Interpreting Random Forest Predictions for Firearm Identification Using LIME

Katherine Goode and Heike Hofmann

October 22, 2018

1 Introduction

The discipline of firearm identification examines bullets to determine the likelihood that a bullet found in a criminal case was fired from a particular gun. This is a procedure that has traditionally been performed by hand. Specially trained examiners visually assess the striations on the unknown bullet from the crime that were created when the bullet passed through the gun barrel to those on a known bullet that was fired from the gun under evalution. Often a comparison microscope is used that allows the examiners to view both bullets at the same time (National Research Counci 2009). The examiners use the ability to compare the two bullets to determine whether the two bullets were fired from the same gun barrel.

Recently, the scientific community has been encouraging the inclusion of more data driven techniques to be used in forensic investigations that would allow for the reporting of a measure of uncertainty in addition to the conclusion drawn from the analysis. This led Hare, Hofmann, and Carriquiry (2017) to propose a new computer automated method of bullet matching. The method involves obtaining bullet signatures from the scans of the two bullets, computing a handful of variables that measure the similarity between two signatures, and fitting a random forest model to the similarity variables. The model can be trained on a set of known bullet matches and non-matches. Then it can be used to predict on a new set of bullet signature comparisons. The results from the paper suggest that the random forest model leads to highly accurate bullet matching predictions. The authors found that when their model was applied to a testing dataset (?), the resulting error rate was 0%.

While random forest models often result in good predictions, it is well known that a disadvantage of random forest models and other machine learning techniques is that it is difficult to interpret the models (reference?). For example, it is not possible to tell which variables played an important role in the creation of individual predictions. This issue led to the development of LIME (reference), which is an algorithm that examines the behavior of the complicated model on a local scale around a new prediction using a linear model. This allows for the ability to understand which were the driving variables in a prediction of interest.

Since firearm identification identification is commonly used as evidence for convictions in court cases, it is important to not only provide accurate predictions and an associated uncertainty level but also to understand which variables were important in making the prediction. Firearm examiners could use LIME to check whether or not the predictions created by LIME are reasonable. This paper provides an example of the application of LIME to an updated version of the random forest model from the Hare paper used for bullet matching.

2 Data

2.1 Training Data: The Hamby...

(fill in once we know which data will be used to train the new random forest model)

2.2 Testing Data: The Hamby 224 Clone

The Hamby 224 Clone is organized as a test set of a cloned (sub-)set of the Hamby 224 bullets. As with all Hamby sets (Hamby, Brundage, and Thorpe 2009), Hamby set 224, is a collection of 35 bullets, organized as 20 known bullets and 15 questioned bullets. The known bullets are fired in pairs of two through one of ten consecutively manufactures P-85 barrels. Clone set 224 is arranged as a test set of fifteen tests, one for each questioned bullet. Each test set is arranged as a combination of three bullets: two known bullets and a questioned bullet. The test asks for a decision on whether the questioned bullet comes from the same source as the two known bullets or from a different source. This situation is similar to what a Firearms and Toolmarks Examiner might encounter in case work.

