Performance Metrics

Deep Analytics and Visualization

Evaluate Techniques for Wi-Fi Locationing

Greg Hepguler

DATA ANALYTICS OBJECTIVES

- Develop a Wi-Fi Fingerprinting model to locate people indoors in 3 large buildings in a campus

- Use Wi-Fi fingerprint characterized by Wireless Access Points (WAPs) and the corresponding RSSI: Received Signal Strength Intensity

- Compare of the models produced by at least three different algorithms

- Recommend the algorithm that is best suited for this data and a justification why it is the preferred choice

- Recommend improvements that can be achieved, based on research on indoor locationing or experimentation with the dataset.

METRICS for the RESULTS

- Accuracy and Kappa are gathered as the two performance metrics from each model

  ACCURACY is calculated as the proportion that represents the number of true positives and

  true negatives, divided by the total number of predictions

$$Accuracy = (TP+TN)/(TP+TN+FP+FN)$$

		Predictions	
		no	**yes**
Actuals	**no**	TN	FP
	yes	FN	TP

- **KAPPA** adjusts ACCURACY by accounting for the possibility of correct prediction by chance

  alone. A common interpretation is shown as follows*:

Kappa:			
Poor agreement	< 0.20		
Fair agreement	0.20	to	0.40
Moderate agreement	0.40	to	0.60
Good agreement	0.60	to	0.80
Very Good agreement	0.80	to	1.00

* Machine Learning with R, Brett Lantz

- The criteria for the "best model" is the highest value

  for Accuracy and Kappa when predicting the LOCATION

- Accuracy and Kappa are obtained by post resampling predicted values against actuals in

  the validation/testing subset, which is 30% of the original Training dataset in this project.
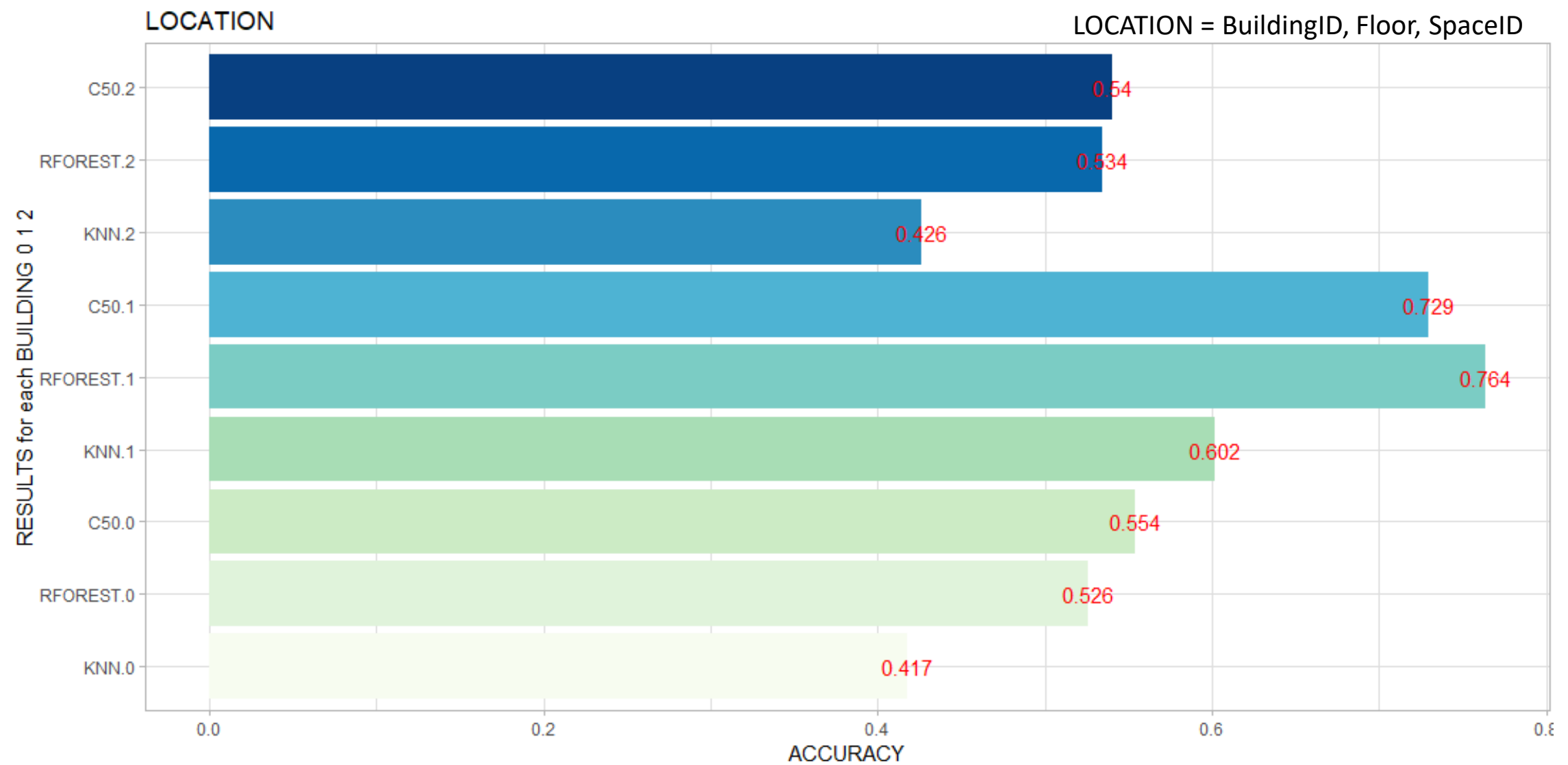
RESULTS:

|		Model Parameter (tunegrid)	Parameter Values	Parameter Value Selected Model	Accuracy	Kappa
MODEL ALGORITHMS, PARAMETERS and METRICS of PREDICTING LOCATION						
Algorithms	**Run**					
kNN	Last	k values	k = 1, 2, 3, 4	k = 1	0.575	0.573
RF	Last	mtry	mtry = 16, 32, 48	mtry = 32	0.701	0.700
C50	Last	trials	trials = 48, 96	trials = 96	0.602	0.600
C50	Next	winnow, trials (Boosting Iterations)	winnow = (TRUE, FALSE), trials=(32, 48, 64), model=("tree", "rules")	winnow = FALSE, trials=24, model="rules"	0.601	0.599
kNN	First	k values	k = 5, 7, 9, 11, 13, 15, 17	k = 5	0.465	0.462
RF	First	mtry	mtry = 1, 2, 3, 5	mtry = 5	0.587	0.585
C50	First	winnow, trials (Boosting Iterations)	winnow = (TRUE, FALSE), trials=c(2, 8, 24),	winnow = FALSE, trials=24	0.587	0.584

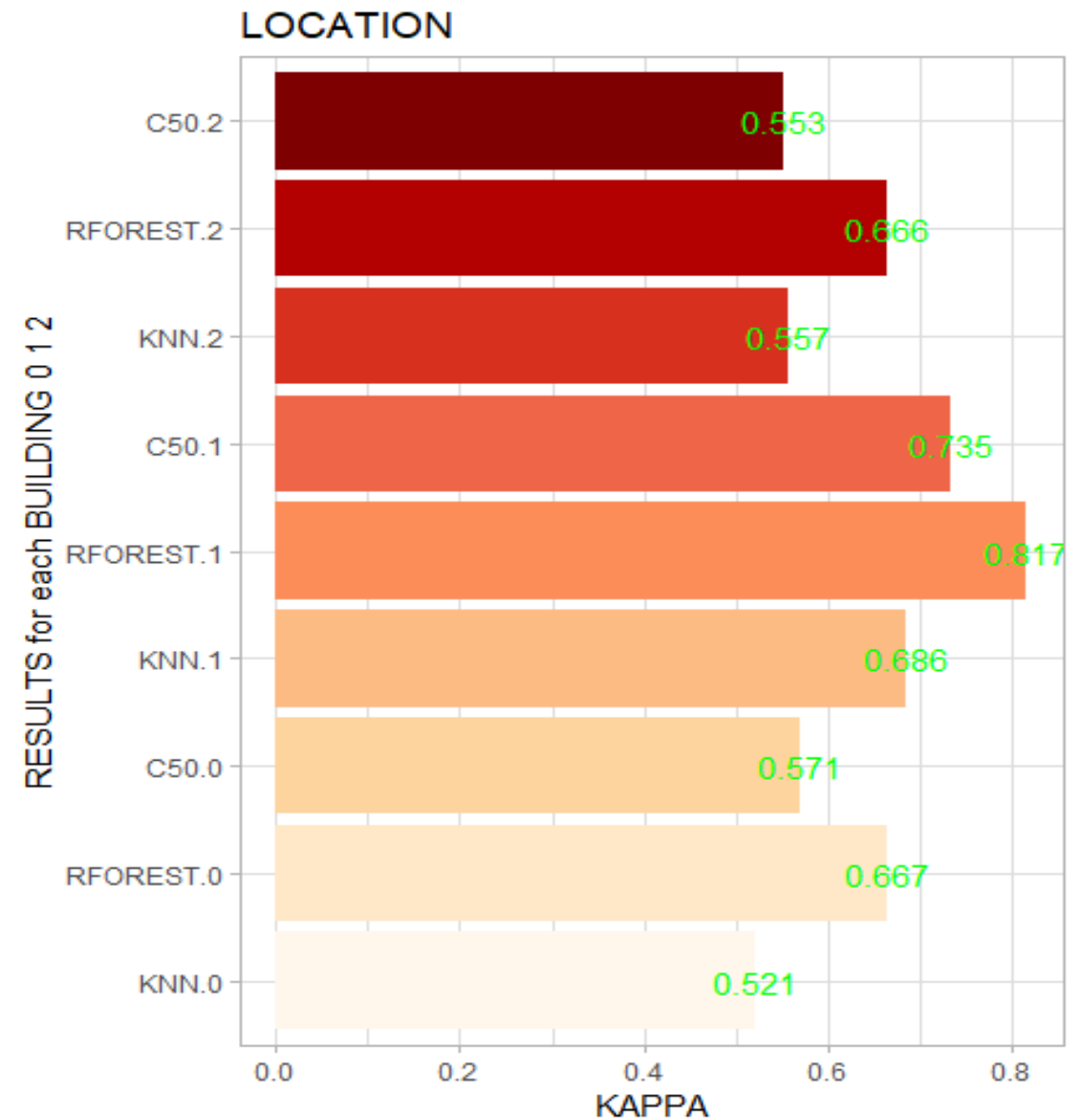# ACCURACY of each Model on the Prediction of LOCATION by Building (1st run)

