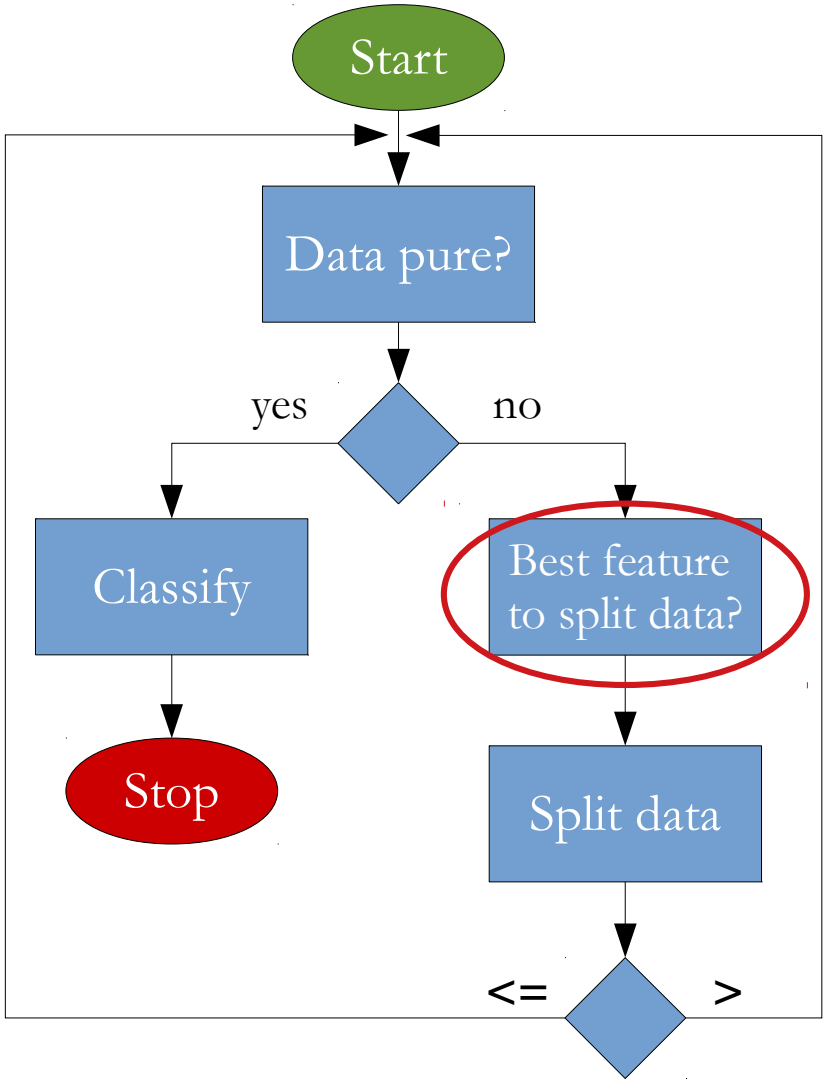
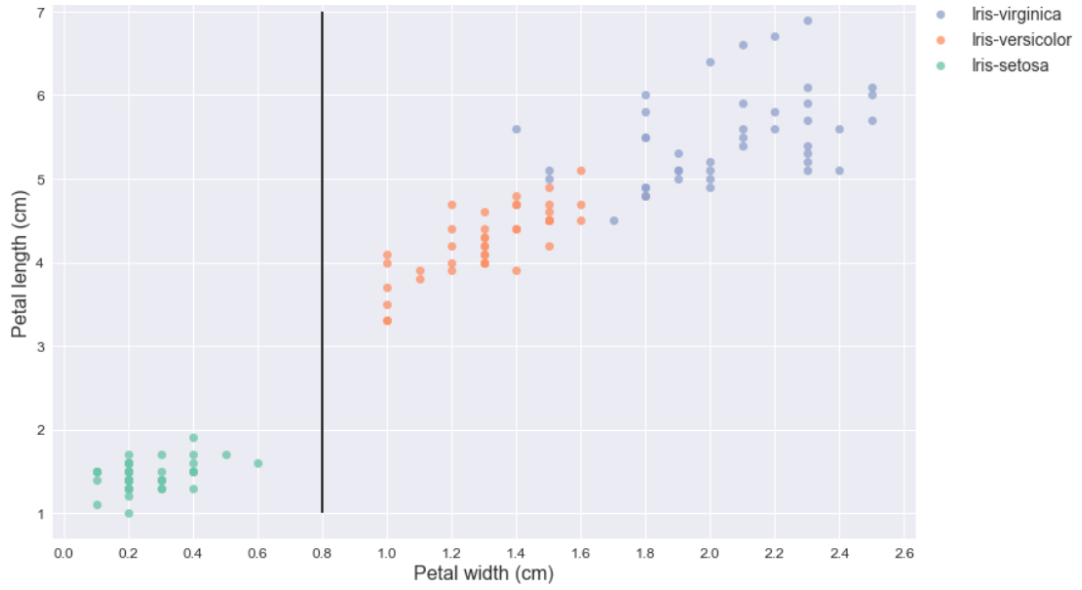
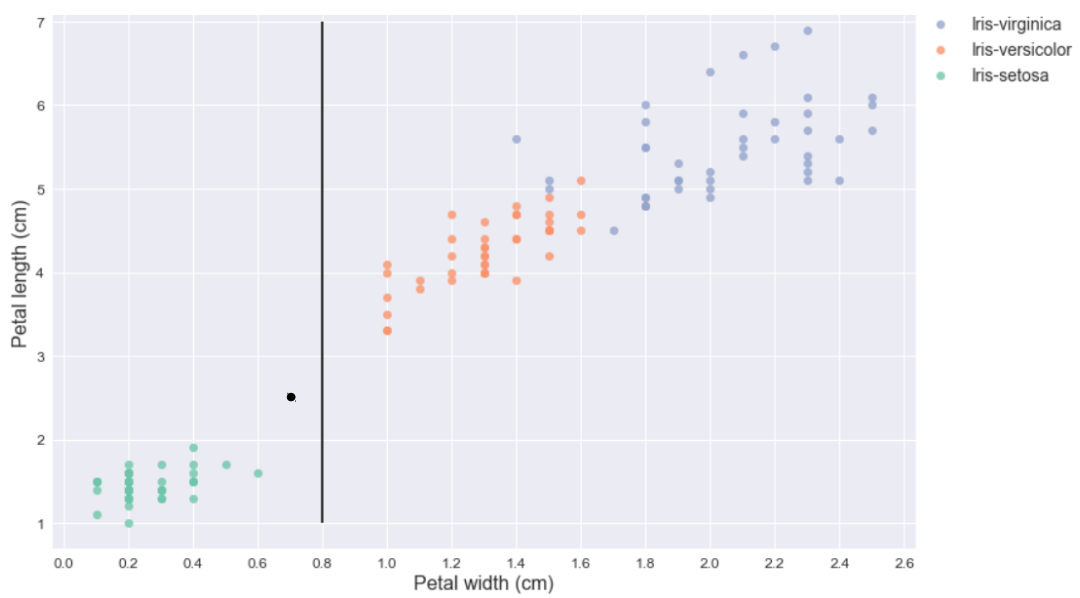
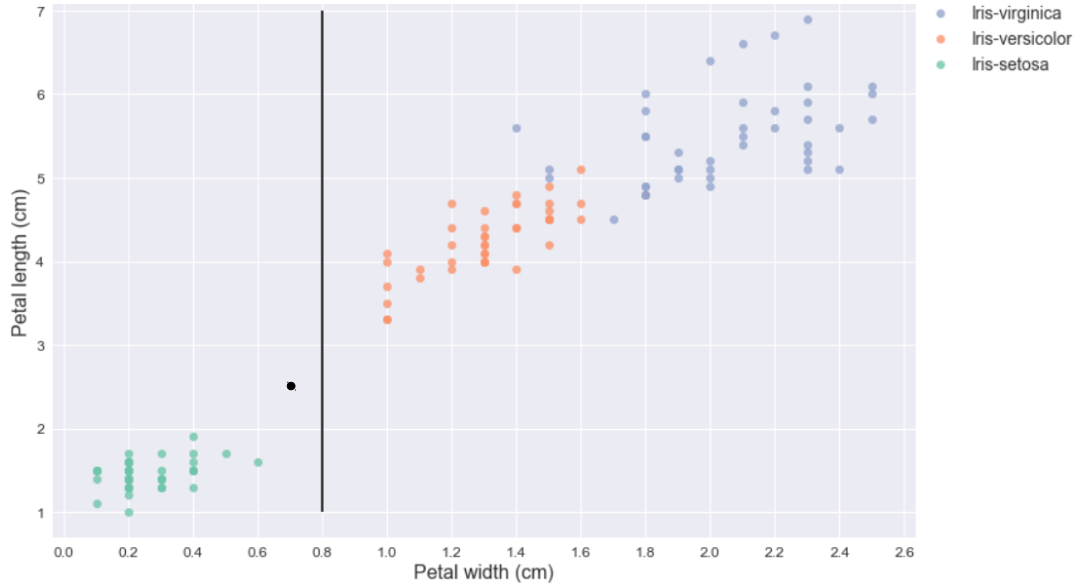
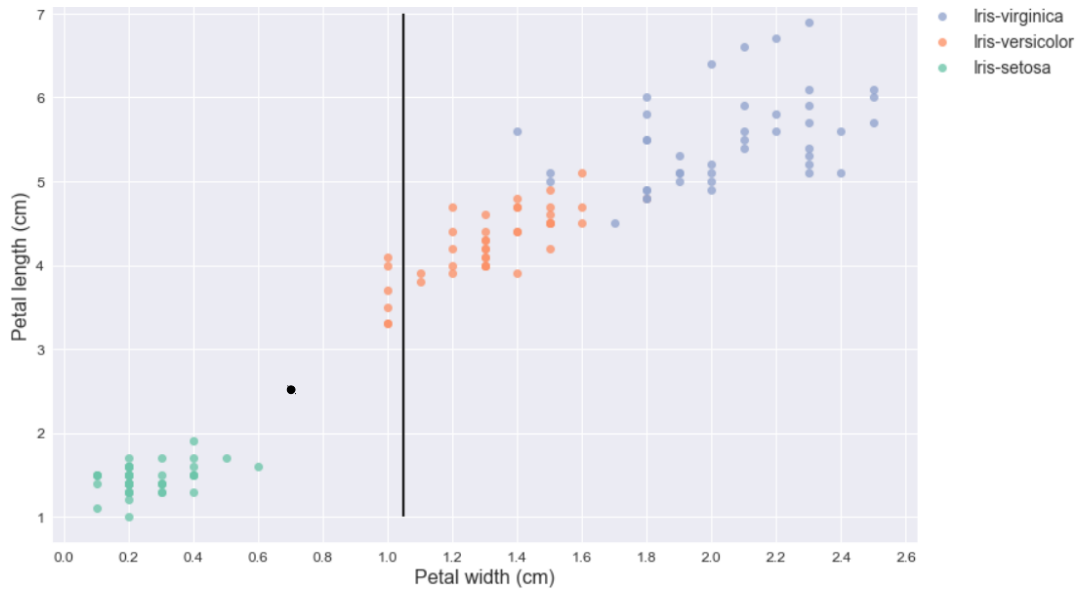
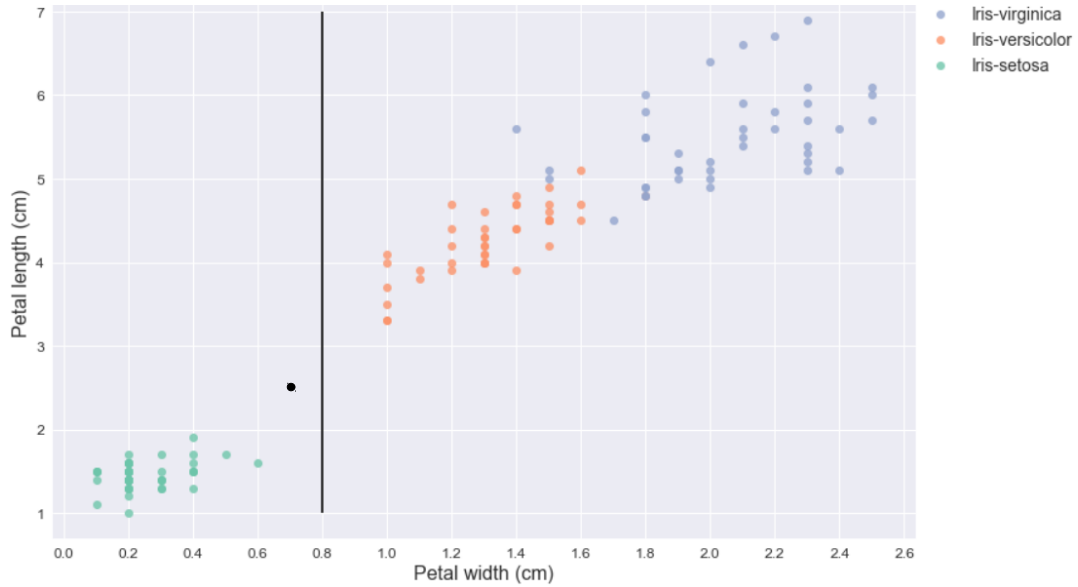

Start
Data pure?
yes
no
Classify
Best feature to split data?
Stop
Split data
<=
>

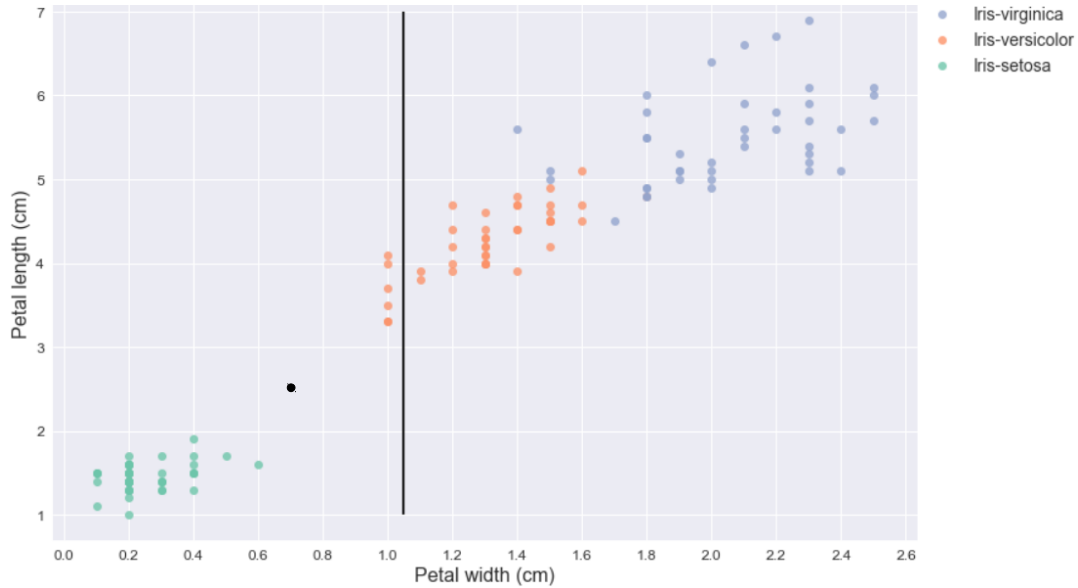

Iris-virginica
Iris-versicolor
Iris-setosa
Petal length (cm)
Petal width (cm)
7
6
5
4
3
2
1
0.0
0.2
0.4
0.6
0.8
1.0
1.2
1.4
1.6
1.8
2.0
2.2
2.4
2.6


Iris-virginica
Iris-versicolor
Iris-setosa
Petal length (cm)
Petal width (cm)
7
6
5
4
3
2
1
0.0
0.2
0.4
0.6
0.8
1.0
1.2
1.4
1.6
1.8
2.0
2.2
2.4
2.6

