Interpreting Random Forest Predictions for Firearm Identification Using LIME

Katherine Goode and Heike Hofmann

November 02, 2018

1 Introduction

(need to finish adding sources)

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. To do this, the bullet from the crime will be compared with a bullet that was known to be fired from the gun under evaluation. Traditionally, this is a procedure that has been performed by hand. Specially trained examiners visually compare the microscopic bullet striations that were created when the bullets passed through the gun barrel. Often a comparison microscope is used that allows the examiners to view both bullets at the same time (National Research Council 2009). The examiners use this ability 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. These methods 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 that could supplement the visual inspection by the firearm examiners. Their method involves obtaining bullet signatures of the striations from the scans of the two bullets, computing variables that measure the similarity of the two signatures, and using a trained random forest model to determine the probability of a match based on the similarity variables. They trained their model on a set of known bullet matches and non-matches from the Hamby study. They demonstrated how the random forest model can be used to make prediction 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 as seen in Hare, Hofmann, and Carriquiry, 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 regression model. This allows for the ability to understand which were the driving variables that led to a prediction of interest.

Since firearm identification identification is commonly used as evidence for convictions in court cases, it is important to be able to understand and assess the model that is being used to quantify the probability that a bullet was fired from a gun. LIME provides the ability to understand which were the key variables used by the random forest model to make a prediction, which would allow firearm examiners to check whether or not the predictions created by the random forest are based on reasonable variables.

This paper provides an example of the application of LIME to a bullet matching problem. (provide more details about what is contained in the papers - i.e. section 2 describes the Hamby data; section 3 describes the random forest model and LIME; . . .)

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

3.2 Overview of LIME

Step 0: Split iris data into training and testing datasets Step 1: Fit a complex model to the training data Step 2: Obtain distributions of the variables from the training data Step 3: Sample from each of the variable distributions n times Step 4: Obtain predictions for sampled values using the complex model For each testing case use the random forest model to make a prediction for each of the $n = 5000$ samples In the iris data, the predictions are represented by the probability that a flower is a particular species Step 5: Obtain similarity score between data observation and sampled values LIME uses exponential kernel function ... where x_{obs}: observed data vector to predict $x_{sampled}$: sampled data vector from distribution of training variables $D(\cdot, \cdot)$: distance function such as euclidean distance, cosine distance, etc. σ: width (default set to 0.75 in lime) Step 6: Perform feature selection by fiting a model to the sampled data and associated predictions (weighted by the similarity scores) The user can specify the number of variables (features) they would like to select: m With the iris data, the following three models will be fit to perform variable selection to select $m = 2$ features: To perform variable selection lime supports: forward selection with ridge regression highest weight with ridge regression LASSO tree models auto: forward selection if $m \leq 6$, highest weight otherwise Step 7: Fit a simple model to regress the predictions on the m selected predictor variables (weighted by the similarity scores) Currently, `lime` is programmed to use ridge regression as the simple model If the response is categorical, the user can select how many categories they want to explain In this example, only setosa will be explained If petal length and sepal length were selected as the most important features for the first case in the testing data, then the simple model is Step 8: Extract the feature weights and use them as the explanations