LOCATION

LOCATION = BuildingID, Floor, SpaceID

RESULTS for each BUILDING 0 1 2

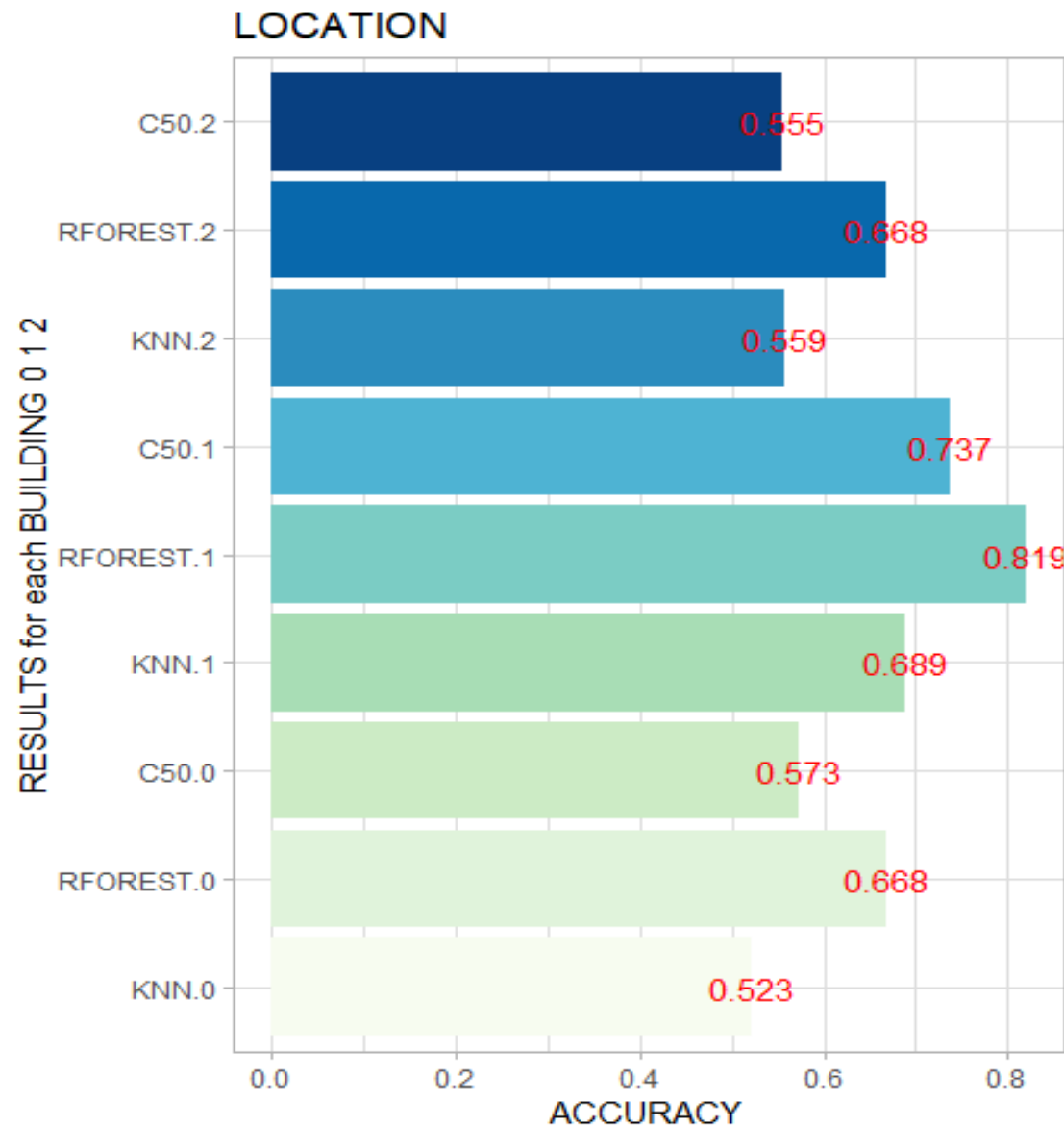| Model	Accuracy
C50.2	0.54
RFOREST.2	0.534
KNN.2	0.426
C50.1	0.729
RFOREST.1	0.764
KNN.1	0.602
C50.0	0.554
RFOREST.0	0.526
KNN.0	0.417

