

Neural Net Classification of Biopsy Tissue

Cynthia O'Rourke





Goal: Use ML to Classify Tissue Biopsies as Malignant or Benign

Labeled data set: 699 unique biopsy samples collected from 1989 through 1991 in 8 batches by one oncology group



Variables

9 descriptive features, all ordinally scaled from 1-10

Split 60/20/20 train/validate/test



Considerations

Is there a batch effect?

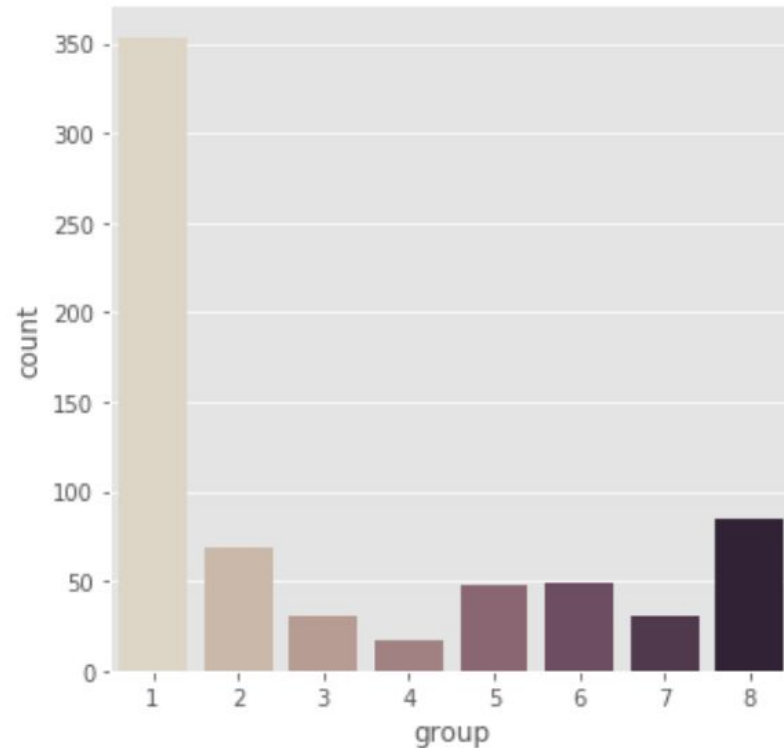
Are the features all predictive?

Do the features covary?



Considerations

Is there a batch effect?



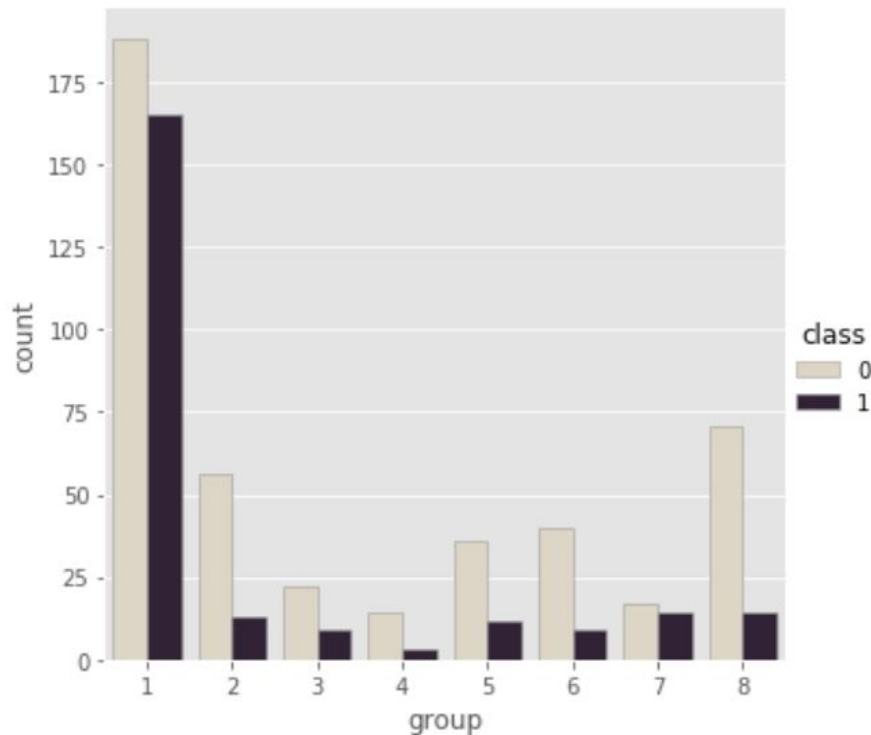


Considerations

Is there a batch effect?

Seems likely;

retain “batch”





Considerations

Is there a batch effect?

Are the features all predictive?



