



# **ANÁLISIS DEL DATASET CITIBIKE DE NUEVA YORK 2020 + IDEA DE NEGOCIO 'BIWHEELS'**

Máster Big Data & Data Science. Aplicaciones al Comercio, Empresa y Finanzas.  
Ainara Celaya Landa

# Resultados de los modelos

## 1º Modelo

Generalized Linear Model

1336 samples  
20 predictor  
2 classes: 'No', 'Si'

No pre-processing  
Resampling: Cross-validated (3 fold)  
Summary of sample sizes: 890, 891, 891  
Resampling results:

| ROC       | Sens      | Spec      |
|-----------|-----------|-----------|
| 0.8385956 | 0.4907407 | 0.9446503 |

## 2º Modelo

Ranger result

Call:  
ranger(fe\_usertype ~ ., data = my\_train2, num.trees = 100, importance = "impurity", write.forest = TRUE, min.node.size = 1, splitrule = "gini", verbose = TRUE, classification = TRUE)

|                                  |                |
|----------------------------------|----------------|
| Type:                            | Classification |
| Number of trees:                 | 100            |
| Sample size:                     | 1336           |
| Number of independent variables: | 23             |
| Mtry:                            | 4              |
| Target node size:                | 1              |
| Variable importance mode:        | impurity       |
| Splitrule:                       | gini           |
| OOB prediction error:            | 14.37 %        |

## 3º Modelo

Ranger result

Call:  
ranger(fe\_usertype ~ ., data = my\_train3, num.trees = 100, importance = "impurity", write.forest = TRUE, min.node.size = 1, splitrule = "gini", verbose = TRUE, classification = TRUE)

|                                  |                |
|----------------------------------|----------------|
| Type:                            | Classification |
| Number of trees:                 | 100            |
| Sample size:                     | 1336           |
| Number of independent variables: | 13             |
| Mtry:                            | 3              |
| Target node size:                | 1              |
| Variable importance mode:        | impurity       |
| Splitrule:                       | gini           |
| OOB prediction error:            | 13.85 %        |

## 4º Modelo

Ranger result

Call:  
ranger(fe\_usertype ~ ., data = my\_train3, num.trees = 200, importance = "impurity", write.forest = TRUE, min.node.size = 1, splitrule = "gini", verbose = TRUE, classification = TRUE)

|                                  |                |
|----------------------------------|----------------|
| Type:                            | Classification |
| Number of trees:                 | 200            |
| Sample size:                     | 1336           |
| Number of independent variables: | 13             |
| Mtry:                            | 3              |
| Target node size:                | 1              |
| Variable importance mode:        | impurity       |
| Splitrule:                       | gini           |
| OOB prediction error:            | 13.10 %        |

## Evaluación modelo ganador

```
> valor_pred<- predict(fit4, data= my_test4)
> table(my_test4$fe_usertype, valor_pred$predictions)

      No  Si
No    50  31
Si    10 242
> library(MLmetrics)
> Accuracy(y_pre = valor_pred$predictions, y_true= my_test4$fe_usertype)
[1] 0.8768769
```

- Precisión del 87.68%
- Las diagonales de 'No-Si' y 'Si-No' serían los valores que conocemos como falsos positivos y negativos.



### Plan Mayores de 65

No se espera una fidelización.

Descuento fijo correspondiente de cada ciudad por pertenecer a este grupo de edad.



### Plan Familia

- Introducción de bicicletas con sillitas y bicicletas infantiles.
- Para largos recorridos, alternativas uniendo bicicleta con otro medio de transporte.

Por una movilidad de diferentes medios de transporte, trasbordos gratis y descuentos en el comercio local.



### Plan Joven

El sedentarismo es un porcentaje muy alto en este grupo. Enmarcado en la educación de hábitos saludables se premia por kms recorridos.

10% descuento a < 25 años y kms canjeables en descuentos en establecimientos de ocio y eventos culturales.



### Plan Comerciantes

Apoyo al comercio local facilitando bicicletas con opción de carga para poder introducir la opción de 'delivery'.

Los Ayuntamientos subvencionarán parte del coste.