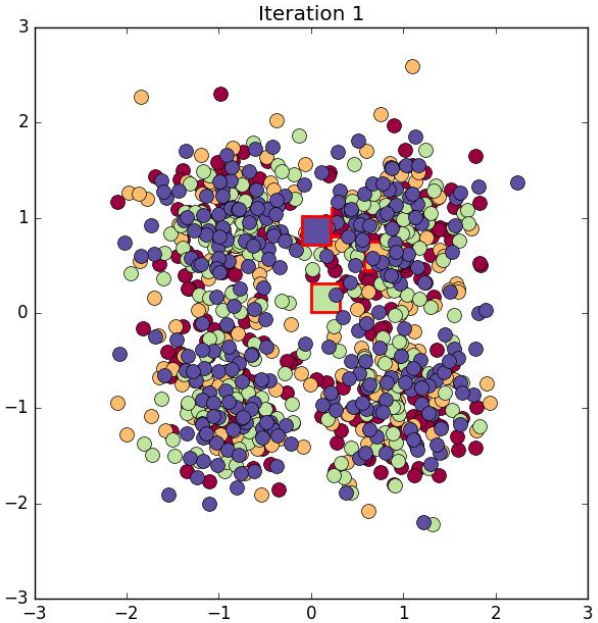
K-Means

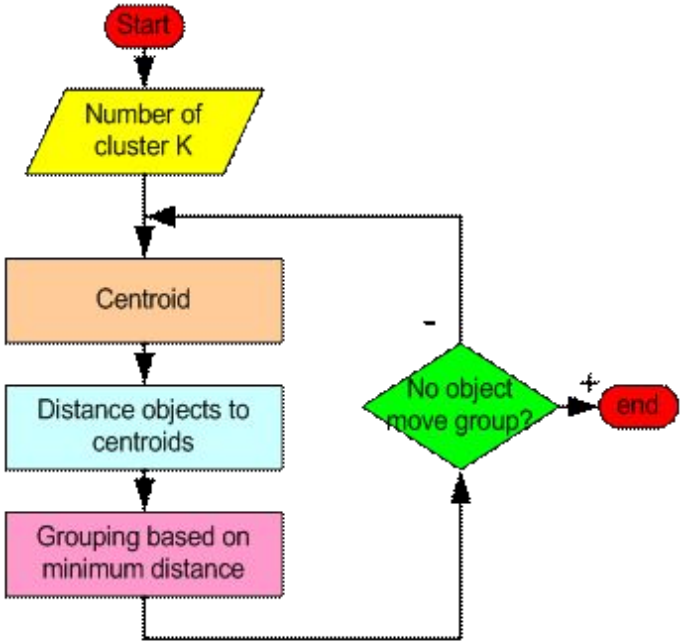
Presentation by Berk Sudan

K-Means Simulation



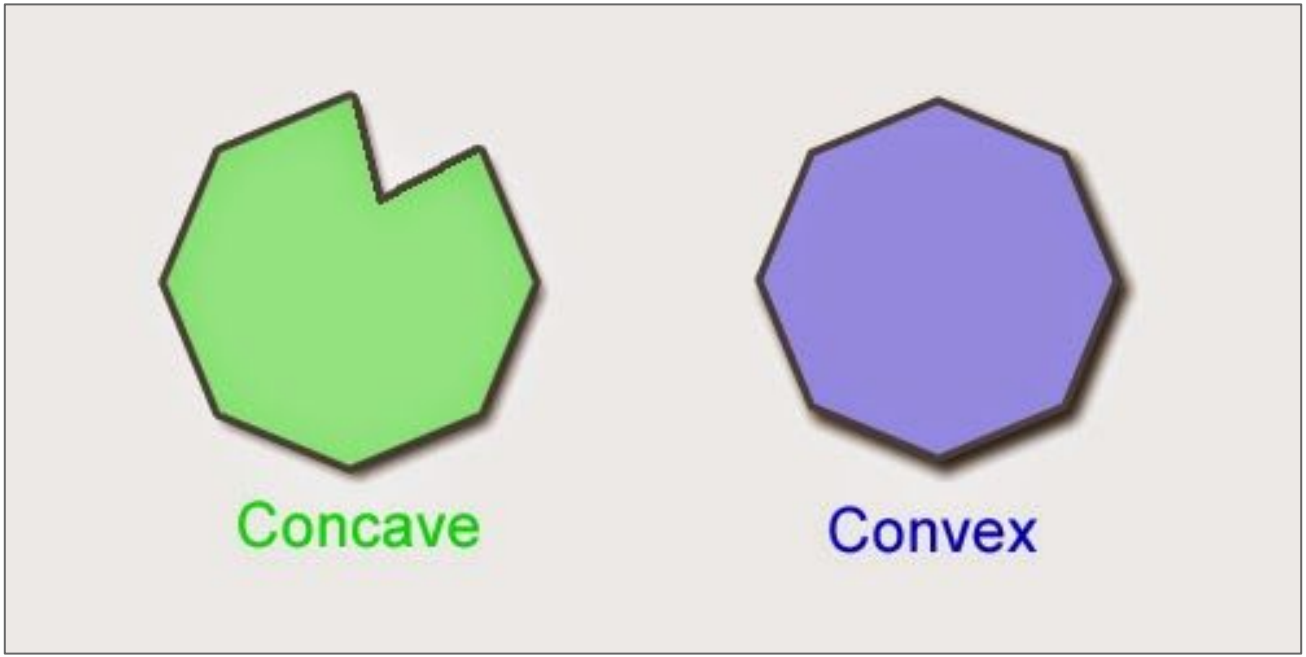
