Project 2 Writeup

Andrea Bajcsy Michelle Cody Charles Parker

04/20/2016

WU1 (A) The words most indicative of being Sauvignon-Blanc are "citrus", "lime", and "grapefruit" for OAA and "citrus", ""crisp", and ""lime" for AVA. The words most indicative of not being Sauvignon-Blanc are "apple" and "flavors" for OAA and "enjoy", "warm", and "apple" for AVA.

The words most indicative of being Pinot-Noir are "cherry" for OAA and "acidity" for AVA. The words most indicative of not being Pinot-Noir are "cassis" and "raspberries" for OAA and 'cassis" and "crisp" for AVA.

(B) The OAA accuracy is 37.29%. The training time is 0.372 seconds. The AVA accuracy is 26.15%. The training time is 0.387 seconds. The words that suggest Viognier is one of your least favorite wines are "lovely" and "enjoy" since the presence of these words indicate that the wine is not Viognier.

(C) The accuracy for OAA using zero/one predictions is 24.30%. The accuracy for AVA using zero/one predictions is 26.35%. The OAA accuracy is much worse than the confidence predictions. Using confidence improves the accuracy by about 50%. The AVA accuracy is roughly the same for zero/one predictions and confidence predictions. The difference could easily be explained by the random choices made by sklearn.

WU2 The test accuracy you get with a balanced tree on the WineData using a DecisionTreeClassifier with max depth 3 is 30.89%.

WU3 Negative values of the step size cause the algorithm to diverge (see Figure 1 for example with step size -5), 0 is constant, and 0.2 converges. We can see in Figure 1 that 0.5 finds the optimum in one step.

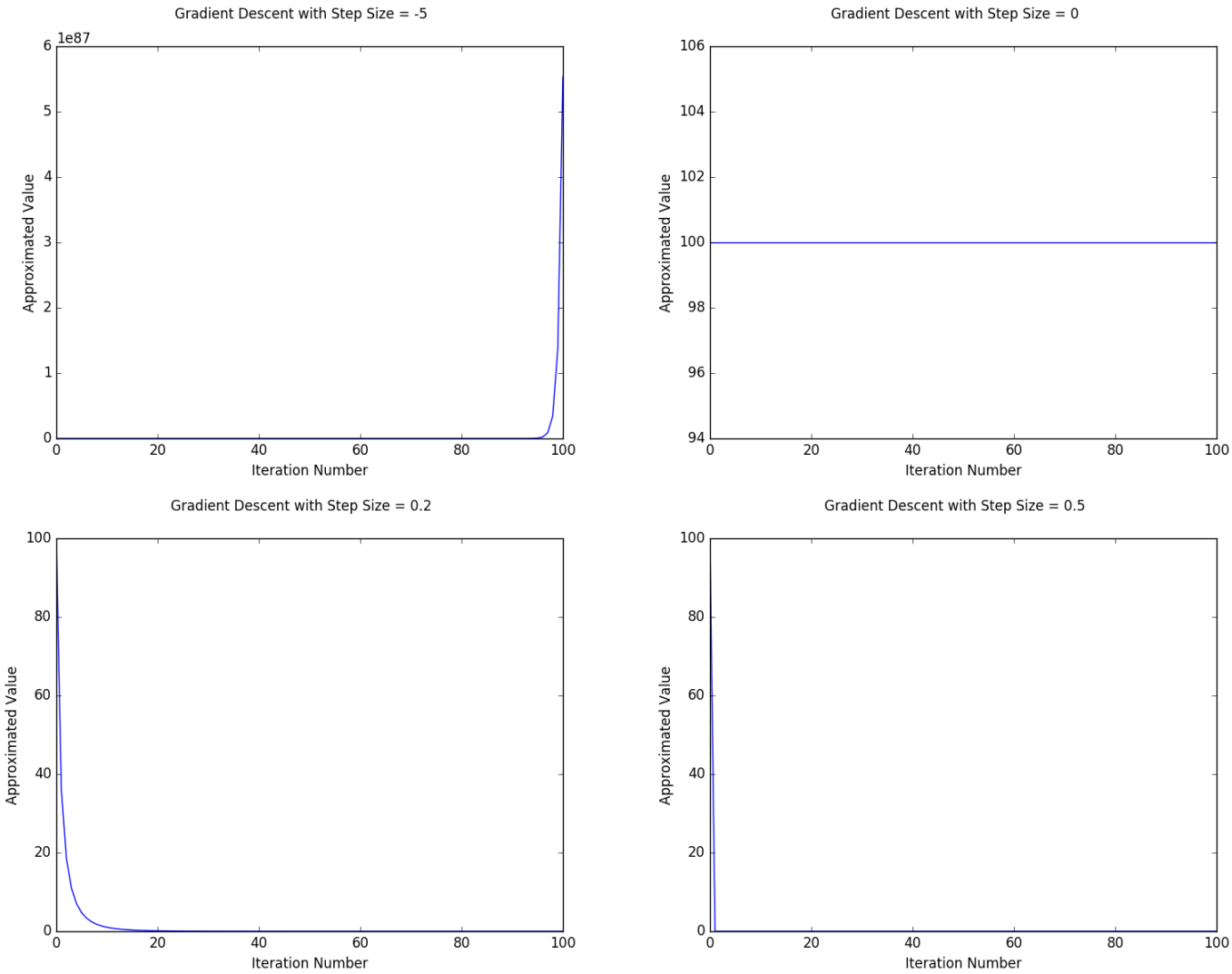

Figure 1: Gradient Descent performed with 100 iterations with step sizes -5, 0, 0.2, and 0.5.

WU4 Come up with a non-convex univariate optimization problem. Plot the function you're trying to minimize and show two runs of gd, one where it gets caught in a local minimum and one where it manages to make it to a global minimum. (Use different starting points to accomplish this.)

WU5

WU6