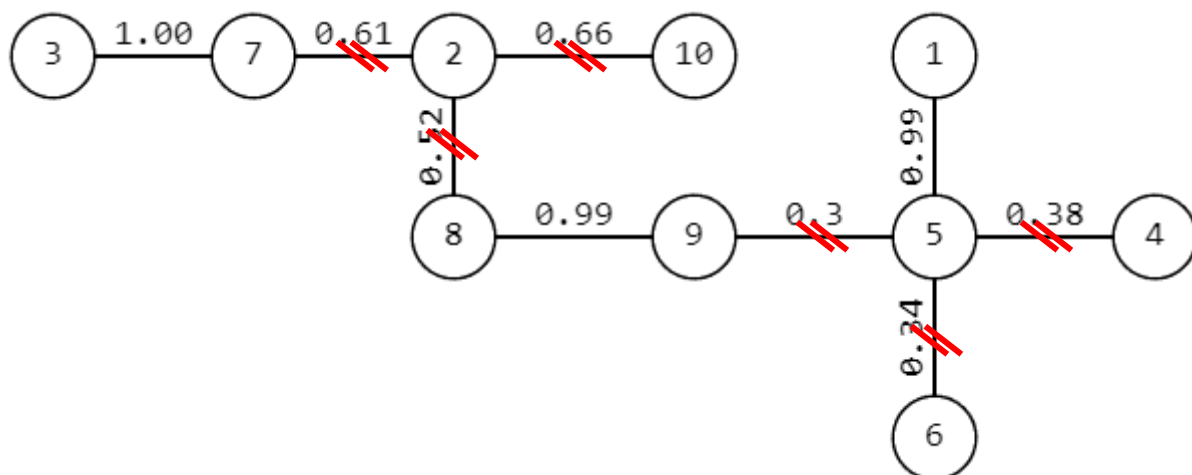


Рисунок Б.1 - Граф взаимных связей признака А

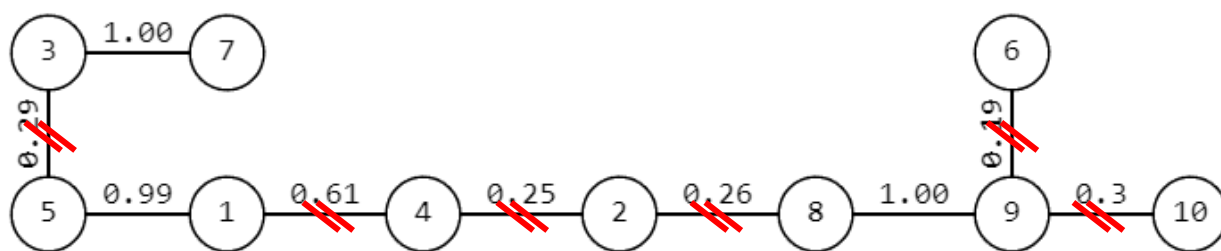


Рисунок Б.2 - Граф взаимных связей признака В

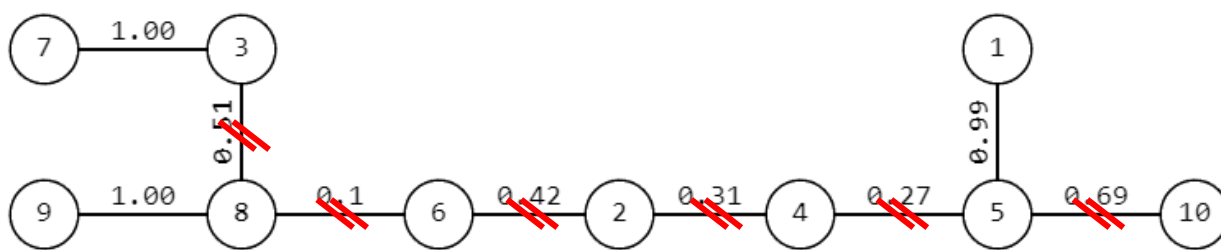


Рисунок Б.3 - Граф взаимных связей признака С

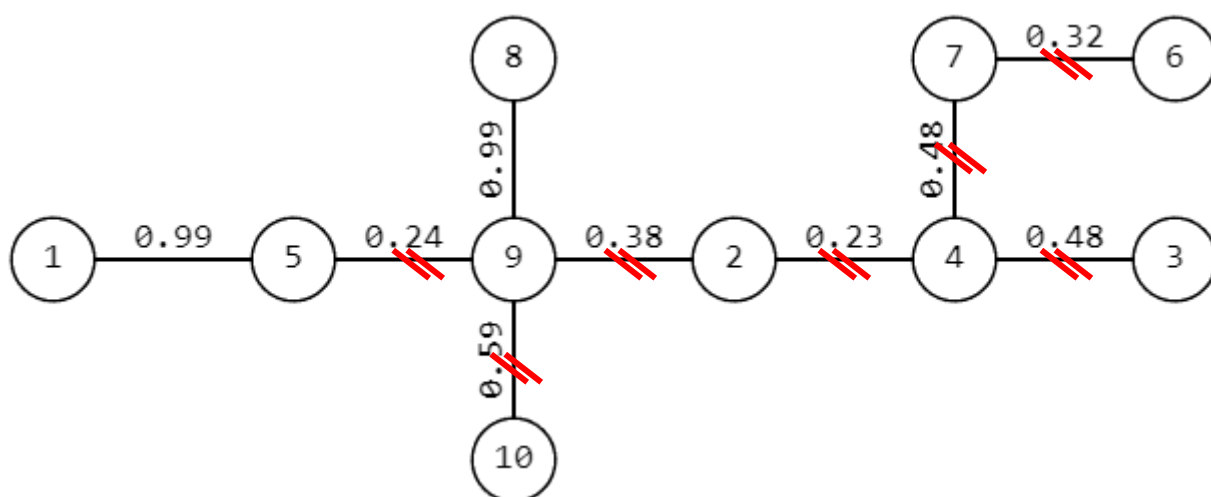


Рисунок Б.4 - Граф взаимных связей признака D

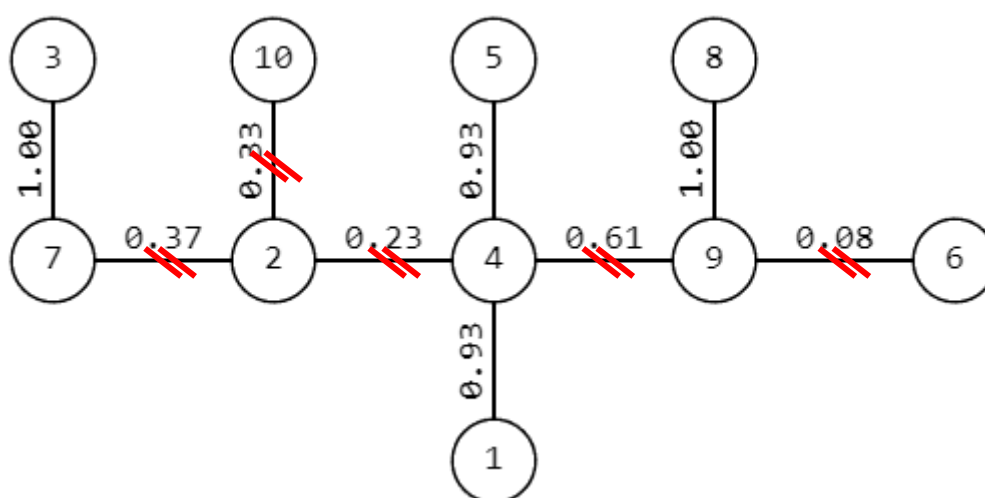


Рисунок Б.5 - Граф взаимных связей признака АВ

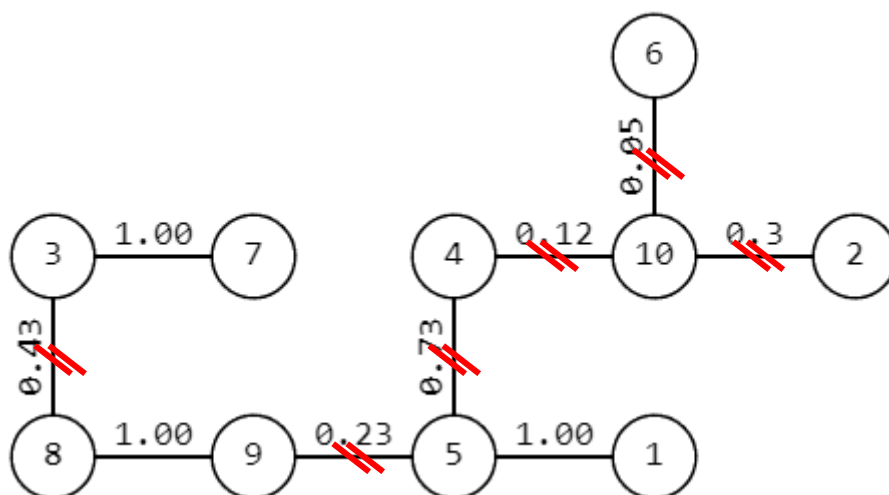


Рисунок Б.6 - Граф взаимных связей признака АС

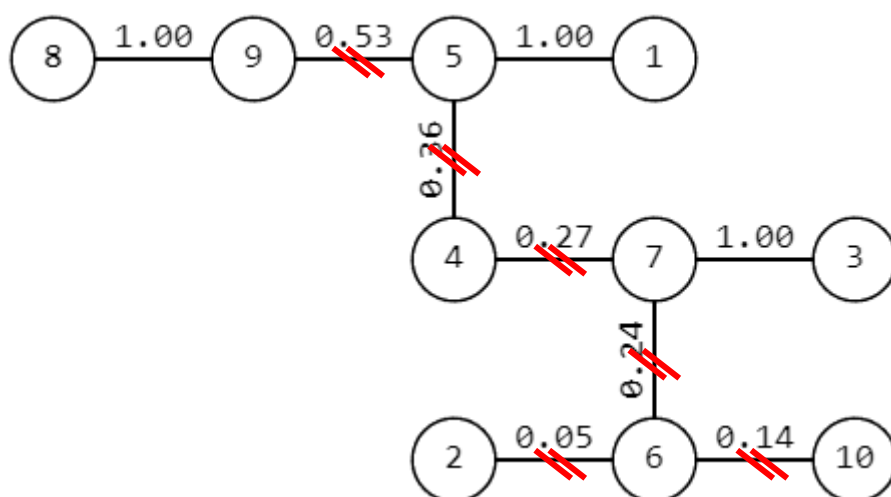


Рисунок Б.7 - Граф взаимных связей признака BD

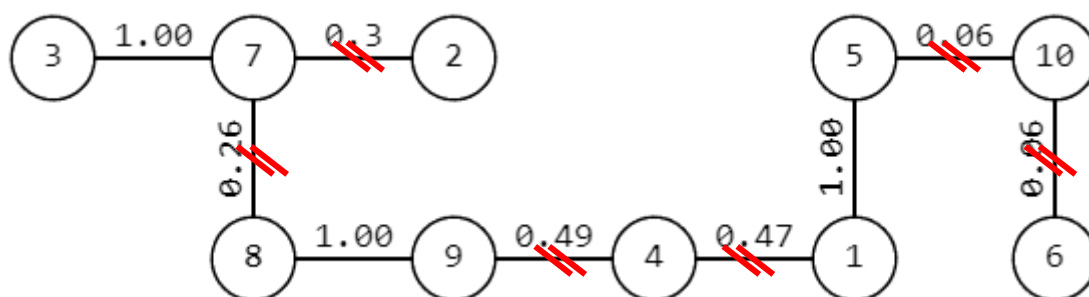


Рисунок Б.8 - Граф взаимных связей признака ABC

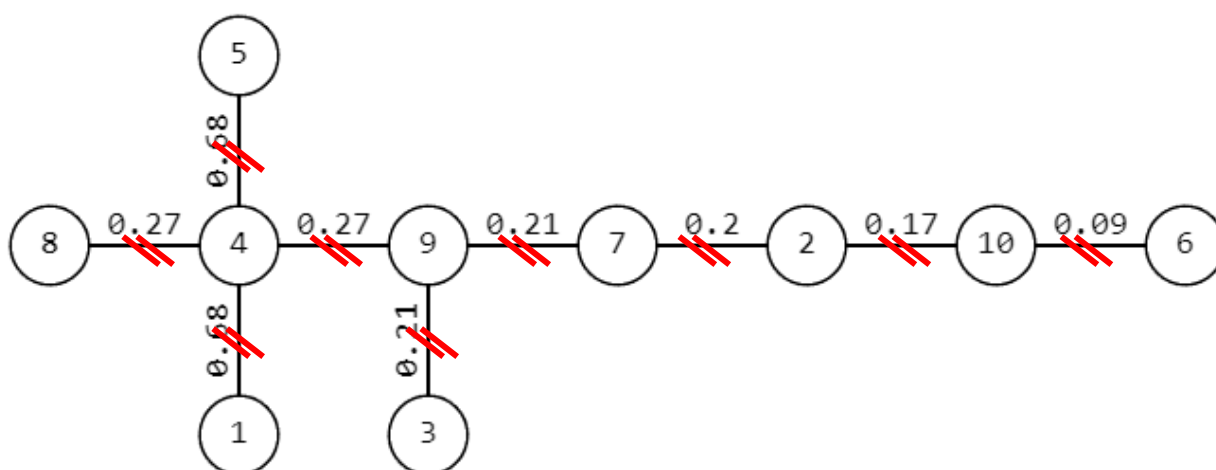


Рисунок Б.9 - Граф взаимных связей признака ABCD

ПРИЛОЖЕНИЕ В

Таблица В.1 - Результат классификации объектов из обучающей выборки для дерева 1 и 2

