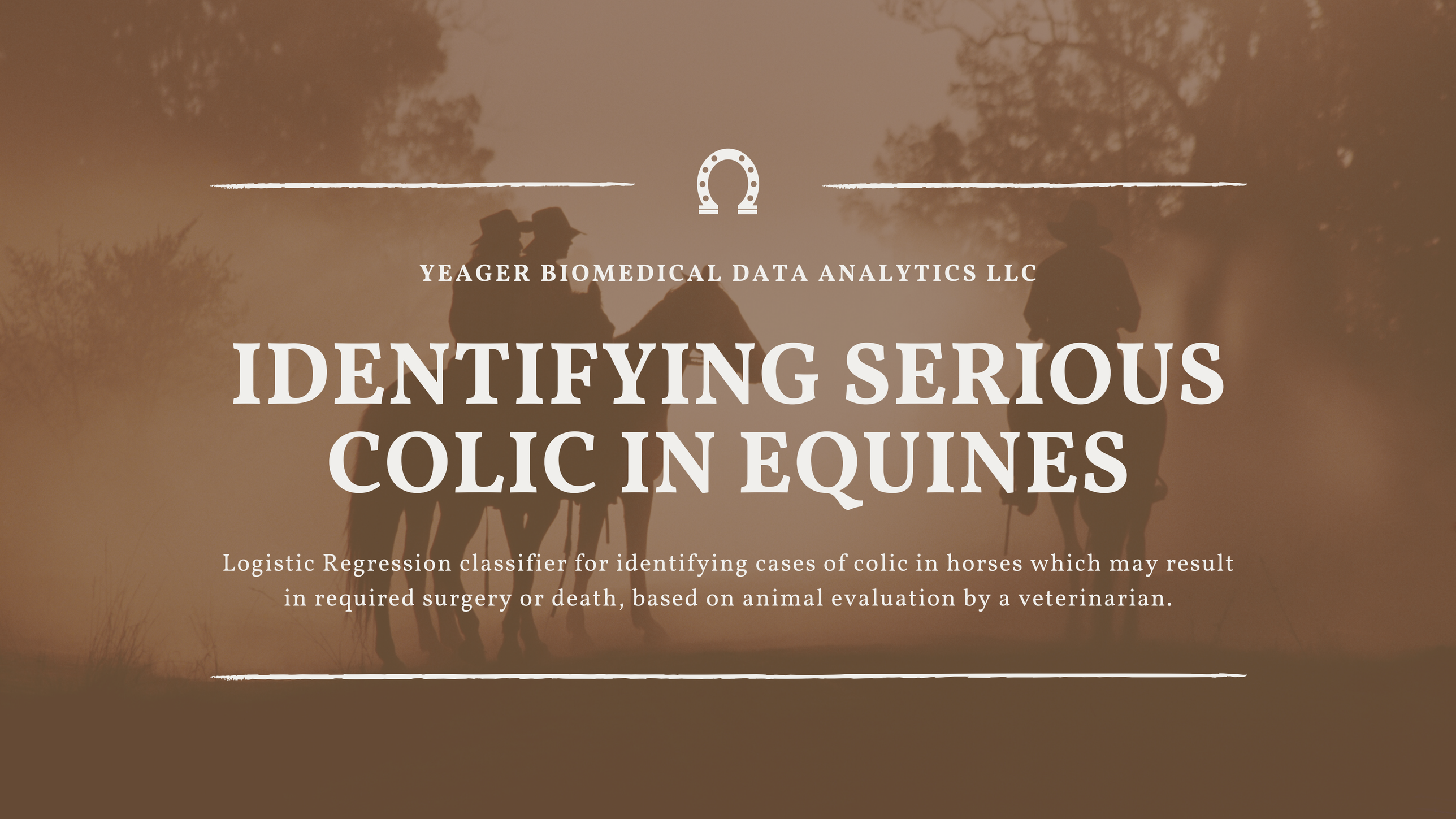


YEAGER BIOMEDICAL DATA ANALYTICS LLC

IDENTIFYING SERIOUS COLIC IN EQUINES

Logistic Regression classifier for identifying cases of colic in horses which may result in required surgery or death, based on animal evaluation by a veterinarian.

