Ch10 - Neural Network

Contents

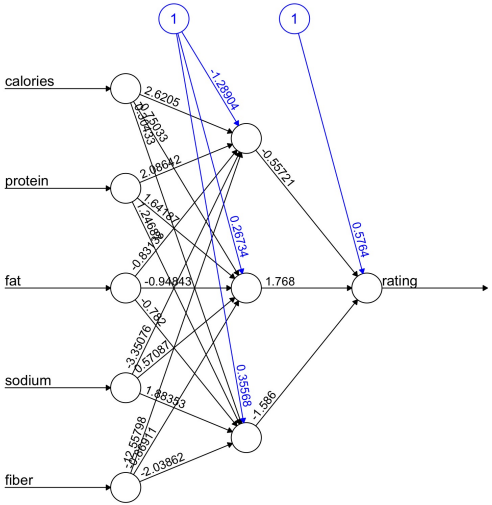
Textbook: James et al. ISLR 2ed.

10A Subsections

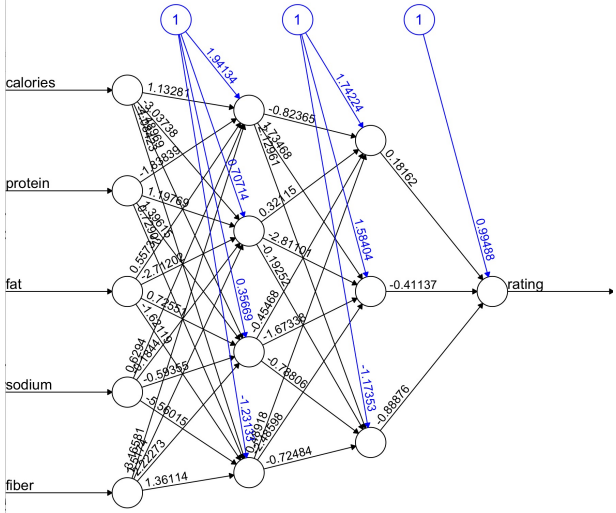
[ToC]

A.1 Neural Network

- First developed in 40's and 50's
- First implementation in 1954 at MIT
- Inspired by how human brain cells are connected.