ACCURACY

KAPPA for each Model on the Prediction of LOCATION by Building (1st run)
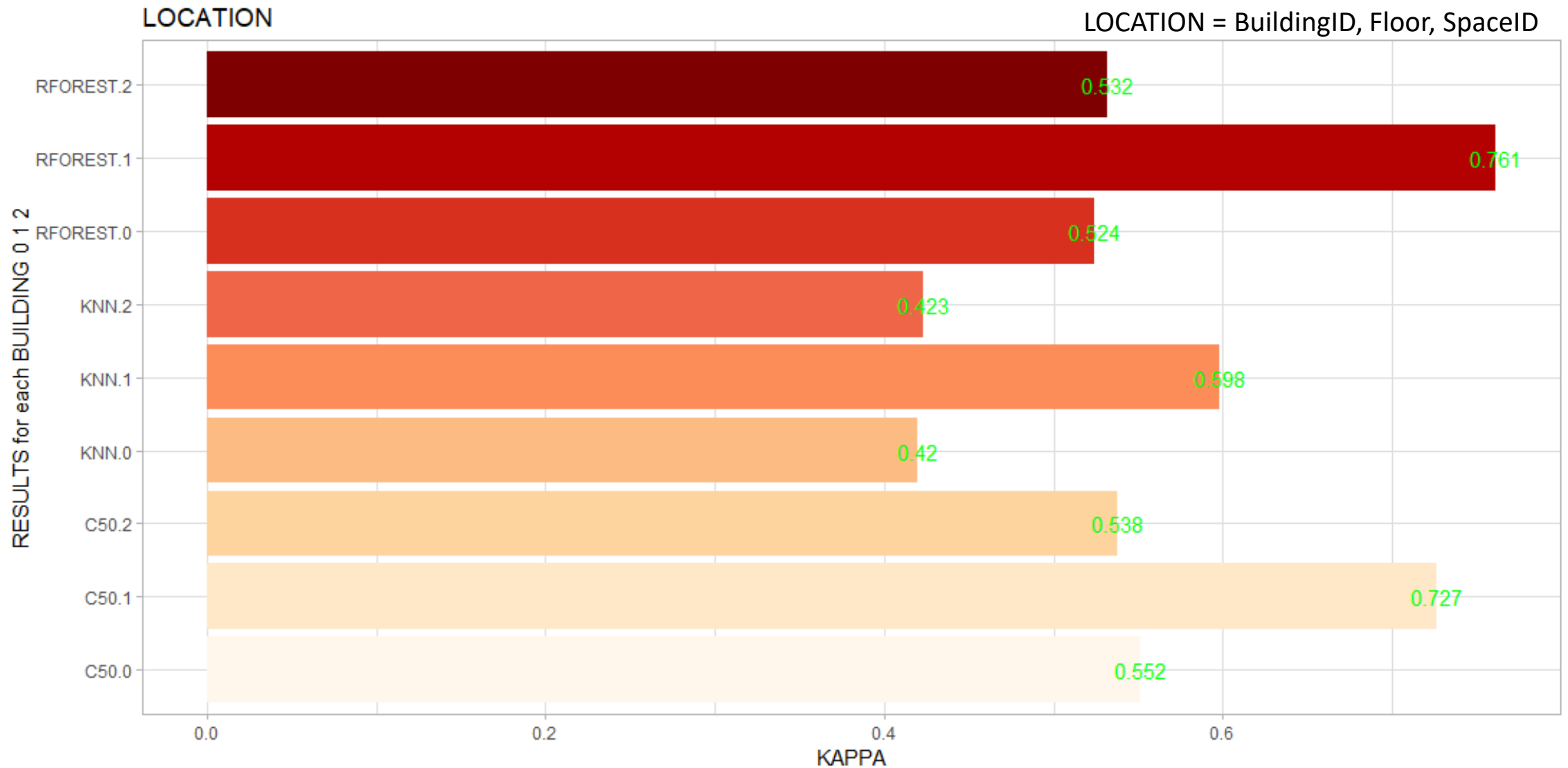