3 Methods

3.1 Random Forest Model

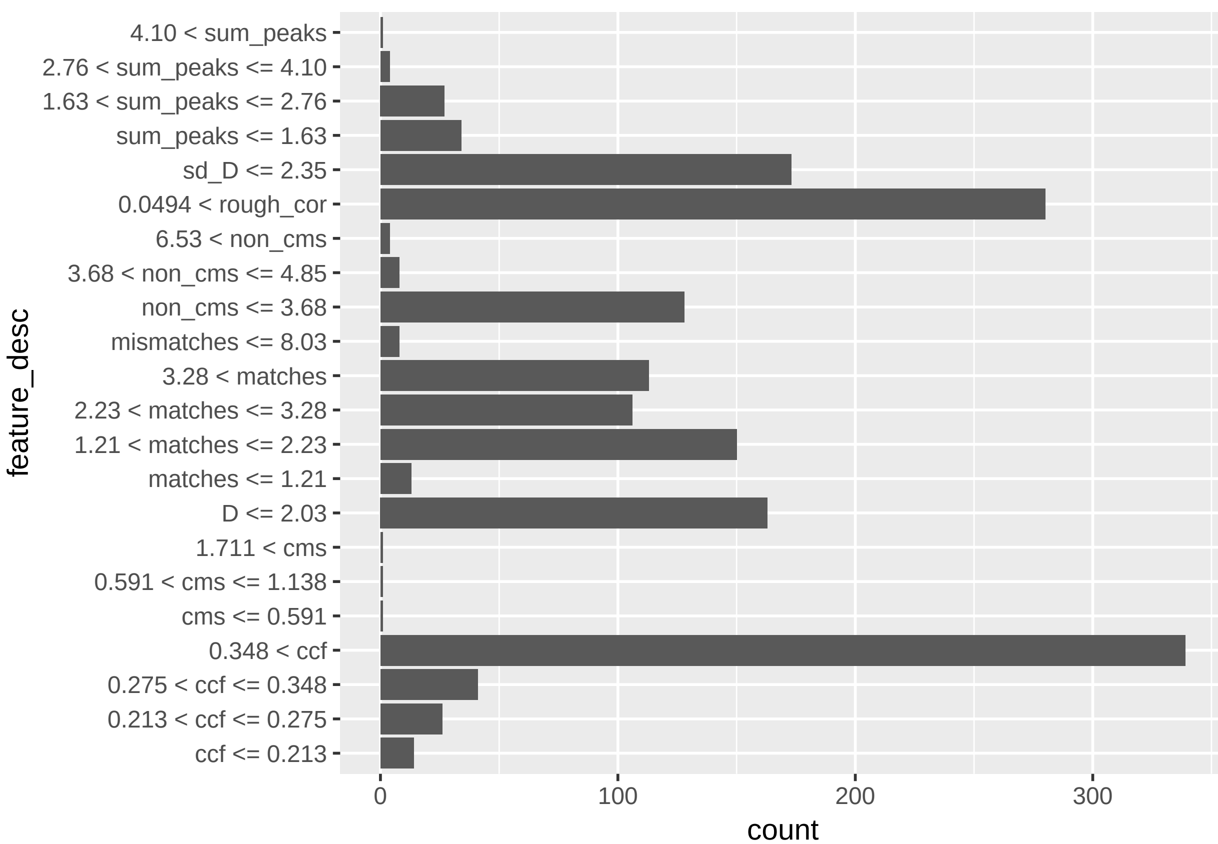
(fill in once the new random forest model has been fit)