Why is this important?

Colic is a broad term used to describe gastrointestinal distress in horses. It often has a sudden onset and can be caused by a myriad of reasons. It is not always easily diagnosed and the seriousness of the problem can often escalate quickly and lead to necessary surgical intervention or death of the animal. The American Horse Council advises that colic causes the death of approximately 64,000 horses in the United States annually.

Overview and Goals

The Data

Modeling, Feature Selection, and Validation

Model 1: Death Outcome

Model 2: Surgical Outcome

Conclusions and Next Steps

OVERVIEW AND GOALS

Yeager Biomedical Data Analytics has created a Logistic Regression Classifier model to assist equine facility staff, as well as veterinarians in identifying cases in which horses are at high risk of death, as well as cases in which surgical intervention may be an optimal solution. The model takes into account medical evaluation data from the subject and produces an output whereby:

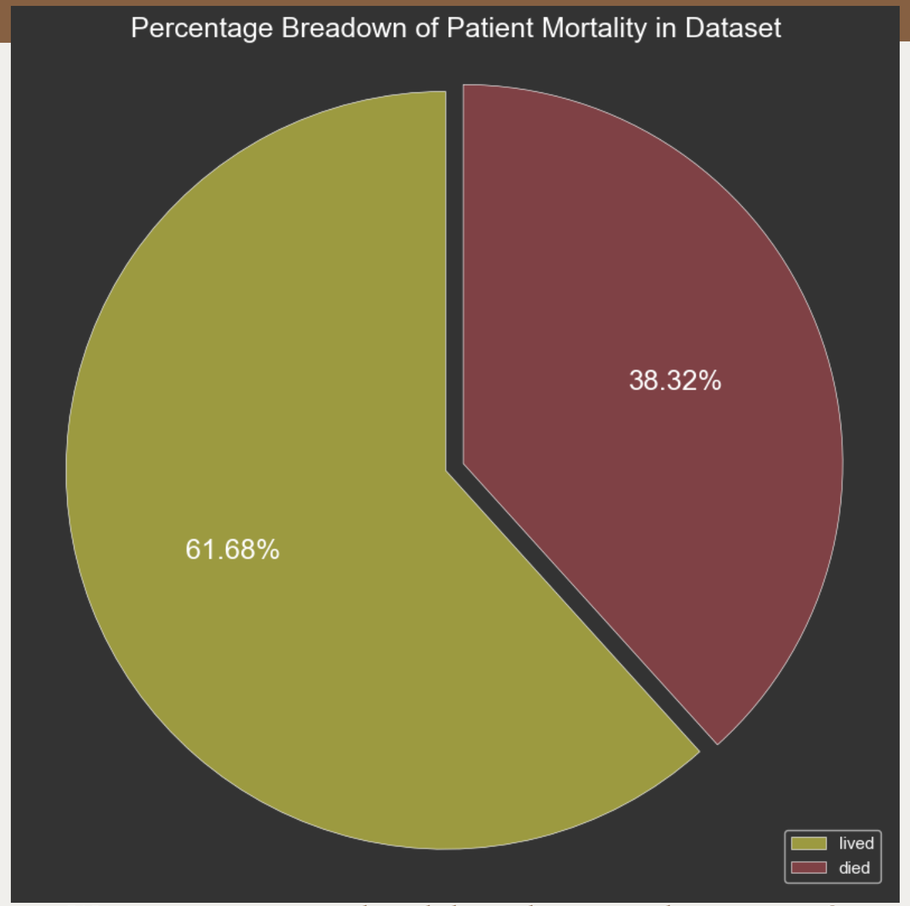
1. - Death is the likely outcome.
2. - Surgical intervention is the required solution to fix the problem.

