1. TODAS LAS MUESTRAS

* **Con batch**

**No dicotomizada**

metrics\_rf\_equipo

metrics\_KNN\_equipo

metrics\_glm\_equipo

metrics\_RANGER\_equipo

metrics\_rf\_fecha

metrics\_KNN\_fecha

metrics\_glm\_fecha

metrics\_ranger\_fecha

**Dicotomizada**

metrics\_rfdico\_equipo

metrics\_GLM\_equipo\_dic\_equipo

metrics\_KNN\_dic\_equipo

metrics\_RANGER\_dico\_equipo

metrics\_RF\_fecha\_dic

metrics\_GLM\_fecha\_dic

metrics\_ran\_fecha\_dic

metrics\_KNN\_fecha\_dic

* **Sin batch**

**Dicotomizada**

metrics\_rf\_dico\_matrix\_tot

metrics\_KNN\_dico

metrics\_GLM\_dico

metrics\_RANGER\_dico

**No dicotomizada**

metrics\_rf\_sin\_correccion

metrics\_GLM\_matrix\_total

metrics\_knn\_matrix\_total

metrics\_RANGER\_matrix\_tot

1. SIN LOS CT low

* **Con batch**

**Dicotomizada**

metricsRF\_equipo\_dico\_carga

metricsGLM\_equipo\_dic\_carga

metricsKNN\_equipo\_dic

metricsRanger\_equipo\_dich\_carga

metricsRANGER\_fecha\_dic\_carga

metricsRF\_fecha\_dic\_carga

metricsGLM\_fecha\_dic\_carga

metricsKNN\_fecha\_dic\_carga

**No dicotomizada**

metricsRF\_equipo\_carga

metricsKNN\_equipo\_carga

metricsGLM\_equipo\_carga

metricsRanger\_equipo\_carga

metricsRanger\_FECHA\_carga

metricsRF\_fecha\_carga

metricsKNN\_fecha\_carga

metricsGLM\_fecha\_carga

* **Sin batch**

**Dicotomizada**

metricsRFd\_carga

metricsKNNd\_carga

metricsGLMd\_carga

metricsRanger\_matrix\_dich\_carga

**No dicotomizada**

metricsRF\_matrix\_carga

metricsGLM\_matrix\_carga

metricsKNN\_matrix\_carga

metricsRANGER\_matrix\_carga

1. SIN otros virus

* **Con batch**

**No dicotomizada**

metrics\_rf\_equipo\_covid

metrics\_KNN\_equipo\_covid

metrics\_GLM\_covid\_equipo

metrics\_RANGER\_equipo\_covid

metrics\_RANGER\_fecha\_covid

metrics\_RF\_fecha\_covid

metrics\_KNN\_fecha\_covid

metrics\_GLM\_fecha\_covid

**Dicotomizada**

metrics\_rf\_dico\_equipo

metrics\_GLM\_dico\_equipo\_covid

metrics\_KNN\_dico\_equipo\_covid

metrics\_ranger\_equipo\_dico\_covid

metrics\_Ranger\_fecha\_covid\_dico

metricsRF\_fecha\_dico\_covid

metricsGLM\_fecha\_dic\_covid

metrics\_KNN\_fecha\_dic\_covid

* **Sin batch**

**Dicotomizada**

metricsRFd\_matrix\_covid

metricsKNNd\_matrix\_dico

metricsGLMd\_matrix\_covid

metricsRanger\_matrix\_dich\_covid

**No dicotomizada**

metricsRF\_fecha\_covid

metrics\_GLM\_fecha\_covid

metricsKNN\_fecha\_covid

metricsRanger\_matrix\_covid

1. SIN otros virus high/int

* **Con batch**

**Dicotomizados**

metricsRF\_equipo\_dic\_cc

metricsGLM\_equipo\_dic\_cc

metricsKNN\_equipo\_dic\_CC

metricsRanger\_equipo\_dich\_cc

metricsRanger\_FECHA\_dich\_cc

metricsRF\_fecha\_dic\_cc

metricsGLM\_fecha\_dic\_cc

metricsKNN\_fecha\_dic\_cc

**No dicotomizados**

metricsRF\_equipo\_cc

metricsKNN\_equipo\_cc

metricsGLM\_equipo\_cc

metricsRanger\_equipo\_cc

metricsRanger\_FECHA\_cc

metricsRF\_fecha\_cc

metricsKNN\_fecha\_cc

metricsGLM\_fecha\_cc

* **Sin batch**

**Dicotomizados**

metricsRFd\_matrix\_cc

metricsKNNd\_matrix\_cc

metricsGLMd\_matrix\_cc

metricsRanger\_matrix\_dich\_cc

**No dicotomizados**

metricsRF\_matrix\_cc

metricsGLM\_matrix\_cc

metricsKNN\_MATRIX\_cc

metricsRANGER\_matrix\_cc

ML\_tot\_WObatch\_dico, ML\_tot\_WObatch, ML\_high\_batch\_dico, ML\_high\_batch, ML\_high\_WObatch, ML\_high\_WObatch\_dico, ML\_high\_WObatch\_dico, ML\_covid\_batch, ML\_covid\_batch\_dico, ML\_covid\_dico, ML\_covid, ML\_cc\_batch\_dico, ML\_cc\_batch, ML\_cc\_dico, ML\_cc

