

МИНОБРНАУКИ РОССИИ
САНКТ-ПЕТЕРБУРГСКИЙ ГОСУДАРСТВЕННЫЙ
ЭЛЕКТРОТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ
«ЛЭТИ» ИМ. В.И. УЛЬЯНОВА (ЛЕНИНА)
Кафедра САПР

ОТЧЕТ
по лабораторной работе №2
по дисциплине «Автоматизация схемотехнического проектирования»
на тему «Классификатор на основе логистической регрессии с
градиентным спуском»

Студенты гр. 1301

Семейкин С.А.

Гальченко М.А.

Преподаватель

Боброва Ю.О.

Санкт-Петербург

2025

Цель:

Разработка модели классификатора на основе логистической регрессии, изучение его свойств и принципов работы, получение навыков программирования на Python и использования модуля scikit-learn.

Ход работы:

1. Создадим переменные, распределение по нормальному закону с незначительно различными средними и дисперсиями.
2. Создадим переменные, соответствующие классам.
3. Обучим классификатор (например, логистическую регрессию) на обучающем наборе данных.
4. Визуализируем распределение вероятностей для обучающей и тестовой выборок с помощью гистограмм.
5. Оценим эффективность классификатора на тестовом наборе данных.

При:

$\mu_0 = [0, 2]$ $\mu_1 = [3, 5]$ $\sigma_0 = [2, 1]$ $\sigma_1 = [1, 2]$

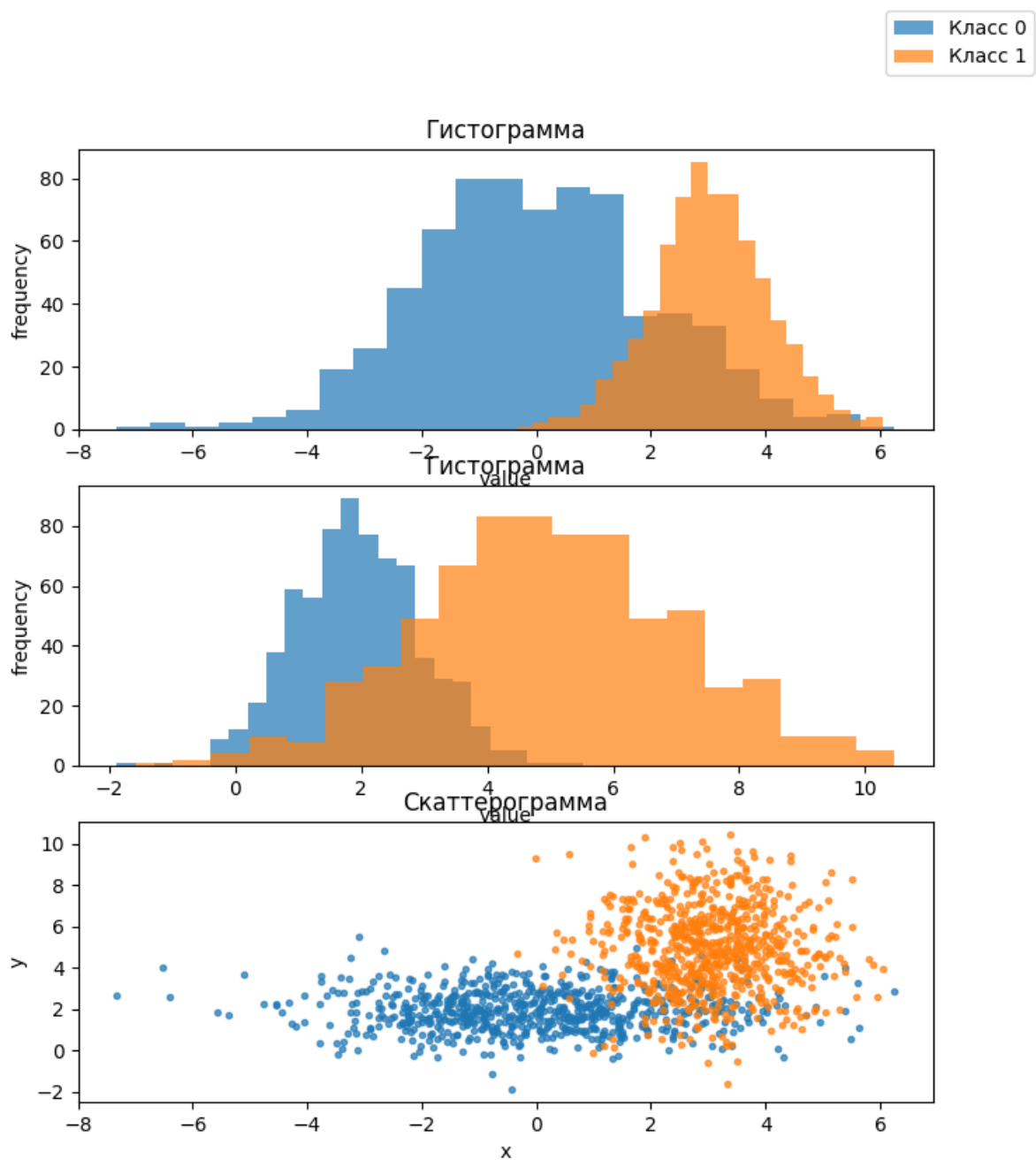


Рисунок 1 Распределение классов

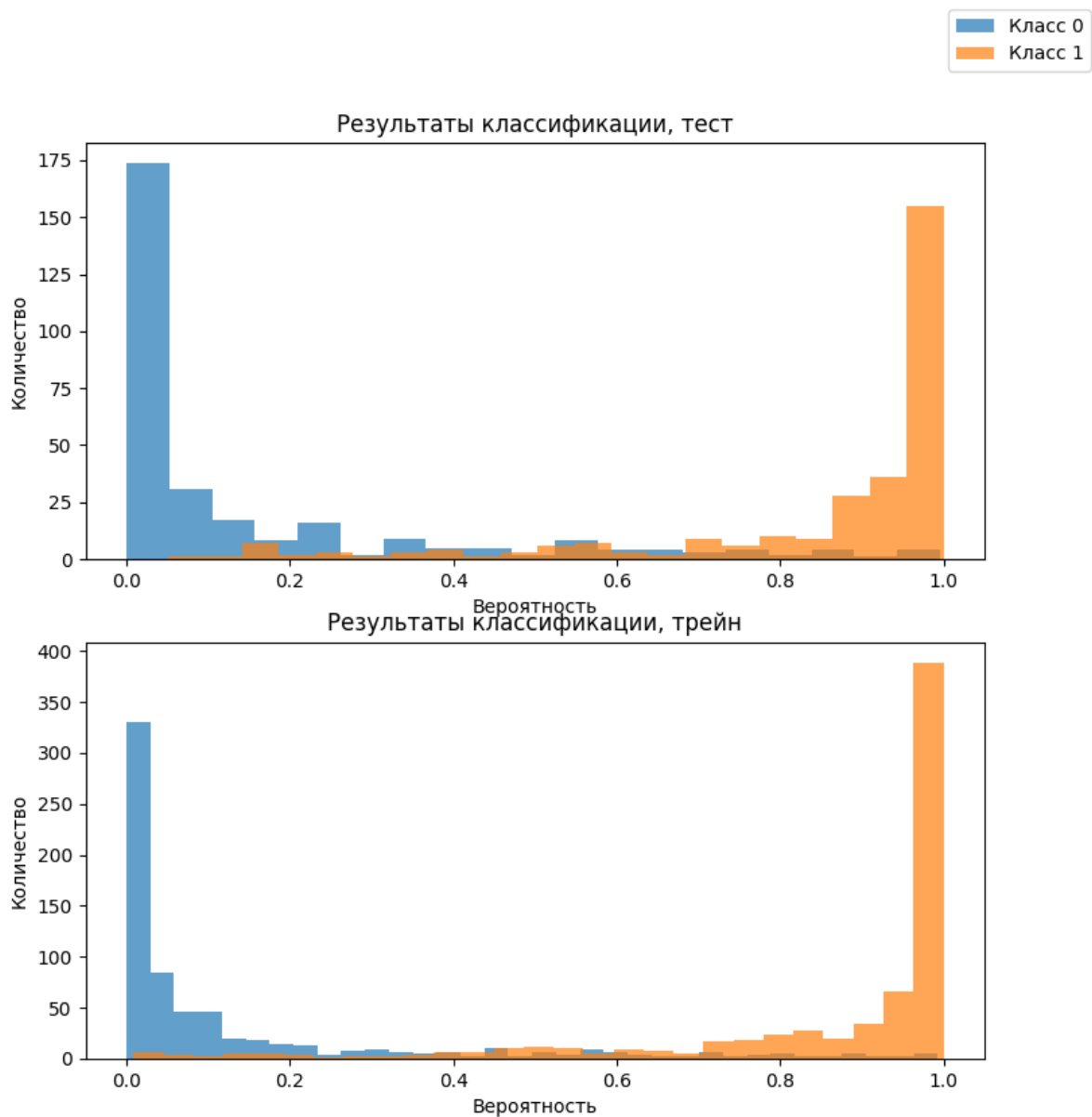


Рисунок 2 — Вероятности принадлежности объектов классам для обучающей и тестовой выборки

