CPSC 383

Enter group UCIDs and names here (if you worked with others) [submit your own copy]

|  |  |
| --- | --- |
| Name | UCID |
|  |  |
|  |  |
|  |  |

Week 11 (10th week of tutorials)

Tutorial 2

**Clustering**

1. Modify the code to set the cluster starting points to

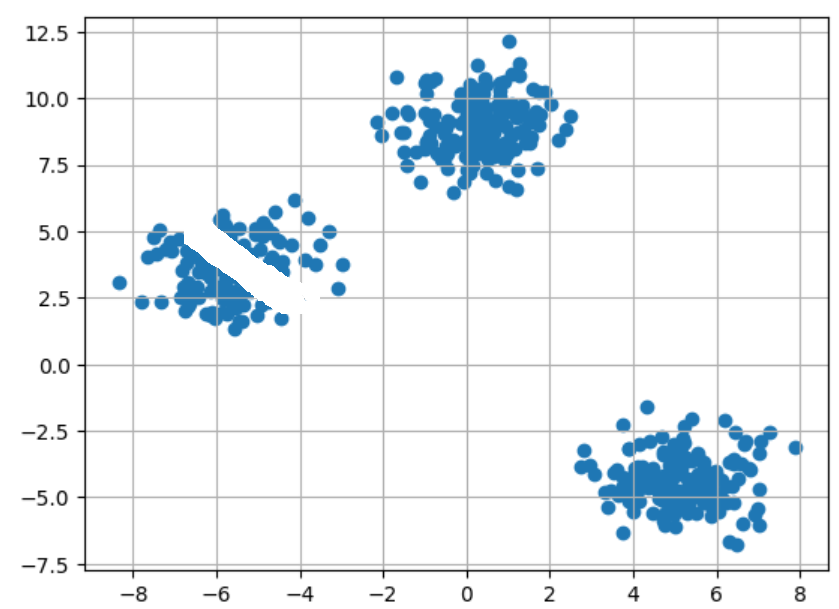
clusters[0]['center'] = (-8, - 2.5)

clusters[1]['center'] = (-6, -5)

clusters[2]['center'] = (-8, -5)

Describe the result.

1. Remove the above center. Change k=4. Describe the result.
2. You don’t need to run code here, but what do you believe would have if you used the code from 2 where k= 4 but there were 4 generated clusters that look like the following image.



Describe the result.

1. Using the X created in the provided code for clustering. Use SciKit to Create a KMeans solution to fit on and predict X. What are the 3 line of code you created?

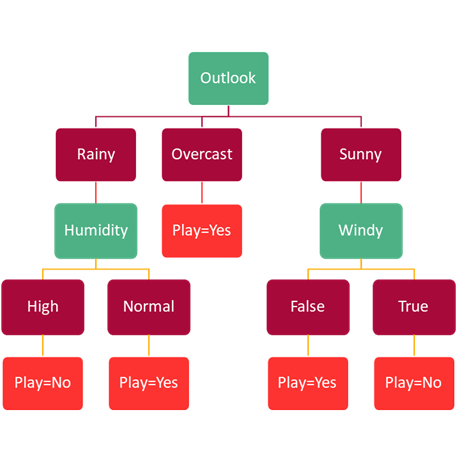
<https://scikit-learn.org/stable/modules/generated/sklearn.cluster.KMeans.html>

**Association Rules**

1. What are all the rules that involve 3 items?
2. What is the highest confidence level of any rule (how many rules share that confidence)?
3. Why might we be not interested in a rule with the highest of confidence (given in 6) versus one of the other rules with 66% confidence instead?
4. If we adjust the threshold to 0.6 what is the highest support we get for any rule?

**Decision Trees**

1. Why does SciKit produce a decision tree that is different from the one made in class? (lecture slide you’ll see in class included)



1. What are the 7 rules that match the decision tree produced by the provided code (not the image provided in 9)?