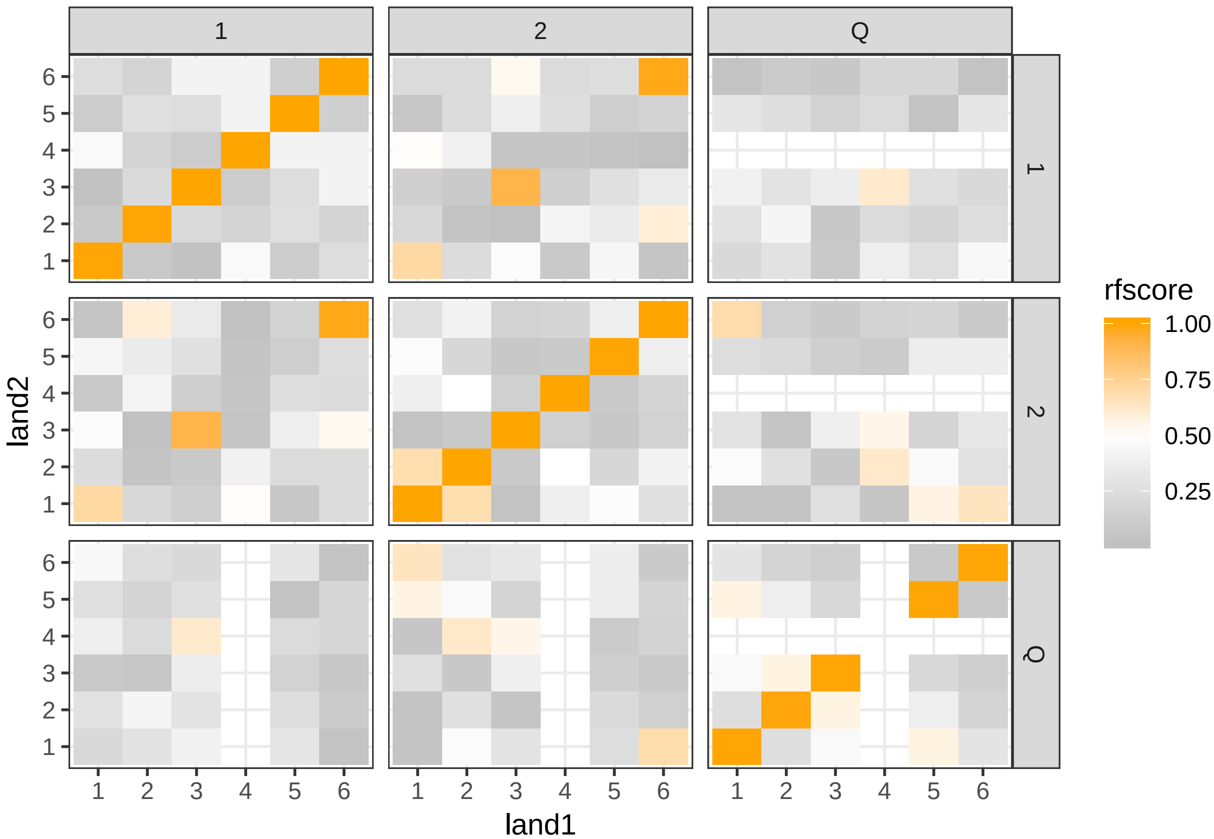
3.2 Overview of LIME

3.3 Applying LIME

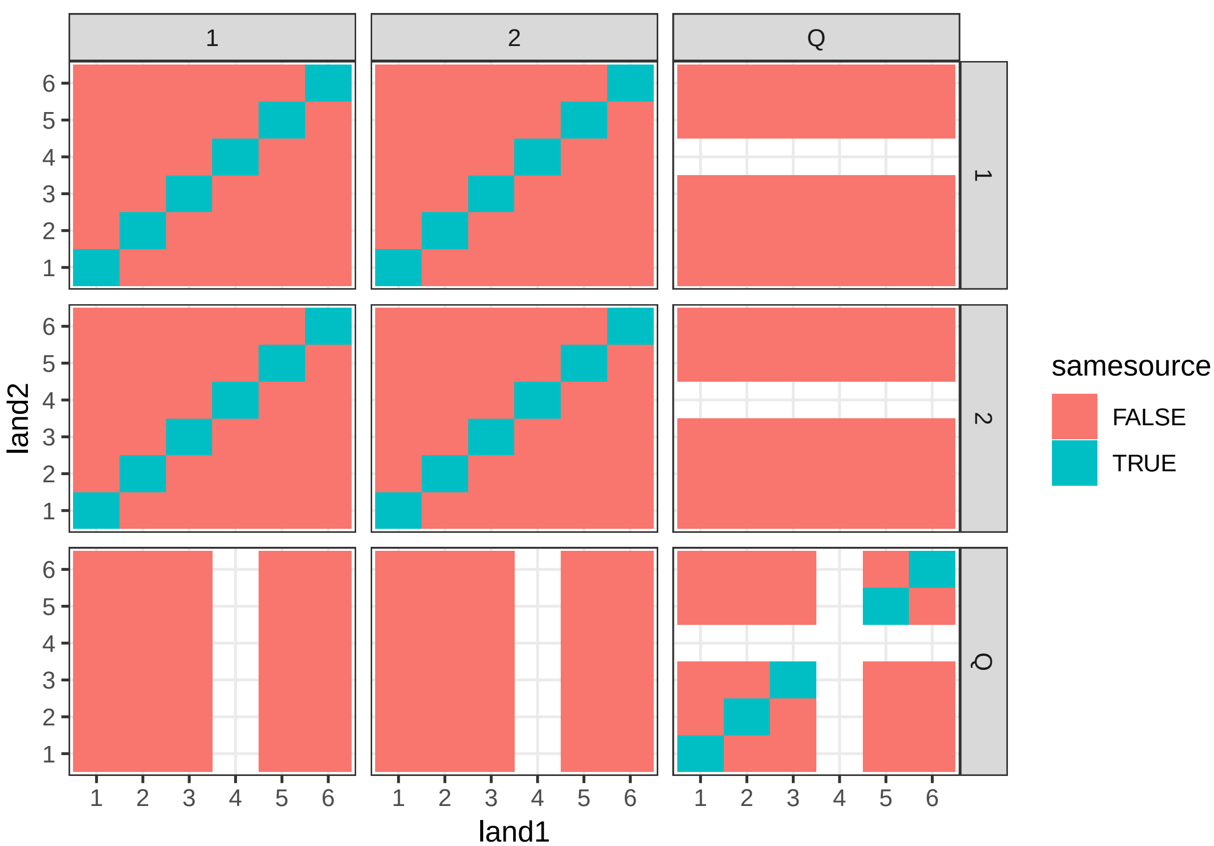
3.4 Visualizing the LIME Explanations

Some plots to tryout working with the new explanations. . .

