

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

Katherine Goode and Heike Hofmann

September 24, 2018

1 Introduction

- background
- motivation
- literature review
- problem statement

2 Data

2.1 Training Data: The Hamby...

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

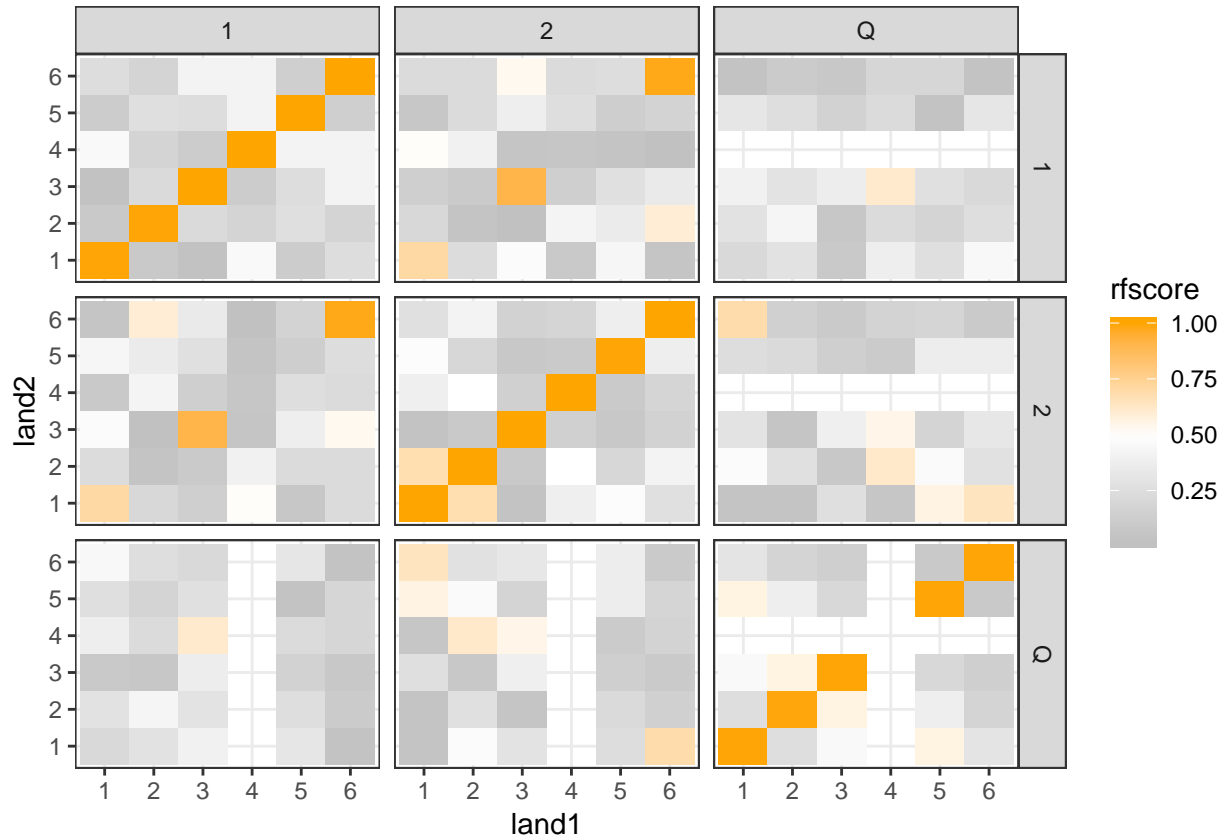
