Modeling Section

12/5/2022

Modeling

Replicate Research Findings

First, we implemented the clinical decision rule found in Leonard et al. This study determined a clinical decision rule by selecting features using forward selection with logistic regression. For each iteration, the algorithm would add the feature whose p-value from the Chi-squared statistic was smallest when that feature is added to the model, versus when any other feature was added. The algorithm stopped adding features when no additional feature resulted in a p-value less than 0.05. 1000 samples were bootstrapped and a new logistic regression model was computed each time. A covariate was included in the final model if it appeared in over 50% of the bootstrapped models. This process was repeated for each control group: random, EMS, and mechanism of injury controls. The forward selection logistic regression models identified 6 common covariates between these 3 models: altered mental status, focal neurological deficit, complaint of neck pain, substantial injury to the torso, high-risk motor vehicle crash, and diving. The decision rule then classified a patient as likely to have cervical spine injury if any one of these factors was true, otherwise the patient was ruled to not have a cervical spine injury.

Although the report did not specify how missing values are handled, we removed samples where any of the covariates were missing, since our protocol in R was to remove NA values in logistic regression. We calculated this decision rule to have the following metrics, where patients from all 3 controll groups were included:

Metric	Estimate	Lower CI Bound	Upper CI Bound
Sensitivity	0.9064588	0.8795248	0.9333928
Specificity	0.4053537	0.3843153	0.4263922

Although the value for sensitivity is 2% lower and the value for specificity is 5% higher than presented in the Leonard et al. paper, we were not able to exactly replicate results due to assumptions we had to make about handling missing values, and which control groups to include in the final calculations. However, this replication gave us a baseline to compare our own model results to.

Modeling Approach

In our own modeling efforts we tried several classification methods coupled with two main approaches for features selection. Here we present two of the modeling approaches that we found amenable to interpretability and stability analysis, namely a single decision tree and linear logistic regression. For all of our modeling experiments we employed k-fold Cross Validation, with the folds determined by site (leaving one site out for each fold). We first outline our feature selection approaches, then present out classification results.

Feature Selection

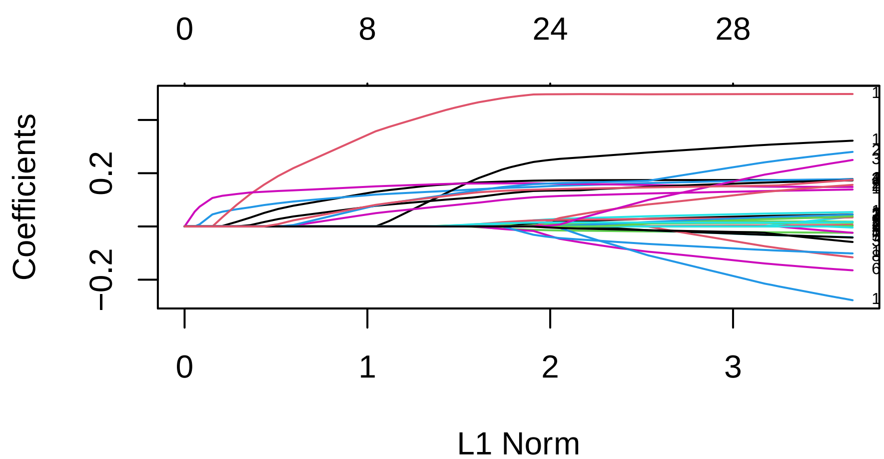
Bootstrapped Forward Selection

First, we selected features using forward selection with logistic regression, similar to the method used by Leonard et al. We started with an empty model, and added features sequentially, including the feature with the smallest p-value each. However, we stopped when there was no feature with a p-value less than 0.15, instead of 0.05 used in the feature selection method by Leonard et al. We then proceeded with the same bootstrapping procedure, selecting features that appeared in over 50% of the bootstrapped models.

Lasso Logistic Regression (L1 regularization)

Next, we selected features from a Lasoo logistic regression model. First, we completed 10 fold cross validation to find the value of *lambda* that minimizes the $L1$ loss for the training data. We then selected the features from this model that had non-zero regression coefficients.

In the following graph, you can see the order coefficients are added to the Lasso model, which provide insights into which features are most important to the final probability.



