Министерство образования Республики Беларусь

Учреждение образования

"Белорусский Государственный университет информатики

и радиоэлектроники"

Лабораторная работа № 2

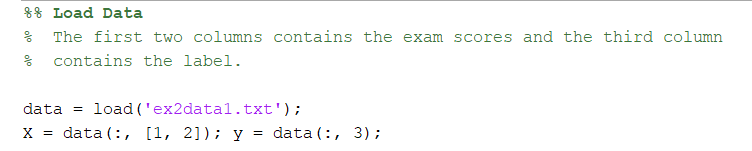
по учебной дисциплине “Машинное обучение”

|  |  |
| --- | --- |
| Выполнил: | Студент гр. 956241 Дубовик Н.О. |

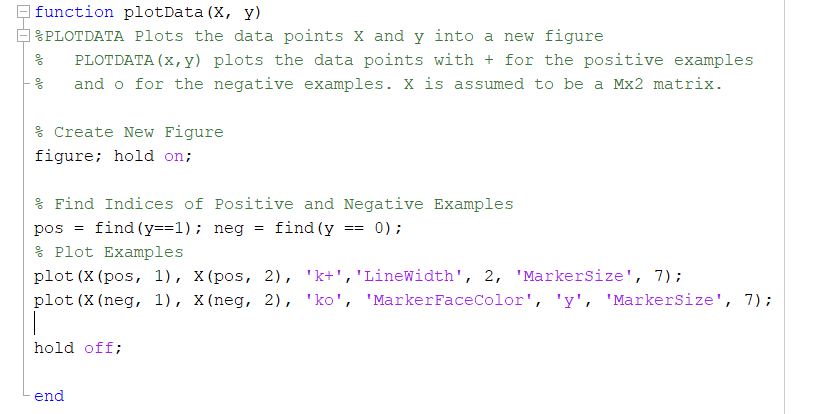
Минск 2019

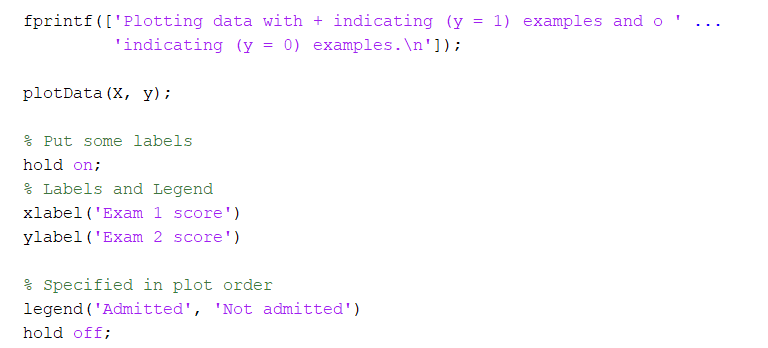
**Задание.**

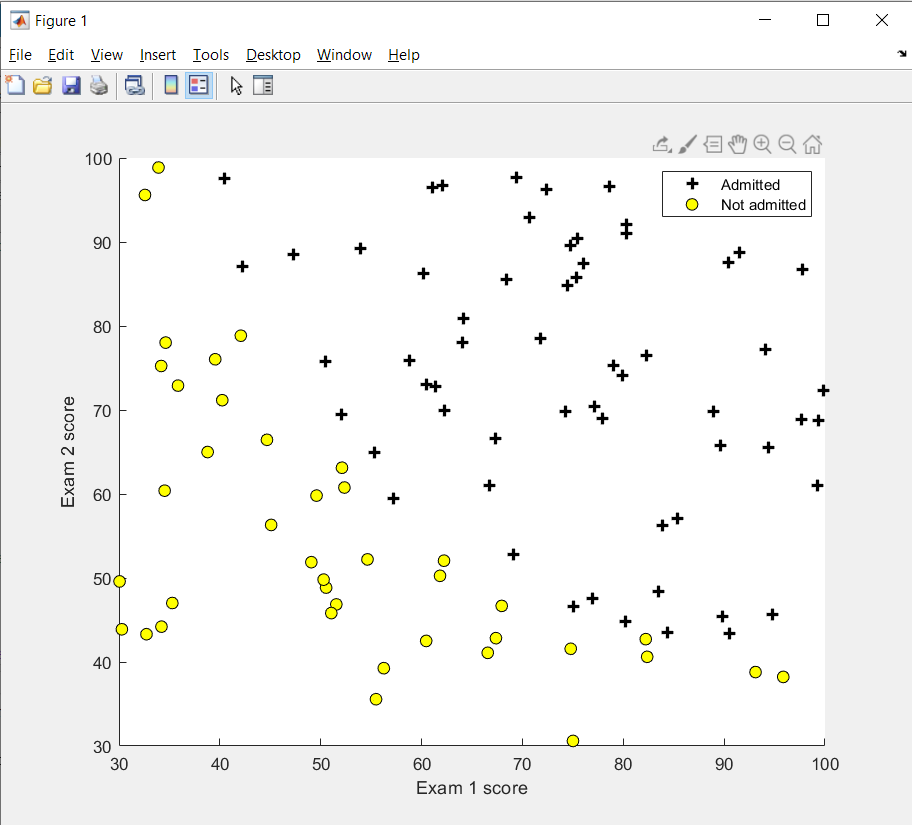
Загрузите данные **ex2data1.txt** из текстового файла.



