CM6 (Wheat Seeds Dataset)

In [41]:
```
print("Feature Means per class: \n", np.round(nb_1.theta_, 2))
```

```
Feature Means per class:
 [[14.34 14.34  0.88  5.53  3.24  2.73  5.08]
 [18.36 16.15  0.88  6.17  3.68  3.67  6.03]
 [11.88 13.26  0.85  5.24  2.85  4.66  5.12]]
```

In [42]:
```
print("Feature Variance per class: \n", np.round(nb_1.sigma_, 3))
```

```
Feature Variance per class:
 [[0.945 0.249 0.009 0.051 0.036 1.185 0.075]
 [2.089 0.366 0.009 0.062 0.046 1.424 0.054]
 [0.512 0.131 0.009 0.027 0.029 0.994 0.036]]
```

In [43]:
```
print("Class Priors: \n",  np.round(nb_1.class_prior_, 3))
```

```
Class Priors:
 [0.327 0.365 0.308]
```

Learned Parameters and Naive Bayes

The Feature Means and Variances best describe the classifier

Perfomance of Decision Tree vs Naive Bayes

Accuracy	Decision Tree	Naive Bayes
Training Accuracy	100	96.23
Test Accuracy	92.5	95
**Wall Time	364ms	4.59ms

Naive Bayes has a slightly higher Test Accuracy and is much faster than Decision Tree. Looking at the difference between the training accuracy and test accuracy of the Decision Tree we can also infer that the Decision Tree is more likely to overfit.

In [37]:
```
from sklearn.inspection import permutation_importance
r = permutation_importance(nb_1, X_train_1, y_train_1, n_repeats=30, random_state=0)

for i in r.importances_mean.argsort()[::-1]:
    if r.importances_mean[i] - 2 * r.importances_std[i] > 0:
        print(f"{dataset.columns[i]:<8}  "
              f" Mean: + {r.importances_mean[i]:.3f}  "
               f" STD: +/- {r.importances_std[i]:.3f}")
```

```
area        Mean: + 0.185   STD: +/- 0.025
perimeter    Mean: + 0.156   STD: +/- 0.023
length_of_kernel_groove    Mean: + 0.078   STD: +/- 0.019
length_kernel    Mean: + 0.077   STD: +/- 0.019
width_kernel    Mean: + 0.075   STD: +/- 0.016
asymmetry_coeff    Mean: + 0.040   STD: +/- 0.011
```

Relation to Decision Tree Splitting Rules

The most important features learned by the Naive Bayes Classifier; Area, length of Kernel groove, width Kernel and Asymmetry Coefficient were the features used in all the splitting rules of the Decision Tree.

CM6 (Covid Dataset)

In [55]:
```
print("Feature Means per class: \n", nb_2.theta_)
```

```
Feature Means per class:
 [[ 2.72566151e+00  4.37274733e+01 -7.95909554e+01  4.96505242e-01
   2.49625562e-04  4.97254119e-01  5.99101348e-03  4.40339491e-01
   1.59760359e-01  2.05941088e-01  1.75736395e-01  1.72241638e-02
   9.98502247e-04  7.90314528e-01  2.09685472e-01]
 [ 2.92826253e+00  4.37725211e+01 -7.96265344e+01  4.99872806e-01
   2.54388196e-04  4.90714831e-01  9.15797507e-03  3.14932587e-01
   3.56652251e-01  1.32790639e-01  1.85703383e-01  9.92113966e-03
   0.00000000e+00  7.83770033e-01  2.16229967e-01]
 [ 6.82482864e+00  4.37284837e+01 -7.94626338e+01  5.18151815e-01
   0.00000000e+00  4.74739782e-01  7.10840315e-03  7.64153338e-02
   7.53998477e-02  9.92637725e-02  7.36481340e-01  1.24397055e-02
   0.00000000e+00  2.37877634e-01  7.62122366e-01]]
```

In [56]:
```
print("Feature Variance per class: \n", nb_2.sigma_)
```

```
Feature Variance per class:
 [[4.82678204 1.19182136 3.06681906 0.93460757 0.68486935 0.93461225
  0.69057491 0.93106041 0.81885677 0.84814914 0.8294729  0.70154728
  0.68561729 0.85033726 0.85033726]
 [5.09463506 1.31843911 3.15977427 0.93461977 0.68487411 0.93453357
  0.69369389 0.90036984 0.91407121 0.79977707 0.83583742 0.6944425
  0.68461979 0.85409435 0.85409435]
 [2.00453138 1.22660619 3.25730642 0.9342903  0.68461979 0.93398171
  0.69167766 0.75519582 0.7543345  0.77403026 0.87869636 0.69690475
  0.68461979 0.86591165 0.86591165]]
```

In [57]:
```
print("Class Priors: \n", nb_2.class_prior_)
```

```
Class Priors:
 [0.33731896 0.3310037  0.33167733]
```

Learned Parameters and Naive Bayes

The Feature Means and Variances best describe the classifier

Perfomance of Decision Tree vs Naive Bayes

Accuracy	Decision Tree	Naive Bayes
Training Accuracy	66.14	65.68
Test Accuracy	65.91	65.34
**Wall Time	364ms	4.59ms

Naive Bayes seems to perform almost as well as Decision Tree and is much faster than Decision Tree in computation time.

In [58]:
```
from sklearn.inspection import permutation_importance
r = permutation_importance(nb_2, X_train_2, y_train_2, n_repeats=30, random_state=0)

for i in r.importances_mean.argsort()[::-1]:
    if r.importances_mean[i] - 2 * r.importances_std[i] > 0:
        print(f"{dataset2.columns[i]:<8}  "
              f" Mean: + {r.importances_mean[i]:.3f}  "
               f" STD: +/- {r.importances_std[i]:.3f}")
```

```
Age_Group    Mean: + 0.263   STD: +/- 0.004
Case_AcquisitionInfo_MISSING INFORMATION    Mean: + 0.046   STD: +/- 0.002
Case_AcquisitionInfo_CC    Mean: + 0.012   STD: +/- 0.001
Case_AcquisitionInfo_NO KNOWN EPI LINK    Mean: + 0.006   STD: +/- 0.001
```

Relation to Decision Tree Splitting Rules

The most important features learned by the Naive Bayes Classifier; Age Group and Case_AcquisitionInfo were also used in the splitting rules of the Decision Tree.