# Método 2: Crear un nuevo factor con los niveles correctos

predRF\_equipo <- factor(as.character(predRF\_equipo),

levels = c("X1", "X2"),

labels = c("1", "2"))

predRF\_equipo <- factor(predRF\_equipo, levels = c("1", "2"), labels = c("Cov.Neg", "Cov.Pos"))

y\_test <- factor(y\_test, levels = c("1", "2"), labels = c("Cov.Neg", "Cov.Pos"))

# Crea la matriz de confusión usando caret

RF\_metrics\_equipo <- caret::confusionMatrix(predRF\_equipo, y\_test, positive = "Cov.Pos")

# Carga los paquetes necesarios

library(caret)

accuracy <- RF\_metrics\_equipo$overall["Accuracy"]

kappa <- RF\_metrics\_equipo$overall["Kappa"]

# Métricas por clase

sensitivity <- RF\_metrics\_equipo$byClass["Sensitivity"]

specificity <- RF\_metrics\_equipo$byClass["Specificity"]

precision <- RF\_metrics\_equipo$byClass["Pos Pred Value"]

recall <- RF\_metrics\_equipo$byClass["Sensitivity"] # Igual a sensitivity

f1\_score <- RF\_metrics\_equipo$byClass["F1"]

npv <- RF\_metrics\_equipo$byClass["Neg Pred Value"]

prevalence <- RF\_metrics\_equipo$byClass["Prevalence"]

detection\_rate <- RF\_metrics\_equipo$byClass["Detection Rate"]

balanced\_accuracy <- RF\_metrics\_equipo$byClass["Balanced Accuracy"]

# Calcular LR+ y LR- que no vienen directamente en confusionMatrix

LR\_plus <- sensitivity / (1 - specificity)

LR\_minus <- (1 - sensitivity) / specificity

# Para manejar valores especiales

LR\_plus <- ifelse(is.nan(LR\_plus) | is.infinite(LR\_plus), NA, LR\_plus)

LR\_minus <- ifelse(is.nan(LR\_minus) | is.infinite(LR\_minus), NA, LR\_minus)

# Crear un dataframe con todas las métricas

metrics\_rf\_equipo <- data.frame(

Metric = c("Accuracy", "Kappa", "Sensitivity", "Specificity", "Precision",

"F1\_Score", "NPV", "Prevalence", "Detection\_Rate",

"Balanced\_Accuracy", "LR+", "LR-", "AUC"),

Value = c(accuracy, kappa, sensitivity, specificity, precision,

f1\_score, npv, prevalence, detection\_rate,

balanced\_accuracy, LR\_plus, LR\_minus, auc\_value))

# Mostrar los resultados

print(metrics\_rf\_equipo)

**Estandarizar los dico**

**Desafiar ciegos**

**Funciones donde se pueda**

**x <- trainData\_equipo[, -ncol(trainData\_equipo)] # se incluyen todas las columnas excepto la última**

1. Tengo que crear ese dataframe donde también pongo el nombre de la institución, equipo, y qué hago con PCR.Cov? también?
2. Tengo que aplicar combat por cada base de datos que usé? TOTAL, SIN LOW, SIN OTROS VIRUS Y COMBINADOS?

**Correcion efecto batch EQUIPO: Combat**

```{r}

combat\_corrected\_equipo\_tot <- ComBat(

dat = t(data\_matrix),

batch = batch\_info, # Información de batch

par.prior = TRUE,

prior.plots = FALSE)

```

```{r}

# Transponer de vuelta los datos corregidos

combat\_corrected\_equipo\_tot <- t(combat\_corrected\_equipo\_tot)

**Correcion efecto batch FECHA: Combat**

```{r}

# Ejemplo de datos de entrada

data\_matrix <- featureMatrix\_num # Tu matriz de datos

metadata <- Datos\_actualizados # Dataframe de metadatos con información de batch

batch\_info <- metadata$fecha # Vector con la información de batch

```

```{r}

# Verificar si 'batch\_info' es un factor o un vector de caracteres

str(batch\_info)

# Si es necesario, convertirlo a un factor

batch\_info <- factor(batch\_info)

```

```{r}

sum(is.na(data\_matrix)) # Revisa si hay valores NA en la matriz

```

```{r}

combat\_corrected\_fecha\_tot <- ComBat(

dat = t(data\_matrix),

batch = batch\_info, # Información de batch

par.prior = TRUE,

prior.plots = FALSE

)

```

```{r}

# Transponer de vuelta los datos corregidos

combat\_corrected\_fecha\_tot\_tot <- t(combat\_corrected\_fecha\_tot)

# Guardar la matriz corregida

write.csv(combat\_corrected\_fecha\_tot, "corrected\_featureMatrix.csv")

```

1. Donde coloco el dataset para la predicción?
2. Cómo hago para ejecutar flexdashboard si tenog archivos RMD? No me deja traer la info con source
3. Como hago esto en flexdashboard?



1. Cómo hago para que se ejecuten los 24 scripts uno detrás del otro?
2. Lo del CV<40 lo hago para los mejores modelos o antes de cada vez que dicotomizo y luego ahí saco los mejores?