Error: 0.064839 Steps: 311



Error: 0.096123 Steps: 57

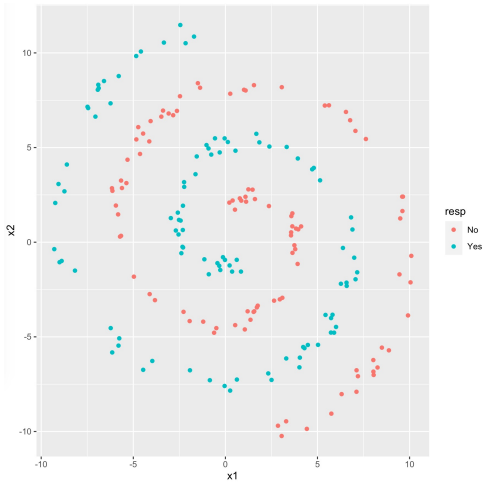
A.2 Forward and Backward Propagation

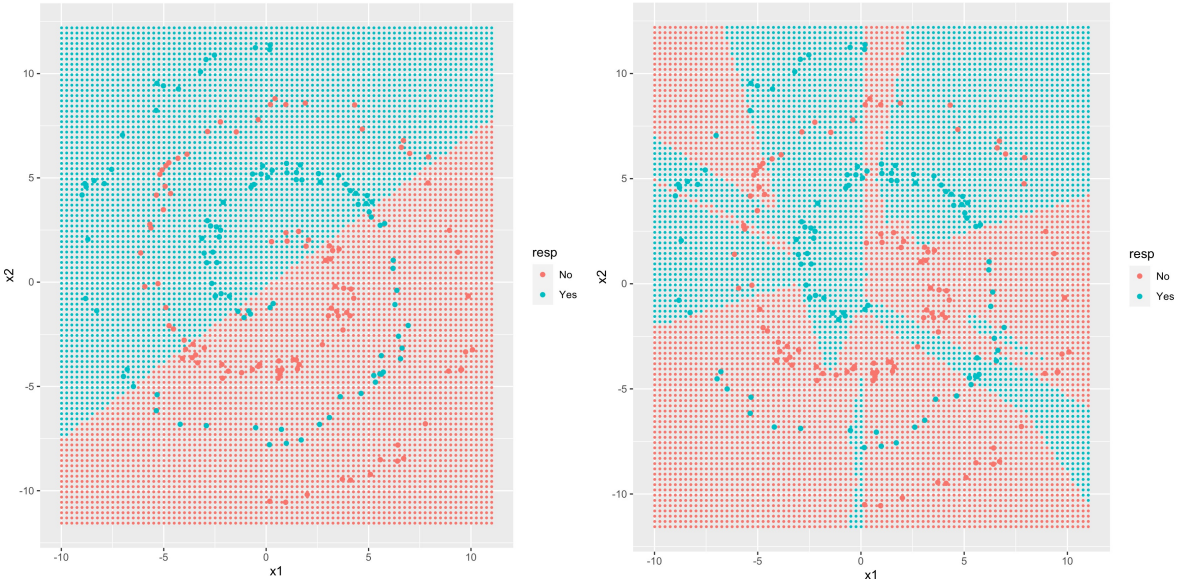
A.3 Gradient Descent vs Newton-Raphson

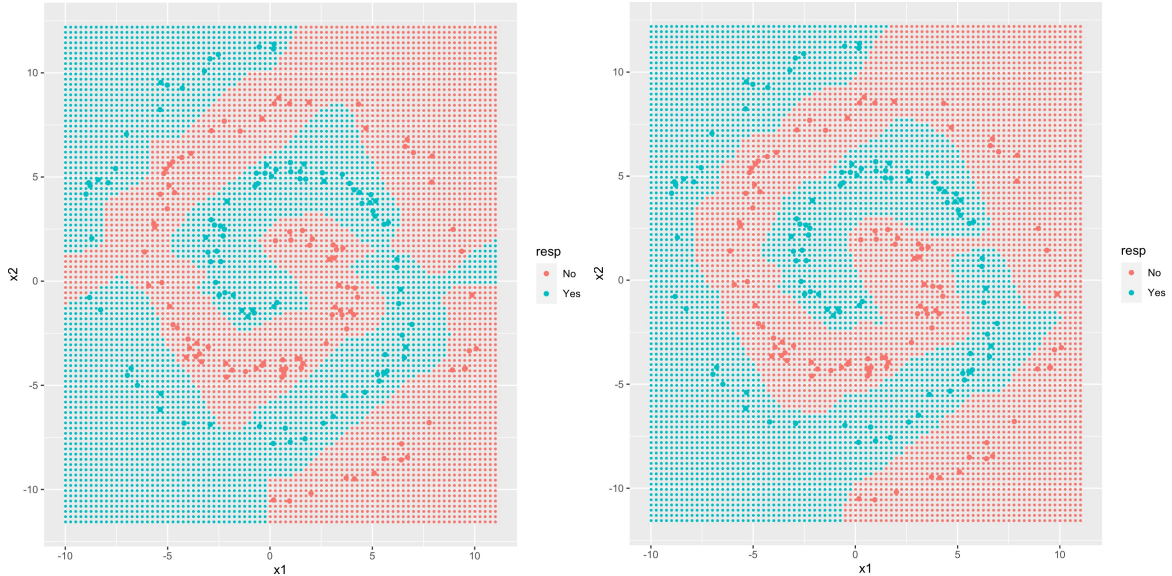
- Gradient Descent (1st order method)
- Newton-Raphson (2nd order method)