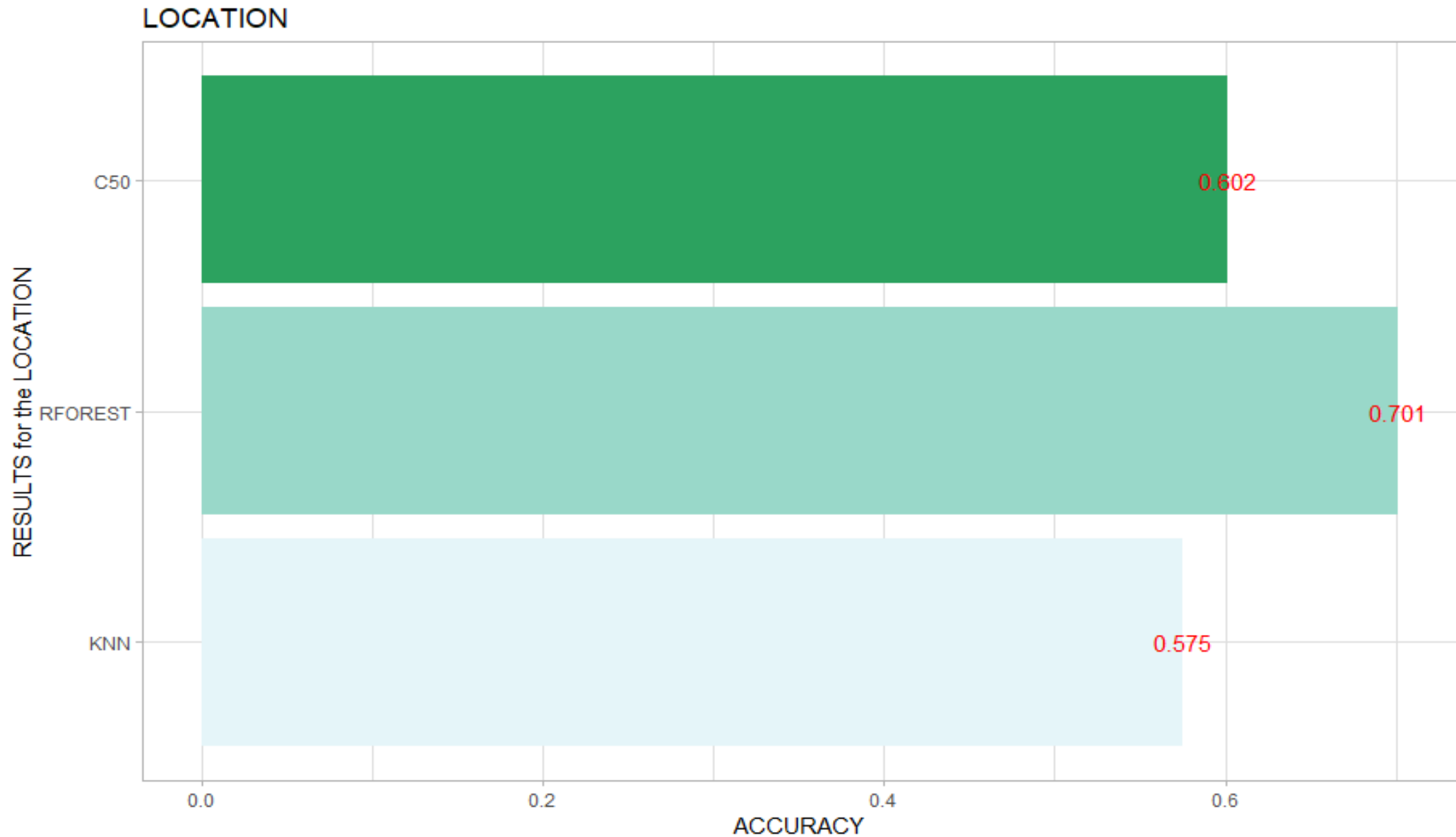# Metrics for each Model on the Prediction of LOCATION by Building (1ˢᵗ run)
