

Ahora les pido estudiar otro tema en el cual se trata de evaluación de eficiencia del clasificador. Por favor estudien de Lecture 11.1 hasta Lecture 11.5. Allí van a aprender las métricas que permitan obtener una idea de que tan bueno es el clasificador. Las métricas se calculan a partir de una tabla que se llama matriz de confusión.

Claro, una manera de evaluar el clasificador ya vimos: es el valor de la función de coste. Sin embargo, este valor es el promedio sobre todos los datos en el conjunto de prueba o el de entrenamiento, es decir, no mide el error para cada ejemplo. Además, este valor no nos permite ver, cuantos ejemplos fueron clasificados correctamente y cuantos no. En los videos van a ver como se puede evaluar el clasificador de manera más clara y precisa.

Para programación, vamos a usar el programa de la regresión logística donde la regresión se programó desde cero. Por favor programen (DESDE CERO, NO USEN FUNCIONES IMPLEMENTADAS EN SCIKIT-LEARN O DE OTRO PAQUETE), para el conjunto de prueba, la generación de la matriz de confusión, el cálculo de las métricas: precisión, recall y la medida F1.

Envíenme por favor el programa y la captura de pantalla con la matriz de confusión y las tres métricas mencionadas.