3.3 Applying LIME

3.4 Visualizing the LIME Explanations

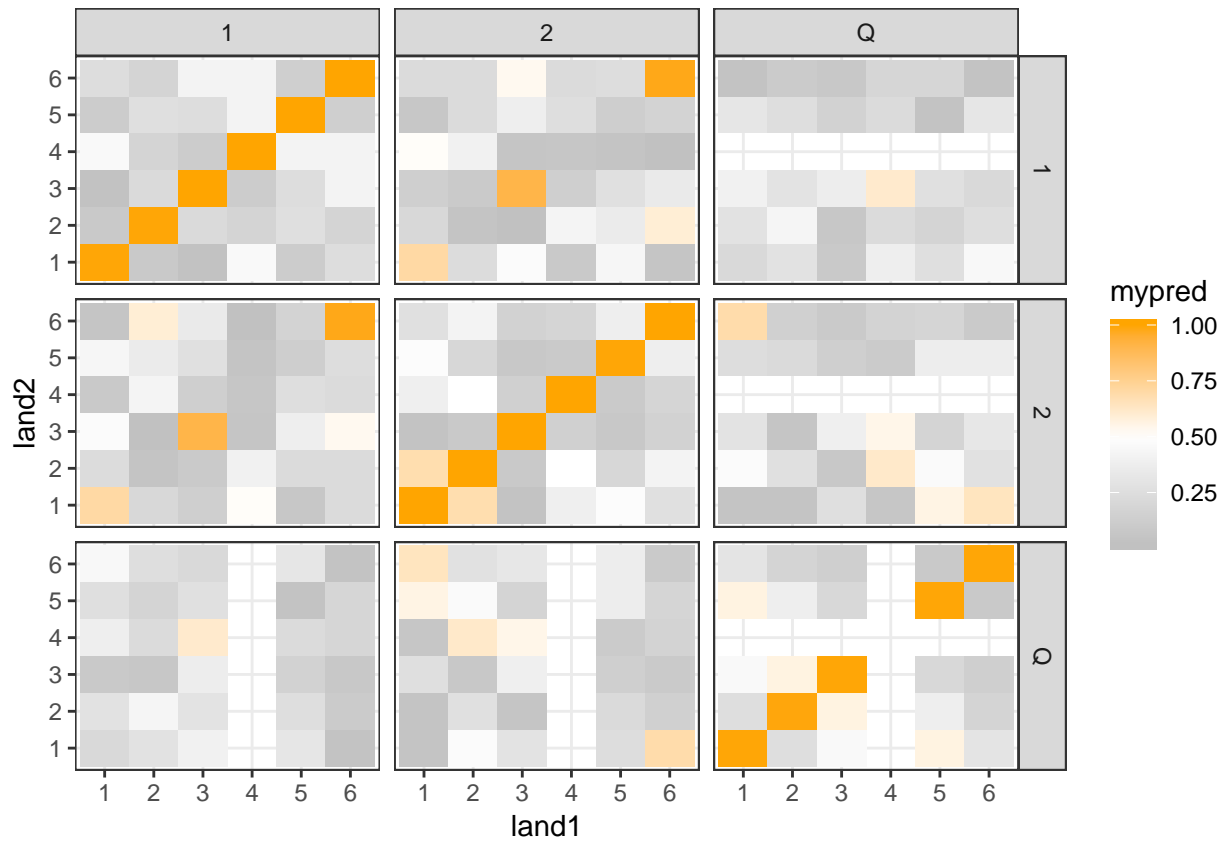
```
# Read in the Hamby 224 combined test data and explanations
hamby224_test_explain <- readRDS("../data/hamby224_test_explain.rds")

hamby224_test_explain %>%
  filter(set == "1") %>%
  select(case, bullet1:land2, rfscore) %>%
  distinct() %>%
```