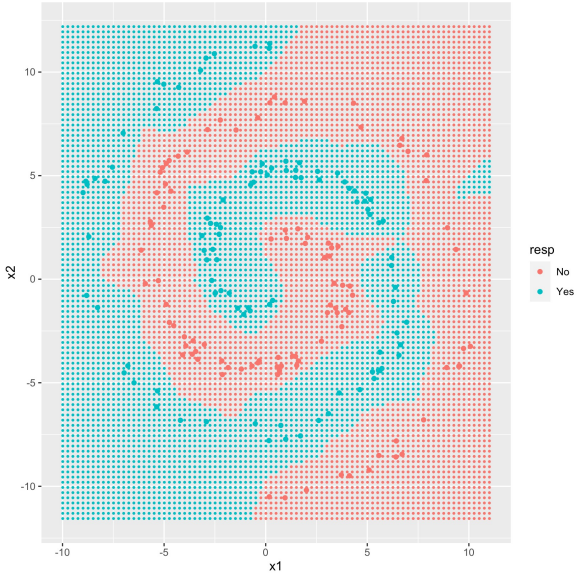
A.4 Spiral Classification

Suppose this is your data:

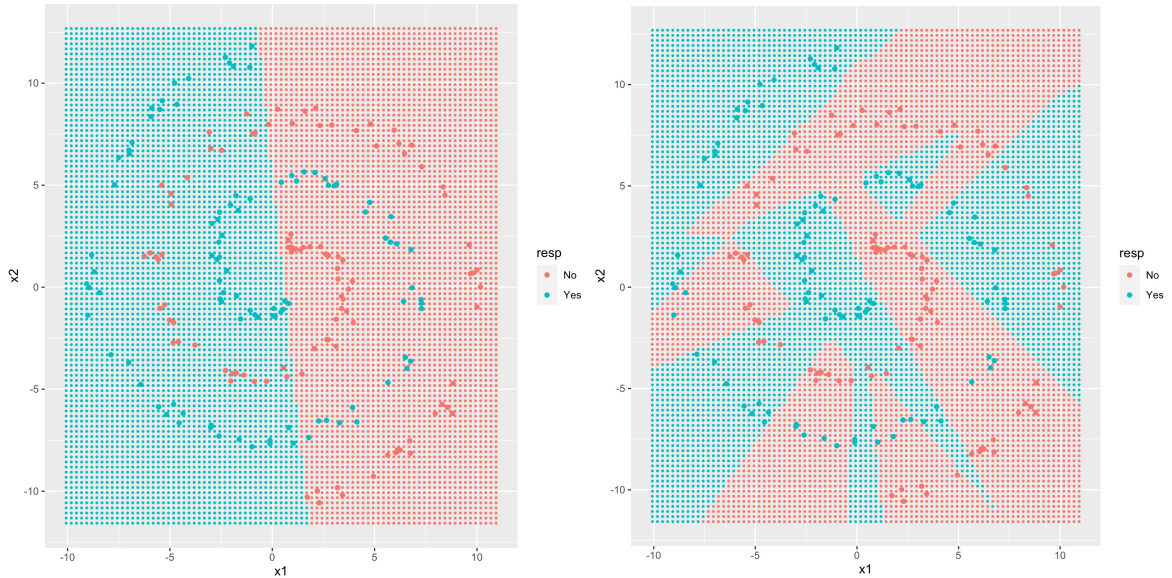


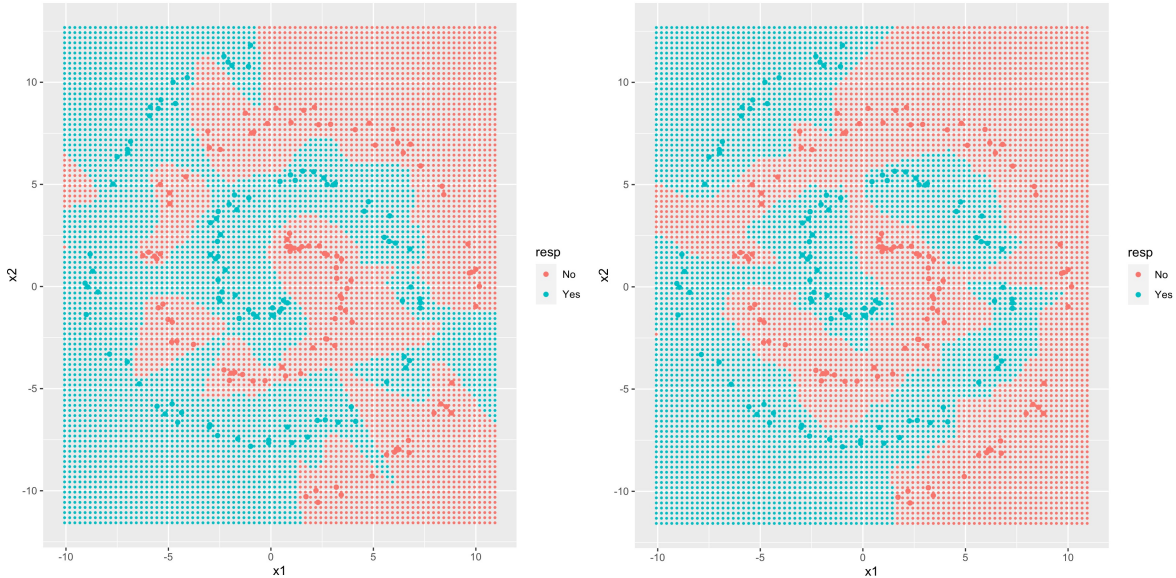

10
5
0
-5
-10
x2
-10
-5
0
5
10
x1
resp
No
Yes

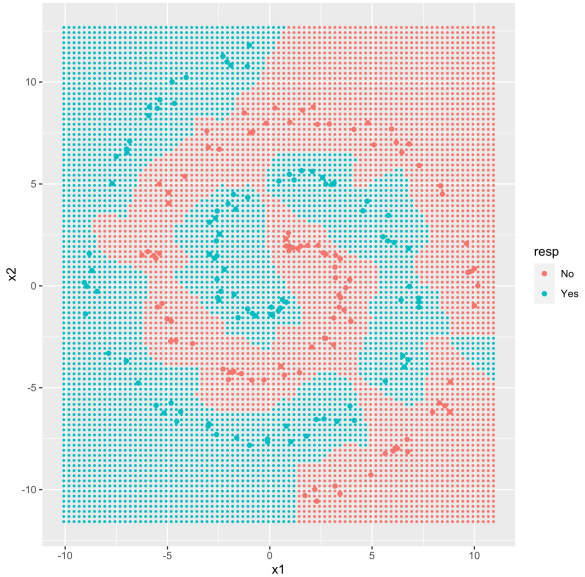

x2
10
5
0
-5
-10
-10
-5
0
5
10
x1
resp
No
Yes
x2
10
5
0
-5
-10
-10
-5
0
5
10
x1
resp
No
Yes


10
5
0
-5
-10
x2
-10
-5
0
5
10
x1
resp
No
Yes

A.5 Different Seed




x2
10
5
0
-5
-10
-10
-5
0
5
10
x1
resp
No
Yes
x2
10
5
0
-5
-10
-10
-5
0
5
10
x1
resp
No
Yes


10
5
0
-5
-10
x2
-10
-5
0
5
10
x1
resp
No
Yes

A.6 Data Prep

- All input must be numeric
- Data must be scaled

A.7 Pros and Cons

- Can do both regression and classification.
- Very flexible. can be trained with any number of inputs and layers.
- Any data which can be made numeric can be used in the model.

  Neural networks are good to model with nonlinear data with large number of inputs; for example, images. It is reliable in an approach of tasks involving many features. It works by splitting the problem of classification into a layered network of simpler elements.
- Once trained, the predictions are pretty fast.
- Often need large training data.

- No interpretability. (Black Box)
- Computationally very expensive. (Graphic Cards)

- Over-fitting and generalization. Must be very careful in tuning.