Name: Ashish Verma Course:CS773 HW#6

Solution

Given Training Data

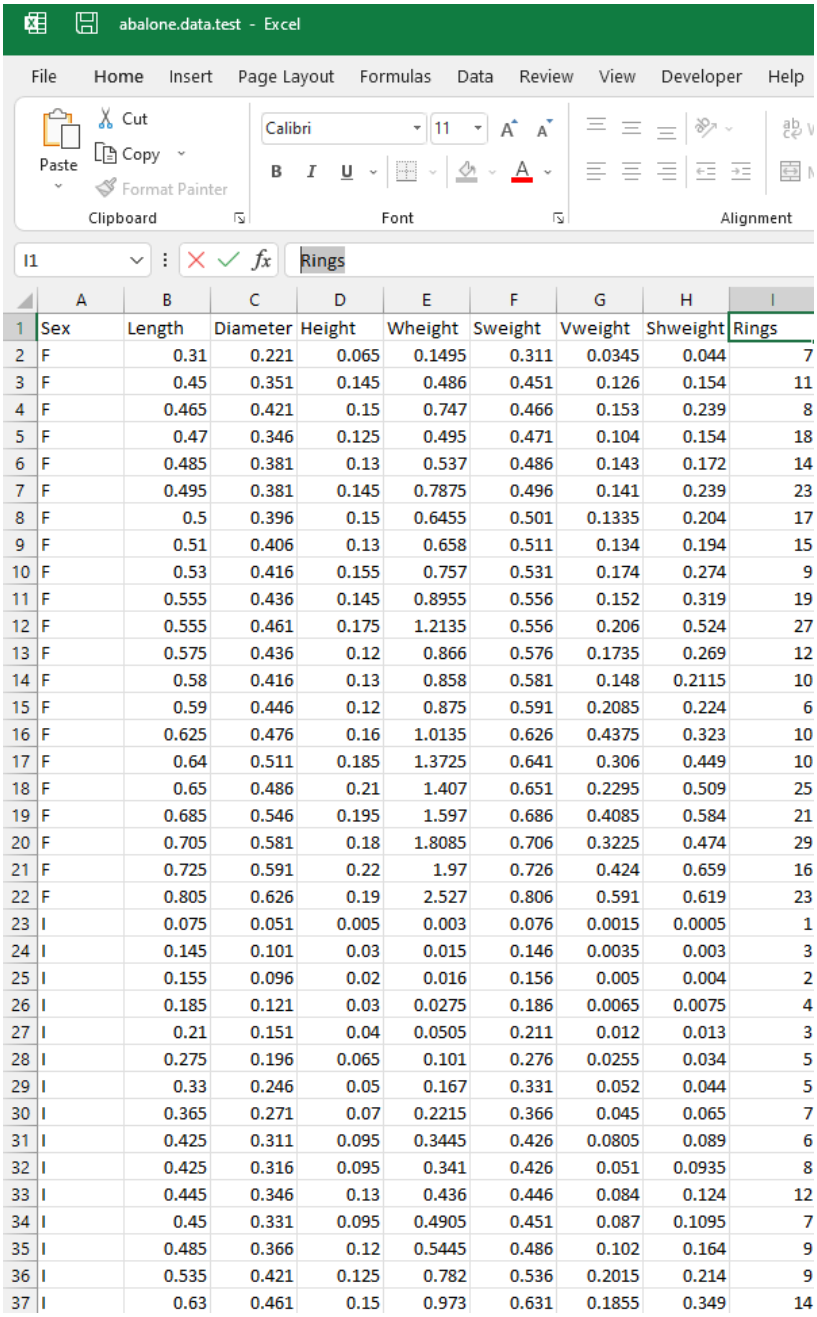
File Edit View ARFF-Viewer - C:\Users\Ashish\Documents\ODU\CS773\HW6\abalone data.arff

abalone data.arff

Relation: abalone

No.	1: Sex Nominal	2: Length Numeric	3: Diameter Numeric	4: Height Numeric	5: Whole weight Numeric	6: Shucked weight Numeric	7: Viscera weight Numeric	8: Shell weight Numeric	9: Class_Rings Numeric
4143	M	0.655	0.525	0.18	1.402	0.624	0.2935	0.365	13.0
4144	F	0.655	0.5	0.22	1.359	0.642	0.3255	0.405	13.0
4145	M	0.67	0.535	0.19	1.669	0.7465	0.2935	0.508	11.0
4146	M	0.67	0.525	0.2	1.7405	0.6205	0.297	0.657	11.0
4147	M	0.695	0.53	0.21	1.51	0.664	0.4095	0.385	10.0
4148	M	0.695	0.55	0.195	1.6645	0.727	0.36	0.445	11.0
4149	M	0.77	0.605	0.175	2.0505	0.8005	0.526	0.355	11.0
4150	I	0.28	0.215	0.07	0.124	0.063	0.0215	0.03	6.0
4151	I	0.33	0.23	0.08	0.14	0.0565	0.0365	0.046	7.0
4152	I	0.35	0.25	0.075	0.1695	0.0835	0.0355	0.041	6.0
4153	I	0.37	0.28	0.09	0.218	0.0995	0.0545	0.0615	7.0
4154	I	0.43	0.315	0.115	0.384	0.1885	0.0715	0.11	8.0
4155	I	0.435	0.33	0.095	0.393	0.219	0.075	0.0885	6.0
4156	I	0.44	0.35	0.11	0.3805	0.1575	0.0895	0.115	6.0
4157	M	0.475	0.37	0.11	0.4895	0.2185	0.107	0.146	8.0
4158	M	0.475	0.36	0.14	0.5135	0.241	0.1045	0.155	8.0
4159	I	0.48	0.355	0.11	0.4495	0.201	0.089	0.14	8.0
4160	F	0.56	0.44	0.135	0.8025	0.35	0.1615	0.259	9.0
4161	F	0.585	0.475	0.165	1.053	0.458	0.217	0.3	11.0
4162	F	0.585	0.455	0.17	0.9945	0.4255	0.263	0.2845	11.0
4163	M	0.385	0.255	0.1	0.3175	0.137	0.068	0.092	8.0
4164	I	0.39	0.31	0.085	0.344	0.181	0.0695	0.079	7.0
4165	I	0.39	0.29	0.1	0.2845	0.1255	0.0635	0.081	7.0
4166	I	0.405	0.3	0.085	0.3035	0.15	0.0505	0.088	7.0
4167	I	0.475	0.365	0.115	0.499	0.232	0.0885	0.156	10.0
4168	M	0.5	0.38	0.125	0.577	0.269	0.1265	0.1535	9.0
4169	F	0.515	0.4	0.125	0.615	0.2865	0.123	0.1765	8.0
4170	M	0.52	0.385	0.165	0.791	0.375	0.18	0.1815	10.0
4171	M	0.55	0.43	0.13	0.8395	0.3155	0.1955	0.2405	10.0
4172	M	0.56	0.43	0.155	0.8675	0.4	0.172	0.229	8.0
4173	F	0.565	0.45	0.165	0.887	0.37	0.239	0.249	11.0
4174	M	0.59	0.44	0.135	0.966	0.439	0.2145	0.2605	10.0
4175	M	0.6	0.475	0.205	1.176	0.5255	0.2875	0.308	9.0
4176	F	0.625	0.485	0.15	1.0945	0.531	0.261	0.296	10.0
4177	M	0.71	0.555	0.195	1.9485	0.9455	0.3765	0.495	12.0

Given Test Data



	A	B	C	D	E	F	G	H	I
1	Sex	Length	Diameter	Height	Wheight	Sweight	Vweight	Shweight	Rings
2	F	0.31	0.221	0.065	0.1495	0.311	0.0345	0.044	7
3	F	0.45	0.351	0.145	0.486	0.451	0.126	0.154	11
4	F	0.465	0.421	0.15	0.747	0.466	0.153	0.239	8
5	F	0.47	0.346	0.125	0.495	0.471	0.104	0.154	18
6	F	0.485	0.381	0.13	0.537	0.486	0.143	0.172	14
7	F	0.495	0.381	0.145	0.7875	0.496	0.141	0.239	23
8	F	0.5	0.396	0.15	0.6455	0.501	0.1335	0.204	17
9	F	0.51	0.406	0.13	0.658	0.511	0.134	0.194	15
10	F	0.53	0.416	0.155	0.757	0.531	0.174	0.274	9
11	F	0.555	0.436	0.145	0.8955	0.556	0.152	0.319	19
12	F	0.555	0.461	0.175	1.2135	0.556	0.206	0.524	27
13	F	0.575	0.436	0.12	0.866	0.576	0.1735	0.269	12
14	F	0.58	0.416	0.13	0.858	0.581	0.148	0.2115	10
15	F	0.59	0.446	0.12	0.875	0.591	0.2085	0.224	6
16	F	0.625	0.476	0.16	1.0135	0.626	0.4375	0.323	10
17	F	0.64	0.511	0.185	1.3725	0.641	0.306	0.449	10
18	F	0.65	0.486	0.21	1.407	0.651	0.2295	0.509	25
19	F	0.685	0.546	0.195	1.597	0.686	0.4085	0.584	21
20	F	0.705	0.581	0.18	1.8085	0.706	0.3225	0.474	29
21	F	0.725	0.591	0.22	1.97	0.726	0.424	0.659	16
22	F	0.805	0.626	0.19	2.527	0.806	0.591	0.619	23
23	I	0.075	0.051	0.005	0.003	0.076	0.0015	0.0005	1
24	I	0.145	0.101	0.03	0.015	0.146	0.0035	0.003	3
25	I	0.155	0.096	0.02	0.016	0.156	0.005	0.004	2
26	I	0.185	0.121	0.03	0.0275	0.186	0.0065	0.0075	4
27	I	0.21	0.151	0.04	0.0505	0.211	0.012	0.013	3
28	I	0.275	0.196	0.065	0.101	0.276	0.0255	0.034	5
29	I	0.33	0.246	0.05	0.167	0.331	0.052	0.044	5
30	I	0.365	0.271	0.07	0.2215	0.366	0.045	0.065	7
31	I	0.425	0.311	0.095	0.3445	0.426	0.0805	0.089	6
32	I	0.425	0.316	0.095	0.341	0.426	0.051	0.0935	8
33	I	0.445	0.346	0.13	0.436	0.446	0.084	0.124	12
34	I	0.45	0.331	0.095	0.4905	0.451	0.087	0.1095	7
35	I	0.485	0.366	0.12	0.5445	0.486	0.102	0.164	9
36	I	0.535	0.421	0.125	0.782	0.536	0.2015	0.214	9
37	I	0.63	0.461	0.15	0.973	0.631	0.1855	0.349	14

Since the attribute names of the training and test data are different, we will use Arffviewer to rename them as below

File Edit View ARFF-Viewer - C:\Users\Ashish\Documents\ODU\CS773\HV

abalone data.arff

Relation: abalone

No.	1: Sex Nominal	2: Length Numeric	3: Diameter Numeric	4: Height Numeric	5: Wheight Numeric	6: Sweight Numeric	7: Vweight Numeric	8: Shweight Numeric	9: **Rings** Numeric
1	M	0.455	0.365	0.095	0.514	0.2245	0.101	0.15	15.0
2	M	0.35	0.265	0.09	0.2255	0.0995	0.0485	0.07	7.0
3	F	0.53	0.42	0.135	0.677	0.2565	0.1415	0.21	9.0
4	M	0.44	0.365	0.125	0.516	0.2155	0.114	0.155	10.0
5	I	0.33	0.255	0.08	0.205	0.0895	0.0395	0.055	7.0
6	I	0.425	0.3	0.095	0.3515	0.141	0.0775	0.12	8.0
7	F	0.53	0.415	0.15	0.7775	0.237	0.1415	0.33	20.0
8	F	0.545	0.425	0.125	0.768	0.294	0.1495	0.26	16.0
9	M	0.475	0.37	0.125	0.5095	0.2165	0.1125	0.165	9.0
10	F	0.55	0.44	0.15	0.8945	0.3145	0.151	0.32	19.0
11	F	0.525	0.38	0.14	0.6065	0.194	0.1475	0.21	14.0
12	M	0.43	0.35	0.11	0.406	0.1675	0.081	0.135	10.0
13	M	0.49	0.38	0.135	0.5415	0.2175	0.095	0.19	11.0
14	F	0.535	0.405	0.145	0.6845	0.2725	0.171	0.205	10.0
15	F	0.47	0.355	0.1	0.4755	0.1675	0.0805	0.185	10.0
16	M	0.5	0.4	0.13	0.6645	0.258	0.133	0.24	12.0
17	I	0.355	0.28	0.085	0.2905	0.095	0.0395	0.115	7.0
18	F	0.44	0.34	0.1	0.451	0.188	0.087	0.13	10.0
19	M	0.365	0.295	0.08	0.2555	0.097	0.043	0.1	7.0
20	M	0.45	0.32	0.1	0.381	0.1705	0.075	0.115	9.0
21	M	0.355	0.28	0.095	0.2455	0.0955	0.062	0.075	11.0
22	I	0.38	0.275	0.1	0.2255	0.08	0.049	0.085	10.0
23	F	0.565	0.44	0.155	0.9395	0.4275	0.214	0.27	12.0
24	F	0.55	0.415	0.135	0.7635	0.318	0.21	0.2	9.0
25	F	0.615	0.48	0.165	1.1615	0.513	0.301	0.305	10.0
26	F	0.56	0.44	0.14	0.9285	0.3825	0.188	0.3	11.0
27	F	0.58	0.45	0.185	0.9955	0.3945	0.272	0.285	11.0
28	M	0.59	0.445	0.14	0.931	0.356	0.234	0.28	12.0
29	M	0.605	0.475	0.18	0.9365	0.394	0.219	0.295	15.0
30	M	0.575	0.425	0.14	0.8635	0.393	0.227	0.2	11.0
31	M	0.58	0.47	0.165	0.9975	0.3935	0.242	0.33	10.0
32	F	0.68	0.56	0.165	1.639	0.6055	0.2805	0.46	15.0

