

Interestingly, if you look in the middle of the plot, you may notice that some of the points appear to be classified incorrectly--not so according to our data. What's happened is a problem with the scale of our data. The Waiting time variable has more significance than the Eruption time variable due to differences in scale. Don't worry, your algorithm works fine. We'll learn about normalizing and standardizing data soon.