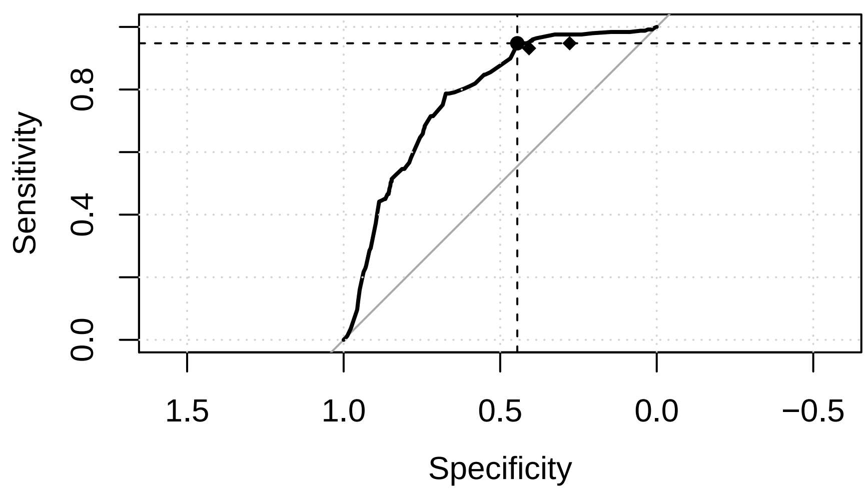
OrderVariable
1FocalNeuroFindings2
2FocalNeuroFindings
3AlteredMentalStatus
4HighriskDiving
5PainNeck2
6subinj_TorsoTrunk2
7HighriskMVC
8Torticollis2

OrderVariable
9Predisposed
10SubInj_Head
11subinj_Head2
12AxialLoadAnyDoc
13ambulatory
14PosMidNeckTenderness
15HighriskHitByCar
16HighriskHanging

Classification Experiments

The first classification approach we will present is the single decision tree approach. For that we used the rpart R package [citation], and compared the cross validation results with the results from the Leonard et al paper.

Single Decision Tree Results



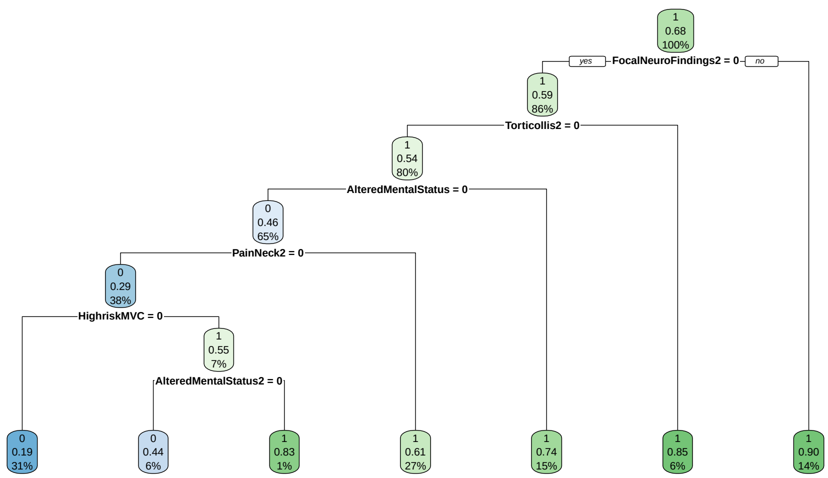
As we can see, classification tree prodcues an ROC curve with AUC = 0.78. The curve is strictly above the decision rule replicated from the published decision rule. Further selection of an appropriate decision threshold produces the following sensitivity/specificity metrics:

```
##               Estimage Lower CI Bound Upper CI Bound
## Sensitivity 0.9477912      0.9201614      0.9754209
## Specificity 0.4453507      0.4175303      0.4731711
```

The final list of predictors used for the decision tree is:

```
##   FocalNeuroFindings2    FocalNeuroFindings            PainNeck2
##            83.5373650            60.9955364           46.8222912
##   AlteredMentalStatus              PainNeck  AlteredMentalStatus2
##            42.7969303            41.3090863           34.7271957
##          Torticollis2           Torticollis           TenderNeck2
##            33.8839654            29.0046743           24.4498654
##            TenderNeck           HighriskMVC PosMidNeckTenderness2
##            23.4111456            21.6577192           20.8542969
##  PosMidNeckTenderness           SubInj_Head          subinj_Head2
##            19.4959711             6.2439290            2.4715552
##        HighriskDiving       HighriskHanging
##             0.5303960             0.1300819
```

