

Benemérita Universidad Autónoma de Puebla
Facultad de Ciencias de la Computación

Máquinas de aprendizaje

Reporte: Evaluación del rendimiento de un modelo, parte II



BUAP

Docente: Abraham Sánchez López

Alumno

Taisen Romero Bañuelos

Matrícula

202055209

Evaluación de modelos pt. 2

El documento de hoy lo sentí como un repaso del documento anterior. Me hubiera gustado ver alguna aplicación de MCC. Aún así, la aplicación de kappa fue ilustrativa aunque de repente llegaba a confundir las cosas con la matriz cruzada de validación por la concordancia entre valores predichos y los reales. Por ejemplo, en la siguiente línea

```
kappa <- kappa2(data.frame(credit_actual, credit_pred))$value
```

De repente me parecía que eran casi lo mismo que la validación cruzada, pero leyendo de nuevo las cosas entendí que aunque no son lo mismo, si están relacionadas ambas cosas. Mientras que la matriz de confusión muestra las frecuencias de las predicciones correctas e incorrectas en cada categoría, kappa ajusta ese acuerdo para tener en cuenta la probabilidad de que la coincidencia ocurra por casualidad.

El rendimiento del modelo fue moderado, la medida kappa (0.294) apenas fue mejor que el azar. Si sólo nos quedamos con los resultados de kappa no sería descabellado descartar el modelo, pero aquí me pregunto si debemos descartar un modelo porque es el modelo el que está mal y no porque la medida de precisión que la evalúa no es representativa de los datos. Kappa es útil cuando tienes clases desbalanceadas, ya que mide el acuerdo entre las predicciones y los resultados reales ajustados por el azar, sin embargo, cojea si los datos están demasiado desbalanceados.

```
> table(credit$default)
  no yes
700 300
```

Como se ve, hay una diferencia significativa en las observaciones de la clase no, lo que significa que los datos están muy desbalanceados. Creo que esto es evidencia suficiente para demostrar que kappa no es la opción más adecuada para evaluar a nuestro modelo.

Desde el documento pasado (e incluso más antes) ha sido imperativo el no dejarse llevar por las primeras impresiones de los modelos de evaluación, y creo que este ejercicio es prueba de ello. Pese a que no se insinúa la pregunta de si kappa es adecuado para el ejemplo (omitendo el resumen final, en todo caso), creo que es necesario dudar en todo momento sobre las medidas de rendimiento que obtenemos. Creo que en esta clase de problemas hay que considerar siempre que uno puede estar equivocado y también que las cosas pueden ocultar más información de la que se nos presenta.