

# Interpreting Random Forest Predictions for Bullet Matching Using LIME (?)

*Katherine Goode and Heike Hofmann*

*September 20, 2018*

```
library(tidyverse)
library(bulletr)
library(lime)
```

## Background/Literature Review

## Problem Statement

## Overview of LIME

## 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.

## Applying LIME

```
# Load in the training data (Hamby Data 173 and 252)
hamby173and252_train <- read.csv("../data/hamby_train.csv")

# Obtain features used when fitting the rtrees random forest
rf_features <- rownames(rtrees$importance)

# Load in the Hamby 224 datasets
hamby224_set1 <- readRDS("../data/h224-set1-features.rds")
hamby224_set11 <- readRDS("../data/h224-set11-features.rds")

# Clean the Hamby 224 set 1 data
hamby224_set1_cleaned <- hamby224_set1 %>%
  select(-bullet_score, -land1, -land2, -aligned, -striae, -features) %>%
  rename(bullet1 = bulletA,
         bullet2 = bulletB,
         land1 = landA,
```

```

    land2 = landB) %>%
mutate(study = factor("Hamby 224"),
       set = factor("1"),
       bullet1 = factor(bullet1),
       bullet2 = factor(bullet2),
       land1 = factor(land1),
       land2 = factor(land2)) %>%
select(study, set, bullet1:land2, rf_features, rf_score, same_source)

# Clean the Hamby 224 set 11 data
hamby224_set11_cleaned <- hamby224_set11 %>%
  select(-bullet_score, -land1, -land2, -aligned, -striae, -features) %>%
  rename(bullet1 = bulletA,
         bullet2 = bulletB,
         land1 = landA,
         land2 = landB) %>%
mutate(study = factor("Hamby 224"),
       set = factor("2"),
       bullet1 = recode(factor(bullet1), "Bullet 1" = "1", "Bullet 2" = "2", "Bullet I" = "I"),
       bullet2 = recode(factor(bullet2), "Bullet 1" = "1", "Bullet 2" = "2", "Bullet I" = "I"),
       land1 = recode(factor(land1), "Land 1" = "1", "Land 2" = "2", "Land 3" = "3",
                        "Land 4" = "4", "Land 5" = "5", "Land 6" = "6"),
       land2 = recode(factor(land2), "Land 1" = "1", "Land 2" = "2", "Land 3" = "3",
                        "Land 4" = "4", "Land 5" = "5", "Land 6" = "6")) %>%
select(study, set, bullet1:land2, rf_features, rf_score, same_source)

# Join the two cleaned Hamby 224 sets into one testing set
hamby224_test <- suppressWarnings(bind_rows(hamby224_set1_cleaned, hamby224_set11_cleaned))

# Export the test data as a .csv file
#write.csv(hamby224_test, "../data/hamby224_test.csv", row.names = FALSE)

# Set a seed
set.seed(84902)

# Apply the lime function from the lime package
# Note that the as_classifier must be added since rtrees is from the
# randomForest package and not fit using caret or one of the other
# available models specified in the lime package. Additionally, the
# randomForest package must be loaded in order to run the functions
# from the lime package. (I should check on exactly how this works.)
hamby224_lime <- lime(x = hamby_train %>% select(rf_features),
                    model = as_classifier(rtrees))

# Save the lime object
saveRDS(hamby224_lime, "../data/hamby224_lime.rds")

```

## Results (using LIME on a TS)

## 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.