Since the training data label (Rings) is in continuous format, in order to classifier work we need to discretize the data using binning method by setting "**useEqualFrequency**" we will use python code to transform data.

```
import pandas as pd

# Load the abalone dataset
url = '/kaggle/input/abalone-1-data/abalone-1.data.csv'
column_names = ['Sex','Length','Diameter','Height','Wheight','Sweight','Vweight','Shweight','Rings']
abalone = pd.read_csv(url, names=column_names)

# Examine the Rings attribute
print(abalone['Rings'].describe())

abalone['Length'] = pd.to_numeric(abalone['Length'], errors='coerce')
abalone['Diameter'] = pd.to_numeric(abalone['Diameter'], errors='coerce')
abalone['Height'] = pd.to_numeric(abalone['Height'], errors='coerce')
abalone['Wheight'] = pd.to_numeric(abalone['Wheight'], errors='coerce')
abalone['Sweight'] = pd.to_numeric(abalone['Sweight'], errors='coerce')
abalone['Vweight'] = pd.to_numeric(abalone['Vweight'], errors='coerce')
abalone['Shweight'] = pd.to_numeric(abalone['Shweight'], errors='coerce')
abalone['Rings'] = pd.to_numeric(abalone['Rings'], errors='coerce')

# Choose a discretization method - e.g., equal-width binning
num_bins = 5

# Alternatively, use equal-frequency binning
abalone['Rings_binned_quantile'] = pd.qcut(abalone['Rings'], q=num_bins, labels=False)

# Display the discretized attribute
print(abalone[['Rings', 'Rings_binned_quantile']].head())

# Define category labels
labels = ["Very Low", "Low", "Medium", "High", "Very High"]

# Map the bins to these labels
abalone['Rings_category'] = abalone['Rings_binned_quantile'].map({
    0: "Very Low",
    1: "Low",
    2: "Medium",
    3: "High",
    4: "Very High"
})

# Display the categorized attribute
print(abalone[['Rings', 'Rings_binned_quantile', 'Rings_category']].head())

# Save the updated dataset
abalone.to_csv('abalone_discretized.csv', index=False)
```

Dataset after discretization looks as below

Relation: abalone_discretized

No.	1: Sex Nominal	2: Length Numeric	3: Diameter Numeric	4: Height Numeric	5: Wheight Numeric	6: Sweight Numeric	7: Vweight Numeric	8: Shweight Numeric	9: Rings Numeric	10: Rings_binned_quantile Numeric	11: **Rings_category** Nominal
1	M	0.455	0.365	0.095	0.514	0.2245	0.101	0.15	15.0	4.0	Very High
2	M	0.35	0.265	0.09	0.2255	0.0995	0.0485	0.07	7.0	0.0	Very Low
3	F	0.53	0.42	0.135	0.677	0.2565	0.1415	0.21	9.0	1.0	Low
4	M	0.44	0.365	0.125	0.516	0.2155	0.114	0.155	10.0	2.0	Medium
5	I	0.33	0.255	0.08	0.205	0.0895	0.0395	0.055	7.0	0.0	Very Low
6	I	0.425	0.3	0.095	0.3515	0.141	0.0775	0.12	8.0	1.0	Low
7	F	0.53	0.415	0.15	0.7775	0.237	0.1415	0.33	20.0	4.0	Very High
8	F	0.545	0.425	0.125	0.768	0.294	0.1495	0.26	16.0	4.0	Very High
9	M	0.475	0.37	0.125	0.5095	0.2165	0.1125	0.165	9.0	1.0	Low
10	F	0.55	0.44	0.15	0.8945	0.3145	0.151	0.32	19.0	4.0	Very High
11	F	0.525	0.38	0.14	0.6065	0.194	0.1475	0.21	14.0	4.0	Very High
12	M	0.43	0.35	0.11	0.406	0.1675	0.081	0.135	10.0	2.0	Medium
13	M	0.49	0.38	0.135	0.5415	0.2175	0.095	0.19	11.0	3.0	High
14	F	0.535	0.405	0.145	0.6845	0.2725	0.171	0.205	10.0	2.0	Medium
15	F	0.47	0.355	0.1	0.4755	0.1675	0.0805	0.185	10.0	2.0	Medium
16	M	0.5	0.4	0.13	0.6645	0.258	0.133	0.24	12.0	3.0	High
17	I	0.355	0.28	0.085	0.2905	0.095	0.0395	0.115	7.0	0.0	Very Low
18	F	0.44	0.34	0.1	0.451	0.188	0.087	0.13	10.0	2.0	Medium
19	M	0.365	0.295	0.08	0.2555	0.097	0.043	0.1	7.0	0.0	Very Low
20	M	0.45	0.32	0.1	0.381	0.1705	0.075	0.115	9.0	1.0	Low
21	M	0.355	0.28	0.095	0.2455	0.0955	0.062	0.075	11.0	3.0	High
22	I	0.38	0.275	0.1	0.2255	0.08	0.049	0.085	10.0	2.0	Medium
23	F	0.565	0.44	0.155	0.9395	0.4275	0.214	0.27	12.0	3.0	High
24	F	0.55	0.415	0.135	0.7635	0.318	0.21	0.2	9.0	1.0	Low
25	F	0.615	0.48	0.165	1.1615	0.513	0.301	0.305	10.0	2.0	Medium
26	F	0.56	0.44	0.14	0.9285	0.3825	0.188	0.3	11.0	3.0	High
27	F	0.58	0.45	0.185	0.9955	0.3945	0.272	0.285	11.0	3.0	High
28	M	0.59	0.445	0.14	0.931	0.356	0.234	0.28	12.0	3.0	High
29	M	0.605	0.475	0.18	0.9365	0.394	0.219	0.295	15.0	4.0	Very High
30	M	0.575	0.425	0.14	0.8635	0.393	0.227	0.2	11.0	3.0	High
31	M	0.58	0.47	0.165	0.9975	0.3935	0.242	0.33	10.0	2.0	Medium
32	F	0.68	0.56	0.165	1.639	0.6055	0.2805	0.46	15.0	4.0	Very High
33	M	0.665	0.525	0.165	1.338	0.5515	0.3575	0.35	18.0	4.0	Very High
34	F	0.68	0.55	0.175	1.798	0.815	0.3925	0.455	19.0	4.0	Very High
35	F	0.705	0.55	0.2	1.7095	0.633	0.4115	0.49	13.0	4.0	Very High
36	M	0.465	0.355	0.105	0.4795	0.227	0.124	0.125	8.0	1.0	Low
37	F	0.54	0.475	0.155	1.217	0.5305	0.3075	0.34	16.0	4.0	Very High
38	F	0.45	0.355	0.105	0.5225	0.237	0.1165	0.145	8.0	1.0	Low
39	F	0.575	0.445	0.135	0.883	0.381	0.2035	0.26	11.0	3.0	High
40	M	0.355	0.29	0.09	0.3275	0.134	0.086	0.09	9.0	1.0	Low
41	F	0.45	0.335	0.105	0.425	0.1865	0.091	0.115	9.0	1.0	Low
42	F	0.55	0.425	0.135	0.8515	0.362	0.196	0.27	14.0	4.0	Very High
43	I	0.24	0.175	0.045	0.07	0.0315	0.0235	0.02	5.0	0.0	Very Low
44	I	0.205	0.15	0.055	0.042	0.0255	0.015	0.012	5.0	0.0	Very Low
45	I	0.21	0.15	0.05	0.042	0.0175	0.0125	0.015	4.0	0.0	Very Low
46	I	0.39	0.295	0.095	0.203	0.0875	0.045	0.075	7.0	0.0	Very Low
47	M	0.47	0.37	0.12	0.5795	0.293	0.227	0.14	9.0	1.0	Low
48	F	0.46	0.375	0.12	0.4605	0.1775	0.11	0.15	7.0	0.0	Very Low

Rings_category is now our new class and we use the classifiers.

Similar processing, we need to do for test data

```python
import pandas as pd

# Load the abalone dataset
url = '/kaggle/input/abalone-data-test/abalone.data.test.csv'
column_names = ['Sex','Length','Diameter','Height','Wheight','Sweight','Vweight','Shweight','Rings']
abalone = pd.read_csv(url, names=column_names)

# Examine the Rings attribute
print(abalone['Rings'].describe())

abalone['Length'] = pd.to_numeric(abalone['Length'], errors='coerce')
abalone['Diameter'] = pd.to_numeric(abalone['Diameter'], errors='coerce')
abalone['Height'] = pd.to_numeric(abalone['Height'], errors='coerce')
abalone['Wheight'] = pd.to_numeric(abalone['Wheight'], errors='coerce')
abalone['Sweight'] = pd.to_numeric(abalone['Sweight'], errors='coerce')
abalone['Vweight'] = pd.to_numeric(abalone['Vweight'], errors='coerce')
abalone['Shweight'] = pd.to_numeric(abalone['Shweight'], errors='coerce')
abalone['Rings'] = pd.to_numeric(abalone['Rings'], errors='coerce')

# Choose a discretization method - e.g., equal-width binning
num_bins = 5

# Alternatively, use equal-frequency binning
abalone['Rings_binned_quantile'] = pd.qcut(abalone['Rings'], q=num_bins, labels=False)

# Display the discretized attribute
print(abalone[['Rings', 'Rings_binned_quantile']].head())


# Display the categorized attribute
print(abalone[['Rings', 'Rings_binned_quantile']].head())

# Save the updated dataset
abalone.to_csv('abalone_discretized_test.csv', index=False)
```

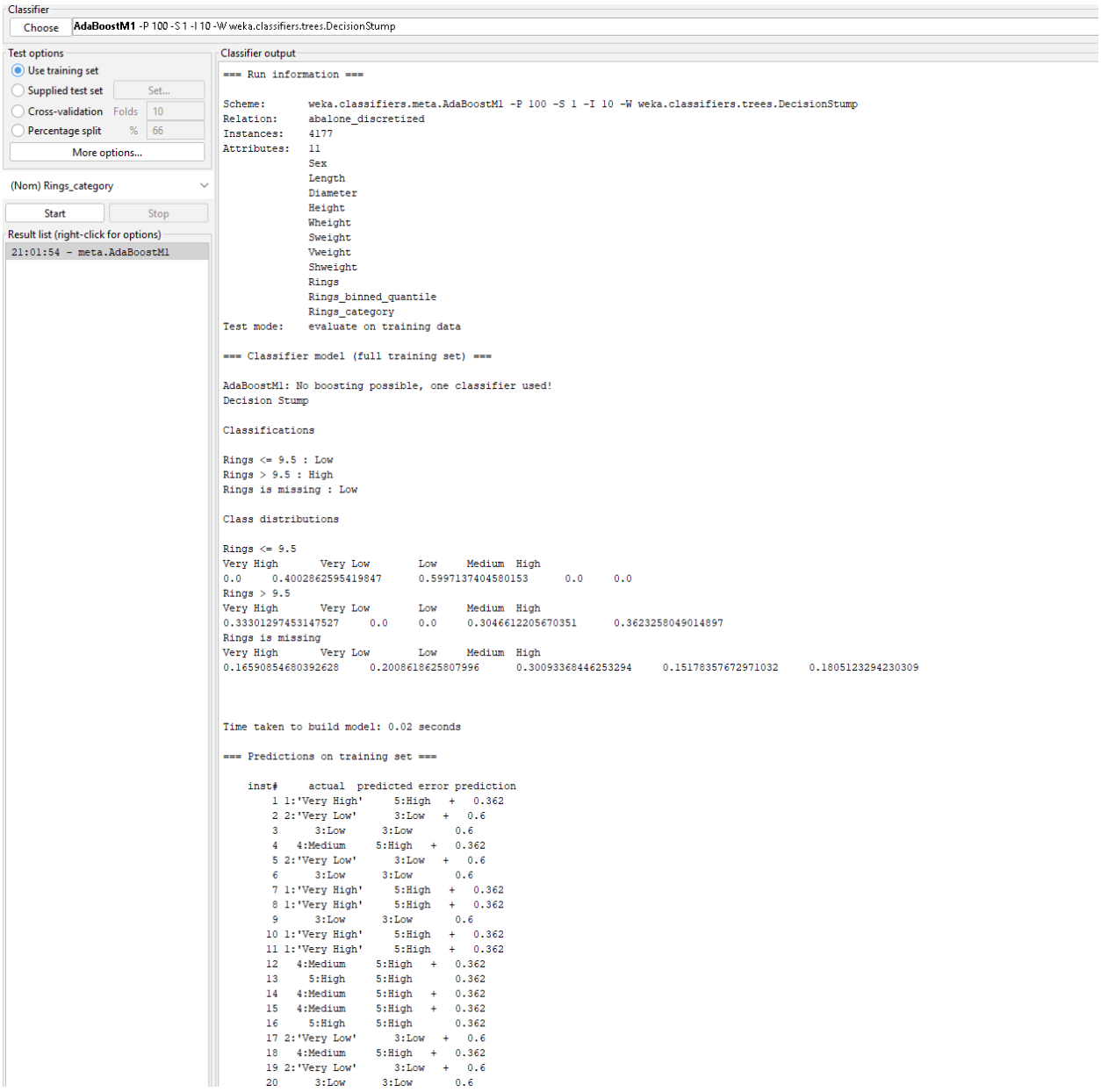
The test data will not have Rings_category which we need to predict.