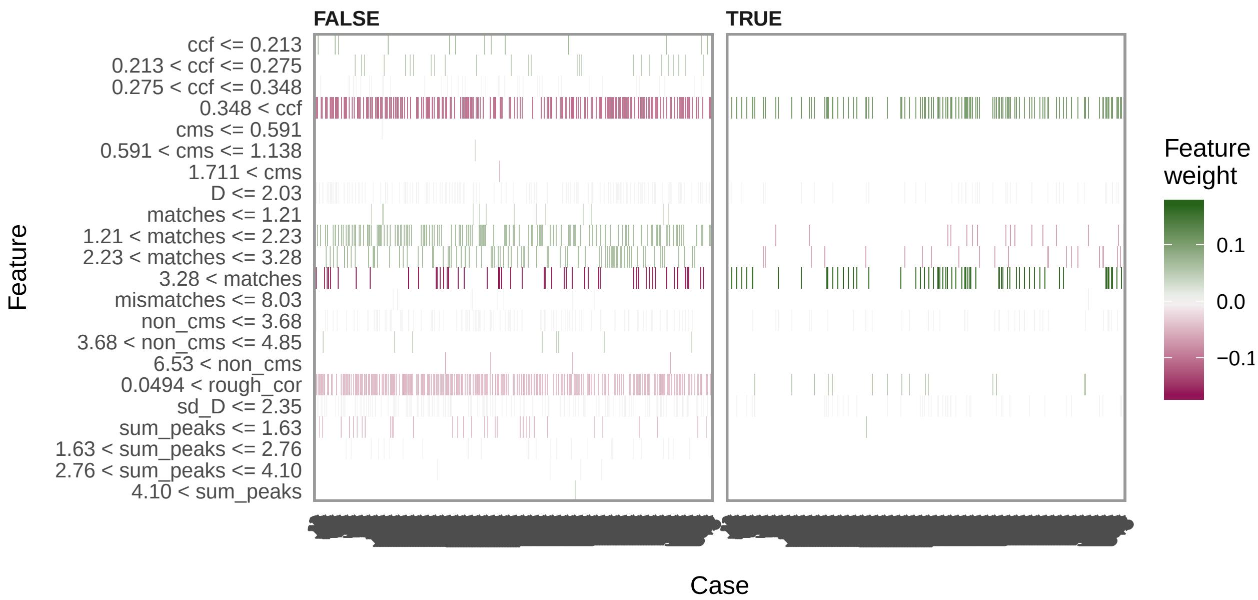


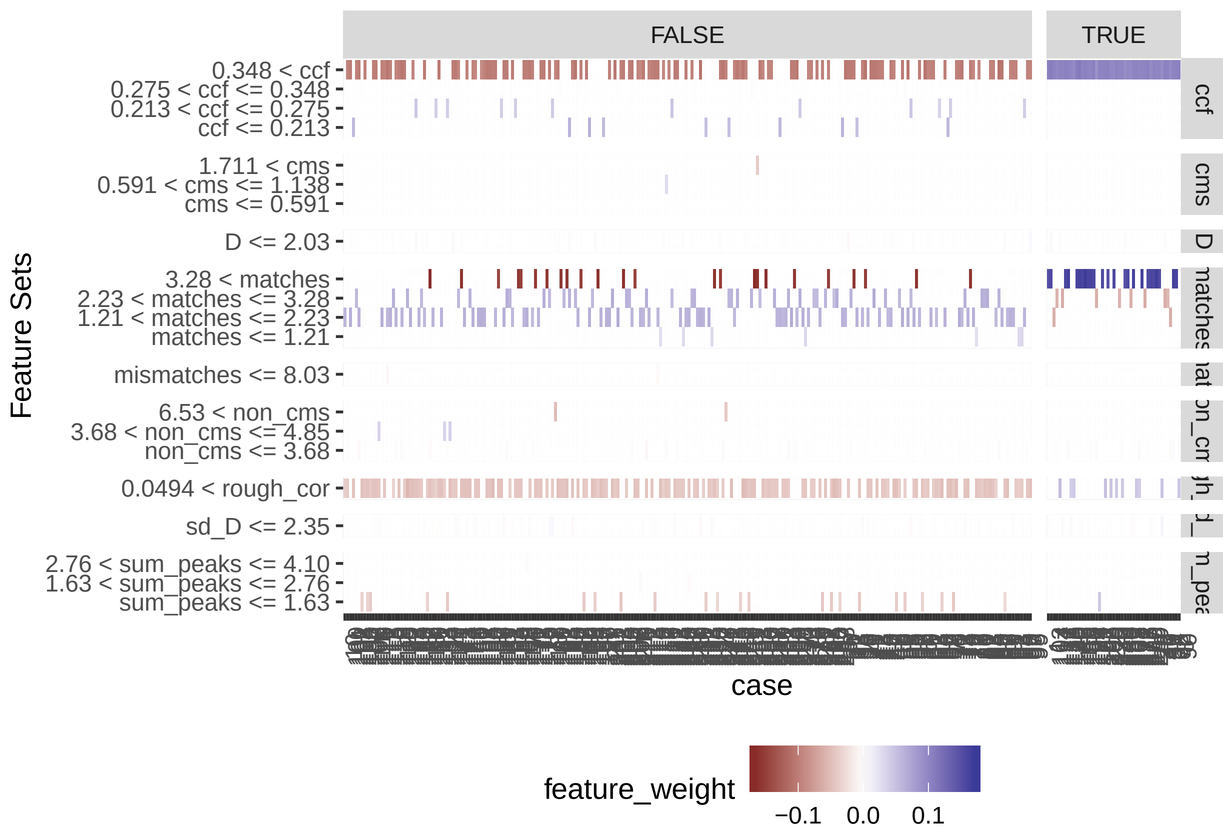