№	1	2	3	4	5	6	7	8	9	10	Дис .ф- ия	Дер ево 1	к сос- ей	Дер ево 2
1	18	19	22	55	35	17	11	14	29	22	А	А	А	А
2	17	16	26	56	34	18	13	11	23	34	А	А	А	А
3	15	12	38	57	30	13	19	16	32	36	А	А	А	А
4	16	11	32	57	32	19	16	19	38	26	А	А	А	А
5	19	17	24	57	37	16	12	13	26	22	А	А	А	А
6	15	18	22	59	30	12	11	15	31	30	А	А	А	А
7	17	13	34	60	33	11	17	12	24	32	А	А	А	А
8	18	11	36	60	36	17	18	21	43	38	А	А	А	А
9	18	15	26	61	35	18	13	19	37	34	А	А	А	А
10	14	16	22	62	28	13	11	13	25	24	А	А	А	А
11	17	19	30	62	33	11	15	13	26	26	А	А	А	А
12	15	17	32	63	30	15	16	12	24	22	А	А	А	А
13	18	12	38	63	35	16	19	20	39	30	А	А	А	А
14	16	13	34	64	32	19	17	11	21	32	А	А	А	А
15	15	11	24	65	29	17	12	15	31	38	А	А	А	А
16	27	15	26	37	53	12	13	17	35	34	С	В	В	В
17	28	16	22	38	55	19	11	19	39	26	С	В	В	В
18	25	19	30	39	50	16	15	16	32	38	В	В	В	В
19	29	17	32	39	57	12	16	24	48	32	Д	В	В	В
20	24	13	38	41	47	11	19	23	45	24	В	В	В	В
21	27	19	34	41	54	17	17	20	41	22	С	В	В	В