abalone_discretized_test.csv

Relation: abalone_discretized_test

No.	1: Sex Nominal	2: Length Numeric	3: Diameter Numeric	4: Height Numeric	5: Wheight Numeric	6: Sweight Numeric	7: Vweight Numeric	8: Shweight Numeric	9: Rings Numeric	10: Rings_binned_quantile Numeric	11: **Ring Category** Nominal
1	F	0.31	0.221	0.065	0.1495	0.311	0.0345	0.044	7.0	0.0	Very Low
2	F	0.45	0.351	0.145	0.486	0.451	0.126	0.154	11.0	2.0	Medium
3	F	0.465	0.421	0.15	0.747	0.466	0.153	0.239	8.0	1.0	Low
4	F	0.47	0.346	0.125	0.495	0.471	0.104	0.154	18.0	3.0	High
5	F	0.485	0.381	0.13	0.537	0.486	0.143	0.172	14.0	2.0	Medium
6	F	0.495	0.381	0.145	0.7875	0.496	0.141	0.239	23.0	4.0	Very High
7	F	0.5	0.396	0.15	0.6455	0.501	0.1335	0.204	17.0	3.0	High
8	F	0.51	0.406	0.13	0.658	0.511	0.134	0.194	15.0	2.0	Medium
9	F	0.53	0.416	0.155	0.757	0.531	0.174	0.274	9.0	1.0	Low
10	F	0.555	0.436	0.145	0.8955	0.556	0.152	0.319	19.0	3.0	High
11	F	0.555	0.461	0.175	1.2135	0.556	0.206	0.524	27.0	4.0	Very High
12	F	0.575	0.436	0.12	0.866	0.576	0.1735	0.269	12.0	2.0	Medium
13	F	0.58	0.416	0.13	0.858	0.581	0.148	0.2115	10.0	1.0	Low
14	F	0.59	0.446	0.12	0.875	0.591	0.2085	0.224	6.0	0.0	Very Low
15	F	0.625	0.476	0.16	1.0135	0.626	0.4375	0.323	10.0	1.0	Low
16	F	0.64	0.511	0.185	1.3725	0.641	0.306	0.449	10.0	1.0	Low
17	F	0.65	0.486	0.21	1.407	0.651	0.2295	0.509	25.0	4.0	Very High
18	F	0.685	0.546	0.195	1.597	0.686	0.4085	0.584	21.0	4.0	Very High
19	F	0.705	0.581	0.18	1.8085	0.706	0.3225	0.474	29.0	4.0	Very High
20	F	0.725	0.591	0.22	1.97	0.726	0.424	0.659	16.0	3.0	High
21	F	0.805	0.626	0.19	2.527	0.806	0.591	0.619	23.0	4.0	Very High
22	I	0.075	0.051	0.005	0.003	0.076	0.0015	5.0E-4	1.0	0.0	Very Low
23	I	0.145	0.101	0.03	0.015	0.146	0.0035	0.003	3.0	0.0	Very Low
24	I	0.155	0.096	0.02	0.016	0.156	0.005	0.004	2.0	0.0	Very Low
25	I	0.185	0.121	0.03	0.0275	0.186	0.0065	0.0075	4.0	0.0	Very Low
26	I	0.21	0.151	0.04	0.0505	0.211	0.012	0.013	3.0	0.0	Very Low
27	I	0.275	0.196	0.065	0.101	0.276	0.0255	0.034	5.0	0.0	Very Low
28	I	0.33	0.246	0.05	0.167	0.331	0.052	0.044	5.0	0.0	Very Low
29	I	0.365	0.271	0.07	0.2215	0.366	0.045	0.065	7.0	0.0	Very Low
30	I	0.425	0.311	0.095	0.3445	0.426	0.0805	0.089	6.0	0.0	Very Low
31	I	0.425	0.316	0.095	0.341	0.426	0.051	0.0935	8.0	1.0	Low
32	I	0.445	0.346	0.13	0.436	0.446	0.084	0.124	12.0	2.0	Medium
33	I	0.45	0.331	0.095	0.4905	0.451	0.087	0.1095	7.0	0.0	Very Low
34	I	0.485	0.366	0.12	0.5445	0.486	0.102	0.164	9.0	1.0	Low
35	I	0.535	0.421	0.125	0.782	0.536	0.2015	0.214	9.0	1.0	Low
36	I	0.63	0.461	0.15	0.973	0.631	0.1855	0.349	14.0	2.0	Medium
37	M	0.16	0.111	0.02	0.025	0.161	0.006	0.0065	5.0	0.0	Very Low
38	M	0.24	0.156	0.055	0.0555	0.241	0.0105	0.014	4.0	0.0	Very Low
39	M	0.28	0.196	0.075	0.1	0.281	0.025	0.029	5.0	0.0	Very Low
40	M	0.305	0.236	0.085	0.162	0.306	0.04	0.049	6.0	0.0	Very Low
41	M	0.33	0.226	0.085	0.148	0.331	0.035	0.044	4.0	0.0	Very Low
42	M	0.34	0.256	0.07	0.221	0.341	0.041	0.084	6.0	0.0	Very Low
43	M	0.44	0.331	0.105	0.4395	0.441	0.0725	0.1305	7.0	0.0	Very Low

Training AdaBoostM1



Classifier

Choose **AdaBoostM1** -P 100 -S 1 -I 10 -W weka.classifiers.trees.DecisionStump

Test options

- Use training set
- Supplied test set — Set...
- Cross-validation — Folds 10
- Percentage split — % 66

More options...

(Nom) Rings_category

Start | Stop

Result list (right-click for options)

21:01:54 - meta.AdaBoostM1

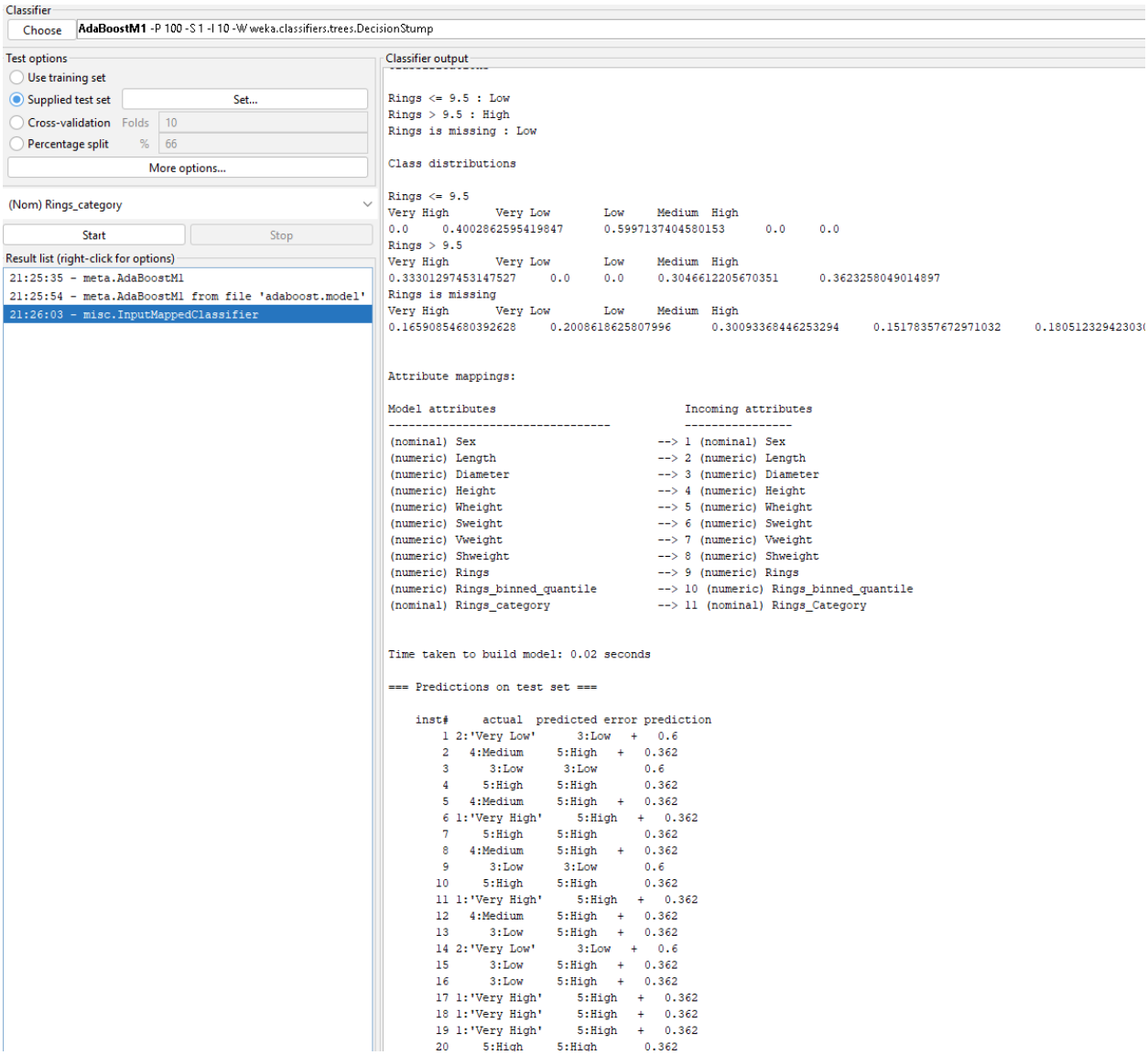
Classifier output

```
=== Run information ===

Scheme:       weka.classifiers.meta.AdaBoostM1 -P 100 -S 1 -I 10 -W weka.classifiers.trees.DecisionStump
Relation:     abalone_discretized
Instances:    4177
Attributes:   11
              Sex
              Length
              Diameter
              Height
              Wheight
              Sweight
              Vweight
              Shweight
              Rings
              Rings_binned_quantile
              Rings_category
Test mode:    evaluate on training data

=== Classifier model (full training set) ===

AdaBoostM1: No boosting possible, one classifier used!
Decision Stump

Classifications

Rings <= 9.5 : Low
Rings > 9.5 : High
Rings is missing : Low

Class distributions

Rings <= 9.5
Very High       Very Low        Low     Medium  High
0.0     0.4002862595419847      0.5997137404580153      0.0     0.0
Rings > 9.5
Very High       Very Low        Low     Medium  High
0.33301297453147527     0.0     0.0     0.3046612205670351      0.3623258049014897
Rings is missing
Very High       Very Low        Low     Medium  High
0.1659085468392628      0.2008618625807996      0.30093368446253294     0.15178357672971032     0.1805123294230309


Time taken to build model: 0.02 seconds

=== Predictions on training set ===

    inst#     actual  predicted error prediction
        1 1:'Very High'     5:High   +   0.362
        2 2:'Very Low'      3:Low    +   0.6
        3      3:Low      3:Low        0.6
        4   4:Medium     5:High   +   0.362
        5 2:'Very Low'      3:Low    +   0.6
        6      3:Low      3:Low        0.6
        7 1:'Very High'     5:High   +   0.362
        8 1:'Very High'     5:High   +   0.362
        9      3:Low      3:Low        0.6
       10 1:'Very High'     5:High   +   0.362
       11 1:'Very High'     5:High   +   0.362
       12   4:Medium     5:High   +   0.362
       13     5:High     5:High       0.362
       14   4:Medium     5:High   +   0.362
       15   4:Medium     5:High   +   0.362
       16     5:High     5:High       0.362
       17 2:'Very Low'      3:Low    +   0.6
       18   4:Medium     5:High   +   0.362
       19 2:'Very Low'      3:Low    +   0.6
       20      3:Low      3:Low        0.6
```

```
=== Evaluation on training set ===

Time taken to test model on training data: 1.13 seconds

=== Summary ===

Correctly Classified Instances        2011               48.1446 %
Incorrectly Classified Instances      2166               51.8554 %
Kappa statistic                          0.3168
Mean absolute error                      0.2289
Root mean squared error                  0.3383
Relative absolute error                 72.806  %
Root relative squared error             85.3273 %
Total Number of Instances             4177

=== Detailed Accuracy By Class ===

               TP Rate  FP Rate  Precision  Recall   F-Measure  MCC      ROC Area  PRC Area  Class
               0.000    0.000    ?          0.000    ?          ?        0.801     0.333     Very High
               0.000    0.000    ?          0.000    ?          ?        0.812     0.400     Very Low
               1.000    0.287    0.600      1.000    0.750      0.654    0.856     0.600     Low
               0.000    0.000    ?          0.000    ?          ?        0.796     0.305     Medium
               1.000    0.388    0.362      1.000    0.532      0.471    0.806     0.362     High
Weighted Avg.  0.481    0.156    ?          0.481    ?          ?        0.820     0.428

=== Confusion Matrix ===

    a    b    c    d    e   <-- classified as
    0    0    0    0  693 |    a = Very High
    0    0  839    0    0 |    b = Very Low
    0    0 1257    0    0 |    c = Low
    0    0    0    0  634 |    d = Medium
    0    0    0    0  754 |    e = High
```

Save the model and load the model

Test AdaBoostM1



Prediction Output from test sample

inst#	actual	predicted error	prediction
1	2:'Very Low'	3:Low	0.6
2	4:Medium	5:High +	0.362
3	3:Low	3:Low	0.6
4	5:High	5:High	0.362
5	4:Medium	5:High +	0.362
6	1:'Very High	' 5:High	0.362