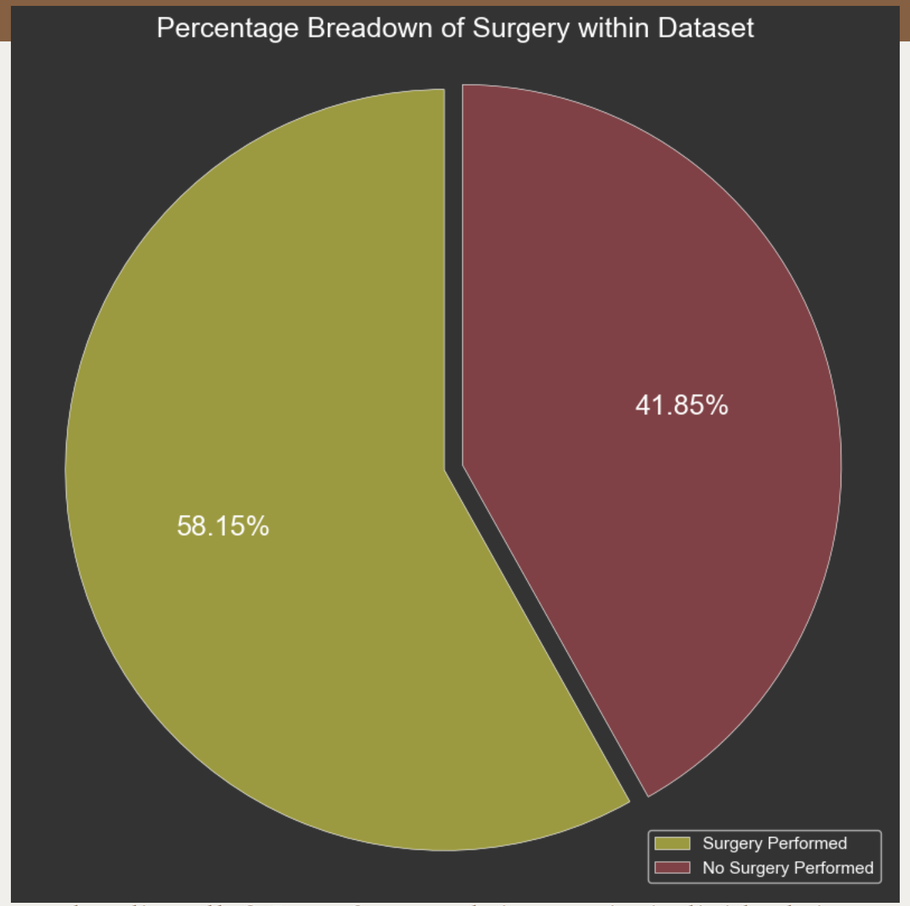
Using the outcome of these models, it is our hope that horse care staff will be able to validate their suspicions and more accurately, aggressively, and quickly treat the conditions of horses experiencing colic.

EXPLORING THE DATA

- The dataset used in this analysis is located on the UCI machine learning repository https://archive.ics.uci.edu/ml/datasets/Horse+Colic, and consists of 368 records of horses which were experiencing colic and evaluated by veterinarians. While small, a total of 27 features are included with each record to assist in bolstering the strength of the dataset for model creation. These include various medical evaluations from pulse to respiratory rate.