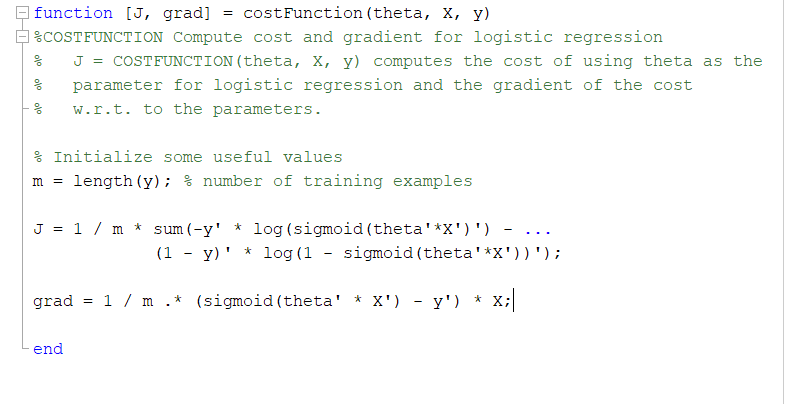
Постройте график, где по осям откладываются оценки по предметам, а точки обозначаются двумя разными маркерами в зависимости от того, поступил ли данный студент в университет или нет.

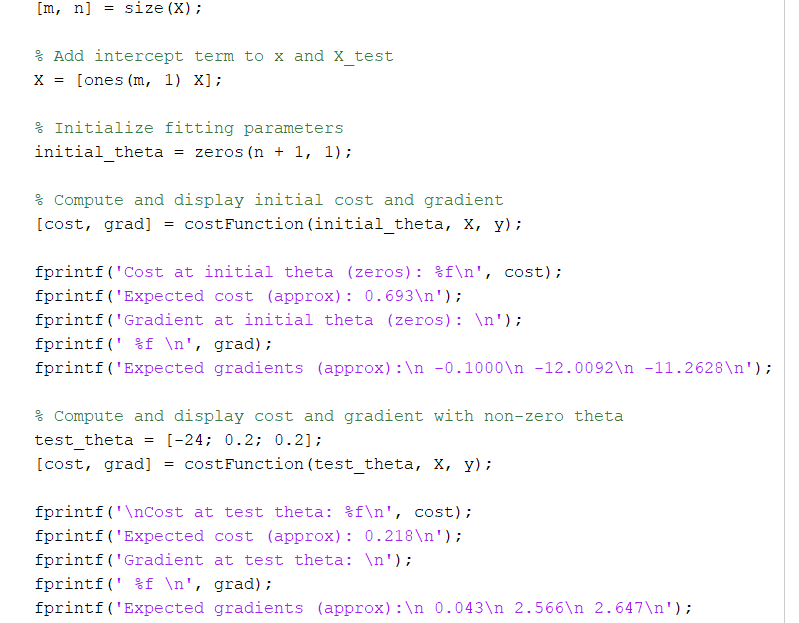


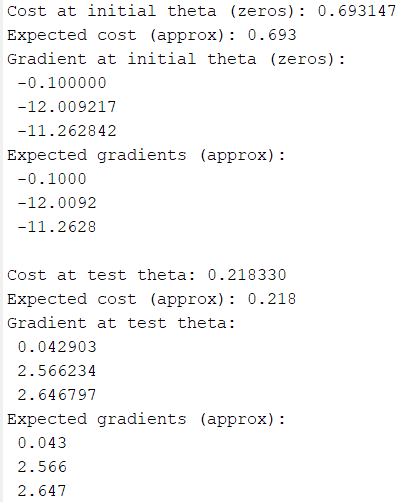




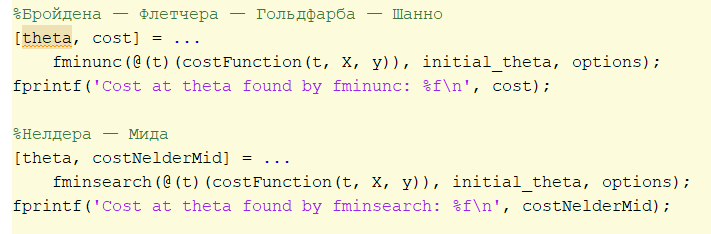
Реализуйте функции потерь J(θ) и градиентного спуска для логистической регрессии с использованием векторизации.





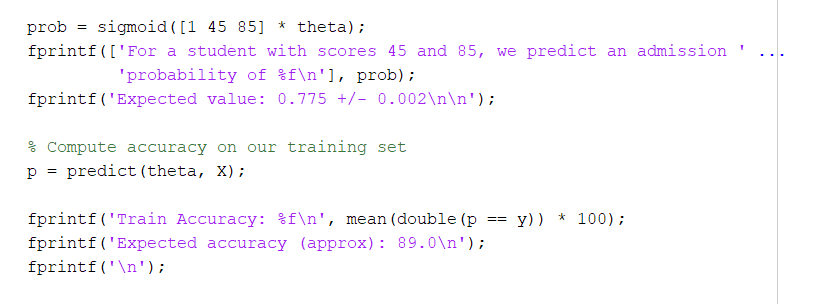