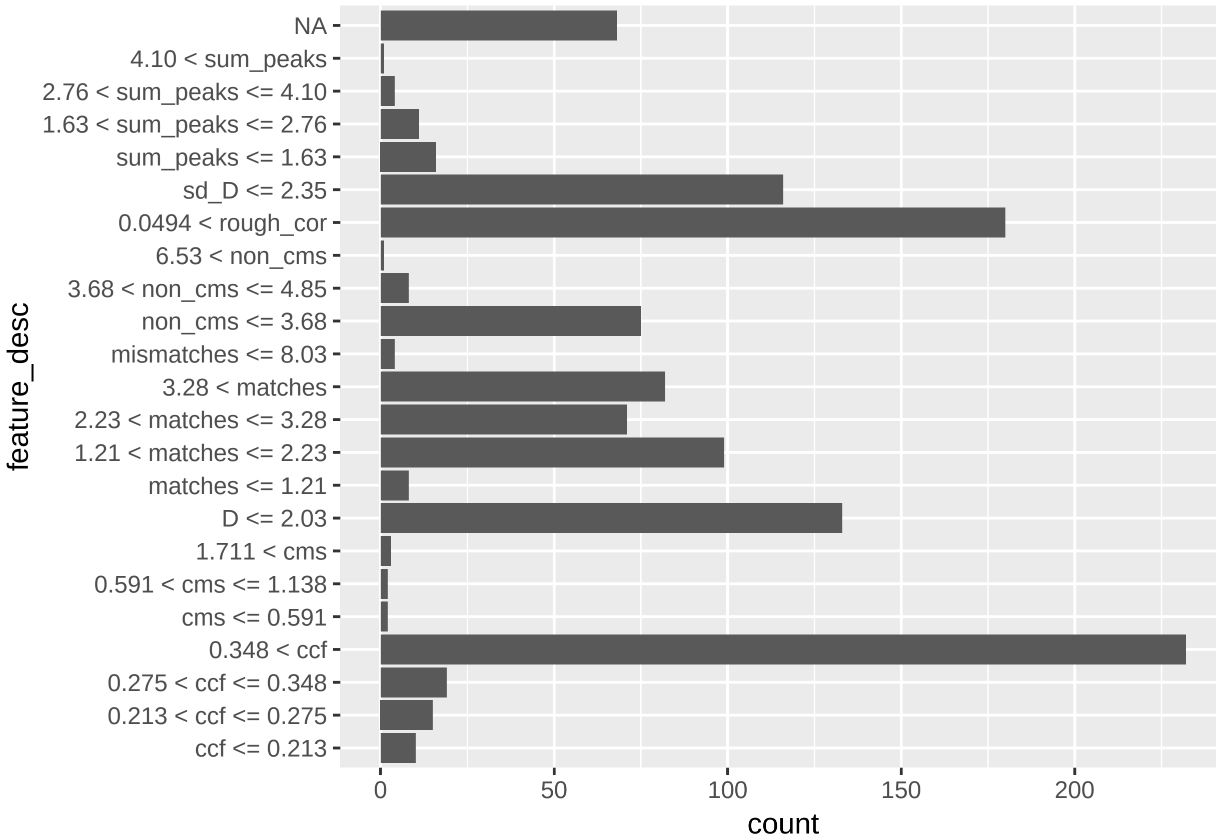
3.3 Applying LIME

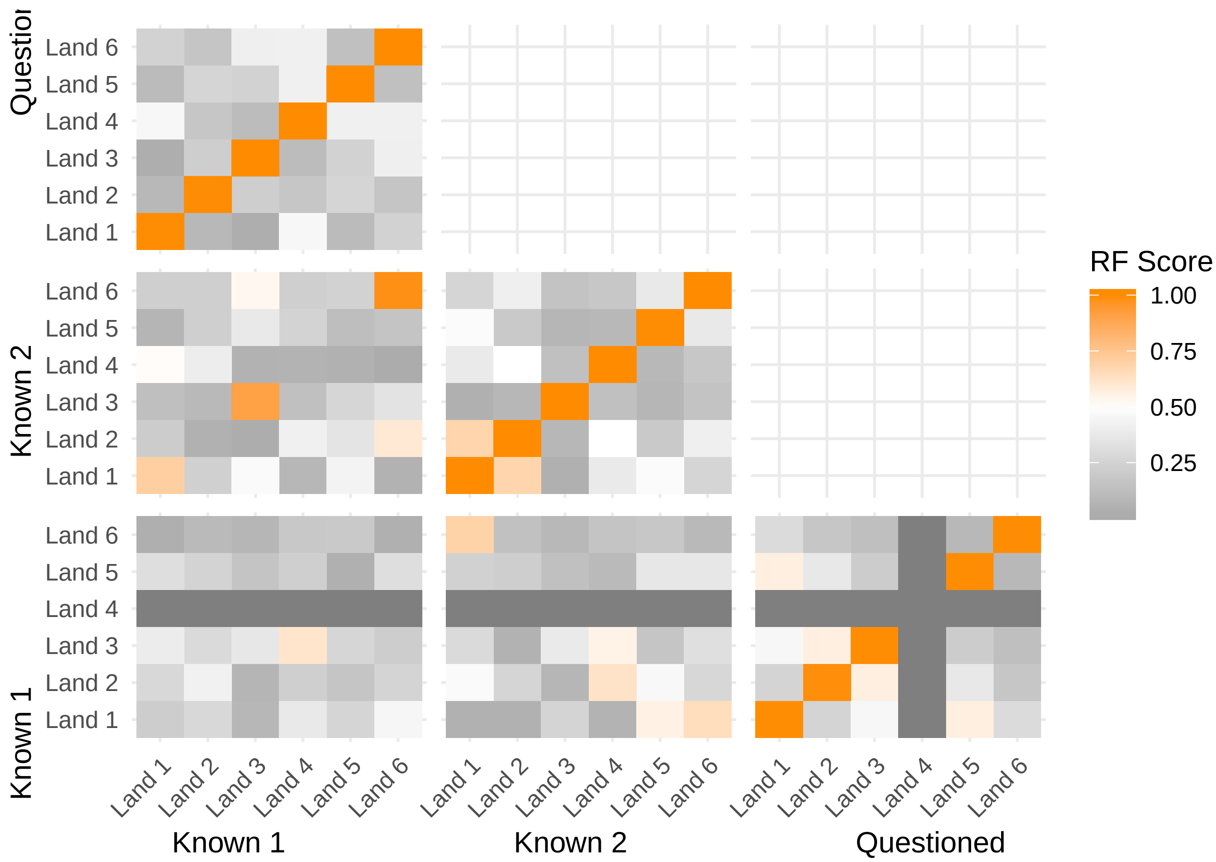
3.4 Visualizing the LIME Explanations

Thoughts on the visualizations:

- It would be interesting to see the "magnitude" associated with a variable for a prediction, but this would require obtaining the scaled values of the perturbations.

- This seems much more meaningful to me than just the model coefficient associated with the variable.
- The current version shows which variables in the model have the largest coefficients.
- However, the importance of the variable might change depending on the observed value associated with the prediction of interest.





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.