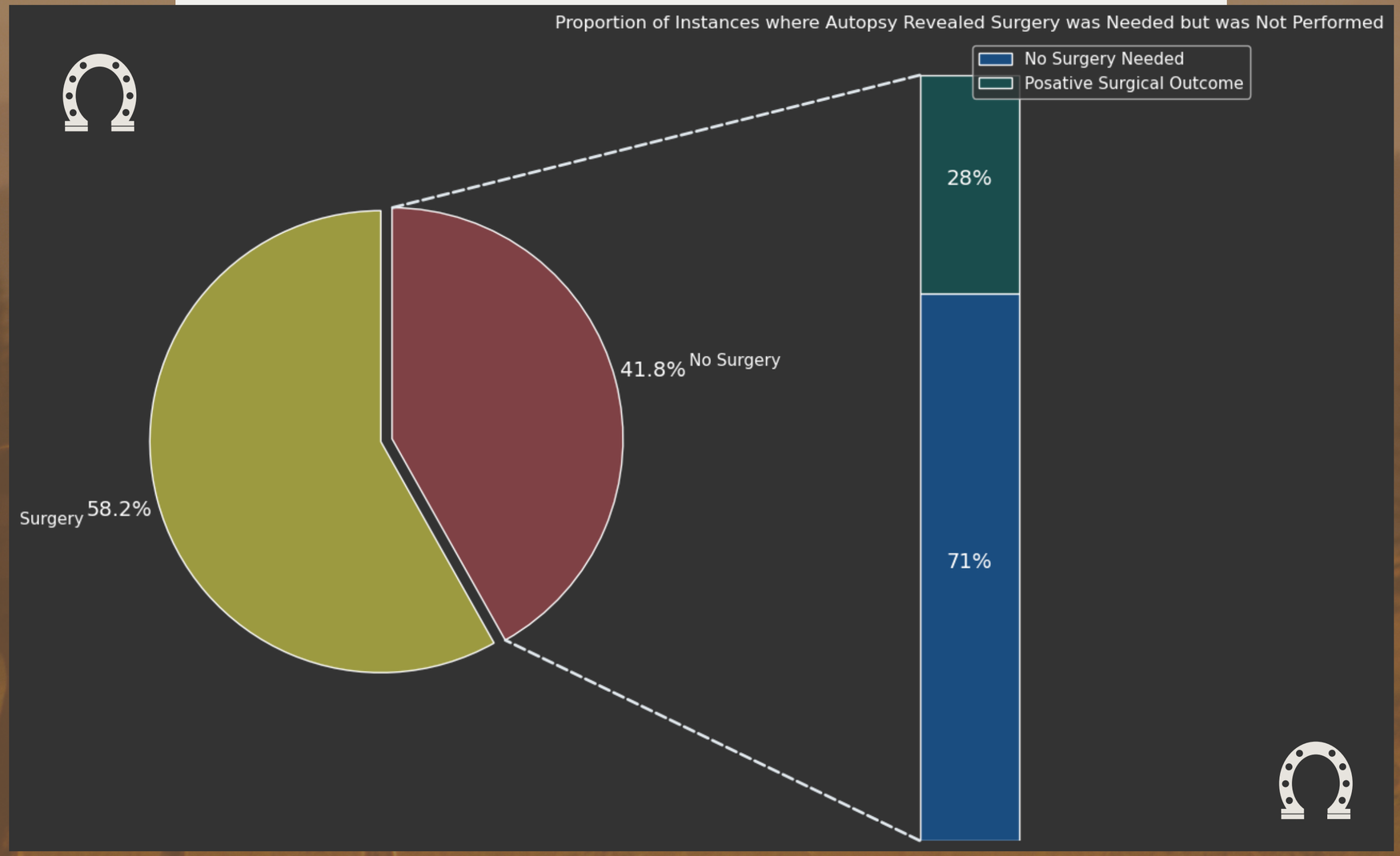
- Our target variables in the creation of these models will be:
  - Outcome - whether the horse lived or died.
  - Surgical Lesion - A retrospective finding, made either as a result of surgery or via autopsy, which states whether or not the horse's life could have been saved by surgery.