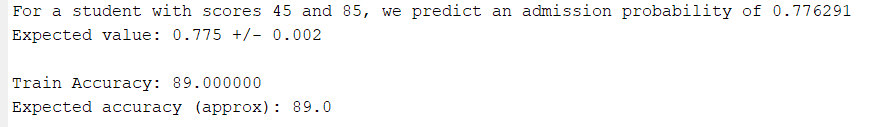
Реализуйте другие методы (как минимум 2) оптимизации для реализованной функции стоимости (например, Метод Нелдера — Мида, Алгоритм Бройдена — Флетчера — Гольдфарба — Шанно, генетические методы и т.п.). Разрешается использовать библиотечные реализации методов оптимизации (например, из библиотеки scipy).



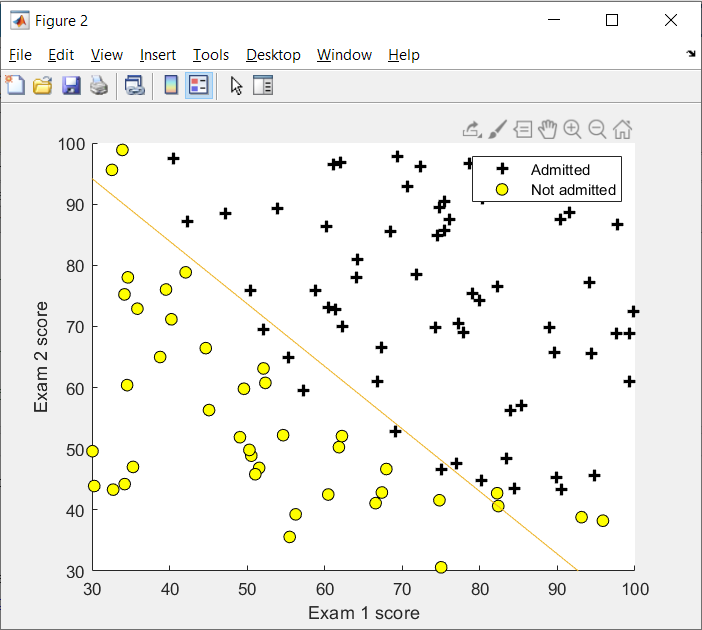


Реализуйте функцию предсказания вероятности поступления студента в зависимости от значений оценок по экзаменам.