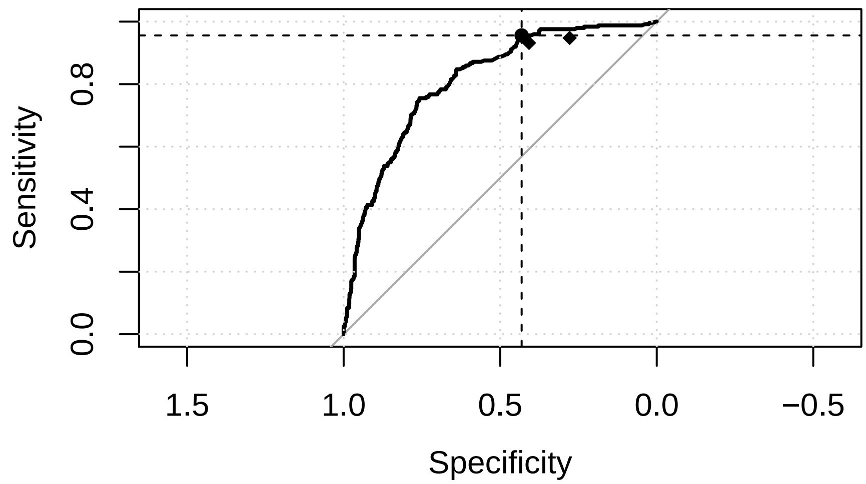
And the tree itself can be illustrated as follows:



One notable advantage to using decisin tree induction, is that most tree induction algorithms can handle missing data, so for the purpose of this classification model we used the whole dataset without having to remove rows with missing data, which should lend the algorithm more statistical power.

Logistic Regression Results

Next we present the results from simple logistic regression, performed on a set of covariates chosen by the feature selection approaches we outlined above. We combined the sets of covariates selected by forward selection with the set of covariates selected by LASSO and used the intersection of the two sets for higher stability.



Again, the logistic regression produced an ROC curve that was strictly dominant to the published decision rule (according to our CV) with AUC = 0.81. A choice of a suitable decision threshold produced the following sensitivity/specificity metrics:

```
##             Estimage Lower CI Bound Upper CI Bound
## Sensitivity 0.9558233      0.9303002      0.9813464
## Specificity 0.4314845      0.4037605      0.4592085
```

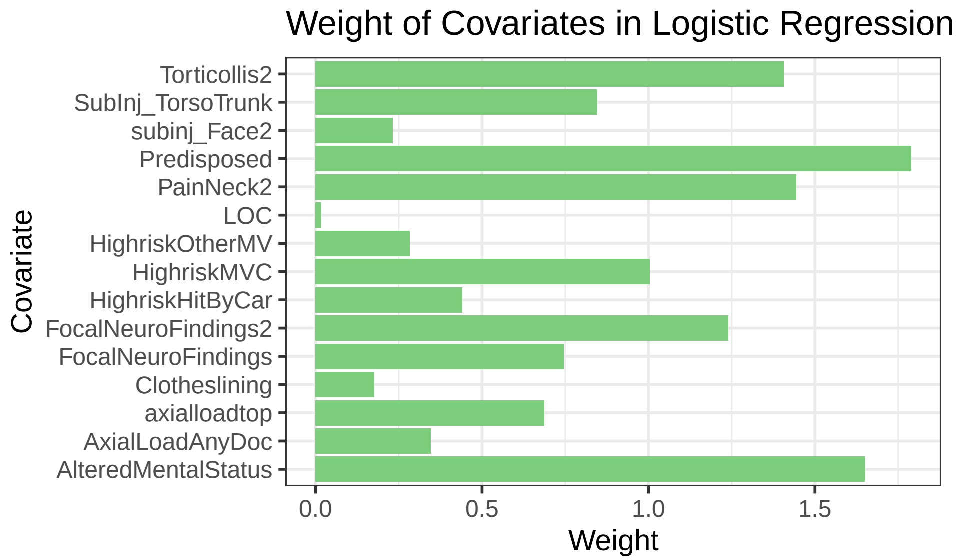
The list of predictors used for the logistic regression model is:

```
##         (Intercept) AlteredMentalStatus  FocalNeuroFindings    HighriskHitByCar
##       -3.8554828845        1.6802206843        0.8701366919        0.5783185541
##         HighriskMVC           PainNeck2         Predisposed   SubInj_TorsoTrunk
##        1.1881910273        1.5023096743        2.1246147166        1.0240226218
## FocalNeuroFindings2        axialloadtop         Torticollis2                 LOC
##        1.1562696439        0.4142044335        1.2980725423        0.0006391999
##     HighriskOtherMV        Clotheslining        subinj_Face2     AxialLoadAnyDoc
##        0.4061711249        0.5891214496        0.2875059958        0.4713429568
```