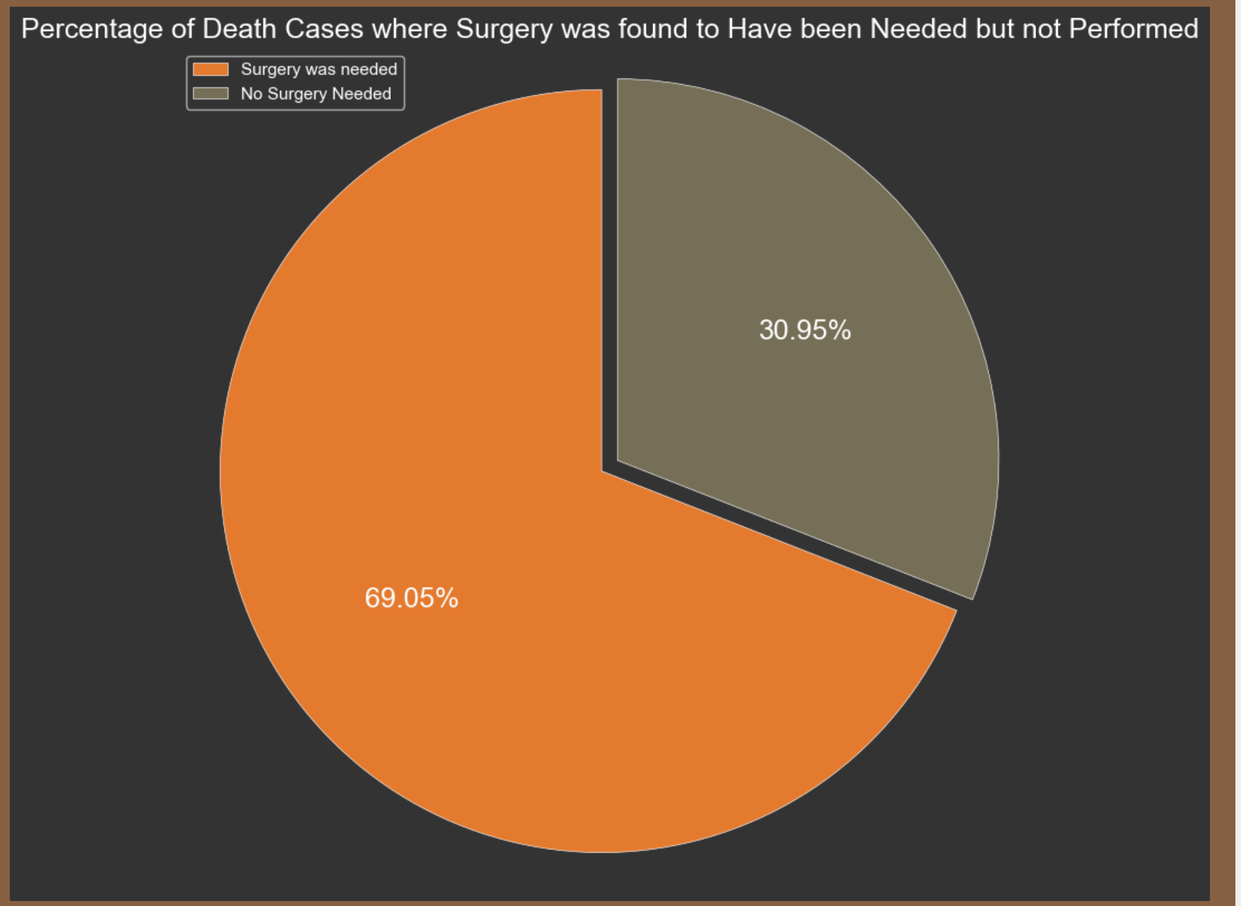




We see a pretty even breakdown between the counts of surgery and death, but does this translate literally? One of our goals is to assist individuals in deciding to conduct surgery as an intervention. Lets see how that breaks down further to understand if there is room for improvement.

Here we see that out of all cases in which surgery was not performed, that it was later revealed (usually via autopsy) that a surgical intervention would have produced a positive outcome for the animal in 28% of cases. This 28% comprises the cases which we will be attempting to improve upon.





Looking into this even further, we filter our dataset to see specifically those cases in which the patient dies after surgery is NOT performed, but it is later found out that surgery was, in fact, needed. The accompanying chart tells us that in 69% of these cases, surgery could have potentially saved the animal. Additional room for our model to assist these medical personnel in decision making.