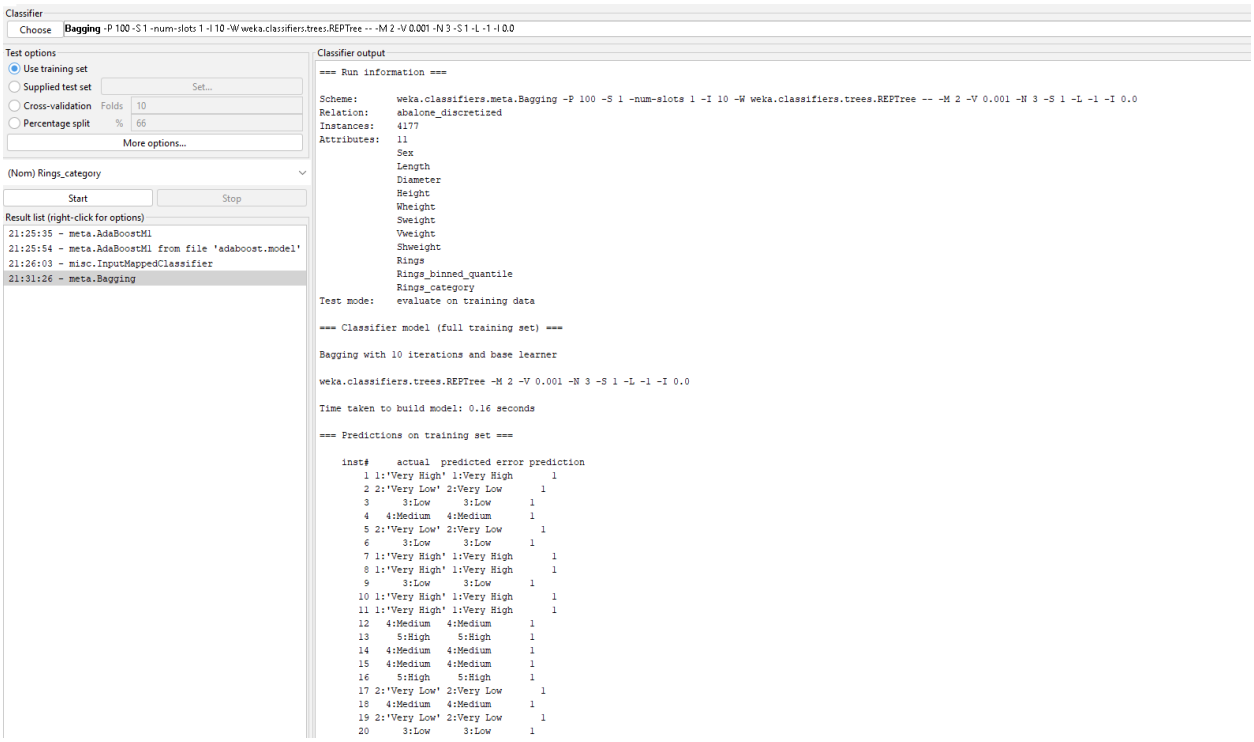
7	5:High	5:High	0.362
8	4:Medium	5:High +	0.362
9	3:Low	3:Low	0.6
10	5:High	5:High	0.362
11	1:'Very High	' 5:High	0.362
12	4:Medium	5:High +	0.362
13	3:Low	5:High +	0.362
14	2:'Very Low'	3:Low	0.6
15	3:Low	5:High +	0.362
16	3:Low	5:High +	0.362
17	1:'Very High	' 5:High	0.362
18	1:'Very High	' 5:High	0.362
19	1:'Very High	' 5:High	0.362
20	5:High	5:High	0.362
21	1:'Very High	' 5:High	0.362
22	2:'Very Low'	3:Low	0.6
23	2:'Very Low'	3:Low	0.6
24	2:'Very Low'	3:Low	0.6
25	2:'Very Low'	3:Low	0.6
26	2:'Very Low'	3:Low	0.6
27	2:'Very Low'	3:Low	0.6
28	2:'Very Low'	3:Low	0.6
29	2:'Very Low'	3:Low	0.6
30	2:'Very Low'	3:Low	0.6
31	3:Low	3:Low	0.6
32	4:Medium	5:High +	0.362
33	2:'Very Low'	3:Low	0.6
34	3:Low	3:Low	0.6
35	3:Low	3:Low	0.6

36	4:Medium	5:High +	0.362
37	2:'Very Low'	3:Low	0.6
38	2:'Very Low'	3:Low	0.6
39	2:'Very Low'	3:Low	0.6
40	2:'Very Low'	3:Low	0.6
41	2:'Very Low'	3:Low	0.6
42	2:'Very Low'	3:Low	0.6
43	2:'Very Low'	3:Low	0.6
44	3:Low	3:Low	0.6
45	5:High	5:High	0.362
46	3:Low	5:High +	0.362
47	5:High	5:High	0.362
48	4:Medium	5:High +	0.362
49	3:Low	3:Low	0.6
50	5:High	5:High	0.362
51	4:Medium	5:High +	0.362
52	4:Medium	5:High +	0.362
53	5:High	5:High	0.362
54	5:High	5:High	0.362
55	2:'Very Low'	3:Low	0.6
56	5:High	5:High	0.362
57	4:Medium	5:High +	0.362
58	1:'Very High	' 5:High	0.362
59	4:Medium	5:High +	0.362
60	4:Medium	5:High +	0.362
61	5:High	5:High	0.362
62	5:High	5:High	0.362
63	4:Medium	5:High +	0.362
64	1:'Very High	' 5:High	0.362
65	1:'Very High	' 5:High	0.362
66	4:Medium	5:High +	0.362
67	3:Low	3:Low	0.6

68	1:'Very High	' 5:High	0.362
69	1:'Very High	' 5:High	0.362
70	5:High	5:High	0.362
71	1:'Very High	' 5:High	0.362
72	5:High	5:High	0.362
73	1:'Very High	' 5:High	0.362
74	1:'Very High	' 5:High	0.362
75	4:Medium	5:High +	0.362
76	1:'Very High	' 5:High	0.362
77	4:Medium	5:High +	0.362
78	1:'Very High	' 5:High	0.362
79	4:Medium	5:High +	0.362
80	4:Medium	5:High +	0.36

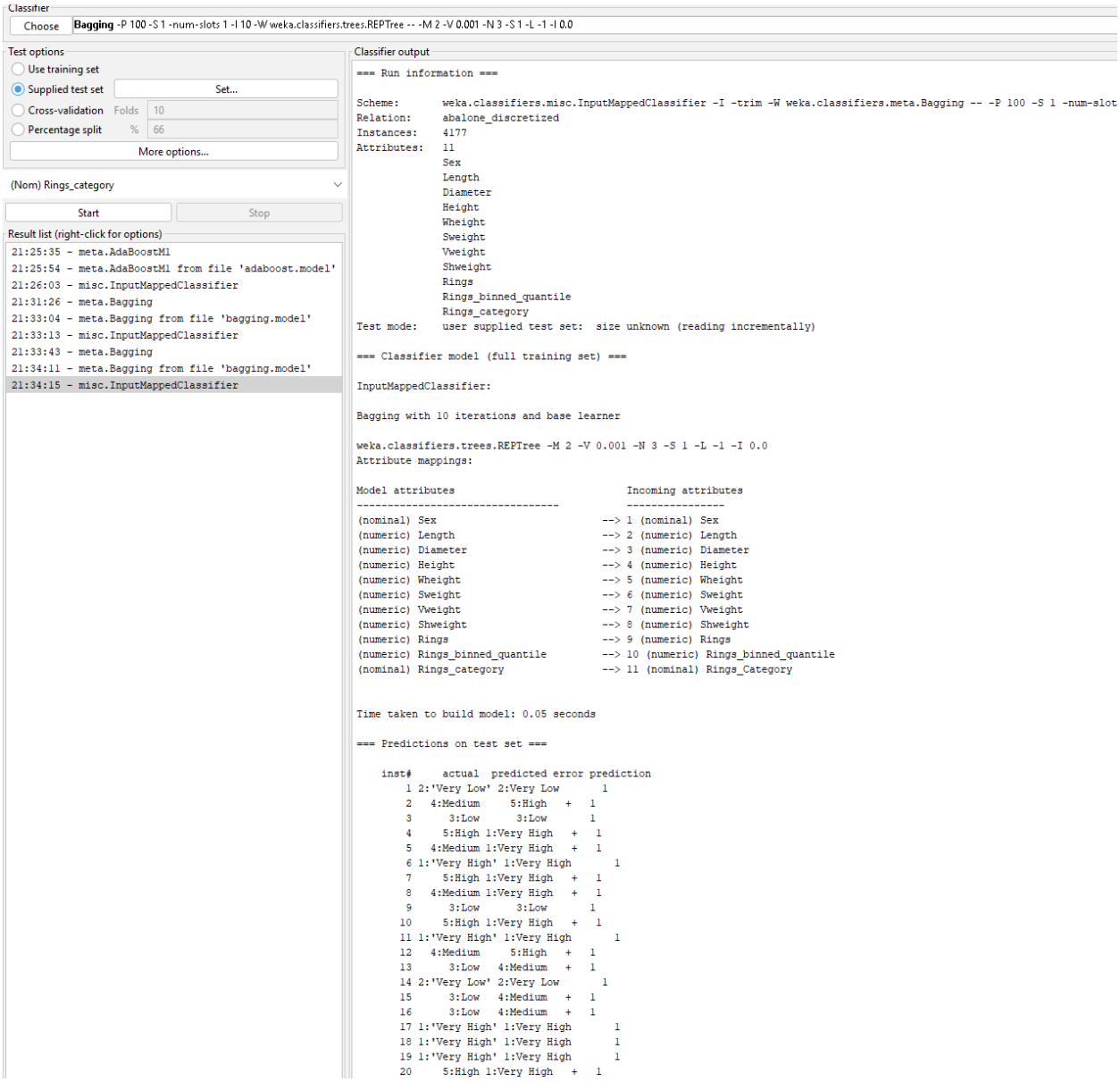
Correctly Classified Instances 22 27.5 %

Training Bagging




```
=== Evaluation on training set ===

Time taken to test model on training data: 0.85 seconds

=== Summary ===

Correctly Classified Instances        4177               100      %
Incorrectly Classified Instances         0                 0      %
Kappa statistic                          1
Mean absolute error                      0
Root mean squared error                  0
Relative absolute error                  0      %
Root relative squared error              0      %
Total Number of Instances             4177

=== Detailed Accuracy By Class ===

                 TP Rate  FP Rate  Precision  Recall   F-Measure  MCC      ROC Area  PRC Area  Class
                 1.000    0.000    1.000      1.000    1.000      1.000    1.000     1.000     Very High
                 1.000    0.000    1.000      1.000    1.000      1.000    1.000     1.000     Very Low
                 1.000    0.000    1.000      1.000    1.000      1.000    1.000     1.000     Low
                 1.000    0.000    1.000      1.000    1.000      1.000    1.000     1.000     Medium
                 1.000    0.000    1.000      1.000    1.000      1.000    1.000     1.000     High
Weighted Avg.    1.000    0.000    1.000      1.000    1.000      1.000    1.000     1.000

=== Confusion Matrix ===

    a    b    c    d    e   <-- classified as
  693    0    0    0    0 |    a = Very High
    0  839    0    0    0 |    b = Very Low
    0    0 1257    0    0 |    c = Low
    0    0    0  634    0 |    d = Medium
    0    0    0    0  754 |    e = High
```

