

Machine learning on Spark

INTRODUCTION TO SPARK WITH SPARKLYR IN R



Richie Cotton

Data Evangelist at DataCamp

Gradient boosted tree models and random forests

```
model <- training_data %>%  
  ft_gradient_boosted_trees_model("response",  
                                   feature_cols)  
  
results <- testing_data %>%  
  collect() %>%  
  mutate(predicted_response = predict(model,  
                                       testing_data))
```

Let's practice!

INTRODUCTION TO SPARK WITH SPARKLYR IN R