22	22	16	26	42	43	18	13	27	53	34	В	В	В	В
23	26	12	38	42	51	13	19	25	51	36	В	В	В	В
24	27	19	32	43	55	11	16	23	47	22	С	В	В	В
25	21	16	24	44	41	15	12	28	56	28	В	В	В	В
26	25	12	22	44	49	16	11	21	42	34	В	В	В	В
27	23	11	34	45	46	19	17	18	36	26	В	В	В	В
28	26	17	36	45	52	17	18	16	31	38	С	В	В	В
29	22	18	22	46	43	12	11	15	30	32	В	В	В	В
30	24	13	28	47	47	13	14	26	51	24	В	В	В	В
31	30	11	34	51	60	11	17	18	36	22	С	С	В	С
32	28	15	26	52	56	15	13	23	47	34	С	С	С	С
33	32	16	38	52	64	16	19	25	50	36	Д	С	С	С
34	26	19	32	53	52	19	16	14	28	22	С	С	С	С
35	30	17	24	53	59	17	12	13	26	28	С	С	С	С
36	28	12	22	54	55	13	11	11	23	30	С	С	С	С
37	31	13	34	54	61	19	17	19	37	26	С	С	С	С
38	33	11	36	54	65	16	18	10	21	38	Д	С	С	С
39	25	15	22	55	50	12	11	12	25	22	С	С	С	С
40	29	16	28	56	58	11	14	15	30	34	С	С	С	С

Таблица В.2 – Окончание таблицы В.1 - результат классификации объектов из обучающей выборки для дерева 1 и 2

