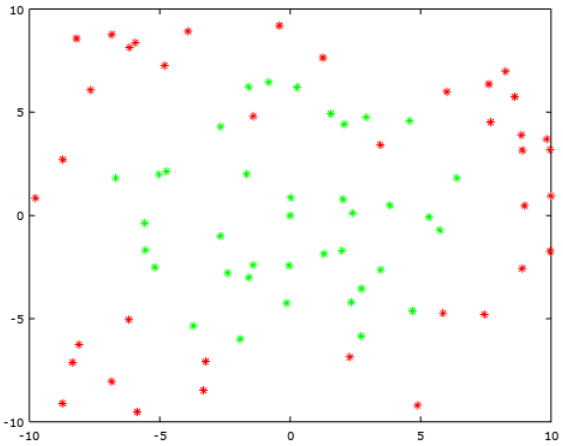
EXERCISE 5

DR. VICTOR UC CETINA

1. Binary Classification through AdaBoost



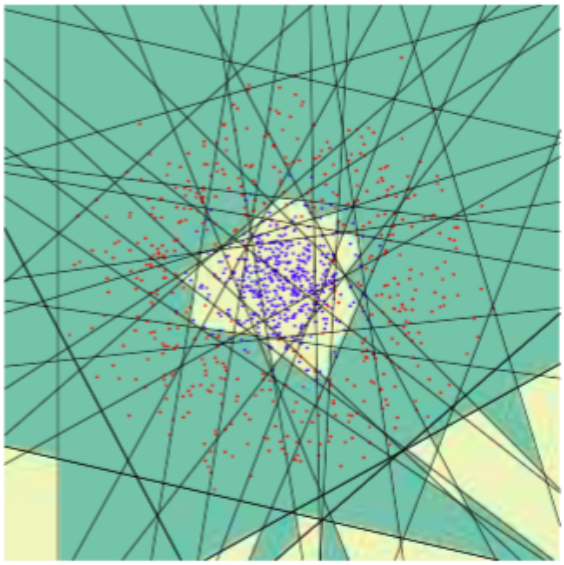
(1) Download the data file "dataCircle.mat" (or "dataCircle.txt" if you are not using matlab) which contains a matrix of size 103×3. The first 40 rows are positive examples (label 1) of points in 2 dimensions. The last 62 rows are negative examples (label 0) also in 2 dimensions.
(2) Implement in your favorite programming language the AdaBoost algorithm using as weak classifiers the Logistic Regression algorithm, so that you classify correctly both types of data.
(3) Plot the data points using one color for each class of data. Also, plot the classifier line that you found using logistic regression.
(4) Prepare a report containing your final model (including parameters), and your graph.

2. Exercise Submission

- Send your report to `cetina@informatik.uni-hamburg.de` with the email title string: MLEX5 Lastname1-Lastname2-Lastname3
- Deadline: A day before your next exercise session.
- Note: Do not forget to include your names in the report!

3. Hints

- Negative examples: replace every label 0 by label -1
- Weak classifiers: use linear models of the form $z = \theta_0 + \theta_1 x_1 + \theta_2 x_2$ (straight lines).
- Each weak classifier should be evaluated in the sigmoid function $g(\mathbf{x}) = \frac{1}{1+e^{-z}}$.
- The kind of strong classifier that we are looking for is similar to this one:



Note: This solution belongs to another data set.