Machine Learning
Exercise sheet 4

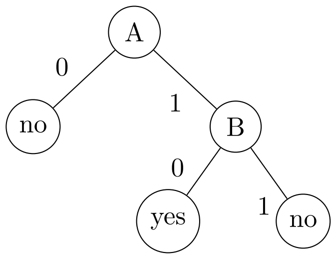
Gruppe 9:
Hauke Wree and Fridtjof Schulte Steinberg

3. Juni 2015

Exercise 1 (Decision trees for Boolean functions):

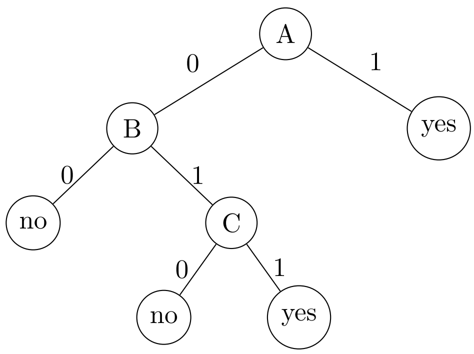
a)

A	B	$A \wedge \neg B$
0	0	0
0	1	0
1	0	1
1	1	0



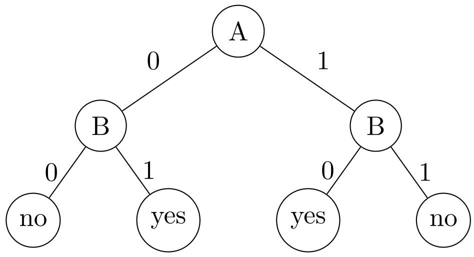
b)

A	B	C	$B \wedge C$	$A \vee [B \wedge B]$
0	0	0	0	0
0	0	1	0	0
0	1	0	0	0
0	1	1	1	1
1	0	0	0	1
1	0	1	0	1
1	1	0	0	1
1	1	1	1	1



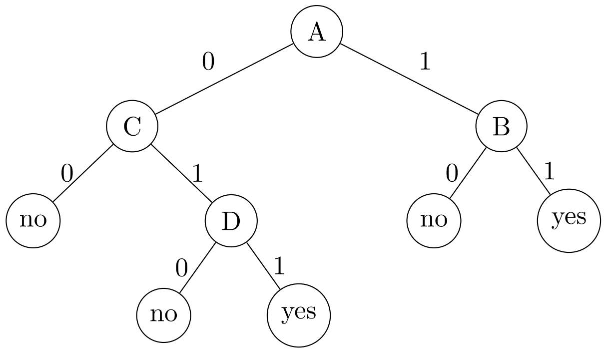
c)

A	B	$AXORB$
0	0	0
0	1	1
1	0	1
1	1	0



d)