Ref: https://github.com/vinhkhuc/VanillaML

Flow Chart

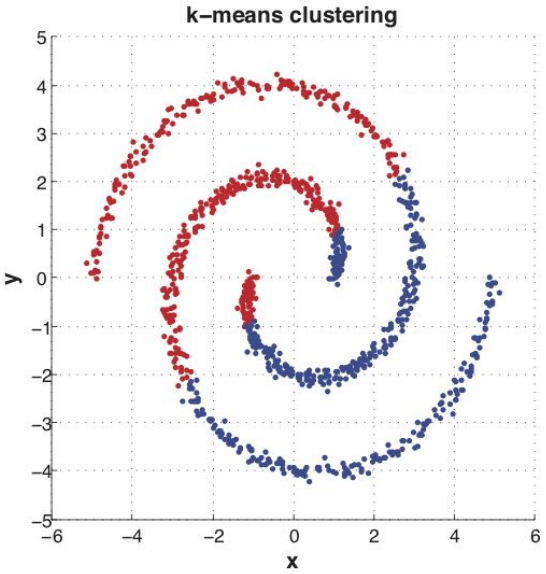


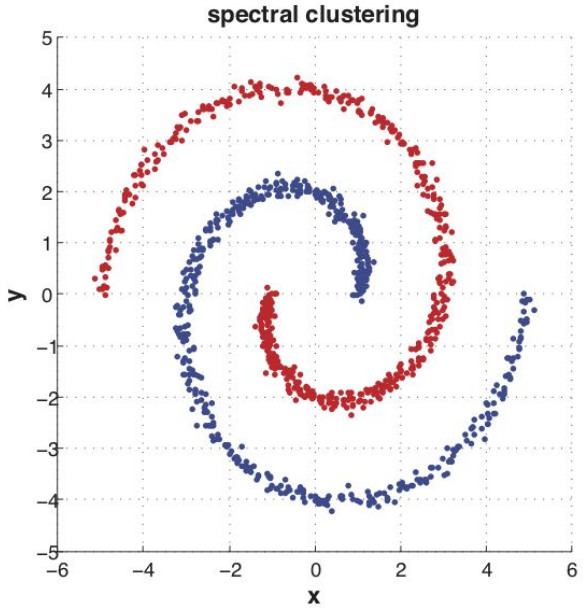
Ref: *Doç. Dr. Songül Varlı, Introduction to Data Mining Lecture Slides, 2018*