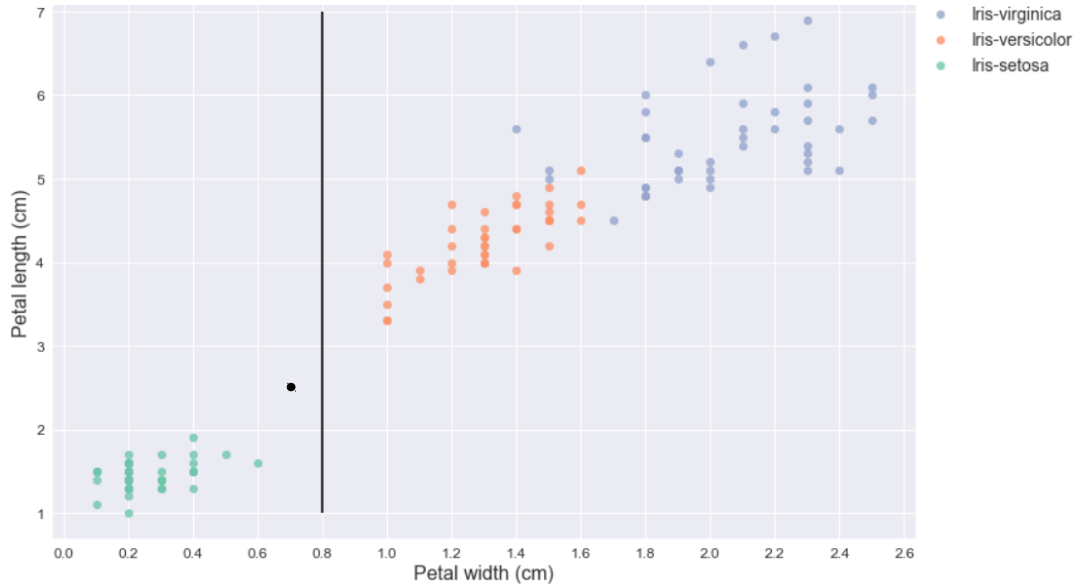

Iris-virginica
Iris-versicolor
Iris-setosa
Petal length (cm)
Petal width (cm)
1
2
3
4
5
6
7
0.0
0.2
0.4
0.6
0.8
1.0
1.2
1.4
1.6
1.8
2.0
2.2
2.4
2.6

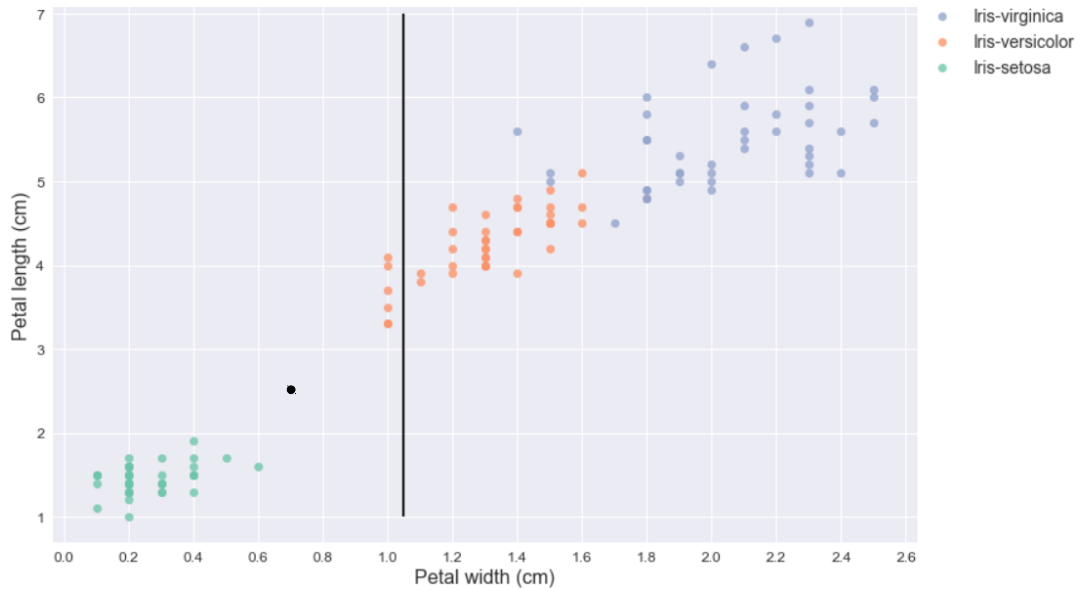

Iris-virginica
Iris-versicolor
Iris-setosa
Petal length (cm)
Petal width (cm)
1
2
3
4
5
6
7
0.0
0.2
0.4
0.6
0.8
1.0
1.2
1.4
1.6
1.8
2.0
2.2
2.4
2.6



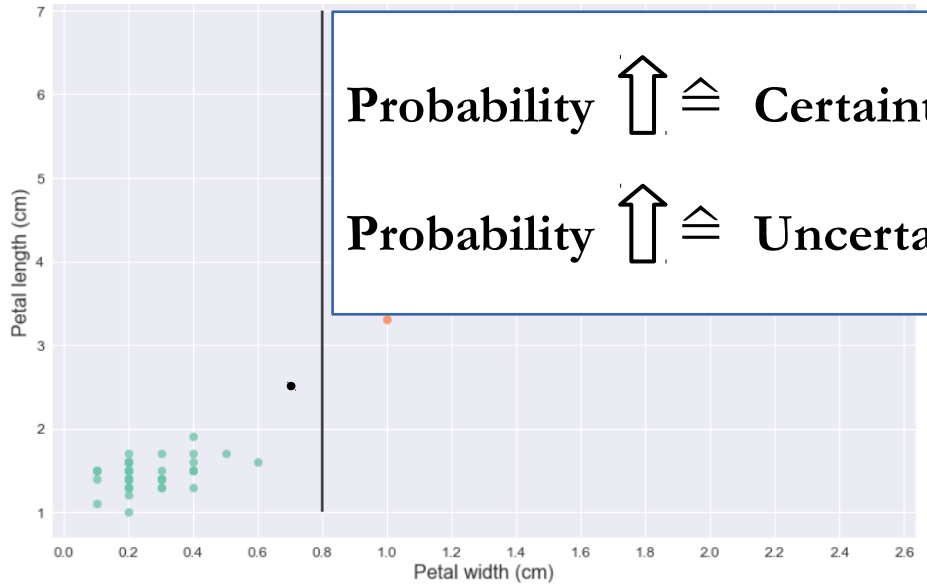


Probability ⇧ ≙ Certainty ⇧





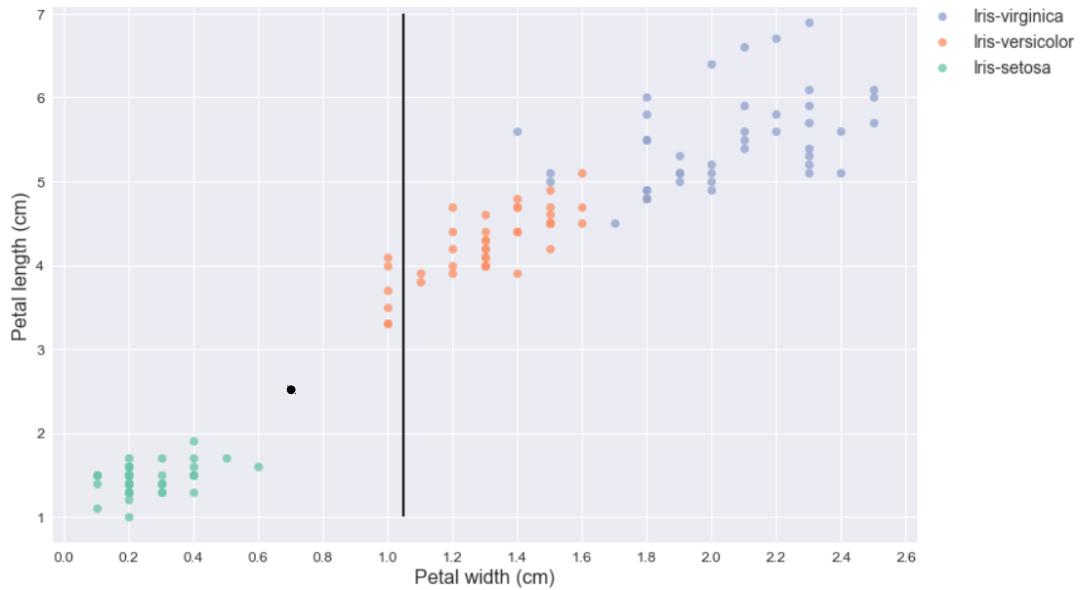
$$\text{Entropy} = -\sum_{i=1}^{c} p_i \cdot \log_2 p_i$$