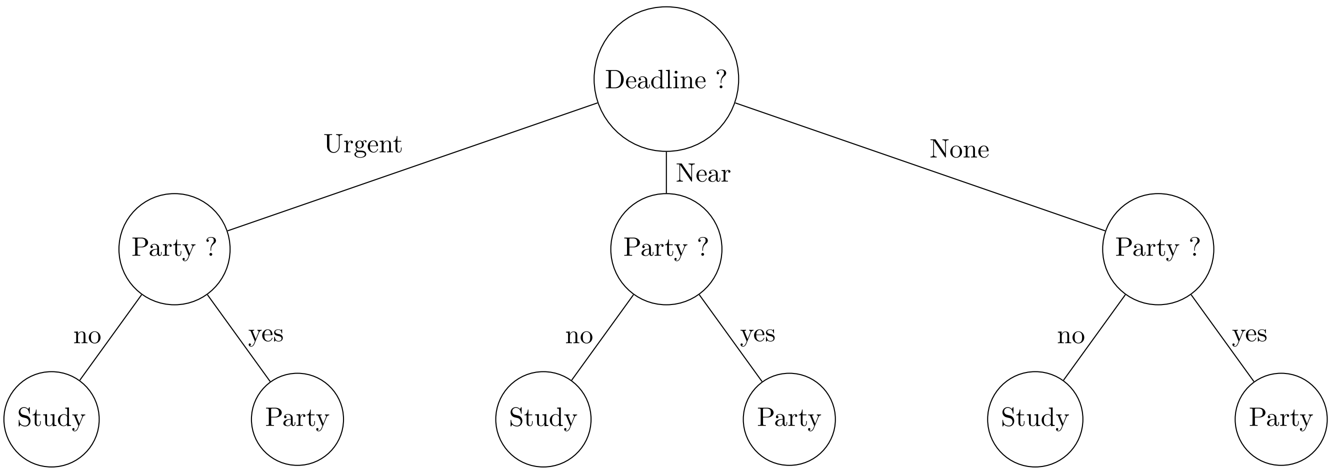
A	B	C	D	$A \wedge B$	$C \wedge D$	$[A \wedge B] \vee [C \wedge D]$
0	0	0	0	0	0	0
0	0	0	1	0	0	0
0	0	1	0	0	0	0
0	0	1	1	0	1	1
0	1	0	0	0	0	0
0	1	0	1	0	0	0
0	1	1	0	0	0	0
0	1	1	1	0	1	1
1	0	0	0	0	0	0
1	0	0	1	0	0	0
1	0	1	0	0	0	0
1	0	1	1	0	1	1
1	1	0	0	1	0	1
1	1	0	1	1	0	1
1	1	1	0	1	0	1
1	1	1	1	1	1	1



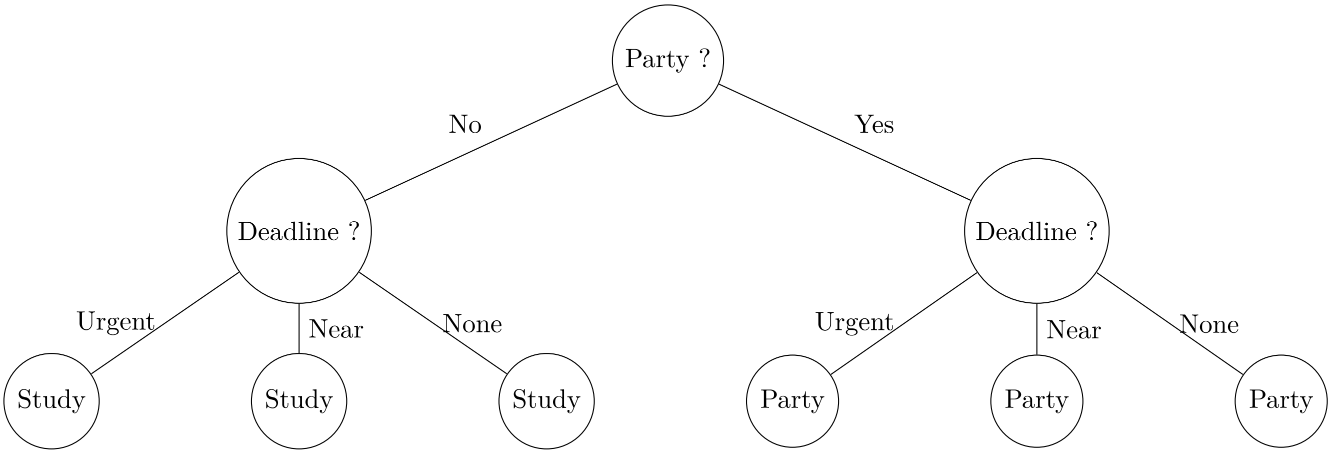
Exercise 2 (Decision tree learning)

a)

i.

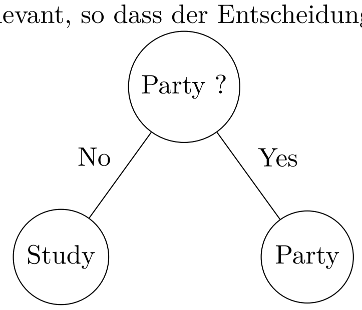


ii.