№	1	2	3	4	5	6	7	8	9	10	Дис . ф- ия	Дер ево 1	К сос -ей	Дер ево 2
41	32	19	30	56	63	17	15	13	26	36	С	С	С	С
42	27	17	26	57	54	18	13	11	22	26	С	С	С	С
43	26	13	38	58	51	17	19	25	51	22	С	С	С	С
44	30	19	34	58	60	18	17	21	42	30	С	С	С	С
45	28	16	24	59	56	13	12	19	38	32	С	С	С	С
46	37	12	36	35	74	15	18	16	33	38	D	D	D	D
47	33	19	26	41	67	16	13	14	27	34	D	D	D	D
48	34	16	22	37	68	19	11	11	21	24	D	D	D	D
49	35	12	30	47	69	17	15	19	37	26	D	D	D	D
50	35	11	32	43	70	12	16	10	21	22	D	D	D	D
51	36	17	38	41	71	13	19	15	31	30	D	D	D	D
52	35	18	34	36	71	11	17	17	33	32	D	D	D	D
53	37	13	24	43	73	15	12	21	42	38	D	D	D	D
54	37	11	26	38	74	16	13	23	45	34	D	D	D	D
55	38	15	22	48	75	19	11	16	32	26	D	D	D	D
56	39	16	30	43	77	17	15	14	27	38	D	D	D	D
57	39	19	32	38	77	13	16	13	26	22	D	D	D	D
58	39	17	38	34	78	19	19	18	36	34	D	D	D	D
59	40	12	34	38	80	16	17	21	42	36	D	D	D	D
60	41	13	26	42	81	12	13	13	26	26	D	D	D	D

Таблица В.2 - Результат классификации объектов из тестовой выборки для дерева 1 и 2

№	1	2	3	4	5	6	7	8	9	10	Дис. ф-ия	Дер ево 1	Дер ево 2
1	18	13	26	57	35	17	13	11	23	22	А	А	А
2	16	11	38	59	32	18	19	16	33	34	А	А	А
3	18	15	32	59	35	13	16	19	39	36	А	А	А
4	15	16	24	61	29	11	12	15	31	26	А	А	А
5	16	19	22	62	32	15	11	11	22	22	А	А	А
6	24	17	28	39	48	16	14	16	32	30	В	В	В
7	26	18	30	40	52	19	15	17	35	32	В	В	В
8	28	13	26	40	55	17	13	13	25	38	С	В	В
9	25	11	38	42	49	12	19	21	41	34	В	В	В
10	23	15	34	43	45	13	17	19	37	24	В	В	В
11	29	16	34	54	57	11	17	12	24	26	С	С	С
12	32	19	26	54	63	15	13	15	30	22	С	С	С
13	27	13	38	55	53	16	19	11	22	30	С	С	С
14	31	11	32	56	61	19	16	17	34	32	С	С	С
15	28	15	24	57	56	17	12	18	37	38	С	С	С
16	36	16	22	38	71	13	11	15	29	34	Д	Д	Д
17	37	19	28	45	73	11	14	11	21	26	Д	Д	Д
18	37	17	30	41	74	15	15	14	28	38	Д	Д	Д
19	38	18	26	36	76	16	13	18	36	32	Д	Д	Д
20	40	13	38	41	79	19	19	19	37	24	Д	Д	Д

Таблица В.3 - Результат классификации объектов из проверяющей выборки для дерева 1 и 2

№	1	2	3	4	5	6	7	8	9	10	Д-во 1	Д-во 2
1	16	14	29	60	33	15	15	15	30	30	А	А
2	16	12	30	59	32	17	17	16	32	31	А	А
3	18	15	31	59	35	13	16	17	32	32	А	А
4	25	16	30	42	50	15	15	21	42	30	В	В
5	23	18	28	44	49	15	13	20	40	32	В	В
6	24	17	28	40	48	16	14	19	43	30	В	В
7	29	15	30	55	58	15	15	17	33	29	С	С
8	28	13	28	53	56	17	13	15	32	31	С	С
9	29	17	32	57	60	13	17	19	35	28	С	С
10	37	15	30	40	74	15	15	16	32	31	Д	Д
11	36	16	32	42	72	14	17	14	28	30	Д	Д
12	39	17	28	38	71	15	13	15	30	32	Д	Д

Таблица В.4 - Результат подсчета дискриминантных функций для деревьев 1 и 2

F_{15} (Отделение класса D от ABC для первого дерева дихотомии)
$F_{15}=(37-23,53)x_1+(73,67-46,58)x_5-(37^2 - 23,57^2 + 73,67^2 - 46,58^2)/2$
$F_{15}=13,47x_1+27,09x_5-2\,035,51$
F_{45} (Отделение класса C от AB для первого дерева дихотомии)
$F_{45}=(54,8-51,13)x_4+(57,6-41,07)x_5-(54,8^2 - 51,13^2 + 57,6^2 - 41,07^2)/2$
$F_{45}=3,67x_4+16,53x_5-1\,009,89$
F_{49} (Отделение класса A от B для первого дерева дихотомии)
$F_{49}=(60,07-42,2)x_4+(29,93-42,47)x_9-(60,07^2 - 42,2^2 + 29,93^2 - 42,47^2)/2$
$F_{49}=17,87x_4-12,54x_9-459,83$
F_{45} (Отделение класса D от ABC для первого дерева дихотомии)
$F_{45}=(40,27-52,36)x_4+(73,67-46,58)x_5-(40,27^2 - 52,36^2 + 73,67^2 - 46,58^2)/2$
$F_{45}=-12,09x_4+27,09x_5-1\,068,83$
F_{58} (Отделение класса C от AB для первого дерева дихотомии)
$F_{58}=(57,60-41,07)x_5+(16,6-18,07)x_8-(57,60^2 - 41,07^2 + 16,6^2 - 18,07^2)/2$
$F_{58}=16,53x_5-1,47x_8-790,03$
F_{14} (Отделение класса A от B для первого дерева дихотомии)
$F_{14}=(16,53-25,07)x_1+(60,07-42,2)x_4-(16,53^2 - 25,07^2 + 60,07^2 - 42,2^2)/2$
$F_{14}=-8,54x_1+17,87x_4-736,15$
F_{14} (Отделение класса AC от BD для второго дерева дихотомии)
$F_{14}=(22,77-31,03)x_1+(57,43-41,23)x_4-(22,77^2 - 31,03^2 + 57,43^2 - 41,23^2)/2$
$F_{14}=-8,26x_1+16,2x_4-576,95$
F_{45} (Отделение класса A от C для второго дерева дихотомии)
$F_{45}=(60,07-54,8)x_4+(32,6-57,6)x_5-(60,07^2 - 54,8^2 + 32,6^2 - 57,6^2)/2$
$F_{45}=5,27x_4-25x_5+824,82$
F_{59} (Отделение класса B от D для второго дерева дихотомии)
$F_{59}=(49,53-73,67)x_5+(42,47-31,93)x_9-(49,53^2 - 73,67^2 + 42,47^2 - 31,93^2)/2$
$F_{59}=-24,14x_5+10,54x_9+1094,94$