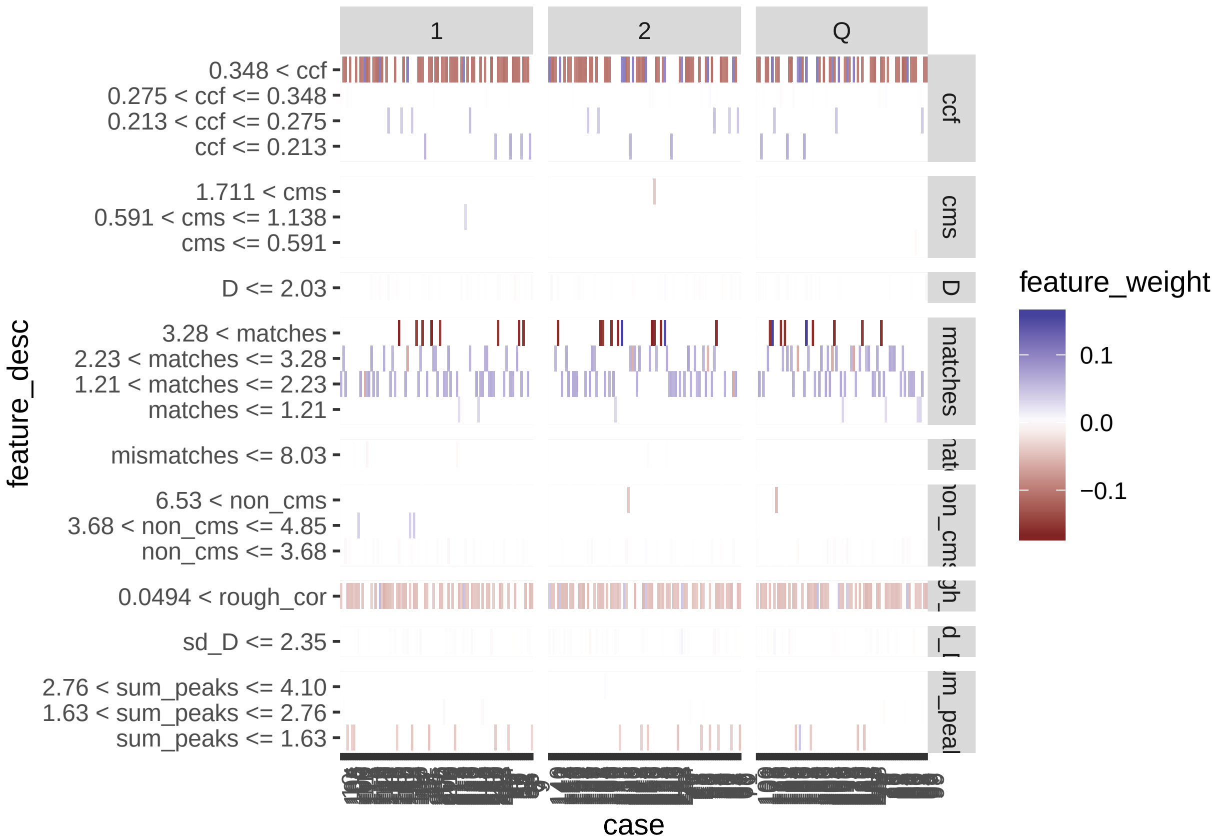
1
2
Q
rfscore
1.00
0.75
0.50
0.25
land2
land1