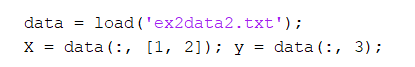




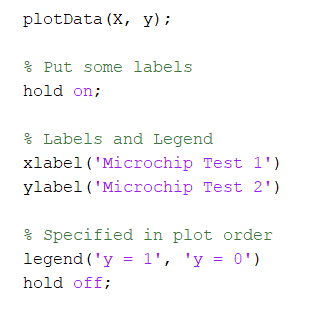
Постройте разделяющую прямую, полученную в результате обучения модели. Совместите прямую с графиком из пункта 2.

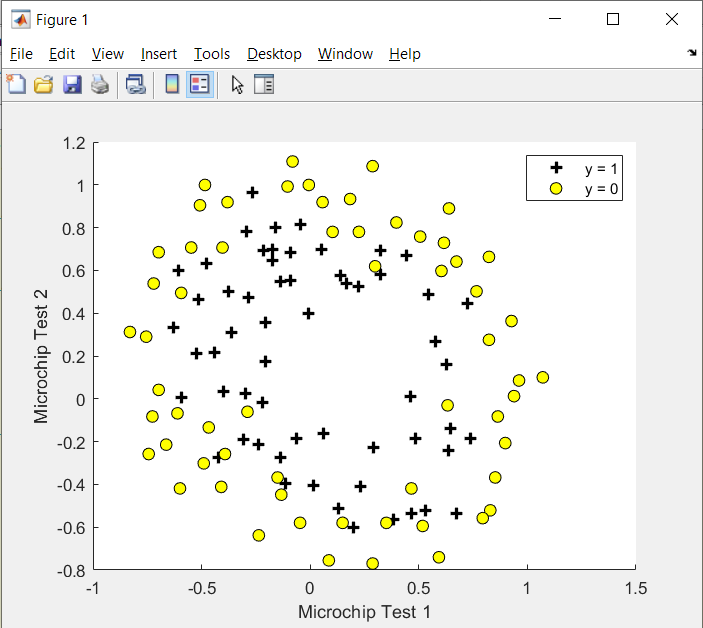


Загрузите данные **ex2data2.txt** из текстового файла.

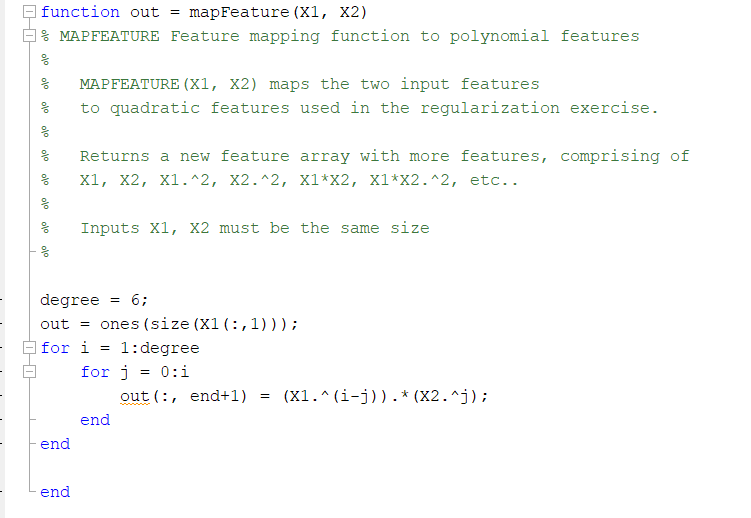


Постройте график, где по осям откладываются результаты тестов, а точки обозначаются двумя разными маркерами в зависимости от того, прошло ли изделие контроль или нет.