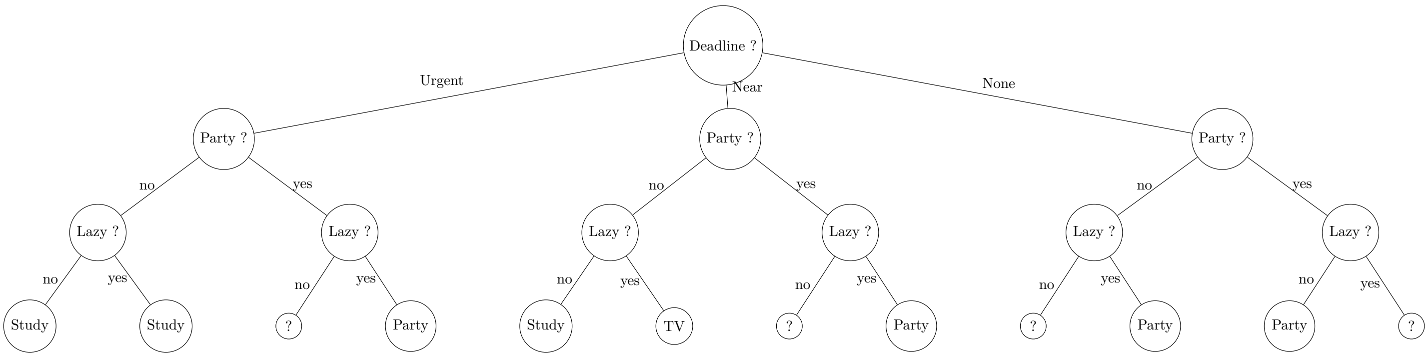


Compare both decision trees.

Der Wert der Deadline ist für das Ergebnis irrelevant nur der Wert von *Is there a party?* ist für das Ergebnis relevant, so dass der Entscheidungsbaum auch so sein kann.



c)



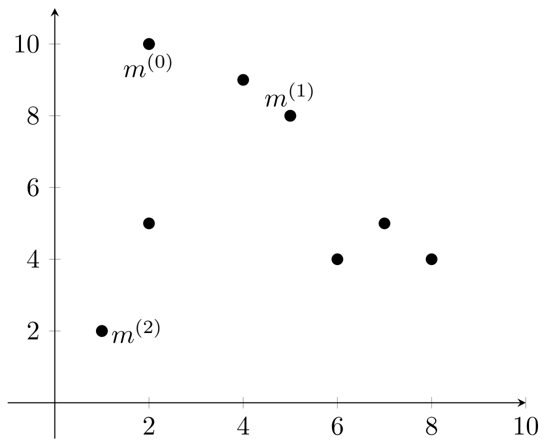
Exercise 3 (Unsupervised learning – k-means clustering):

a)

	$x^{(1)}$	$x^{(2)}$	$x^{(3)}$	$x^{(4)}$	$x^{(5)}$	$x^{(6)}$	$x^{(7)}$	$x^{(8)}$
$x^{(1)}$	0.0	5.0	8.48528137424	3.60555127546	7.07106781187	7.21110255093	8.0622577483	2.2360679775
$x^{(2)}$	5.0	0.0	6.0827625303	4.24264068712	5.0	4.12310562562	3.16227766017	4.472135955
$x^{(3)}$	8.48528137424	6.0827625303	0.0	5.0	1.41421356237	2.0	7.28010988928	6.40312423743
$x^{(4)}$	3.60555127546	4.24264068712	5.0	0.0	3.60555127546	4.12310562562	7.21110255093	1.41421356237
$x^{(5)}$	7.07106781187	5.0	1.41421356237	3.60555127546	0.0	1.41421356237	6.7082039325	5.0
$x^{(6)}$	7.21110255093	4.12310562562	2.0	4.12310562562	1.41421356237	0.0	5.38516480713	5.38516480713
$x^{(7)}$	8.0622577483	3.16227766017	7.28010988928	7.21110255093	6.7082039325	5.38516480713	0.0	7.61577310586
$x^{(8)}$	2.2360679775	4.472135955	6.40312423743	1.41421356237	5.0	5.38516480713	7.61577310586	0.0