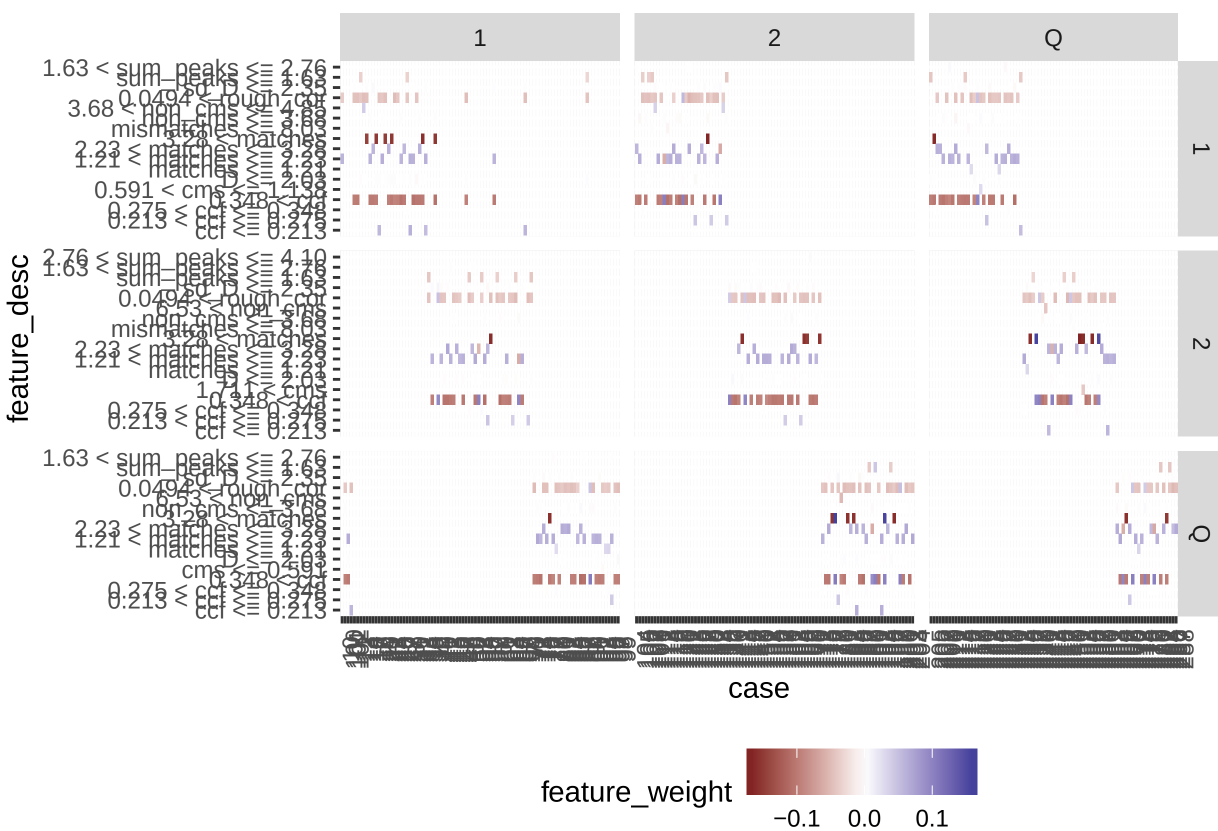


1
2
Q
samesource
FALSE
TRUE
land2
land1


FALSE
TRUE
ccf <= 0.213
0.213 < ccf <= 0.275
0.275 < ccf <= 0.348
0.348 < ccf
cms <= 0.591
0.591 < cms <= 1.138
1.711 < cms
D <= 2.03
matches <= 1.21
1.21 < matches <= 2.23
2.23 < matches <= 3.28
3.28 < matches
mismatches <= 8.03
non_cms <= 3.68
3.68 < non_cms <= 4.85
6.53 < non_cms
0.0494 < rough_cor
sd_D <= 2.35
sum_peaks <= 1.63
1.63 < sum_peaks <= 2.76
2.76 < sum_peaks <= 4.10
4.10 < sum_peaks
Feature
Case
Feature weight
0.1
0.0
−0.1


FALSE
TRUE
0.348 < ccf
0.275 < ccf <= 0.348
0.213 < ccf <= 0.275
ccf <= 0.213
1.711 < cms
0.591 < cms <= 1.138
cms <= 0.591
D <= 2.03
3.28 < matches
2.23 < matches <= 3.28
1.21 < matches <= 2.23
matches <= 1.21
mismatches <= 8.03
6.53 < non_cms
3.68 < non_cms <= 4.85
non_cms <= 3.68
0.0494 < rough_cor
sd_D <= 2.35
2.76 < sum_peaks <= 4.10
1.63 < sum_peaks <= 2.76
sum_peaks <= 1.63
ccf
cms
D
Feature Sets
case
feature_weight
−0.1
0.0
0.1


1
2
Q
feature_desc
0.348 < ccf
0.275 < ccf <= 0.348
0.213 < ccf <= 0.275
ccf <= 0.213
1.711 < cms
0.591 < cms <= 1.138
cms <= 0.591
D <= 2.03
3.28 < matches
2.23 < matches <= 3.28
1.21 < matches <= 2.23
matches <= 1.21
mismatches <= 8.03
6.53 < non_cms
3.68 < non_cms <= 4.85
non_cms <= 3.68
0.0494 < rough_cor
sd_D <= 2.35
2.76 < sum_peaks <= 4.10
1.63 < sum_peaks <= 2.76
sum_peaks <= 1.63
ccf
cms
D
matches
non_cms
case
feature_weight
0.1
0.0
−0.1


1
2
Q
feature_desc
case
feature_weight
−0.1
0.0
0.1

4 Results

5 Discussion

References

Hamby, James E., David J. Brundage, and James W. Thorpe. 2009. "The Identification of Bullets Fired from 10 Consecutively Rifled 9mm Ruger Pistol Barrels: A Research Project Involving 507 Participants from 20 Countries." *AFTE Journal* 41 (2): 99–110.