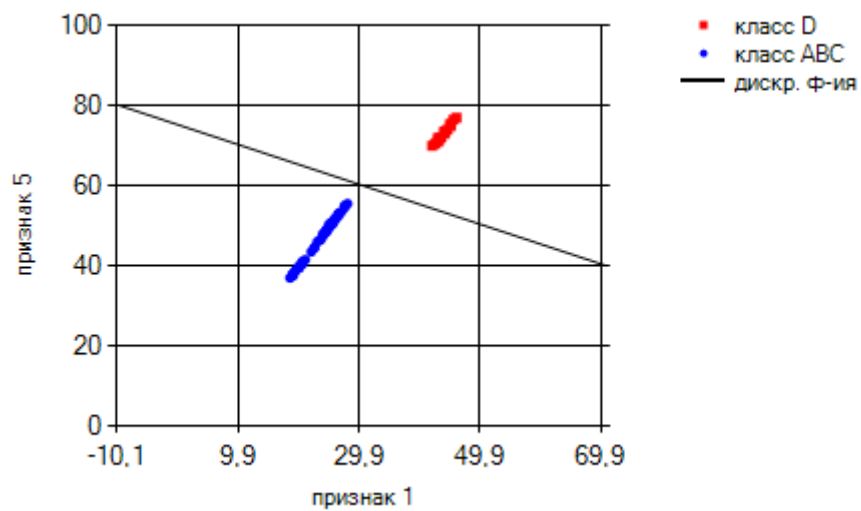


Рисунок В.1 – График, построенный по признакам 1 и 5 первым способом, распределения классов D и ABC для 1-го узла 1-го дерева дихотомии

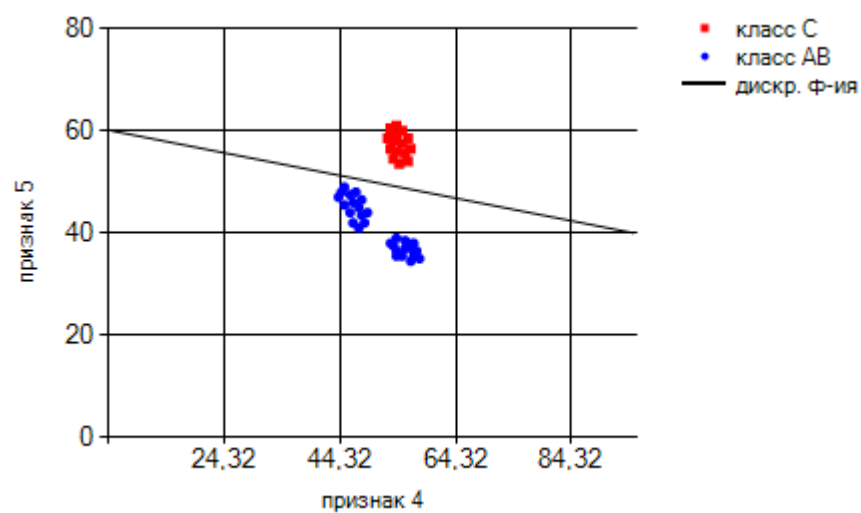


Рисунок В.2 – График, построенный по признакам 4 и 5 первым способом, распределения классов C и AB для 2-го узла 1-го дерева дихотомии

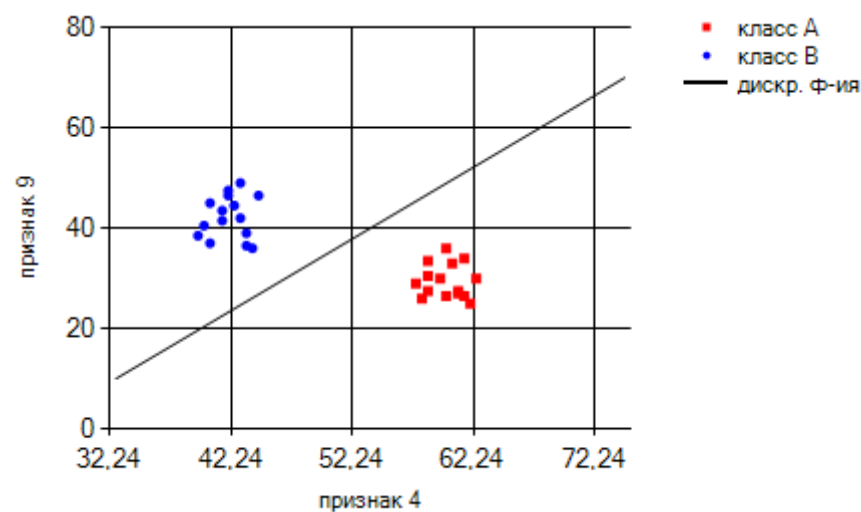


Рисунок В.3 – График, построенный по признакам 4 и 9 первым способом, распределения классов A и B для 3-го узла 1-го дерева дихотомии