1. Wähle k Cluster Mittelwerte hier gegeben.

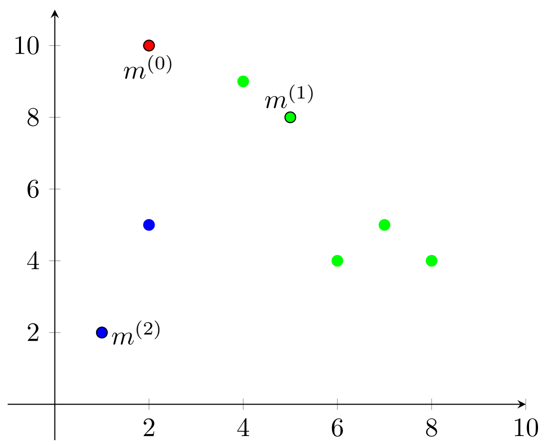
$$x^{(1)} = m^{(0)} = \begin{pmatrix} 7 \\ 4 \end{pmatrix}, x^{(4)} = m^{(1)} = \begin{pmatrix} 3.5 \\ 9 \end{pmatrix}, x^{(7)} = m^{(2)} = \begin{pmatrix} 1.5 \\ 3,5 \end{pmatrix},$$



2. Erstelle cluster

   Ermittel welche Datenpunkte zu welchem Cluster die Datenpunkte gehören.

	$x^{(1)}$	$x^{(2)}$	$x^{(3)}$	$x^{(4)}$	$x^{(5)}$	$x^{(6)}$	$x^{(7)}$	$x^{(8)}$
$m^{(0)}$	**0.0**	5.0	8.48528137424	**3.60555127546**	7.07106781187	7.21110255093	8.0622577483	2.2360679775
$m^{(1)}$	**3.60555127546**	4.24264068712	**5.0**	**0.0**	3.60555127546	**4.12310562562**	7.21110255093	**1.41421356237**
$m^{(2)}$	8.0622577483	**3.16227766017**	7.28010988928	7.21110255093	6.7082039325	5.38516480713	**0.0**	7.61577310586



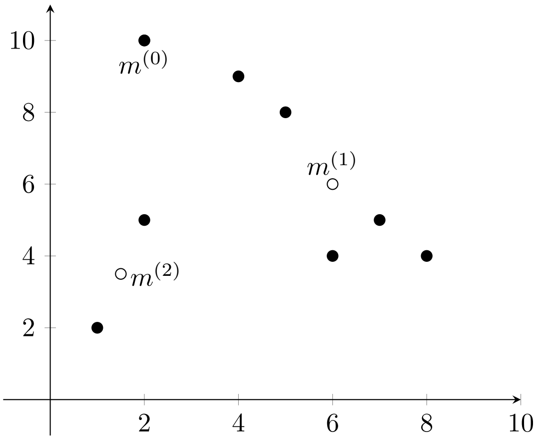
3. Neue Mittelwerte für k Cluster berechnen. Neu Mittelwerte sind:

$$\tilde{m}^{(0)} = \begin{pmatrix} 2 \\ 10 \end{pmatrix}$$

$$\tilde{m}^{(1)} = \frac{1}{5}\left(\begin{pmatrix} 8 \\ 4 \end{pmatrix} + \begin{pmatrix} 5 \\ 8 \end{pmatrix} + \begin{pmatrix} 7 \\ 5 \end{pmatrix} + \begin{pmatrix} 6 \\ 4 \end{pmatrix} + \begin{pmatrix} 4 \\ 9 \end{pmatrix}\right) = \begin{pmatrix} 6 \\ 6 \end{pmatrix}$$

$$\tilde{m}^{(2)} = \frac{1}{2}\left(\begin{pmatrix} 2 \\ 5 \end{pmatrix} + \begin{pmatrix} 1 \\ 2 \end{pmatrix}\right) = \begin{pmatrix} 3/2 \\ 7/2 \end{pmatrix} = \begin{pmatrix} 1,5 \\ 3,5 \end{pmatrix}$$