	Число объектов	Точность, %	Чувствительность, %	Специфичность, %
Train	1400	90.16	90.96	89.36
Test	600	91.92	91.72	92.13

При более плотном пересечении
 $\mu_0 = [0, 2]$ $\mu_1 = [1, 3]$ $\sigma_0 = [2, 2]$ $\sigma_1 = [2, 2]$

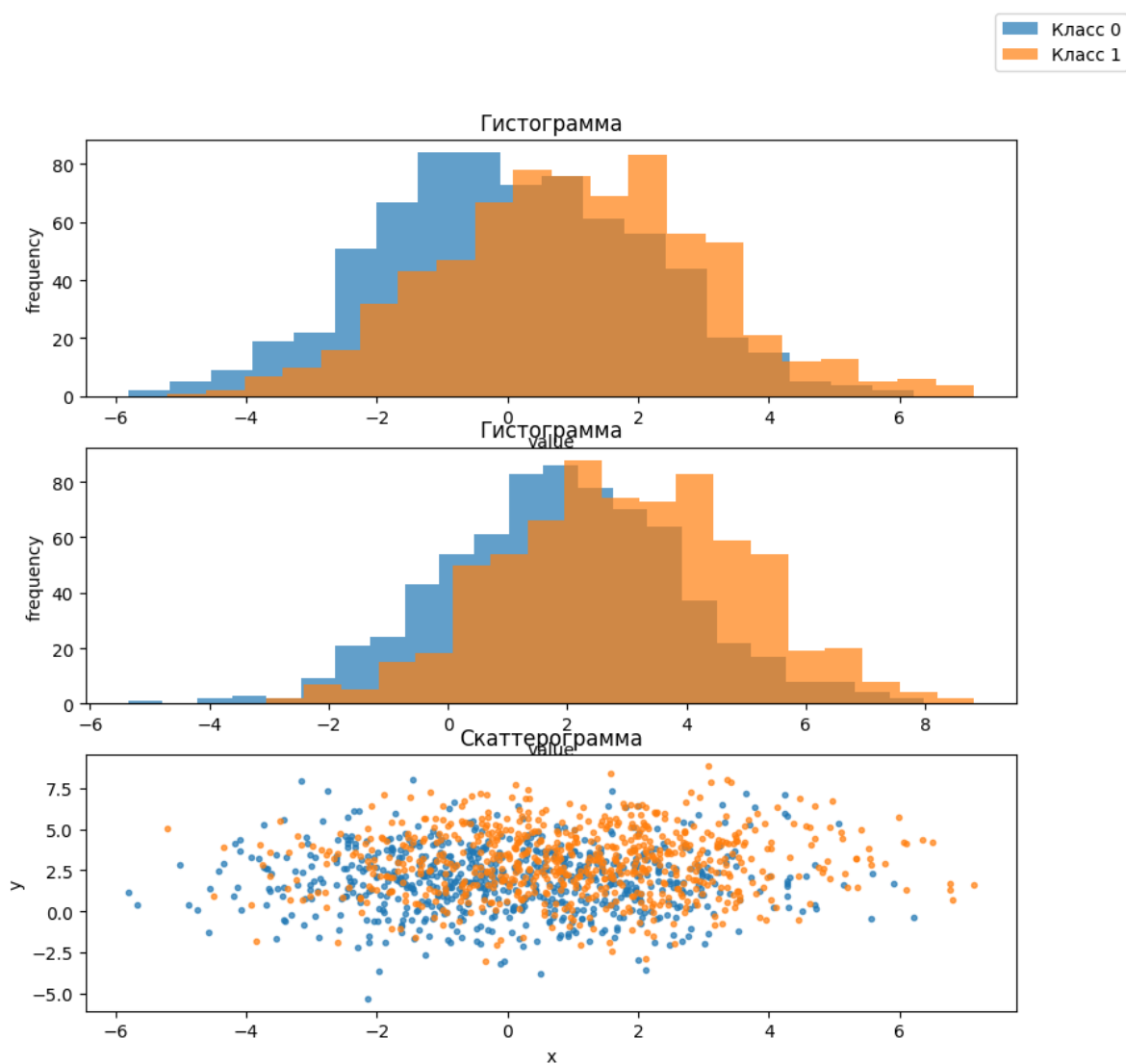


Рисунок 3 Распределение классов

