08-06-2021

Algoritmos de clasificación(aprendizaje supervisado) . Variables categóricas

Clasificación binaria

Regresión logística. No se considera regresión sino clasificación. La variable que se predice es categórica las que se utilizan para predecir pueden ser categóricas o no

Es un tipo de análisis de regresión para predecir variables categóricas. Valores entre 0 y 1. Nos da una probabilidad

Max\_iter=700 🡪 número de iteraciones que te permitan encontrar el mínimo de error

Las categorías han de pasarse a número

Validación cruzada: obtener un conjunto de entrenamiento independiente (extraído del conjunto de entrenamiento total) y conocer el error que se comete en cada conjunto de validación de la iteración. Se utiliza para evitar el sobreentrenamiento

En el sobreaprendizaje los errores de entrenamiento continúan bajando y los de validación después de ir bajando vuelven a subir porque se empieza a aprender detalles no relevantes en el aprendizaje que nos interesa

Matriz de confusión

Falsos positivos y negativos. Nos permite identificar en que sentido se está equivocando el modelo creado

Entrenamiento y clasificación de dígitos con regresión logística