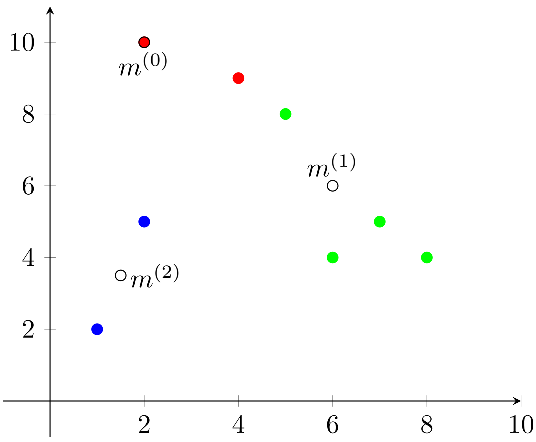
4. Mittelwert einzeichnen



5. Datenpunkte in Cluster einteilen.

	$x^{(1)}$	$x^{(2)}$	$x^{(3)}$	$x^{(4)}$	$x^{(5)}$	$x^{(6)}$	$x^{(7)}$	$x^{(8)}$
$m^{(1)}$	0.0	5.0	8.48528137424	3.60555127546	7.07106781187	7.21110255093	8.0622577483	2.2360679775
$m^{(2)}$	5.65685424949	4.12310562562	2.82842712475	2.2360679775	1.41421356237	2.0	6.40312423743	3.60555127546
$m^{(3)}$	6.5192024052	1.58113883008	6.5192024052	5.7008771255	5.7008771255	4.52769256907	1.58113883008	6.0415229868