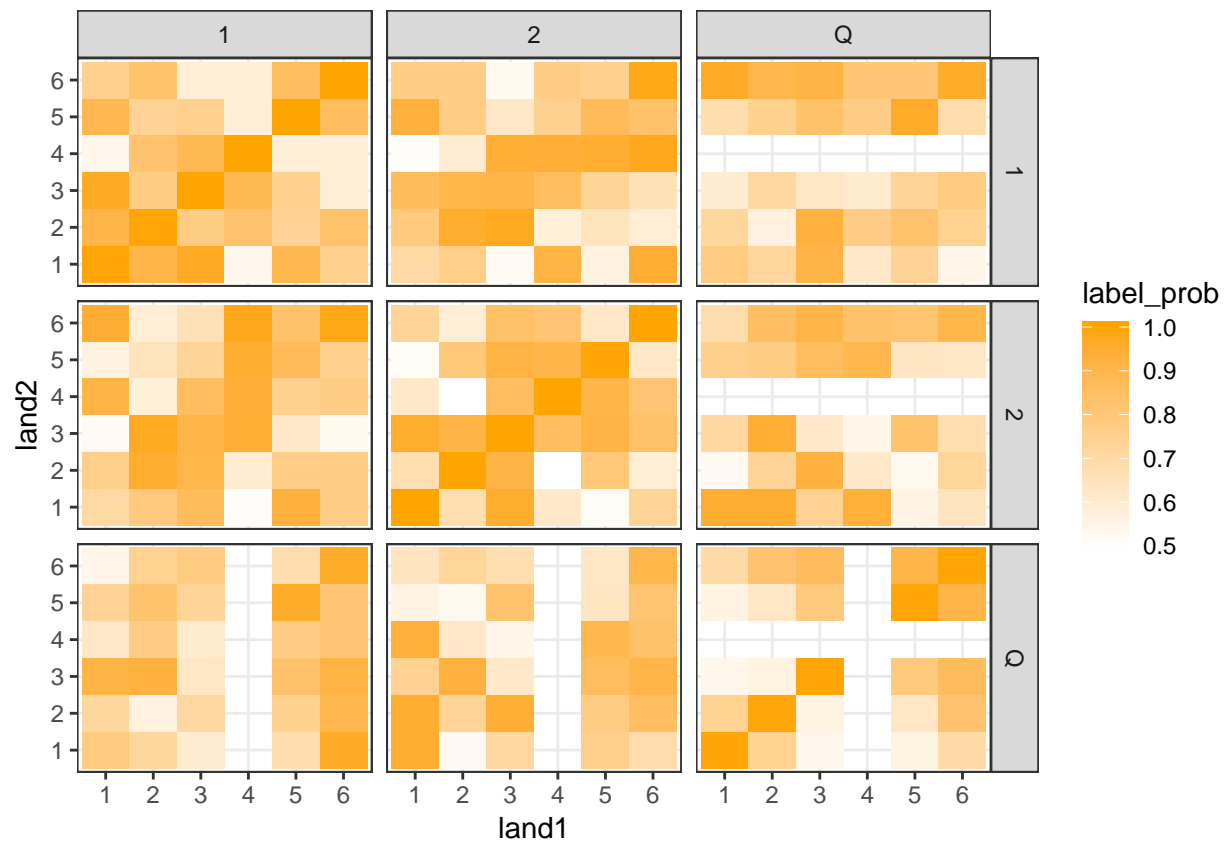
```
ggplot(aes(x = land1, y = land2)) +
  geom_tile(aes(fill = rfscore)) +
  facet_grid(bullet1 ~ bullet2) +
  theme_bw() +
  scale_fill_gradient2(low = "grey", high = "orange", midpoint = 0.5)
```