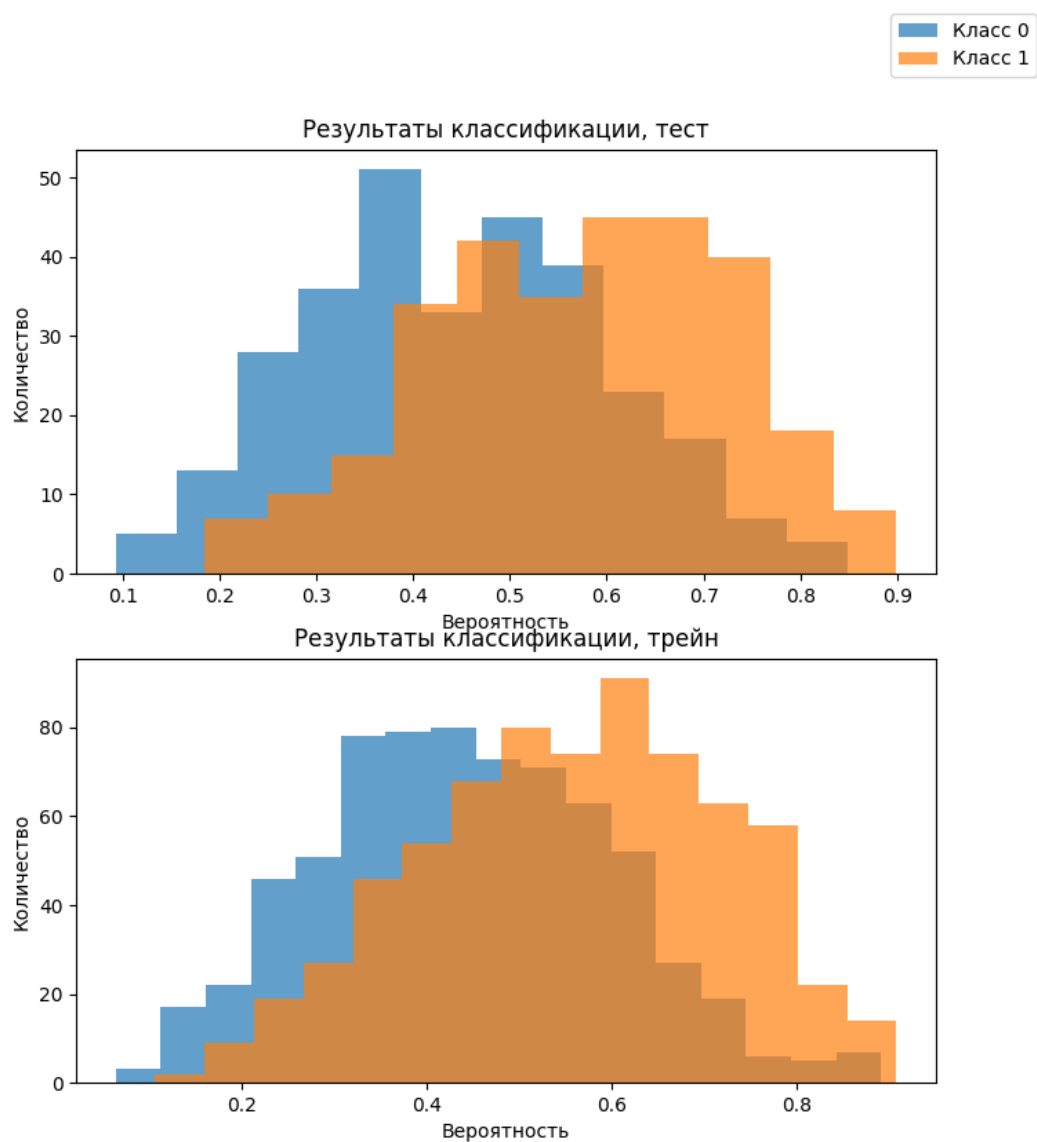


Рисунок 4 — Вероятности принадлежности объектов классам при более плотном пересечении для обучающей и тестовой выборок

	Число объектов	Точность, %	Чувствительность, %	Специфичность, %
Train	1400	63.85	64.05	63.66
Test	600	64.66	66.88	62.45