6. Datenpunkte einzeichnen

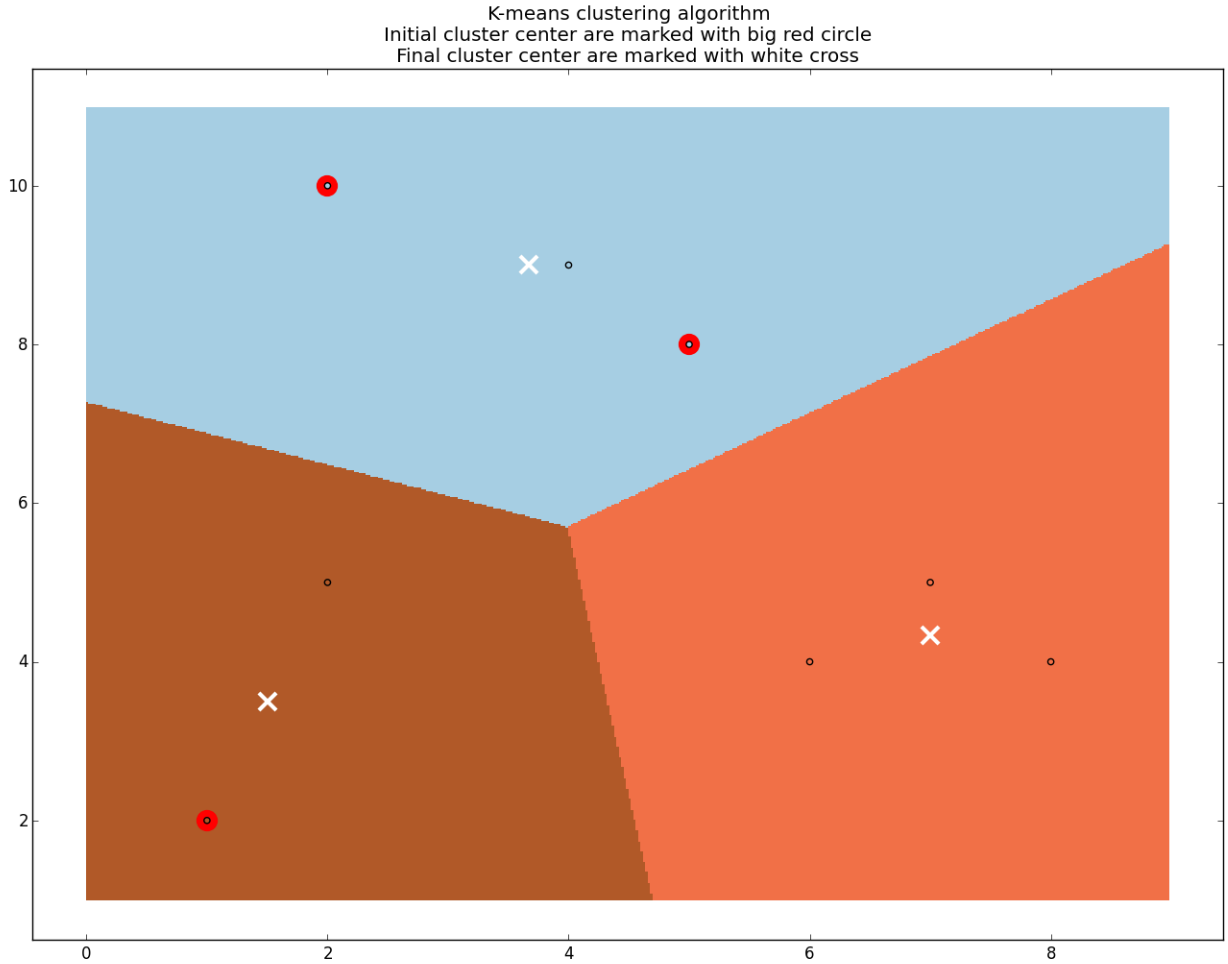


7. Solange fortführen bis Clusterzugehörigkeit der Punkte sich nicht mehr ändern.

b)

```
number of features: 2
number of samples: 8


Iteration: 1
cluster center:
[[  2.   10. ]
 [  6.    6. ]
 [  1.5   3.5]]

labels of training patterns:
[0 2 1 1 1 1 2 1]

Iteration: 2
cluster center:
[[ 3.    9.5 ]
 [ 6.5   5.25]
 [ 1.5   3.5 ]]

labels of training patterns:
[0 2 1 1 1 1 2 0]

Iteration: 3
cluster center:
[[ 3.66666667  9.        ]
 [ 7.          4.33333333]
 [ 1.5         3.5       ]]

labels of training patterns:
[0 2 1 0 1 1 2 0]

Iteration: 4
cluster center:
[[ 3.66666667  9.        ]
 [ 7.          4.33333333]
 [ 1.5         3.5       ]]

labels of training patterns:
[0 2 1 0 1 1 2 0]

max_iter
4
```

Abbildung 5: Erwiterte Ausgabe von exercise3b.py



Das Ergebnis aus a) geht in die richtige Richtung, wenn die Iterationsanzahl in a) höher wäre, wäre das Ergebnis besser. Der minimale Wert von *max_iter* ist 3, wie in Abbildung 5 zu entnehmen ist.

,jedoch wird eine Iterationen mehr benötigt, um festzustellen, dass sich die Mittelwerte nicht ändern.

c)

K-Means ++ Initialisierung:

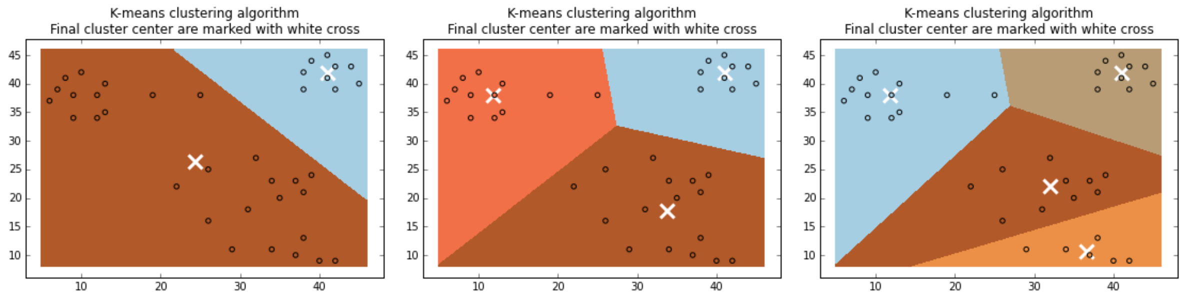


Abbildung 6: Verschiedene Cluster Anzahlen im Vergleich

Um ein gutes Clustering zu erhalten muss die Anzahl der Cluster stimmen. Es ist deutlich zu sehen, dass 3 die ideale Anzahl ist. Im Fall von 2 Clustern werden Punkte zusammengefasst, die nicht direkt zusammengehören und bei 4 Clustern werden Punkte getrennt, welche offensichtlich zusammengehören.