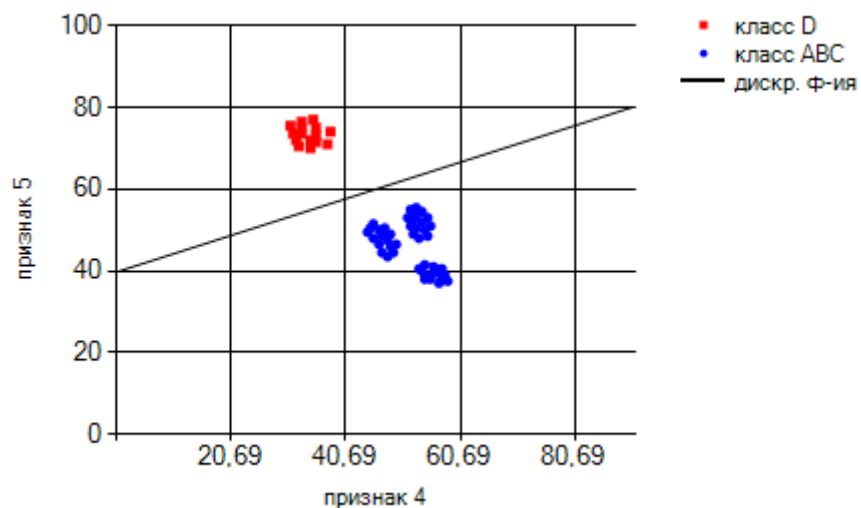


Рисунок В.4 – График, построенный по признакам 4 и 5 вторым способом, распределения классов D и ABC для 1-го узла 1-го дерева дихотомии

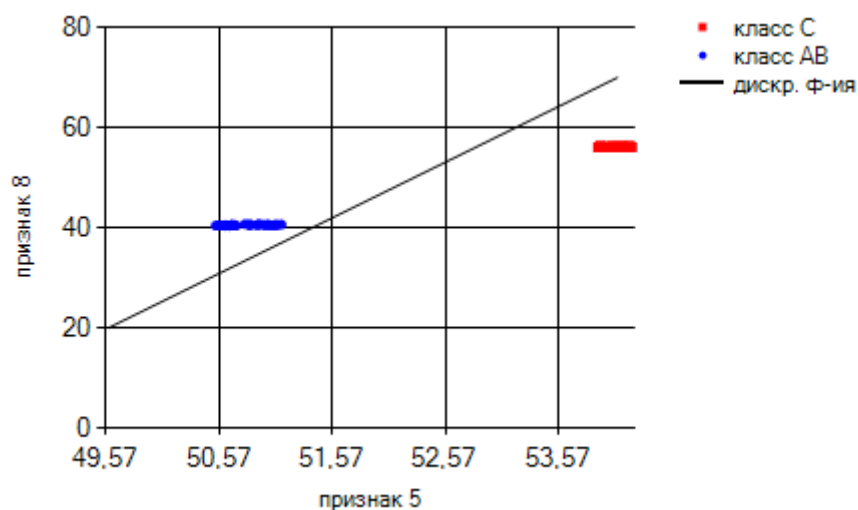


Рисунок В.5 – График, построенный по признакам 5 и 8 вторым способом, распределения классов C и AB для 2-го узла 1-го дерева дихотомии

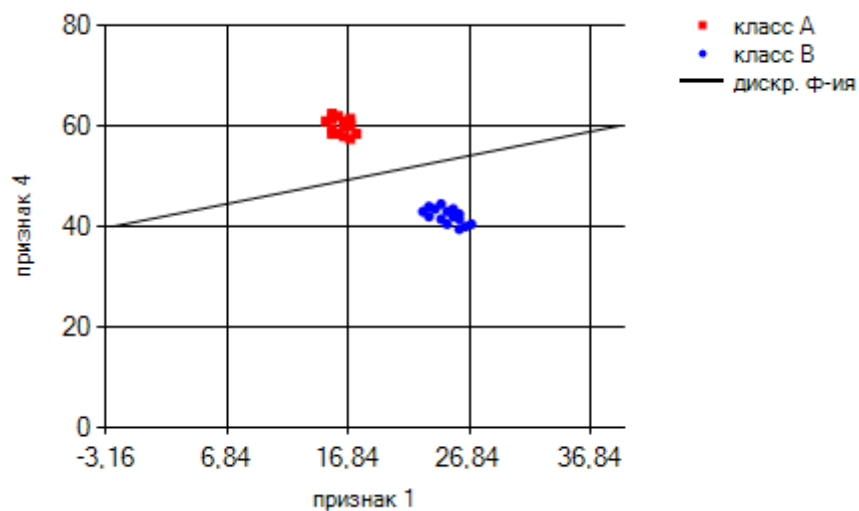


Рисунок В.6 – График, построенный по признакам 1 и 4 вторым способом, распределения классов A и B для 3-го узла 1-го дерева дихотомии

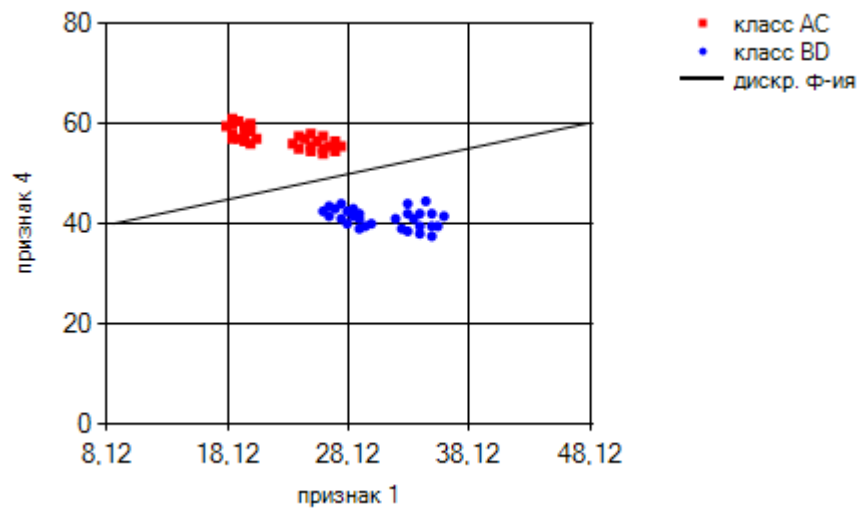


Рисунок В.7 – График, построенный по признакам 1 и 4 первым и вторым способами, распределения классов AC и BD для 1-го узла 2-го дерева дихотомии