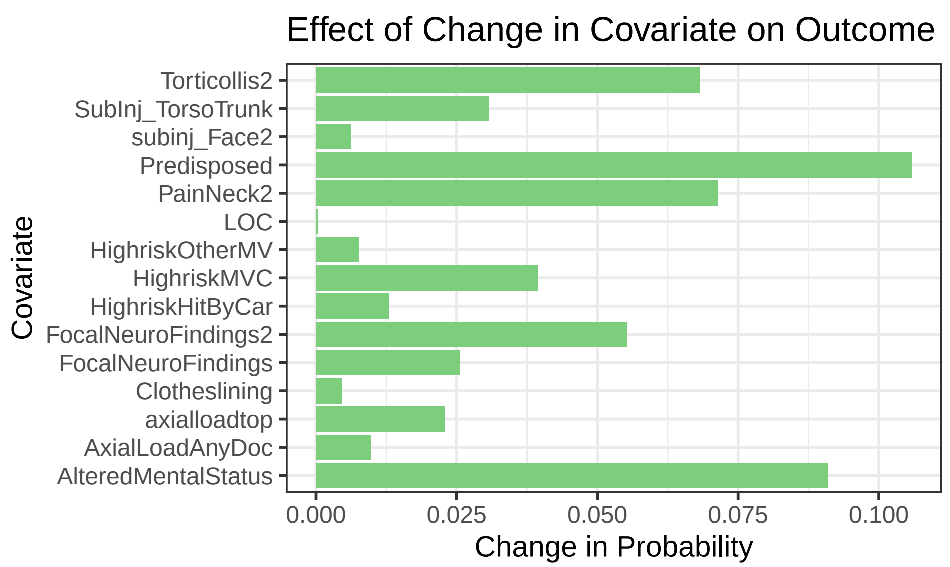
In addition to these two approaches we have experimented with neural networks and gradient boosting (on stubs and trees). Both approaches yielded results that were not much better than the approaches we presented here, and produced models that were more difficult to interpret, so we chose to omit them from this report.

Interpretation

Our model benefits from the interpretability of logistic regression. When a patient is inferenced using the model, the model gives a probability of that patient having a cervical spine injury. That said, due to the necessity for high sensitivity in the model, we declare any patient with a probability of an injury greater than 7.9% as a patient who needs further imaging. It is also possible to view the weights of the features, which demonstrate which features contribute the most to the final probability of injury:



It is important to note that the values of the coefficients in logistic regression cannot be interpreted the same way the coefficients in linear regression can be interpreted. An increase in any covariate by 1, does not results in linear change in the output by the value of B_j, where B_j is the value of the coefficient for that variable in the model. Hence, an increase in any of the covariates by one would then yield an increase in the odds ratio by $exp(B_j)$, where B_j is the weight for that particular covariate. However, coefficients with greater weight still can have a greater effect on the output probability from logistic regression, which we will examine below:



In the above plot, we find the difference in the outcome probability of logistic regression when the value of each covariate is changed from 0 to 1, while all other covariates are held at 0. We chose to hold all other covariates at 0 as the mode for each variable is 0. As you can see, the largest coefficients from logistic regression have the largest changes in probability when that particular covariate is present in the patient.

However, these particular changes in probability are dependent on all of the other covariates being held at 0, so a medical professional could not easily state that a patient being predisposed for a cervical spine injury would have a 10% increase in probability of having a cervical spine injury versus if they were not predisposed. However, these values do serve to give a sense on how certain variables affect the outcome more than others.

Stability with Model Perterbation

Introduce a perturbation to your final model, and summarize the effects of this perturbation on the predictions of your model.