
Data Mining

Under Guidance: Dr.BFM

Upload dataset here

Drag and drop file here

Limit 200MB per file • CSV

Browse files

iris.csv 3.9KB

Dataset loaded Table

	sepal.leng	sepal.wid	petal.leng	petal.wid	variety
40	5.0000	3.5000	1.3000	0.3000	Setosa
41	4.5000	2.3000	1.3000	0.3000	Setosa
42	4.4000	3.2000	1.3000	0.2000	Setosa
43	5.0000	3.5000	1.6000	0.6000	Setosa
44	5.1000	3.8000	1.9000	0.4000	Setosa
45	4.8000	3.0000	1.4000	0.3000	Setosa
46	5.1000	3.8000	1.6000	0.2000	Setosa
47	4.6000	3.2000	1.4000	0.2000	Setosa
48	5.3000	3.7000	1.5000	0.2000	Setosa
49	5.0000	3.3000	1.4000	0.2000	Setosa
50	7.0000	3.2000	4.7000	1.4000	Versico
51	6.4000	3.2000	4.5000	1.5000	Versico
52	6.9000	3.1000	4.9000	1.5000	Versico
53	5.5000	2.3000	4.0000	1.3000	Versico

Select Assignment

- Assignment No.1
- Assignment No.2
- Assignment No.3 and 4
- Assignment No.5

Assignment 3

Information gain

Information Gain

target feature: variety

descriptive_feature: sepal.length

split criterion: entropy

remaining impurity: `0.7091279999999999`

information gain: `0.8758720000000001`

sepal.length information gain: 0.8758720000000001

==========

target feature: variety

descriptive_feature: sepal.width

split criterion: entropy

remaining impurity: `1.067731`

information gain: `0.517269`

sepal.width information gain: 0.517269

==========

target feature: variety

descriptive_feature: petal.length

split criterion: entropy

remaining impurity: 0.138604

information gain: 1.446396

petal.length information gain: 1.446396

==========

target feature: variety

descriptive_feature: petal.width

split criterion: entropy

remaining impurity: 0.148849

information gain: 1.436151

petal.width information gain: 1.436151

==========

Maximum information gain is : 1.446396 for feature petal.length

Gini Index

Decision Tree

Made with Streamlit