Test Bagging



Prediction Output from test sample

inst# actual predicted error prediction

1 2:'Very Low' 2:Very Low 1

2 4:Medium 5:High + 1

3 3:Low 3:Low 1

4 5:High 1:Very High + 1

5 4:Medium 1:Very High + 1

6 1:'Very High' 1:Very High 1

7 5:High 1:Very High + 1

8 4:Medium 1:Very High + 1

9 3:Low 3:Low 1

10 5:High 1:Very High + 1

11 1:'Very High' 1:Very High 1

12 4:Medium 5:High + 1

13 3:Low 4:Medium + 1

14 2:'Very Low' 2:Very Low 1

15 3:Low 4:Medium + 1

16 3:Low 4:Medium + 1

17 1:'Very High' 1:Very High 1

18 1:'Very High' 1:Very High 1

19 1:'Very High' 1:Very High 1

20 5:High 1:Very High + 1

21 1:'Very High' 1:Very High 1

22 2:'Very Low' 2:Very Low 1

23 2:'Very Low' 2:Very Low 1

24 2:'Very Low' 2:Very Low 1

25 2:'Very Low' 2:Very Low 1

26 2:'Very Low' 2:Very Low 1

27 2:'Very Low' 2:Very Low 1

28 2:'Very Low' 2:Very Low 1

29 2:'Very Low' 2:Very Low 1

30 2:'Very Low' 2:Very Low 1

31 3:Low 3:Low 1

32 4:Medium 5:High + 1

33 2:'Very Low' 2:Very Low 1

34 3:Low 3:Low 1

35 3:Low 3:Low 1

36 4:Medium 1:Very High + 1

37 2:'Very Low' 2:Very Low 1

38 2:'Very Low' 2:Very Low 1

39 2:'Very Low' 2:Very Low 1

40 2:'Very Low' 2:Very Low 1

41 2:'Very Low' 2:Very Low 1

42 2:'Very Low' 2:Very Low 1

43 2:'Very Low' 2:Very Low 1

44 3:Low 3:Low 1

45 5:High 1:Very High + 1

46 3:Low 4:Medium + 1

47 5:High 1:Very High + 1

48 4:Medium 1:Very High + 1

49 3:Low 3:Low 1

50 5:High 1:Very High + 1

51 4:Medium 1:Very High + 1

52 4:Medium 5:High + 1

53 5:High 1:Very High + 1

54 5:High 1:Very High + 1

55 2:'Very Low' 2:Very Low 1

56 5:High 1:Very High + 1

57 4:Medium 1:Very High + 1

58 1:'Very High' 1:Very High 1

59 4:Medium 1:Very High + 1

60 4:Medium 1:Very High + 1

61 5:High 1:Very High + 1

62 5:High 1:Very High + 1

63 4:Medium 1:Very High + 1

64 1:'Very High' 1:Very High 1

65 1:'Very High' 1:Very High 1

66 4:Medium 1:Very High + 1

67 3:Low 3:Low 1

68 1:'Very High' 1:Very High 1

69 1:'Very High' 1:Very High 1

70 5:High 1:Very High + 1

71 1:'Very High' 1:Very High 1

72 5:High 1:Very High + 1

73 1:'Very High' 1:Very High 1

74 1:'Very High' 1:Very High 1

75 4:Medium 5:High + 1

76 1:'Very High' 1:Very High 1

77 4:Medium 5:High + 1

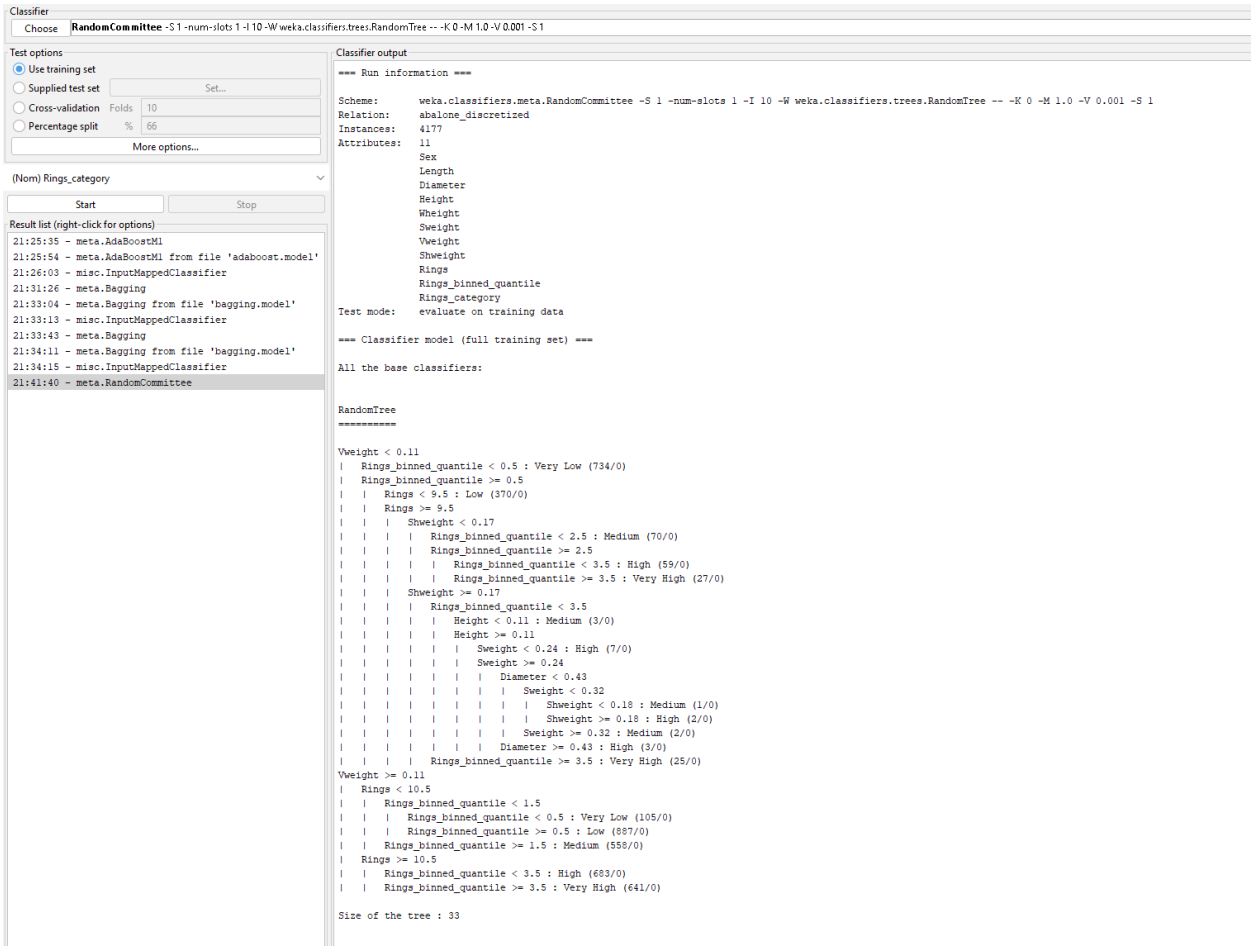
78 1:'Very High' 1:Very High 1

79 4:Medium 5:High + 1

80 4:Medium 1:Very High + 1

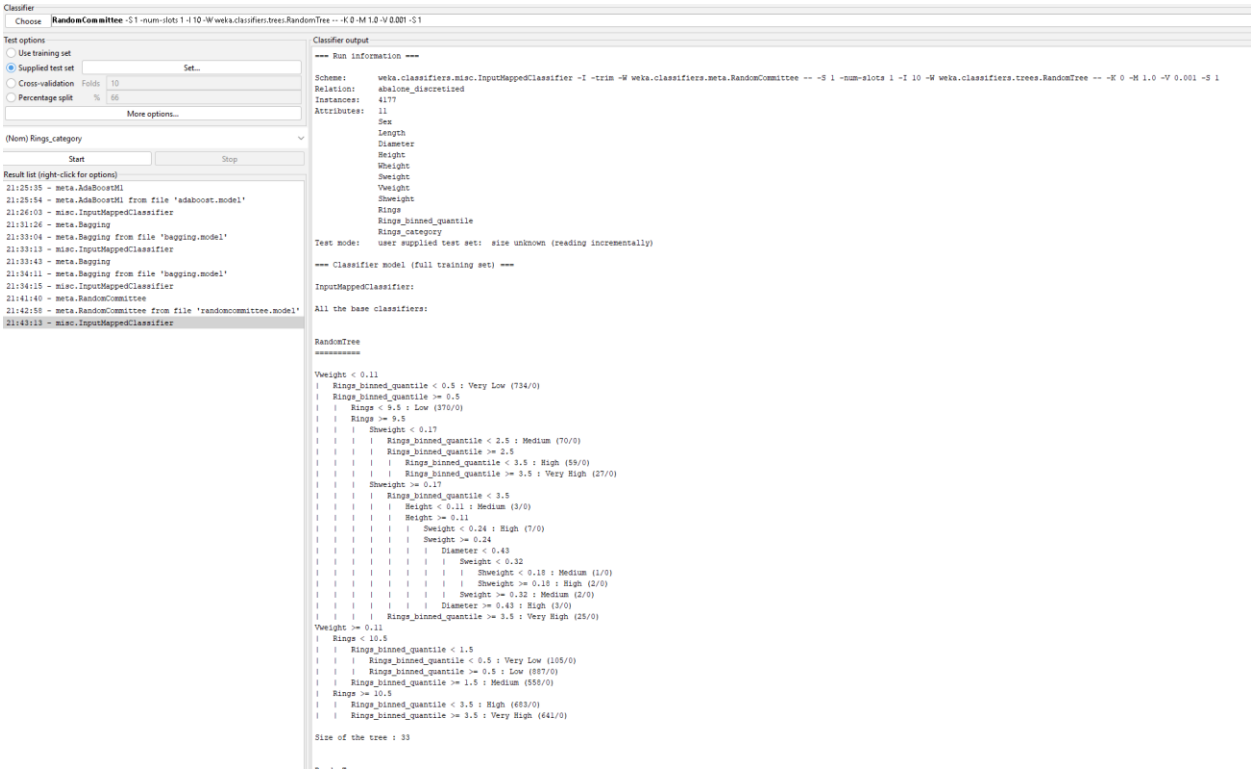
Correctly Classified Instances 44 55 %

Train Randomcommittee



```
Time taken to test model on training data: 0.93 seconds

=== Summary ===

Correctly Classified Instances        4177               100      %
Incorrectly Classified Instances         0                 0      %
Kappa statistic                          1
Mean absolute error                      0
Root mean squared error                  0
Relative absolute error                  0      %
Root relative squared error              0      %
Total Number of Instances             4177

=== Detailed Accuracy By Class ===

                 TP Rate  FP Rate  Precision  Recall   F-Measure  MCC      ROC Area  PRC Area  Class
                 1.000    0.000    1.000      1.000    1.000      1.000    1.000     1.000     Very High
                 1.000    0.000    1.000      1.000    1.000      1.000    1.000     1.000     Very Low
                 1.000    0.000    1.000      1.000    1.000      1.000    1.000     1.000     Low
                 1.000    0.000    1.000      1.000    1.000      1.000    1.000     1.000     Medium
                 1.000    0.000    1.000      1.000    1.000      1.000    1.000     1.000     High
Weighted Avg.    1.000    0.000    1.000      1.000    1.000      1.000    1.000     1.000

=== Confusion Matrix ===

   a   b   c   d   e   <-- classified as
 693   0   0   0   0 |   a = Very High
   0 839   0   0   0 |   b = Very Low
   0   0 1257   0   0 |   c = Low
   0   0   0 634   0 |   d = Medium
   0   0   0   0 754 |   e = High
```

Test Randomcommittee



Prediction Results from randomcommitte

=== Predictions on test set ===

inst#	actual	predicted	error	prediction
1	2:'Very Low'	2:Very Low		1
2	4:Medium	4:Medium		0.5
3	3:Low	3:Low		1
4	5:High	5:High		0.6
5	4:Medium	4:Medium		0.4
6	1:'Very High'	1:Very High		1
7	5:High	5:High		0.6
8	4:Medium	4:Medium		0.4
9	3:Low	3:Low		1
10	5:High	1:Very High	+	0.5
11	1:'Very High'	1:Very High		1
12	4:Medium	4:Medium		0.6
13	3:Low	4:Medium	+	0.6
14	2:'Very Low'	2:Very Low		1
15	3:Low	4:Medium	+	0.7
16	3:Low	4:Medium	+	0.7
17	1:'Very High'	1:Very High		1
18	1:'Very High'	1:Very High		1
19	1:'Very High'	1:Very High		1
20	5:High	1:Very High	+	0.5
21	1:'Very High'	1:Very High		1
22	2:'Very Low'	2:Very Low		1
23	2:'Very Low'	2:Very Low		1
24	2:'Very Low'	2:Very Low		1