Model 1 Predicting Death Outcome

Test Accuracy

.79 Achieved

10-fold cross validation score

.74 Achieved

Features Used

- Pulse
- Respiratory rate
- Pain level
- Packed cell vol
- Type of lesion

Holdout Accuracy

.84 Achieved

Model 2 Predicting Required Surgery

Test Accuracy

1 Achieved

10-fold cross validation score

1 Achieved

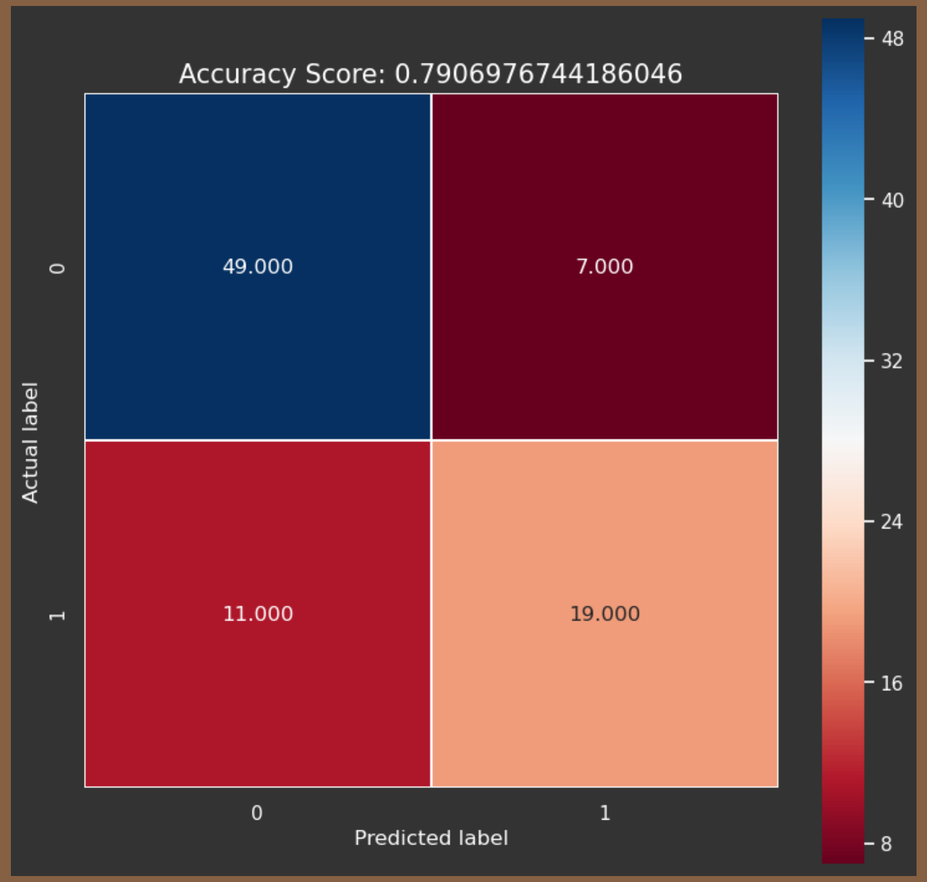
Features Used

- surgery
- pulse
- temperature of extremities
- peripheral pulse
- mucous membranes
- capillary refill time
- pain level
- peristalsis
- abdominal distension
- packed cell volume
- type of lesion
- location of lesion