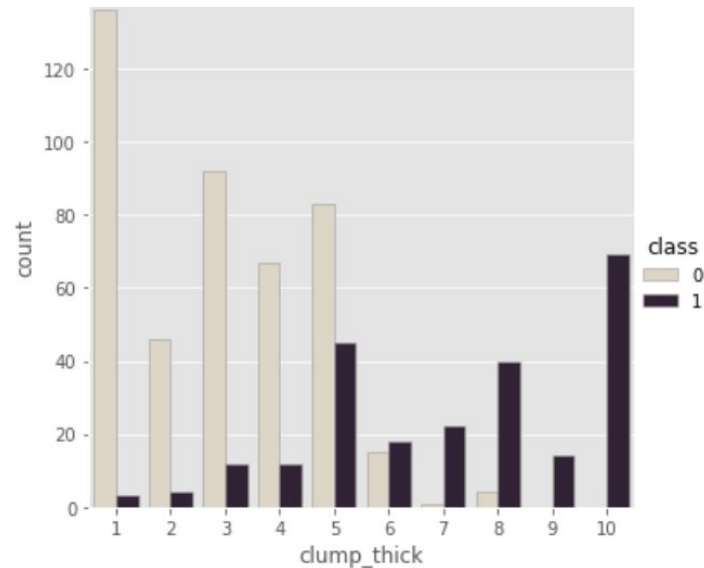
Considerations

Is there a batch effect?

Are the features all predictive?

Yes - all 9 features appear predictive

“Clump Thickness” example shown





Considerations

Is there a batch effect?

Are the features all predictive?

Do the features covary?



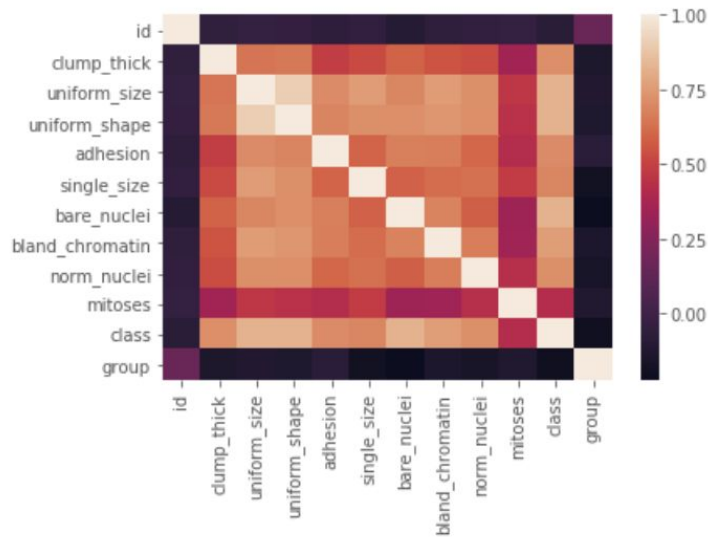
Considerations

Is there a batch effect?

Are the features all predictive?

Do the features covary?

Yes, some strongly





Algorithm Choice

A multi-layer perceptron **classifier neural net** will **tolerate multicollinearity** and a **non-linear relationship** between the features and the response variable



Metrics of Success

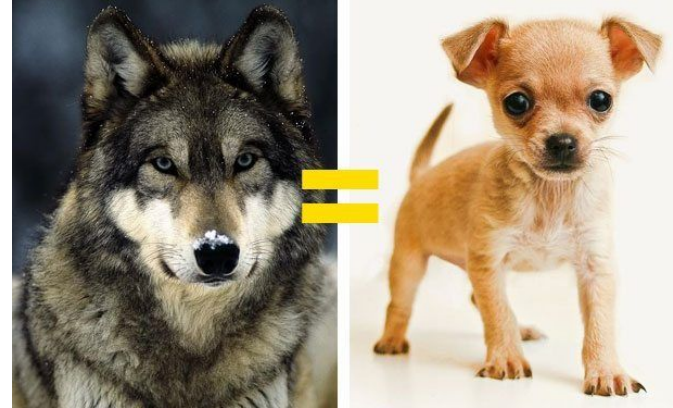


Accuracy - prop TP+TN out of all results

How often did the model get it Right?



Metrics of Success



Precision - prop TP out of all TP+FP

How often did the model cry wolf?



Metrics of Success



Recall - prop TP out of TP+FN

How often did the model miss a malignancy?



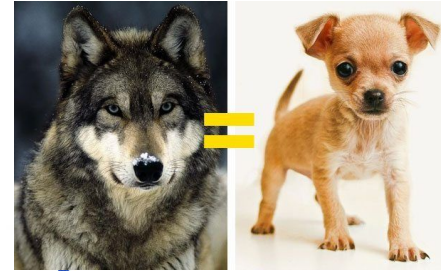
Algorithm Tuning

Increasing the number of neurons in the hidden layers had limited effect

Increasing the number of hidden layers from 5 to 7 had a *huge effect*

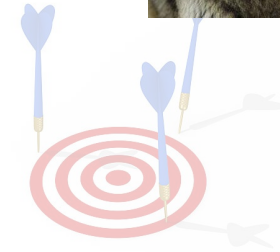
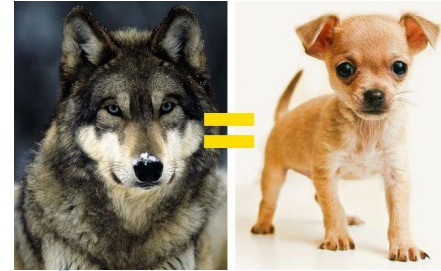
Algorithm Tuning: Train vs. Validation

	5 layers	7 layers
Accuracy	94.2%	100%
Precision	95.2%	100%
Recall	87.0%	100%



Algorithm *Testing*: Test vs. Validation

	Validation	Test
Accuracy	100%	97%
Precision	100%	100%
Recall	100%	91.7%



Conclusion

Trade-Off: False alarms will be rare under this model, but some cancers will go undiagnosed at early stages