25 2:'Very Low' 2:Very Low 1

26 2:'Very Low' 2:Very Low 1

27 2:'Very Low' 2:Very Low 1

28 2:'Very Low' 2:Very Low 1

29 2:'Very Low' 2:Very Low 1

30 2:'Very Low' 2:Very Low 1

31 3:Low 3:Low 1

32 4:Medium 5:High + 0.7

33 2:'Very Low' 2:Very Low 1

34 3:Low 3:Low 1

35 3:Low 3:Low 1

36 4:Medium 4:Medium 0.6

37 2:'Very Low' 2:Very Low 1

38 2:'Very Low' 2:Very Low 1

39 2:'Very Low' 2:Very Low 1

40 2:'Very Low' 2:Very Low 1

41 2:'Very Low' 2:Very Low 1

42 2:'Very Low' 2:Very Low 1

43 2:'Very Low' 2:Very Low 1

44 3:Low 3:Low 1

45 5:High 5:High 0.6

46 3:Low 4:Medium + 0.6

47 5:High 5:High 0.6

48 4:Medium 4:Medium 0.4

49 3:Low 3:Low 1

50 5:High 1:Very High + 0.5

51 4:Medium 4:Medium 0.4

52 4:Medium 4:Medium 0.5

53 5:High 1:Very High + 0.5

54 5:High 1:Very High + 0.5

55 2:'Very Low' 2:Very Low 1

56 5:High 1:Very High + 0.5

57 4:Medium 4:Medium 0.4

58 1:'Very High' 1:Very High 1

59 4:Medium 4:Medium 0.5

60 4:Medium 4:Medium 0.4

61 5:High 1:Very High + 0.5

62 5:High 1:Very High + 0.5

63 4:Medium 4:Medium 0.6

64 1:'Very High' 1:Very High 1

65 1:'Very High' 1:Very High 1

66 4:Medium 4:Medium 0.5

67 3:Low 3:Low 1

68 1:'Very High' 1:Very High 1

69 1:'Very High' 1:Very High 1

70 5:High 1:Very High + 0.6

71 1:'Very High' 1:Very High 1

72 5:High 1:Very High + 0.6

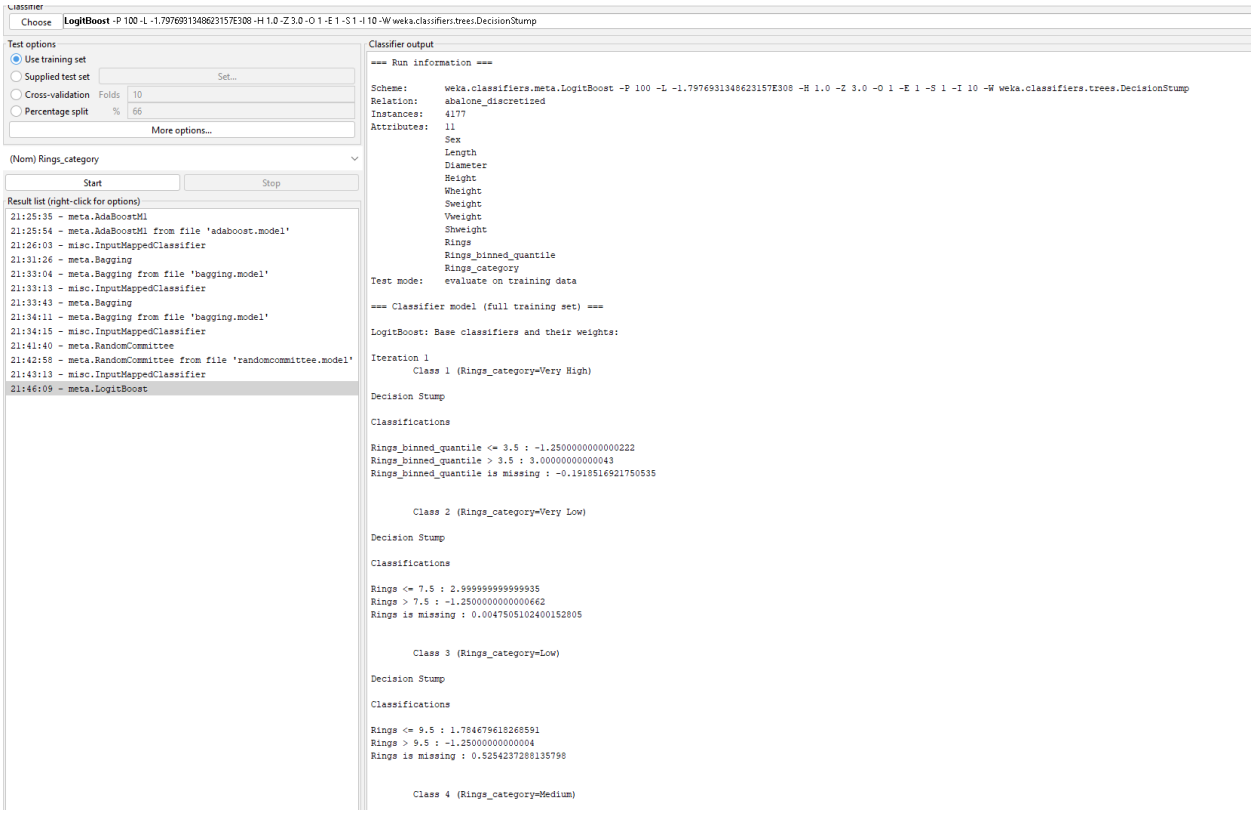
73 1:'Very High' 1:Very High 1

74 1:'Very High' 1:Very High 1

75	4:Medium	4:Medium		0.5
76	1:'Very High'	1:Very High		1
77	4:Medium	4:Medium		0.6
78	1:'Very High'	1:Very High		1
79	4:Medium	4:Medium		0.5
80	4:Medium	1:Very High	+	0.4

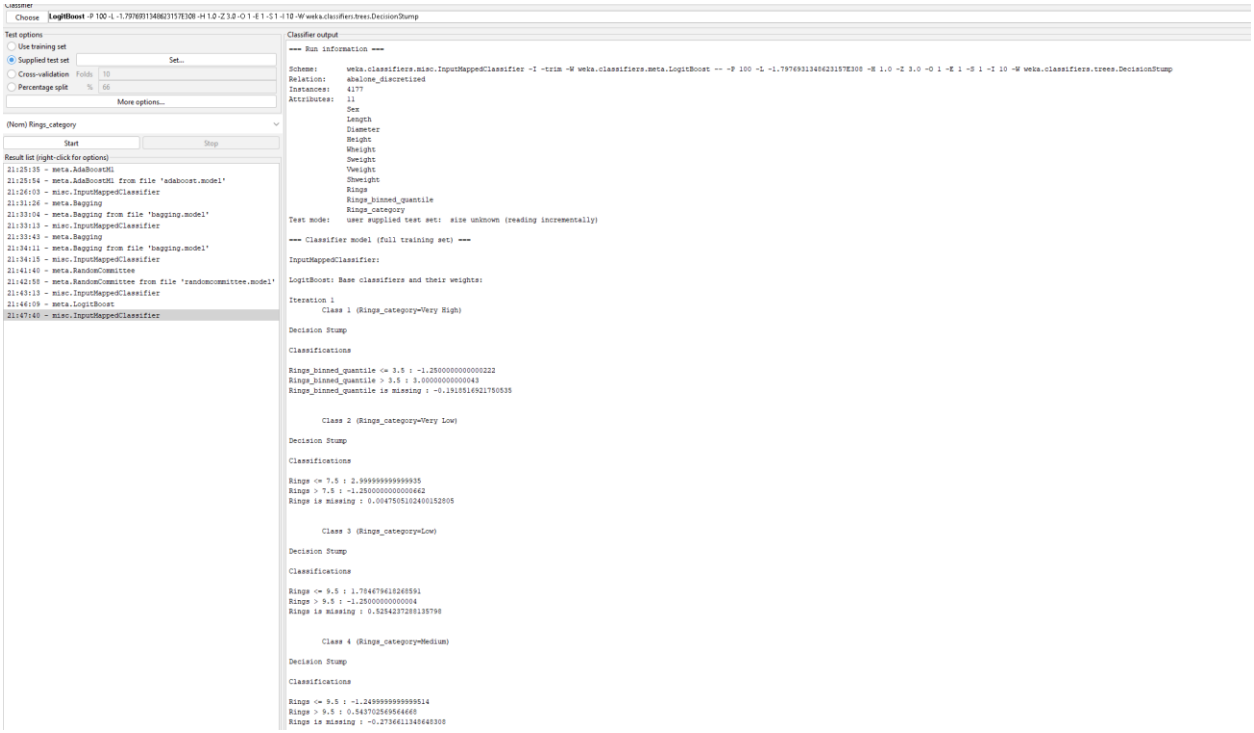
Correctly Classified Instances 64 80 %

Training LogitBoost



```
Time taken to test model on training data: 0.89 seconds

=== Summary ===

Correctly Classified Instances        4177               100      %
Incorrectly Classified Instances         0                 0      %
Kappa statistic                          1
Mean absolute error                      0.0002
Root mean squared error                  0.0003
Relative absolute error                  0.0503 %
Root relative squared error              0.0676 %
Total Number of Instances             4177

=== Detailed Accuracy By Class ===

               TP Rate  FP Rate  Precision  Recall   F-Measure  MCC      ROC Area  PRC Area  Class
               1.000    0.000    1.000      1.000    1.000      1.000    1.000     1.000     Very High
               1.000    0.000    1.000      1.000    1.000      1.000    1.000     1.000     Very Low
               1.000    0.000    1.000      1.000    1.000      1.000    1.000     1.000     Low
               1.000    0.000    1.000      1.000    1.000      1.000    1.000     1.000     Medium
               1.000    0.000    1.000      1.000    1.000      1.000    1.000     1.000     High
Weighted Avg.  1.000    0.000    1.000      1.000    1.000      1.000    1.000     1.000

=== Confusion Matrix ===

   a   b   c   d   e   <-- classified as
 693   0   0   0   0 |   a = Very High
   0 839   0   0   0 |   b = Very Low
   0   0 1257   0   0 |   c = Low
   0   0   0 634   0 |   d = Medium
   0   0   0   0 754 |   e = High
```

Test LogitBoost



Prediction Results

=== Predictions on test set ===

inst# actual predicted error prediction

1	2:'Very Low'	2:Very Low			1
2	4:Medium	5:High	+		0.797
3	3:Low	3:Low			1
4	5:High	5:High			0.895
5	4:Medium	1:Very High	+		0.7
6	1:'Very High'	1:Very High			1
7	5:High	5:High			0.895
8	4:Medium	1:Very High	+		0.7
9	3:Low	3:Low			1
10	5:High	5:High			0.895
11	1:'Very High'	1:Very High			1
12	4:Medium	5:High	+		0.797
13	3:Low	4:Medium	+		0.847
14	2:'Very Low'	2:Very Low			1
15	3:Low	4:Medium	+		0.847
16	3:Low	4:Medium	+		0.847
17	1:'Very High'	1:Very High			1
18	1:'Very High'	1:Very High			1
19	1:'Very High'	1:Very High			1
20	5:High	5:High			0.895
21	1:'Very High'	1:Very High			1
22	2:'Very Low'	2:Very Low			1
23	2:'Very Low'	2:Very Low			1
24	2:'Very Low'	2:Very Low			1
25	2:'Very Low'	2:Very Low			1