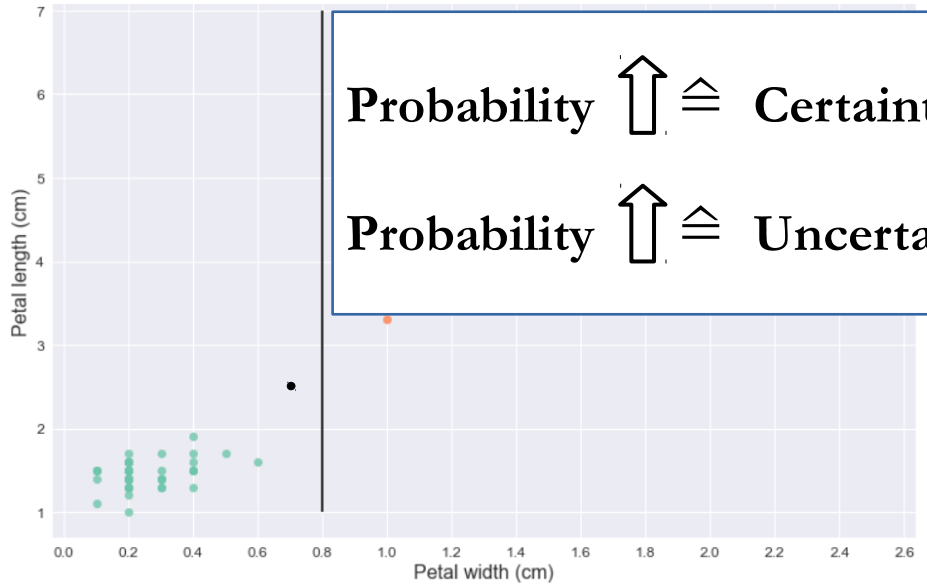


Probability ⇧ ≙ Certainty ⇧

Probability ⇧ ≙ Uncertainty ⇩

$$\text{Entropy} = -\sum_{i=1}^{c} p_i \cdot \log_2 p_i$$

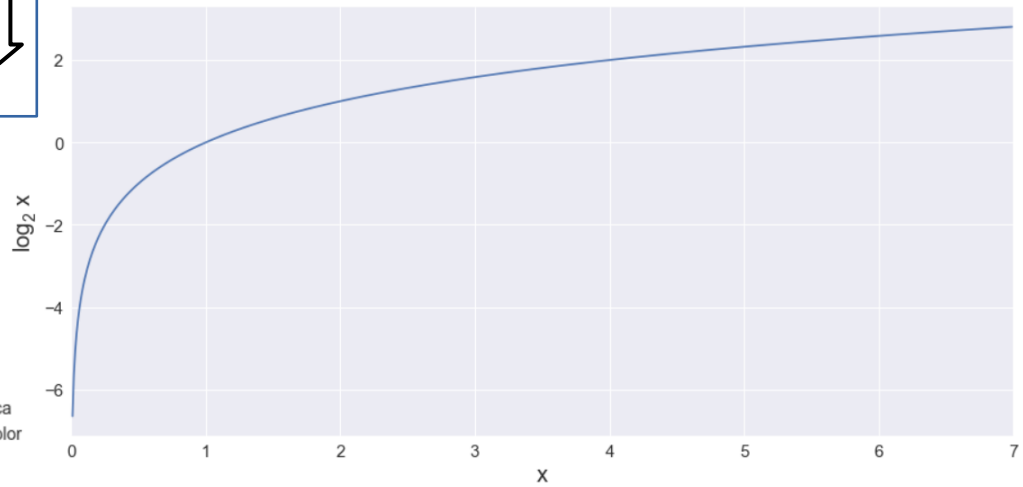


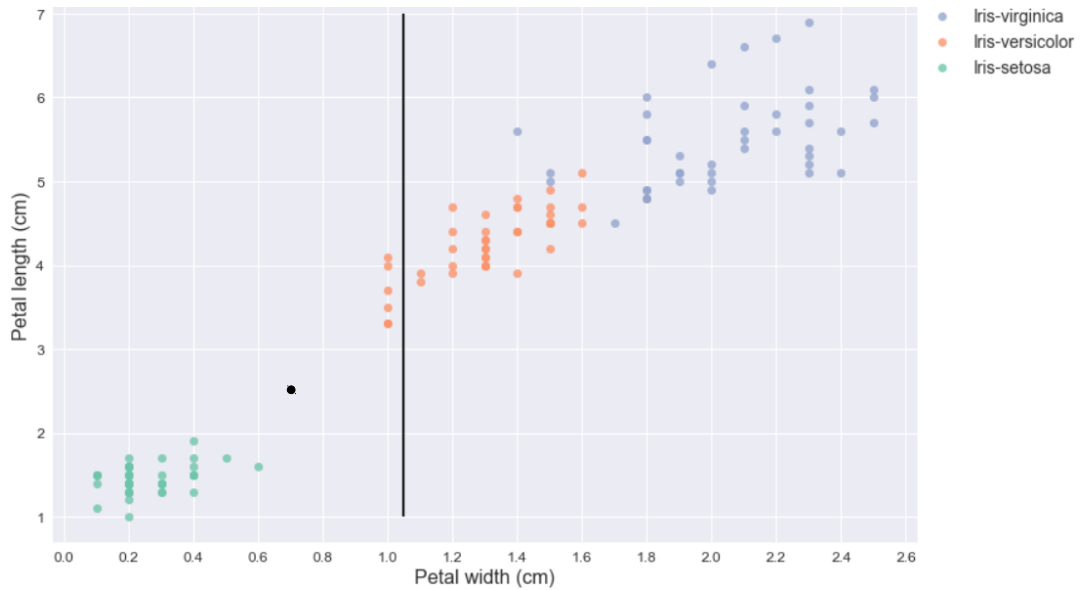


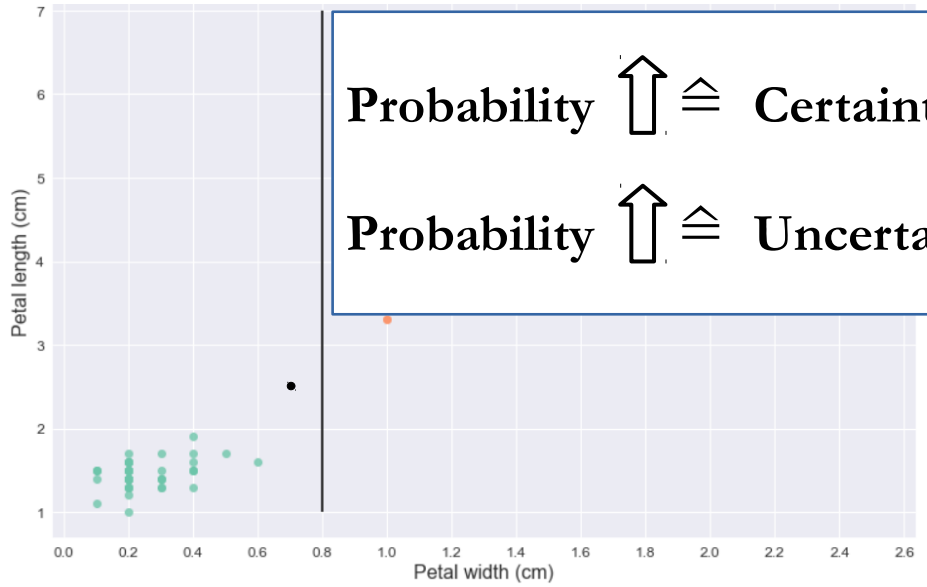
Probability ⇧ ≙ Certainty ⇧

Probability ⇧ ≙ Uncertainty ⇩

$$\text{Entropy} = -\sum_{i=1}^{c} p_i \cdot \log_2 p_i$$



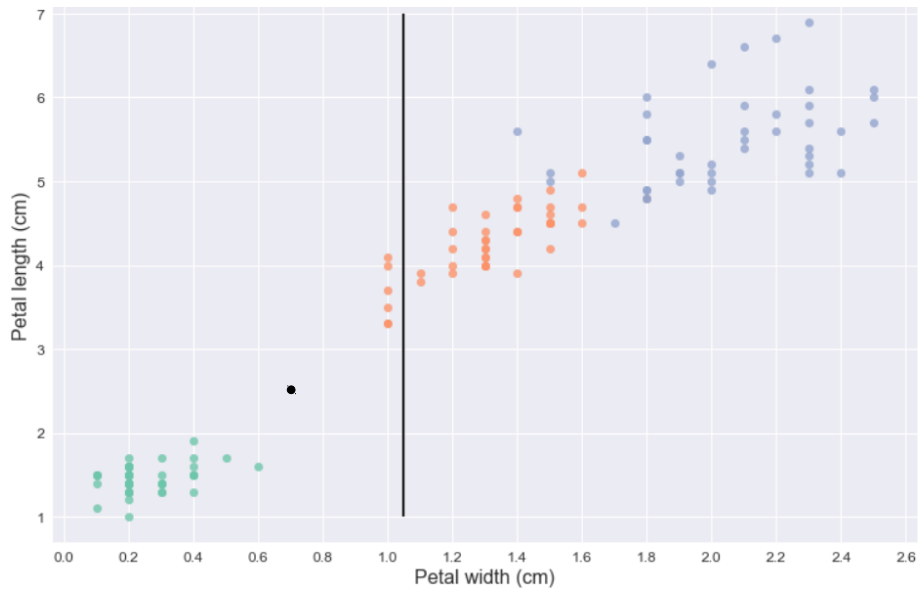


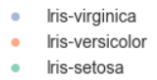


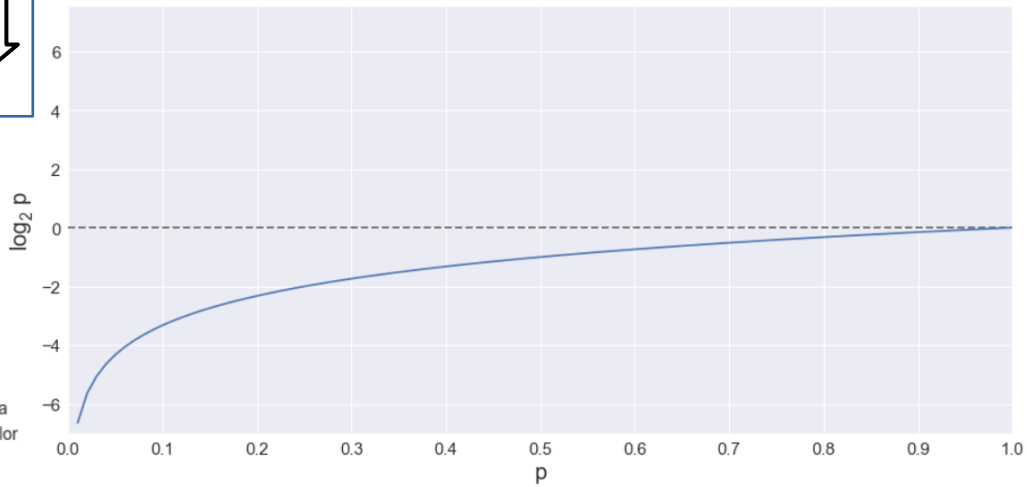
Probability ⇧ ≙ Certainty ⇧

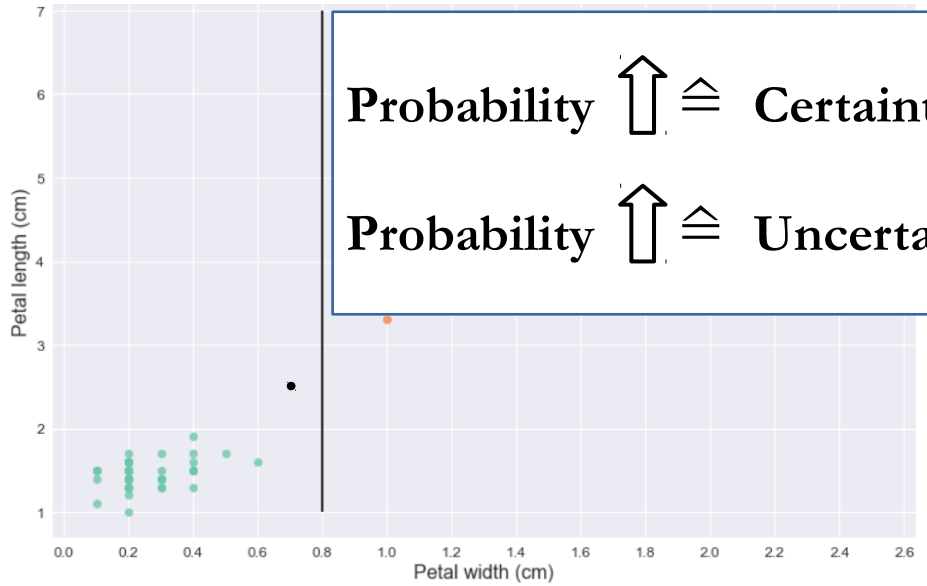
Probability ⇧ ≙ Uncertainty ⇩

$$\text{Entropy} = -\sum_{i=1}^{c} p_i \cdot \log_2 p_i$$





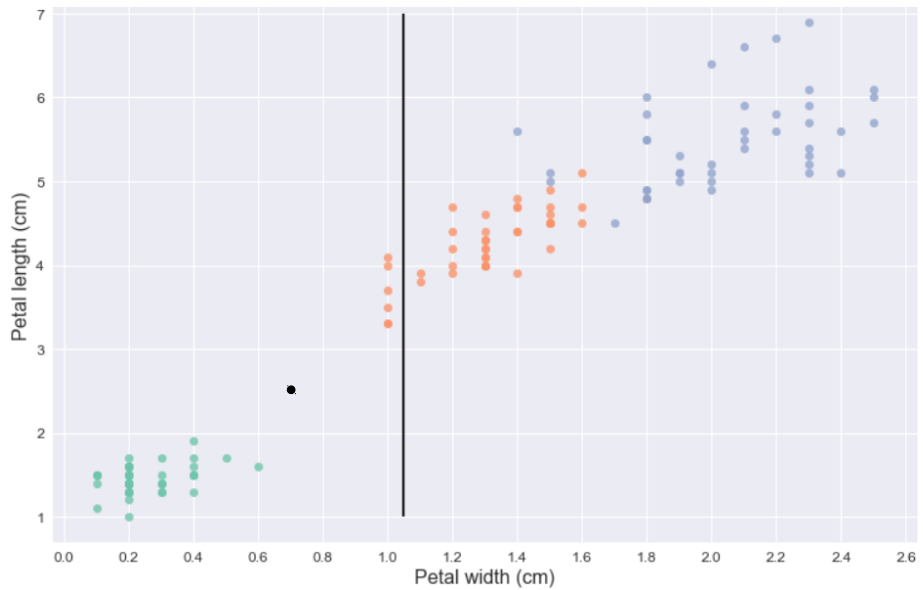


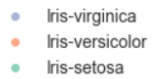


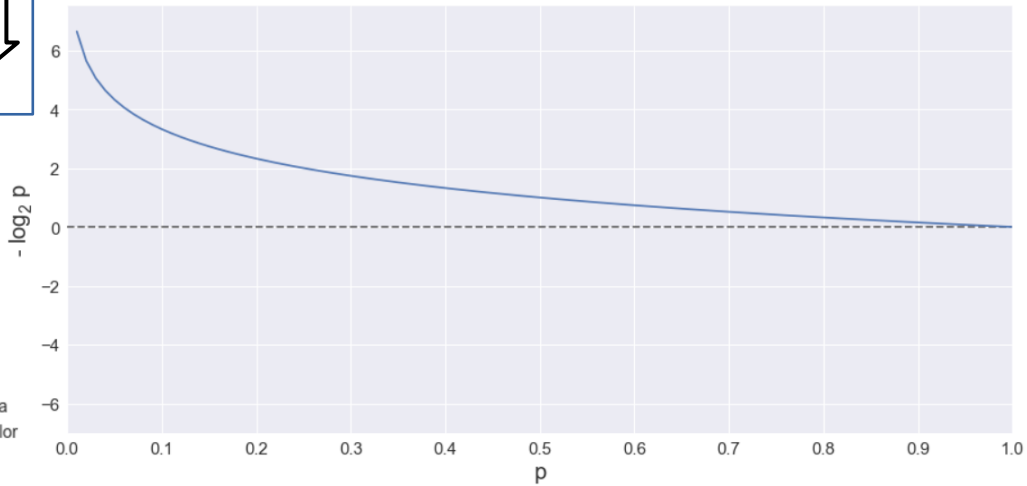
Probability ⇧ ≙ Certainty ⇧