При нелинейно пересекаемой выборке:

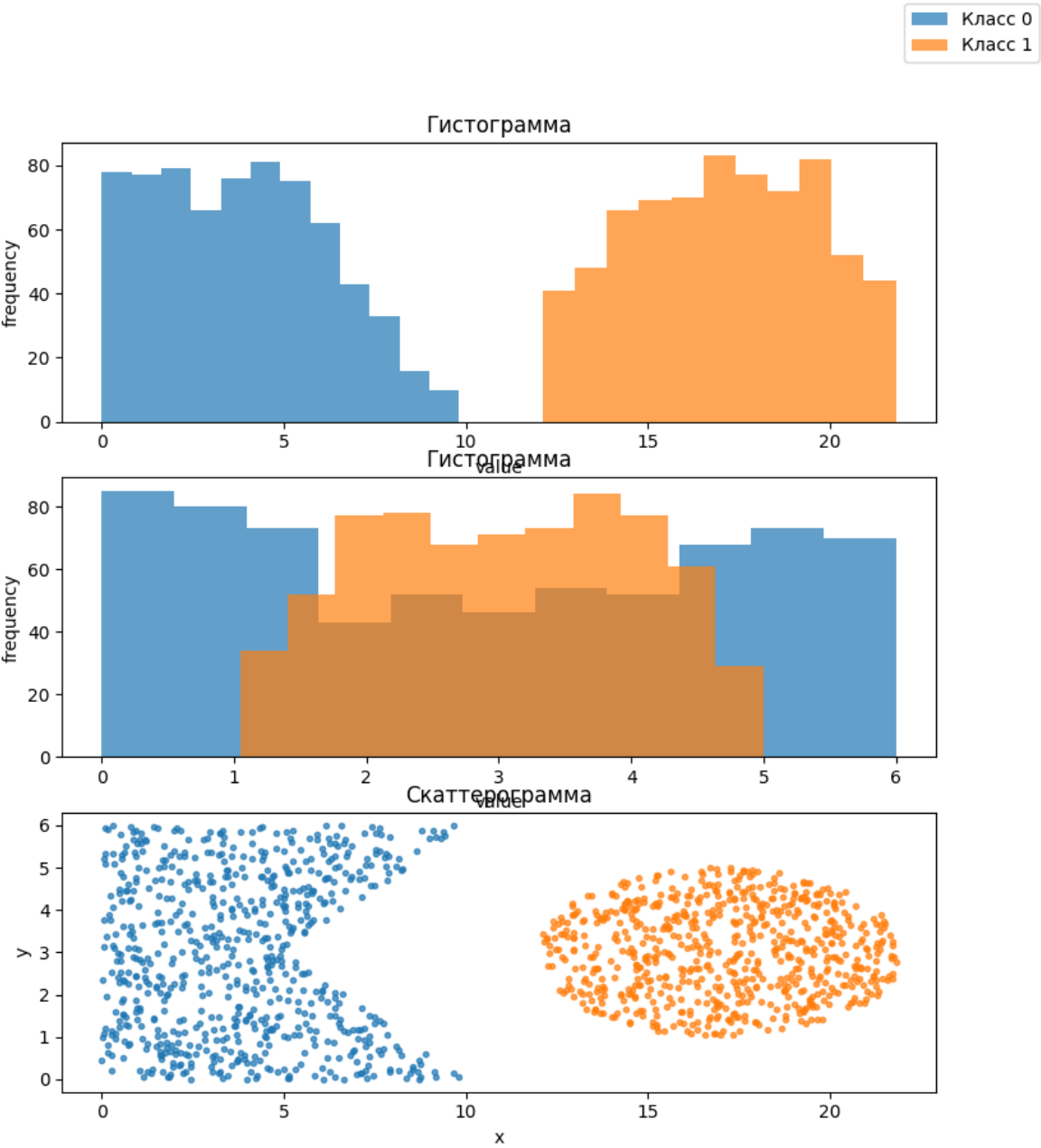


Рисунок 5 Распределение классов

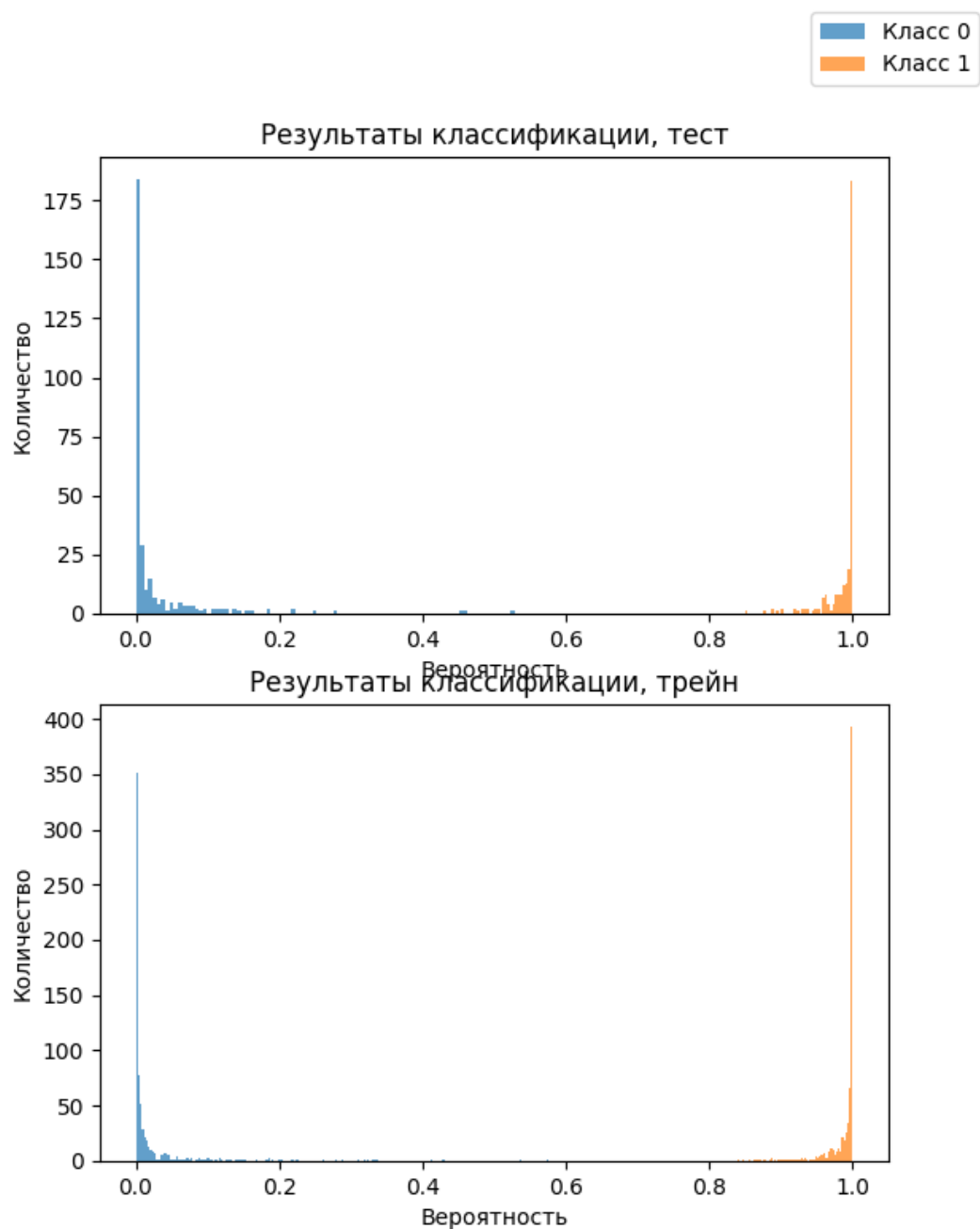


Рисунок 6 Вероятности принадлежности объектов классам при более плотном пересечении для обучающей и тестовой выборок

	Число объектов	Точность, %	Чувствительность, %	Специфичность, %
Train	1400	99.85	100	99.71
Test	600	99.83	100	99.67

Вывод:

В ходе лабораторной работы была разработана модель классификатора на основе логистической регрессии, изучены его свойства и принципы работы, получены навыки программирования на Python и использования модуля scikit-learn.

Изменяя параметры при генерации данных, мы получили более или менее пересекающиеся классы и оценили, как это влияет на эффективность классификатора — при более плотном пересечении классов точность классификатора ухудшается.

Это связано с тем, что при плотном пересечении классов граница между классами становится менее очевидной. Модель сталкивается с трудностями в поиске правильного разделения пространства, что приводит к увеличению ошибок классификации. Кроме того, может потребоваться больше обучающих данных для построения более точной модели.

При нелинейно пересекаемой выборке значения метрик модели стремились к 100%, что означает ее хорошую работу, ведь в данных явно просматривается принадлежность к разным классам.