

Instructions: More information on expectations for the report can be found in the general directions for the data analysis sets. In this document, the specifics for the individual assignment will be discussed. Students are responsible for both the requirements in the general directions, and for the specific directions discussed below.

Topic 1: Binary Classification

Install the package {mlbench}, and download the Ionosphere dataset.

	V1	V2	V3	V4	V5	V6	V7	V8	V9	V10	V11	V12	V13	V14	V15
1	1	0	0.99539	-0.05889	0.85243	0.02306	0.83398	-0.37708	1.00000	0.03760	0.85243	-0.17755	0.59755	-0.44945	0.60536
2	1	0	1.00000	-0.18829	0.93035	-0.36156	-0.10868	-0.93597	1.00000	-0.04549	0.50874	-0.67743	0.34432	-0.69707	-0.51685
3	1	0	1.00000	-0.03365	1.00000	0.00485	1.00000	-0.12062	0.88965	0.01198	0.73082	0.05346	0.85443	0.00827	0.54591
V16	V17	V18	V19	V20	V21	V22	V23	V24	V25	V26	V27	V28	V29	V30	
-0.38223	0.84356	-0.38542	0.58212	-0.32192	0.56971	-0.29674	0.36946	-0.47357	0.56611	-0.51171	0.41078	-0.46168	0.21266	-0.34090	
-0.97515	0.05499	-0.62237	0.33109	-1.00000	-0.13151	-0.45300	-0.18056	-0.35734	-0.20332	-0.26569	-0.20468	-0.18401	-0.19040	-0.11593	
0.00299	0.83775	-0.13644	0.75535	-0.08540	0.70887	-0.27502	0.43385	-0.12062	0.57528	-0.40220	0.58984	-0.22145	0.43100	-0.17365	
V31	V32	V33	V34	Class											
0.42267	-0.54487	0.18641	-0.45300	good											
-0.16626	-0.06288	-0.13738	-0.02447	bad											
0.60436	-0.24180	0.56045	-0.38238	good											

```
data(Ionosphere)
view(Ionosphere)
```

Compare two binary classifiers (such as logistic regression and SVM). Use a confusion matrix to compare the results. Describe the process, and why you chose these particular classifiers. Also discuss any hyperparameters that had to be set, any required rescaling, etc. You may use a classification method that is not strictly binary, however, if you use it here, you must select a different one for Topic 2. Create appropriate graphs to visualize the results.

Topic 2: Non-binary classification

Also from the {mlbench} package, download the Glass dataset. More information can be found here: <https://machinelearningmastery.com/machine-learning-datasets-in-r/>

	RI	Na	Mg	Al	Si	K	Ca	Ba	Fe	Type
1	1.52101	13.64	4.49	1.10	71.78	0.06	8.75	0.00	0.00	1
2	1.51761	13.89	3.60	1.36	72.73	0.48	7.83	0.00	0.00	1
3	1.51618	13.53	3.55	1.54	72.99	0.39	7.78	0.00	0.00	1

```
data(Glass)
view(Glass)
```

Select two classification methods that can handle more than two classes (for example, KNN and Random Forest). Set up both models and compare the results with a confusion matrix. Create appropriate graphs. Describe any hyperparameters that had to be set and anything you did to improve the results of your model.

Consider one additional model, in addition to the two you used. You don't have to run this third model, but consider the potential advantages or disadvantages of this alternative approach over the two you used.

You may include your code in an appendix for separate file, but the report of approximately 10 pages should focus on the analysis. It should look professionally formatted. Raw code and raw output is frowned upon.