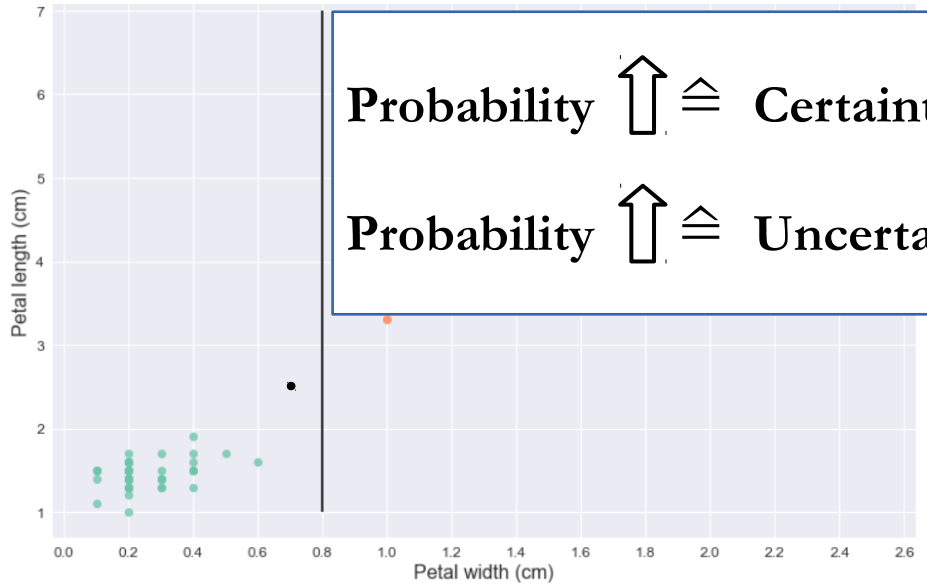
Probability ⇧ ≙ Uncertainty ⇩

$$\text{Entropy} = -\sum_{i=1}^{c} p_i \cdot \log_2 p_i$$





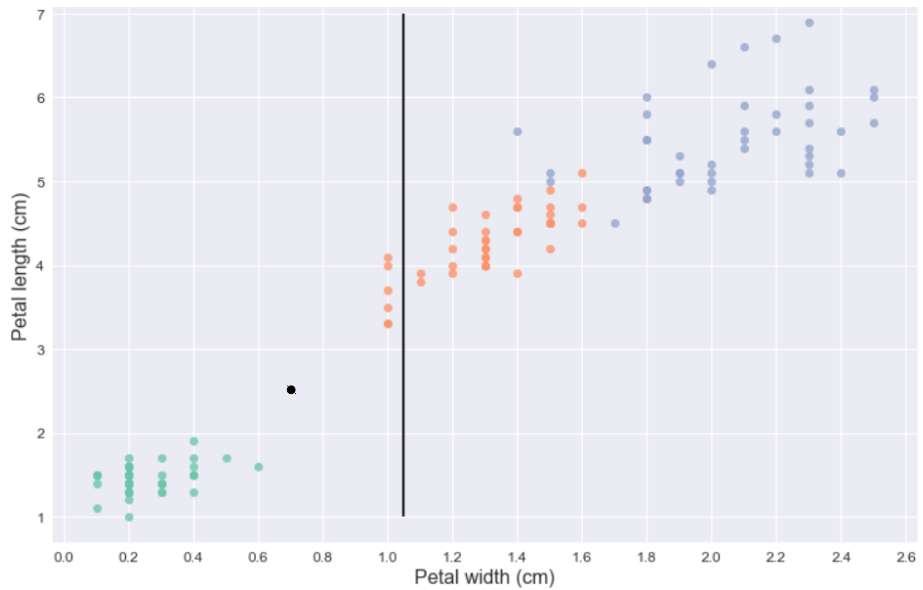


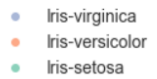


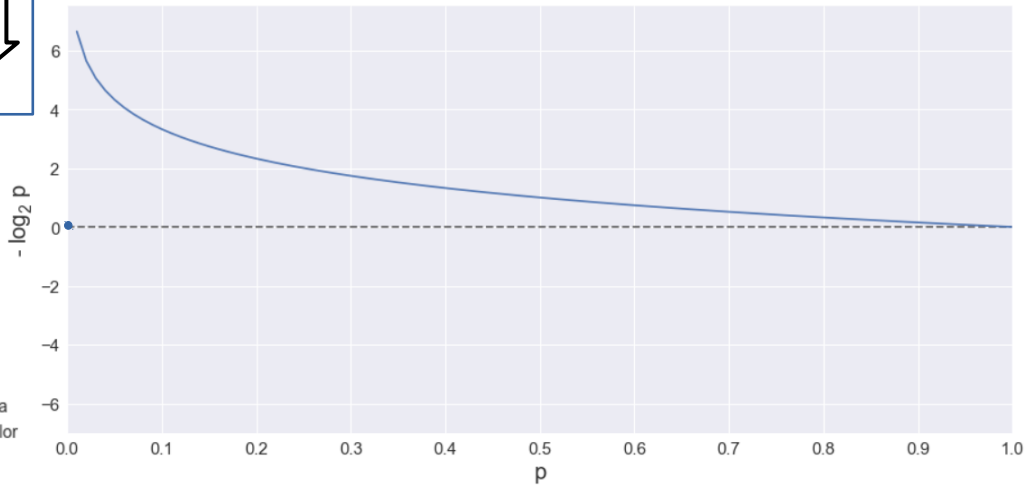
Probability ⇧ ≙ Certainty ⇧

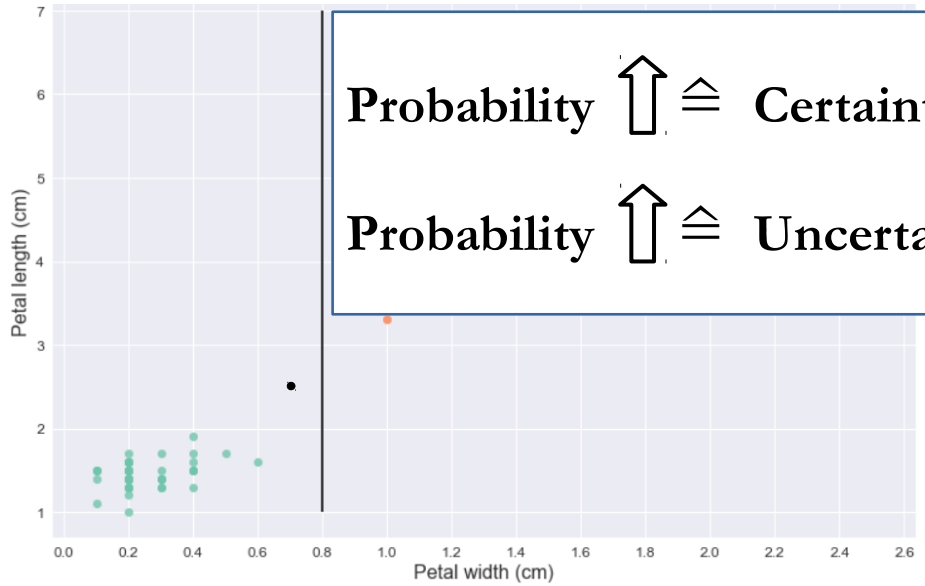
Probability ⇧ ≙ Uncertainty ⇩

$$\text{Entropy} = -\sum_{i=1}^{c} p_i \cdot \log_2 p_i$$





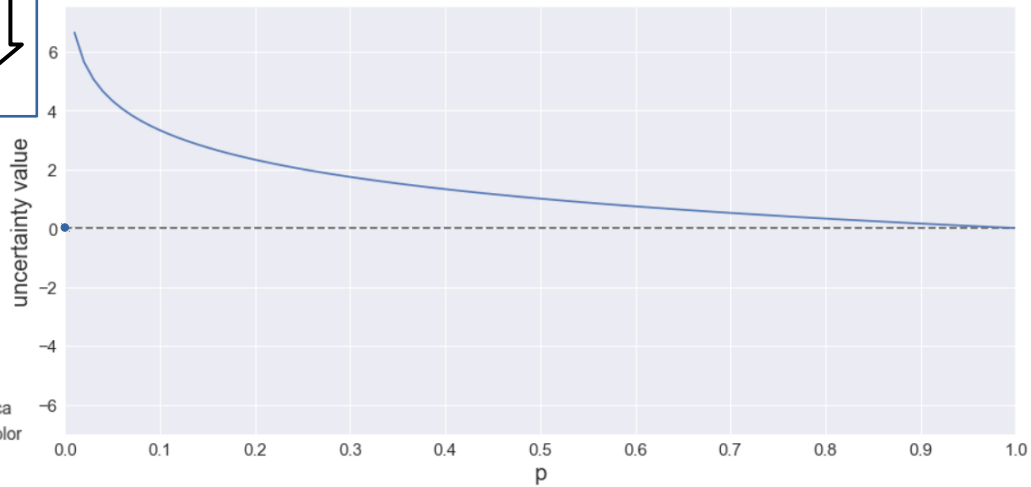


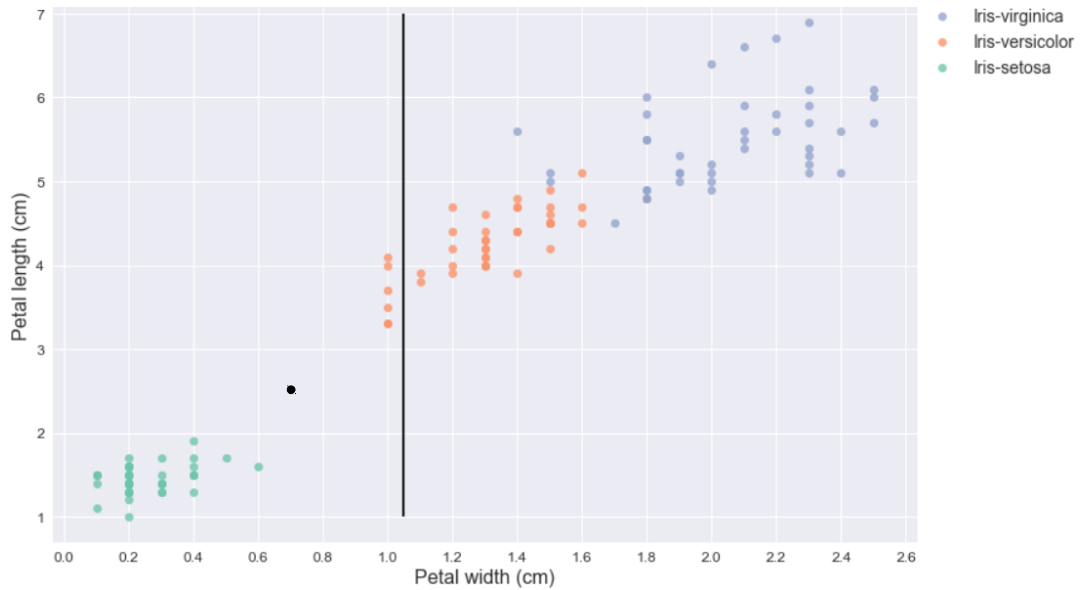


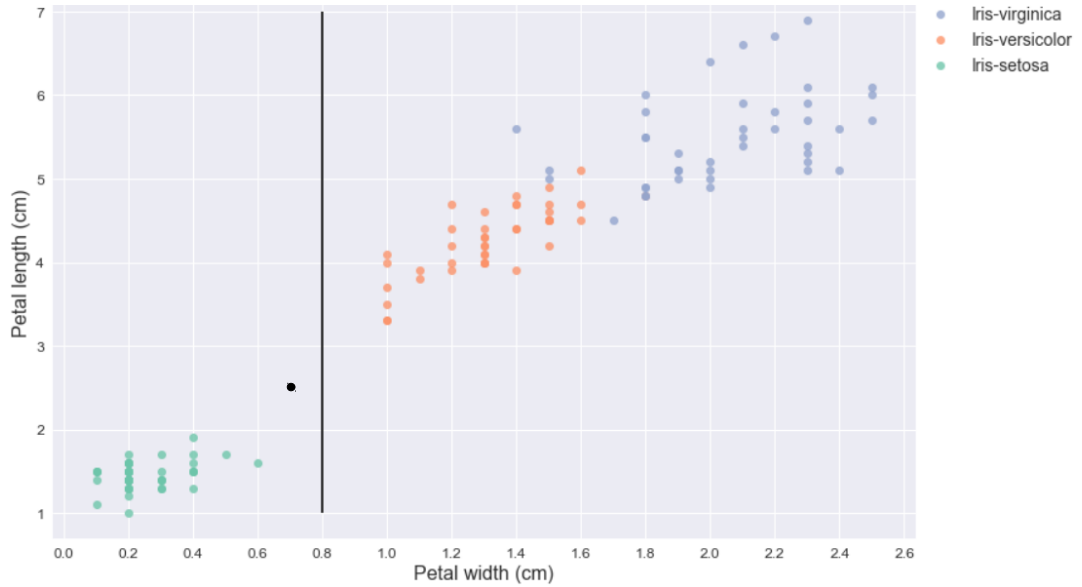
Probability ⇧ ≙ Certainty ⇧

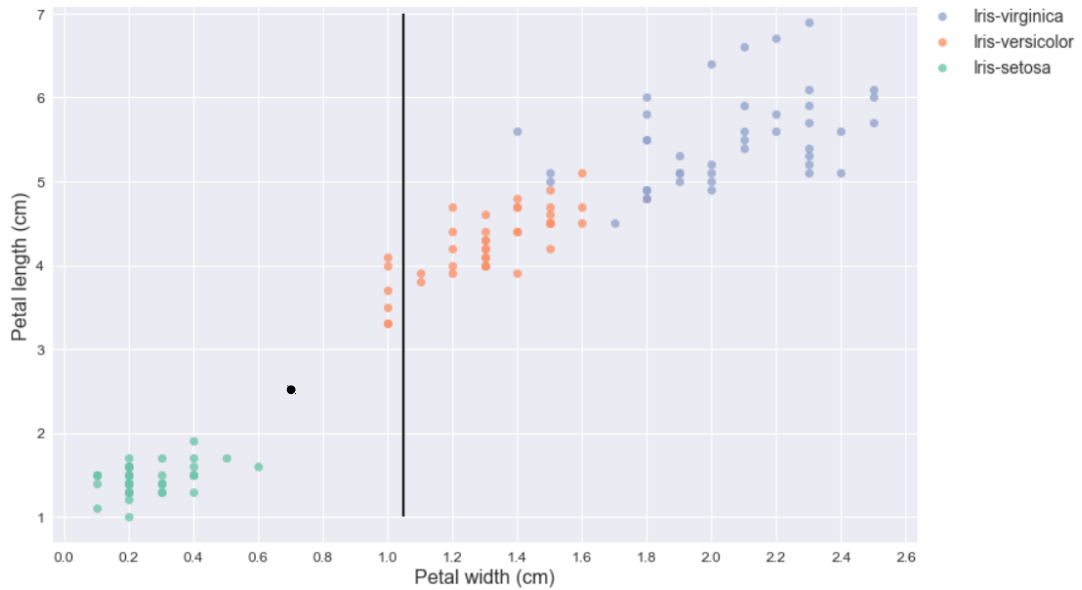
Probability ⇧ ≙ Uncertainty ⇩

$$\text{Entropy} = -\sum_{i=1}^{c} p_i \cdot \log_2 p_i$$







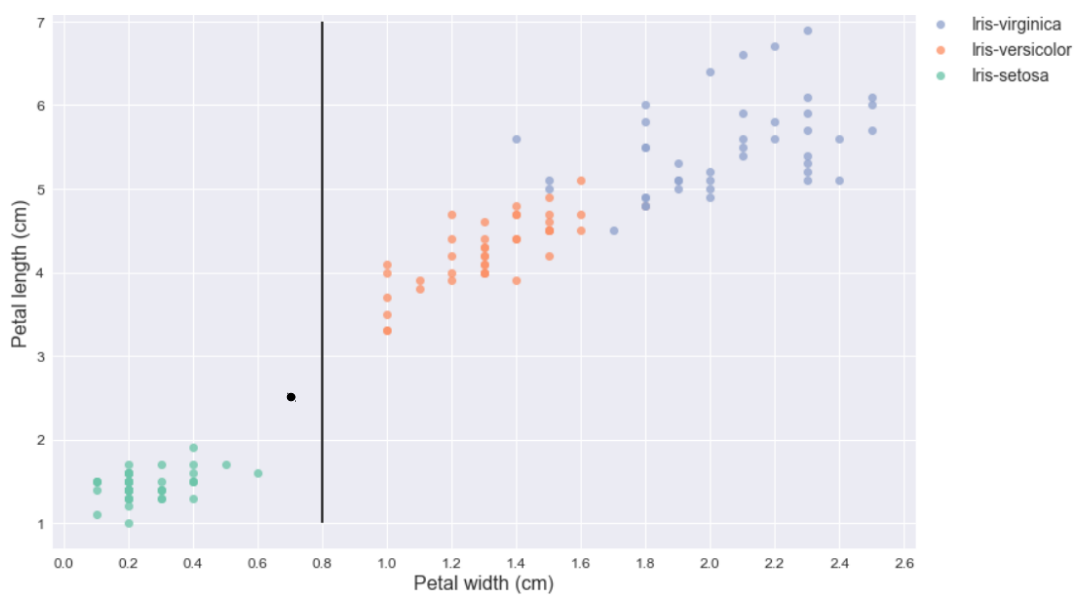


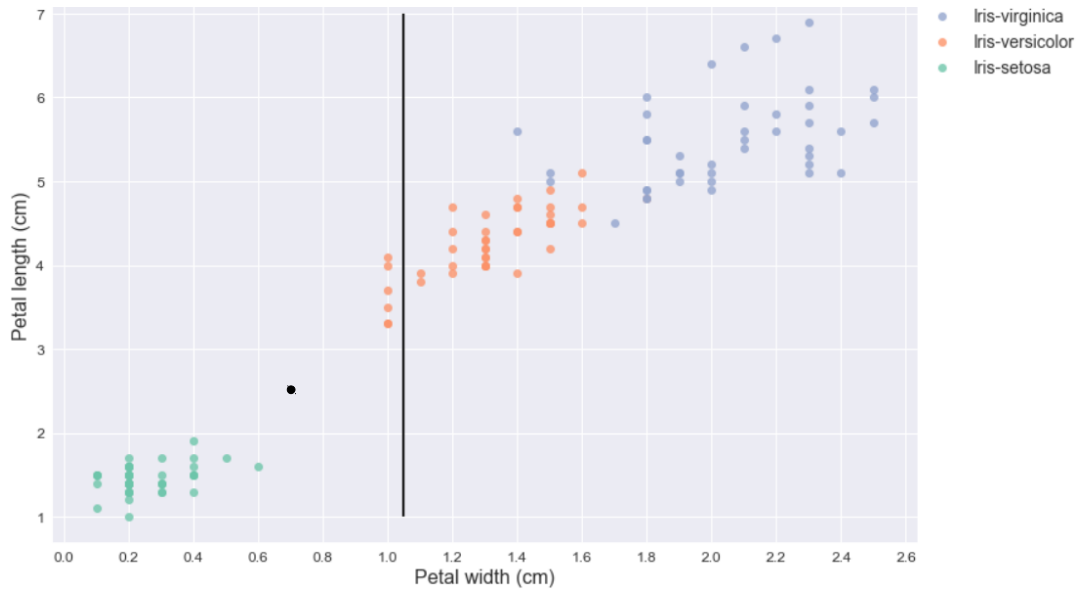
$$\text{Entropy} = -\sum_{i=1}^{c} p_i \cdot \log_2 p_i$$