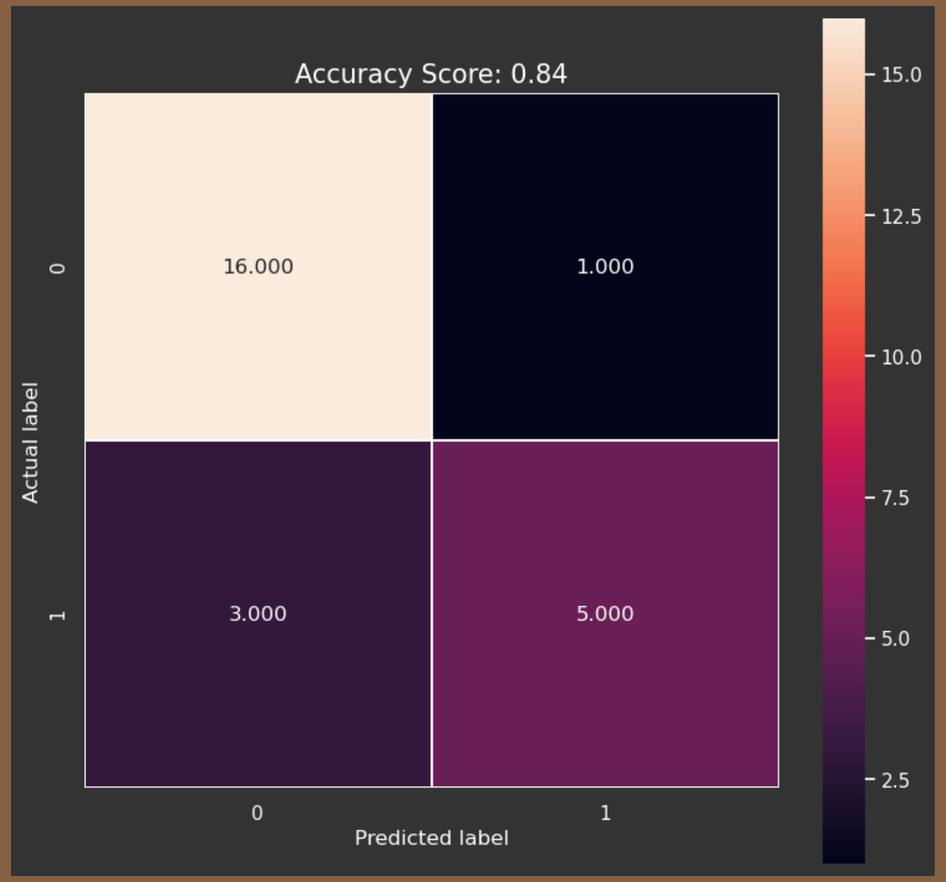
Holdout Accuracy

.52 Achieved

Model 1 Predicting Death Outcome Contd.