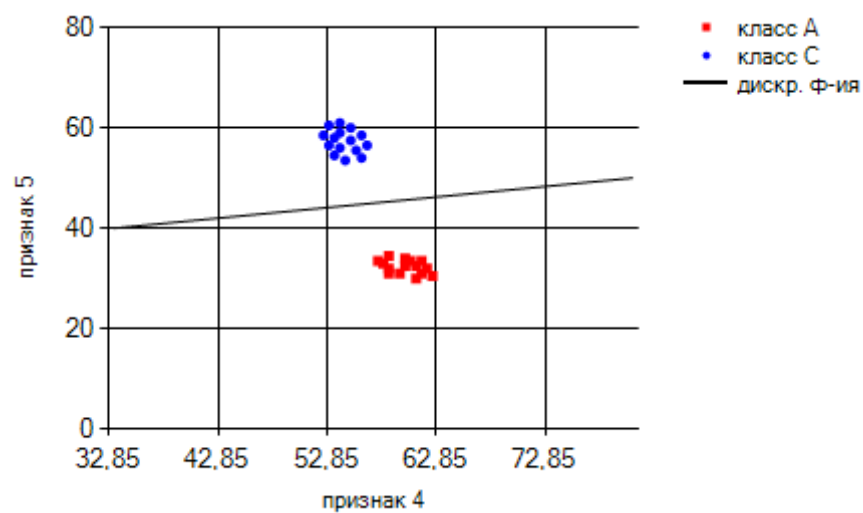


Рисунок В.8 – График, построенный по признакам 4 и 5 первым и вторым способами, распределения классов A и C для 2-го узла 2-го дерева дихотомии

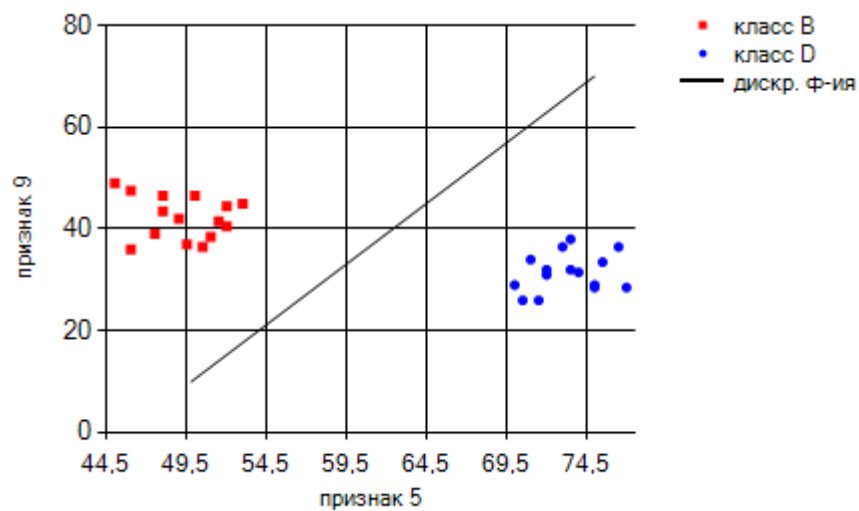


Рисунок В.9 – График, построенный по признакам 5 и 9 первым и вторым способами, распределения классов B и D для 3-го узла 2-го дерева дихотомии