26 2:'Very Low' 2:Very Low 1

27 2:'Very Low' 2:Very Low 1

28 2:'Very Low' 2:Very Low 1

29 2:'Very Low' 2:Very Low 1

30 2:'Very Low' 2:Very Low 1

31 3:Low 3:Low 1

32 4:Medium 5:High + 0.797

33 2:'Very Low' 2:Very Low 1

34 3:Low 3:Low 1

35 3:Low 3:Low 1

36 4:Medium 1:Very High + 0.7

37 2:'Very Low' 2:Very Low 1

38 2:'Very Low' 2:Very Low 1

39 2:'Very Low' 2:Very Low 1

40 2:'Very Low' 2:Very Low 1

41 2:'Very Low' 2:Very Low 1

42 2:'Very Low' 2:Very Low 1

43 2:'Very Low' 2:Very Low 1

44 3:Low 3:Low 1

45 5:High 5:High 0.895

46 3:Low 4:Medium + 0.847

47 5:High 5:High 0.895

48 4:Medium 1:Very High + 0.7

49 3:Low 3:Low 1

50 5:High 5:High 0.895

51 4:Medium 1:Very High + 0.7

52 4:Medium 5:High + 0.797

53 5:High 5:High 0.895

54 5:High 5:High 0.895

55 2:'Very Low' 2:Very Low 1

56 5:High 5:High 0.895

57 4:Medium 1:Very High + 0.7

58 1:'Very High' 1:Very High 1

59 4:Medium 1:Very High + 0.7

60 4:Medium 1:Very High + 0.7

61 5:High 5:High 0.895

62 5:High 5:High 0.895

63 4:Medium 1:Very High + 0.7

64 1:'Very High' 1:Very High 1

65 1:'Very High' 1:Very High 1

66 4:Medium 1:Very High + 0.7

67 3:Low 3:Low 1

68 1:'Very High' 1:Very High 1

69 1:'Very High' 1:Very High 1

70 5:High 5:High 0.895

71 1:'Very High' 1:Very High 1

72 5:High 5:High 0.895

73 1:'Very High' 1:Very High 1

74 1:'Very High' 1:Very High 1

75 4:Medium 5:High + 0.797

76 1:'Very High' 1:Very High 1

77 4:Medium 5:High + 0.797

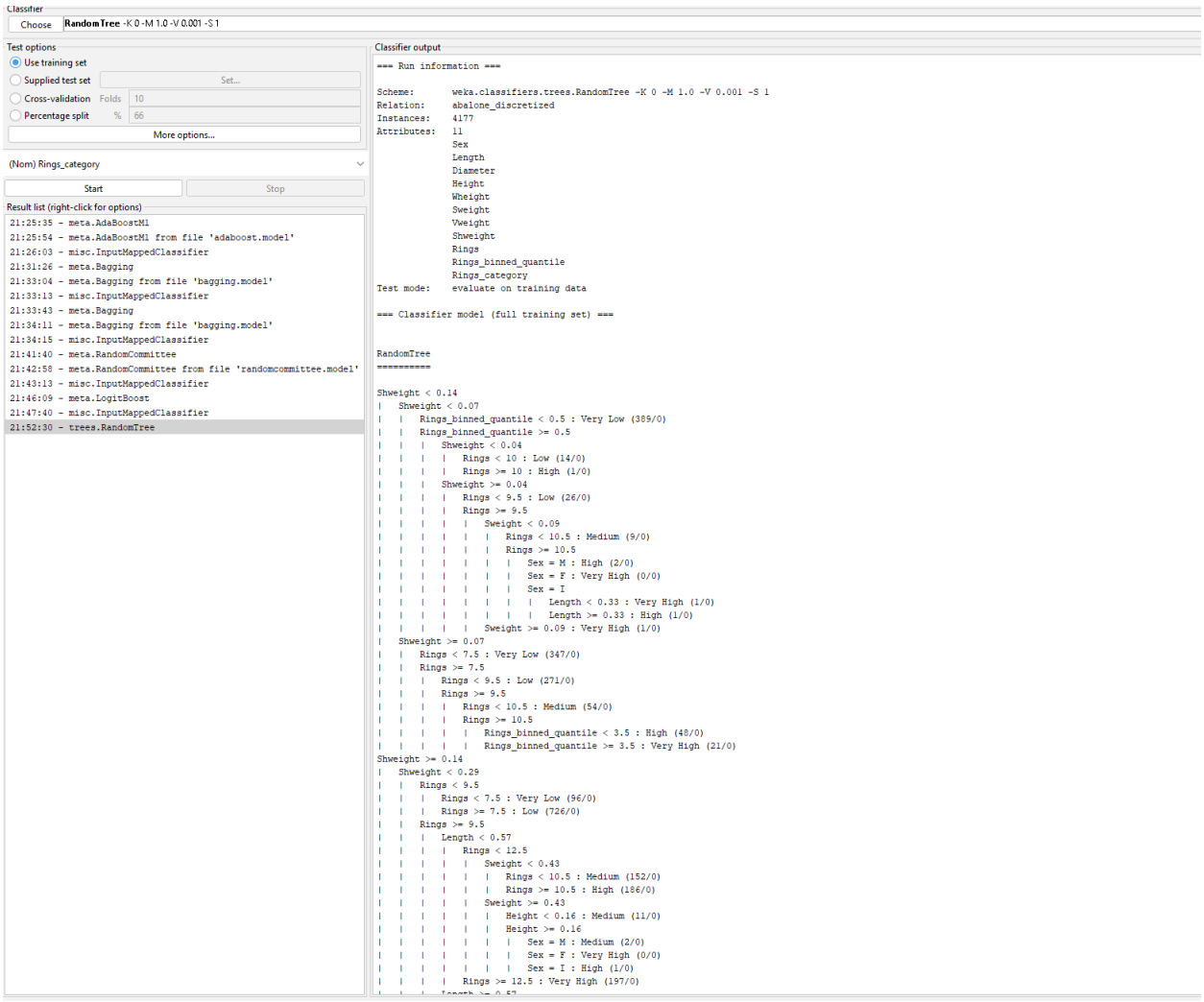
78 1:'Very High' 1:Very High 1

79 4:Medium 5:High + 0.797

80 4:Medium 1:Very High + 0.7

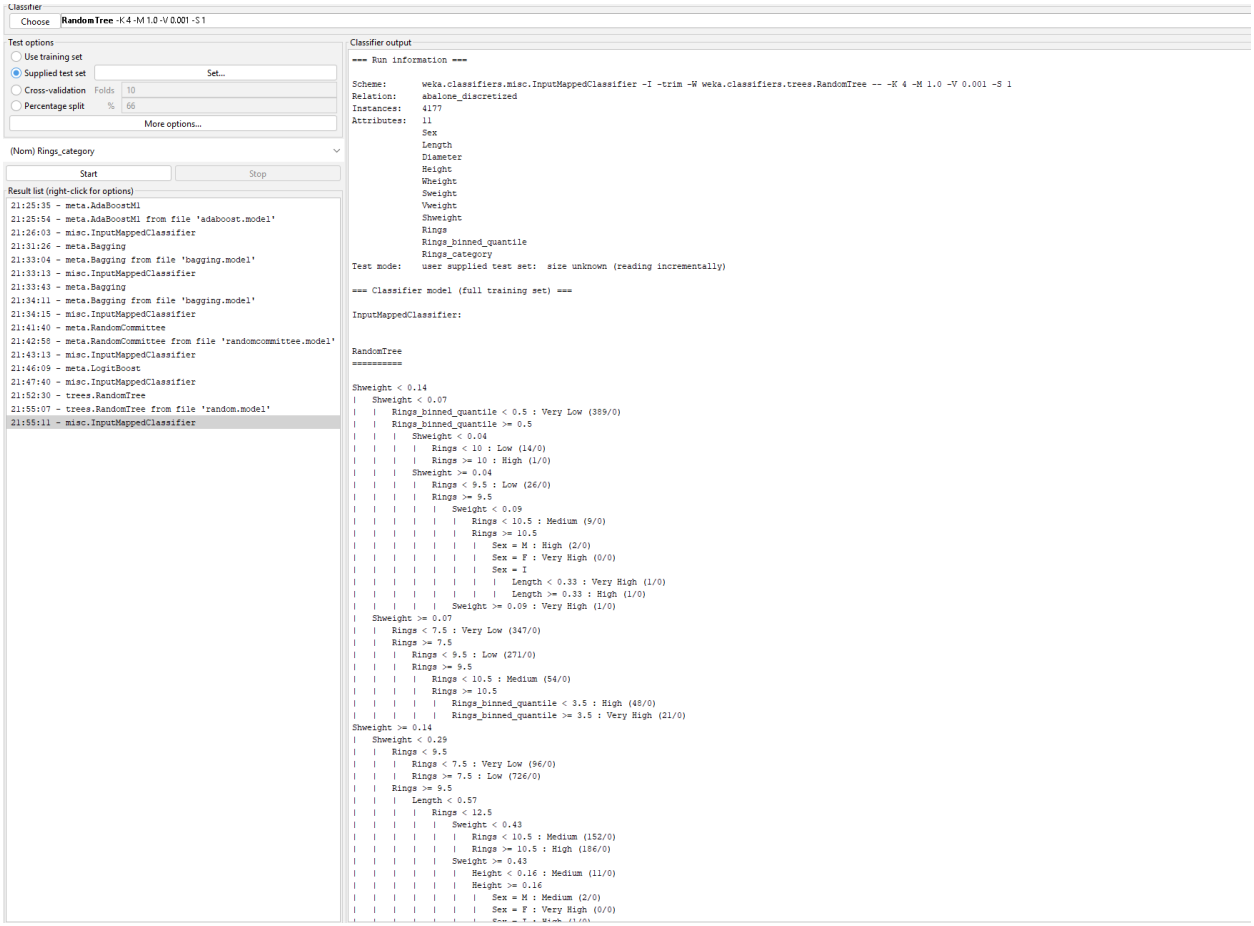
Correctly Classified Instances 58 72.5 %

Train Random Tree



```
=== Evaluation on training set ===

Time taken to test model on training data: 0.89 seconds

=== Summary ===

Correctly Classified Instances        4177               100      %
Incorrectly Classified Instances         0                 0      %
Kappa statistic                          1
Mean absolute error                      0
Root mean squared error                  0
Relative absolute error                  0      %
Root relative squared error              0      %
Total Number of Instances             4177

=== Detailed Accuracy By Class ===

                 TP Rate  FP Rate  Precision  Recall   F-Measure  MCC      ROC Area  PRC Area  Class
                 1.000    0.000    1.000      1.000    1.000      1.000    1.000     1.000     Very High
                 1.000    0.000    1.000      1.000    1.000      1.000    1.000     1.000     Very Low
                 1.000    0.000    1.000      1.000    1.000      1.000    1.000     1.000     Low
                 1.000    0.000    1.000      1.000    1.000      1.000    1.000     1.000     Medium
                 1.000    0.000    1.000      1.000    1.000      1.000    1.000     1.000     High
Weighted Avg.    1.000    0.000    1.000      1.000    1.000      1.000    1.000     1.000

=== Confusion Matrix ===

   a   b   c   d   e   <-- classified as
 693   0   0   0   0 |   a = Very High
   0 839   0   0   0 |   b = Very Low
   0   0 1257   0   0 |   c = Low
   0   0   0 634   0 |   d = Medium
   0   0   0   0 754 |   e = High
```

Test Random Tree



Prediction Results

=== Predictions on test set ===

inst#	actual	predicted	error	prediction
1	2:'Very Low'	2:Very Low		1
2	4:Medium	4:Medium		1
3	3:Low	3:Low		1
4	5:High	1:Very High	+	1
5	4:Medium	1:Very High	+	1
6	1:'Very High'	1:Very High		1
7	5:High	1:Very High	+	1
8	4:Medium	1:Very High	+	1
9	3:Low	3:Low		1
10	5:High	5:High		1
11	1:'Very High'	1:Very High		1
12	4:Medium	4:Medium		1
13	3:Low	4:Medium	+	1
14	2:'Very Low'	2:Very Low		1
15	3:Low	3:Low		1
16	3:Low	3:Low		1
17	1:'Very High'	1:Very High		1
18	1:'Very High'	1:Very High		1
19	1:'Very High'	1:Very High		1
20	5:High	5:High		1
21	1:'Very High'	1:Very High		1

22 2:'Very Low' 2:Very Low 1

23 2:'Very Low' 2:Very Low 1

24 2:'Very Low' 2:Very Low 1

25 2:'Very Low' 2:Very Low 1

26 2:'Very Low' 2:Very Low 1

27 2:'Very Low' 2:Very Low 1

28 2:'Very Low' 2:Very Low 1

29 2:'Very Low' 2:Very Low 1

30 2:'Very Low' 2:Very Low 1

31 3:Low 3:Low 1

32 4:Medium 5:High + 1

33 2:'Very Low' 2:Very Low 1

34 3:Low 3:Low 1

35 3:Low 3:Low 1

36 4:Medium 4:Medium 1

37 2:'Very Low' 2:Very Low 1

38 2:'Very Low' 2:Very Low 1

39 2:'Very Low' 2:Very Low 1

40 2:'Very Low' 2:Very Low 1

41 2:'Very Low' 2:Very Low 1

42 2:'Very Low' 2:Very Low 1

43 2:'Very Low' 2:Very Low 1

44 3:Low 3:Low 1

45 5:High 1:Very High + 1

46 3:Low 4:Medium + 1

47 5:High 1:Very High + 1

48 4:Medium 1:Very High + 1

49 3:Low 3:Low 1

50 5:High 5:High 1

51 4:Medium 1:Very High + 1

52 4:Medium 4:Medium 1

53 5:High 1:Very High + 1

54 5:High 5:High 1

55 2:'Very Low' 2:Very Low 1

56 5:High 5:High 1

57 4:Medium 1:Very High + 1

58 1:'Very High' 1:Very High 1

59 4:Medium 4:Medium 1

60 4:Medium 1:Very High + 1

61 5:High 5:High 1

62 5:High 1:Very High + 1

63 4:Medium 4:Medium 1

64 1:'Very High' 1:Very High 1

65 1:'Very High' 1:Very High 1

66 4:Medium 4:Medium 1

67 3:Low 3:Low 1

68 1:'Very High' 1:Very High 1

69 1:'Very High' 1:Very High 1

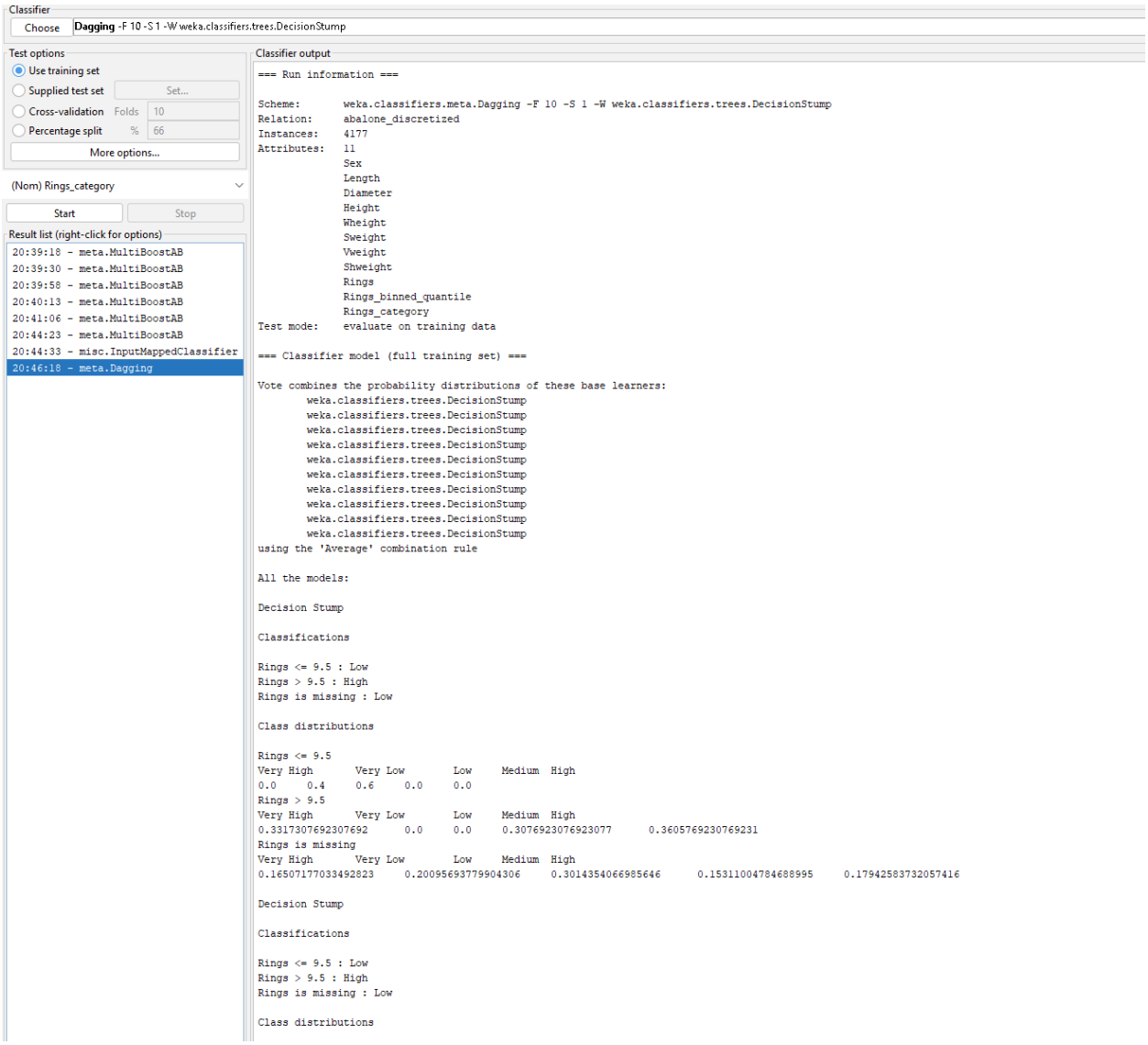
70 5:High 5:High 1

71 1:'Very High' 1:Very High 1

72	5:High	5:High	1
73	1:'Very High'	1:Very High	1
74	1:'Very High'	1:Very High	1
75	4:Medium	4:Medium	1
76	1:'Very High'	1:Very High	1
77	4:Medium	4:Medium	1
78	1:'Very High'	1:Very High	1
79	4:Medium	4:Medium	1
80	4:Medium	4:Medium	1