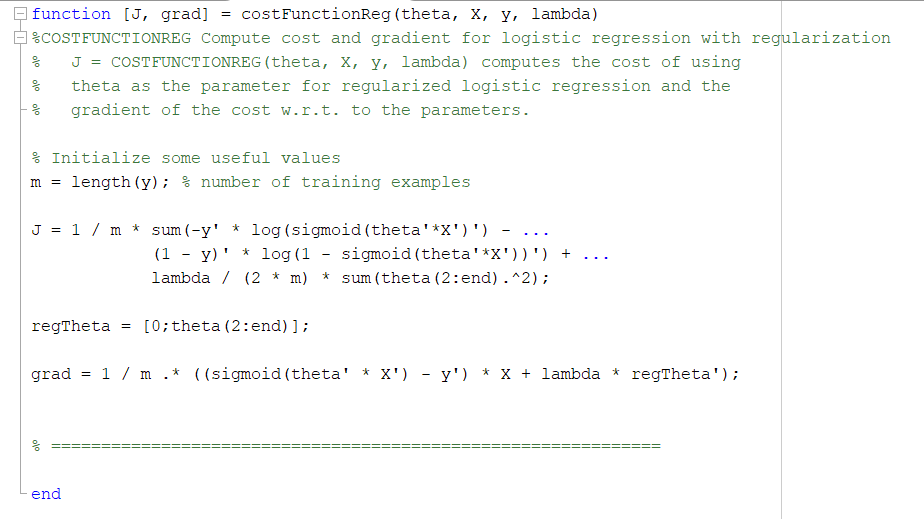


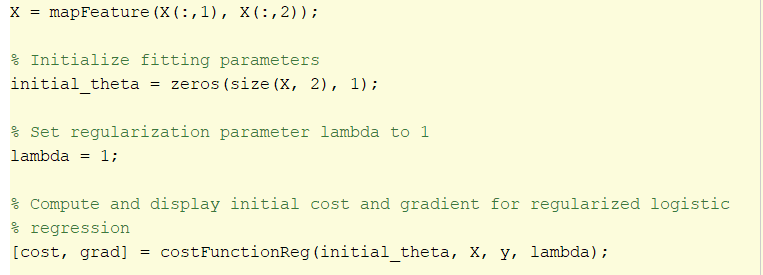


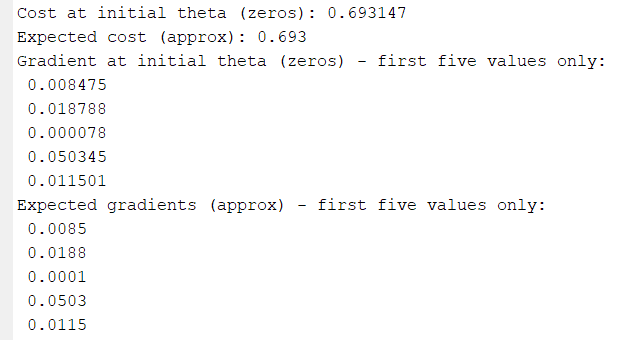
Постройте все возможные комбинации признаков x1 (результат первого теста) и x2 (результат второго теста), в которых степень полинома не превышает 6, т.е. 1, x1, x2, x12, x1x2, x22, …, x1x25, x26 (всего 28 комбинаций).



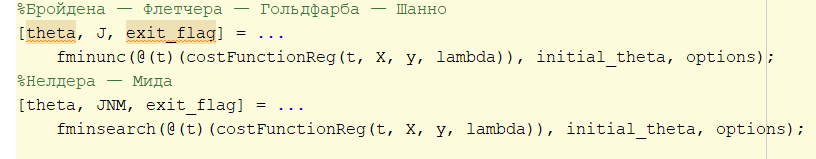
Реализуйте L2-регуляризацию для логистической регрессии и обучите ее на расширенном наборе признаков методом градиентного спуска.



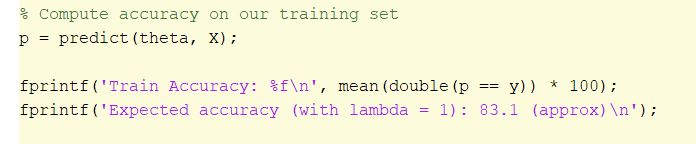




Реализуйте другие методы оптимизации.