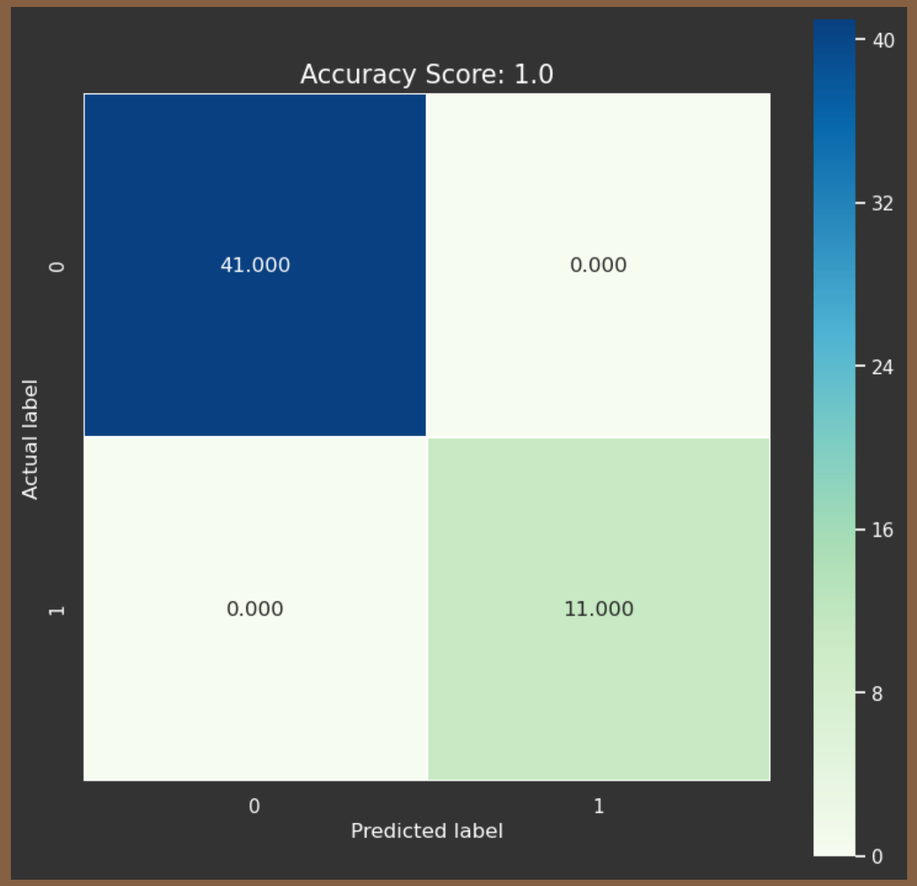


Test Data Confusion Matrix

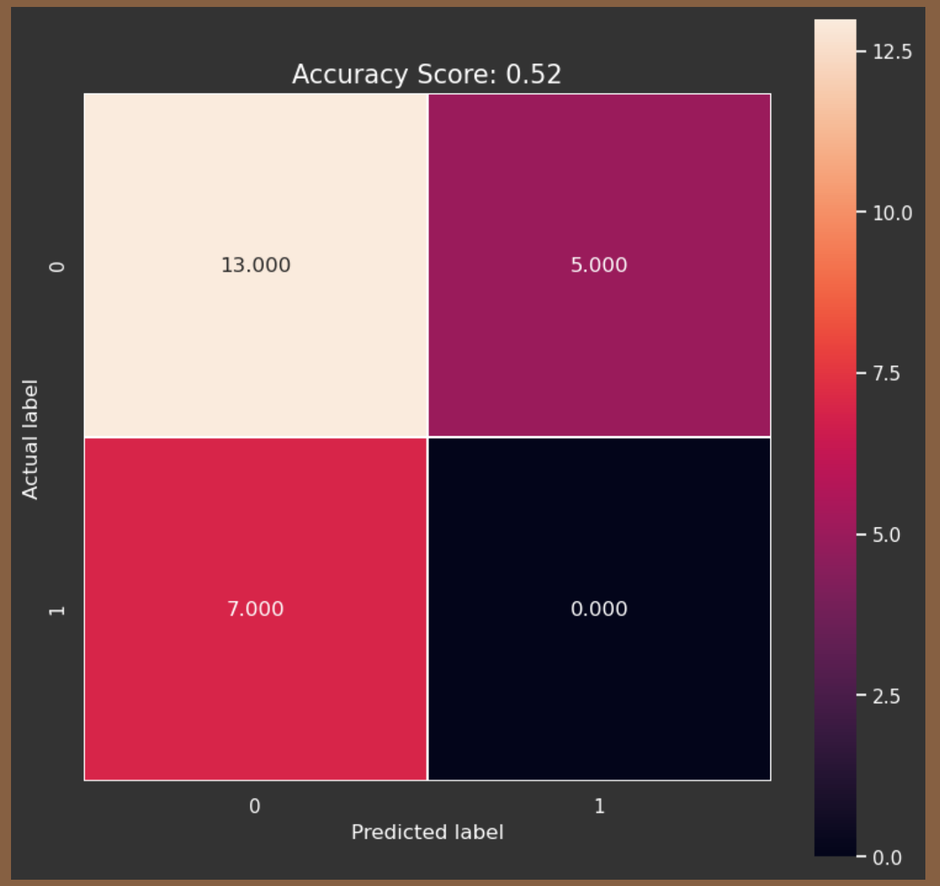


Holdout Data Confusion Matrix

Model 2Predicting Required Surgery Contd.



Test Data Confusion Matrix



Holdout Data Confusion Matrix

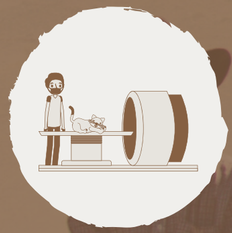
NEXT STEPS

INCREASED SIZE OF DATA SAMPLES

Provider Participation.

MORE INSIGHT

Using feature importance to incorporate new information into the model.



MORE MODERN TEQNIUQES

Incorporate things such as ultrasound image classification.

CONCLUSIONS

The goal in creating our models was to give veterinarians and equine care staff an additional tool in gauging the severity of suspected colic in horses under their care. A reasonable level of accuracy has been achieved, given our limited data set, in attaining that goal. A rapid evaluation of the sick animal can be entered into our models to determine the severity of the illness and validate the decision to aggressively treat, and hopefully save the patient. The tough decision to commit to surgery and reduce the percentage of horses which perish from a lack of this intervention is our hope. With adoption of this model and an increase in observations, its strength in making accurate predictions can only increase.