$$- (-3 + (-7)) = 10$$

$$\text{Entropy} = \sum_{i=1}^{c} p_i \cdot (-\log_2 p_i)$$

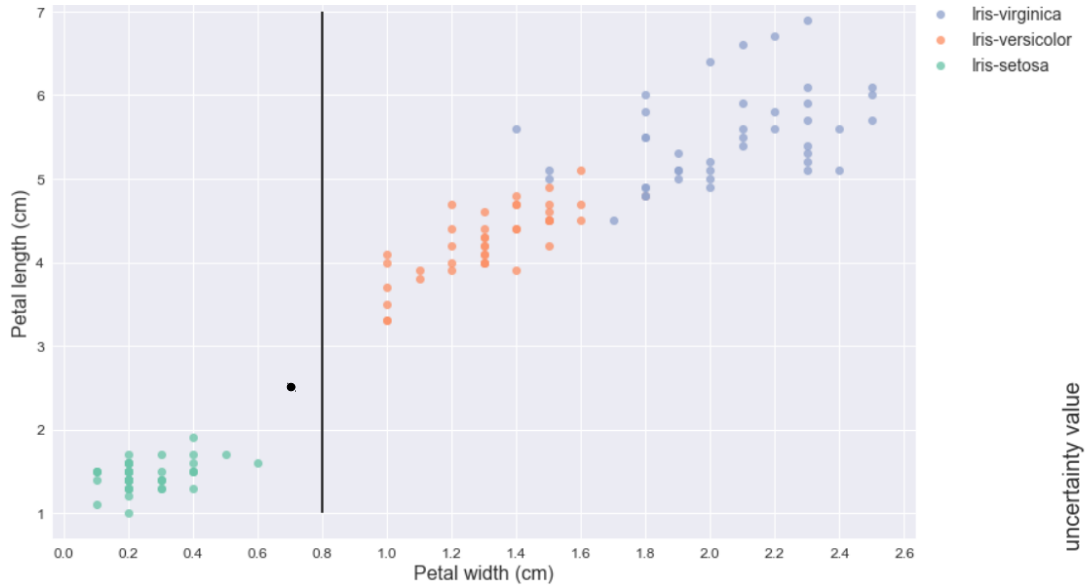
$$(3 + 7) = 10$$

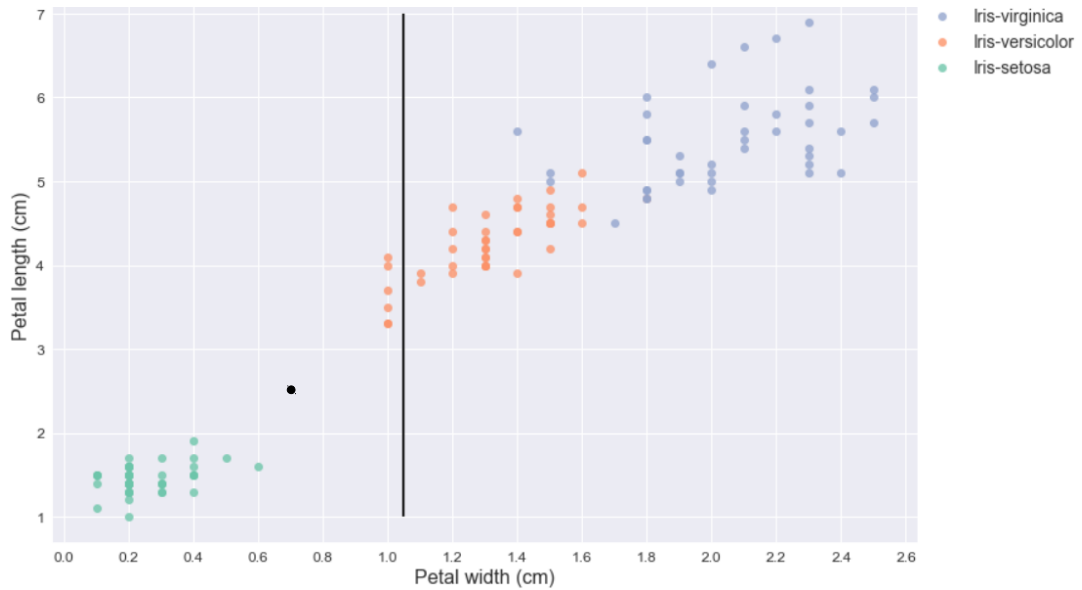




Uncertainty Value

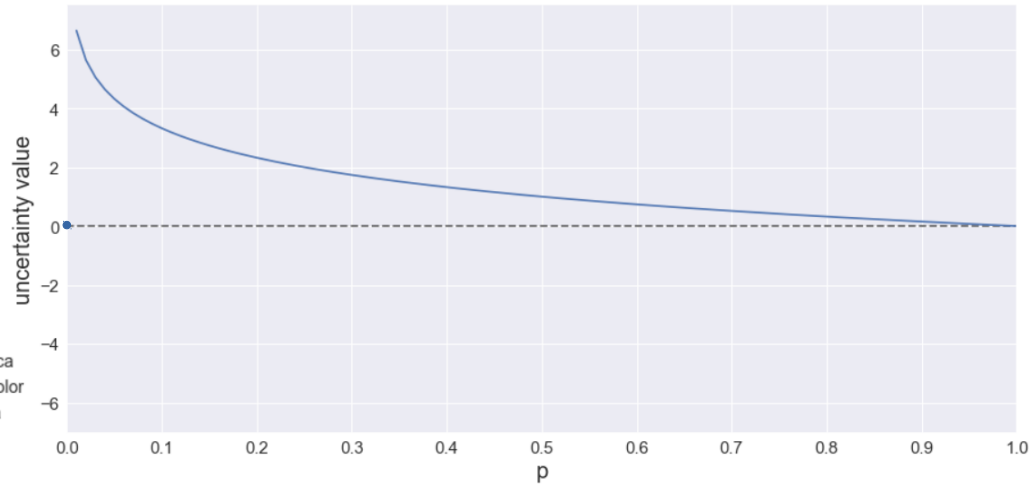
$$\text{Entropy} = \sum_{i=1}^{c} p_i \cdot (-\log_2 p_i)$$