Random Initialisierung:

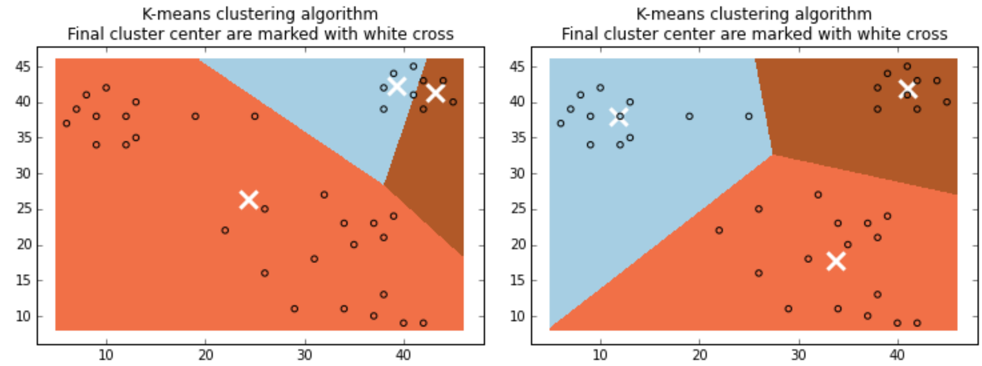


Abbildung 7: np.random.seed(1) und np.random.seed(42) im Vergleich

Die initialen Bedingung scheinen eine große Rolle zu spielen für das Ergebnis des Clusterings. Bei schlechten initialen Zentren erhält meine eine schlechtes Clustering. K-Means-Clustering liefert nur ein lokales Optimum.

Als Folgerung ergibt sich, dass man die richtige Anzahl von Clustern benötigt und vernünftige Startzentren, anderen Falls kann man schlechte Lösungen erhalten.

d)

Die Punkte lassen sich mit dem K-Means-Clustering nicht gut clustern, weil es sich um Punkte handelt, die entlang von Geraden verteilt sind und nicht konzentrisch um einen Punkt.

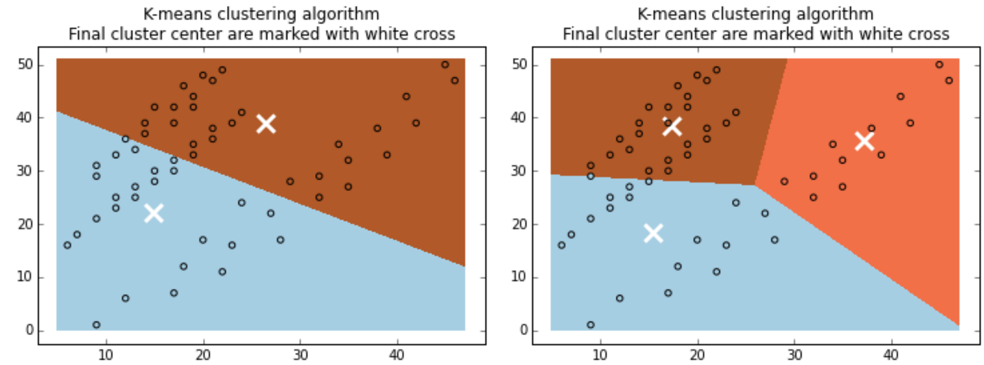


Abbildung 8: Spaeth05 mit 2 und 3 Clustern

e)

Abbildung 9: KMeans-Clustering mit 2 und 16 Farben