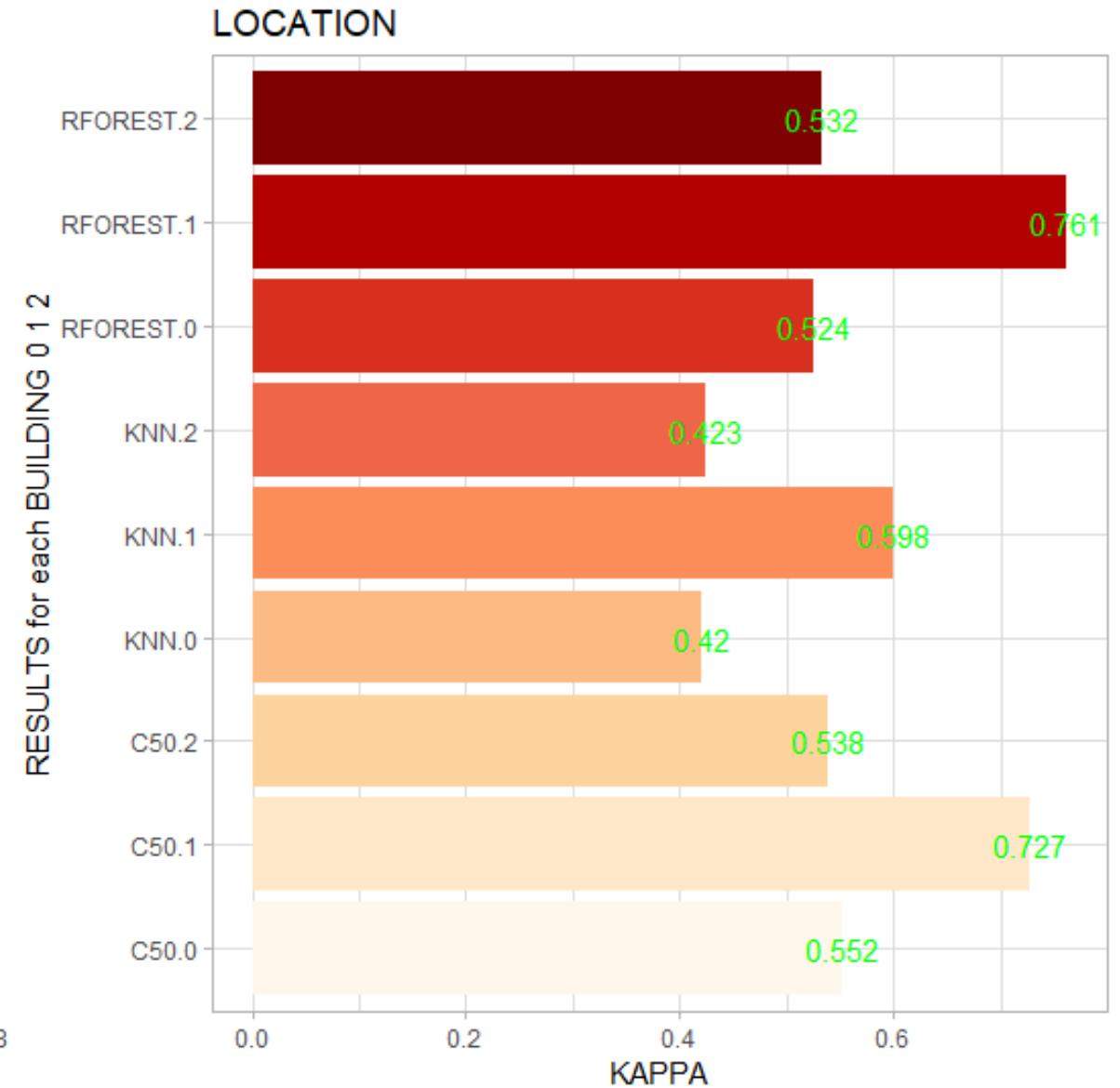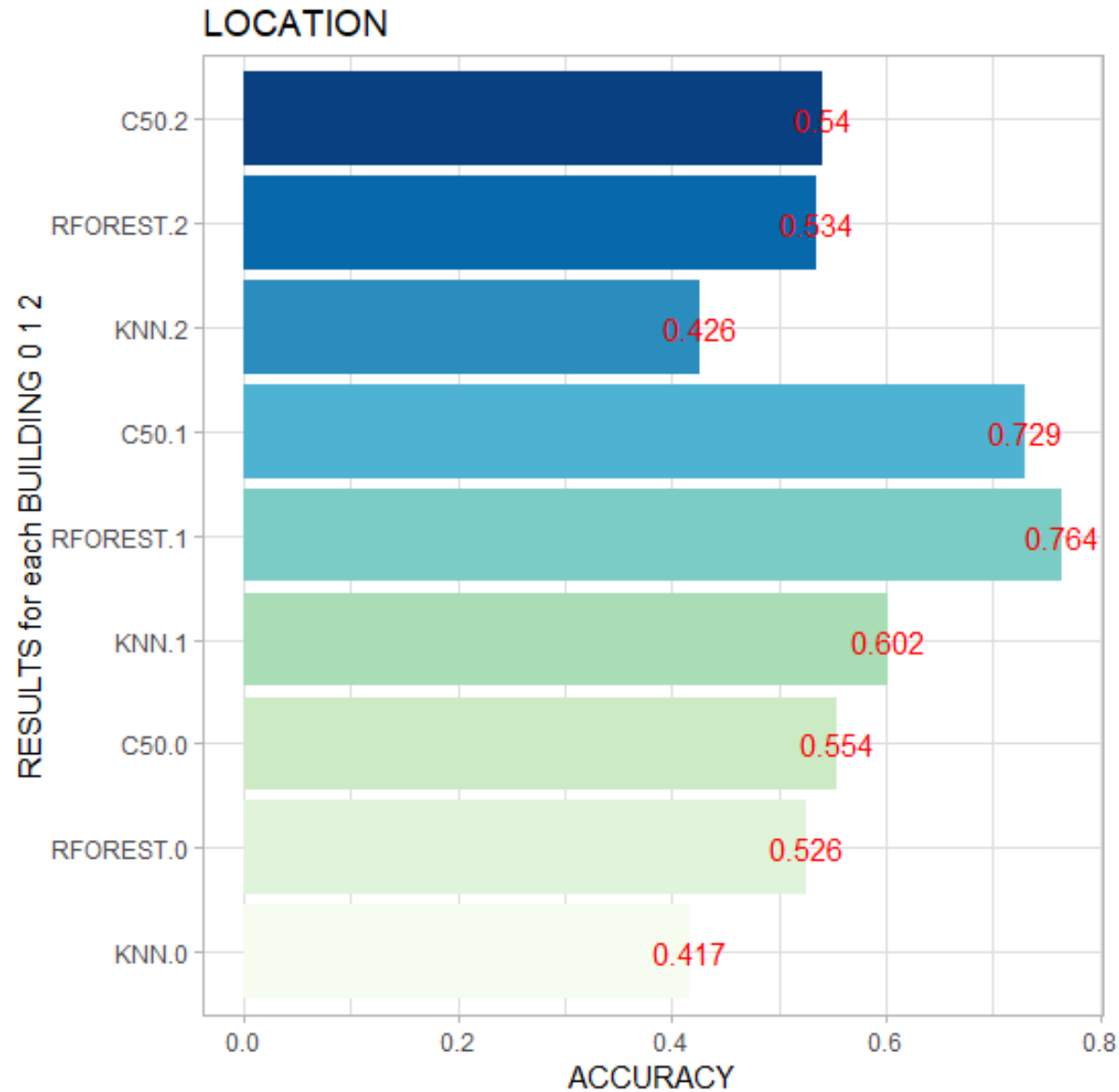
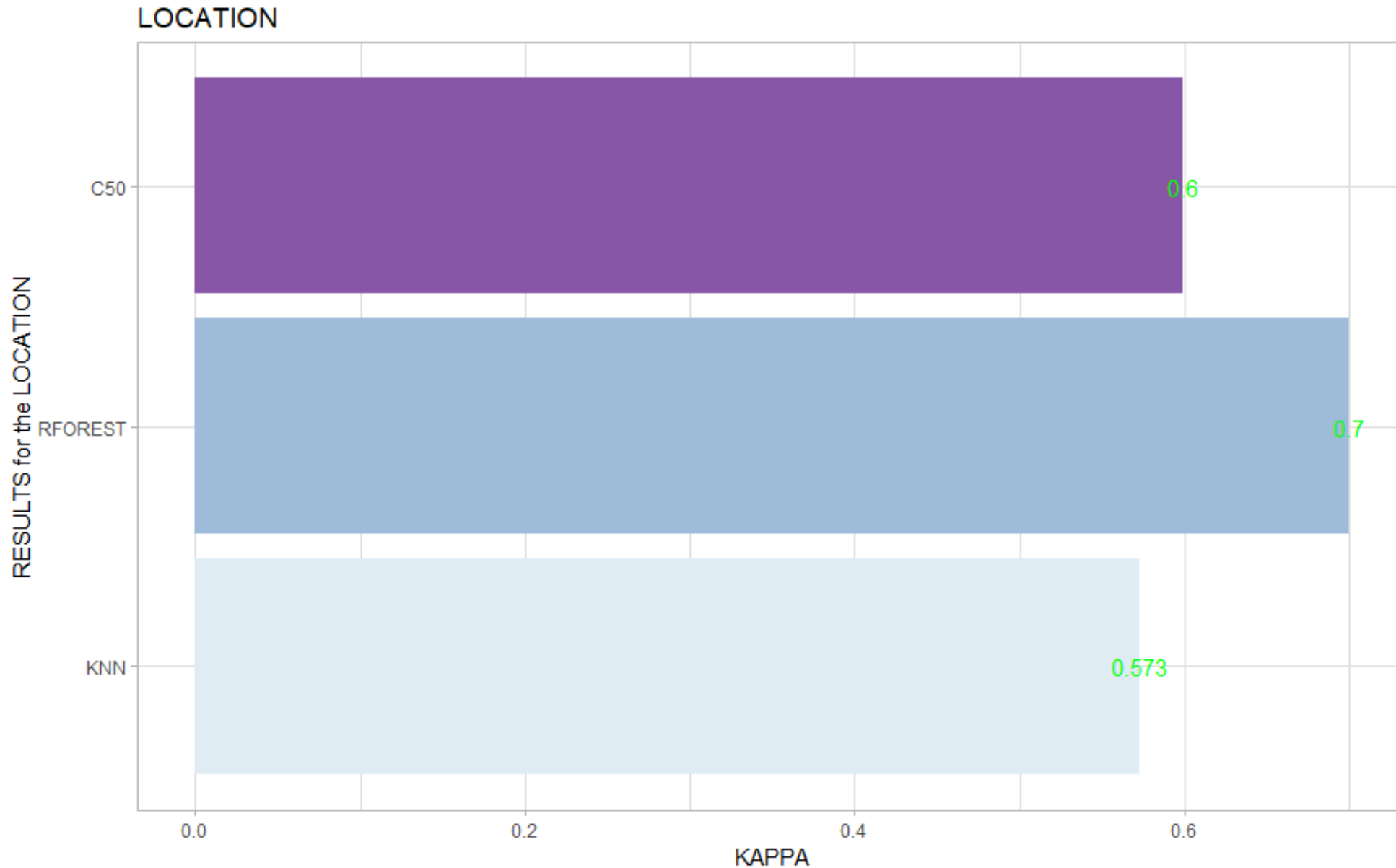# Metrics for each Model on the Prediction of LOCATION by Building (Final run)

ACCURACY for each Model on the Prediction of LOCATION (Final run)

KAPPA for each Model on the Prediction of LOCATION (Final run)

Accuracy & KAPPA for Prediction of LOCATION (Final Results)