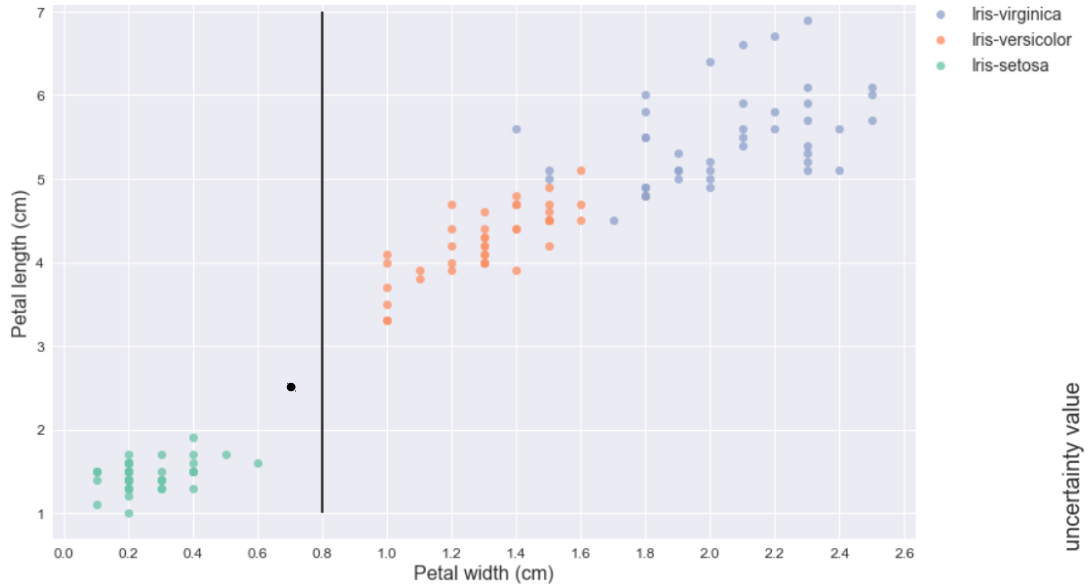


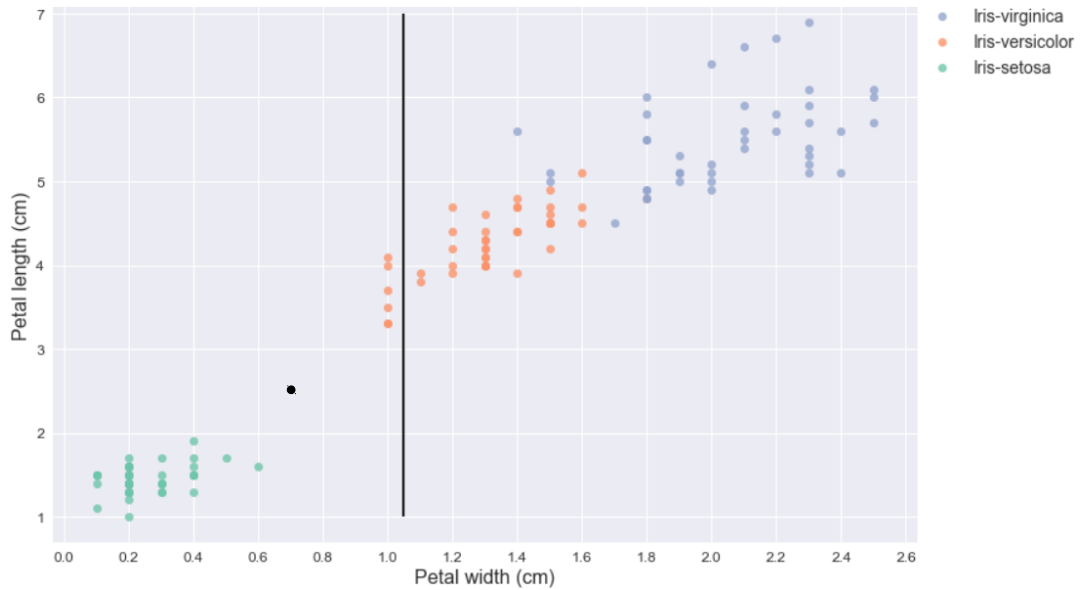


Uncertainty Value

$$\text{Entropy} = \sum_{i=1}^{c} p_i \cdot (-\log_2 p_i)$$

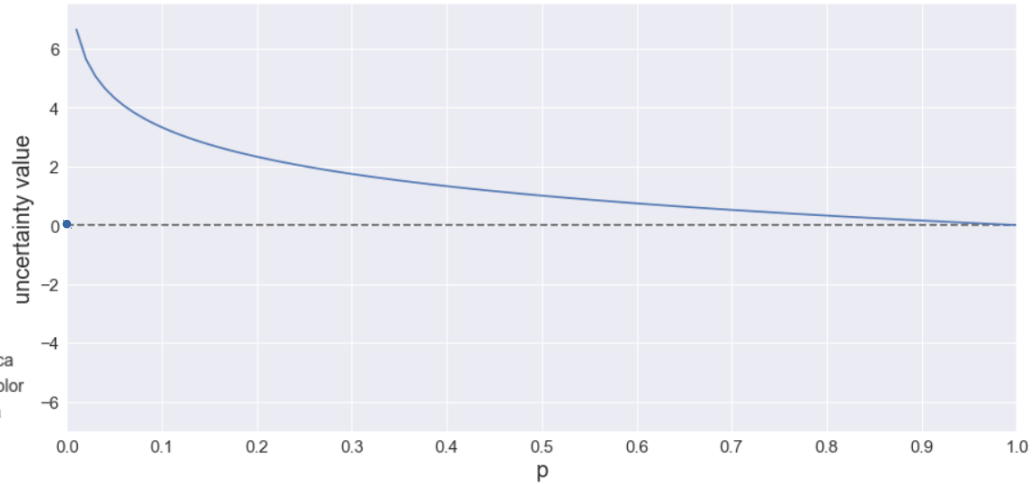


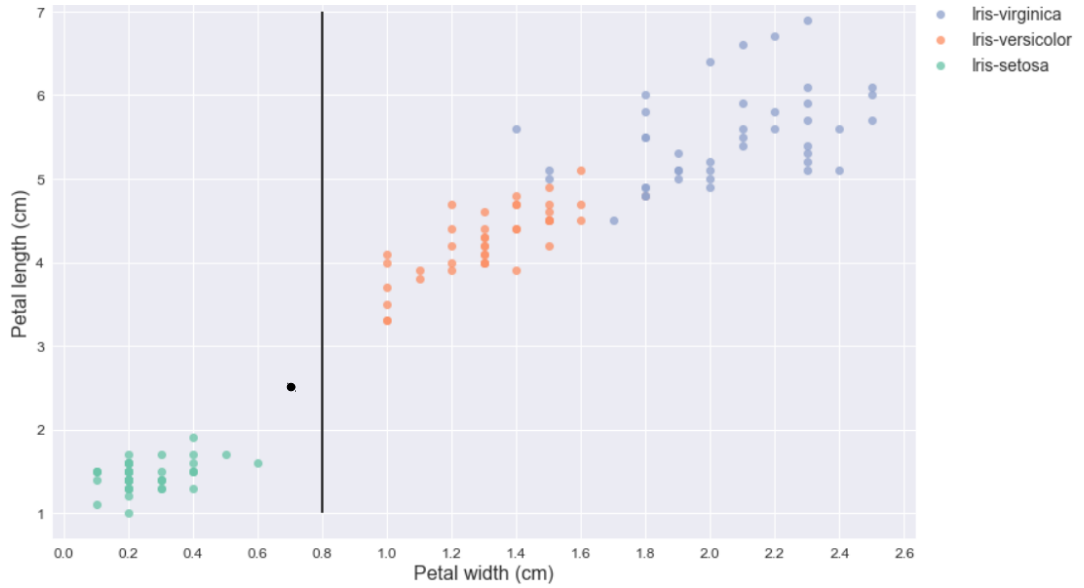


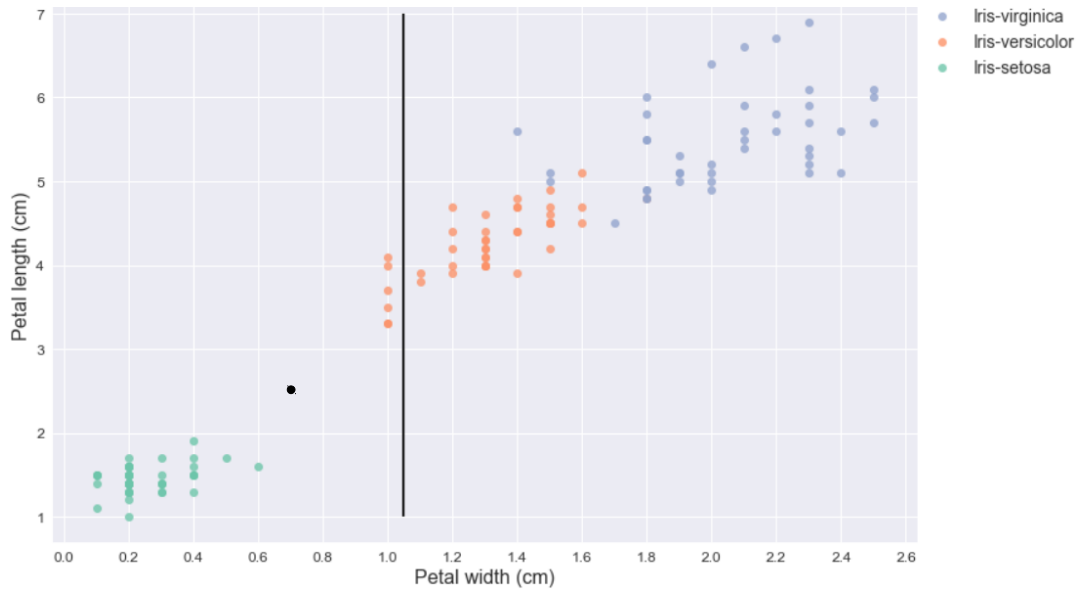


Weighted Sum Uncertainty Value

$$\text{Entropy} = \sum_{i=1}^{c} p_i \cdot (-\log_2 p_i)$$

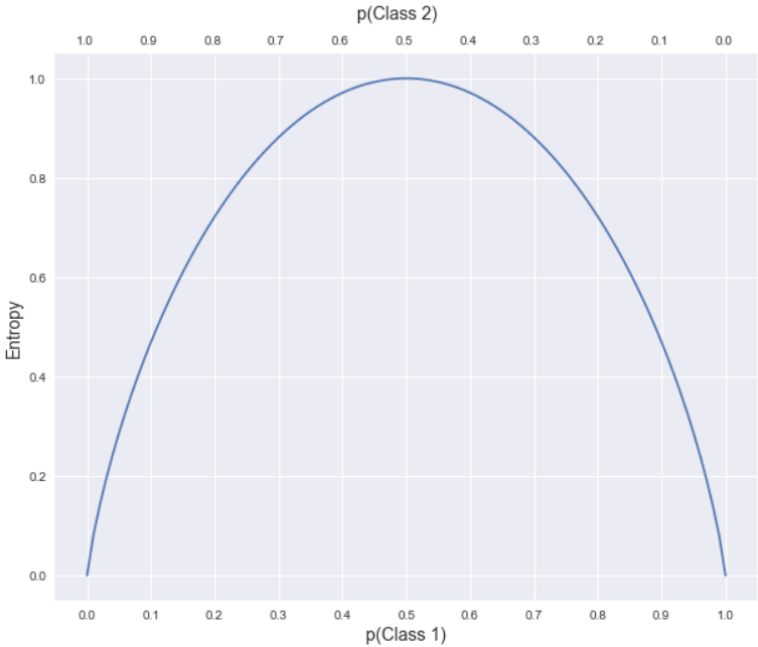


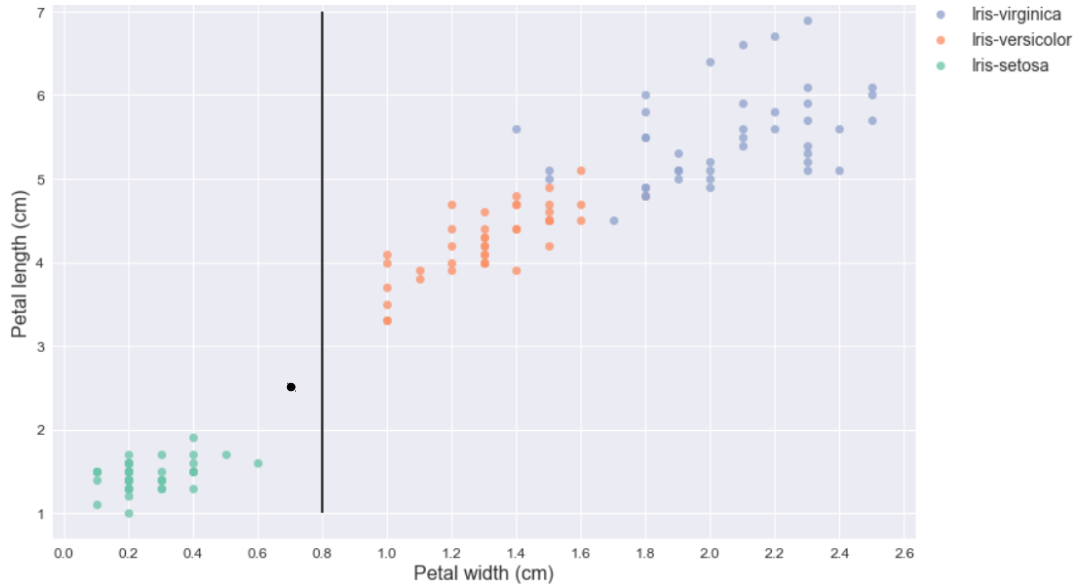


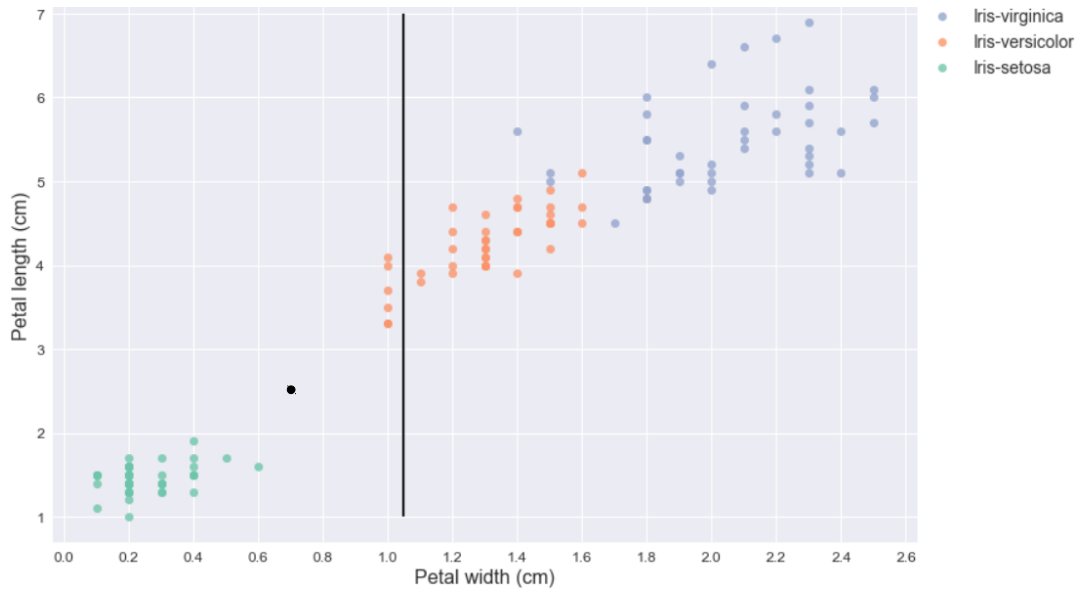


Weighted Sum Uncertainty Value

$$\text{Entropy} = \sum_{i=1}^{c} p_i \cdot (-\log_2 p_i)$$

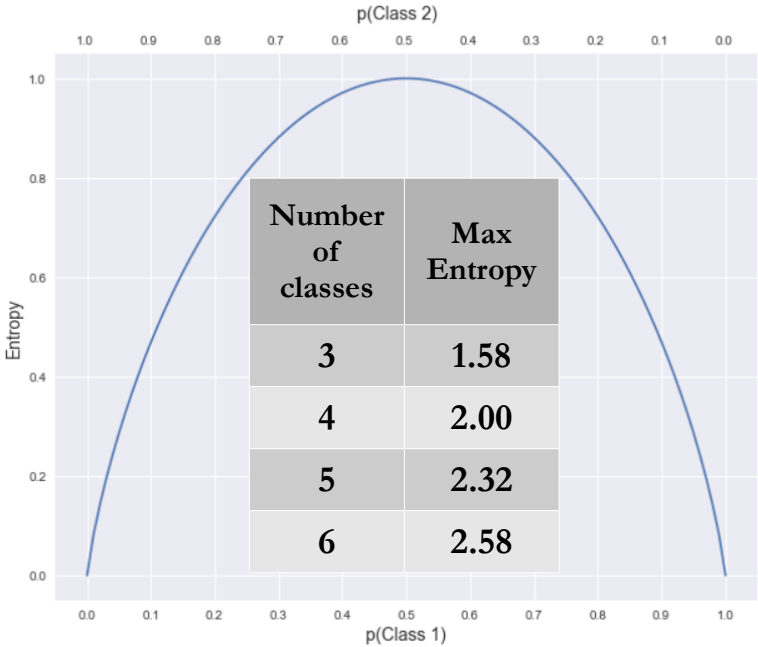




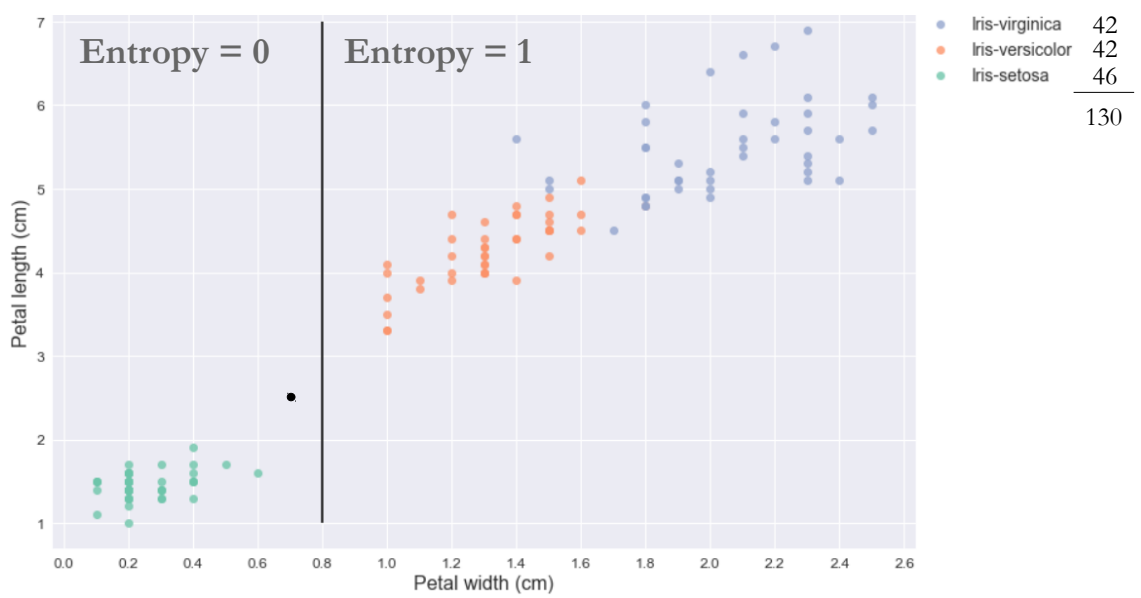


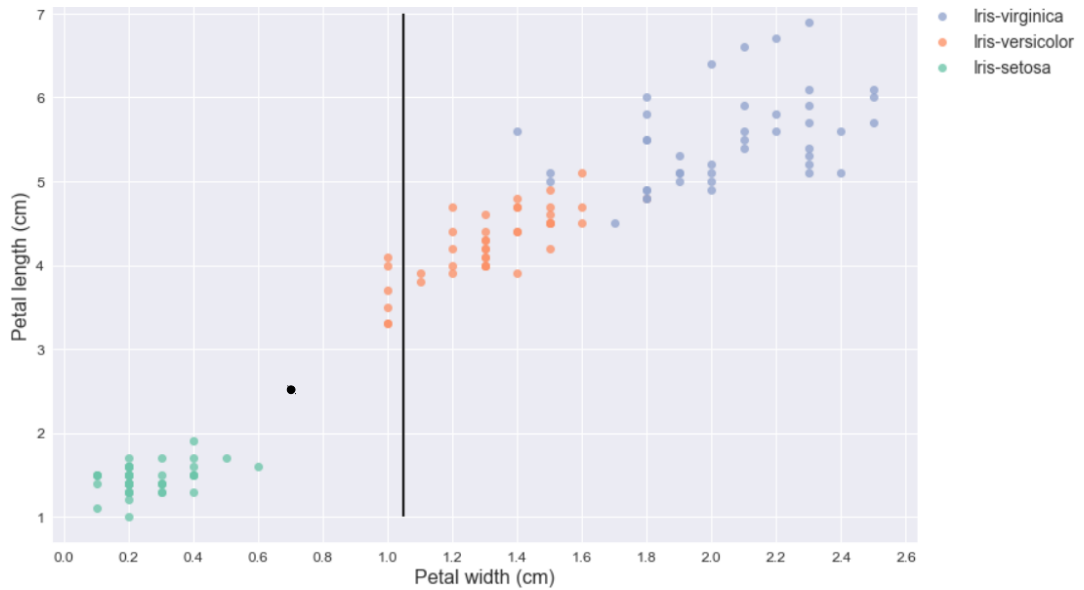
Weighted Sum Uncertainty Value

$$\text{Entropy} = \sum_{i=1}^{c} p_i \cdot (-\log_2 p_i)$$



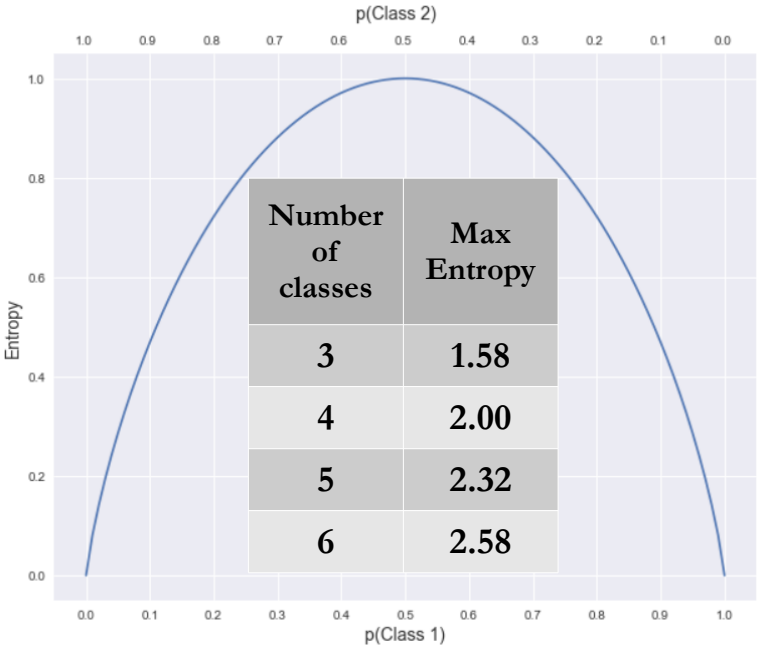
Number of classes	Max Entropy
3	1.58
4	2.00
5	2.32
6	2.58