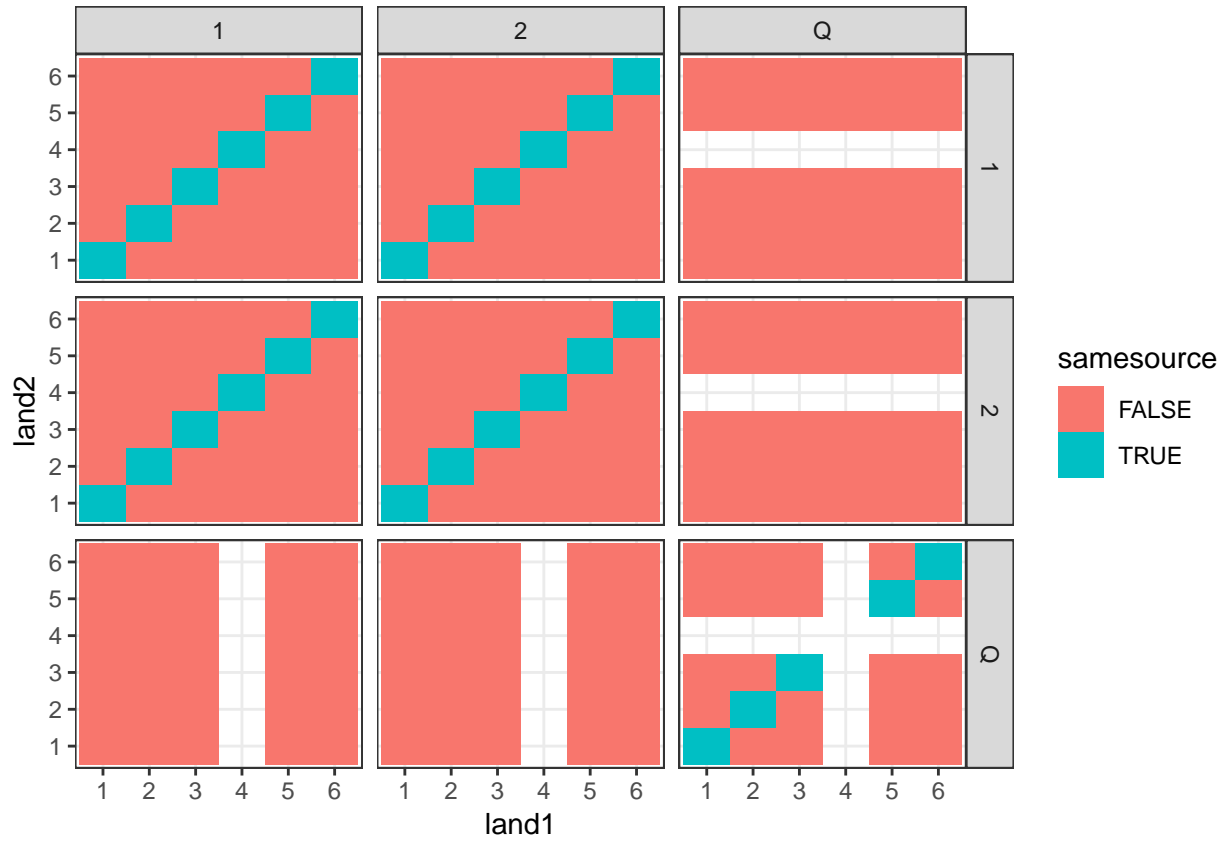
```
hamby224_test_explain %>%
  filter(set == "1") %>%
  select(case, bullet1:land2, mypred) %>%
  distinct() %>%
  ggplot(aes(x = land1, y = land2)) +
  geom_tile(aes(fill = mypred)) +
  facet_grid(bullet1 ~ bullet2) +
  theme_bw() +
  scale_fill_gradient2(low = "grey", high = "orange", midpoint = 0.5)
```



```
hamby224_test_explain %>%
  filter(set == "1") %>%
  select(case, bullet1:land2, label_prob) %>%
  distinct() %>%
  ggplot(aes(x = land1, y = land2)) +
  geom_tile(aes(fill = label_prob)) +
  facet_grid(bullet1 ~ bullet2) +
  theme_bw() +
  scale_fill_gradient2(low = "grey", high = "orange", midpoint = 0.5)
```



```
hamby224_test_explain %>%
  filter(set == "1") %>%
  select(case, bullet1:land2, samesource) %>%
  distinct() %>%
  ggplot(aes(x = land1, y = land2)) +
  geom_tile(aes(fill = samesource)) +
  facet_grid(bullet1 ~ bullet2) +
  theme_bw()
```



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.