Fails on Concave Shapes



Fails on Spectral Shapes





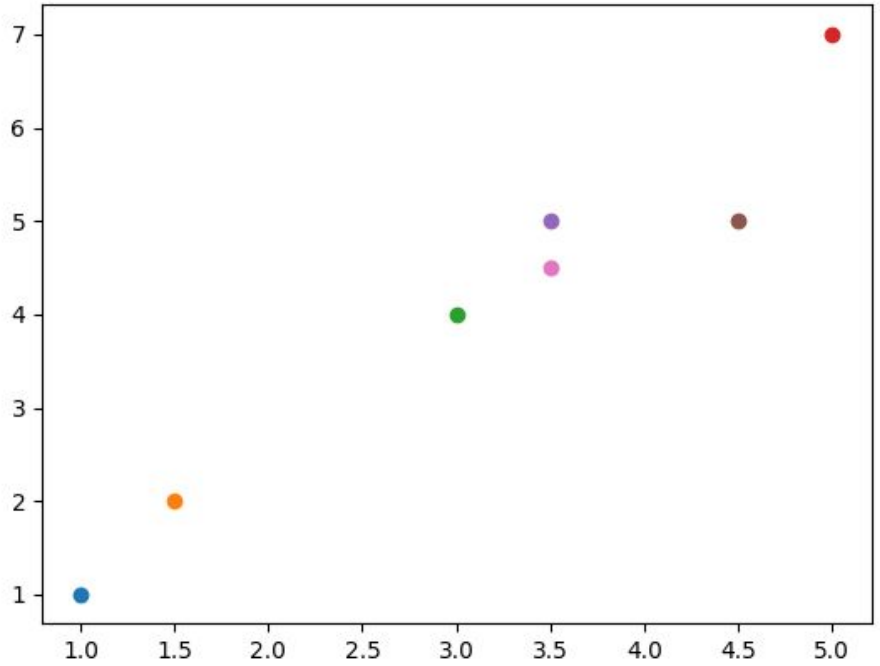
Ref: *Machine Learning – A Probabilistic Perspective, K. P. Murphy, The MIT Press, 2012, syf. 893.*

A Clustering Example

Individual	Variable 1	Variable 2
1	1.0	1.0
2	1.5	2.0
3	3.0	4.0
4	5.0	7.0
5	3.5	5.0
6	4.5	5.0
7	3.5	4.5

Ref: *Doç. Dr. Songül Varlı, Introduction to Data Mining Lecture Slides, 2018*

A Clustering Example



Ref: *Doç. Dr. Songül Varlı, Introduction to Data Mining Lecture Slides, 2018*

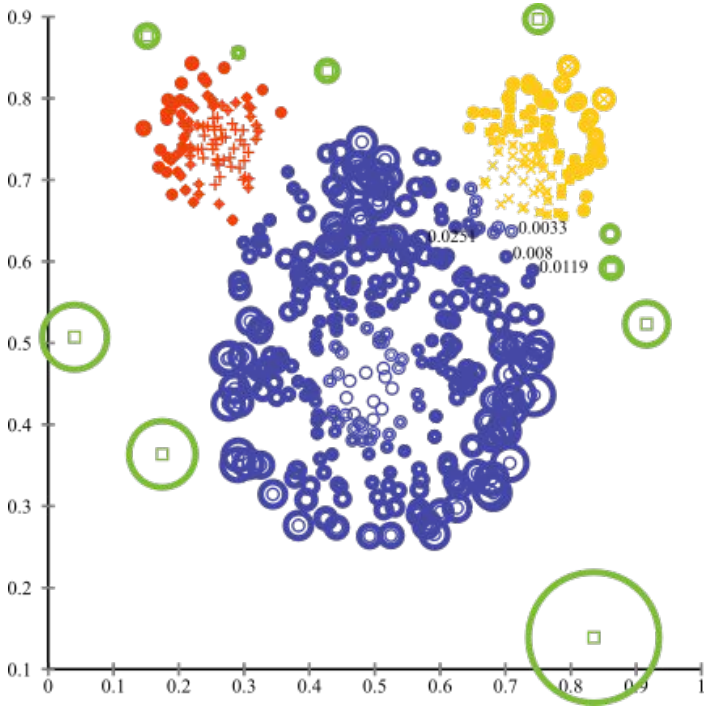
Brainstorming

- What if there is so many **outliers**?

- What if there is so many **clusters**?

- What if there is so many **dimensions/features**?

Many Outliers



Ref: *https://stats.stackexchange.com/questions/160260/anomaly-detection-based-on-clustering*

End of Presentation

Presented by Berk Sudan