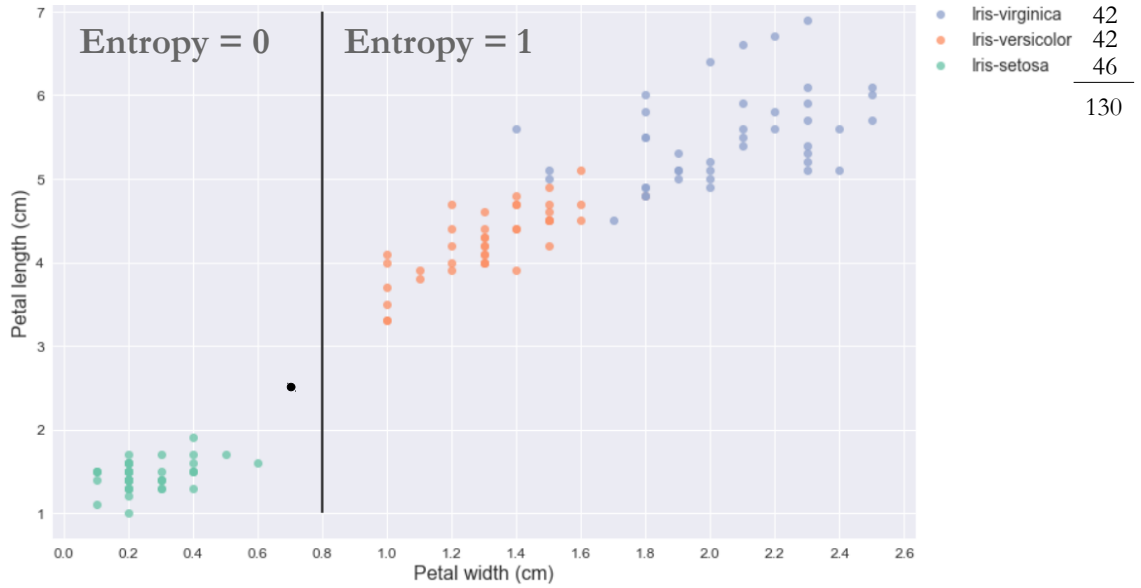


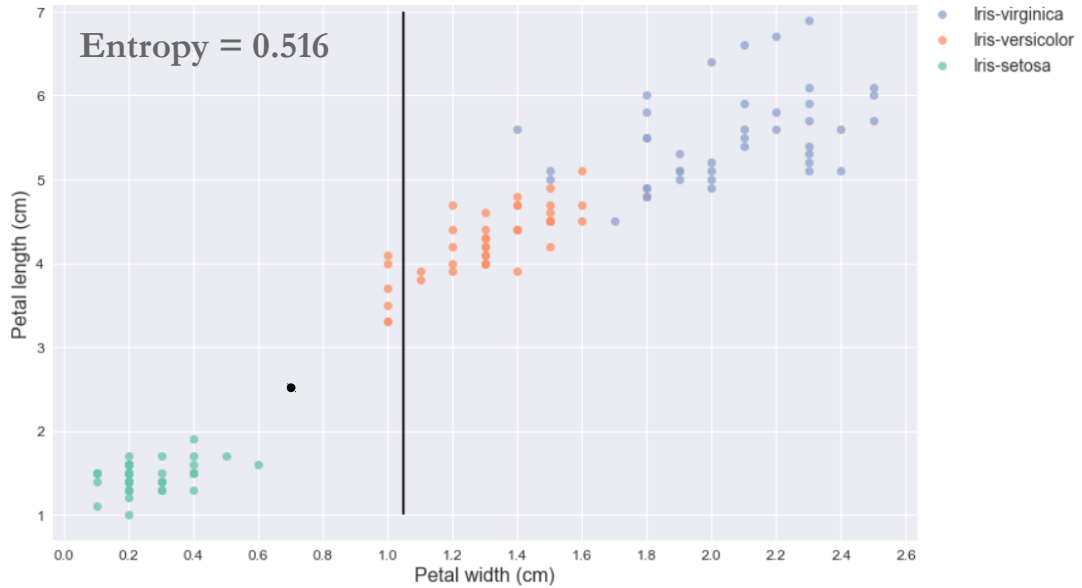
Weighted Sum Uncertainty Value

$$\text{Entropy} = \sum_{i=1}^{c} p_i \cdot (-\log_2 p_i)$$



Number of classes	Max Entropy
3	1.58
4	2.00
5	2.32
6	2.58



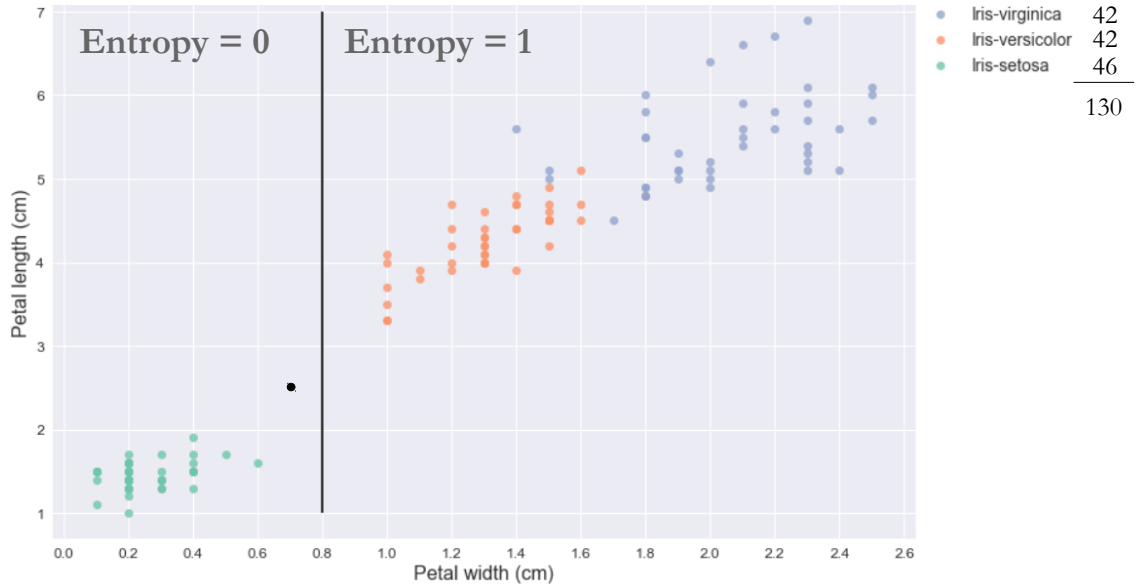


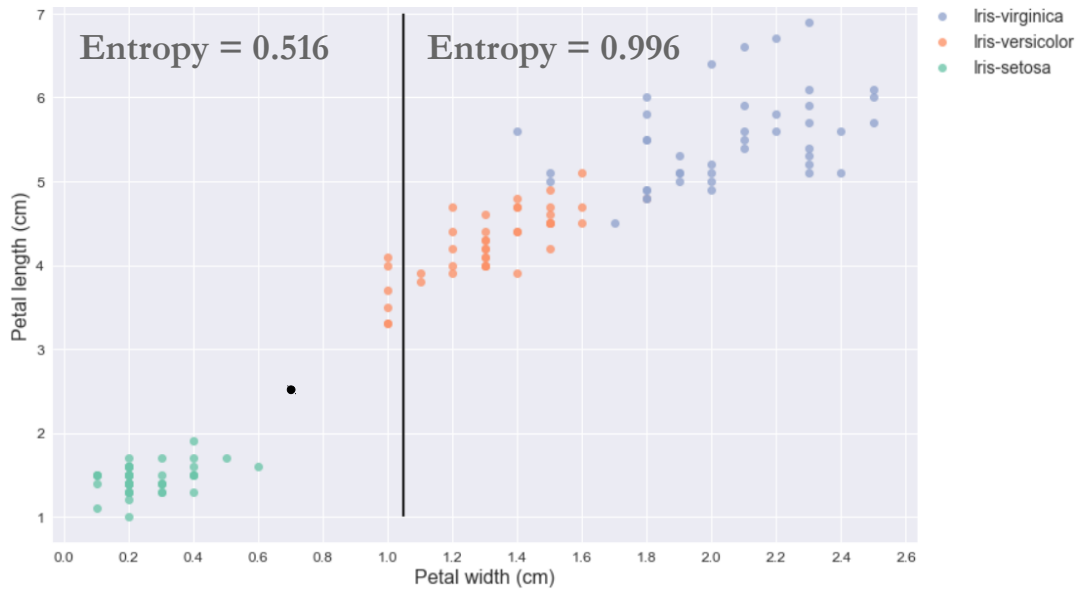
Weighted Sum Uncertainty Value

$$\text{Entropy} = \sum_{i=1}^{c} p_i \cdot (-\log_2 p_i)$$

$$\text{Entropy} = {}^{46}/_{52} \cdot (-\log_2 {}^{46}/_{52}) + {}^{6}/_{52} \cdot (-\log_2 {}^{6}/_{52})$$

$$= 0.516$$



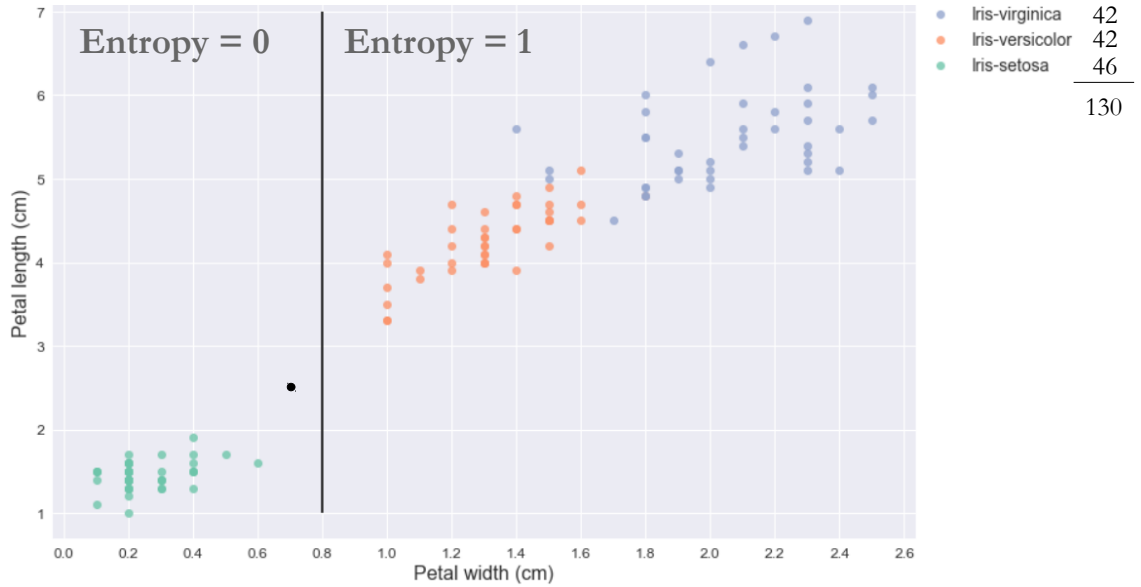


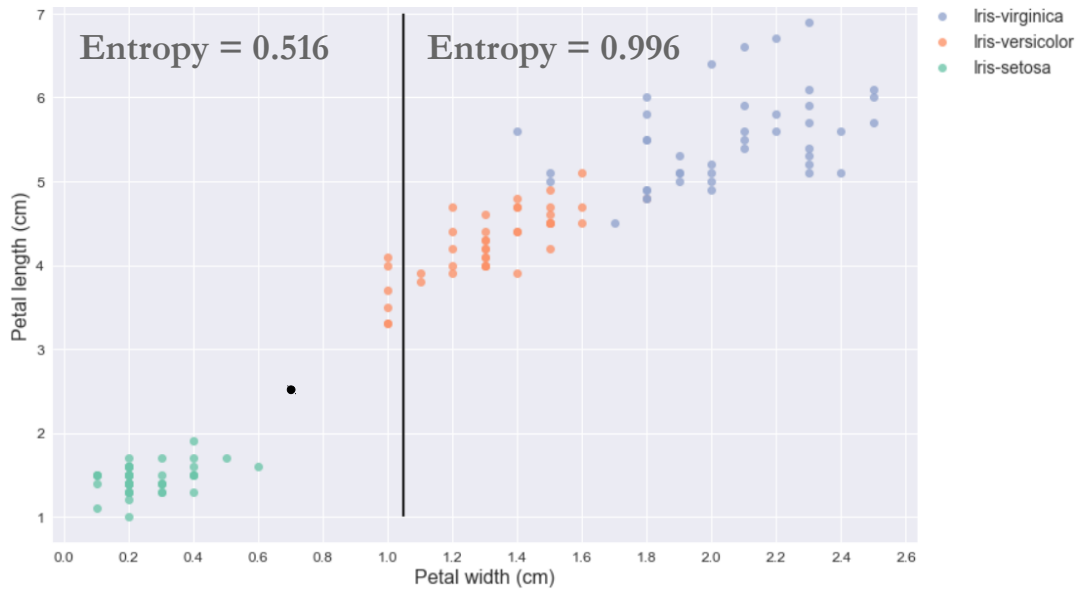
Weighted Sum Uncertainty Value

$$\text{Entropy} = \sum_{i=1}^{c} p_i \cdot (-\log_2 p_i)$$

$$\text{Entropy} = {}^{36}/_{78} \cdot (-\log_2 {}^{36}/_{78}) + {}^{42}/_{78} \cdot (-\log_2 {}^{42}/_{78})$$

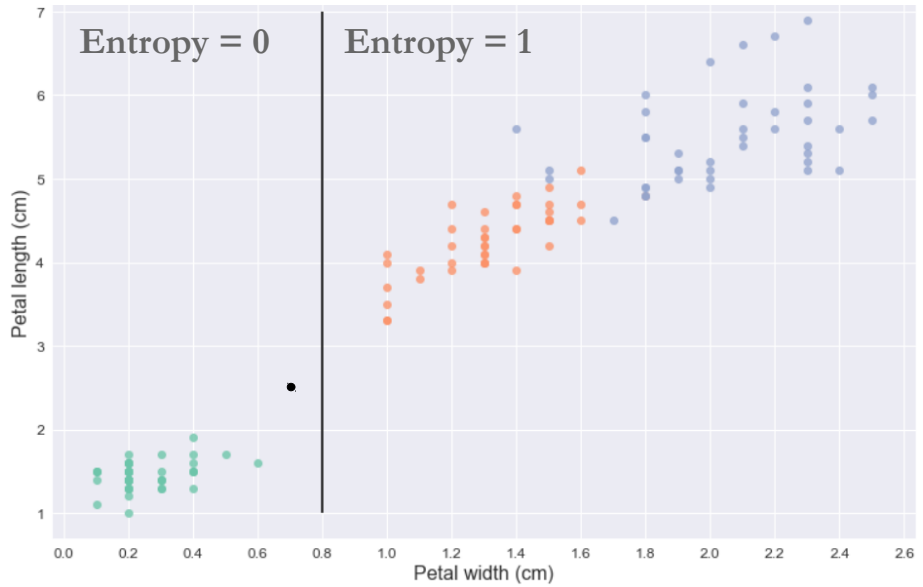
$$= 0.996$$





Weighted Sum Uncertainty Value

$$\text{Entropy} = \sum_{i=1}^{c} p_i \cdot (-\log_2 p_i)$$

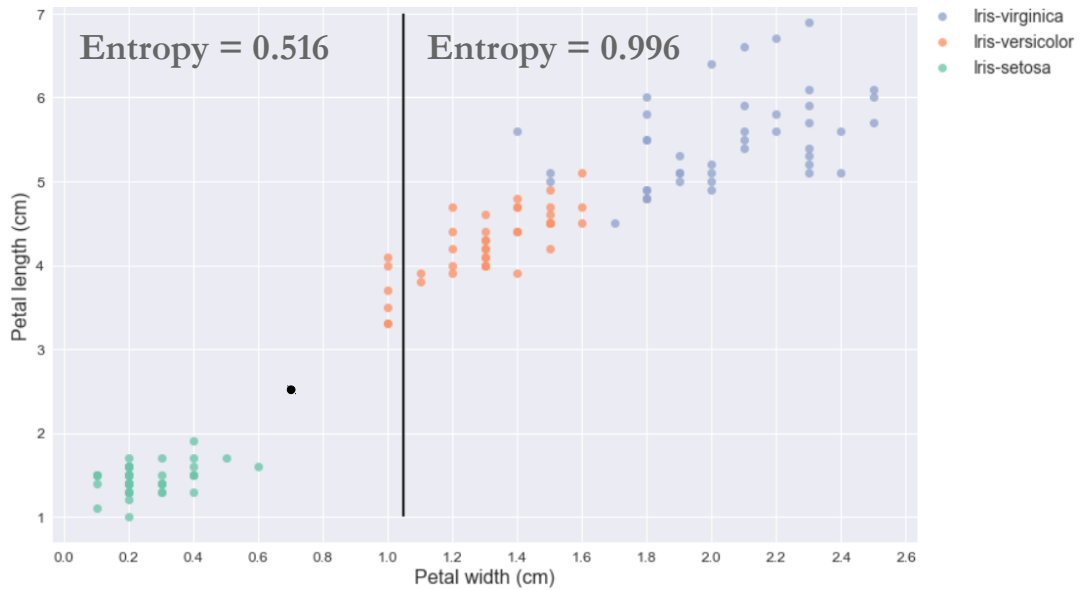


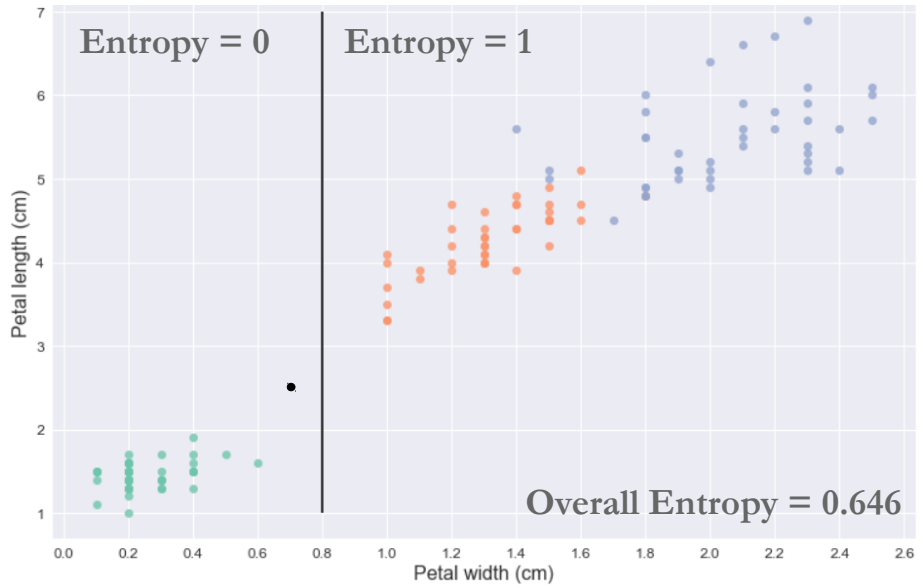
Iris-virginica	42
Iris-versicolor	42
Iris-setosa	46
	130

