

Project 2: Classification with Logistic Regression (10% of final grade)

Due: Before midnight on March 4 (-3⁽ⁿ⁻¹⁾ for late programs)

For this project we will apply Logistic Regression to the Divorce Predictors Scale to classify whether married couples will eventually divorce. Each of the questions in the appendix were answered on a scale of 0 (Strongly Disagree) to 4 (Strongly Agree). Our data set has 54 questions (features, n) and 170 sets of data (m). Your data file is formatted with the first line containing m and n , tab separated). Then each line following has 55 entries, the first 54 are the responses to the questions below (0-4) followed by a y coded as either 0 (divorce class) or 1 (did not divorce class). All data is tab separated.

Your assignment is to use what you have learned in class and the notes to create (from scratch in Python, not using a Logistic Regression library function!) a binary classifier to predict whether a person will divorce (0) or stay married (1). You are free to use any model variation and any testing or training approach we have discussed for logistic regression.

What to Turn In Via Canvas

1. **A single py file (*lastname_firstname_P2.py*)** that prompts for a training file name, computes weights using gradient descent, prints out a plot of iterations vs. J , and then prompts for a test file name, and using the computed weights prints out final J , accuracy, FP, FN, TP, TN, accuracy, precision, recall and F1 for the test set. All values should be clearly labelled.
2. **A pdf file (*lastname_firstname_P2.pdf*)** that includes
 - A description of your model and testing procedure, including
 - Size of your training and test sets.
 - Initial values that you chose for your weights, alpha, and the initial value for J .
 - Final values for alpha, your weights, how many iterations your learning algorithm went through and your final value of J on your training set.
 - Include a plot of J (vertical axis) vs. number of iterations (horizontal axis).
 - If you did feature scaling, describe what you did.
 - Value of J on your test set.
 - A confusion matrix showing your results on your test set.
 - A description of your final results that includes accuracy, precision, recall and F1 values.

Note that we have covered many possible tools you could use such as plotting, model selection with training, validation and test sets, feature scaling, and/or regularization. However, a simple linear approach that works is better than banging your head against your desk because a more sophisticated approach is not working!

Do NOT zip, tar or feather your files! Upload them to canvas with the proper names.

Appendix

1. If one of us apologizes when our discussion deteriorates, the discussion ends.
2. I know we can ignore our differences, even if things get hard sometimes.
3. When we need it, we can take our discussions with my spouse from the beginning and correct it.
4. When I discuss with my spouse, to contact him will eventually work.
5. The time I spent with my wife is special for us.
6. We don't have time at home as partners.
7. We are like two strangers who share the same environment at home rather than family.
8. I enjoy our holidays with my wife.
9. I enjoy traveling with my wife.
10. Most of our goals are common to my spouse.
11. I think that one day in the future, when I look back, I see that my spouse and I have been in harmony with each other.
12. My spouse and I have similar values in terms of personal freedom.
13. My spouse and I have similar sense of entertainment.
14. Most of our goals for people (children, friends, etc.) are the same.
15. Our dreams with my spouse are similar and harmonious.
16. We're compatible with my spouse about what love should be.
17. We share the same views about being happy in our life with my spouse
18. My spouse and I have similar ideas about how marriage should be
19. My spouse and I have similar ideas about how roles should be in marriage
20. My spouse and I have similar values in trust.
21. I know exactly what my wife likes.
22. I know how my spouse wants to be taken care of when she/he sick.
23. I know my spouse's favorite food.
24. I can tell you what kind of stress my spouse is facing in her/his life.
25. I have knowledge of my spouse's inner world.
26. I know my spouse's basic anxieties.
27. I know what my spouse's current sources of stress are.
28. I know my spouse's hopes and wishes.
29. I know my spouse very well.
30. I know my spouse's friends and their social relationships.
31. I feel aggressive when I argue with my spouse.
32. When discussing with my spouse, I usually use expressions such as 'you always' or 'you never'.
33. I can use negative statements about my spouse's personality during our discussions.
34. I can use offensive expressions during our discussions.
35. I can insult my spouse during our discussions.
36. I can be humiliating when we discussions.
37. My discussion with my spouse is not calm.
38. I hate my spouse's way of open a subject.
39. Our discussions often occur suddenly.
40. We're just starting a discussion before I know what's going on.
41. When I talk to my spouse about something, my calm suddenly breaks.

42. When I argue with my spouse, I only go out and I don't say a word.
43. I mostly stay silent to calm the environment a little bit.
44. Sometimes I think it's good for me to leave home for a while.
45. I'd rather stay silent than discuss with my spouse.
46. Even if I'm right in the discussion, I stay silent to hurt my spouse.
47. When I discuss with my spouse, I stay silent because I am afraid of not being able to control my anger.
48. I feel right in our discussions.
49. I have nothing to do with what I've been accused of.
50. I'm not actually the one who's guilty about what I'm accused of.
51. I'm not the one who's wrong about problems at home.
52. I wouldn't hesitate to tell my spouse about her/his inadequacy.
53. When I discuss, I remind my spouse of her/his inadequacy.
54. I'm not afraid to tell my spouse about her/his incompetence.