ПРИЛОЖЕНИЕ Г

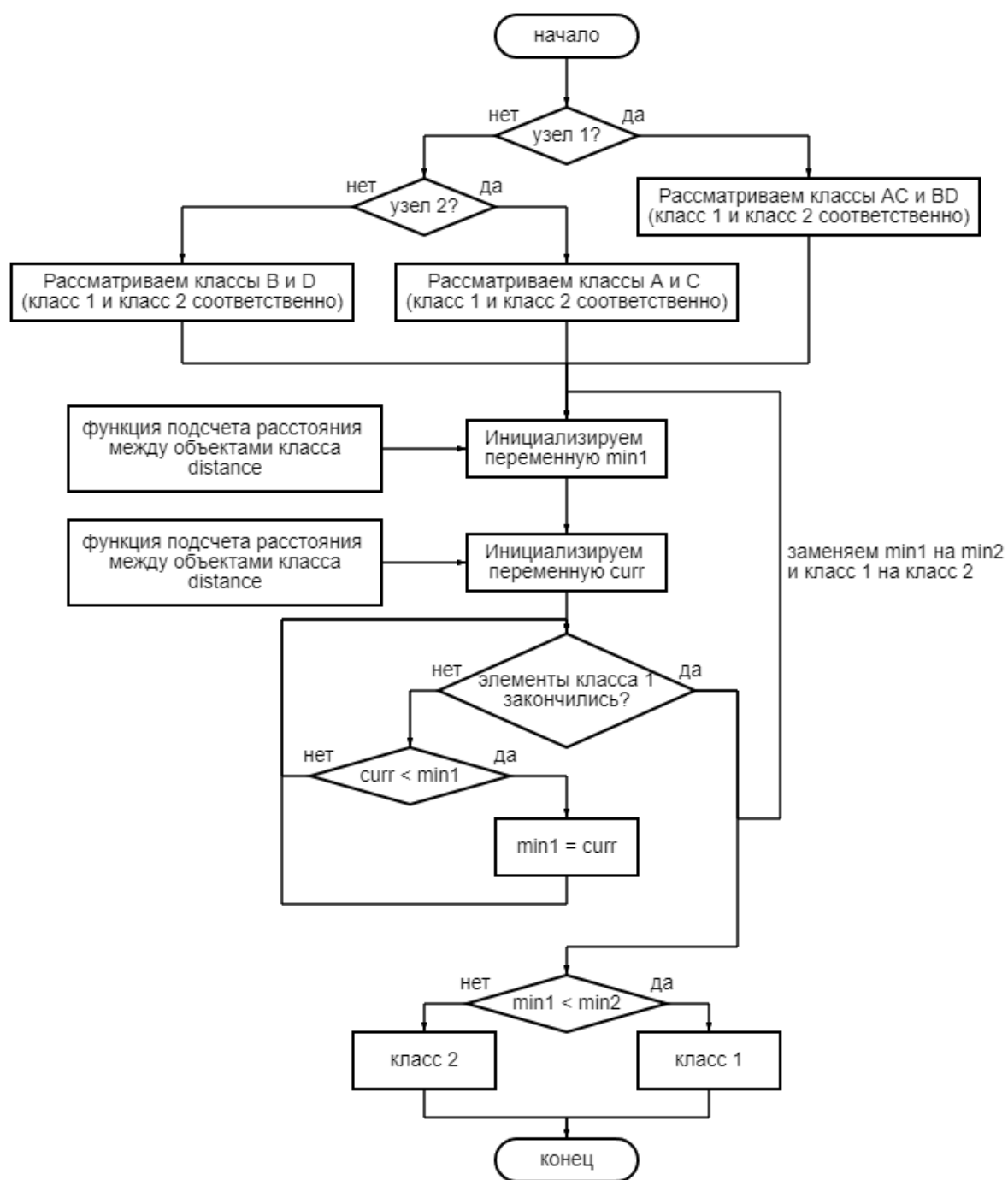


Рисунок Г.1 – Алгоритм, реализующий метод одного ближайшего соседа

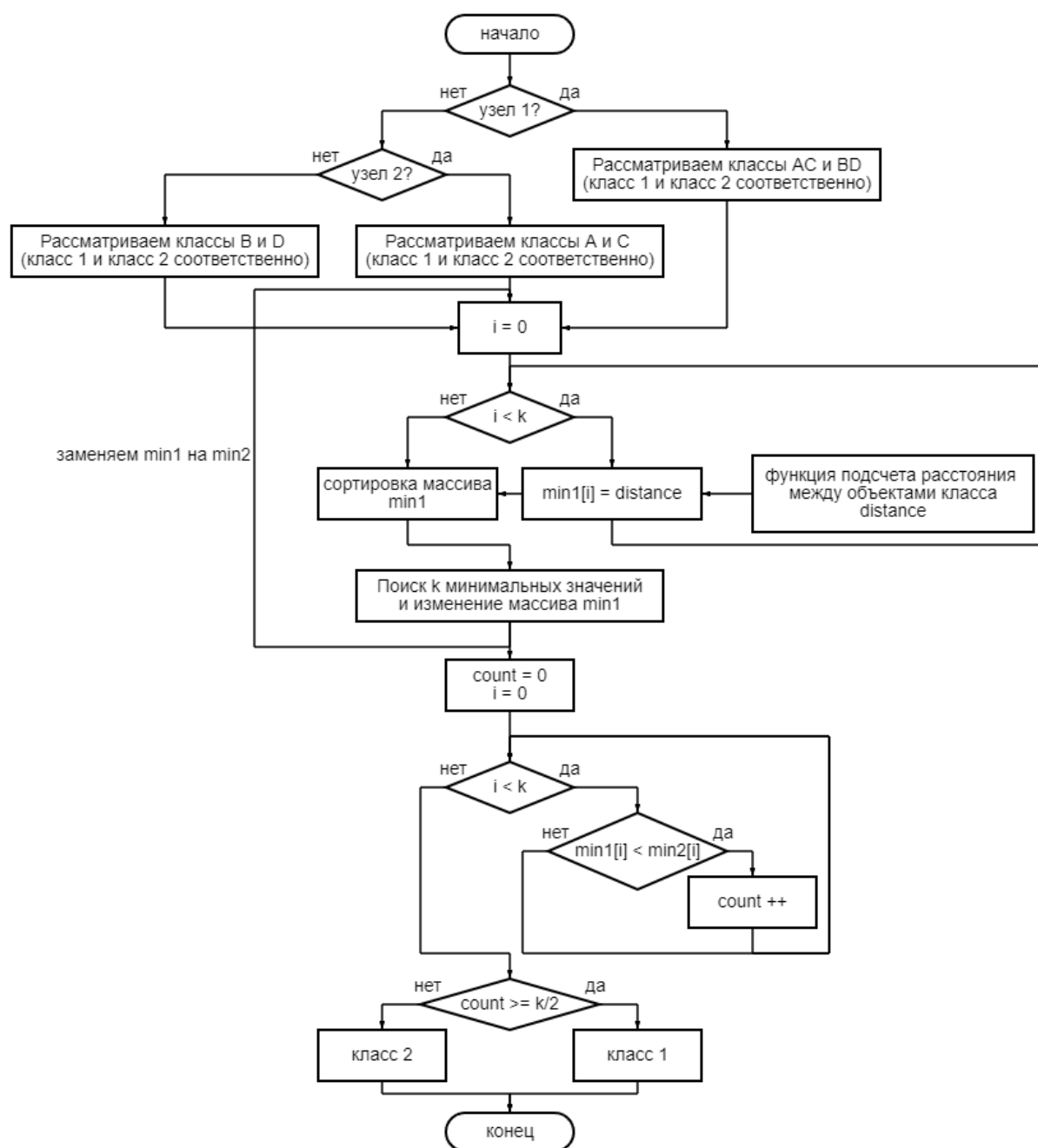


Рисунок Г.2 – Алгоритм, реализующий метод k ближайших соседей