Weighted Sum Uncertainty Value

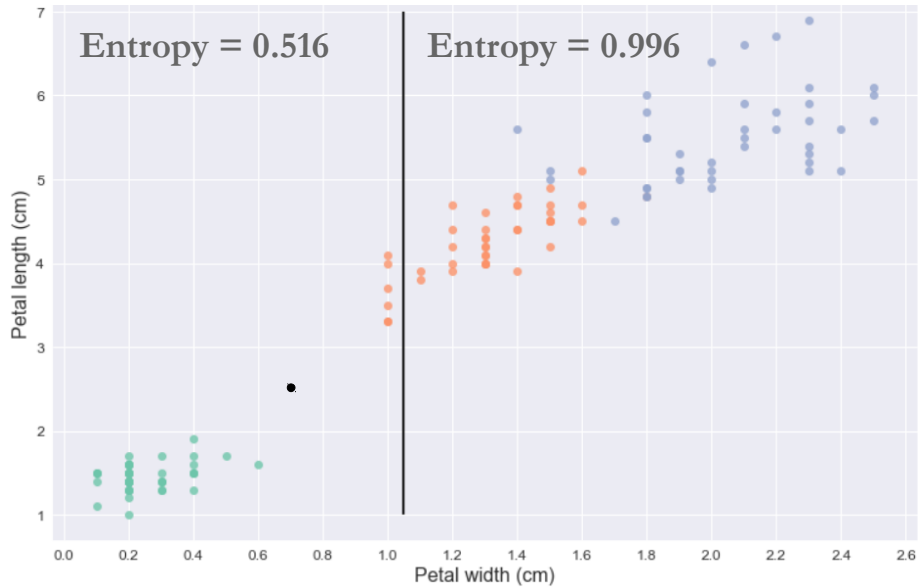
$$\text{Entropy} = \sum_{i=1}^{c} p_i \cdot (-\log_2 p_i)$$

$$\text{Overall Entropy} = \sum_{j=1}^{2} p_j \cdot \text{Entropy}_j$$





Iris-virginica	42
Iris-versicolor	42
Iris-setosa	46
	130



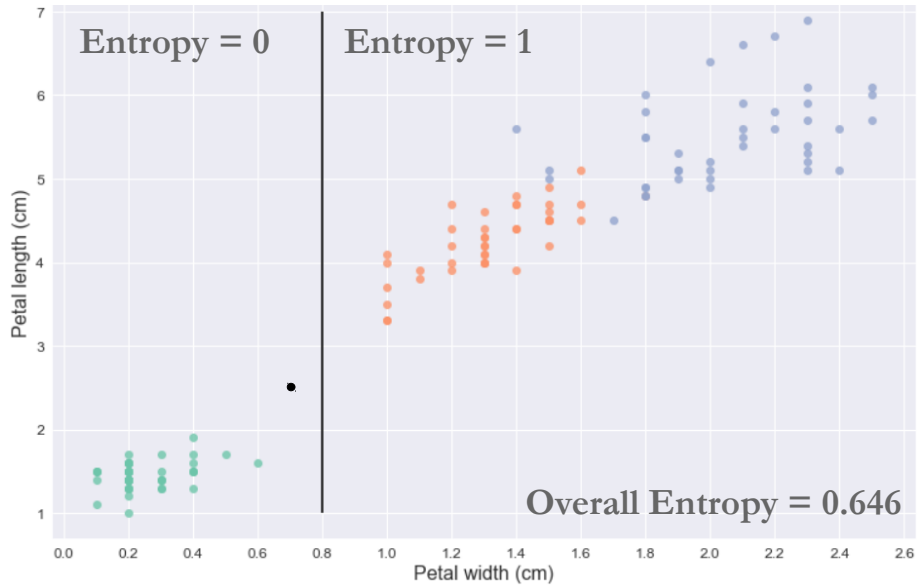
Iris-virginica
Iris-versicolor
Iris-setosa

Weighted Sum Uncertainty Value

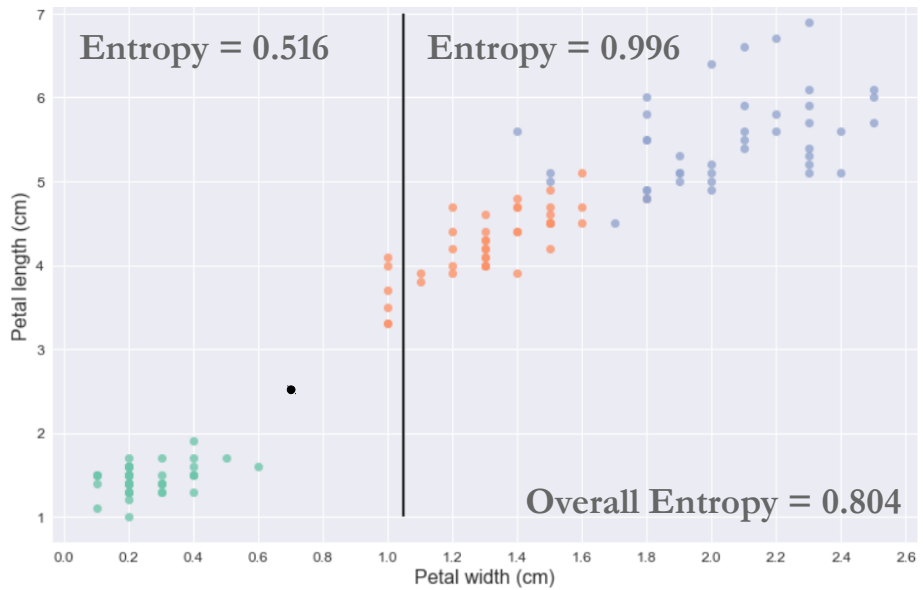
$$\text{Entropy} = \sum_{i=1}^{c} p_i \cdot (-\log_2 p_i)$$

$$\text{Overall Entropy} = \sum_{j=1}^{2} p_j \cdot \text{Entropy}_j$$

$$\text{Overall Entropy} = {}^{46}/_{130} \cdot 0 + {}^{84}/_{130} \cdot 1 = 0.646$$



Iris-virginica	42
Iris-versicolor	42
Iris-setosa	46
	130

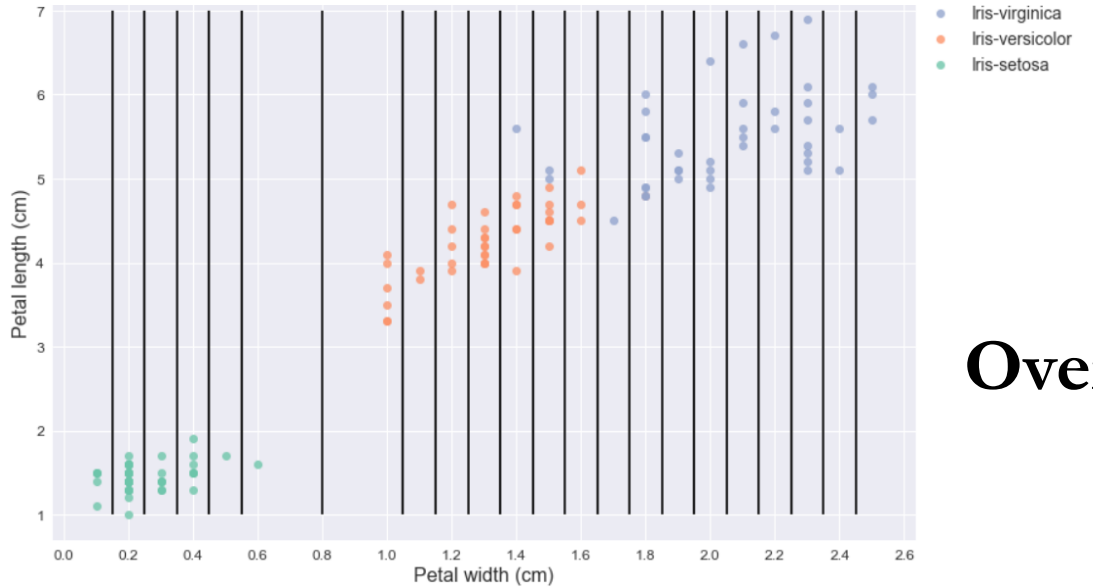


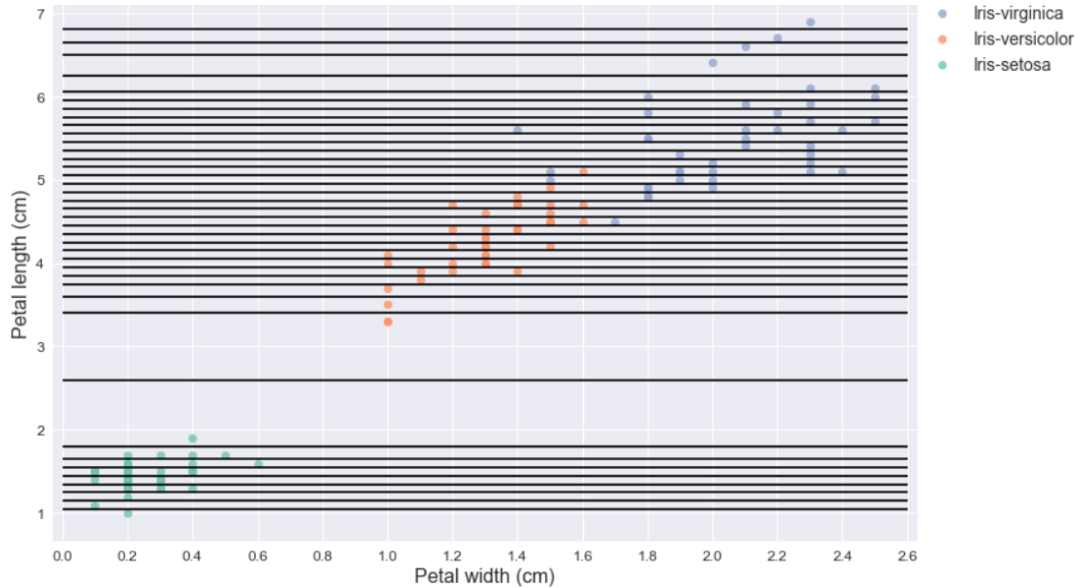
Weighted Sum Uncertainty Value

$$\text{Entropy} = \sum_{i=1}^{c} p_i \cdot (-\log_2 p_i)$$

$$\text{Overall Entropy} = \sum_{j=1}^{2} p_j \cdot \text{Entropy}_j$$

$$\text{Overall Entropy} = {}^{52}/_{130} \cdot 0.516 + {}^{78}/_{130} \cdot 0.996 = 0.804$$

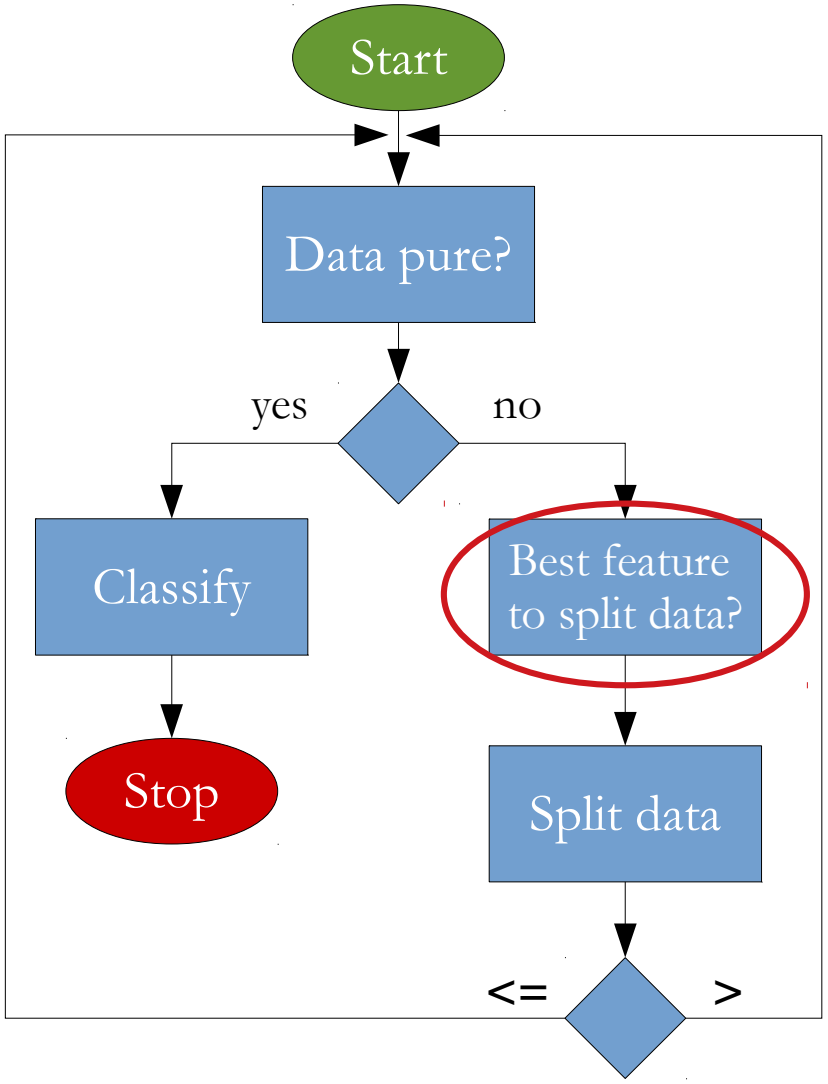


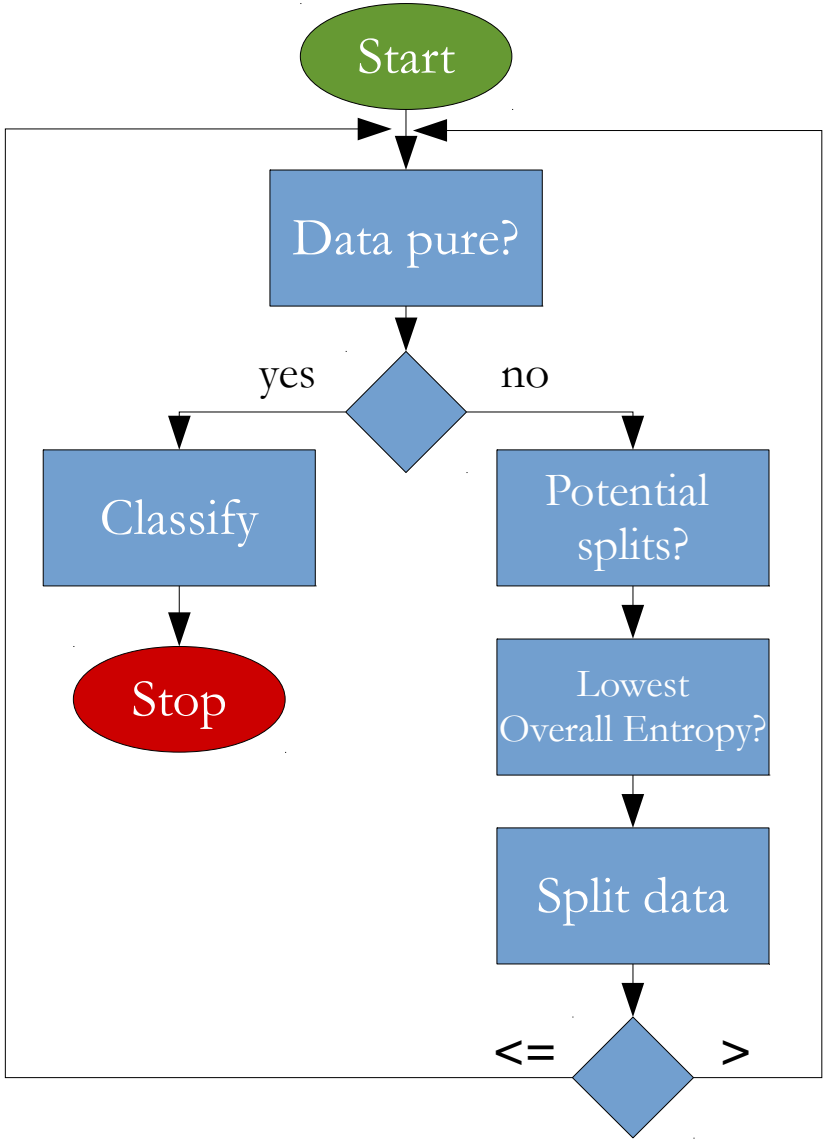


Weighted Sum Uncertainty Value

$$\text{Entropy} = \sum_{i=1}^{c} p_i \cdot (-\log_2 p_i)$$

$$\text{Overall Entropy} = \sum_{j=1}^{2} p_j \cdot \text{Entropy}_j$$


Start
Data pure?
yes
no
Classify
Best feature to split data?
Stop
Split data
<=
>


Start
Data pure?
yes
no
Classify
Stop
Potential splits?
Lowest Overall Entropy?
Split data
<=
>