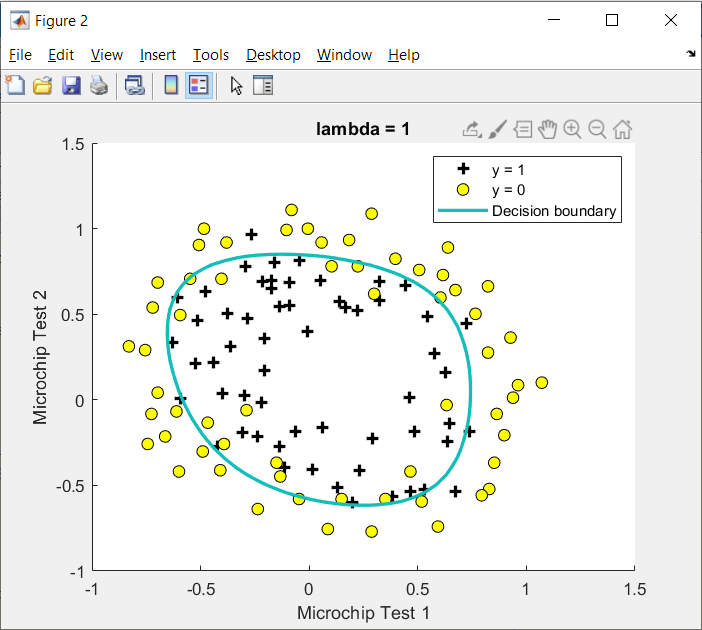


Реализуйте функцию предсказания вероятности прохождения контроля изделием в зависимости от результатов тестов.



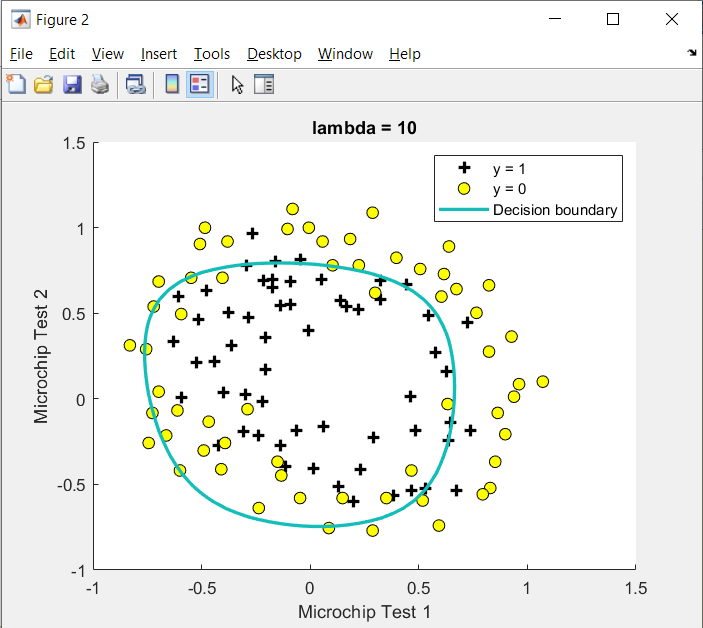


Постройте разделяющую кривую, полученную в результате обучения модели. Совместите прямую с графиком из пункта 7.

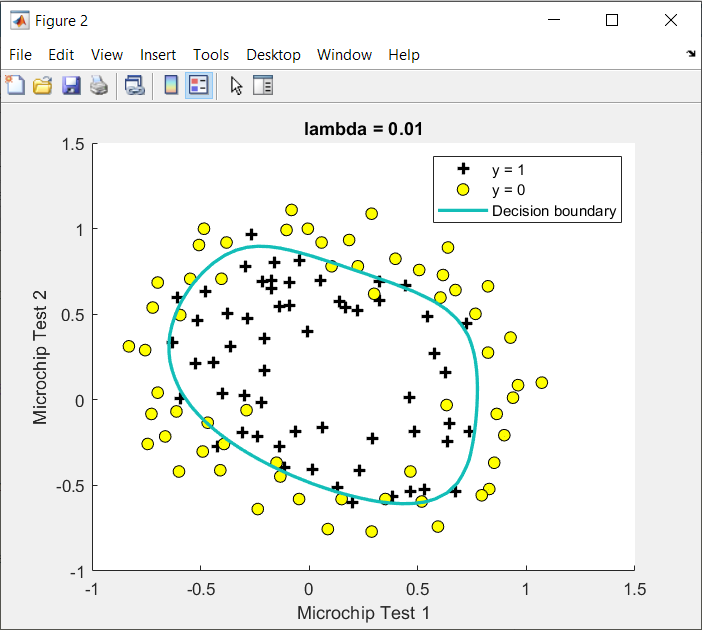


Попробуйте различные значения параметра регуляризации λ. Как выбор данного значения влияет на вид разделяющей кривой? Ответ дайте в виде графиков.