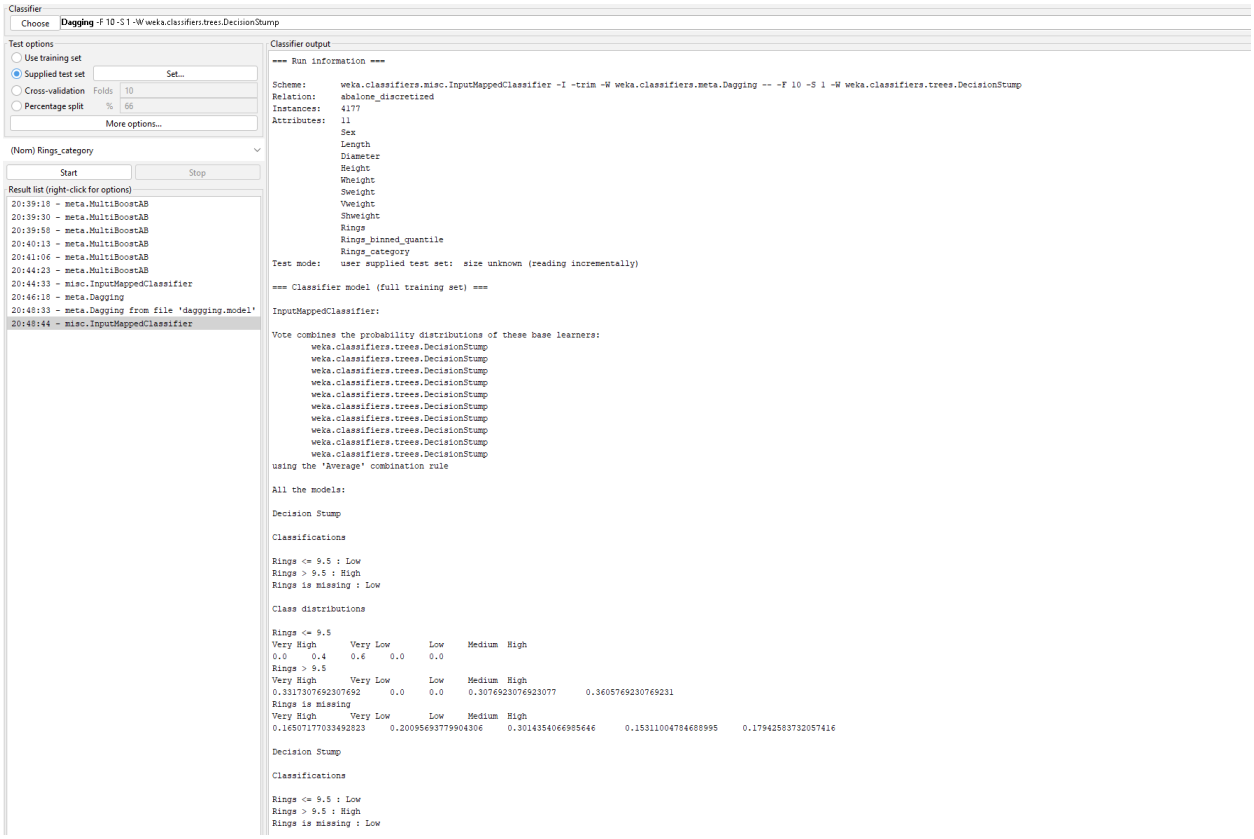
Correctly Classified Instances 65 81.25 %

Training Dagging


Classifier
Choose
Dagging -F 10 -S 1 -W weka.classifiers.trees.DecisionStump
Test options
Use training set
Supplied test set
Set...
Cross-validation
Folds
10
Percentage split
%
66
More options...
(Nom) Rings_category
Start
Stop
Result list (right-click for options)
20:39:18 - meta.MultiBoostAB
20:39:30 - meta.MultiBoostAB
20:39:58 - meta.MultiBoostAB
20:40:13 - meta.MultiBoostAB
20:41:06 - meta.MultiBoostAB
20:44:23 - meta.MultiBoostAB
20:44:33 - misc.InputMappedClassifier
20:46:18 - meta.Dagging
Classifier output
=== Run information ===
Scheme: weka.classifiers.meta.Dagging -F 10 -S 1 -W weka.classifiers.trees.DecisionStump
Relation: abalone_discretized
Instances: 4177
Attributes: 11
Sex
Length
Diameter
Height
Wheight
Sweight
Vweight
Shweight
Rings
Rings_binned_quantile
Rings_category
Test mode: evaluate on training data
=== Classifier model (full training set) ===
Vote combines the probability distributions of these base learners:
weka.classifiers.trees.DecisionStump
weka.classifiers.trees.DecisionStump
weka.classifiers.trees.DecisionStump
weka.classifiers.trees.DecisionStump
weka.classifiers.trees.DecisionStump
weka.classifiers.trees.DecisionStump
weka.classifiers.trees.DecisionStump
weka.classifiers.trees.DecisionStump
weka.classifiers.trees.DecisionStump
weka.classifiers.trees.DecisionStump
using the 'Average' combination rule
All the models:
Decision Stump
Classifications
Rings <= 9.5 : Low
Rings > 9.5 : High
Rings is missing : Low
Class distributions
Rings <= 9.5
Very High Very Low Low Medium High
0.0 0.4 0.6 0.0 0.0
Rings > 9.5
Very High Very Low Low Medium High
0.3317307692307692 0.0 0.0 0.3076923076923077 0.3605769230769231
Rings is missing
Very High Very Low Low Medium High
0.16507177033492823 0.20095693779904306 0.3014354066985646 0.15311004784688995 0.17942583732057416
Decision Stump
Classifications
Rings <= 9.5 : Low
Rings > 9.5 : High
Rings is missing : Low
Class distributions

```
Time taken to test model on training data: 1.25 seconds

=== Summary ===

Correctly Classified Instances        2011               48.1446 %
Incorrectly Classified Instances      2166               51.8554 %
Kappa statistic                          0.3168
Mean absolute error                      0.2289
Root mean squared error                  0.3383
Relative absolute error                 72.806  %
Root relative squared error             85.3273 %
Total Number of Instances             4177

=== Detailed Accuracy By Class ===

                 TP Rate  FP Rate  Precision  Recall   F-Measure  MCC      ROC Area  PRC Area  Class
                 0.000    0.000    ?          0.000    ?          ?        0.801     0.333     Very High
                 0.000    0.000    ?          0.000    ?          ?        0.812     0.400     Very Low
                 1.000    0.287    0.600      1.000    0.750      0.654    0.856     0.600     Low
                 0.000    0.000    ?          0.000    ?          ?        0.796     0.305     Medium
                 1.000    0.388    0.362      1.000    0.532      0.471    0.806     0.362     High
Weighted Avg.    0.481    0.156    ?          0.481    ?          ?        0.820     0.428

=== Confusion Matrix ===

    a    b    c    d    e   <-- classified as
    0    0    0    0  693 |    a = Very High
    0    0  839    0    0 |    b = Very Low
    0    0 1257    0    0 |    c = Low
    0    0    0    0  634 |    d = Medium
    0    0    0    0  754 |    e = High
```

Test Dagging



Prediction Results

=== Predictions on test set ===

inst#	actual	predicted	error	prediction
1	2:'Very Low'	3:Low	+	0.6
2	4:Medium	5:High	+	0.362
3	3:Low	3:Low		0.6
4	5:High	5:High		0.362
5	4:Medium	5:High	+	0.362
6	1:'Very High'	5:High	+	0.362
7	5:High	5:High		0.362
8	4:Medium	5:High	+	0.362
9	3:Low	3:Low		0.6
10	5:High	5:High		0.362
11	1:'Very High'	5:High	+	0.362
12	4:Medium	5:High	+	0.362
13	3:Low	5:High	+	0.362
14	2:'Very Low'	3:Low	+	0.6
15	3:Low	5:High	+	0.362
16	3:Low	5:High	+	0.362
17	1:'Very High'	5:High	+	0.362
18	1:'Very High'	5:High	+	0.362
19	1:'Very High'	5:High	+	0.362
20	5:High	5:High		0.362
21	1:'Very High'	5:High	+	0.362

22 2:'Very Low' 3:Low + 0.6

23 2:'Very Low' 3:Low + 0.6

24 2:'Very Low' 3:Low + 0.6

25 2:'Very Low' 3:Low + 0.6

26 2:'Very Low' 3:Low + 0.6

27 2:'Very Low' 3:Low + 0.6

28 2:'Very Low' 3:Low + 0.6

29 2:'Very Low' 3:Low + 0.6

30 2:'Very Low' 3:Low + 0.6

31 3:Low 3:Low 0.6

32 4:Medium 5:High + 0.362

33 2:'Very Low' 3:Low + 0.6

34 3:Low 3:Low 0.6

35 3:Low 3:Low 0.6

36 4:Medium 5:High + 0.362

37 2:'Very Low' 3:Low + 0.6

38 2:'Very Low' 3:Low + 0.6

39 2:'Very Low' 3:Low + 0.6

40 2:'Very Low' 3:Low + 0.6

41 2:'Very Low' 3:Low + 0.6

42 2:'Very Low' 3:Low + 0.6

43 2:'Very Low' 3:Low + 0.6

44 3:Low 3:Low 0.6

45 5:High 5:High 0.362

46 3:Low 5:High + 0.362

47 5:High 5:High 0.362

48 4:Medium 5:High + 0.362

49 3:Low 3:Low 0.6

50 5:High 5:High 0.362

51 4:Medium 5:High + 0.362

52 4:Medium 5:High + 0.362

53 5:High 5:High 0.362

54 5:High 5:High 0.362

55 2:'Very Low' 3:Low + 0.6

56 5:High 5:High 0.362

57 4:Medium 5:High + 0.362

58 1:'Very High' 5:High + 0.362

59 4:Medium 5:High + 0.362

60 4:Medium 5:High + 0.362

61 5:High 5:High 0.362

62 5:High 5:High 0.362

63 4:Medium 5:High + 0.362

64 1:'Very High' 5:High + 0.362

65 1:'Very High' 5:High + 0.362

66 4:Medium 5:High + 0.362

67 3:Low 3:Low 0.6

68 1:'Very High' 5:High + 0.362

69 1:'Very High' 5:High + 0.362

70 5:High 5:High 0.362

71 1:'Very High' 5:High + 0.362

72 5:High 5:High 0.362

73 1:'Very High' 5:High + 0.362

74 1:'Very High' 5:High + 0.362

75 4:Medium 5:High + 0.362

76 1:'Very High' 5:High + 0.362

77 4:Medium 5:High + 0.362

78 1:'Very High' 5:High + 0.362

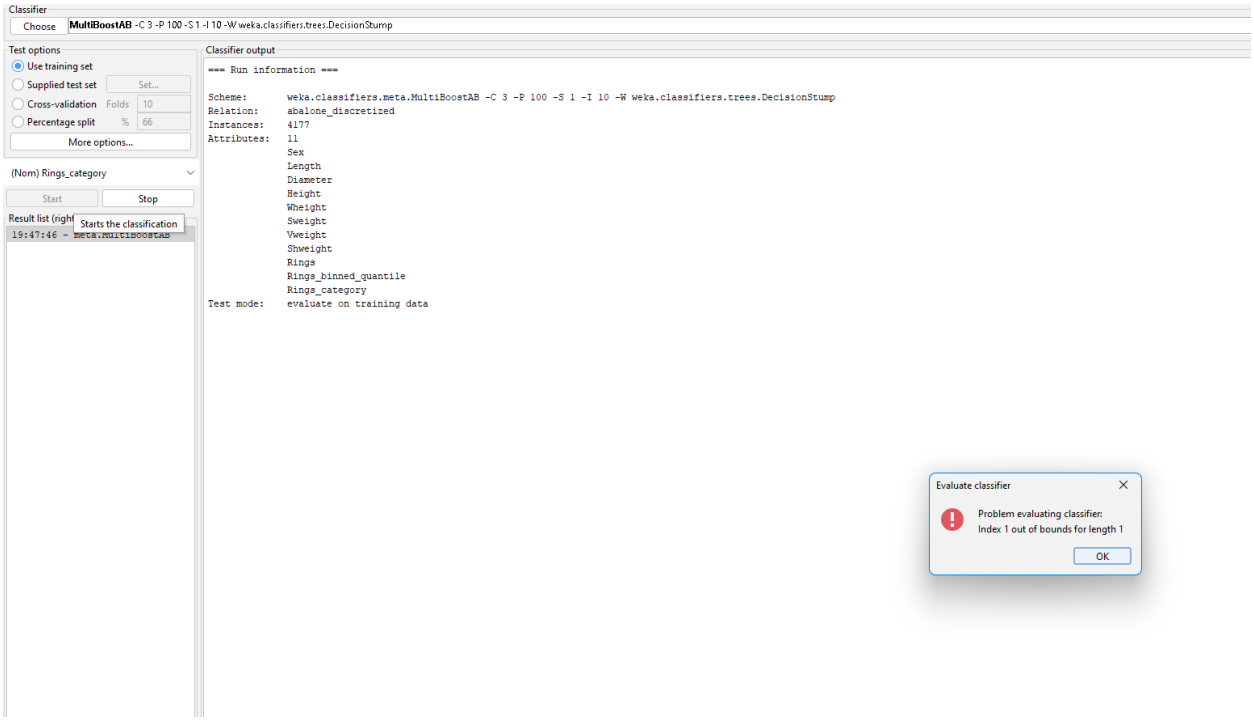
79 4:Medium 5:High + 0.362

80 4:Medium 5:High + 0.362

Correctly Classified Instances 22 27.5 %

Training MultiBoostAB

Weka tool showing an error for same training and test data.



Final Accuracy table

Classifier	Accuracy(%)	Correct Instance
RandomTree	81.25	65
RandomCommittee	80	64
Logitboost	72.5	58
Bagging	55	44
AdaBoostM1	27.5	22
Dagging	27.5	22