

Assignment 8

1. Use an SVM-based model to predict the Portuguese outcome of the battle from the number of ships involved on all sides and Spanish involvement.
2. Try solving the same problem using two other classifiers that you know.
3. Report and compare their results with those from SVM.

SVM based approach produced a fairly inaccurate result. The approach generated a **33.3% prediction accuracy** when using the factors (number of ships and Spanish involvement) to predict Portuguese outcome. One notable piece of information is that the sample size is relatively small, with only 28 battles to learn from. If the train/test split were larger, results would likely have a prediction accuracy score that is a bit better than 33%.

```
In [186]: print(classification_report(y_test, predictions))
```

	precision	recall	f1-score	support
-1	0.50	0.50	0.50	2
0	0.33	0.50	0.40	4
1	0.00	0.00	0.00	3
accuracy			0.33	9
macro avg	0.28	0.33	0.30	9
weighted avg	0.26	0.33	0.29	9

Random forest produced about the same results on average. The initial random forest classification was a bit lower, at a **22.2% prediction accuracy**. The same issue with a small sample size likely played a large factor in such a low score.

```
In [200]: print(classification_report(y_test, tree_predictions))
```

	precision	recall	f1-score	support
-1	0.00	0.00	0.00	2
0	0.40	0.50	0.44	4
1	0.00	0.00	0.00	3
accuracy			0.22	9
macro avg	0.13	0.17	0.15	9
weighted avg	0.18	0.22	0.20	9

KNN produced pretty similarly bad accuracy prediction scores. The initial result was **33.3% prediction accuracy**. Rerunning the method multiple times produced a range of between **11%-33%**. With this assignment, it reaffirms the notion that an appropriate sample size is necessary in order to run classification algorithms.

```
In [207]: print(classification_report(y_test,knn_predictions))
```

	precision	recall	f1-score	support
--	-----------	--------	----------	---------

-1	0.20	0.50	0.29	2
----	------	------	------	---

0	0.50	0.50	0.50	4
---	------	------	------	---

1	0.00	0.00	0.00	3
---	------	------	------	---

accuracy			0.33	9
----------	--	--	------	---

macro avg	0.23	0.33	0.26	9
-----------	------	------	------	---

weighted avg	0.27	0.33	0.29	9
--------------	------	------	------	---