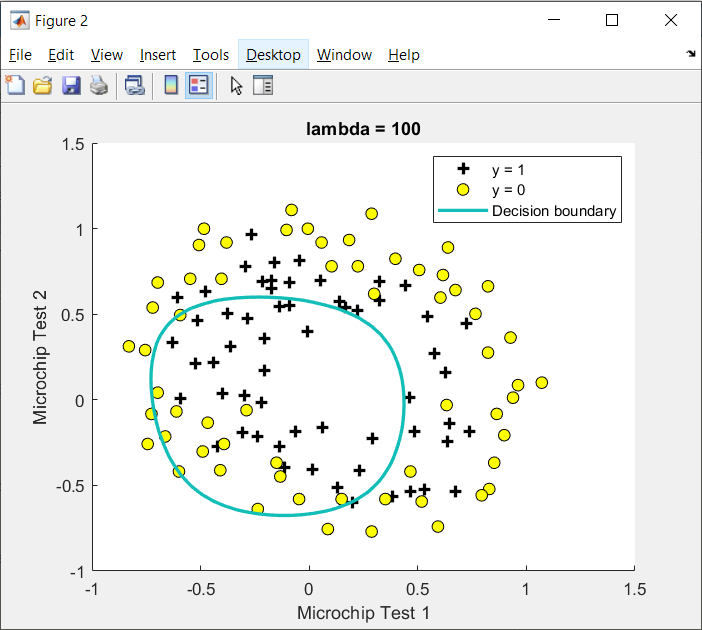
λ=10



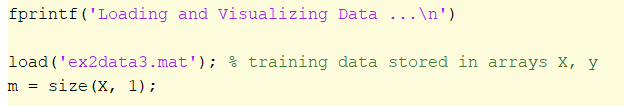
λ=0.01



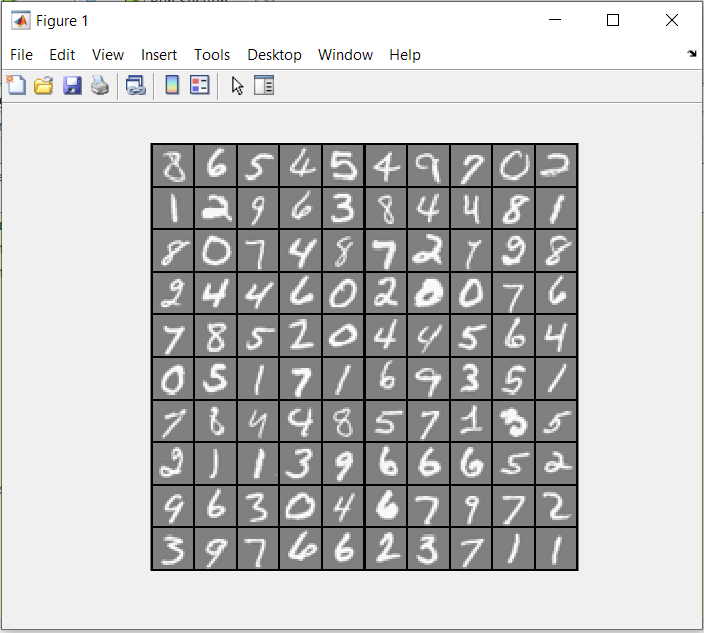
λ=100



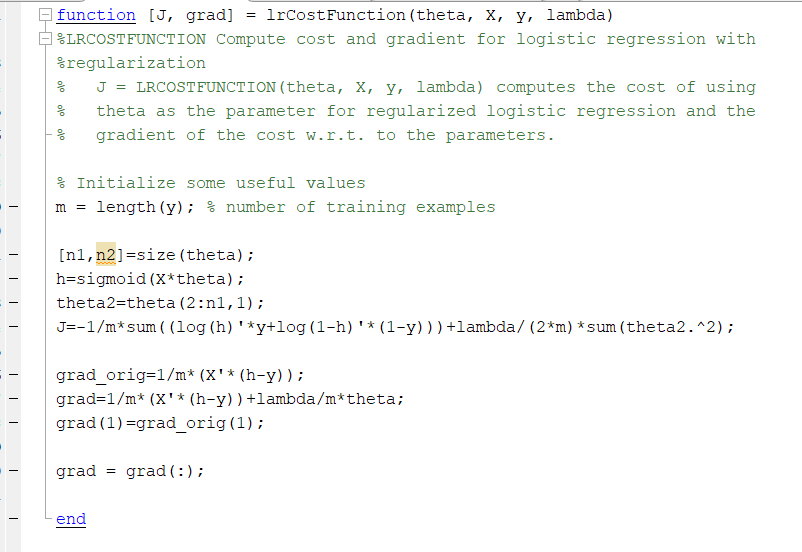
Загрузите данные **ex2data3.mat** из файла.



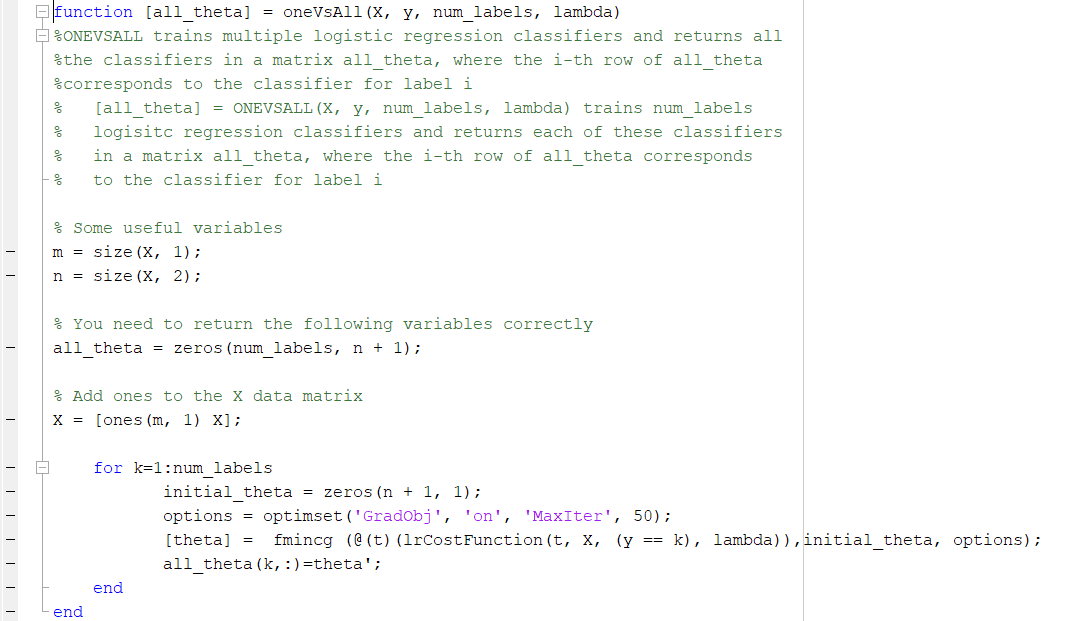
Визуализируйте несколько случайных изображений из набора данных. Визуализация должна содержать каждую цифру как минимум один раз.



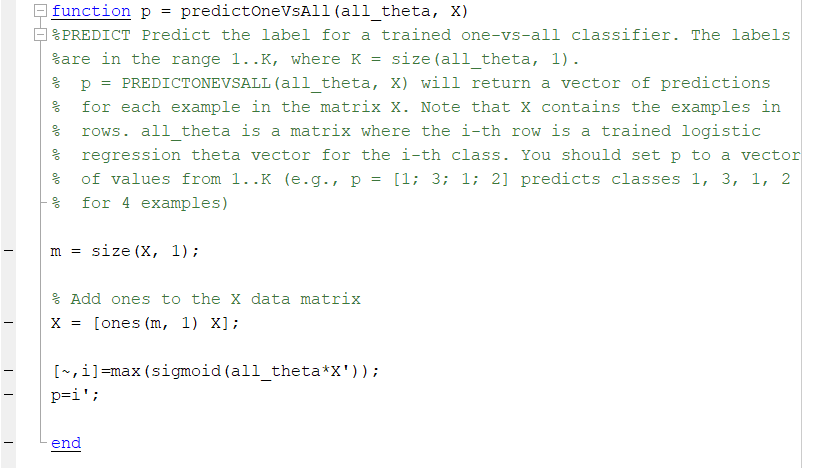
Реализуйте бинарный классификатор с помощью логистической регрессии с использованием векторизации (функции потерь и градиентного спуска). Добавьте L2-регуляризацию к модели.



Реализуйте многоклассовую классификацию по методу “один против всех”.



Реализуйте функцию предсказания класса по изображению с использованием обученных классификаторов.



Процент правильных классификаций на обучающей выборке должен составлять около 95%.

